Hydraulic-hydrologic model for the Zambezi River using satellite data and artificial intelligence techniques

THÈSE N° 6225 (2014)

PRÉSENTÉE LE 11 JUILLET 2014
À L'ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE DE CONSTRUCTIONS HYDRAULIQUES

ET

À L'INSTITUTO SUPERIOR TÉCNICO (IST) DA UNIVERSIDADE DE LISBOA
DEPARTAMENTO DE ENGENHARIA CIVIL, ARQUITECTURA E GEORRECURSOS (DECIVIL)

PROGRAMME DOCTORAL EN ENVIRONNEMENT

ET

DOUTORAMENTO EM ENGENHARIA CIVIL

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES (PhD)

PAR

José Pedro GAMITO DE SALDANHA CALADO MATOS

acceptée sur proposition du jury:

Prof. J.-L. Scartezzini, président du jury
Prof. A. Schleiss, Prof. M. M. Portela Correia dos Santos Ramos da Silva, directeurs de thèse
Dr J.-M. Fallot, rapporteur
Dr F. Jordan, rapporteur
Dr B. Schaefli, rapporteuse

EPFL
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2014

Man searches for answers and finds only questions

Unknown author

ABSTRACT

The Zambezi River Basin (ZRB), in Africa, spreads over some unfathomable 1 370 000 km². In all its magnificence it is home to approximately 30 million inhabitants, harbors a number of priceless wildlife sites and has an estimated hydropower production capacity of 13 000 MW, of which only about 5 000 MW are currently exploited. The Zambezi River is central to both the culture and the economy of riparian countries. With a heterogeneous landscape and a semi-arid climate, the basin faces great challenges brought about by growing populations, soaring economic growth, and climate change. In the future, increasing pressure on water resources is inescapable.

Hydrological modeling will certainly support decision makers in all levels of decision as they rise to meet the forthcoming challenges. While historically the basin's size, heterogeneity, political situation, and constraining lack of hydrological data have conditioned the scope and success of hydrological models of the basin, relatively new technologies such as satellite remote sensing or machine learning present promising tools with which some of these problems can be addressed.

The present work set out to develop a performing hydraulic-hydrological model of the ZRB at a daily time scale, envisaging future use in dam operation optimization and synchronization, environmental impact assessments, evaluation of future scenarios (predicting responses to climate change and increased demands) and a broad range of other studies related to themes such as wildlife, water chemistry, sediment transport, and integrated water management.

In order for this to be successful, constraining issues, mostly related to input data, would have to be addressed before the actual modeling stages; resorting to satellite remote sensed data would be mandatory and the most had to be made out of the few good quality discharge series available. Also, it was early recognized that no single model could be a "best" choice for such a wide array of uses and that a large emphasis would have to be placed on model calibration and validation.

Aiming to extend the time scope of the analysis, the novel Pattern-Oriented Memory (POM) historical rainfall interpolation methodology was introduced. Based on machine learning models, POM was shown to be superior to competing methods in data scarce environments and when the true rainfall field shows high variability. Over the ZRB, errors in the POM interpolated rainfall series were observed to be on par with those of state-of-the-art satellite rainfall estimates. Still, POM presents additional advantages worth noticing: its performance improves as more satellite data becomes available; and POM interpolated rainfall can be directly combined with satellite estimates as forcing for hydrological models leading to minimal "change of support" problems.

The use of machine learning models for discharge forecast was developed in four fronts: the comparison of alternative models (e.g. Autoregressive Moving-Average (ARMA), Artificial Neural Networks (ANN) and Support-Vector Regression (SVR)); the enhancement of rainfall aggregation techniques; the study of limitations inherent to SVR forecasting models; and the development of a non-parametric empirical uncertainty post-processor. Going beyond the development of deterministic forecasting models with promising accuracies, even for long lead times of up to 60 days at Victoria Falls, the conducted research most notably motivated a reevaluation of previous

findings reported in literature by showing that SVR models are particularly hazardous when used for discharge forecasting purposes; a conclusion based on their underlying theoretical principles and easily observable in practice.

A novel non-parametric empirical uncertainty post-processor was developed. The proposed methodology is able to effectively generate probabilistically correct uncertainties of detrended series given a representative set of training patterns. It is an (informal) technique that, unlike Bayesian methods (formal), does not require the definition of likelihood functions nor an external "conceptual" model of the phenomenon being modeled. Performing, extremely versatile, and straightforward to set up, it can be easily adapted to incorporate new information. The potential range of applicability of the methodology goes well beyond discharge forecasting and even hydrology.

The Soil and Water Assessment Tool (SWAT) was used in order to prepare a continuous-time hydrological model of the whole ZRB. Recognizing the importance of sound calibration and validation phases, investments were made on the development of a flexible and computationally efficient calibration interface. In parallel, an analysis of the Soil and Water Assessment Tool (SWAT) hydrological model in its application to the ZRB has evidenced inadequacies in the source code which should be taken into account, particularly in catchments with relatively large waterbodies.

Resorting to millions of simulations, the full calibration of daily hydrological models covering the whole basin, from the Upper Zambezi to a few kilometers upstream from the Delta (Marromeu) was accomplished – it is believed – for the first time. Heterogeneity was shown to play a noticeable role in the basin's hydrology and it is recommended that the calibration of future models allows for the definition of regional parameters. Among four tested calibration schemes, best results were obtained using a regional-regularized calibration approach due to its capacity of approximating contributions, not only of subbasins along the main reach of the Zambezi, but also along its tributaries.

The most important outcomes of the research have been, along with original data and works from fellow (African Dams Project) ADAPT researchers, be conveyed to stakeholders through the ongoing ADAPT online database project (http://zambezi.epfl.ch), initially developed in the scope of this thesis.

Keywords: artificial neural networks, calibration, discharge forecasting, hydrological modeling, machine learning, optimization, Pattern-Oriented Memory, rainfall interpolation, support-vector machines, SWAT, TRMM, uncertainty, Zambezi.

RÉSUMÉ

Le bassin du fleuve Zambèze (ZRB), en Afrique, s'étend sur quelques 1 370 000 km². Dans toute sa splendeur, il abrite environ 30 millions d'habitants, héberge plusieurs sites naturels d'une valeur inestimable et a une potentiel de production hydroélectrique estimé à 13 000 MW, dont seulement environ 5000 MW sont actuellement exploités. Le Zambèze est central à la fois pour la culture et l'économie des pays riverains. Avec un paysage hétérogène et un climat semi-aride, le bassin fait face à de grands défis posés par la croissance démographique, l'enrichissement économique, et le changement climatique. Dans l'avenir, la pression croissante sur les ressources en eau est incontournable.

La modélisation hydrologique va certainement aider les parties prenantes à répondre aux défis à venir. Historiquement, la situation politique, la taille et l'hétérogénéité du bassin, ainsi que le manque contraignant de données hydrologiques ont conditionné la précision de modèles hydrologiques du bassin. Les technologies récentes et prometteuses telles que la télédétection par satellite ou l'apprentissage automatique peuvent présenter des alternatives pour résoudre certains de ces problèmes.

Le présent travail a entrepris d'élaborer un modèle hydrologique du ZRB à un pas de temps journalier, envisagé pour des utilisations futures telles que des évaluations d'impact environnemental, l'optimisation de l'exploitation et la synchronisation des barrages, ainsi qu'un large éventail d'autres études portant sur des thèmes tels que la faune, de la chimie de l'eau, le transport des sédiments et de la gestion intégrée des eaux.

Dans ce but, des problèmes, la plupart liées aux données d'entrée, devront être abordées avant les étapes de modélisation. Le recours aux données de télédétection par satellite est obligatoire et les séries de débits de bonne qualité disponibles devront être prises en compte. En outre, il a été reconnu qu'il n'existe pas de «meilleur» modèle pour un tel éventail d'utilisations et qu'un grand accent devrait être mis sur l'étalonnage du modèle et sa validation.

Le *Pattern-Oriented Memory* (POM), une méthodologie d'interpolation pour précipitations historiques, a été introduit pour prolonger la longueur des simulations. Basé sur les modèles d'apprentissage automatique, le POM est supérieur aux méthodes concurrentes dans des environnements pauvres en données et surtout quand la distribution spatiale de la pluie montre une forte variabilité. Sur le ZRB, les erreurs sur l'interpolation des séries de précipitation obtenues avec POM sont comparables à celles des précipitations par satellite. En outre, le POM présente des avantages supplémentaires intéressants : sa performance s'améliore à mesure que davantage de données satellite sont disponibles et les précipitations interpolées par le POM peuvent être directement associées aux estimations par satellite comme données pour des modèles hydrologiques avec un effet de « changement de soutien » minimal.

L'utilisation de modèles d'apprentissage automatique pour la prévision des débits a été développé sur quatre fronts: (1) la comparaison de modèles alternatifs (par exemple des modèles autorégressifs et moyenne mobile (ARMA), des réseaux de neurones artificiels (ANN) et des

machines à vecteurs de support pour régression (SVR)); (2) l'amélioration des techniques d'agrégation des précipitations; (3) l'étude des limites inhérentes aux modèles SVR de prévision et (4) le développement d'un post-processeur non paramétrique pour l'incertitude. Au-delà de l'élaboration de modèles de prévision déterministes avec des précisions prometteuses, même pour de longs délais (jusqu'à 60 jours à Victoria Falls), les recherches menées ont notamment motivé une réévaluation des résultats antérieurs rapportés dans la littérature ont montré que les modèles SVR sont particulièrement dangereux lorsqu'ils sont utilisés à des fins de prévision de débits. Cette conclusion est fondée sur les principes théoriques sous-jacents et facilement observables dans la pratique.

Un post-processeur non paramétrique pour l'incertitude a été développé. La méthodologie proposée est capable de générer efficacement des incertitudes autour des séries donnant un ensemble représentatif de cas historiques. Il s'agit d'une technique informelle qui, contrairement aux méthodes Bayésiennes (formelles), ne nécessite ni la définition de fonctions de vraisemblance, ni d'un modèle externe « conceptuel » du phénomène modélisé. D'une exécution très polyvalente et simple à mettre en place, il peut être facilement adapté pour incorporer de nouvelles informations. Le potentiel d'application de la méthodologie va bien au-delà des prévisions des débits et même de l'hydrologie.

L'outil *Soil and Water Assessment Tool* (SWAT) a été utilisé afin de préparer un modèle hydrologique continu du ZRB. Reconnaissant l'importance de l'étalonnage et des phases de validation, des investissements ont été faits sur le développement d'une interface d'étalonnage flexible et efficace. En parallèle, une analyse de SWAT dans son application au ZRB a mis en évidence des insuffisances dans le code source qui doivent être prises en compte, en particulier dans les bassins versants avec des plans d'eau relativement étendus.

Grace à des millions de simulations, le calage complet de modèles hydrologiques journaliers couvrant le bassin, du Haut-Zambèze jusqu'à quelques kilomètres à l'amont du Delta (Marromeu) a été réalisé – vraisemblablement – pour la première fois. L'hétérogénéité a démontré jouer un rôle notable dans l'hydrologie du bassin et il est recommandé que l'étalonnage de futurs modèles prend en compte la définition de paramètres régionaux. Parmi les quatre types d'étalonnage testés, les meilleurs résultats ont été obtenus en utilisant une approche régionale-régularisée en raison de sa capacité pour simuler non seulement les contributions des sous-bassins le long du Zambèze, mais aussi de ses affluents.

Les résultats les plus importants de la recherche, avec les données originales et les travaux d'autres chercheurs du « African Dams Project » (ADAPT), ont été transmis aux parties prenantes au travers d'une base de données en ligne (http://zambezi.epfl.ch), initialement développée dans le champ d'application de cette thèse.

Mots-clés: apprentissage automatique, calage, incertitude, interpolation de précipitations, machines à vecteurs de support, modélisation hydrologique, optimisation, Pattern-Oriented Memory, prévision des débits, réseaux de neurones artificiels, SWAT, TRMM, Zambèze.

RESUMO

A bacia do rio Zambeze (ZRB), em África, desenvolve-se sobre uma vasta superfície de 1 370 000 km^2. Em toda a sua grandeza, ela alberga aproximadamente 30 milhões de habitantes, inclui várias zonas de valor ecológico ímpar e tem um potencial hidroeléctrico estimado em 13 000 MW, dos quais apenas 5 000 MW se encontram explorados. O rio Zambeze assume um papel central tanto na cultura como na economia locais. Com uma paisagem heterogénea e clima semiárido, a bacia enfrenta desafios significativos motivados pelo crescimento demográfico, o rápido desenvolvimento económico e alterações climáticas. No futuro, a pressão crescente sobre os recursos hídricos locais será inevitável.

A modelação hidrológica irá certamente servir de apoio a decisores de todos os níveis no seu esforço para enfrentar desafios vindouros. Embora, historicamente, a dimensão, a heterogeneidade, a situação política e a escassez de dados hidrológicos tenham limitado a abrangência e o sucesso dos modelos hidrológicos desenvolvidos para a bacia, tecnologias relativamente novas tais como a detecção remota por satélite ou a aprendizagem máquina constituem ferramentas cujo potencial poderá servir para ultrapassar tais limitações.

O presente trabalho foi encetado com o objectivo de produzir um modelo hidrológico da ZRB à escala diária e passível de ser utilizado na optimização e sincronização das operações de barragens, em avaliações de impactos ambientais e num vasto leque de outros estudos relacionados com a vida selvagem, a química aquática, o transporte de sedimentos ou a gestão integrada de recursos hídricos.

Por forma a cumprir os objectivos propostos, algumas limitações, principalmente associadas aos dados disponíveis, teriam que ser estudadas antes da fase de modelação propriamente dita; o recurso a dados de satélite seria incontornável e o máximo proveito deveria ser retirado das séries de caudais observados disponíveis. Adicionalmente, cedo se verificou que não seria possível apontar um modelo único como a "melhor" escolha para o conjunto de utilizações desejado e que grande ênfase teria que ser dada às etapas de calibração e validação.

Procurando alargar o período analisado, o método de interpolação *Pattern-Oriented Memory* (POM), orientado para a reprodução espacial de séries históricas de precipitação, foi introduzido. Com base em modelos de aprendizagem máquina, o POM demonstrou ser superior a métodos alternativos em situações de escassez de dados e na presença de variabilidade elevada do campo espacial da precipitação. Na ZRB, os erros associados à interpolação POM são semelhantes aos apresentados pelas melhores estimativas de precipitação por satélite. Para além do referido, o POM tem vantagens adicionais: o seu desempenho melhora à medida que mais e melhores dados de satélite vão ficando disponíveis e os mapas interpolados podem ser directamente combinados com observações de satélite mais recentes para efeitos de modelação hidrológica, observando-se apenas efeitos mínimos relacionados com a "alteração de suporte".

A aplicação de modelos de aprendizagem máquina à previsão de caudais foi dividida em quatro frentes: a comparação de modelos alternativos (e.g. modelos auto-regressivos de médias móveis (ARMA), redes neuronais artificiais (ANN) e máquinas de vectores de suporte para regressão (SVR)); o melhoramento de técnicas de agregação da precipitação; o estudo de limitações associadas com a utilização de SVR e o desenvolvimento de um pós-processador de incerteza não paramétrico. Indo para além da preparação de modelos de previsão do caudal determinísticos com desempenhos promissores, mesmo para horizontes temporais alargados de 60 dias nas cataratas Vitória, a investigação motiva a reavaliação da literatura publicada sobre a utilização de modelos SVR quando utilizados para previsão de caudais, tendo evidenciado com base em princípios teóricos e demonstrado na prática perigos a eles associados.

Um novo pós-processador de incerteza não paramétrico foi desenvolvido. A metodologia proposta é capaz de gerar distribuições de probabilidade correctas associadas a realizações de séries temporais com base num conjunto de observações passadas representativo. Trata-se de uma técnica informal que, ao contrário de técnicas Bayesianas (formais), não requere a definição de funções de verosimilhança ou de um modelo "conceptual" do fenómeno subjacente. Evidenciando bons desempenhos, extremamente versátil e de simples utilização, a metodologia pode ser facilmente adaptada para incorporar nova informação. O seu potencial vai bem para além da previsão de caudais e mesmo do campo da hidrologia.

O modelo hidrológico *Soil and Water Assessment Tool* (SWAT) foi utilizado com vista à modelação da totalidade da ZRB. Reconhecendo a importância das etapas de calibração e validação, investiu-se no desenvolvimento de uma interface de calibração flexível e computacionalmente eficiente. Em paralelo, uma análise do SWAT na sua aplicação à bacia evidenciou algumas falhas no código-fonte que deveriam ser tidas em conta, particularmente em bacias em que grandes corpos de água estejam presentes.

Apoiada em milhões de simulações, a calibração de modelos hidrológicos da bacia à escala diária, desde o Alto Zambeze praticamente até ao Delta (Marromeu), foi realizada – crê-se – pela primeira vez. Revelou-se que a heterogeneidade desempenha um papel importante na hidrologia local e recomenda-se que a calibração de modelos futuros possibilite a definição regional de parâmetros. Entre os quatro esquemas de calibração testados, os melhores resultados foram obtidos utilizando uma metodologia regional-regularizada devido à sua capacidade de simular, não só contribuições ao longo do trecho principal do Zambeze, como ao nível dos principais tributários.

Os resultados mais importantes decorrentes da presente investigação, assim como os dados de base e trabalhos de outros investigadores do *African Dams Project* (ADAPT), foram transmitidos a partes interessadas através da base de dados *online* do ADAPT (http://zambezi.epfl.ch), um projeto inicialmente desenvolvido no âmbito da presente tese cujo desenvolvimento irá continuar no âmbito de um novo projeto de investigação.

Palavras-chave: aprendizagem máquina, calibração, incerteza, interpolação da precipitação, máquinas de vectores de suporte, modelação hidrológica, optimização, Pattern-Oriented Memory, previsão de caudal, redes neuronais artificiais, SWAT, TRMM, Zambeze.

ACKNOWLEDGMENTS

For a few days now I have been intermittently pondering on the best ways to go about the acknowledgements. By now, even without a satisfactory strategy, the time has come to put thoughts into words.

Memory is a funny thing. How it mystifyingly decides on what to forever keep and what to soon forget; how it tends to shamefully overvalue the insignificant happenings of yesterday in comparison to life-changing events gone by; how unjust and volatile behavior it might display. Despite my efforts, the reliance on memories of the last four years means that these acknowledgements will not be fair to all of the great people that have come across my path and deserve to be mentioned. To those whose important role slipped my mind I beg understanding, for it was against my wishes that it did so. Also, I would like to apologize for making these acknowledgements as much work related as personal; having both planes contributed to the completion of this endeavor, I neither could nor wanted to bring myself to draw a line between them.

Heartfelt recognition is due to both my advisors which, despite their different characters, continuously supported me along similar lines. Prof. Maria Manuela Portela was firstly responsible for my application to the IST-EPFL joint doctoral program. Accompanying me closely, she has always found the time to comment, review and consolidate my findings. Her friendship drove me forward and I would be hard-pressed to understate her role in the work that has been accomplished. Prof. Anton Schleiss welcomed me to his laboratory, allowed me freedom of thought, and gave me priceless opportunities to enrich my knowledge. Moderate and always informed in his opinions, he truly led me by example with is relentless working habits and thoughtful trait. The confidence they transmitted – many times undeserved – motivated me to overcome my many limitations (a process still in its early stages) and, in a very positive way, obliged me to grow closer to their expectations. I am not able to thank them enough. I also thank the president of the jury, Prof. J.-L. Scartezzini, and the examiners Prof. J.-M. Fallot, Prof. B. Schaefli, and Dr. F. Jordan for their availability to discuss the work herein presented.

A word of gratitude must be addressed to the ones who have contributed to setting up the IST-EPFL joint doctoral program. It was remarkable to be part of this initiative and having the change of conducting my studies at such reputable institutions.

Théodora Cohen Liechti walked a path similar to mine, also working on the Zambezi's hydrology. I greatly benefitted from her genral judgment and particular advice regarding French; even more from her friendship. She was a great colleague and many of the accomplishments of this thesis are owed to her valuable contribution.

aquilo que, infelizmente, tem que ficar por dizer. Obrigado por, apesar disso, me teres dado sempre autonomia para seguir o meu caminho e me teres apoiado incondicionalmente nas decisões que tomei.

Moreninha, és verdadeiramente especial; objecto da minha admiração constante. A tua ajuda foi inestimável e deste-me mais durante este período do que seria legítimo pedir. Os nossos primeiros seis meses, o nosso segundo ano. As muitas horas (sofridas) a escrever código. Os tantos conselhos sobre o caminho a seguir. Os volumes de fraldas, papas, roupas e brinquedos que carregaste. Os sorrisos trocados à distância. Os abraços apertados a cada reencontro. O desejo recorrente e insuportável de voltar a casa. Obrigado por tudo isso e obrigado por acreditares mim (mal) muito mais do que eu.

JPM, 2014

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ACE2	Altimeter corrected elevations 2
ADAPT	African Dams Project
AEDNMP	Average expected distance to the nearest measured pixel
AIRAD	Average interval-relative absolute deviation
AMALGAM	A multi-algorithm genetically adaptive multi-objective method
ANN	Artificial neural network
ANOVA	Analysis of variance
API	Application programming interface
ARA-Zambeze	*Administração Regional de Águas do Zambeze* (Water Regional Administration of the Zambezi)
ARMA	Autoregressive moving-average
ARMAX	Autoregressive moving-average with exogenous inputs
ASTER	Advanced spaceborne thermal emission and reflection
BI	Bayesian-inspired
CCES	Competence Center Environment and Sustainability
CEHIDRO	*Centro de Estudos de Hidrossistemas* (Centre for Hydrosystems Research)
CI	Confidence interval
CMA-ES	Covariance matrix adaptation evolution strategy
CMORPH	Climate Prediction Center morphing method
CSI	Consortium for Spatial Information
DEM	Digital elevation model
DHV	Dutch consulting company. Today named Royal HaskoningDHV.
DOE	Department of Energy
DREAM	Differential evolution adaptive metropolis
ECMWF	European Centre for Medium-Range Weather Forecasts
EF	Exceedance fraction
EKF	Extended Kalman filter
EnKF	Ensemble Kalman filter
EPFL	École Polytechnique Fédérale de Lausanne (Swiss Federal Institute of Technology – Lausanne)
EROS	US National Center for Earth Resources Observation and Science
ESA	European Space Agency
FAO	Food and Agriculture Organization
FAST	Fourier amplitude sensitivity test
FCT	*Fundação para a Ciência e a Tecnologia* (Foundation for Science and Technology)
FEWS	Famine Early Warning System
GA	Genetic algorithms
GDEM	Global digital elevation model
GE	Google Earth
GIS	Geographic information system
GLUE	Generalized likelihood uncertainty estimation
GMTED 2010	Global multi-resolution terrain elevation data 2010
GPCC	Global Precipitation Climatology Centre
GPM	Global precipitation measurement
GRDC	Global Runoff Data Centre
GSMaP	Global satellite mapping of precipitation

GSOD	Global surface summary of the day
HCB	Cahora Bassa Hydroelectric (*Hidroeléctrica de Cahora Bassa*)
HRU	Hydrologic response unit
HydroSHEDS	Hydrological data and maps based on shuttle elevation derivatives at multiple scales
IDW	Inverse distance weighting
IGeoE	Portuguese Army Geographical Institute (*Instituto Geográfico do Exército)*
InSAR	Interferometric synthetic aperture radar
IR	Infrared
IRBFN	Improved radial basis network functions
IRRMSE	Interval-relative root mean square error
IST	*Instituto Superior Técnico*
JSTL	JavaServer Pages Standard Tag Library
KAFRIBA	Kafue River basin
LAI	Leaf area index
LCH	*Laboratoire de Constructions Hydrauliques* (Hydraulic Constructions Laboratory)
LEW	Lumped elementary watershed
LS-SVM	Least-squares support-vector machine
LS-SVR	Least-squares support-vector regression
MAE	Mean absolute error
MaxAE	Maximum absolute error
MCMC	Markov chain Monte Carlo
MLP	Multilayer perceptron
MO	Multiple-objective
MOSCEM	Multi-objective shuffled complex evolution metropolis
MROK	Multiple-realization ordinary Kriging
MSE	Mean squared error
MVK	Moving vector with Kalman filter method
MW	Microwave
NARX	Non-linear autoregressive with exogenous inputs
NCDC	US National Climatic Data Center
NCEP	US National Centers for Environmental Prediction
NE	Nash-Sutcliffe efficiency coefficient
NOAA	US National Oceanic and Atmospheric Administration
NRMSE	Normalized root mean square error
NSGA-II	Non-dominated sorting genetic algorithm II
OK	Ordinary Kriging
OSPER	Online Support Platform for Environmental Research
ParaSol	Parameter solutions
PCA	Principal components analysis
PERSIANN	Precipitation estimation from remotely sensed information using artificial neural networks
PEST	Model-independent parameter estimation
PET	Potential evapotranspiration
PKR	Peak ratio
POM	Pattern-oriented memory
PR	Precipitation radar
PSO	Particle swarm optimization
RB	Relative bias
RBF	Radial basis function
RBFN	Radial basis function network
RFE2.0	Rainfall estimate 2.0
RIBASIM	River basin simulation model
RMSE	Root mean square error
SARIMA	Seasonal autoregressive integrated moving-average
SARMA	Seasonal autoregressive moving-average
SCE	Shuffled complex evolution
SCS	US Soil Conservation Service
SO	Single-objective
SOM	Self-organizing map
SROK	Single-realization ordinary Kriging
SRTM	Shuttle Radar Topography Mission
STREAM	Spatial tools for river basin environmental analysis and management

SUFI-2	Sequential uncertainty fitting – ver. 2
SVM	Support-vector machine
SVR	Support-vector regression
SWAT	Soil and Water Assessment Tool
SWAT-CUP	SWAT calibration and uncertainty procedures 2
TCI	TRMM Combined Instrument
TMI	TRMM microwave imager
TMPA	TRMM multisatellite precipitation analysis
TRMM	Tropical Rainfall Measuring Mission
ULisboa	University of Lisbon
USGS	United States Geological Survey
VAR	Variance ratio
VIS	Visible
VR	Volume ratio
WWF	World Wildlife Fund
ZRA	Zambezi River Authority
ZRB	Zambezi River basin

LIST OF SYMBOLS

For an easier reference, the symbols are associated with the document's chapters or sections through a color code:

General	S3.2.3	S3.2.4	S3.3	S3.4.2	S3.4.3	S3.4.4 & S3.4.5	C5	C6	C7

Lower case Latin letters

Symbol	Units	Description
af		Activation function in a MLP node.
a_r	[m³]	Regression intercept (used in the computation of streamflow volume after transmission losses; subbasin process).
a_x	[m³]	Regression intercept (used in the computation of streamflow volume after transmission losses; subbasin process).
$\mathbf{b}$		Vector of intercepts in the scope of RFBN models.
b_k^j		Bias added to the input signals to node k of layer j in a MLP.
b_r	[-]	Regression slope (used in the computation of streamflow volume after transmission losses; subbasin process).
b_x	[-]	Regression slope (used in the computation of streamflow volume after transmission losses; subbasin process).
c_p		Coordinates of point p .
c^{evap}	[-]	Coefficient for the estimation of channel evaporation rates.
co	[m²/m²]	Soil cover index.
cv	[kg·ha⁻¹]	Aboveground biomass and residue.
d	[m]	Main channel's depth.
$d_{k,p}$	[m]	Distance from point k to point p .
$d^{rsv,thr}$	[m]	Threshold water depth for upper flow to occur used in the computation of wetland and lake outflows.
e		Prediction error.
$esco$	[-]	Soil uptake compensation coefficient.
f		General function.
$\mathbf{f}$		Vector of radial basis functions.
fr^{deep}	[-]	Fraction of the main channel's infiltration that contributes to the deep aquifer.
fr^{rch}	[-]	Fraction of the time step in which water is flowing into the main channel.
$f_{soil}(z)$	[mm·day⁻¹]	Evaporative demand at depth z [m].
$\mathbf{h}_p$		Vector of hyperparameters controlling the POM regression.
h_t^{rsv}	[m.a.s.l.]	Elevation of the water surface at the wetland or reservoir during day t .
$h^{rsv,inv}$	[m.a.s.l.]	Elevation of the wetland of lake outlet's invert.

Symbol	Units	Description
$h^{soil,imp}$	[m]	Distance between the bottom of the soil profile and an impervious layer.
if		Input function of a MLP node.
k_r	[m^{-1}·km^{-1}]	Decay factor (used in the computation of streamflow volume after transmission losses; subbasin process).
k^{rsv}	[m^2·s^{-1}]	Flow coefficient used in the computation of wetland and lake outflows.
m	[m/m]	Slope of the main channel's banks.
m_c	[%]	The soil's clay contents.
n_{ch}	[m$^{-1/3}$·s]	Manning's roughness coefficient of the channel.
n_{sb}	[m$^{-1/3}$·s]	Manning's roughness coefficient for the subbasin.
of		Output function of a MLP node.
o		Degree of the polynomial trends in ARIMA / SARIMA models.
p		Autoregressive order of an ARMA model.
q		Moving-average order of an ARMA model.
r		Pearson correlation coefficient.
sc	[-]	Storage coefficient of the main channel.
s		Seasonality period of SARMA and SARIMA models.
s_{ch}	[m/m]	Average channel slope.
s_{sb}	[m/m]	Average subbasin slope.
$surlag$	[-]	Surface lag coefficient.
t_c	[hr]	Time of concentration.
t_c^{ch}	[hr]	Channel maximum travel time.
t^{dur}	[hr]	Duration of the runoff flow.
t^{lag}	[days]	Lateral flow travel time.
t_c^{ov}	[hr]	Overland maximum flow travel time.
t^{perc}	[hr]	Travel time for percolating water.
t^{travel}	[s]	Time of travel within the main channel.
u	[m·s^{-1}]	Mean flow velocity.
w	[m]	Main channel width.
$\mathbf{W}$		Vector defining the regression hyperplane in SVR and LS-SVR models.
w_1	[-]	Shape coefficient of the retention parameter correction (S^c).
w_2	[-]	Shape coefficient of the retention parameter correction (S^c).
w^{btm*}	[m]	First estimate for the main channel's bottom width.
$w^{btm,fld}$	[m]	Bottom width of main channel floodplains.
w		Kriging weight.
w_n		Weight associated to fitness value n in the scope of the calibration interface.
$\mathbf{w}_p$		Vector characterizing the POM regression at point p.
$\hat{\mathbf{w}}_p$		Satellite-based estimate of $\mathbf{w}_p$.
w_{ch}	[m]	Average width of the largest tributary channel in a subbasin.
w_{ik}^j		Weight of the connection from node i of layer $j-1$ to node k of layer j in a MLP.
w^{wtr}	[m]	Width of the main channel's inundated area.
x		Input / independent variable.
x^*		Input vector to machine learning streamflow forecast models.
x_k^j		Signal output from node k of layer j in a MLP.
y		Observed value (also used as output or the dependent variable in simple equations).
$\mathbf{y}$		Observed series.

Symbol	Units		Description
$\hat{y}$			Estimated value.
$\hat{\mathbf{y}}$			Estimated series.
$\mathbf{y}^{obs}$			Vector of point observations in the same time frame and in the vicinity of the point of interest (regressors for the POM interpolation).
$\mathbf{y}_p^{sat}$			Series of satellite observations at point p.
z	[m]		Depth within the soil profile.
z^{imp}	[m]		Depth of the first impervious layer in the ground profile.
z_i^{lower}	[m]		Lower limiting depth of the i th soil layer.
z_i^{upper}	[m]		Upper limiting depth of the i th soil layer.

Upper case Latin letters

Symbol	Units		Description
A	[m²]		Main channel's inundated area.
A_t^{rsv}	[m²]		Reservoir's surface area at day t.
A_{sb}	[km²]		Area of the subbasin.
B			Delay operator used in ARMA models.
$\mathbf{B}$			Matrix of biases in a MLP.
C			Capacity hyperparameter for SVM, SVR, LS-SVM, and LS-SVR models.
CN	[-]		SCS's curve number.
CN_k	[-]		SCS's curve number corresponding to antecedent conditions of class k.
E	[mm·day⁻¹]		Real evapotranspiration.
E^*	[mm·day⁻¹]		Remaining evaporative water demand (after E^{can} is deduced).
E_0	[mm·day⁻¹]		Potential evapotranspiration.
E^{can}	[mm·day⁻¹]		Contribution from canopy interception to evapotranspiration.
E^{rch}	[m³·day⁻¹]		Evaporation losses from the main channel.
$D_{(n,\text{length})}$			Observed series' length in the scope of the calibration interface.
$D_{(n,\text{magnitude})}$			Observed series' average value in the scope of the calibration interface.
E^{revap}	[mm·day⁻¹]		Real upflow from the shallow aquifer.
$E^{revap,max}$	[mm·day⁻¹]		Water reaching unsaturated layers upwards from the shallow aquifer through capillarity and evaporation.
E_t^{rsv}	[m³]		Evaporation from the reservoir at day t.
E^{soil}	[mm·day⁻¹]		Soil and snow evaporation/sublimation.
E_i^{soil}	[mm·day⁻¹]		Real evaporation from the i th layer of soil.
E_i^{soil*}	[mm·day⁻¹]		Estimate of the evaporation from the i th layer of soil conditioned on the layer's actual water content.
E_i^{soil**}	[mm·day⁻¹]		Uncorrected estimate of the evaporation from the i th layer of soil.
$E^{soil,pot}$	[mm·day⁻¹]		Potential evaporation from the soil.
$E^{soil,adj}$	[mm·day⁻¹]		Evaporation from the soil adjusted for plant water use.
E_t	[mm]		Evapotranspiration during day t.
E^{tra}	[mm·day⁻¹]		Plant transpiration.
$E^{tra,pot}$	[mm·day⁻¹]		Maximum transpiration on a given day.
$\mathbf{F}$			Matrix containing realizations of $\mathbf{f}$ for all patterns used for RBFN training.
H_t^{rsv}	[m³]		Other abstractions (adding to regular outflows) from the reservoir during day t.
H_t^{rsv*}	[m³]		Desired abstractions (adding to regular outflows) from the reservoir during day t.
I_a	[mm]		Initial abstractions (SCS curve number approach).
K			Kernel function in SVR and LS-SVR models.

Symbol	Units		Description
K^{rch}	[mm·hr⁻¹]		Hydraulic conductivity of the channel alluvium.
K^{rsv}	[mm·hr⁻¹]		Hydraulic conductivity of the bottom of the reservoir.
K_i^{sat}	[mm/hr]		Hydraulic conductivity for soil layer i.
LAI	[m²/m²]		Leaf area index.
L_{rch}	[km]		Length of the main reach within a given subbasin.
L^{hill}	[m]		HRU hillslope length.
L_{ov}	[m]		Subbasin slope length.
L_{sb}	[km]		Channel length from the most distant point of the subbasin to the outlet.
L_{tr}	[km]		Length of the longest tributary within a subbasin.
N^{layers}	[-]		Total number of soil layers.
O			Seasonal trend parameter in SARMA and SARIMA models.
P	[mm]		Precipitation depth (SCS curve number approach). See alternative use below.
P			Seasonal autoregressive parameter in SARMA and SARIMA models. See alternative use above.
P_{rch}	[m]		Main channel's wetted perimeter.
P_t^{rsv}	[m³]		Direct precipitation on the reservoir during day t.
P_t	[mm]		Precipitation during day t.
Q			Seasonal moving-average parameter in SARMA and SARIMA models.
$Q^{bank,in}$	[m³·day⁻¹]		Part of the of the main channel's infiltration that contributes to the river banks.
$Q^{bank,revap}$	[m³·day⁻¹]		Amount of water flowing upwards from bank storage through "revap".
$Q^{bank,return}$	[m³·day⁻¹]		Return flow from the main channel's bank storage.
$Q^{bank,revap,max}$	[m³·day⁻¹]		Potential "revap" from bank storage.
Q^{gen}	[mm]		Generated runoff (SCS curve number approach).
Q_t^{gw}	[mm]		Base flow during day t.
$\bar{Q}^{in}$	[m³·s⁻¹]		Average inflows to the main channel.
Q^{lat}	[mm]		HRU's lateral flow reaching the river network.
Q^{lat*}	[mm]		Generated lateral flow movement.
Q_i^{lat*}	[mm]		Generated lateral flow movement for soil layer i.
Q^{lost}	[m³·day⁻¹]		Loss rate from the main channel due to infiltration.
Q_i^{perc}	[mm]		Water percolating downwards from soil layer i.
Q_t^{rchrg}	[mm]		Combined recharge to shallow and deep aquifers at day t.
$Q^{rchrg,deep}$	[mm/day]		Recharge of the deep aquifer.
$Q^{rchrg,sh}$	[mm/day]		Recharge of the shallow aquifer.
Q_t^{seep}	[mm]		Percolation flow exiting the soil profile bottom during day t.
Q_{t-1}^{stor}	[mm]		Accumulated runoff in transit from day $t-1$.
$Q_{t-1}^{stor,lat}$	[mm]		Amount of lateral water in transit at the end of day $t-1$.
Q_t^{surf}	[mm]		Surface runoff during day t.
Q_t			Observed discharge at time step t.
$\hat{Q}_t$			Estimated discharge at time step t.
R	[MJ·m⁻¹·day⁻¹]		Extraterrestrial radiation.
R_{rch}	[m]		Main channel's hydraulic radius.
S	[mm]		Retention parameter (SCS curve number approach).

Symbol	Units	Description
S_1	[mm]	Retention parameter value for antecedent dry conditions (SCS curve number approach).
S^c	[mm]	Retention parameter corrected for water content.
S_k	[mm]	Retention parameter corresponding to antecedent conditions of a class k (SCS curve number approach).
SW_t	[mm]	Soil water content at day t.
SW^{awc}	[mm]	Soil available water capacity.
SW_i^{exc}	[mm]	Soil water contents available for percolation at layer i.
SW^{fc}	[mm]	Soil water content at field capacity.
SW^{sat}	[mm]	Soil water content at saturation.
SW^{wp}	[mm]	Soil water content at the wilting point.
T	[mm·day^{-1}]	Real transpiration.
$\bar{T}$	[°C]	Average daily temperature.
T^{max}	[°C]	Maximum daily air temperature.
T^{min}	[°C]	Minimum daily air temperature.
V_1	[m^3]	Initial volume stored in the main reach on a given day.
V_t^{in}	[m^3]	Volume entering the reservoir at day t.
V^{bank}	[m^3]	Volume of the main channel's bank storage.
V^{out}	[m^3]	Bulk volume of water leaving the main reach.
V_1^{out}	[m^3]	Volume leaving the main reach at the beginning of a given day.
V_2^{out}	[m^3]	Volume leaving the main reach at the end of a given day.
V_t^{out}	[m^3]	Volume being released from the reservoir at day t.
$V^{out,ds}$	[m^3]	Volume flowing downstream from the main channel at each subbasin.
V_t^{out*}	[m^3]	Revised reservoir outflow estimates at day t.
V_t^{out**}	[m^3]	Preliminary reservoir outflow estimates at day t.
$V_t^{out,lower}$	[m^3]	Lower flow component used in the computation of V_t^{out*}.
$V_t^{out,min}$	[m^3·s^{-1}]	User-specified monthly minimum reservoir outflow at day t.
$V_t^{out,upper}$	[m^3]	Upper flow component used in the computation of V_t^{out*}.
V_t^{rsv}	[m^3]	Volume held in the reservoir at the end of day t.
$V_t^{rsv,av}$	[m^3]	Water available in the reservoir, before the computation of discharges other human abstractions, at the end of day t.
$V_t^{rsv,av*}$	[m^3]	Corrected available volume in a controlled reservoir.
V_t^{seep}	[m^3]	Volume lost from the reservoir through seepage during day t.
$V^{sub,rch}$	[m^3]	Streamflow volume after transmission losses (subbasin process).
$V^{sub,rch*}$	[m^3]	Streamflow volume before transmission losses (subbasin process).
$V^{sub,rch,thr}$	[m^3]	Threshold value intervening in transmission losses computation (subbasin process).
V_t^{target}	[m^3]	User-defined reservoir monthly volume target.
$\mathbf{W}$		Matrix of connection weights in a MLP / matrix of linear regression coefficients in a RBFN.
W_t^{int}	[mm]	Volume of intercepted water at day t.
W^{sh}	[mm]	Water contents of the shallow aquifer.
$W^{sh,revap,thr}$	[mm]	Threshold minimum water content in the shallow aquifer for "revap" to occur.
$W^{sh,thr}$	[mm]	Threshold minimum water content in the shallow aquifer for groundwater flow to occur.

Symbol	Units		Description
$\mathbf{Y}^{sat}$			Matrix containing series of satellite estimates over the gauged points in the vicinity of the point of interest for POM interpolation.
$\mathbf{X}$			Matrix representing input data series.
$Z(\cdot)$			True random field in the scope of Kriging.
$\hat{Z}(\cdot)$			Estimated random field in the scope of Kriging.

Lower case Greek letters

Symbol	Units		Description
α^{bank}	[days]		Bank flow recession constant.
α^{gw}	[days]		Baseflow recession constant.
α^{rsv}	[m·s⁻¹]		Upper flow coefficient used in the computation of wetland and lake outflows.
α_{tc}	[-]		Fraction of daily rainfall that occurs during the subbasin's time of concentration.
β			Free parameter of the IDW model.
β^{deep}	[-]		Aquifer percolation coefficient.
β^{rsv}	[-]		Upper flow exponent used in the computation of wetland and lake outflows.
β^{rev}	[-]		"Revap" coefficient that defines $E^{revap,max}$ as a fraction of E_0.
β^{revap}	[-]		"Revap" coefficient for bank storage.
χ			Contribution of each observation to the fading-memory's filter estimate.
δ			Power in the flow-contributing area relation.
δ_{gw}	[days]		Ground water delay time.
δ^{target}	[days]		Parameter affecting the reaction time of controlled reservoirs when attempting to meet monthly targets.
ε			Half-width of a band in which errors are neglected.
$\phi(\cdot)$			Function that spans the feature space (SVR and LS-SVR models).
ϕ_i^{drain}	[-]		Drainable porosity of soil layer i.
γ			Variable characterizing the RBF width.
η	[m/m]		Evaporation coefficient at the reservoir.
λ	[MJ·kg⁻¹]		Latent heat of vaporization.
ρ_b	[g·cm³]		Soil bulk density.
σ			$1/\gamma$. Variable characterizing the RBF width.
$\boldsymbol{\theta}$			Vector of model parameters.
φ			Fading constant of the fading memory-filter.
$\mathbf{v}$			Coordinates of a RBF's center / training input for SVR and LS-SVR models.
π			Year of interest in the computation of the peak ratio error measure.
ω			Operator in ARMA models related to the autoregressive components.
ξ			Value of a given error measure function.
ψ			Operator in ARMA models related to the moving-average components.

Upper case Greek letters

Symbol	Units		Description
ΔQ_{max}			Range of observed discharges within a given time series.
$\mathbf{\Phi}$			Matrix representing model structure.
Θ			Domain of possible parameter sets.
Π			Number of years with valid peaks for the computation of the peak ratio error measure.

Symbol	Units	Description
Ω		Seasonal operator in SARMA and SARIMA models related to the autoregressive components.
Ξ		Aggregated fitness value within the calibration interface.
Ψ		Seasonal operator in SARMA and SARIMA models related to the moving-average components.
Z		White noise process.

1   INTRODUCTION

1.1   Background

It is only fair that the first words of the body of the manuscript acknowledge the Portuguese *Fundação para a Ciência e a Tecnologia* (Foundation for Science and Technology) and the *École Polytechnique Fédérale de Lausanne* (Swiss Federal Institute of Technology Lausanne, EPFL) for providing the lion's share of the funding and resources that made this research possible.

The research was hosted at the *Centro de Estudos de Hidrossistemas* (Centre for Hydrosystems Research, CEHIDRO) of *Instituto Superior Técnico* (IST), affiliated to the University of Lisbon (ULisboa), and the *Laboratoire de Constructions Hydrauliques* (Hydraulic Constructions Laboratory, LCH) of EPFL. Although a mostly individual endeavor, the work herein presented is also a product of the unique combination of skills and resources achieved by the establishment of the IST-EPFL joint doctoral initiative, of which it makes part.

The African Dams Project (ADAPT) was a multidisciplinary project set on enhancing the scientific basis for integrated water resources management in the Zambezi River basin (ZRB) (Mertens 2013). Financed by the Swiss Competence Center Environment and Sustainability (CCES), ADAPT covered topics as diverse as ecology, economics, biogeochemistry, hydrology and governance. The present work is closely related to the activities developed under ADAPT and attempted to take advantage of synergies with the project, resulting that although it was not directly funded by ADAPT, it greatly benefited from its financial, material and, most importantly, human resources.

Finally, the development of the work profited from some synergies with Théodora Cohen Liechti's Ph.D. research, entitled "Influence of dam operation on water resources management under different scenarios in the Zambezi River Basin considering environmental objectives and hydropower" (Cohen Liechti 2013), which was also developed at the LCH.

1.2   Motivation and objectives

Africa is a truly bewildering Continent. Full of potential, teeming with life, it has for some time hinged between explosive, all-promising development and economic, political, and humanitarian disaster. As this text is written, investment surges and it seems clear that the Continent is on the right path. Sadly, even on the right path, populations in this part of the World endure difficulties downright inconceivable by European standards and which is crucial to relieve. Given such a background, research that can contribute to alleviate the situation, whether directly or not, is particularly worthwhile pursuing.

Making for more pragmatic motivation for research related to Africa are the still too long lists of infrastructure and capacitation needs. In both fields research is needed, either to improve

efficiency, enhance effectiveness, better evaluate environmental impacts and rally public awareness, or to create, foster and enlarge local scientific communities.

From a foreigner's point of view some of Southern Africa can appear deeply anachronistic, with GSM[1] technology, cutting-edge cellphones, dirt roads, and slums appearing more often than not combined. From a researcher's perspective, such an environment requires adaptation and provides unique challenges and opportunities. With so much to be done, one might be led to believe it is easy to make a difference – it is not. Tales of efforts whose fruits have never been reaped abound and objectives should be defined with this in mind.

Africa is still to face remarkable and new challenges in the future. They are related to population growth, unprecedented migration towards urban centers, higher pressure on natural resources, and climate change. Water resources, in particular, are regarded has a pivotal issue in the stability of the region.

The Zambezi River and its main tributaries jointly cross several countries and are the basis for ecological, social and economic systems of great value. Notwithstanding, with few improvements in the management of water resources over the last decades (The World Bank 2010), the Zambezi River Basin (ZRB) still lacks an integrated water resources policy.

The present research focuses on the hydrology of the ZRB, where the riparian countries depend strongly on the river and main tributaries and, consequently, water resources issues have profound economic, ecological, social and political impacts.

The ultimate aim of the research was to prepare a hydraulic-hydrological model of the ZRB at a daily time scale with potential to be applied in parallel and subsequent researches being developed under ADAPT and directly by stakeholders, envisaging future use in dam operation optimization and synchronization, environmental impact assessments, evaluation of future scenarios (predicting responses to climate change and increased demands) and a broad range of other studies related to themes such as wildlife, water chemistry, sediment transport, and integrated water management. Being these very general aims, it became apparent to the author at an early stage that:

- constraining issues, mostly related to input data, would have to be addressed before the actual modeling stages;
- no single model could be a "best" choice for such a wide array of uses;
- a large emphasis would have to be placed on model calibration and validation; and
- in order to engage stakeholders – responsible for providing data and general knowledge about the basin – the research should produce appealing results.

Besides other features, better discussed in Chapter 2, one of the main characteristics of the ZRB is hydrological data scarcity. As will be described in Chapter 3, past modeling attempts often downplayed the importance of calibration and validation; partly, it is believed, due to the difficult access to information. One of the objectives of the research was, therefore, to study how to best overcome or alleviate such a constraint, either through the use of appropriate remote sensing alternatives or by developing new data processing strategies.

It was recognized that no hydrological model could perform adequately tasks as diverse as streamflow forecasting and environmental impact assessments. Having in mind the limited access to data and knowledge about the basin's hydrology, it was decided that the development of a traditional physically based semi-distributed hydrologic model should be complemented with more flexible machine learning tools. In fact, while the former should be capable of providing insights about the basin's behavior, return results at several points of interest and be fit for simulating alternative future scenarios, the latter have the potential to perform better at determinate

---

[1] Global System for Mobile Communications

locations and the advantage of requiring much less resources to set up, calibrate, validate, operate and maintain.

Another goal was to endeavor, likely for the first time, the calibration and validation of a basin-wide daily model from the upper catchment down to the delta. In order to achieve this, several "lesser" objectives were addressed, such as the development of a calibration strategy capable of coping with multi-site performance evaluation and parameter heterogeneity, maximizing the computational performance of the model, and preparing it to simulate features and processes that, while not particular to the ZRB, have there an unusual weight (e.g. relatively large wetlands and large reservoirs).

Finally, acknowledging the importance of organizing findings and publishing results, an online platform for data storage and sharing was developed. This platform, which can be visited at http://zambezi.epfl.ch, is likely to harbor real-time streamflow forecasts based on the findings of this and other ADAPT researches, as well as to enable the online modification and operation of the developed hydrological models in the near future. Going further than the present research, this effort was supported by the ADAPT project and plans are in place to boost its development and expansion.

1.3   General description of and structure of the document

By virtue of the broad-scoped objectives the research set out to attain, the contents of the document cover a wide range of subjects, going from the choice and pre-processing of appropriate remote sensing datasets to the development of an efficient calibration strategy for a basin-wide hydrological model; passing still in between by advances in machine learning streamflow forecasting tools.

Due to the vast size of the ZRB and the little existing knowledge about its hydrology – particularly at smaller scales – modeling was more often interpreted from a mathematical point of view than from a process-focused perspective. In practice, this led to a document that mostly drifts away from the description and exploration of physical hydrological processes. Albeit an arguable choice, this was the only way of keeping the research tractable and goal oriented. Indeed, it is believed that by placing too much emphasis on physical processes, the sheer scale of the problem, combined with the insurmountable lack of detailed data, would unavoidably lead to an undesired detachment between theory and practical results.

The statement of goals having been made and being given the justification for the nature of the developed research, the document is structured as follows:

- Chapter 2 provides details about the ZRB, including a physical description of its main subbasins and main features. Challenges and opportunities related to water resources where hydrological modeling is likely to play a pivotal role are also briefly laid out.
- Chapter 3 is devoted to literature review and focuses on four main themes: past hydrological modeling efforts on the ZRB; the Soil and Water Assessment Tool (SWAT – the chosen classic hydrological modeling tool); calibration and validation of hydrological models; and machine learning algorithms for regression.
- Chapter 4 provides a thorough introduction to data sources, with emphasis on Digital Elevation Models (DEM) and satellite rainfall estimates. The diversity among datasets, as well as their applicability to the ZRB is discussed.
- Chapter 5 describes a novel rainfall interpolation methodology, named Pattern-Oriented Memory (POM), which makes use of modern satellite rainfall estimates in order to interpolate historical rainfall areal maps, a crucial step in order to make the best use of the discharge series available for the calibration of the hydrological models.
- Chapter 6 focuses on the application of data-driven machine learning models for daily streamflow forecast. The chapter covers three related, but distinct topics. Firstly, different data-driven models are compared for several forecast lead times and input sets. In order

to convey rainfall information to the models, a novel approach to pre-process areal rainfall maps is introduced. Secondly, behaviors specifically associated to support vector regression models are shown to potentially lead to a substantial underestimation of extreme discharges. Finally, a new, flexible approach to produce empiric probabilistic forecast ensembles is developed.

- In Chapter 7 the preparation of the SWAT hydrological model of the ZRB, as well as changes made to its source code are discussed. Additionally, the development of the model calibration framework is described.
- Chapter 8 returns to the POM interpolation methodology by exploring one of its advantages beyond the good comparative performance with state-of-the-art techniques. In this chapter it is shown that models calibrated directly with satellite rainfall estimates can assimilate POM interpolated rainfall maps seamlessly or, at least, much better than maps interpolated resorting to other techniques such as Kriging.
- Chapter 9 dwells on the calibration of basin-wide daily SWAT models. Limitations of global calibration approaches are discussed, a cascading single-objective calibration process is analyzed, and multi-objective global and regional calibration schemes are introduced and debated. Issues such as model detail *vs.* computation time trade-off and the influence of the number of free model parameters on training performance and reliability are also focused.
- Finally, conclusions and prospects for future work are summarized in Chapter 10.

Appendix I contains information supporting the main text. Not comprising scientific work *per se*, put also a relevant part of this thesis' accomplishments, Appendix II is dedicated to a brief description of the ADAPT online database for data sharing and publication of results.

2   CASE STUDY

2.1   The Zambezi River basin

The ZRB is located in Southern Africa, being bounded by the 20°30'S and 9°S latitudes and 18°20'E and 36°25'E longitudes[1]. The Zambezi River, stretching over nearly 2 600 km, is the fourth longest river in Africa, after the Nile, the Congo and the Niger. The basin covers approximately 1 370 000 km[2] (The World Bank 2010), which is roughly the equivalent of 15 times the area of Portugal, 30 times the area of Switzerland, or 30% of the area of the 28 member states' European Union. It is home to approximately 30 million inhabitants (The World Bank 2010) and its area is shared among 9 riparian countries: Angola (18.3%), Botswana (1.2%), Democratic Republic of the Congo (0.1%), Malawi (8.3%), Mozambique (11.8%), Namibia (1.2%), United Republic of Tanzania (2.0%), Zambia (41.7%) and Zimbabwe (15.5%)[2]. The basin is positioned in Figure 2.1.
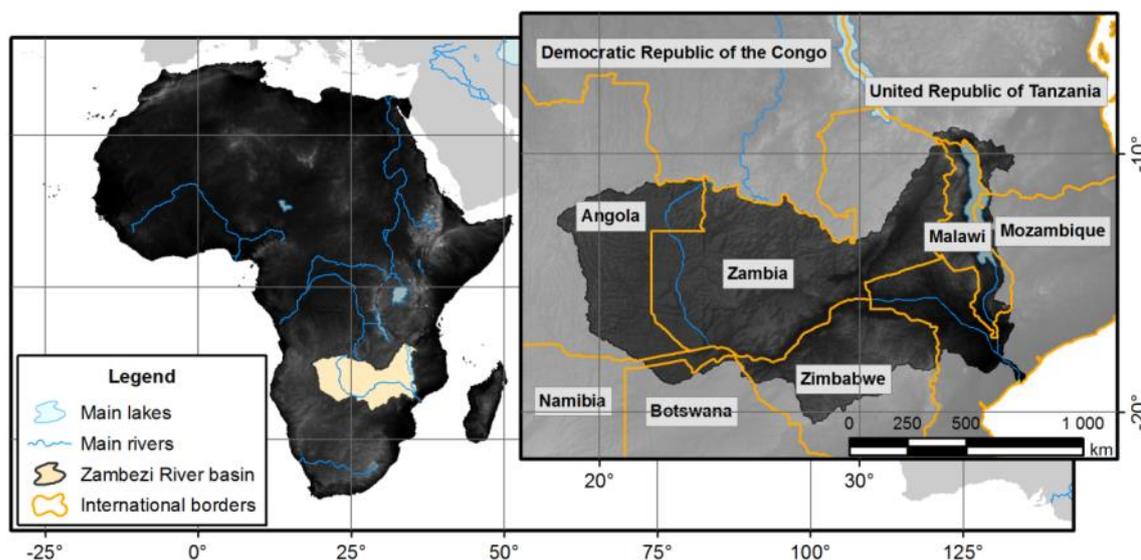
Figure 2.1. Location of the Zambezi River basin.

The ZRB encompasses humid, semi-arid and arid regions dominated by seasonal rainfall patterns associated with the Inter-Tropical Convergence Zone. The dominant rainfalls over the basin occur

---

[1] In the remainder of the text and contained figures coordinates are referenced in the WGS84 datum.

[2] Other area distributions have been estimated. An example is Angola (18.3%), Botswana (2.8%), Democratic Republic of the Congo (negligible), Malawi (7.7%), Mozambique (11.4%), Namibia (1.2%), Tanzania (2.0%), Zambia (40.7%) and Zimbabwe (15.9%) (Vörösmarty and Moore 1991).

during the Southern hemisphere summer (from October to April) and the winter months are generally dry (Cohen Liechti 2013).

Three distinct seasons can be identified in the Zambezi. A cold dry season with temperatures from 15 to 27°C lasts from May to September. From October to November a dry hot season where temperatures reach up to 32°C can be identified. Finally, a hot rainy season from December to April is characterized by high temperature and high humidity (Meier 2012).

The average yearly rainfall over the basin displays a high spatial variability and is of about 1000 mm/yr, being the potential evapotranspiration close to 2000 mm/yr (Meier 2012, Cohen Liechti 2013). The hydrology is not uniform, generally prone to higher rainfall rates in the northern regions. In some areas such of the Upper Zambezi and the Malawi Lake (see Section 2.2 for their location) rainfall can amount to as much as 1400 mm/yr, while in the southern part of Zimbabwe it can be as little as 500 mm/yr (The World Bank 2010).

Climate variations are particularly strong in the basin, although difficult to assess. An extensive analysis of Southern African climate by Tyson et al. (2002) revealed variability patterns with main components of 80 and 18 years (Cohen Liechti 2013).

The runoff is also affected by long cycles which have been reported to depend mainly on the rainfall cycles (Cohen Liechti 2013). Referring to the period from 1924 to 2004, Mazvimavi and Wolski (2006) estimated the period of the main runoff cycle to be of 40 years. Values for the mean annual discharge at the Delta vary appreciably. Examples are The World Bank (2010), which gives a precise figure of 4134 m^3/s, Tilmant et al. (2010), which point towards 3800 m^3/s, and Matondo and Mortensen (1998), that estimate 3251 m^3/s.

The river and its main tributaries are vital to the riparian populations from cultural and economic standpoints. They are sources of hydropower, havens of ecological diversity and essential for the region's food security. The ZRB is rich in natural resources. The main economic activities are fisheries, mining, agriculture, tourism, and manufacturing (The World Bank 2010).

The ZRB is one of the most valuable natural resources in Africa. However, it is also one of Africa's most heavily dammed river systems. Dam-induced ecological changes have already had consequences on wildlife and ecosystem-based livelihood of downstream residents (Meier 2012). The largest consumptive water user besides dams (evaporation through impoundment, approx. 13 km^3/yr) is irrigated agriculture (approx. 2 km^3/yr). Domestic water use amounts to approx. 1 km^3/yr and industrial water to about 0.2 km^3/yr. In total, consumptive water use is presently around 15-20% of total annual runoff (Beck 2010).

In the second half of the 20th century the political situation within the basin became highly unstable. Zambia and Malawi gained their independence from the British Empire in 1964. Following independence, both countries instated single-party systems which ruled until the 1990's. Zimbabwe declared its own independence in 1965 under a white minority rule. Independence however, was not to be granted until 1980, being Zimbabwe targeted by international sanctions in the midterm. In parallel, guerilla actions against the white minority rule were conducted eventually escalating into a civil war, between 1964 and 1979, until an agreement leading to majority rule was achieved. From 1982 to 1985 internal opposition led to further conflicts. In Angola, guerilla actions towards independence from Portugal started in 1961. Shortly after independence, in 1975, civil war broke out, ravaging the country from 1976 to 2002. Mozambique followed an identical path. Armed action against the Portuguese started in 1964, lasting until the independence, granted also in 1975. The political instability that followed degenerated into a civil war which took place from 1977 to 1992.

2.2 Physical description of the Zambezi River basin

2.2.1 Overview

A brief climatological, hydrological, economic and political introduction to the ZRB has already been made in Section 2.1. Here, a description of its main physical features is made.

As illustrated in Figure 2.2 the ZRB can be split into three main regions: the Upper Zambezi, the Middle Zambezi, and the Low Zambezi. Additionally, there are 13 main subbasins identified in Figure 2.3.
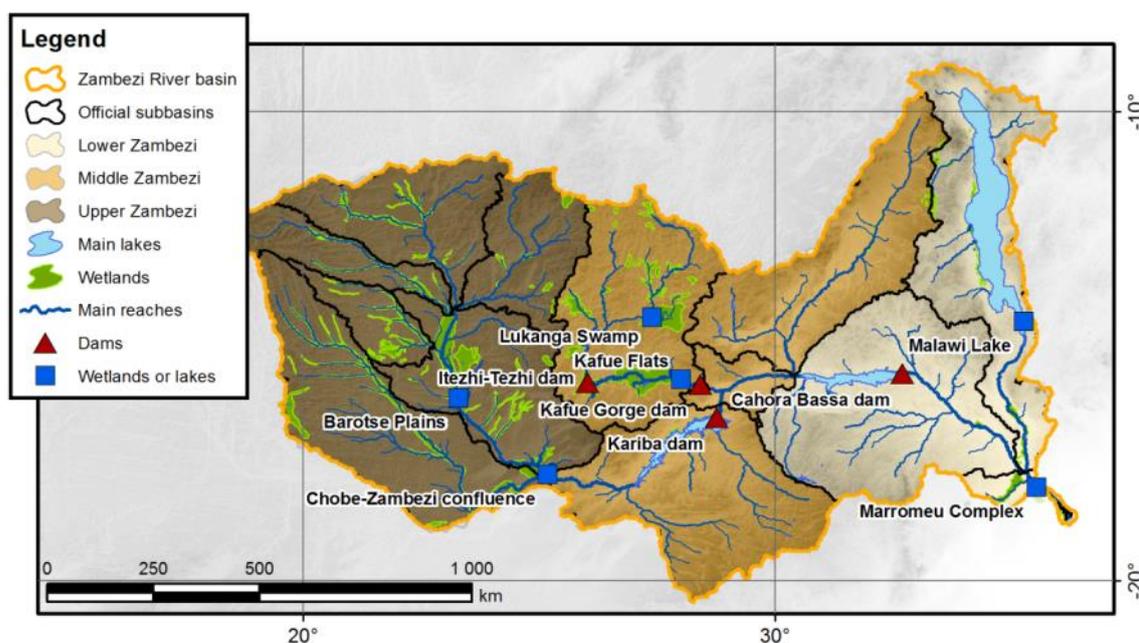
Figure 2.2. Main regions and features of the ZRB.

The Upper Zambezi is marked by steep slopes in the northern area and, going south, large wetlands such as the Barotse Plains. A distinctive feature of this part of the basin is also the Chobe-Zambezi confluence, where water can flow both ways depending on water levels.

The Middle Zambezi develops between the World-renowned Victoria Falls and the Cahora Bassa reservoir. To major affluents of the Zambezi (the Kafue and Luangwa rivers) join the main reach in this area. The Kariba dam and the Kafue hydropower system (Itezhi-Tezhi and Kafue Gorge dams), as well as the Kafue Flats, are its most noticeable features.

Finally, the Lower Zambezi is dominated by the Malawi Lake and the Cahora Bassa dam. In the lowlands near the outlet into the Indian Ocean, the Marromeu Complex wetlands and the Delta can be found.

The approximate areas and mean annual rainfall over the ZRB's 13 main subbasins of Figure 2.3 are shown in Table 2.1 and Table 2.2.

2.2.2 Main subbasins

Following, a brief description of the main subbasins is presented. In Appendix I.A detailed maps of the different subbasins along with the identification of their most relevant features are presented.

Upper Zambezi, Lungue Bungo and Kabompo

The Upper Zambezi, Lungue Bungo and Kabompo subbasins are the most upstream catchments of the ZRB. Due to high rainfalls, a great share of the Zambezi's runoff is generated in this area. With no major impoundments or wetland areas, streamflows are mostly unregulated. Two main discharge

gauging stations could be identified: Chavuma falls, in the Upper Zambezi (Figure 2.4) and Watopa Pontoon, in the Kabompo.
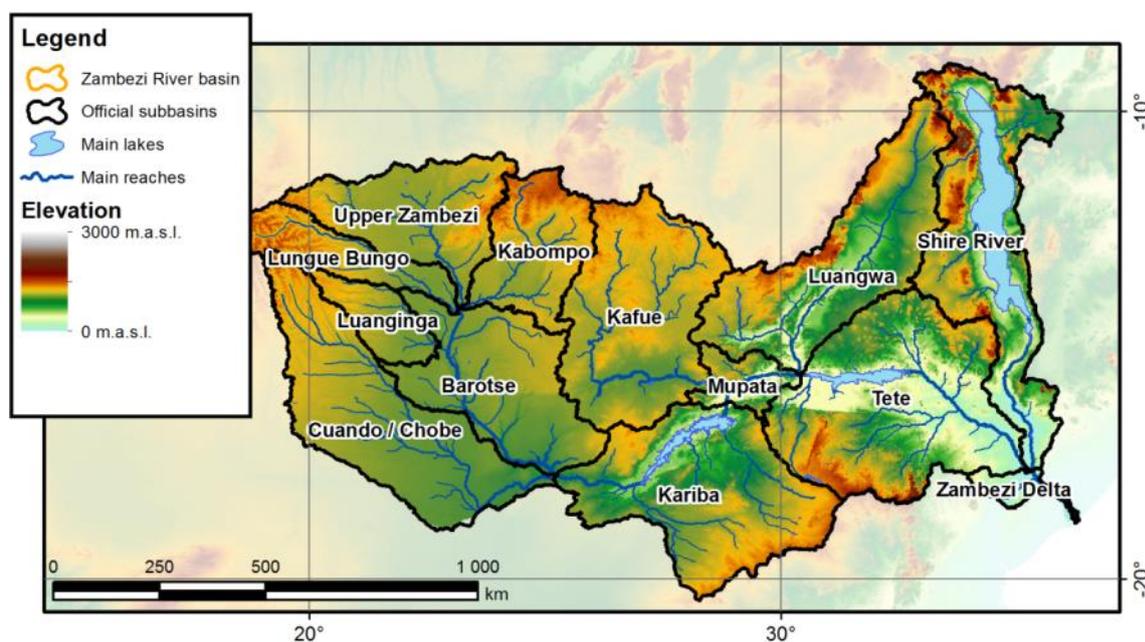
Figure 2.3. Elevation and main subbasins of the ZRB.

Table 2.1. Approximate areas of the Zambezi's main subbasins.

Subbasin	Area [km^2]
Barotse	116 000
Cuando / Chobe	155 000
Kabompo	72 000
Kafue	155 000
Kariba	165 000
Luanginga	33 000
Luangwa	153 000
Lungue Bungo	47 000
Mupata	25 000
Shire River	169 000
Tete	204 000
Upper Zambezi	94 000
Zambezi Delta	12 000

Table 2.2. Estimated mean annual rainfall over the Zambezi's main subbasins (The World Bank 2010).

Subbasin	Mean annual rainfall [mm]
Barotse	820
Cuando / Chobe	800
Kabompo	1 220
Kafue	1 040
Kariba	700
Luanginga	960
Luangwa	1 020
Lungue Bungo	1 100
Mupata	820
Shire River	1 120
Tete	880
Upper Zambezi	1 220
Zambezi Delta	1 060

Barotse and Luanginga

Moving downstream, the Zambezi passes through the Barotse subbasin and receives contributions from the Luanginga River. Here, the characteristics of the main reaches are substantially different due to the presence of large wetlands (the most notable being the Barotse Plains and the Chobe/Zambezi confluence). The main discharge gauging stations assessed for the region were Lukulu, immediately upstream of the Barotse Plains, Kalabo, near the outlet of the Luanginga River, Senanga, close to the outlet of the Plains, and Sesheke, between the Plains and the Chobe/Zambezi confluence.

Figure 2.4. Chavuma Falls (*Arthur Taute*, from Panoramio:
http://www.panoramio.com/photo/53299941).

Cuando/Chobe

The Chobe (also named Cuando or Kwando) is a very particular subbasin. Despite its large area, its contribution to the Zambezi is negative on average and highly dependent on water levels (Matondo and Mortensen 1998, The World Bank 2010). This happens due to the low slopes of the main stem of the Chobe River, which make most of its flow disperse into an inland delta. The delta itself is only linked to the Zambezi trough a narrow strip with another sizeable wetland at the confluence.

Kariba and Mupata

The Kariba and Mupata subbasins are mostly dominated by the Kariba reservoir – the largest in the World by volume – which heavily regulates the flows from upstream. The Kariba subbasin also marks the transition into the Middle Zambezi at Victoria Falls (Figure 2.5).

Most of the Zimbabwean share of the ZRB lies south of the Kariba reservoir. Although average annual rainfall is relatively low in the region, the difficult access to discharge series and the presence of numerous small impoundments renders the evaluation of actual runoff extremely difficult.

The Mupata subbasin, where the Kafue and Zambezi rivers meet, is largely ungauged area of reduced slopes where the river branches and widens.

Figure 2.5. Victoria Falls (on the background).

Kafue

The Kafue subbasin, in Zambia, is perhaps the most well-known part of the ZRB due to a large hydropower scheme and the Kafue Flats ecosystem. The Kafue river basin roughly divides into two major sections: the headwaters upstream of the Itezhi-Tezhi reservoir and the lower Kafue basin downstream of Itezhi-Tezhi. While at its headwaters the flow is still mostly unregulated, in the lower basin the hydrology is strongly influenced by the Itezhi-Tezhi and Kafue Gorge dams (Meier 2012).

The hydrological processes in the Kafue Basin are particularly complex with the influence of massive floodplains (the Lukanga and the Kafue flats) and the aforementioned dams (Cohen Liechti 2013). The subbasin is relatively well monitored, with long discharge series at key locations such as the Kafue Hook Bridge, upstream of the Itezhi-Tezhi reservoir.

Luangwa

The Luangwa River flows into the Zambezi just upstream of the Cahora Bassa reservoir. Its subbasin presents a relevant hydrologic modeling challenge due to being mostly ungauged[3], having a small storage volume and displaying quick response times to rainfall events (Meier et al. 2011). Despite having a limited contribution to the Zambezi's average discharge (about 500 m³/s (The World Bank 2010)), its quick hydrological response often leads to relevant high peaks. Consequently, the subbasin is of crucial importance to the operation of the Cahora Bassa dam.

Shire River

The Shire River's subbasin is marked by its most prominent feature: the Malawi Lake. Despite a high average annual rainfall in the subbasin, the Shire River contributes only with a relatively small flow to the Zambezi River. This is mainly due to the large lake's surface, where great water losses through evaporation take place (averaging almost 290 m³/s according to The World Bank (2010)).

Several gauging stations along the Shire exist, although the effects of mobile beds and backwater curves are hard to assess.

Tete and Zambezi Delta

The Tete subbasin, including the Cahora Bassa dam and reservoir, witnesses the fast descent of the river to low altitude and mostly flat regions. Discharges within the main reach of the Zambezi are hard to assess downstream of the dam due to mobile river beds and increased branching as the river approaches the Ocean (Figure 2.6).

Figure 2.6. Zambezi River near Tete (photo by Théodora Cohen Liechti).

2.2.3 Main dams

The main dams within the ZRB are briefly described below. In Table 2.3 their key characteristics are presented and, in Appendix I.B, their height-volume, volume-surface, and maximum discharge curves are presented.

---

[3] Records of two discharge gauging stations were used in this work. Notwithstanding, recent records are scares, conditioning most validation efforts.

Table 2.3. Characteristics of the main dams within the ZRB and associated reservoirs. Adapted from Cohen Liechti (2013).

Reservoir name		Commissioning year	Operation level [m a.s.l.]	Op. volume [km^3]	Reservoir area [10^{6}m^2]	Turbine capacity [m^3/s]	[MW]	Spillway capacity and associated water level [m^3/s]	[m a.s.l.]
Itezhi-Tezhi	Max.	1977	1030.5	6	380	160	-	4425	1030.5
	Min.		1006.0	0.78	90			402	1020.0
Extension		2013				312	120		
Kafue Gorge	Max.	1972	976.6	0.90	750	252	900	3600	978.0
	Min.		975.4	0.13	180			780	972.3
Kariba	Max.	1961	489.0	191	5627	1800	1470	9402	488.6
	Min.		475.5	116	5300			8502	484.0
North bank extension		2012-2014				430	360		
Cahora Bassa	Max.	1974	329.0	63.0	2974	2250	2075	15683	331.0
	Min.		295.0	12.2	838			6760	295.0

Kariba

The Kariba dam was completed in 1959. With a storage capacity of 180 km^3 (of which about 65 km^3 are active storage) and having a surface area of approximately 5500 km^2, it is among the great artificial reservoirs in the World, being the largest one by volume (Figure 2.7). The dam is managed by the Zambezi River Authority (ZRA) a binational company owned by the states of Zambia and Zimbabwe.

Figure 2.7. Kariba dam and reservoir (left photo by *Sean Ross*, from Panoramio: http://www.panoramio.com/photo/65334783; right image from Google Earth – approx. 210 km across).

The mean annual runoff in the section of the dam is of about 1300 m^3/s (Meier 2012). The dam has a considerable installed capacity of 1450 MW, which equates to a discharge of 1800 m^3/s. Its spillway capacity is close to 9500 m^3/s. Due to the development of a scour hole which threatened the dam's stability, however, the reservoir has not been operated near its maximum capacity for some time[4]. Consequently, since the 1980's operations have mostly focused in reducing the use of the spillways (Cohen Liechti 2013). The reservoir has large impacts on the flows downstream, greatly reducing seasonal variability (Beilfuss and Dos Santos 2001, Matos et al. 2010)

---

[4] The development of this scour hole, as well as measures to control it and safeguard the dam's operation were recently subject to an indepth study by EPFL's Laboratody of Hydraulic Constructions (LCH).

Cahora Bassa

The Cahora Bassa dam (Figure 2.8) was completed in 1974 with the primary objective of exporting power to South Africa. It is managed by the Mozambican state-owned company *Hidroeléctrica de Cahora Bassa* (Cahora Bassa Hydroelectric, HCB). However, due to the civil war, the transmission line was destroyed and for 20 years, nearly no electricity was produced. This continued until the hydropower station was reactivated in 2000 (The World Bank, 2010).

Figure 2.8. Cahora Bassa dam and reservoir (left photo by *Théodora Cohen Liechti*; right image from Google Earth – top facing East, approx. 150 km across).

Having a 60 km^3 reservoir, with 51 km^3 of net storage capacity, the dam can presently produce up to 2075 MW of electricity, which corresponds to a flow of roughly 2250 m^3/s – very close to the average discharge. The spillway's capacity is of about 14 000 m^3/s.

Plans to safely rise the reservoir's operating levels through additional spillways are being developed. In parallel, the development of an additional powerhouse is expected to increase the dam's installed capacity to approximately 3000 MW.

Itezhi-Tezhi and Kafue Gorge

The Kafue hydropower scheme, comprised by the Itezhi-Tezhi and Kafue Gorge dams, is quite singular. The Kafue Gorge dam, downstream, takes advantage of a large hydraulic head but has no substantial storage capacity. This storage is provided by the Itezhi-Tezhi reservoir's 6 km^3 which, in spite of comprising a considerable volume in absolute terms, pale in comparison to the volume of the Kariba or Cahora Bassa reservoirs. Amidst Itezhi-Tezhi and Kafue Gorge lay the Kafue Flats, a wetland area of enormous ecological value that stretches for approximately 200 km.

The Itezhi-Tezhi dam (Figure 2.9) was completed in 1978 and, being mainly built in order to provide extended storage capacity for the Kafue Gorge dam, did not have turbines for electricity production installed until 2010. Itezhi-Tezhi is especially designed to allow managed flood releases to maintain a certain flooding in the floodplain downstream. However, the operation rules of the dam are very strict and usually the benefits of increased power production were chosen in detriment of a distinct flood release for the floodplain. The Kafue Gorge reservoir was built in 1972 immediately downstream of the Kafue Flats. It has an installed capacity of 900 MW and plans exist to extended it to 990 MW (Meier 2012).

2.2.4 Malawi Lake and the main wetlands

Malawi Lake

The Malawi Lake (Figure 2.10), also known as Nyasa or Niassa Lake, is by far the largest waterbody in the ZRB. It lies at 474 m above sea level in the African rift valley, with a surface area of 28 000 km^2, a volume of 8000 km^3, and a length of 550 km (Jury and Gwazantini 2002). Its bottom goes well below sea level with depths reaching 706 m. The lake is remarkable for the unusually high ratio between its surface and that of the contributing catchments, which is close to 1/3.

Figure 2.9. Detail of the Itezhi-Tezhi dam (photograph by *Richard Beilfuss*, from the ADAPT webpage: http://www.cces.ethz.ch/projects/nature/adapt/itezhitezhi.jpg?hires.

Figure 2.10. Malawi Lake (left picture by *Bene Zoltan*, from Panoramio: http://www.panoramio.com/photo/7085360; right image from Google Earth – approx. 480 km across).

Barotse Plains

The Barotse Plains (Figure 2.11) are the most distinctive feature of the Upper Zambezi. Having a maximum width over 30 km, this wetland covers an area of 7500 km^2. Being that underlying formation consists mainly of Kalahari sands, which form an enormous ground water reservoir (Winsemius et al. 2006a, Moore et al. 2008, Meier 2012). The plains substantially affect the shape of the annual streamflow hydrographs.

Figure 2.11. Details of the Barotse Plains (left picture by *Petri Viljoen*, from Panoramio: http://www.panoramio.com/photo/19578337; right image from Google Earth – approx. 32 km across).

Chobe-Zambezi Confluence

The Chobe-Zambezi confluence (also named Chobe Swamps, Figure 2.12) does not exert an influence on the Upper Zambezi's hydrographs comparable to that of the Barotse Plains. Nonetheless, it is interesting due to the complex hydraulics. Depending on water levels, flow from the Chobe River can contribute to the Zambezi or, conversely, a portion of the Zambezi can flow into the Chobe. Net contributions are from the Zambezi into the Chobe and have been quantified at 20 m^3/s (Matondo and Mortensen 1998). A substantial reduction of peak discharges, however, can be observed in this area.

Figure 2.12. Detail of the Chobe-Zambezi Confluence in May 2010 (from Wikimedia Commons: http://commons.wikimedia.org/wiki/File:Zambezi_Flood_Plain,_Namibia_(EO-1).jpg – approx. 39 km across).

Kafue Flats

The Kafue Flats are an extensive, 250 km long and up to 90 km wide wetland, which spreads from downstream of Itezhi-Tezhi down to the Kafue Gorge reservoir and covers an area of 6500 km^2. The Kafue river meanders through this extremely flat area with an average slope of 3 cm/km, hence the travel time of the water through the Flats is up to two months (Meier 2012).

The combined effect of the Itezhi-Tezhi dam and the Flats on the attenuation of annual flood is potentially complete (removing seasonal fluctuations). Because of the ecological value of the wetland and the fact that many endogenous species depend on the seasonal flood cycles, current Itezhi-Tezhi dam operations attempt to recreate a yearly artificial flood.

Figure 2.13. Detail of the Kafue Flats near Monze (*Janus Jansen*, from Panoramio: http://www.panoramio.com/photo/22050168).

2.3   Challenges and opportunities

Several paths for improvement that are associated with water resources can be identified in the ZRB. To greater or lesser degree, endeavors to seize such opportunities would benefit from sound hydrological models upon which to base decisions.

In spite of the already considerable installed capacity put in place (almost 5000 MW mostly located at Cahora Bassa, Kariba and Kafue Gorge, but also present in several smaller schemes),

there is an acknowledgeable potential for it to be expanded. The World Bank (2010) assed a total hydropower potential approaching 13 000 MW.

Of the many proposed schemes, the following can be highlighted due to their capacity and stage: the Mphanda Nkuwa run-of-river scheme, 70 km downstream Cahora Bassa (2000 MW), the Batoka Gorge dam, just upstream of Kariba (1600 MW), the Kafue Gorge Lower dam (downstream of Kafue Gorge, 750 MW), the Kariba extension (660 MW), and the Cahora Bassa North Bank (850 MW) (The World Bank 2010).

Related issues are the coordinated operation of the hydropower facilities, which would yield substantial economic gains (The World Bank 2010, SWRSD Zambezi Basin Joint Venture 2011), as well as the drive towards a more demanding operation of the existing reservoirs by virtue of higher demands, powerhouse extensions, and the expected development of more competitive power markets in the region[5].

Another pressing issue is flood management, which affects several regions of the basin such as the Barotse Plains, the Kafue Flats, the Chobe-Zambezi confluence and the Delta. Although some populations of the ZRB are culturally adapted and relatively resilient to floods, they are still vulnerable owing to difficult communications and transport, the existence of few shelters, and reluctance to leave the affected areas with fear of losing their little possessions (SWRSD Zambezi Basin Joint Venture 2011).

The hydropower sector is by far the highest water user due to the evaporation from the larger reservoirs (Euroconsult Mott MacDonald 2007). The total annual evaporation from the existing reservoirs has been valued at 13 to 17 km^3, of which 85% are lost by the Kariba and Cahora Bassa reservoirs alone. In total, consumptive water uses amount to 15-20% of the total annual runoff. In addition to hydropower production, this relatively moderate water consumption, compared to the river flow today, offers many development opportunities for irrigated agriculture (Beck 2010, Cohen Liechti 2013).

One problem with such a development is that the more densely populated areas with bigger needs are located in the medium to low rainfall areas. The asymmetry between water availability and population density and the density of economic activity is likely to become more pronounced in the future, especially with growing consumptive agricultural use in some parts of the basin. This heterogeneity and uneven expansion of water use is likely to become a source of conflict within and among the eight riparian countries (Beck 2010).

Growing demands on water resources, motivated by population increases and developments in economic activity, will increase the pressure imposed on ecologically valuable areas. The definition of ecological flows is, therefore, a major and difficult issue. Many riparian ecosystems depend on seasonal flood patterns of complex definition. Also, due to the yield losses that such environmental flows can potentially force upon the hydropower sector, the topic is sensitive not only technically but also politically.

Finally, better management and cooperative development of the basin's water resources could significantly increase agricultural yields, hydropower outputs, and economic opportunities. Collaboration has the potential to increase the efficiency of water use, strengthen environmental sustainability, improve regulation of the demands made on natural resources, and enable greater mitigation of the impact of droughts and floods (The World Bank 2010).

The assertion of the long term success of new hydropower schemes can only go through hydrological modeling. Similarly, it is only through the proper evaluation of future scenarios that an adequate balance between hydropower, irrigation and the preservation of natural ecosystems

---

[5] Today, the Kariba and Cahora Bassa dams operate nerly at full load as the turbining capacity is similar to the mean discharge. With increased hydropower generation capabilities, however, operators will necessarily face added complexity regarding reservoir management.

can be achieved (Cohen Liechti 2013). On another plane, joint reservoir management, the planning of environmental flows, flood forecasting, and flow control depend on sound hydrologic models as well.

Technically, hydrological modeling of the ZRB presents specific difficulties. Perhaps the most relevant are the basin's size and heterogeneity. If, on the one hand, such a large basin has a "smooth" hydrological response where the effects of particular features are averaged out, on the other hand, size and heterogeneity are in the way of a detailed understanding of the hydrological processes taking place and both put into question the validity of lumped approaches and pose acknowledgeable computational difficulties to distributed models.

Another challenge is presented by the difficult access to hydrological data; firstly, due to the relatively low investment that has been made on measurement infrastructure – a legacy of the basin's recent political history – and, secondly, because the lack of a consolidated supranational interlocutor makes the access to existing data a time consuming and intricate process.

A source of concern is the complex hydraulics in some of the basin's wetlands which, while having acknowledgeable non-linear effects on streamflow, are not easily reproduced within many established hydrological modeling frameworks. Additionally, dams such as Kariba, Cahora Bassa and Itezhi-Tezhi can condition downstream flows to a large extent; particularly so if small time scales are analyzed. In particular, given the author's trouble in recognizing the influence of provided rating curves on historical discharges, in order to successfully model the whole basin these infrastructures' outflow series had to be explicitly accounted for.

The predominant semi-arid climate of the ZRB should not be neglected. In fact, semi-arid climates have usually higher spatial rainfall variability; transmission losses are greater than those observed in temperate regions; and seasonal differences in vegetation cover have a larger influence on the hydrological response (Pilgrim et al. 1988).

Finally, the ZRB is marked by climate variations of appreciable amplitude that affect rainfall and runoff. If accounting for such variations can be particularly challenging in light of the only recent history of hydrometric and meteorological measurements in most parts of the basin, in practical terms they raise questions about the validity of analyses based on relatively limited periods of time.

3   LITERATURE REVIEW

3.1   Hydrological modeling in the Zambezi River basin

3.1.1   Introduction

Given its size, importance to local populations, and potential – taken in a wide context, from economical to ecological – the Zambezi has been the subject of only a small fraction of the studies one would expect were the basin located in Europe or North America. The oldest relevant hydrological studies focused on the ZRB have mostly been associated with the planning and construction of the major impoundments such as Kariba or Cahora Bassa. These were mainly consulting works, partly affected by the lack of historic hydrological data.

The troubled politics and conflicts that scared the second half of the 20th century in the region, along with plummeting to stagnant economies, had large impacts on the investments in water resources, on the collection of hydrological data, on the number of local consultants and researchers, and on the interest of the international research community.

The resolution of the armed conflicts and a more stable, more open, political phase spurred renewed international interest in the area. Several development aid agencies and other institutions such as The World Bank and the World Wildlife Fund have recently been sponsoring initiatives in the field of water resources, whether financing reports and raising awareness to strategic issues or promoting exchanges and capacity building. This state of affairs has, naturally, generated interest on the part of researchers, all the more given the standing issues of new hydropower schemes to be developed, challenges of dam synchronization and increasing pressure on natural ecosystems, just to name a few.

Below a brief description of the most relevant hydrological models developed on the Zambezi to date is endeavored. Several model types, corresponding to different aims and constraints, are presented. With goals as distinct as broad future scenario assessments or real-time flood forecasts, they range from simple lumped to complex distributed approaches adapting daily to monthly time steps. This description, necessarily incomplete, is mainly built upon information presented on the works of King (2012) and Katiyo (2012), on the SWRSD Zambezi Basin Joint Venture (2011) report, and on own review. A similar version of this text is contained in the *Recommendations on modeling for the implementation of environmental flows project in the Zambezi River basin* report (Juízo et al. 2012), part of the Joint Zambezi River Basin Environmental Flows Programme, commissioned by the Worldwide Wildlife Fund and other partners (the author participated in the report's preparation). Being the Zambezi a sizeable basin, some of the models worth referencing are limited in their extent. As such, adding to basin-wide models, the description of local works is split among Upper, Middle and Lower Zambezi.

Overall, one of the major concerns among authors is model calibration and validation; a problem mostly bound to short and unevenly spatially distributed discharge time series. In the majority of the models, particularly so those developed with research purposes, remote sensing data sources are used extensively. Also striking is the focus of most works, which seems to gravitate towards the large impoundments and the regions where more data is available, virtually neglecting some other hydrologically relevant areas of the basin.

3.1.2   Basin-wide modeling

Given the specificities of the different areas of the ZRB and the lack of data (which adds to the need of retrieving it separately from several stakeholders), basin wide modeling can become a daunting challenge. Despite this, several researches and consultants have undertaken this effort in the past.

Vörösmarty and Moore (1991) proposed an early distributed model for the whole Zambezi. Conclusions of the work outlined the challenge of quantifying seasonal discharges and clearly define the lack of reliable historical discharge records as a major issue, conditioning model validation.

Later, water balances of the Zambezi Basin have been carried out by Denconsult (1998) and Matondo and Mortensen (1998), with results differing by as much as 50% in the runoff estimates at the delta. Among both studies, discrepancies can be identified even on the definition of sub-basins, evidencing the challenges faced.

Of late several master theses also attempted to model the whole Zambezi. Examples are the work from Gerrits (2005), who applied the Spatial Tools for River Basin Environmental Analysis and Management, STREAM (Aerts et al. 1999), and Zambezi-script models in order to evaluate the feasibility of water balances in the basin using gravity disturbances data. The calibration was carried out only for the western part of the basin and only based on two locations. Uncertainty analysis was performed using the Generalized Likelihood Uncertainty Estimation (GLUE) methodology. Michailovsky (2008) also worked with gravity disturbances data, using the Soil and Water Assessment Tool (SWAT). Calibration was performed manually, not having been verified in all the available discharge locations, namely due to the uneven distribution of the considered gauging stations. Landert (2008) also used SWAT, but with the purpose of modeling biochemistry. Emphasis was placed on model calibration, done resorting to the SWAT Calibration and Uncertainty Procedures software bundle (SWAT-CUP) 2 software (Abbaspour 2012). Although the calibration was once again only accomplished for a restricted and unevenly distributed set of subbasins, conclusions raised the issues of varying sensitivity to parameters with space and time and of the difference in behavior between wet and dry conditions. Perrin (2013) modeled irrigation based on outputs from a SWAT model calibrated by Cohen Liechti (2013), having difficulty interpreting the impacts of distinct water demand scenarios.

A set of Ph.D. researches also include detailed models for the whole Zambezi Basin. Meier et al. (2011) worked on three watersheds: the Upper Zambezi, the Kafue and the Luangwa. That study introduced the novelty of using remotely sensed soil moisture data and including a real-time forecasting application, using a conceptual model updated by an Ensemble Kalman Filter (EnKF) to that end. Michailovsky (2013) addressed the issue of insufficient discharge data by resorting to satellite radar altimetry in order to derive discharge estimates at key ungauged locations (Michailovsky et al. 2012). Also, with the goal of producing operational reservoir inflows at Kariba and Itezhi-Tezhi, the same author applied SWAT for simulating the rainfall-runoff process and coupled it to the Muskingum routing scheme and a custom floodplain model in order to represent transport. Updates were made using an Extended Kalman Filter (EKF) (Michailovsky and Bauer-Gottwein 2013). Calibration of the SWAT model was manual, being that results were satisfactory for Itezhi-Tezhi but lacking accuracy for Kariba due to the complex influence of the Barotse plains.

In a work closely related to the present one, Cohen Liechti (2013) addressed the challenge of calibrating a detailed model of the whole ZRB with the application of A Multi-Algorithm Genetically

Adaptive Multi-objective (AMALGAM) optimization algorithm to records of river discharge and stored volumes in reservoirs from 1998 to 2003. Having obtained a non-homogeneous performance, Cohen Liechti underpinned uncertainties in the input and calibration data as a major challenge, highlighting also the model's structure and the lack of knowledge about floodplain processes as sources of uncertainty. Using longer series for calibration was pointed out as a possible strategy to improve results.

Another set of models have been developed to address IWRM issues such as optimizing water allocation, analyzing investment opportunities, future scenarios, or operation synchronization. Of major interest is the *Multi-sector Investment Opportunity Analysis* by The World Bank (2010), where a modified version of the HEC-3 model was applied. Tilmant et al. (2010) focused on optimal water allocation in the basin, having used a semi-distributed conceptual model. Beck (2010) developed a simplified hydrological model of the basin with the objective of quantifying water availabilities in future scenarios. Another relevant work, *The Dam Synchronization and Flood Releases in the Zambezi River Basin Project* (SWRSD Zambezi Basin Joint Venture 2011) used spreadsheets and HDAM graphs developed by the WRNA consultant to conduct hydrological and statistical analyses.

3.1.3   Upper Zambezi
The Upper Zambezi Basin, going from the headwaters of the river in Zambia and Angola to Victoria Falls, is the less studied part of the basin. There are no major dams in the area and the Barotse plains, along with the Chobe-Zambezi confluence, are its most distinctive features.

There is no knowledge of operational hydrologic models in use by the main stakeholders that target specifically the Upper Zambezi Basin. Besides the regression model in use by the Zambezi River Authority (ZRA) to forecast inflows into Kariba reservoir, which obviously depends heavily on the Upper Zambezi Basin but is technically a Middle Zambezi Basin model, the work By Winsemius et al. (2006a), comparing the Lumped Elementary Watershed (LEW) and STREAM models, can be highlighted. The work was however, focused on a broad comparison of model outputs with seasonal gravity measurements and, as such, should not be directly applicable to the simulation of daily discharges.

3.1.4   Middle Zambezi
The Middle Zambezi Basin stretches from Victoria Falls down to Cahora Bassa dam. It comprises the Kariba reservoir and the Kafue hydropower system. Along with these infrastructures, the Kafue Flats wetland and the Luangwa River, which reaches the Zambezi just upstream of Cahora Bassa, can be highlighted for their hydrological significance.

In terms of water resources, the Middle Zambezi is by far the most studied area of the basin. Within it are three distinct regions of interest: the Kafue Flats on the North, the Zambezi River and Kariba dam on the South, and, to the East, the practically ungauged Luangwa River.

The "Kafue River Basin" (KAFRIBA) is an historical model for the Kafue Flats (DHV 2004). It was based on RIBASI, an earlier model also developed by DHV (1980). It uses inflows at Itezhi-Tezhi estimated through the Pitman model (Mwelwa 2004), followed by the calculation of water levels at the reservoir. It is also able to compute the flood propagation through the Flats and compare the situation with a simulation of the "natural" flood event, removing the influence from the reservoir. Today, the KAFRIBA model is not in use due, namely, to the lack of confidence by stakeholders. Aduah (2007) used remote sensing to evaluate and quantify errors in the model's flooded area estimates, reaching mean values of 18%. DHV, Mwelwa (2004) and Aduah (2007) highlighted the issue of model calibration, both in the Flats and in the Upper Kafue.

Work on the Kafue Flats continues today. Of particular relevance is the 2-D Info Works RS model developed by the Institute of Hydrology of Wallingford to be used by ZESCO and the Department of Water Affairs of Zambia. Along with this new hydraulic model, a fine resolution DEM of the Kafue Flats was commanded, being now in the latest stages of validation. Research was also conducted

under the scope of the ADAPT, namely by Meier (2012) and Köck (Meier and Kinzelbach 2010, Köck et al. 2011).

Other models considered for the Kafue Flats were HEC ResSim, HEC-RAS, HEC-3 and HYDRO-PC (Katiyo 2012).

Chen (2010) applied the River Basin Simulation Model (RIBASIM) in order to study the flow releases from the Itezhi-Tezhi reservoir, but with limited interest. As the author states: *"(…) it was not the intention to develop a flow forecast, improve the operational management or execute an environmental flow assessment for the local situation of the Kafue River (...). This reduces the significance of the case study results for the Kafue River."*

Less work was identified focusing on the Zambezi River. The Zambezi River Authority (ZRA) makes use of regression and HEC-3 models in order to forecast inflows into the Kariba reservoir. Also of notice are the past attempts to use MIKE BASIN which *"were not successful because of some installation problems which seemed mainly to stem from issues of capacity in personnel"* (Katiyo 2012).

Also in Katiyo (2012) there is reference to the master's thesis of Ekandjo, entitled "Hydrological analysis of the Middle Zambezi and impacts of the operation of hydropower dams on flow regime in the Mana Pools National Park". In this work, a rainfall-runoff model based on the U.S. Soil Conservation Service's (SCS) equations was applied in order to estimate contributions from six ungauged catchments into the Mana Pools.

Inflows into the Cahora Bassa reservoir are conditioned by operations at the Kariba dam and the Kafue hydropower system. An additional – very relevant – inflow is the unregulated discharge coming mostly from the Luangwa catchment. The managing company of the reservoir, *Hidroeléctrica de Cahora Bassa* (HCB), uses a spreadsheet hydrological model to aid the decision process regarding dam operations.

Research work has been carried out focusing on the Luangwa subbasin (Winsemius et al. 2009, Michailovsky et al. 2012, Michailovsky 2013). Indeed, being mostly ungauged and having a fast hydrologic response, the Luangwa poses significant challenges to modelers. Here, efforts are being made on the use of remote sensed hydrological data, such as rainfall, evaporation and soil moisture, but also water levels, in order to overcome the lack of a well distributed network of discharge gauging stations.

Proceeding downstream, Beilfuss (2001) modeled inflows to Cahora Bassa as part of a broad effort to prescribe flood releases from the reservoir in a report that served as a reference for several posterior assessments. In this work a HEC-5 model was used, assimilating inflows from the Upper Zambezi and Upper Kafue Catchments, and incorporating the Kafue Hydropower System and Kariba Dam.

3.1.5   Lower Zambezi

The Lower Zambezi Basin spreads from Cahora Bassa to the delta in Mozambique. Here, the flow is heavily dependent on the discharges from the Cahora Bassa dam, but a contribution from the Shire River, bringing water from the Malawi Lake, is also important.

Being the flows, either coming from Cahora Bassa or the Malawi Lake, essentially regulated, the main efforts in this lower part of the ZRB fall towards flood routing and the forecast of flooded areas. The most relevant model here is perhaps the version of MIKE FLOOD in use by the *Administração Regional de Águas do Zambeze* (Water Regional Administration of the Zambezi, ARA-Zambeze).

Additional efforts were undertaken by Ronco (2008) and Ronco et al. (2010), who focused on sediment transport, and Beilfuss (2001), whose landmark work, presented for the Middle Zambezi Basin, also includes inflows from the Moravia-Angona and Manica Plateaus and the Shire River that contribute to a flood routing model for the lower reaches of the Zambezi.

This description covers no models of the Shire River subbasin. The lack of interest in the area might partly be due to a difficult access to sound historical discharge records. Also, the substantial flow regulation effect that the Malawi Lake possesses, given the unusual ratio between its surface and the area of the contributing basin, might be a justifying factor.

No account of researches addressing the modeling of the Zambezi River's hydrology downstream of Cahora Bassa was made. Notwithstanding, the development of Tete's mining industry and the studies related to the planned Mphanda Nkuwa dam motivated some consultancy works which might be a motivation and a basis for such efforts in the future.

3.2 The Soil and Water Assessment Tool

3.2.1 Context

Today, a plethora of hydrological models exist, being SWAT but one within their ever growing list (e.g. Moriasi et al. 2012). In order to position it in relation to other models, an effort of classification according to key characteristics is made below.

Domain

On a conceptual level, a model's domain is a relevant characteristic. Common models can be focused on hydrologic or water quantity issues alone, encompass water quality sub-models, or account for a whole set of ecological, sociological, economic and political aspects. SWAT encompasses water quantity and quality.

Goals

Regarding model goals, a basic distinction can be made between models whose goal is to accurately simulate observed phenomena and models that aim to perform optimization tasks. In general, simulation and optimization-oriented models differ substantially in complexity, being that it is not uncommon that some simplifying assumptions are made in the "optimization" models in order to haste their execution.

Although SWAT is not prepared to undertake water resources management related optimizations, it also (arguably) falls short of a detailed simulation model. Due to this, it could be place halfway on the goals scale.

Process representation

On a more technical plane, water resources models can be differentiated by how they represent processes. Regarding this representation a model can be labeled:

- theoretical or physically based;
- conceptual;
- empirical.

A theoretical model is based on physical principles. In such a model, all the governing physical processes are described by mathematical functions. A different approach is followed in empirical (or statistical) models, which omit the physics and are representations of observed data. Amid theoretical and empirical, conceptual models do not include the ruling physical principles of the phenomena but, alternatively, implement simplified equations that have been observed to approximate them adequately.

While SWAT is often advertised and referred to as a physically based model, current development of physically based models should relegate it to the conceptual category.

Results

Another relevant distinction of models is the character of the results obtained. These can be broadly classified as deterministic or stochastic. If one or more of the variables in the model are assumed to behave as random, changing unpredictably over time, the model is classified as stochastic, being otherwise branded deterministic. The majority of models are deterministic.

Notwithstanding, even a deterministic model can produce pseudo-stochastic results if run several times with parameters subject to variations. SWAT is deterministic. With appropriately sampled inputs, however, in can naturally produce pseudo-stochastic results as well.

Spatial discretization

According to their spatial discretization models are customarily classified under three categories:

- lumped;
- distributed; or
- semi-distributed.

Spatially lumped models treat the modeled area as a single unit and average the effects of variability over that unit. On the contrary, spatially distributed models separate the region to be modeled into discrete units, enabling different model inputs or parameters to be used to represent spatial variability (CRC Catchment Hydrology 2005).

Between the extremes of lumped and fully distributed models, a class of semi-distributed models can be conceived. Such semi-distributed models are widely used and represent the basin being modeled through a series of smaller conceptual units. SWAT should be classified in this latter category.

Time scope

The time scope is important as it can have substantial implications on the model's structure and adopted sub-models. Regarding this aspect, models can be event-based or continuous-time representations. As an example, models fit to simulate one or a restricted set of events can safely assume inputs such as land use to be constant. In long continuous-time representations this might not be a valid assumption. In respect to this category, SWAT is a continuous-time model.

Time step

The calculation step of water resources models can vary considerably and finding models operating with steps from minutes to months is not uncommon. The calculation step depends essentially on desired results, computational resources and available data, but may have influence on model structure. Again resorting to an example, at a monthly step and depending on the catchment considered, the use of a flow routing sub-model might be unnecessary if results achieved by simple mass balance equations prove to be accurate. With smaller calculation steps, this is seldom the case. Several time steps, for hourly to yearly, can be chosen to conduct SWAT simulations.

3.2.2 Justification and overview

The preparation and calibration of a hydrological model for the ZRB was one of the main drivers of this research. On the one hand, work focused on achieving a sound and reliable model while understanding and acting upon the key drivers of its performance and, on the other hand, it was desired that the resulting model should be easily adapted by third parties for further scientific and applied uses.

In the initial stages of the research the Routing System hydraulic/hydrological model (Garcia Hernandez et al. 2007, Jordan and Schleiss 2007) was chosen. This was an obvious choice as the model, ubiquitously applied in Switzerland, was developed at the LCH and maintained by several researchers based in the Lausanne area. The facilitated access to experts was definitely a point for Routing System, but not the only one. Modeling experience in Switzerland has proven Routing System as a computationally efficient tool which can achieve very good performances. It is a particularly strong model regarding the reproduction of dam operations and the snow melting process.

Notwithstanding these advantages, it soon became apparent that modeling in Southern Africa and at very large scales presented particular challenges. One of such challenges derived from the heterogeneity of the whole region (for example in terms of climate, topography and soil

characteristics). Being in the presence of a large and heterogeneous area demanded that a fairly adaptable hydrological model was used but, more importantly, it emphasized the difficulties associated with calibration. At the critical time for model choice, Routing System was being mainly supported by a private consulting company and all model calibration was manual and based on expert knowledge[1]. In practice, this meant that: i) the resulting model, once prepared, would not be easily passed on to interested stakeholders; ii) there was no expert knowledge on the calibration of Routing System in Southern African catchments and no guarantee that model components would perform well in the region and; iii) building and manually calibrating a Zambezi basin model without detailed knowledge of the area could be associated with a great deal of uncertainty regarding time and the scalability of the results (of particular concern was the possibility that if data inputs changed, or the model's topology had to be altered, the manual calibration process would have to be repeated).

At the same time, the SWAT had been applied several times to the ZRB and some subcatchments. It is an open-source software with great international projection, a thriving user community and extensive use in semi-arid climates and large basins. Reports of promising results in related applications abounded and, being programmed in FORTRAN, SWAT was promoted as a computationally efficient model.

The SWAT model was developed to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use, and management conditions over long periods of time (Neitsch et al. 2011). Associated with its genesis, another appealing feature of SWAT is the wide range of processes it simulates, from discharges to plant growth, soil erosion, or sediment, nutrient and pollutant transport. With these and many more processes already coded, a calibrated SWAT model is relatively easily adapted as a decision support tool regarding a plethora of integrated water management related issues. Since the late 90's SWAT has been modified to work appropriately in the Southern Hemisphere, a feature particularly important concerning plant growth cycles. Additionally, the model was enhanced in its representation of large wetlands, an important feature in the ZRB, by Cohen Liechti et al. (2014).

It should be stated that although SWAT is a very broad model, quite adapted to evaluate the impact of future scenarios on a river basin, it is not very detailed in the representation of the processes and on their interaction. This is both a positive and a negative aspect. On the positive side, a simpler model is faster to set up and run, less prone to human errors, easier to interpret and, with a parsimonious set of parameters, usually easier to calibrate. On the negative side, simplifications of the processes that take place in the river basin might lead to inaccurate simulations and/or results that cannot be safely generalized. As such, in a data rich environment, SWAT would not usually be the best model choice with the goal of reproducing very detailed hydrologic responses.

Several researchers having studied the calibration of SWAT presented a clear advantage. In that line, a number of implementations of automatic calibration algorithms have already been developed, most notably those contained in SWAT-CUP[2] (Abbaspour 2012).

SWAT also presented some drawbacks, which will be laid out in further detail later. Among these, the fact that there is no graphical user interface associated with it outstood. In fact, SWAT is a command line program whose execution is entirely controlled by a list of text files which contain information related to the model's topology, physical characteristics of subcatchments, modeling choices, input data, etc. While promoting flexibility (the text files are quite straightforward to modify) this also renders changes cumbersome due to the associated number of operations to

---

[1] At this time there are also open source versions of the software available and some testing with automated calibration tools has been endeavored.

[2] SWAT-CUP includes the following optimization tools: sequential uncertainty fitting – ver. 2, SUFI-2 (Abbaspour et al. 2007), generalized likelihood uncertainty estimation, GLUE (Beven and Binley 1992), parameter solutions, ParaSol (Griensven and Meixner 2006), Markov chain Monte Carlo, MCMC (Vrugt et al. 2008, Yang et al. 2008), and particle swarm optimization, PSO (Eberhart and Kennedy 1995, Zhang et al. 2009).

perform and files to modify. In practice, without a program managing such a procedure, it is very easy to make mistakes and, overall, quite hard to form a reasonably accurate mind scheme of the models.

Dividing the basin into smaller units or subbasins, linked through the river network, SWAT is a semi-conceptual hydrological model. Also described as physically based (Neitsch et al. 2011), SWAT's account of basin processes is, although not often going into small scale details, heavily based on their physical nature and, as much as possible, parameterized with measurable variables. From this results that, using an interface tool, a SWAT model can be prepared with the help of a Geographic Information System (GIS) based on a set of raster files with information about the region's topography, soils and land use. In the present work the model's preparation was accomplished using the ArcSWAT (Winchell et al. 2010) interface for the ArcMap GIS[3]. A description of the model's preparation and associated issues is made in chapter 7.

As written before, SWAT has a very encompassing water quality component. The successful simulation of water quality parameters demands, however, a correct representation of the basin's hydrologic and hydraulic responses. The focus of this thesis was on the latter components, being water quality modeling of the Zambezi, possibly based on the results herein reported, left for future research. As in most hydrologic software, water quality and water quantity components are largely decoupled, with the information flow going mainly from the water quantity processes to the water quality ones. This means that the water quality processes can be ignored to great extent without substantial losses in the quality and range of future uses of SWAT.

3.2.3 Hydraulic/hydrologic working principles and main components – land phase

As in most hydrological models, SWAT's representation of the basin can be divided into two main components:

- the computation of effective precipitation, infiltration and runoff, along with the subsequent generation of hydrographs; and
- routing of the hydrographs through the river network.

In a typical SWAT model the basin is divided into a number of subbasins according to local topography and other relevant features such as flow control structures or gauging stations. These subbasins can be further divided into Hydrologic Response Units (HRU). As their name indicates, HRUs are lumped areas with similar hydrologic response. In SWAT they are defined according to slope, soil profile type and land cover. Within each subbasin, the consideration of different HRUs enables a certain refinement in the calculation of overland processes while not impairing computational efficiency.

For each HRU a water balance equation (3.1) is used to update the soil water content from one day to the next. Defining the terms of the equation effectively characterizes the land phase computations of a SWAT simulation.

$$SW_{t+1} = SW_t + P_t - Q_t^{surf} - E_t - Q_t^{seep} - Q_t^{gw} \tag{3.1}$$

In equation (3.1), adapted from Neitsch et al. (2011), SW_t represents the final soil water content at day t; P_t is the precipitation[4] during day t; Q_t^{surf} stands for the surface runoff; E_t is the evapotranspiration; the percolation flow exiting the soil profile bottom is represented by Q_t^{seep}; and Q_t^{gw} is the base flow. The units for the variables in equation (3.1) are water depths [mm].

---

[3] ArcMap is a commercial software developed by ESRI. Although it is presently the most convenient software to prepare SWAT models, open-source alternatives exist, for example MWSWAT based on the MapWindow GIS.
[4] In the Zambezi, owing to its warm climate, precipitation mostly occurs in the form of rainfall. In the remainder of the document both terms are used interchangeably.

The description below resumes and highlights the workings behind the applied models (objects of Chapters 7, 8 and 9) and pinpoints the role of their main parameters. Although SWAT can be run at different time steps, from sub-daily to yearly, in the present work the daily time step was adopted. For simplicity, the account given in the next subsections is based on the daily equations.

Precipitation

In SWAT two ways of defining P_t have been foreseen. While one relies on a weather generation model, the other relies on direct inputs. In the present work precipitation data was taken from external sources[5] and, therefore, the weather generation model will not be looked into. In most SWAT implementations, precipitation inputs are identical for distinct HRUs in the same subbasin.

Runoff volume

The runoff volume, associated with Q_t^{surf}, can be computed according to the SCS curve number approach or the Green and Ampt infiltration method. Because it requires sub-daily rainfall data, the latter method was not used and is not described here. The SCS curve number approach is an empirical model based on the curve number equation (3.2):

$$Q^{gen} = \frac{\left(P - I_a\right)^2}{P - I_a + S}$$
(3.2)

Where Q^{gen} is the generated runoff [mm], P is the precipitation depth [mm], I_a represents the initial abstractions [mm] and S is the retention parameter [mm].

S depends on soils, land use and management, slope and changes in the soil water content. It is defined by equation (3.3) and varies according to the curve number parameter, CN, from which the method derives its designation.

$$S = 25.4\left(\frac{1000}{CN} - 10\right)$$
(3.3)

CN is a function of the soil's permeability, land use and antecedent soil water conditions. Its values can be found according to land use and management for four different hydrologic soil groups in tables published by the SCS Engineering Division. These tables reference CN values for fixed slopes and average moisture antecedent soil moisture conditions. When circumstances stray away from those portrayed in the tables, correction factors must be applied. Regarding antecedent moisture conditions, three classes can be considered:

1. dry (wilting point);
2. average moisture;
3. wet (field capacity).

The conversions between CN's for different moisture conditions are expressed in equations (3.4) and (3.5).

$$CN_1 = CN_2 - \frac{20 \cdot \left(100 - CN_2\right)}{100 - CN_2 + e^{2.533 - 0.0636 \cdot \left(100 - CN_2\right)}}$$
(3.4)

---

[5] Daily precipitation data was taken from several satellite rainfall estimation products and, indirectly, from gauging stations through spatial interpolation techniques. Both sources are described in detail in Chapters 4 and 5.

$$CN_3 = CN_2 \cdot e^{0.00673(100 - CN_2)}$$ (3.5)

SWAT does not internally correct CN for different land slopes, being suggested by Neitsch et al. (2011) that this can be done by defining already correct CN values in the input files. Such a procedure, however, severely affects the workflow of the standard SWAT model preparation as CN values are associated to soils and slopes, themselves related to the HRU definition. Correcting CN values outside of the model would therefore entail defining HRUs, changing the soil information accordingly, and redefining revised HRUs. Due to the complexity of the task CN was not modified accounting for slope.

As stated before, S depends on the soil water content. SWAT can correct S in two ways. The first is according to the soil profile water content. The second relies on plant evapotranspiration and past S values. Being the one used in this work, only the first alternative will be described. According to it a corrected S parameter, S^c, can be computed through equation (3.6):

$$S^c = S_1 \cdot \left(1 - \frac{SW}{SW + e^{w_1 - w_2 \cdot SW}} \right)$$ (3.6)

where SW represents the soil water content, S_1 is the retention parameter value for antecedent dry conditions, and w_1 and w_2 are shape coefficients.

S_k is obtained by equation (3.7), which results from equation (3.3) using a CN for antecedent conditions of class k, CN_k.

$$S_k = 25.4 \left(\frac{1000}{CN_k} - 10 \right)$$ (3.7)

The shape coefficients can be calculated from equations (3.8) and (3.9):

$$w_2 = \frac{\ln\left(\frac{SW^{fc}}{1 - S_3 \cdot S_1^{-1}} - SW^{fc} \right) - \ln\left(\frac{SW^{sat}}{1 - 2.54 \cdot S_1^{-1}} - SW^{sat} \right)}{SW^{sat} - SW^{fc}}$$ (3.8)

$$w_1 = \ln\left(\frac{SW^{fc}}{1 - S_3 \cdot S_1^{-1}} - SW^{fc} \right) + w_2 \cdot SW^{fc}$$ (3.9)

where SW^{fc} and SW^{sat} represent the soil water contents at field capacity and saturation, respectively. They are calculated according to equations from (3.10) to (3.12) introducing the notions of available water capacity, SW^{awc}, wilting point, SW^{wp}, soil bulk density, ρ_b [g·cm³], and the soil's clay contents, m_c [%], all of these, with the exception of SW^{wp} input parameters to the model.

$$SW^{sat} = 1 - \frac{\rho_b}{2.65}$$ (3.10)

$$SW^{fc} = SW^{wp} + SW^{awc}$$ (3.11)

$$SW^{wp} = 0.4 \cdot \frac{m_c \cdot \rho_b}{100}$$
(3.12)

The equations down to (3.9) make the estimation of surface runoff volumes possible. Considering areas with relatively large times of concentration, it is not expected that all the runoff makes it to the outlet in a single daily time step. Accounting for this, the surface runoff reaching the river network in a given day, Q_t^{surf}, is computed by equation (3.13):

$$Q_t^{surf} = \left(Q_t^{gen} + Q_{t-1}^{stor} \right) \cdot \left(1 - e^{-\frac{surlag}{t_c}} \right)$$
(3.13)

where Q_t^{gen} is the amount of runoff generated, Q_{t-1}^{stor} is the accumulated runoff in transit from the previous day (3.14), t_c represents the time of concentration [hr], and *surlag* is the surface lag coefficient [-] – a model parameter.

$$Q_t^{stor} = \left(Q_t^{gen} + Q_{t-1}^{stor} \right) \cdot e^{-\frac{surlag}{t_c}}$$
(3.14)

There are a large number of formulae to calculate t_c for a natural subbasin. While it goes well beyond the scope of this work to enumerate or describe them, it is a well-known fact that different watershed characteristics are used in different formulations and that results vary substantially. In agreement with SWAT's simplified process representation, t_c's calculation includes some assumptions worth mentioning.

The basis for the calculation of t_c in SWAT is equation (3.15) in which it appears as the sum of overland (t_c^{ov}) and channel (t_c^{ch}) travel times [hr].

$$t_c = t_c^{ov} + t_c^{ch}$$
(3.15)

The first component, t_c^{ov}, is computed using equation (3.15), which results from the application of Manning's equation to a 1 m wide strip of slope and SWAT's assumption of an average flow rate of 6.35 mm/hr.

$$t_c^{ov} = \frac{L_{ov}^{0.6} \cdot n_{sb}^{0.6}}{18 \cdot s_{sb}^{0.3}}$$
(3.16)

Above, n_{sb} is Manning's roughness coefficient for the subbasin [$m^{-1/3} \cdot s$], s_{sb} represents the average slope in the subbasin [m/m], and L_{ov} stands for the subbasin slope length [m].

The calculation of t_c^{ch} is based on a similar procedure, culminating in equation (3.17):

$$t_c^{ch} = \frac{0.62 \cdot L_{sb} \cdot n_{ch}^{0.75}}{A_{sb}^{0.125} \cdot s_{ch}^{0.375}}$$
(3.17)

Where L_{sb} represents the channel length from the most distant point of the subbasin to the outlet [km], n_{ch} is Manning's roughness coefficient of the channel [m$^{-1/3}$·s], A_{sb} is the subbasin's area [km²] and s_{ch} is the average channel slope [m/m].

As before, equation (3.17) implies the assumption of a unit source area flow rate of 6.35 mm/hr. Additionally, the average flow channel length for the subbasin, directly related to t_c^{ch} 's definition, is taken as $0.71L_{sb}$.

Evapotranspiration

Evapotranspiration is concept with arguable physical representation. Nonetheless, serving as an "umbrella" definition encompassing evaporation, sublimation and transpiration processes, its use is often convenient.

Albeit relevant in general, roughly accounting for 66% of the water that falls on the Continents (Chahine 1992), evapotranspiration processes are extremely important on the ZRB as only less than 10% of the precipitation ever reaches the outlet in the Indian Ocean.

The first step in order to compute evapotranspiration is to estimate its potential. SWAT code implements three distinct methods to do so: Penman-Monteith, Priestley-Taylor, and Hargreaves. All methods have been extensively documented in specialized literature (e.g. Brutsaert 2005) and will not be described in detail in this manuscript. After preliminary evaluations (Cohen Liechti 2013), it was observed that, for the Zambezi basin, the Hargreaves method led to more accurate historical discharge simulations[6]. As a result of these findings and given the greater simplicity of the Hargreaves formulation, it was adopted in this work. According to it, the potential evapotranspiration, E_0 [mm/day], is defined by equation (3.18):

$$E_0 = \frac{0.0023 \cdot R \cdot \left(T^{max} - T^{min}\right) \cdot \left(\bar{T} + 17.8\right)}{\lambda} \tag{3.18}$$

where R represents the extraterrestrial radiation [MJ·m^{-1}·day^{-1}], T^{max} and T^{min} are maximum and minimum daily air temperatures [°C], $\bar{T}$ is the mean daily temperature [°C], and λ is the latent heat of vaporization [MJ·kg^{-1}].

At a given time step, once the potential evapotranspiration is defined, it is important to quantify the real evapotranspiration, E [mm/day]. As mentioned previously, the evapotranspiration includes several processes which are broadly aggregated in equation (3.19):

$$E = E^{can} + E^{tra} + E^{soil} \tag{3.19}$$

---

[6] Commenting the results that point out the Hargreaves method as being the best performing for the Zambezi, several aspects should be taken into account. At first glance, and given the very structured physically-based approach of the Penman-Monteith formulation, it is surprising that such a method has led to worse discharge simulations. Looking further into the matter, however, the Penman-Montheith equations require data of solar radiation, air temperature, relative humidity and wind speed, while the Hargreaves method makes only use of air temperature. Historical climate records for the Zambezi area are sparse and relatively hard to attain. As such, climate input data for the SWAT model was derived from climate reanalysis models (looked into in chapter 4), which have a relatively low spatial resolution and limited accuracy, for what it is likely that imperfect input data was the main driver of the poorer performance of the Penman-Monteith method. Also, calibration issues might have taken a toll on the quality of the simulations that were compared, being that the superior performance of the Hargreaves method might have been related to circumstantially favorable optimization runs.

where E^{can} is the contribution from canopy interception [mm/day], E^{tra} is the plant transpiration component [mm/day], and E^{soil} is the soil and snow evaporation/sublimation [mm/day].

In SWAT, the first contribution to the real evapotranspiration comes from the canopy interception. While the SCS curve number approach is used in order to compute surface runoff, SWAT estimates canopy interception (along with surface storage and infiltration prior to runoff) as 20% of the retention parameter $\left(S \right)$. E^{can} at day t is computed according to equation (3.20), in which E_0 stands for the potential evapotranspiration [mm/day], calculated earlier, and W_t^{int} represents the volume of intercepted water [mm].

$$E_t^{can} = \min\left[E_0, W_t^{int} \right] \tag{3.20}$$

The remaining evaporative water demand, E^* [mm/day], is then computed as the difference between the potential evapotranspiration and the contribution from canopy interception (3.21):

$$E^* = E_0 - E^{can} \tag{3.21}$$

Next, if $E^* > 0$, the transpiration contribution is accounted for. When the Hargreaves method is used to estimate potential evapotranspiration, the maximum transpiration on a given day, $E^{tra,pot}$ [7], [mm/day] is function of the leaf area index, LAI (3.22):

$$\begin{cases} E^{tra,pot} = \dfrac{E^* \cdot LAI}{3} & 0 \le LAI \le 3 \\ E^{tra,pot} = E^* & LAI < 3 \end{cases} \tag{3.22}$$

The lack of water in the soil profile might prevent that the maximum transpiration occurs. The computation of the real transpiration, T [mm/day], carried out by SWAT relies on a rather complex plant growth model based on the concept of heat units, radiation-use efficiency, LAI, root development and water uptake[8].

Towards the quantification of the soil's contribution to the evapotranspiration, E^{soil}, SWAT assumes that the potential evaporation from the soil, $E^{soil,pot}$ [mm/day], depends on the remaining evaporative water demand (after discounting the canopy interception component), E^*, and is affected by the soil cover index, co [m²/m²], according to equations (3.23) and (3.24):

$$E^{soil,pot} = E^* \cdot co \tag{3.23}$$

$$co = e^{-5 \cdot 10^{-5} \cdot cv} \tag{3.24}$$

---

[7] Defined as the amount of transpiration that will occur on a given day when the plant is growing under ideal conditions.
[8] This is one of the two parts of this description of SWAT's working principles that are not self-contained. The introduction of concepts leading to the LAI calculation, closely related to plant growth models, would be lengthy and stray away from the main topics of this work. An in-depth discussion can be found on SWAT's theoretical documentation (Neitsch et al. 2011).

where cv represents the aboveground biomass and residue [kg·ha⁻¹].

$E^{soil,pot}$ is then adjusted for plant water use according to equation (3.25):

$$E^{soil,adj} = \min\left\{ E^{soil,pot}, \frac{E^{soil,pot} \cdot E^*}{E^{soil,pot} + E^{tra}} \right\}$$ (3.25)

Where $E^{soil,adj}$ is the adjusted soil evaporation component [mm/day] and E^{tra} is the plant transpiration component [mm/day], as defined earlier.

The desired soil contribution to the real evapotranspiration, E^{soil} [mm/day], is the sum of the real contributions of each soil layer. This is translated by equation (3.26) in which i identifies a given soil layer, N^{layers} is the total number of soil layers, and E_i^{soil} represents the real evaporation taking place at layer i [mm/day].

$$E^{soil} = \sum_{i=1}^{N^{layers}} E_i^{soil}$$ (3.26)

In order to calculate E_i^{soil}, SWAT relies firstly on equation (3.27), a function that returns the evaporative demand [mm/day] at a certain depth, z [m].

$$f_{soil}(z) = E^{soil,adj} \cdot \frac{z}{z + e^{2.374 - 0.00713 \cdot z}}$$ (3.27)

Defining the lower and upper limiting depths of the i th soil layer as z_i^{lower} and z_i^{upper} [m], and introducing $esco$ as a soil uptake compensation coefficient, an uncorrected estimate of the layer's contribution, E_i^{soil**} [mm/day], is obtained by equation (3.28):

$$E_i^{soil**} = f_{soil}\left(z_i^{lower}\right) - f_{soil}\left(z_i^{upper}\right) \cdot esco$$ (3.28)

For each layer, this first estimate is then corrected in agreement with the available water in the soil and its physical properties. The corrected contribution, E_i^{soil*} [mm/day], is conditioned on the layer's actual water content, SW_i [mm], the water content at field capacity, SW_i^{fc} [mm], and the water content at the wilting point, SW_i^{wp} [mm] (3.29).

$$\begin{cases} E_i^{soil*} = E_i^{soil**} \cdot e^{\left(\frac{2.5\left(SW_i - SW_i^{fc}\right)}{SW_i^{fc} - SW_i^{wp}}\right)} & \text{if } SW_i < SW_i^{fc} \\ \\ E_i^{soil*} = E_i^{soil**} & \text{otherwise} \end{cases}$$ (3.29)

In the final equation needed in order to define equation (3.19) and estimate evapotranspiration, SWAT enforces a maximum on the amount of water removed from a given soil layer (3.30):

$$E_i^{soil} = \min\left\{ E_i^{soil*}, 0.8 \cdot \left(SW_i - SW_i^{wp}\right) \right\}$$ (3.30)

Soil water

The next component of equation (3.1) to be quantified is Q_t^{seep}, the amount of water reaching the river network through soil percolation. In order to calculate it one has to model the water budget within the soil. To do so SWAT can consider several soil layers. Percolation is assumed to occur between soil layers when the upper layer's water content is over the field capacity and is described by equation (3.31):

$$Q_i^{perc} = SW_i^{exc} \cdot \left(1 - e^{-24/t^{perc}}\right)$$

(3.31)

where Q_i^{perc} [mm] is the water percolating downwards from layer i, SW_i^{exc} [mm], calculated by equation (3.32), represents the soil water contents available for percolation, and t^{perc}, the travel time for percolating water [hr].

$$SW_i^{exc} = \max\left\{0, SW_i - SW_i^{fc}\right\}$$

(3.32)

t^{perc} itself depends on the saturated layer's water contents, SW_i^{sat}, on its field capacity, SW_i^{fc}, and on its saturated hydraulic conductivity, K_i^{sat} [mm/hr] – another SWAT parameter.

$$t^{perc} = \frac{SW_i^{sat} - SW_i^{fc}}{K_i^{sat}}$$

(3.33)

When the main driver of the water movement is gravity, the percolation in the soil profile is conditioned by the amount of water leaving its lower layer, Q^{seep} [mm/day]. SWAT assumes it to depend on the distance between the bottom of the soil profile and an impervious layer, $h^{soil,imp}$ [m], and the result of equation (3.32) applied to the bottom layer of the soil profile, SW_{bottom}^{exc}, applying equation (3.34) to calculate it:

$$Q^{seep} = SW_{bottom}^{exc} \left(\frac{h^{soil,imp}}{h^{soil,imp} + e^{8.833 - 2.598 \cdot h^{soil,imp}}}\right)$$

(3.34)

The distance of the bottom of the soil profile and an impervious layer, $h^{soil,imp}$, can be found through equation (3.35):

$$h^{soil,imp} = z_{bottom}^{lower} - z^{imp}$$

(3.35)

where z_{bottom}^{lower} [m] is the depth of the lower limit of the soil profile and z^{imp} [m] is the depth of the aforementioned impervious layer[9].

It is worth mentioning that SWAT is also able to model cracks in the soil, which, depending on its composition and water contents, can contribute significantly to store a share of the generated overland flow and add to the water percolating leaving the bottom of soil profile. In the models

---

[9] These equations are valid for the Zambezi model because it does not account for perched water tables. In fact, the impervious layer is considered to be at a depth of 6 m, always below the bottom of the soil profile.

considered in this work no crack flow was modeled and, consequently, the process is not described here.

Adding to its contribution to the aquifer, soil water can also percolate directly into the river network through lateral flow. The generated lateral flow movement for layer i, Q_i^{lat*}, is calculated applying equation (3.36).

$$Q_i^{lat*} = 0.024 \cdot \left[\frac{2 \cdot SW_i^{exc} \cdot K_i^{sat} \cdot s_{sb}}{\left(\phi_i^{drain} - L^{hill} \right)} \right]$$

(3.36)

Intervening in the calculation are already introduced terms as SW_i^{exc}, the available water for percolation in the layer, K_i^{sat}, its saturated hydraulic conductivity, and s_{sb}, the average slope in the subbasin. Additionally, L^{hill} [m] corresponds to the hillslope length in the HRU (a parameter) and ϕ_i^{drain} [-] is the drainable porosity of the layer. Because the water contents are handled per area unit and in mm, the later can be obtained by equation (3.37), where the soil water contents at saturation and field capacity, SW^{sat} and SW^{fc}, have already been defined.

$$\phi_i^{drain} = SW^{sat} - SW^{fc}$$

(3.37)

As for the soil evaporation in equation (3.26), the total generated lateral flow, Q^{lat*} [mm], is aggregated across all the soil layers according to (3.38):

$$Q^{lat*} = \sum_{i=1}^{N^{layers}} Q_i^{lat*}$$

(3.38)

Again recurring to earlier concepts, not all the generated lateral flow reaches immediately the river network. Similarly to equation (3.13), for overland flow, equation (3.39) is used to lag lateral flows.

$$Q_t^{lat} = \left(Q_t^{lat*} + Q_{t-1}^{stor,lat} \right) \cdot \left(1 - e^{-\frac{1}{t^{lag}}} \right)$$

(3.39)

In (3.39) the index t identifies a given day. As can be seen, the HRU's contribution to the river depends on the amount of lateral water in transit from previous days, $Q_{t-1}^{stor,lat}$ [mm], the lateral flow generated in the present day, Q_t^{lat*}, and a model parameter, t^{lag}, which represents the lateral flow travel time [days].

Groundwater

The last component addressed in this section is the ground water. In SWAT, groundwater is modeled with the aid of a shallow and a deep aquifer. First, the recharge of both aquifers at day t, Q_t^{rchrg} [mm], is computed according to equation (3.40):

$$Q_t^{rchrg} = \left(1 - e^{-\frac{1}{\delta_{gw}}} \right) \cdot Q_t^{seep} + e^{-\frac{1}{\delta_{gw}}} \cdot Q_{t-1}^{rchrg}$$

(3.40)

where Q_t^{seep} is the amount of water leaving the lower soil layer calculated by equation (3.34) and δ_{gw} [days] is the delay time, a SWAT parameter.

The partition of the recharge among the deep and shallow aquifers, $Q^{rchrg,deep}$ [mm/day] and $Q^{rchrg,sh}$ [mm/day] respectively, varies according to the aquifer percolation coefficient, β^{deep} (also a SWAT parameter), as indicated by (3.41).

$$\begin{cases} Q^{rchrg,deep} = \beta^{deep} \cdot Q^{rchrg} \\ Q^{rchrg,sh} = \left(1 - \beta^{deep}\right) \cdot Q^{rchrg} \end{cases} \tag{3.41}$$

The water contributing to the deep aquifer is essentially assumed lost by the model unless a pump operation is defined. As such, the shallow aquifer is the sole source of groundwater directly flowing into the river network.

SWAT accounts also for the possibility that, by processes of capillarity and evaporation, water can reach unsaturated layers upwards from the shallow aquifer. Though theoretically a phenomenon more prone to occur in areas where the saturated zone is at low depths or where there is deep-rooted vegetation, in SWAT its potential, $E^{revap,max}$ [mm/day], is simply quantified by equation (3.42):

$$E^{revap,max} = \beta^{rev} \cdot E_0 \tag{3.42}$$

Where β^{rev} is another model parameter, the "revap" coefficient, and E_0 [mm/day] represents the potential evapotranspiration, calculated by equation (3.18).

The real upflow from the shallow aquifer, E^{revap} [mm/day] is, however, also limited by its water contents, W^{sh} [mm], occurring only over a certain threshold, $W^{sh,revap,thr}$ [mm]. This is done according to the equations in (3.43). $W^{sh,revap,thr}$ is a SWAT parameter.

$$\begin{cases} E^{revap} = 0 & \text{if } W^{sh} \leq W^{sh,revap,thr} \\ E^{revap} = \left(1 \cdot E^{revap,max} - W^{sh,revap,thr}\right) \cdot 1^{-1} & \text{if } W^{sh,revap,thr} < W^{sh} < W^{sh,revap,thr} + 1 \cdot E^{revap,max} \\ E^{revap} = E^{revap,max} & \text{if } W^{sh} \geq W^{sh,revap,thr} + 1 \cdot E^{revap,max} \end{cases} \tag{3.43}$$

Finally, the shallow aquifer's contribution to the river network in any given day t, Q_t^{gw} [m], depends on the contribution at the previous day, Q_{t-1}^{gw}, on the recharge of the shallow aquifer, $Q_t^{rchrg,sh}$, and on a model parameter, α^{gw}, the baseflow recession constant [days]. Also, groundwater flow is conditioned on the water contents of the shallow aquifer, W_t^{sh} [mm], happening only if they are over a pre-defined threshold, $W^{sh,thr}$ [mm]. This is translated in equation (3.44).

$$\begin{cases} Q_t^{gw} = Q_{t-1}^{gw} \cdot e^{-\alpha^{gw}} + Q_t^{rchrg,sh} \cdot \left(1 - e^{-\alpha^{gw}}\right) & \text{if } W_t^{sh} > W^{sh,thr} \\ \\ Q_t^{gw} = 0 & \text{otherwise} \end{cases} \tag{3.44}$$

Transmission losses within the subbasin

After contributions from all the HRUs have been aggregated and the streamflow generated in the subbasin is computed, SWAT calculates transmission losses in the tributary reaches. The streamflow volume after transmission losses, $V^{sub,rch}$ [m³], is conditionally computed by (3.45):

$$\begin{cases} V^{sub,rch} = a_x + b_x \cdot V^{sub,rch*} & \text{if} \quad V^{sub,rch*} > V^{sub,rch,thr} \\ V^{sub,rch} = 0 & \text{otherwise} \end{cases} \tag{3.45}$$

where $V^{sub,rch*}$ is the streamflow volume before losses, $V^{sub,rch,thr}$ [m³] is a volume threshold computed through equation (3.46), and a_x and b_x are regression coefficients.

$$V^{sub,rch,thr} = -\frac{a_x}{b_x} \tag{3.46}$$

a_x and b_x are defined by the equations in (3.47). a_r and b_r are additional regression coefficients, k_r is a decay factor [m⁻¹·km⁻¹], L_{tr} [km] represents the length of the longest tributary in the subbasin, and w_{ch} [m] is the average width of that tributary (a model parameter).

$$\begin{cases} a_x = \dfrac{a_r}{1 - b_r} \cdot (1 - b_x) \\ b_x = e^{-k_r \cdot L_{tr} \cdot w_{ch}} \end{cases} \tag{3.47}$$

The coefficients and decay factor intervening in (3.47) are defined in (3.48):

$$\begin{cases} k_r = -2.22 \cdot \ln\left(1 - 2.6466 \cdot \dfrac{K_{ch} \cdot t^{dur}}{V^{sub,rch*}}\right) \\ a_r = -0.2258 \cdot K_{ch} \cdot t^{dur} \\ b_r = e^{-0.4905 \cdot k_r} \end{cases} \tag{3.48}$$

where K_{ch} [mm/hr] is the effective conductivity of the channel alluvium (a model parameter) and t^{dur} [hr] represents the duration of the runoff flow, obtained by equation (3.49).

$$t^{dur} = \frac{t_c}{\alpha_{tc}} \tag{3.49}$$

The definition of t^{dur} builds upon the subbasin's t_c [hr] and α_{tc} [-], the fraction of daily rainfall that occurs during t_c.

Finally, α_{tc}, calculated by SWAT according to equation (3.50), depends itself on t_c and on the fraction of daily rainfall falling in the half-hour of highest intensity, $\alpha_{0.5}$ [-][10].

$$\alpha_{tc} = 1 - e^{\left[2t_c \cdot \ln\left(1 - \alpha_{0.5}\right)\right]} \tag{3.50}$$

3.2.4 Hydraulic/hydrologic working principles and main components – routing
The contribution of the subbasins to the river network can be quantified in is various components (superficial, lateral and groundwater flow) according to the equations described in section 3.2.3. After flowing into the river network, these contributions are routed downstream through channels, lakes, reservoirs and wetlands. During routing, and particularly in basins with semi-arid climate, transmission losses occur and are particularly important (Pilgrim et al. 1988).

---

[10] The calculation of $\alpha_{0.5}$ depends on monthly rainfall records and is not addressed here. Details about the procedure can be found in SWAT's technical documentation (Neitsch et al. 2011).

The routing processes in SWAT are handled in a simplified way, assuming the validity of Manning's equation for trapezoidal shaped channels in the rivers and using an elementary representation of reservoir and lake hydraulics. Also, one single river reach is considered to pass through each subbasin, being that the basin's own contributions are added to inflows from upstream before routing. How these processes are mathematically described is discussed in below.

Channels

SWAT assumes rivers to have composite trapezoidal cross-sections of constant dimensions within each subbasin. These include main channel and floodplain sections, as illustrated in Figure 3.1.
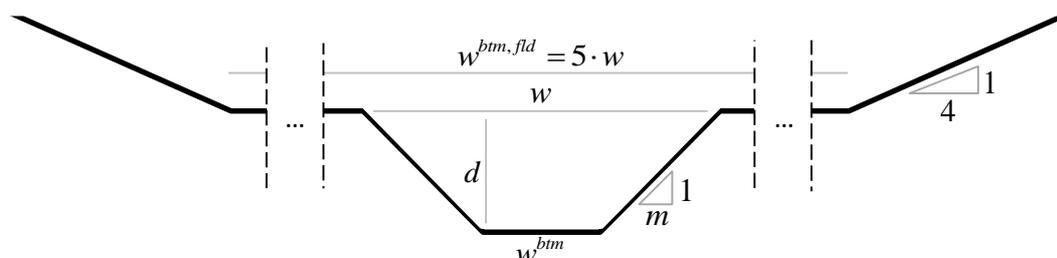
Figure 3.1. Scheme of SWAT's channel geometry.

The main channel slope (m:1) is initially assumed by SWAT to be of 2:1, while the floodplain's is 4:1. Additionally, the bottom of the floodplain, $w^{btm,fld}$ [m], is always 5 times the size of the main channel's width, w [m] – a model parameter. In order to fully define the river's geometry, an additional parameter, the main channel's depth, d [m], needs to be introduced. The first step is then to compute an initial estimate of the main channel's bottom width, w^{btm*} [m], which is done according to equation (3.51):

$$w^{btm*} = w - 2 \cdot 2d \tag{3.51}$$

As equation (3.51) can lead to negative values for w^{btm*}, which are not physically plausible, the condition (3.52) is implemented.

$$\begin{cases} w^{btm} = w^{btm*} & \text{if} \quad w^{btm*} > 0 \\ w^{btm} = 0.5w \\ m = \dfrac{w - w^{btm}}{2d} & \text{otherwise} \end{cases} \tag{3.52}$$

In case a negative estimate of the bottom width is produced, the bottom width is assumed to equal half of w. The main channel's slope, m, needs also to be recomputed.

Having defined the cross-sectional geometry, the inundated area, A [m²], can be used to calculate variables such as the water height, the width of the inundated area, w^{wtr} [m], the wetted perimeter, P_{rch} [m], or the hydraulic radius, R_{rch} [m][11]. The expressions to calculate these variables are of common knowledge among water resources practitioners and engineers and can be derived from basic geometric considerations, not being reproduced here.

---

[11] SWAT does not assume composed cross-sections where flow velocities and roughness coefficients differ between the main channel and the flood plains.

In order to quantify the amount of water exiting the reach on a given day, SWAT offers two options: the variable storage and the Muskingum routing methods. In this work only the simpler one – the variable storage method – was used[12]. According to it and as implied in equation (3.53), the volume leaving the reach on a given day, V_2^{out} [m³] (the subscript identifies the beginning – 1 – and end – 2 – of the time step), depends on the initial volume stored in the reach, V_1 [m³], the average inflows, $\bar{Q}^{in}$ [m³·s⁻¹], and a storage coefficient, sc [-].

$$V_2^{out} = sc\left(\bar{Q}^{in} \cdot 86400 + V_1\right) \tag{3.53}$$

The average inflows, $\bar{Q}^{in}$, are computed as the average between inflows at the beginning and end of the time step, assuming a linear transition.

The storage coefficient for the daily time step can be computed by equation (3.54):

$$sc = \frac{2 \cdot 86400}{2t^{travel} + 86400} \tag{3.54}$$

Where the time of travel, t^{travel} [s], is estimated by equation (3.55) based on stored volumes and outflows from the beginning of the time step.

$$t^{travel} = \frac{V_1}{Q_1^{out}} \tag{3.55}$$

To find the flow's characteristics within the reach SWAT assumes that A depends on the volume of water stored in the reach and its length, L_{rch} [km] – another model parameter. This is done according to equation (3.56).

$$A = \frac{V}{1000 \cdot L_{rch}} \tag{3.56}$$

As previously mentioned, transmission losses can be particularly important in semi-arid climates. SWAT accounts for two types of transmission losses: evaporation and infiltration.

Evaporation losses, E^{rch} [m³·day⁻¹], are computed according to equation (3.57):

$$E^{rch} = c^{evap} \cdot E_0 \cdot L_{rch} \cdot w^{wtr} \cdot fr^{rch} \tag{3.57}$$

where c^{evap} is a evaporation coefficient (model parameter), E_0 [mm·day⁻¹], L_{rch} [km] and w^{wtr} [m] have already been defined, and fr^{rch}, the fraction of the time step in which water is flowing into the channel [-], can be calculated by (3.58):

---

[12] The use of the Muskingum routing method was not endeavored due to the additional parameters its application entails and the very low degree of knowledge about the reaches' geometries and roughness on the ZRB. Without solid physical basis for evaluating these parameter's values, an additional burden would be placed on the automatic calibration process.

$$fr^{rch} = \max\left(1, \frac{t^{travel}}{86400}\right)$$

(3.58)

The volume lost through infiltration, Q^{lost} [m³·day⁻¹], is computed according to equation (3.59):

$$Q^{lost} = \frac{K^{rch} \cdot t^{travel} \cdot P_{rch} \cdot L_{rch}}{3600}$$

(3.59)

where K^{rch} [mm·hr⁻¹] is the hydraulic conductivity of the channel alluvium, and the remaining terms, t^{travel} [s], P_{rch} [m], and L_{rch} [km] have been identified above.

The infiltrated water flows towards the deep aquifer (being lost to the system) and to the river banks. The part contributing to the river banks, $Q^{bank,in}$ [m³·day⁻¹], is parameterized by fr^{deep} [-], the fraction of water flowing to the deep aquifer, as translated in equation (3.60).

$$Q^{bank,in} = Q^{lost} \cdot \left(1 - fr^{deep}\right)$$

(3.60)

The river banks act as a reservoir, eventually returning water to the reach, $Q^{bank,return}$ [m³·day⁻¹]. The process is quantified by equation (3.61):

$$Q^{bank,return} = V^{bank} \cdot \left(1 - e^{-\alpha^{bank}}\right)$$

(3.61)

where V^{bank} [m³] represents the volume of water stored in the river banks and α^{bank} [days] is the bank flow recession constant (a model parameter).

Return flow is not the only means of water to leave the bank storage. In a similar fashion to how the land phase's shallow aquifer is modeled, water can also move up through capillarity and evaporation from the bank storage. To quantify the process, SWAT begins by estimating a potential "revap", $Q^{bank,revap,max}$ [m³·day⁻¹], making use of equation (3.62) and the notion of "revap" coefficient, β^{revap} [13], introduced earlier (3.42). Additional variables also condition the estimate: E_0 [mm·day⁻¹], L_{rch} [km] and w^{wtr} [m].

$$Q^{bank,revap,max} = \beta^{revap} \cdot E_0 \cdot L_{rch} \cdot w^{wtr}$$

(3.62)

Finally, the actual amount of water flowing upwards, $Q^{bank,revap}$ [m³·day⁻¹], can be calculated conditionally according to (3.63):

---

[13] In SWAT the same β^{revap} is used for the land phase and the routing processes. A problem arises, however, because there is only one reach per subbasin, but potentially several HRUs. As there can be one different β^{revap} value for each HRU, it is not straightforward how to define the one used for the bank storage. SWAT is currently assuming it to be the β^{revap} value of the last processed HRU which introduces some degree of "unpredictability" in the calculations.

$$\begin{cases} Q^{bank,revap} = V^{bank} \cdot 1^{-1} & \text{if} \quad V^{bank} < Q^{bank,revap,max} \cdot 1 \\ Q^{bank,revap} = Q^{bank,revap,max} & \text{otherwise} \end{cases} \tag{3.63}$$

where V^{bank} [m³] has been defined previously.

SWAT assumes that the bulk volume flowing out of the reach, V^{out} [m³], is the average of the volumes flowing out of the reach at the beginning, V_1^{out} [m³], and end, V_2^{out} [m³], of each day (3.64). They are quantified with the application of equation (3.53).

$$V^{out} = \frac{V_1^{out} + V_2^{out}}{2} \tag{3.64}$$

The contributions of evaporation (E^{rch}) and infiltration ($Q^{bank,in}$ and $Q^{bank,return}$) are included in V^{out}. As such, the volume flowing downstream, $V^{out,ds}$ [m³], is defined by equation (3.65):

$$V^{out,ds} = V^{out} - 1 \cdot E^{rch} - 1 \cdot Q^{bank,in} + 1 \cdot Q^{bank,return} \tag{3.65}$$

Reservoirs, lakes and wetlands

The routing through reservoirs, lakes and wetlands is based on water balance considerations. In the following, unless indicated, the reservoir term will be used to describe lakes and wetlands as well. According to them, the volume update is ruled by equation (3.66):

$$V_t^{rsv} = V_{t-1}^{rsv} + V_t^{in} - V_t^{out} + P_t^{rsv} - E_t^{rsv} - V_t^{seep} - H_t^{rsv} \tag{3.66}$$

where V_t^{rsv} [m³] is the volume held at the end of day t, V_t^{in} [m³] and V_t^{out} [m³] represent the volumes entering and flowing downstream from the reservoir, P_t^{rsv} [m³] accounts for the precipitation falling over the reservoir, E_t^{rsv} [m³] is the evaporation taking place, V_t^{seep} [m³] values the volume lost through seepage and, finally, H_t^{rsv} [m³] are abstractions for other uses such as irrigation or water supply.

V_t^{out} is obviously limited by the water actually available on the reservoir at the end of the day, $V_t^{rsv,av}$ [m³], which can be obtained by applying equation (3.67):

$$V_t^{rsv,av} = V_{t-1}^{rsv} + V_t^{in} + P_t^{rsv} - E_t^{rsv} - V_t^{seep} \tag{3.67}$$

V_t^{in} is known as it depends of the upstream reach and the flow generated within the subbasin where it lays. The precipitation's contribution is straightforward, being simply obtained by multiplying the daily rainfall over the subbasin, P_t [mm], by the reservoir's surface, A_t^{rsv} [m²], and correcting the unit discrepancy.

E_t^{rsv} is calculated by equation (3.68):

$$E_t^{rsv} = \frac{\eta \cdot E_0 \cdot A_t^{rsv}}{1000}$$

(3.68)

where the only newly introduced variable is η [m/m], the evaporation coefficient of the reservoir and a SWAT model parameter.

Similarly, the seepage is obtained by a simple relationship (3.69) also dependent on the reservoir's area and on the effective hydraulic conductivity of its bottom, K^{rsv} [mm·hr^{-1}].

$$V_t^{seep} = \frac{24 \cdot K^{rsv} \cdot A_t^{rsv}}{1000}$$

(3.69)

Based on $V_t^{rsv,av}$ [m³], the water flowing downstream, V_t^{out} [m³], is then computed. They exact way in which it is computed depends on the nature of the reservoir (or lake). Originally, SWAT includes four distinct methods of assessing outflows:

1. measured daily outflow;
2. measured monthly outflow;
3. average annual release rate; and
4. target release for controlled reservoir.

Whenever possible, it was chosen to quantify outflows based on observations. Nonetheless, this was not possible for the ungauged wetlands or the Malawi Lake. Additionally, the existing information for the main reservoirs contains occasional gaps, being that observed releases are not available for the full length of the simulations.

In the original SWAT code, no gaps are allowed if observations are chosen to define outflows. Other options also present limitations. While the average annual release rate method is (arguably) far too simple to accurately model large wetlands or lakes, the target release controlled reservoir method, depending on a target release, and, as discussed below, on complex considerations such as the subbasin's water contents, is (again arguably) too elaborate. Most importantly, neither method represents well the operations of a large reservoir. For example, the spillway discharge rating is ignored.

Due to this, two new methods were introduced in the code. The first one, aiming at modeling lakes and wetlands, is based on the work of Cohen Liechti et al. (2014). The second one, fit for controlled reservoirs, builds upon the measured outflow and the target release for controlled reservoir original methods.

Lake and wetland outflow estimates, V_t^{out*} [m³], are made taking into account lower, $V_t^{out,lower}$ [m³], and upper, $V_t^{out,upper}$ [m³], flow components (3.70):

$$V_t^{out*} = V_t^{out,lower} + V_t^{out,upper}$$

(3.70)

The lower flow component is calculated by equation (3.71):

$$V_t^{out,lower} = \max\left\{0,86400 \cdot k^{rsv} \cdot \min\left[d^{rsv,thr}, h_t^{rsv} - h^{rsv,inv}\right]\right\}$$

(3.71)

where k^{rsv} [m^2·s^{-1}] and $d^{rsv,thr}$ [m] – a flow coefficient and a threshold water depth for upper flow to occur, are calibration parameters, $h^{rsv,inv}$ [m.a.s.l.] represents the elevation of the outlet's invert, and h_t^{rsv} [m.a.s.l.] the elevation of the water's surface.

The upper flow, $V_t^{out,upper}$ is calculated resorting to two additional calibration parameters, α^{rsv} [m·s^{-1}] and β^{rsv} as illustrated in equation (3.72). $V_t^{out,upper}$ is naturally limited to non-negative values.

$$V_t^{out,upper} = \max \left\{ 0, 86400 \cdot \alpha^{rsv} \left(h_t^{rsv} - d^{rsv,thr} - h^{rsv,inv} \right)^{\beta^{rsv}} \right\} \tag{3.72}$$

h_t^{rsv}, intervening the two previous equations, was considered a function of the available volume in the reservoir, $V_t^{rsv,av}$ (3.73).

$$h_t^{rsv} = f \left(V_t^{rsv,av} \right) \tag{3.73}$$

In the original SWAT code it is also used, being linearly interpolated based on only two points[14]. Of course, linear interpolations are acceptable for limited elevation ranges, but in the presence of large reservoir volume fluctuations they lose validity. As such, alterations were made so that, optionally, a list of volume-elevation pairs can be considered. Values not contained in the user-defied list are interpolated linearly.

The second method, designed for the basin's controlled reservoirs, firstly tries to use past observations to estimate outflows. When such observations are missing, it proceeds to estimate outflows, V_t^{out**}, based on equation (3.74):

$$V_t^{out**} = \frac{V_t^{rsv,av} - V_t^{target}}{\delta^{target}} \tag{3.74}$$

where $V_t^{rsv,av}$ is, as previously defined, the available volume in the reservoir, V_t^{target} [m^3] is a user-defined monthly volume target, and δ^{target} is another model parameter, related to the expected number of days for the reservoir to meet the volume target.

Within the code it is also foreseen that, in the absence of monthly targets, these are estimated based on the flood season and the basin's soil water contents. In the present work monthly targets are always supplied. As such, this part of the code will not be further discussed.

V_t^{out**} is then corrected according to a user-defined volume-maximum discharge curve and monthly values of minimum outflow. The corrected value, V_t^{out*}, is accordingly obtained with the conditions in (3.75):

$$\begin{cases} V_t^{out*} = 86400 \cdot f \left(V_{t-1}^{rsv} \right) & \text{if } V_t^{out**} > 86400 \cdot f \left(V_{t-1}^{rsv} \right) \\ V_t^{out*} = 86400 \cdot V_t^{out,min} & \text{if } V_t^{out**} < 86400 \cdot V_t^{out,min} \\ V_t^{out*} = V_t^{out**} & \text{otherwise} \end{cases} \tag{3.75}$$

---

[14] These points are the height of the principal and emergency spillways, the name being conceptualized. These concepts are not used according to the newly proposed outflow estimation methods.

where $f\left(V_{t-1}^{rsv}\right)$ [m^3·s^{-1}] stands for the maximum discharge obtained from the volume-maximum discharge curve and $V_t^{out,min}$ [m^3·s^{-1}] is the user-specified monthly minimum outflow.

$$V_t^{out*} = \frac{V_t^{rsv,av} - V_t^{target}}{\delta^{t\,arg\,et}} \qquad (3.76)$$

Occasionally, it might happen that $V_t^{rsv,av} \leq 0$, implying a non-positive reservoir volume at the end of the time step. When this occurs $V_t^{out*} = 0$ and water is firstly deduced from the seepage volume, V_t^{seep}. If the reservoir volume remains negative, additional water is deduced from the evaporation estimate, E_t^{rsv}.

The contribution from other abstractions, H_t^{rsv}, is then removed from the available volume according to both equations in (3.77):

$$\begin{cases} V_t^{rsv,av*} = \max\left\{0, V_t^{rsv,av} - H_t^{rsv*}\right\} \\ H_t^{rsv} = V_t^{rsv,av} - V_t^{rsv,av*} \end{cases} \qquad (3.77)$$

where H_t^{rsv*} [m^3] are the desired abstractions, $V_t^{rsv,av*}$ [m^3] is the corrected available volume and H_t^{rsv} are the actual abstractions.

Finally, when $V_t^{rsv,av*} > 0$, the computed outflows, V_t^{out}, are withdrawn from the reservoir. Again, equation (3.78) enforces the condition that there is water available.

$$V_t^{out} = \min\left\{V_t^{out*}, V_t^{rsv,av*} - V_t^{out*}\right\} \qquad (3.78)$$

3.2.5   On the choice of parameters

Being a fairly complex model, SWAT's successful application entails the definition of a large number of parameters. Reviewing the equations above, a great care in order to define processes in terms of physical parameters (which can be measured directly independently of the hydrologic response) is noticeable. Still, a large share of process parameters (which cannot be measured directly and have to be inferred through indirect means) (Moradkhani and Sorooshian 2008) remains.

It is easy to understand that in a river basin such as the Zambezi, still poorly studied, extremely large, and with little investment in hydrological data acquisition and studies, even quantities that could eventually be measured are, in practice, very hard to assess directly. Part of the solution comes from the employment of remote sensing data, particularly that focusing on soil and land cover. Indeed such data can, through the use of pedotransfer functions, be used to estimate some of the physical parameters of the model. Albeit extremely useful, remote sensing data and pedotransfer functions provide only rough parameter estimates and, in some cases, the physical parameters must be addressed in the same fashion as process ones.

Still addressing the size of the Zambezi, one would be hard pressed to agree that a homogeneous distribution of physical or even process parameters exists throughout the basin[15]. This, of course, adds another dimension to the parameter choice problem: spatial distribution.

---

[15] In fact, even the existence of a set of "correct" local parameters is commonly put into question by hydrologists, as discussed in the next section.

Due to the number of parameters and the fact that most of them cannot be measured directly, parameter choice is indeed a difficult task (Beven 1989). In order to choose appropriate sets of parameters, one can then resort to expert knowledge (unfortunately still scarce for the Zambezi), base choices on earlier works (an approach limited by the prevalent difficulties in past models' calibrations – partly a parameter definition issue), rely entirely on remote sensing data complemented with general assumptions or, alternatively, implement sound calibration techniques.

Physically based to some extent, SWAT has been tested without recurring to calibration techniques with reported success. For example, Srinivasan et al. (2010) modeled the Upper Mississippi River basin by using a careful framework for preparing input data, having obtained satisfactory results at the monthly time step without performing any kind of calibration. Such a use of the model, however, is hardy fit for the ZRB due to the lack of information and adequate data covering the area, its variable climate, and the complex hydrology. In order to be reliable, hydrological models for the Zambezi should be extensively validated and undergo careful calibration.

3.3 Calibration and validation of hydrological models

3.3.1 Overview

Hydrological model calibration and validation have been a major issue among the hydrology community for over 30 years, often raising hearted discussions over its concepts and between alternative approaches (e.g. Beven and Binley 1992, Mantovan and Todini 2006, Beven et al. 2007, Mantovan et al. 2007, Montanari 2007, Beven 2009, Vrugt et al. 2009b, Efstratiadis and Koutsoyiannis 2010).

Due to the complexity inherent to hydrologic systems, it is seldom the case that a newly prepared model with a sensible choice of parameters generates predictions which match observations. Even when this happens, the quality of these predictions can always be improved. A response is therefore to solve an inverse problem in which parameters are fitted according to historical output observations (typically, but not only, river discharges).

In general terms, any predictive model balances complexity, calibration performance and reliability. Usually, by the addition of free parameters and the increase of structural model complexity, calibration performance can be arbitrarily improved. This, however, comes at the cost of the model's reliability when processing new inputs – a problem known as overfitting. In a clear exposition of the problem when applied to hydrological modeling, the Cooperative Research Centre for Catchment Hydrology (2005) elaborates on the implications of model complexity and data availability on model choice, highlighting their importance. The risk of overfitting demands that, following calibration, a validation of the model is carried out. In the validation the model's performance when simulating events not used for calibration is evaluated.

If, in its infancy, model calibration was performed manually by experienced hydrologists, with the increasing availability of computational power and the development of evermore sophisticated optimization techniques, automatic calibration methods gained ground; particularly so among the research community. Illustrating the capabilities of automatic methods, more than a decade ago Gupta et al. (1999) compared them against multilevel expert calibration and found that their results were equivalent. Today, it can be argued that, in most cases and if correctly applied, state-of-the-art automatic methods have the potential to surpass manual calibration efforts.

A first approach to the inverse problem was to endeavor the minimization (or maximization) of a chosen objective function in respect to a set of model parameters. This Single-Objective (SO) approach, however, implicitly assumed the existence of an optimal set of parameters that could reproduce the full range of observed hydrological responses. Because hydrological modeling is laden with uncertainty, it is often the case that a single optimal set of parameters is not adequate to model the full range of a basin's hydrological responses.

While the direct problem of hydrological modeling produces unique model outputs from a set of parameters, the inverse problem attempts to characterize sets of parameters from observations. In a sense, such a problem is ill-posed because distinct sets of parameters can lead to equivalent performances, which equates to the concept of equifinality (Beven and Binley 1992, Moradkhani and Sorooshian 2008, Efstratiadis and Koutsoyiannis 2010).

In a groundbreaking contribution Beven and Binley (1992) explored the concept of equifinality, that uncertainty (Montanari 2007) has to be explicitly addressed, and that hydrologic response can only be captured by accounting for multiple sets of parameters. This new, very promising, avenue of research led to the development of quasi-Bayesian and formal Bayesian techniques to characterize the probability distribution of parameters and model errors. The potential of Bayesian-Inspired (BI) methods is enormous, enabling that prior information about the model[16] is taken into account and updated through a measure of the simulations' likelihood.

Either optimizing in regard to an arbitrary objective function or under an assumption of model error distribution (intervening in the likelihood measure), both aforementioned approaches essentially interpret the model's errors from a single point of view. Model performance, however, cannot usually be fully characterized in such a fashion. Arguably, it can be better interpreted by resorting to multiple objective functions and the concept of Pareto optimality and non-dominated solutions (e.g. Srinivas and Deb 1994, Vrugt and Robinson 2007, Efstratiadis and Koutsoyiannis 2010). According to this Multi-Objective (MO) view, solutions are evaluated in respect to several objective functions (or dimensions) and said to be optimal if there is no other solution for which all of the objectives are equally or better met. Like BI techniques, MO optimization does not offer a single solution to the calibration problem, offering a set of solutions instead.

Accompanying the development of these calibration approaches, the belief that there is a correct set of parameters – or correct hydrological model – for a given catchment has lost weight among the hydrological community. Recognizing that uncertainty from different sources is unavoidable, it is generally agreed upon that the way forward will not come from more detailed models, inputs of higher quality, or more effective calibration tools, but rather from explicitly accounting for the various sources of uncertainty.

3.3.2   Single-objective optimization

Early examples of SO optimization (e.g. Dawdy and O'Donnell 1965, Ibbit 1970, Nash and Sutcliffe 1970) applied to hydrological models relied on local methods, which cannot be used reliably to estimate global solutions (Efstratiadis and Koutsoyiannis 2010). This lack of reliability is firstly related to the usually non-linear and rough solution space generated by hydrological models, which contains numerous local optima that "trap" local methods.

An answer was to employ global search methods, less prone to be stopped at local minima. Examples of such models are Genetic Algorithms (GA) (Goldberg and Holland 1988), simulated annealing , Particle Swarm Optimization (PSO) (Kennedy and Eberhart 1995), and Shuffled Complex Evolution (SCE) (Duan et al. 1992), among many others. These methods balance exploration and exploitation (Moradkhani and Sorooshian 2008), using information to progress towards the local minima but also performing random searches of the solution space. While it is not possible to highlight one best method (Wolpert and Macready 1997), each has its strengths and weaknesses, none being able to guarantee a globally optimal set of parameters in rough non-convex solution spaces.

---

[16] Theoretically the prior information can be "hard" or "soft", being capable of including even the modeler's unproven beliefs.

The mathematical problem SO methods solve is to estimate a set of parameters, $\hat{\boldsymbol{\theta}}$, which minimize[17] a given error measure $f(\cdot)$ – equation (3.79).

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\{ f\left(\left[\hat{\mathbf{y}}(\boldsymbol{\theta} \mid \mathbf{X}, \boldsymbol{\Phi}), \mathbf{y} \right] \right) \right\} \tag{3.79}$$

The objective function – in this case minimizing $f(\cdot)$ – depends on observations, $\mathbf{y}$, and simulated values, $\hat{\mathbf{y}}$; the latter being affected by the choice of parameters, the input data, $\mathbf{X}$, and model structure, $\boldsymbol{\Phi}$.

Stemming from this formulation, calibration using SO methods depends on input data, observations, model structure, choice of objective function, initial conditions of the search and, of course, the SO method itself (Moradkhani and Sorooshian 2008).

There are numerous examples of objective functions applied to hydrological studies (e.g. Diskin and Simon 1977, Sorooshian and Gupta 1983, Yew Gan et al. 1997, Gupta et al. 1998). Perhaps the most popular among these is the maximization Nash-Sutcliffe Efficiency coefficient (NE), translated in equation (3.80) (Nash and Sutcliffe 1970):

$$\text{NE} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{3.80}$$

where N represents the number of observations, y_i is the ith observation, $\hat{y}_i$ the corresponding model prediction, and $\bar{y}$ the mean of the observations.

Alternative and popular examples of error measures minimized as an optimization objective are the Mean Squared Error (MSE), shown in equation (3.81), the Mean Absolute Error (MAE), in equation (3.82), the Maximum Absolute Error (MaxAE), in equation (3.83), or the Pearson correlation coefficient. The choices are numerous and beyond the scope of this work to enumerate.

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{3.81}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{3.82}$$

$$\text{MaxAE} = \max_{1 \le i \le N}\left\{ |y_i - \hat{y}_i| \right\} \tag{3.83}$$

Not only is there a vast choice of objective functions, there is also a sizable array of SO methods and different techniques of using them. When correctly applied these methods have been considered effective and efficient against various peculiarities (e.g. multiple peaks at all scales, discontinuous first derivatives, extended flat areas, long and curved multi-dimensional ridges, etc.) of highly non-convex solution spaces of parameter sets (Efstratiadis and Koutsoyiannis 2010). Such effectiveness and efficiency, however, holds only to some degree.

Regardless of the choices and algorithm developments (e.g. Yapo et al. 1996, Zhang et al. 2009, Abbaspour 2012) general critiques can be made of SO optimization. These have been extensively

---

[17] In reality, algorithms can be applied to minimize or maximize objective functions. As both formulations are easily interchangeable, the distinction is not important. For simplicity, in the remainder of the text it is assumed that the algorithms are minimizing the objective function.

discussed (e.g. Beven 1989, Beven and Binley 1992, Refsgaard 1997, Gupta et al. 1998, 1999, Moradkhani and Sorooshian 2008, Efstratiadis and Koutsoyiannis 2010) and dwell over the following points:

- The estimated parameter values may not be realistic.
- A good match to the calibration data might not imply a good match to independent observations (validation).
- The estimated parameters vary widely according to SO algorithm and objective function choice, input data, etc.
- Estimated parameters are not usually fit to reproduce the full range of hydrological responses of the system.
- It is difficult to recognize model structure inadequacies from SO optimizations.
- The full range of amplitude, phase and shape errors (Shamseldin and O'Connor 2001) are seldom captured by a single objective function (Diskin and Simon 1977, Sorooshian and Gupta 1983).

Authors have highlighted that poor model implementation can cause some of these problems (e.g. Kavetski et al. 2006), but the widespread belief among the scientific community is that the next qualitative leap forward will not arrive through the development of either "better" hydrological models or more powerful SO optimization techniques.

3.3.3 Multi-objective optimization

Given the limitation of SO optimization, researchers developed other approaches of model calibration. One of these is the MO optimization (Gupta et al. 1998, Efstratiadis and Koutsoyiannis 2010).

MO attempts to circumvent the inability of one objective function to fully describe model errors by optimizing parameters with regard to several objective functions. Initial attempts of solving this problem were to aggregate different objective functions through a weighting scheme. This, besides entailing obvious difficulties such as the subjective choice of weights and the combination of objective functions with different domains, reverted into a SO problem.

Based on the Pareto-dominance concept, algorithms capable of handling multiple objectives simultaneously began to emerge (Yapo et al. 1998). These, instead of aiming to find a single "best" solution, search for full range of Pareto-optimal solutions.

In order to define Pareto optimality one can start by defining $\xi_i(\boldsymbol{\theta})$ as the value of the ith error measure, $f_i(\cdot)$, formulated as before (3.84):

$$\xi_i(\boldsymbol{\theta}) = f_i\left(\left[\hat{\mathbf{y}}(\boldsymbol{\theta} \mid \mathbf{X}, \boldsymbol{\Phi}), \mathbf{y}\right]\right) \tag{3.84}$$

Considering N objectives, in regard to a set of possible parameter combinations, $\Theta = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M\}$, the combination of parameters $\boldsymbol{\theta} \in \Theta$ is said to be Pareto optimal if $\xi_i(\boldsymbol{\theta}) \leq \xi_i(\boldsymbol{\theta}_j) \ \forall_{i=1\ldots N, \ \boldsymbol{\theta}_j \in \Theta}$, and, for at least one i $\xi_i(\boldsymbol{\theta}) < \xi_i(\boldsymbol{\theta}_j)$ for all $\boldsymbol{\theta}_j \in \Theta$.

MO optimization algorithms try to characterize the hypersurface of Pareto optimal parameter sets. Again, a sizable choice of global search MO algorithms exists, particularly of evolutionary inspiration. The mathematical description of such models is well beyond the scope of this work, but they consist of a balance between parameter attraction towards points with better performance under different objectives and the repulsion of Pareto optimal parameter sets (Srinivas and Deb 1994).

Examples of different algorithms and their application to the calibration of hydrological models are plentiful and extensively listed in Efstratiadis and Koutsoyiannis (2010), having most studies

implemented the non-dominated sorting genetic algorithm II (NSGA-II) (Srinivas and Deb 1994, Deb et al. 2002a) and the multi-objective shuffled complex evolution metropolis (MOSCEM) (Vrugt et al. 2003). Still, other algorithms have emerged. One example is a multi-algorithm genetically adaptive multi-objective method (AMALGAM) (Vrugt and Robinson 2007, Vrugt et al. 2009a) which is a meta MO optimization algorithm that takes its strength from applying NSGA-II, PSO, the Metropolis-Hastings algorithm and differential evolution simultaneously.

MO optimization algorithms face the same computational difficulties as SO optimization algorithms, but present additional issues:

- The problem posed by MOs is much harder to solve than the one posed by SOs as one surface, instead of a point, is being approximated. This implies added computational resources and time.
- Typical MO optimizations are limited to a small number of objectives (e.g. two or three) as the problem's complexity soars with the addition of extra dimensions (Efstratiadis and Koutsoyiannis 2010).
- Most MO optimization algorithms are designed to handle continuous problems. When discrete variables are considered they may face convergence difficulties.

Although the ensemble of parameter sets that approximate the Pareto surface spanned by MOs contain much more information than a "best" solution found by minimizing a SO, they present a greater computational challenge. Additionally, although in principle the ensemble characterizes a range of model responses, it provides no direct indication on the probability of occurrence of each forecast or of the plausibility of a given parameter set.

3.3.4 Bayesian formal and informal approaches

Beyond the natural uncertainty (Montanari 2007) that can be associated with natural processes, epistemic uncertainty in hydrological models stems from numerous sources (Efstratiadis and Koutsoyiannis 2010, Abbaspour 2012):

- measurement errors;
- inaccurate model structures;
- inappropriate spatial and temporal representations of input data;
- poor identification of initial and boundary conditions;
- lack of information about the system's hydrological regime in the calibration data;
- use of statistically inconsistent fitting criteria;
- difficulties of finding global optima in rough and high-dimensional solution spaces;
- parameters assumed to remain constant in time.

Being impossible to perfectly characterize a hydrologic system by simulation and, therefore, being the existence of a single set of "true" parameters a misconception, another type of algorithms used to calibrate hydrological models is based on BI in order to explicitly account for uncertainty in parameter definition and model predictions. In essence, BI can be translated into equation (3.85):

$$p\left(\boldsymbol{\theta} \mid \mathbf{y}, \boldsymbol{\Phi}, \mathbf{X}\right) = \frac{p\left(\boldsymbol{\theta}\right) \cdot p\left(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\Phi}, \mathbf{X}\right)}{\int_{\Theta} p\left(\boldsymbol{\theta}\right) \cdot p\left(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\Phi}, \mathbf{X}\right) d\boldsymbol{\theta}} \tag{3.85}$$

where $p\left(\boldsymbol{\theta} \mid \mathbf{y}, \boldsymbol{\Phi}, \mathbf{X}\right)$ is the posterior probability of the set of parameters $\boldsymbol{\theta}$ given a certain model structure, $\boldsymbol{\Phi}$, input data, $\mathbf{X}$, and output observations, $\mathbf{y}$; $p\left(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\Phi}, \mathbf{X}\right)$ is the likelihood, or the probability that $\mathbf{y}$ is generated by model $\boldsymbol{\Phi}$ using the set of parameters $\boldsymbol{\theta}$ when presented with input data $\mathbf{X}$; $p\left(\boldsymbol{\theta}\right)$ represents prior beliefs about the distribution of $\boldsymbol{\theta}$; and Θ is the domain of the parameter sets.

The formal approach is powerful, yet difficult to implement in hydrological models. The difficulty is mostly associated with the definition of the likelihood function, which entails assumptions about the nature of the model errors.

As an illustration, time dependency in the error series and heteroscedastic behavior, commonly observed in hydrological series, prevent the assumption of normal likelihood distributions. Explicitly considering these characteristics of the errors in hydrological models and, consequently, assuming more elaborate likelihood functions renders equation (3.85) intractable and, in general, solvable only by computational sampling methods (Moradkhani and Sorooshian 2008).

Inspired on BI and recognizing that a single parameter set if often unfit as a base for reliable predictions, Beven and Binley (1992) addressed the problem of uncertainty in hydrological models by developing the Generalized Likelihood Uncertainty Estimation (GLUE) numerical algorithm. Being (arguably) a conceptually simple model (Beven and Binley 1992) which can estimate uncertainty, GLUE has become very popular among the hydrological research community. In order to overcome the difficulty posed by the definition of a formally correct likelihood, GLUE uses informal likelihood functions in order to calculate posterior distributions of model uncertainty.

With the enhancement of formal approaches, GLUE has been heavily criticized for its "incoherencies" (e.g. Mantovan and Todini 2006, Beven et al. 2007, Mantovan et al. 2007). The Differential Evolution Adaptive Metropolis (DREAM) Markov Chain Monte Carlo (MCMC) scheme is an example of a sophisticated formal alternative. In DREAM (Vrugt et al. 2008, Vrugt et al. 2009b), the problem of time dependency of errors is solved by including in the model's parameters the coefficients of an autoregressive series. Also, non-normal error distributions are handled by the adoption of a flexible skew exponential power error distribution and heteroscedasticity by assuming a linear relationship between flow and the error's standard deviation (Schoups and Vrugt 2010).

Despite its potential, BI presents several challenges when applied to hydrological modeling. It is complex and subjective regarding the choice of priors and likelihood functions. Additionally, like MO optimization, it demands vast computational resources (Efstratiadis and Koutsoyiannis 2010). Also, sampling algorithms that probe the solution space for regions of higher likelihood also face problems of convergence, particularly when the number of parameters is high and the solution space complex. Finally, in spite of the strong theoretical background of BI, its superiority to other approaches can still be questioned on the grounds of the validity of the chosen priors (Forster 2005).

3.3.5 Calibration and validation of the SWAT model

Numerous algorithms and strategies have been tested in the calibration of SWAT model parameters. References on the subject are Arnold et al. (2012) which overview the SWAT model's use, calibration and validation, and Abbaspour (2012) which introduces the SWAT-CUP software; a package that includes a number of optimization tools adapted for SWAT, namely:

- Sequential Uncertainty Fitting – ver. 2, SUFI-2 (Abbaspour et al. 2007).
- GLUE (Beven and Binley 1992).
- Parameter solutions, ParaSol (Griensven and Meixner 2006).
- MCMC (Vrugt et al. 2008, Yang et al. 2008).
- PSO (Eberhart and Kennedy 1995, Zhang et al. 2009).

Van Griensven and Bauwens (2003) calibrated a modified version of SWAT (ESWAT) using an aggregated MO function (in practice a SO problem) in conjunction with the SCE optimization algorithm. Among other conclusions, they state that a solution that optimized the different objectives simultaneously was unlikely to be found. In another work, a SWAT model with a total of 12 free parameters was calibrated using manual and the automatic calibration (the manual calibration was limited to six parameters) inbuilt in SWAT 2003 (van Liew et al. 2005). Results were split, being that the manual approach performed better under certain measures. Overall, the author concluded that the automatic approach was useful and particularly sound if

complemented with expert judgment. In an application of the model-independent Parameter Estimation (PEST) software (Doherty 2005) to two small South African catchments, Govender and Everson (2005) limited automatic calibration to two types of parameters and attributed errors to SWAT model inadequacies.

Zhang et al. (2009) compared GA, SCE, PSO, differential evolution and artificial immune system SO optimization algorithms in three North American watersheds while calibrating 16 parameters. Their findings were that GA led to better results if more than 2000 model evaluations were allowed, while PSO was advantageous otherwise.

Abbaspour et al. (2007) applied the SUFI-2 algorithm, which performs an informal estimation of model output uncertainty, to a pre-Alpine/Alpine catchment. Conclusions were positive, highlighting the remarkable adjustment between simulations and observations.

Five alternative uncertainty analysis procedures were tested by Yang et al. (2008). They found GLUE to generate large uncertainty bands for model parameters but cover adequately model prediction errors. ParaSol, resulted in too moderate uncertainty bands due to the violation of normality and error independence assumptions. SUFI-2 was acknowledged as the best performing model when only few model evaluations are possible. MCMC achieved a good agreement with measurements, having the disadvantages of demanding the construction of an appropriate likelihood function and needing a large number of model evaluations to converge. Finally, a primitive importance sampling approach based on an autoregressive error model was deemed too inefficient.

Bekele and Nicklow (2007) used the NSGA-II MO optimization algorithm in a small watershed and found encouraging representations of a 3-D Pareto surface while calibrating 12 streamflow parameters. Cohen Liechti (2013) applied the AMALGAM MO algorithm to the optimization of a SWAT model of the ZRB, adopting a multi-site and computationally intensive calibration strategy. Overall, satisfactory results were attained. AMALGAM was also employed to optimize three parameters for the Upper Rhone River Basin (Rahman et al. 2013), having performed better than manual calibration.

3.4   Machine learning models

3.4.1   Overview

Machine learning models endeavor the difficult task of learning from data. The concepts, motivations and approaches behind such models vary widely. At a basic level they can be differentiated between supervised and unsupervised. In the latter case, models learn from input features without the need for any "targets". Examples are mapping or clustering tools like the Self-Organizing Map (SOM) (Kohonen 1982). The supervised models, more used in the conducted research, learn from sets of input-output patterns which provide a "target" to their learning. Examples of supervised models are Multilayer Perceptrons (MLP) (Rumelhart et al. 1986), Support-Vector Machines (SVM) (Cortes and Vapnik 1995), or Random Forests (Breiman 2001). Illustrations of several artificial neural networks (ANN), following different concepts, are presented in Figure 3.2.

Unsupervised models such as the SOM are useful to map problems with many dimensions into smaller spaces or simply to cluster data. As SOM theory will not be addressed here, a brief example of this model's capabilities was programmed and is illustrated below. In the example, a five dimensions problem will be mapped into one and two dimensions, respectively. Generally, it is not easy to keep track of the variables in a 5-D problem. One exception humans are able to cope with quite remarkably are images. In Figure 3.3 an image of the Cahora Bassa dam is decomposed in 5-D data by considering horizontal positioning (x), vertical positioning (y) and three color channels.
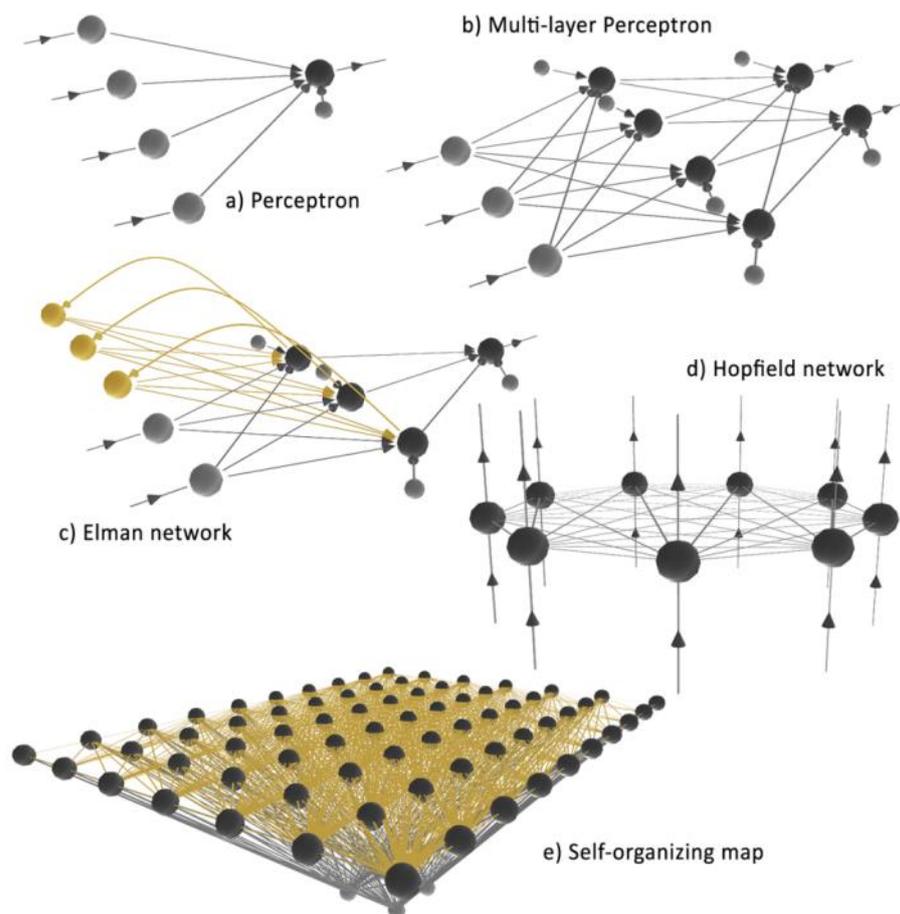
Figure 3.2. Illustration of different types of artificial neural network.

In order to properly interpret the example, one must keep in mind that the SOMs are unaware of the spatial structure or the nature of the image, being only presented with a matrix containing 5 columns and a number of rows equal to the number of pixels in the image (under no special ordering).

In Figure 3.4 the adjustment of a 1-D SOM to the data is shown after one iteration. 50 iterations later, the line has spread to cover most of the image (Figure 3.5). What is special about this mapping is that each node in the line is connected to the next and, therefore, adjacent nodes tend to be spatially close (contribution of x and y information) and cover areas with similar color (contribution of the red, green and blue channels). In Figure 3.6 a representation of the final mapping of the problem into 1-D is displayed. Color is attributed to each node as the composite of the node's coordinates in the color channels of the 5-D space for easier interpretation.

The 2-D mapping follows the same principles, but now the nodes are related through a grid. In Figure 3.7 the first iteration of the SOM's adjustment is shown. In Figure 3.8 the final positions after 50 iterations are depicted.

The interpretation of 2-D maps is far more difficult as the final SOM can fold and bend over itself. Nonetheless, valuable information can be gathered from it. A representation of the 2-D map is made in Figure 3.9. As before, color was added to each node to aid interpretation. In Figure 3.10 the clustering interpretation of the SOM is hinted by the count of the number of pixels attracted by each node.
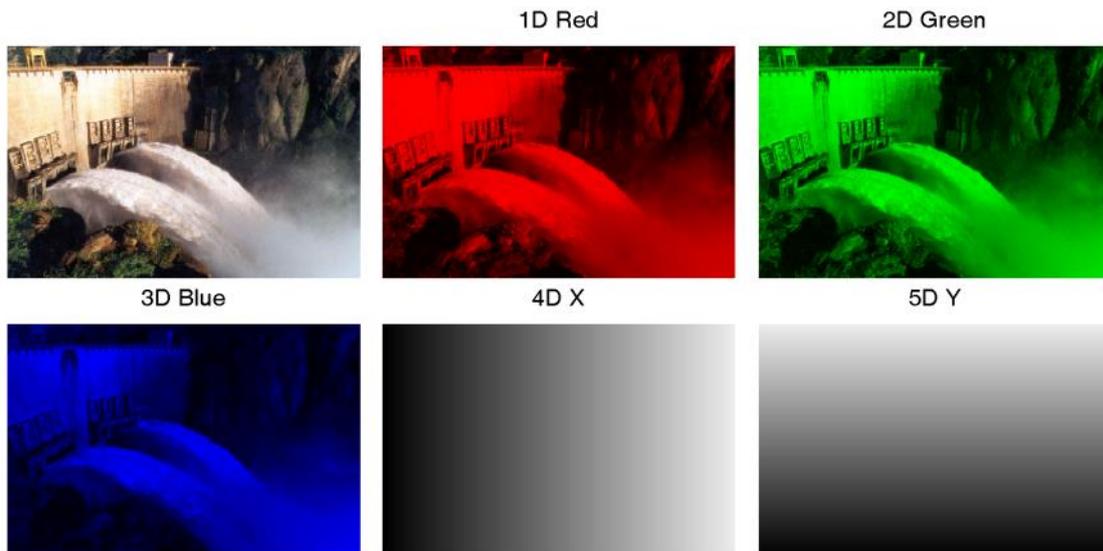
Figure 3.3. Decomposition of an image of the Cahora Bassa dam in five dimensions.

Figure 3.4. First iteration of a 1-D SOM's adjustment to the image.
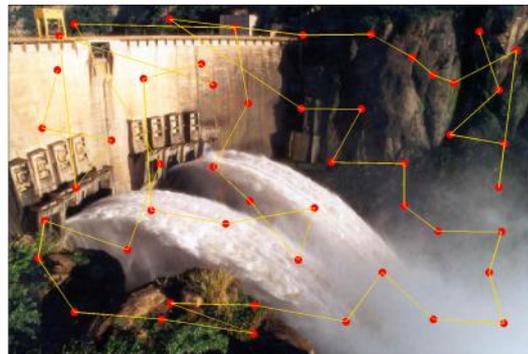
Figure 3.5. Last iteration of a 1-D SOM's adjustment to the image.

Figure 3.6. Illustration of the final 1-D mapping of the image[18].

Supervised models can be aimed at solving classification or regression tasks. In classification problems the model is employed in order to assign a certain input to a given class minimizing misclassification errors. In regression problems the goal is to approximate the value of an unknown law or function whose domain is defined by the inputs. While classification was not used in the scope of this work and, therefore, is no longer debated, some landmark regression models are discussed in detail in the next subsections.

Numerous engineering applications and researches have resorted to machine learning models, in particular to ANNs such as MLPs. The potential of machine learning to approximate unknown, non-linear, multidimensional laws from observed data can be very useful when the physical processes underlying the phenomena of interest are little known or too complex to be accurately described by traditional approaches.

---

[18] The interpretation of such maps is not always easy. In this case, from right to left, the dam's structure can be clearly seen on the edge of the map, followed by the water jets. Then, the map goes over the green areas above the dam and on the bottom-left of the picture, passes through the mist on the bottom right, and finally depicts the rock on the upper right corner.
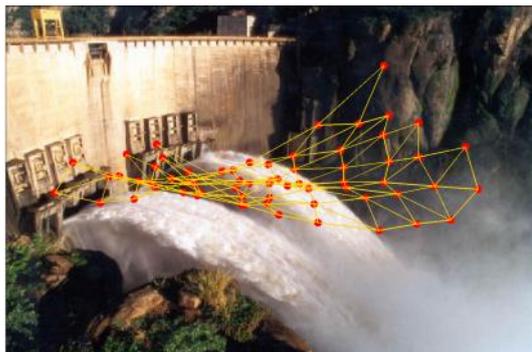
Figure 3.7. First iteration of a 2-D SOM's adjustment to the image.
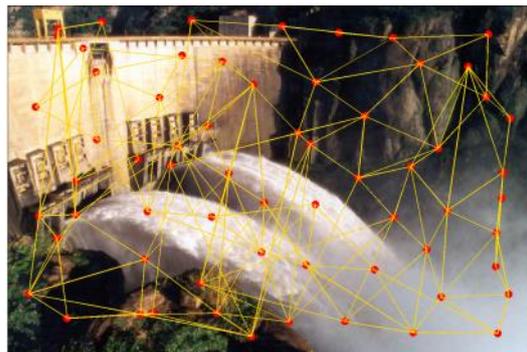
Figure 3.8. First iteration of a 2-D SOM's adjustment to the image.
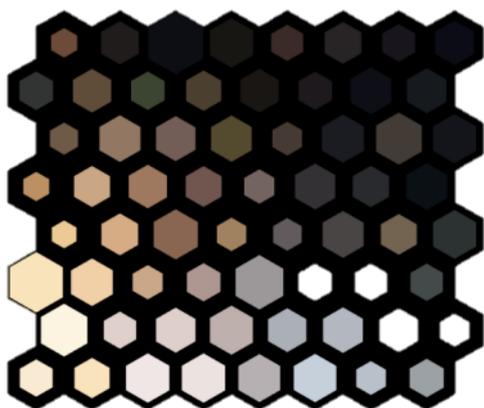
Figure 3.9. Illustration of the final 2-D mapping of the image (rotated 180°).
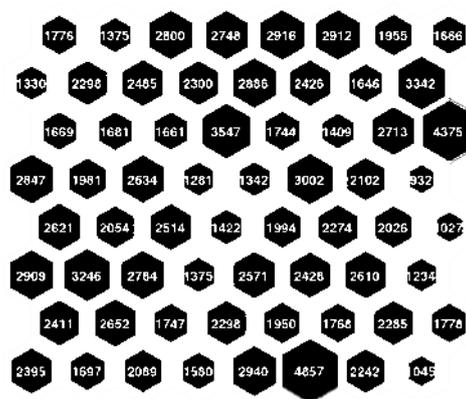
Figure 3.10. Number of hits for each node in the final 2-D map.

In the field of hydrology, for example, ANNs have been extensively tested as a data-driven alternative or complement to conceptual and physically based rainfall-runoff models (e.g. Thirumalaiah and Deo 1998, Govindaraju 2000b, Kumar et al. 2004, De Vos and Rientjes 2005, 2006, Rouhani et al. 2007) or as a means to estimate rainfall (e.g. Hsu et al. 1997, Lin and Chen 2004, Sivapragasam et al. 2010, Tapiador et al. 2010, Matos et al. 2013a).

Machine learning models for regression such as the MLP, the Radial Basis Function Network (RBFN), or Support Vector Regression (SVR) are of a *quasi*-non-parametric nature, reducing the number of parameters that the user should adjust to a minimum.

As for hydrological models – and even more markedly so – the touchstone of producing a good general purpose regression model is the trade-off between training performance, amount of information available in the training set, and reliability. In general, this important trade-off is controlled by the model's complexity (Figure 3.11).

As will be described in the following, the way in which complexity is set differs between models. Illustrating this, if in the MLP it is mostly related to the number of connections (weights) in the networks, in SVR it is set through the choice of a kernel function[19] and the hyperparameter C [20].

---

[19] A function that maps the input space in a high dimensionality "feature space" where the regression is actually performed. Further details can be found in Subsection 3.4.4.

[20] Also addressed in Subsection 3.4.4, the C hyperparameter explicitly controls the SVR response's complexity. C is coined a hyperparameter because it controls how the model's internal parameters – the ones dependent not on the user but on the training patterns – are calculated.
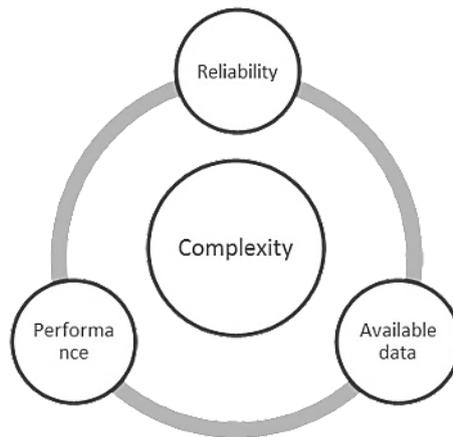
Figure 3.11. Trade-off to be pondered in regression machine learning models.

While it is known that seldom can it be said that there are better or worse models in absolute terms (Wolpert 1996a, b), having each alternative its own strengths and weaknesses, it is considered interesting to illustrate the behavior of the regression models used in this work. To that end, an artificial 1-D problem was devised. The aim is to approximate an unknown sine wave signal, from which samples have been measured. The measurements have been corrupted by Gaussian noise and outliers were artificially introduced. Furthermore, the measurements have been split into training and validation subsets – something needed to perform cross-validation in MLP models and adjust hyperparameters in the RBFN, SVR and Least-Squares SVR (LS-SVR) models[21]. The true signal, along with training and validation points, is illustrated in Figure 3.12.
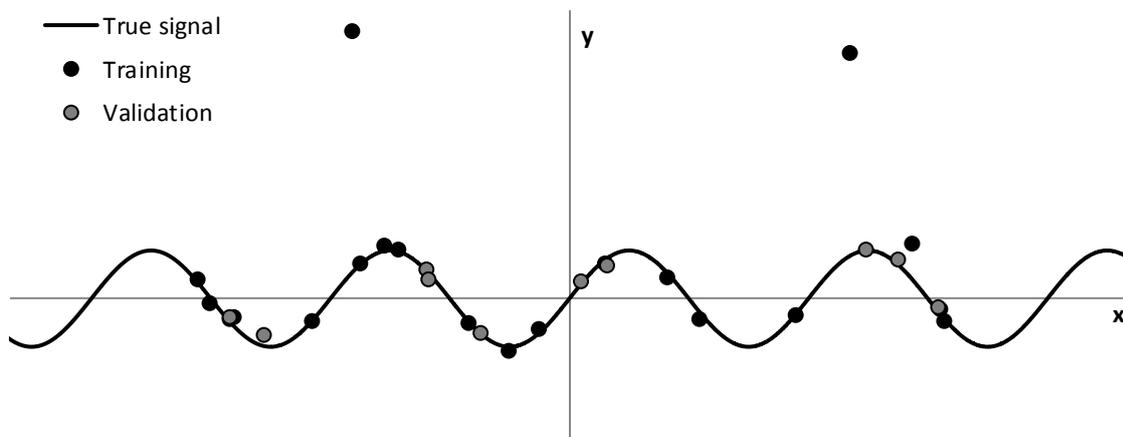
Figure 3.12. Problem of the regression of a sine wave used to illustrate the behavior of different machine learning models.

3.4.2   Multilayer Perceptrons

The most well-known and used ANN is the MLP (Haykin 1994, Govindaraju 2000b, a). Based on the linear perceptron model (Rosenblatt 1958) and heavily inspired in the most basic levels of functioning of biological nervous systems, MLPs have been and are still heavily used in engineering and, in particular, hydrology (e.g. Hsu et al. 1997, Thirumalaiah and Deo 1998, Chen and Adams 2006, Rouhani et al. 2007, Akhtar et al. 2009, Wu and Chau 2010, Mekanik et al. 2013).

The MLP is a feedforward network comprised by links and nodes and organized in layers, being that the designation of "feedforward" stems from the fact that signals propagate from layer to layer always in the same direction.

---

[21] The concept of cross-validation is described below, as are the referenced models.

Within the network, nodes process and modify signals coming from the previous layer by resorting to input, activation and output functions. Non-linearity of the activation functions grants the MLPs the ability to approximate non-linear spaces.

Each MLP has one input layer, one output layer, and one or more hidden layers in between. The hidden layers derive their designation from the fact that, unlike what happens for their input and output counterparts, the signals passing through them are not "shown" to the user and are generally hard to interpret.

MLPs are used for classification and regression tasks, being able to cope well with non-linear systems. With the appropriate choice of activation function and comprised by one hidden layer they have the capability to approximate arbitrary decision regions arbitrarily well (Cybenko 1989).

Part of the interest of MLPs is that they can be adapted or trained to approximate a set of input and output patterns. The adaptation occurs not in the nodes, but in the links between them. In fact, each link has a weight that affects the signals it transmits from layer to layer. It is through changes in these weights that the network's response is controlled.

An example of the topology of a simple MLP is presented in Figure 3.13, where values $x_1 \cdots x_n$ represent inputs; w_{ik}^j is the weight of the link from node i in layer $j-1$ to node k in layer j; b_k^j is a constant or bias added to the signals flowing into node k in layer j; and x_k^j is the signal after being processed by the latter node. The network's outputs are the processed signals of the output node(s).
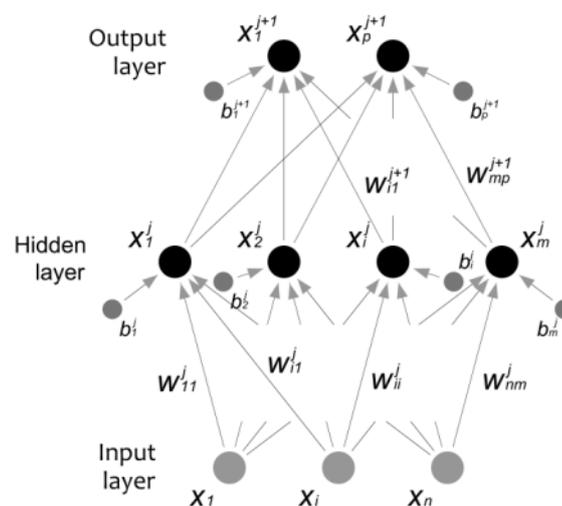
Figure 3.13. Information flow within a multilayer perceptron.

Within each node, inputs from the previous layers are aggregated and transformed as illustrated in Figure 3.14. This is done by input functions (if), which can vary, but in the vast majority of engineering applications a simple sum is employed. The activation functions (af) are the most important part of the nodes, as it is them that control the network's type of response. Finally, output functions (of) can rescale the signals before being passed on to the following layers, but are seldom used in engineering.
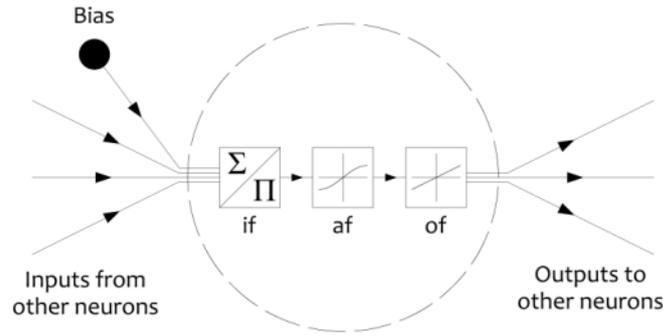
Figure 3.14. Information flow within a multilayer perceptron's node.

Four types of activation function are commonly considered: linear, hyperbolic tangent (3.86), arc-tangent (3.87), and logistic (3.88).

$$y = \frac{e^{2x} - 1}{e^{2x} + 1} \tag{3.86}$$

$$y = \frac{i\left[\log\left(1 - i \cdot x\right) - \log\left(1 + i \cdot x\right)\right]}{2} \tag{3.87}$$

$$y = \frac{1}{1 + e^{-x}} \tag{3.88}$$

Illustrations of these functions are presented from Figure 3.15 to Figure 3.18. The hyperbolic tangent is commonly used in general purpose MLPs. The logistic function can be useful in cases where negative outputs are not plausible as it always displays positive values.
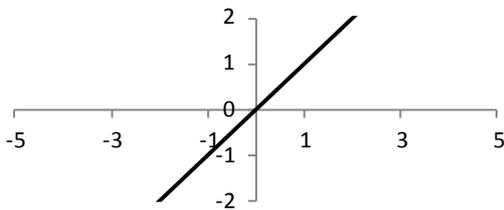
Figure 3.15. Linear activation function.
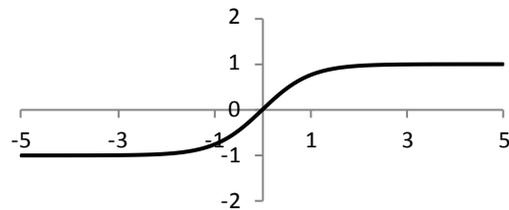
Figure 3.16. Hyperbolic tangent activation function.
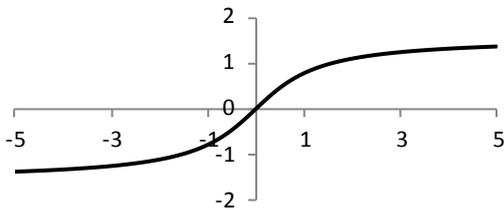
Figure 3.17. Arc-tangent activation function.
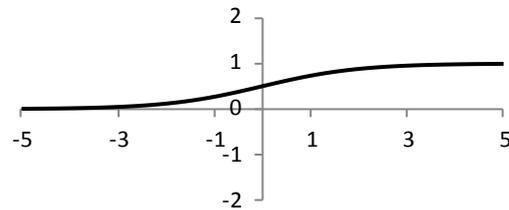
Figure 3.18. Logistic activation function.

According to Figure 3.13, the output of node k of the hidden layer j, x_k^j, is given by equation (3.89):

$$x_k^j = af\left(b_k^j + \sum_{i=1}^{n} w_{ik}^j \cdot x_i^{j-1}\right) \tag{3.89}$$

where n is the number of inputs and all the remaining terms have been previously identified.

The computation can be extended to the output layer, where node p's response is given by equation (3.90).

$$x_p^{j+1} = af\left[b_p^{j+1} + \sum_k^m w_{kp}^{j+1} \cdot af\left(b_k^j + \sum_{i=1}^n w_{ik}^j \cdot x_i^{j-1} \right) \right] \tag{3.90}$$

Of course, for the network to be of any use it is crucial that all the weights and biases are properly adjusted. Doing it is not trivial and the widespread answer to the problem came only with the work of Rumelhart et al. (1985, 1986), which introduced the backpropagation algorithm.

Although the details of backpropagation can be found in numerous publications (e.g. Haykin 1994, Kasabov 1996), its basic idea is to accomplish the improvement of the network's simulation of an observed training set of input patterns in regard to a chosen error measure (such as the mean squared error, MSE). In general terms the problem is, as shown in equation (3.91), to estimate the matrices of weights, $\hat{\mathbf{W}}$, and biases, $\hat{\mathbf{B}}$, which minimize an error function computed with a matrix of observations, $\mathbf{Y}$, and the network's outputs, $f(\mathbf{X} \mid \mathbf{W}, \mathbf{B})$; the latter being defined by the inputs, $\mathbf{X}$, weights and biases.

$$\hat{\mathbf{W}}, \hat{\mathbf{B}} = \underset{\mathbf{W}, \mathbf{B}}{\arg\min} \left\{ MSE\left[\mathbf{Y} - f(\mathbf{X} \mid \mathbf{W}, \mathbf{B}) \right] \right\} \tag{3.91}$$

To perform the backpropagation of the error, that error is firstly computed and attributed to the nodes in the output layer. Then, it is propagated backwards according to the estimated contribution that each of the links made to that error. Backpropagation is an iterative process of optimization and, significantly, one that can and often is trapped in local minima. Many methods have been proposed to optimize the network's weights effectively and efficiently. Among these, which have been extensively compared (e.g. Matos 2008) and is beyond of the scope of this work to enumerate, the most relevant are perhaps the Broyden-Fletcher-Goldfarb-Shanno update (Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970), particularly useful for classification problems, the Levenberg-Marquardt algorithm (Hagan and Menhaj 1994) extremely effective for regression problems, and Bayesian regularization (MacKay 1992a, MacKay 1992b). In addition to such local search algorithms, global search methods as the ones described in Section 3.3 can be used to train the network. In some conditions, particularly for small-sized networks, they can be advantageous.

A common problem which has to be properly addressed when applying MLPs is overfitting: the acquisition of better training performances at the expense of the network's generalization capabilities. In fact, MLPs can be adapted to most functions arbitrarily well given enough network size and training resources are provided, but they can become "specialized" in the training patterns and lose the ability to simulate different inputs. An example of overfitting is presented in Figure 3.19.

Several techniques can be applied in order to prevent overfitting. These range from the Bayesian regularization algorithm, mentioned before, to pruning network weights (e.g. Haykin 1994), or implementing a cross-validation scheme.
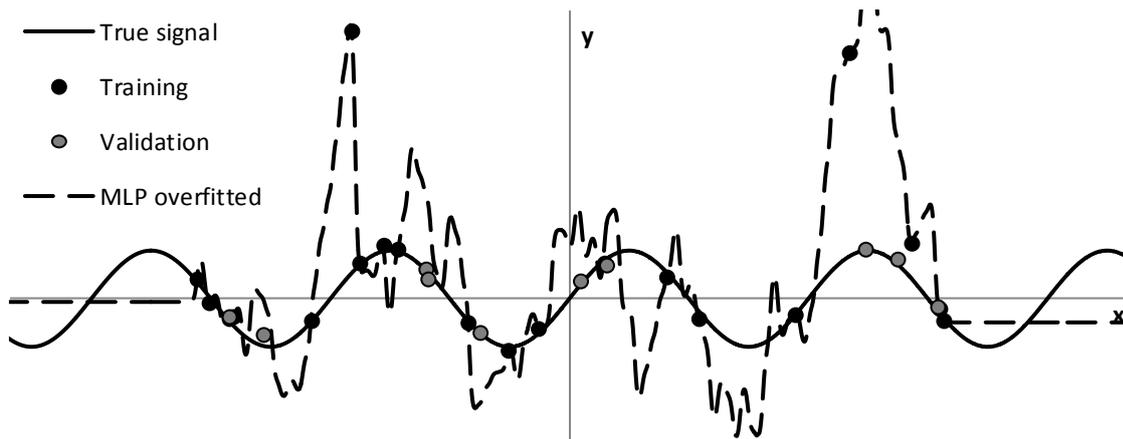
Figure 3.19. Example of an overfitted MLP.

Cross-validation is adopted most commonly due to its conceptual simplicity. For MLPs, the input-output patterns available for training are split among a training and a validation subset. For each iteration of the network's adaptation, based solely on patterns contained in the training subset, its performance is evaluated for both training and validation subsets. As the network's adaptation progresses, performances of both subsets should improve. Past the point of overfitting, however, the training performance continues to improve while validation performance worsens. Beyond this point training should stop and the combination of weights that led to the best validation performance be taken. The cross-validation procedure is illustrated in Figure 3.20.
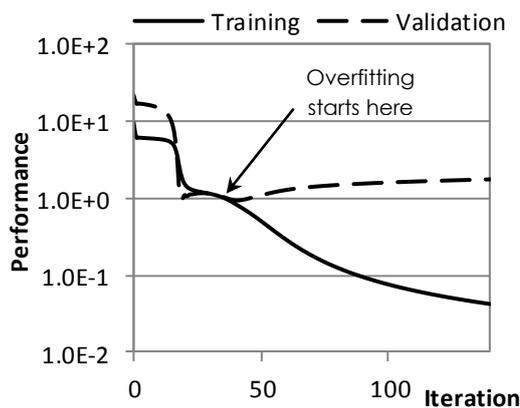
Figure 3.20. Illustration of the cross-validation procedure with indication of the point where overfitting begins.

The optimization process in MLPs is of a stochastic nature, being affected by the arbitrary choice of the initial weights and the partition of training and validation subsets. Stemming from the consequent uncertainty, the validity or performance of MLPs to perform a certain task should never be evaluated by analyzing a single case.

In order to obtain the best results using a MLP the network's architecture (number of hidden layers, number of nodes within each layer, type of activation functions, etc.) should be customized to the problem at hand and an appropriate training function should be selected. As a rule of thumb, the more available training data there is, the more complex a performing architecture can be. The process, however, is not linear, being that the number of training patterns needed to guarantee reliability soars with added complexity.

In Figure 3.21 a simple MLP is tested in the example sine wave regression problem. As can clearly be seen, while the MLP achieves a rough approximation of the sine wave it is affected by outliers and does not extrapolate well.
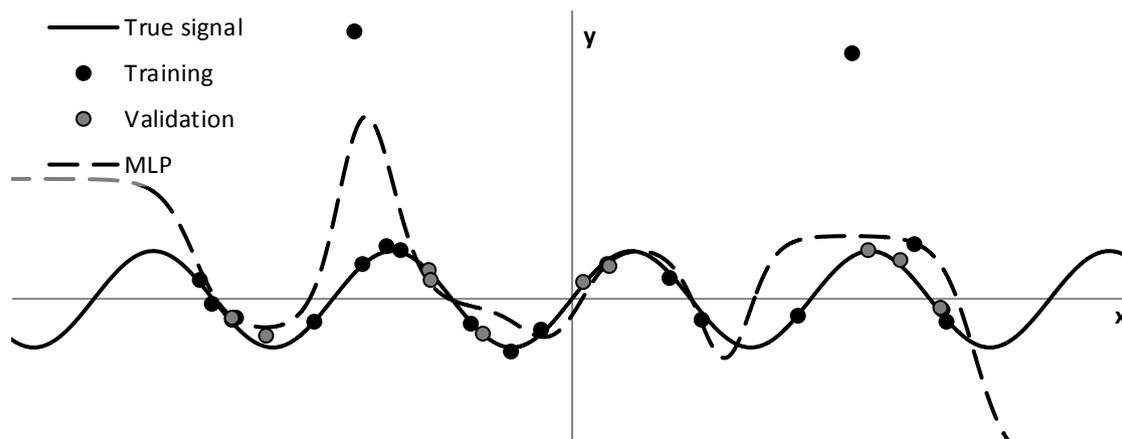
Figure 3.21. Simple MLP adjusted to reproduce a sine wave.

3.4.3 Radial basis function networks

RBFNs are another type of artificial intelligence model which can be applied to regression and classification tasks (Broomhead and Lowe 1988, Haykin 1994). They differ substantially from the MLP concept, having great affinity with multiple regression. A RBFN estimate of pattern p, $\hat{\mathbf{y}}_p$, is obtained by equation (3.92):

$$\hat{\mathbf{y}}_p = \mathbf{W} \cdot \mathbf{f}_p + \mathbf{b}$$
(3.92)

where $\mathbf{W}$ is a matrix of (linear) regression weights, $\mathbf{b}$ is the vector of intercepts, and $\mathbf{f}_p$ is a vector of radial basis functions (RBF), as defined by equations (3.93) and (3.95):

$$\mathbf{f}_p = \left[f_p^1, f_p^2, \ldots, f_p^k \right]^{\mathrm{T}}$$
(3.93)

$$f_p^i = e^{-\gamma \left\| \mathbf{v}_i, \mathbf{x}_p \right\|^2}$$
(3.94)

where k is the number of bases for the RBFs, f_p^i is the value of the i th RBF, $\mathbf{v}_i$ represents the coordinates for the its center, γ defines the RBF's width, and $\mathbf{x}_p$ are the inputs for pattern p.

Once all the $\mathbf{v}_i$'s are defined and taking $\mathbf{F}$ as the matrix containing realizations of $\mathbf{f}_p$ for all patterns used for training, the estimate of $\mathbf{W}$ is obtained as for a multiple regression problem (3.95)[22].

$$\hat{\mathbf{W}} = \left(\mathbf{F}^{\mathrm{T}} \mathbf{F} \right)^{-1} \mathbf{F}^{\mathrm{T}} \mathbf{y}$$
(3.95)

The difficulty of implementing RBFNs lies, therefore, in the choice of the RBF centers, $\mathbf{v}_i$, and width, γ.

---

[22] The notation in this equation is not entirely coherent with equation (3.92)'s as the intercepts ($\mathbf{b}$) are no longer explicit. This was done for simplicity. In matrix implementations of multiple regressions it is easier to estimate the intercepts by adding a row of ones to the input matrix, in this case $\mathbf{F}$.

γ is best chosen by trial and error or through a cross-validation procedure. As for choosing $\mathbf{v}_i$, there are several approaches. One is to consider an exact RBFN in which there is one center for each input pattern. In this approach, the RBFN is said to be exact in the sense that it perfectly simulates all input patterns[23]. An exact RBFN can, however lead to problems of overfitting. A more parsimonious approach is to add the most relevant $\mathbf{v}_i$'s iteratively according to how their inclusion decreases the error and stop at a desired value, or to implement a cross-validation scheme similar to the one described for MLPs. Because $\mathbf{v}_i$ can be chosen arbitrarily, not necessarily being part of the training set, yet another approach would be to implement a clustering technique in order to derive centers which characterize the input space.

Albeit constituting great interpolators, depending on the chosen γ, RBFNs can diverge fast when extrapolating beyond the training domain.

An example of the RBFN's response to the sine wave regression problem is depicted in Figure 3.22 for the cases of exact and iterative approaches to RBF center selection. While both models perform relatively well in the training region, they very quickly diverge when extrapolating.
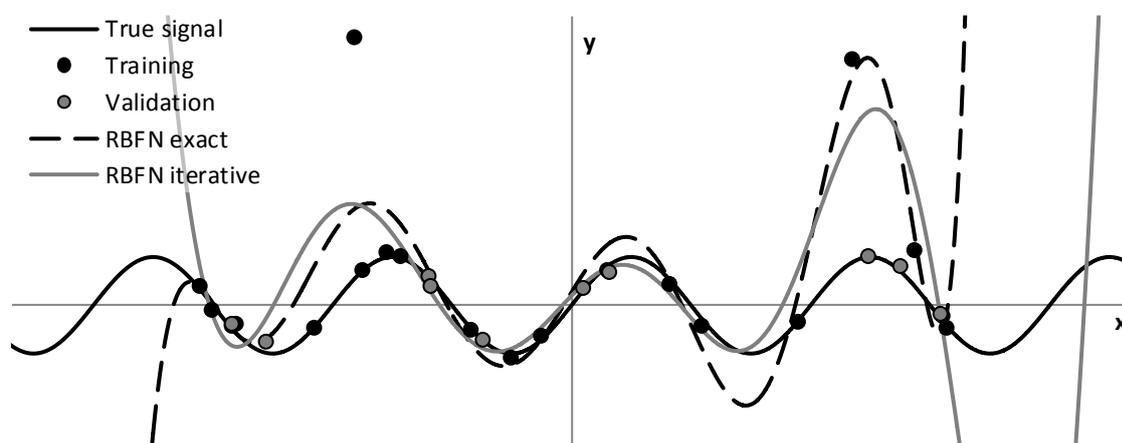
Figure 3.22. Exact and iterative RBFNs adjusted to reproduce a sine wave.

3.4.4 Support vector machines

SVR (Drucker et al. 1996) is a machine learning regression model closely related to the SVM concept (Cortes and Vapnik 1995). Being thoroughly described in some of the aforementioned publications (Liong and Sivapragasam 2002, Lin et al. 2006, Yu et al. 2006), among others (e.g. Haykin 1994), SVR consists of a regularized multiple linear regression performed in a high dimensional space spanned by a potentially non-linear projection of training inputs. Formally, the regression is accomplished by solving equation (3.96):

$$\hat{y}_p = \sum_{i=1}^{N} \left(\alpha_i^* - \alpha_i \right) \left\langle \phi\left(\mathbf{v}_i\right), \phi\left(\mathbf{x}_p\right) \right\rangle + b \qquad (3.96)$$

where $\left\langle \ \right\rangle$ is the internal product operator, $\hat{y}_p$ is the desired prediction, N represents the number of training cases, α_i, α_i^* and b are regression parameters to be optimized $\phi(\cdot)$ is a function which projects the inputs into a high dimensionality feature space, $\mathbf{v}_i$ is the ith training input and $\mathbf{x}_p$ the inputs to be predicted upon.

---

[23] Only valid if the patterns are coherent. If the same inputs can lead to different results the RBFN's response will only be approximate.

The $2N+1$ regression parameters can be determined by minimizing expression (3.97) with respect to $\mathbf{w}$, a vector which defines the regression hyperplane in the feature space.

$$\frac{1}{2}\left(\mathbf{w}^{\mathrm{T}}\mathbf{w}\right)+C\left(\sum_{i=1}^{N}\xi_{i}+\sum_{i=1}^{N}\xi_{i}^{*}\right) \tag{3.97}$$

In expression (3.97) ξ_{i} and ξ_{i}^{*} are parameters associated with model errors and C is a regularization constant that controls the trade-off between model simplicity and training performance. In order to define ξ_{i} and ξ_{i}^{*} one has to attend to the denominated ε-insensitive loss function incorporated in SVR (3.98):

$$L_{i}=\begin{cases}0 & \text{if } |y_{i}-\hat{y}_{i}|\leq\varepsilon \\ |y_{i}-\hat{y}_{i}|-\varepsilon & \text{otherwise}\end{cases} \tag{3.98}$$

where y_{i} is the ith observation used for training and ε the half-width of a band in which errors are neglected.

In such a context, the application of this loss function is interesting because the absolute error is less affected by outliers than the quadratic error and, from a computational perspective, all the training observations falling inside the 2ε band can be neglected. The vectors outside the band, effectively used for computation are the so-called support vectors.

If $y_{i}>\hat{y}_{i}+\varepsilon$, ξ_{i} equals L_{i} and ξ_{i}^{*} is zero. Conversely, if $y_{i}<\hat{y}_{i}-\varepsilon$, ξ_{i}^{*} takes the value of the loss function and ξ_{i} is null. In order to minimize (3.97) N systems of inequalities (3.99) must be taken into account and the resulting Lagrangian minimized. In fact, α_{i} and α_{i}^{*}, used in equation (3.96), are the Lagrange multipliers.

$$\begin{cases}y_{i}-\mathbf{w}^{\mathrm{T}}\phi\left(\mathbf{x}_{i}\right)-b\leq\varepsilon+\xi_{i} \\ \mathbf{w}^{\mathrm{T}}\phi\left(\mathbf{x}_{i}\right)+b-y_{i}\leq\varepsilon+\xi_{i}^{*} \\ \qquad\xi_{i}\geq0 \\ \qquad\xi_{i}^{*}\geq0\end{cases}, \qquad i=1,\ldots,N \tag{3.99}$$

In agreement with this description, the computation of inner products in the high dimensionality feature space is carried out. This can be another computationally costly task and determining $\phi(\cdot)$ complex. Notwithstanding, the problem can be obviated by resorting to appropriate kernel functions, K, defined by equation (3.100).

$$K\left(\mathbf{v}_{i},\mathbf{x}_{p}\right)=\left\langle\phi(\mathbf{v}_{i}),\phi\left(\mathbf{x}_{p}\right)\right\rangle \tag{3.100}$$

In order to be appropriate, the kernel functions should comply with Mercer's condition. In literature a large set of kernels can be found, being the linear, polynomial, sigmoid and RBF the most popular. The present work is based on examples for the linear (3.101) (simpler one) and RBF (3.102) (non-linear, general purpose) kernels.

$$K\left(\mathbf{v}_{i},\mathbf{x}_{p}\right)=\mathbf{v}_{i}^{\mathrm{T}}\mathbf{x}_{p} \tag{3.101}$$

$$K\left(\mathbf{v}_i, \mathbf{x}_p\right) = e^{-\gamma\left\|\mathbf{v}_i, \mathbf{x}_p\right\|^2} \qquad (3.102)$$

C, ε and, in the case of the RBF kernel, γ, are free hyperparameters, so called as they influence the optimization of α_i, α_i^* and b, themselves model parameters, and should be optimized prior to SVR use. Finally, the SVR conceptualization described above (also named ε-SVR) is not unique. Namely, Schölkopf et al. (2000) developed an alternative approach (termed ν-SVR) which substitutes the ε hyperparameter by ν (not related to $\mathbf{v}_i$, the ith training input). Unlike ε $[0,\infty[$, ν has a limited domain, $]0,1]$, and is directly related to the number of support vectors in use. Computations performed in the scope of this research resorted to the *LIBSVM* library (Chang and Lin 2011).

As SVR is used in the sine wave problem (Figure 3.23) the property of robustness to outliers emerges, particularly when relatively small values for the hyperparameter C are chosen. Also, it can be seen that, in this example, not only SVR models performed better than either MLPs or RBFNs in the training region, their extrapolations are much closer to the desired values.
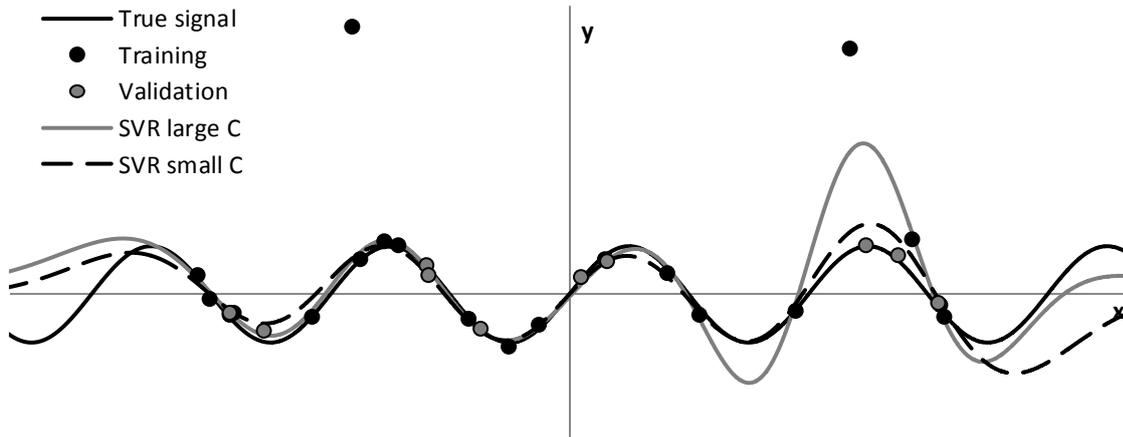
Figure 3.23. Two SVRs adjusted to reproduce a sine wave. Differences in the hyperparameter C control the models' response.

3.4.5 Least squares support vector machines

LS-SVR (Suykens and Vandewalle 1999, Suykens et al. 2002) is a variation of SVR, being the practical difference between the two models the fact that unlike the former, LS-SVR resorts to a quadratic loss function. As a consequence, LS-SVR predictions are computed from equation (3.103).

$$\hat{y}_p = \sum_{i=1}^{N} \alpha_i \cdot K\left(\mathbf{v}_i, \mathbf{x}_p\right) + b \qquad (3.103)$$

The $N+1$ regression parameters (α_i and b) are estimated by minimizing (3.104) subject to a set of N inequalities (3.105). In both expressions e_i represents the prediction error for the ith training observation ($e_i = y_i - \hat{y}_i$)

$$\frac{1}{2}\left(\mathbf{w}^T\mathbf{w}\right) + C\sum_{i=1}^{N} e_i^2 \qquad (3.104)$$

$$y_i \left[\mathbf{w}^{\mathrm{T}} \phi(\mathbf{x}_i) + b \right] \leq 1 - e_i, \qquad i = 1, \ldots, N \tag{3.105}$$

As for SVR, also LS-SVR training depends on the choice of hyperparameters. Here, however, hyperparameters are reduced to C and, in the case of RBF kernels, also γ. Because the model uses a quadratic loss function, it is not as robust to outliers as SVR. Although the issue is not addressed in the present contribution, work has been developed towards robust LS-SVR as well (e.g. Valyon and Horváth 2007). Due to the fact that its loss function does not consider any "insensitive" band, all training observations are actively used for predictions, potentially hindering computational performance. It should be said, however, that the *LS-SVMlab* toolbox (De Brabanter et al. 2010), used in the scope of the research, has shown little problems handling even large datasets.

When applied in the sine wave regression problem (Figure 3.24), LS-SVR achieves a relatively smooth reproduction of the sine wave and does not perform worse than the other models while extrapolating. Unlike SVR, however, it was substantially affected by both outliers.
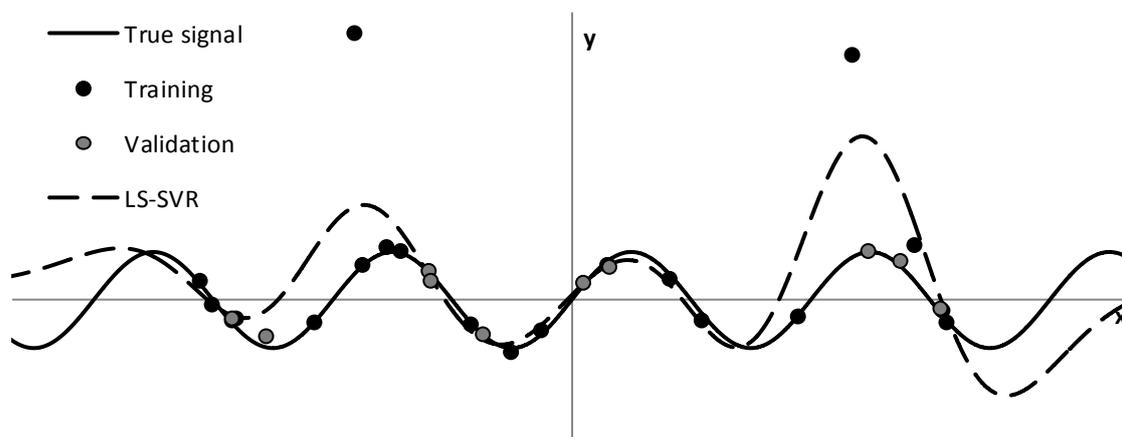
Figure 3.24. LS-SVR adjusted to reproduce a sine wave.

4   MODELING DATA FOR THE ZAMBEZI RIVER BASIN

4.1   On remote sensing data

Remote sensing data is a major asset for hydrology, particularly so when concerning large, poorly characterized remote areas, whose hydrologic features largely remain to be studied in detail. In the Zambezi, it can be said that the use of remote sensing data derived from spaceborne instruments is prevalent in hydrological studies.

Today, a wealth of data can be derived from satellite instruments, most of it not used in this work. Possibilities range from the preparation of Digital Elevation Models (DEM) (e.g. Moreira et al. 2004, Fujisada et al. 2005, Farr et al. 2007, Reuter et al. 2007, Berry et al. 2008, Lehner et al. 2008) to the valuation of basin-scale water storage time-series through the measurement of disturbances in Earth's gravitational field (e.g. Wahr et al. 1998, Klees et al. 2006, Winsemius et al. 2006b, Klees et al. 2008, Michailovsky 2008).

Along the way, other estimates relevant for hydrology can be highlighted, such as soil moisture contents (e.g. Kerr et al. 2001, Scipal et al. 2005, Wagner et al. 2007, Meier et al. 2011, Brown and Escobar 2014), vegetation state (e.g. Bannari et al. 1995, Gao 1996, van Leeuwen et al. 1999, Tucker et al. 2005, Swinnen and Veroustraete 2008), land cover classification (e.g. Tucker et al. 1985, Defries and Townshend 1994, Loveland et al. 2000, Mayaux et al. 2004), rainfall (e.g. Xie and Arkin 1996, Sorooshian et al. 2000, Joyce et al. 2004, Huffman et al. 2007, Kubota et al. 2007), evapotranspiration (e.g. Kustas and Norman 1996, Roerink et al. 2000, Kalma et al. 2008), or water levels and water-covered surfaces (e.g. Birkett 1994, Berry et al. 2005, Brakenridge et al. 2005, Brakenridge and Anderson 2006, Calmant et al. 2008, Bartsch et al. 2009, Birkett et al. 2011, Meier 2012, Michailovsky et al. 2012).

The extent of the research developed on the subject is overwhelming. In the present work special attention was given to DEMs, as the entire topology of the hydrological model depends on accurate topography, and rainfall, the main forcing variable of the land phase of the water cycle.

4.2   Global digital elevation maps

In basins as large as the ZRB, the only practical way of delineating the river network, the correspondent watersheds and subbasin features is to resort to DEMs. With this goal in mind, due to the basin's size and lack of detailed information covering a sizeable portion of it, freely available Global DEMs (GDEM) are the tool of choice.

The technology leading to the production of GDEMs has experienced a great evolution in recent years, with newer, more reliable products being released at a very fast pace. For hydrological

purposes, accurate topography is extremely important: firstly to correctly place rivers and draw watersheds, secondly to perform sound estimates of relevant features such as slopes, and lastly in order to properly estimate flooded areas. Being still limited in their precision (Chang et al. 2010, Karlsson and Arnberg 2011), GDEMs are mostly suitable for the first two tasks.

There are several approaches to the production of GDEMs, which rely on different technology. The GTOPO30 World elevation map, from 1996, was an early effort based on eight distinct data sources as no consistent data sets with global coverage existed at the time (Gesch et al. 1999). It succeeded in mapping the globe with a detail of 30 arc-seconds (approximately 930x930 m)[1] and varying accuracy.

A major quality leap was achieved by the Shuttle Radar Topography Mission (SRTM), which in mapped a large part of the World (±60° latitudes) in a 3 arc-second grid (approximately 90x90m) (van Zyl 2001). The SRTM dataset has some shortcomings related to the interferometry technique used to derive it. Namely, it loses accuracy in densely vegetated regions (as the signal does not penetrate the tree canopies), is subject to shadow effects in landscapes of pronounced slope, and faces problems in urbanized areas, deserts and water bodies due to their influence on the backscattered signal. Today, datasets where most of these problems have been corrected exist.

While products such as Altimeter Corrected Elevation 2 (ACE2) (Berry et al. 2008) or Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) (Danielson and Gesch 2011) are mostly based on SRTM C-band measurements, independent and newer alternatives, such as the Advanced Spaceborne Thermal Emission and Reflection (ASTER) GDEM version 2 (Tachikawa et al. 2011a), a product with a detail of 1 arc-second (close to 30x30m) derived by stereoscopy, exist.

Recently, SRTM X-band measurements (1 arc-second resolution) have been freely released (Marschalk et al. 2004). While displaying many gaps, the X-band data is relatively more accurate than the C-band's. Soon after the time of writing, the product of the TerraSAR-X and TanDEM-X satellites' detailed mapping of the World (0.4 arc-second resolution, about 12x12 m at the equator) with unprecedented accuracy (Krieger et al. 2007, Martone et al. 2013) will be available for paying customers and selected scientific applications (e.g. Köck et al. 2011).

In order to obtain the best hydrological results, several GDEMs were tested over the Zambezi River basin.

HYDRO1k
The HYDRO1k (U.S. Geological Survey 2001) GDEM, produced by the United States Geological Survey (USGS) and the United States National Center for Earth Resources Observation and Science (EROS), is based on the GTOPO30 World elevation map (Gesch et al. 1999) and was corrected for hydrological applications. According to Cohen Liechti (2013), the techniques used in its development are the following: projection of the DEM, identification of the natural sink features, filling of the DEM, and verification (comparison of streamlines generated to existing digitalized data). The vertical accuracy of the DEM is not available (Karlsson and Arnberg 2011).

With the same resolution as GTOPO30 (30-arcseconds) or roughly 930x930 m at the equator, the HYDRO1k GDEM was used in previous Zambezi studies (e.g. Cohen Liechti 2013). An illustration of HYDRO1k elevations covering the Zambezi is shown in Figure 4.1.

Although the use of HYDRO1k to delineate watersheds in the Zambezi leads to overall acceptable results, as will be seen at the end of this section, some areas (relatively small, but of the same order of magnitude of the area of a country like Switzerland) are drained erroneously.

---

[1] The resolution of the products depends on the latitude, being only approximate values given here.
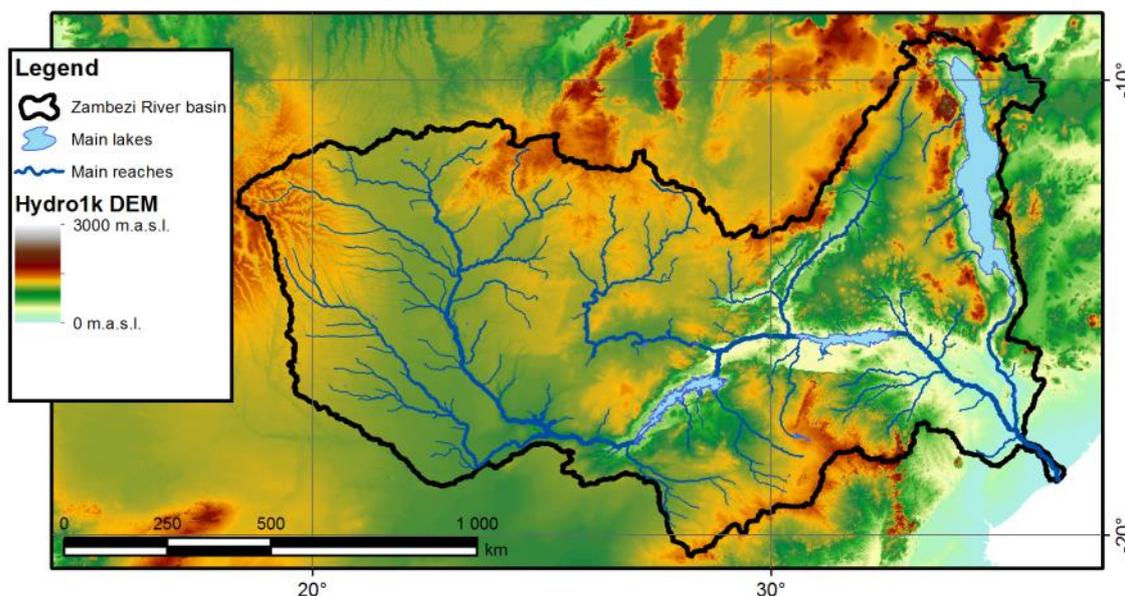
Figure 4.1. HYDRO1k elevation estimates for the Zambezi River basin area.

Shuttle Radar Topography Mission

SRTM has constituted a landmark as far as GDEMs are concerned. The product is based on measurements taken by an Interferometric Synthetic Aperture Radar (InSAR) with two receivers mounted on the Space Shuttle Endeavour and covers over 80% of the globe (coverage between 56°S and 60°N). While for the USA the model has a resolution of 1 arc-second, for the rest of the World it has been released at an artificially degraded resolution of 3 arc-seconds. The estimated 90% absolute error over the Southern African subcontinent is generally less than 5 m (Rodriguez et al. 2005).

The elevation estimates resulting from the interferometry technique that was used are affected by voids and artifacts due to terrain features that either block or otherwise substantially affect the backscattered radar signal. In order to fill voids and correct artifacts, several techniques have been used, resulting in four versions of the product.

Here, the void-filled SRTM DEM version 4.1 (Jarvis et al. 2008) from the CGIAR[2] Consortium for Spatial Information (CSI) is referenced. Using "finished grade" SRTM version 3[3] as a basis, the void-filling methodology, described by Reuter et al. (2007), includes the filling of small holes iteratively, the cleaning of the surface to reduce pits and peaks, and using a range of interpolation through the holes. The interpolation method employed depends on the size and surrounding topography and included information from higher resolution DEMs where possible.

The SRTM DEM version 4.1 elevation estimates for the Zambezi area are shown in Figure 4.2.

ACE 2

Despite the relatively low 90% absolute error, SRTM elevation data does fall outside the accuracy limit of ±16 m in localized areas (Berry et al. 2010) and displays a significant sensitivity to the vertical structure of vegetation as most of the signal received by the InSAR instrument is reflected by the upper canopy in dense rainforest areas (Kellndorfer et al. 2004).

altimeter data is interesting as it reflects most strongly from the ground surface underneath the canopy (Smith and Berry 2010).

---

[2] Originally the acronym for Consultative Group on International Agricultural Research. In 2008 the CGIAR underwent a major transformation, yet kept the CGIAR "designation".

[3] In 2013 NASA released a SRTM version 3 plus product, with all voids filled. The process included data from ASTER GDEM version 2 and GMTED2010 data.
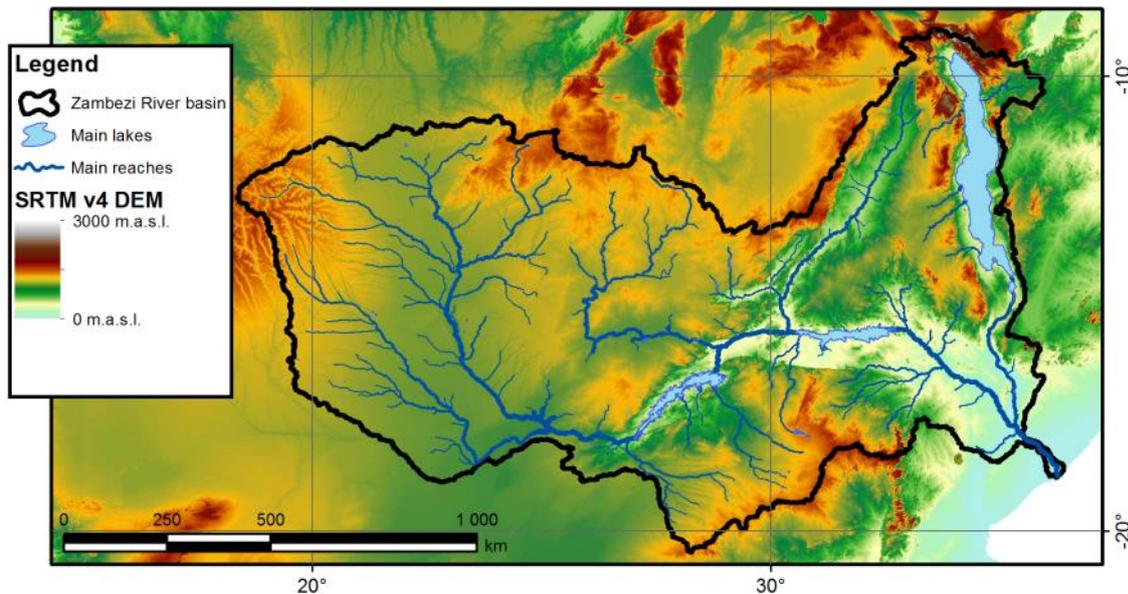
Figure 4.2. SRTM v4.1 elevation estimates for the Zambezi River basin area.

Altimetry point measurements from spaceborne radar instruments can be extremely precise over ocean areas (5 cm or better), but over land the complexity of the return echoes poses a number of problems. Despite this, once the challenges posed by the complex echoes are solved, the

The ACE 2 dataset (Berry et al. 2010) addresses the SRTM deviations over densely vegetated areas by merging its "final grade" product with a DEM resulting from the interpolation of altimetry measurements.

The representation of the ACE 2 elevation over the ZRB area is not shown as differences from Figure 4.2 are hardly perceivable.

HydroSHEDS
Similarly to HYDRO1k, the Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS) DEM aims at hydrological applications. It is based on SRTM version 3 "unfinished data" and earlier DEM products complemented by vector information concerning rivers and water bodies such as the SRTM Water Body Data, the Digital Chart of the World global vectorized river network, the ArcWorld vectorized river network and the Global Lakes and Wetlands Database (Lehner et al. 2008).

The SRTM data was void-filled and corrected for spurious sinks and depressions. Most relevantly, the data was then conditioned by "burning" the known major stream paths into the base DEM and modifying adjacent areas. Also, the open water surfaces were artificially deepened in order to force derived flow paths to stay within them (Lehner et al. 2006).

Again, due to similarities with the SRTM data, HydroSHEDS elevation estimates for the Zambezi area are not displayed.

ASTER
An independent GDEM estimate from SRTM is the ASTER GDEM version 2, released in 2011.

Compared to SRTM, ASTER has a wider global coverage (±83° latitudes), finer resolution (1 arc-second or approximately 30x30 m over the equator) (Tachikawa et al. 2011a), and is based on entirely different technology supported by the ASTER stereoscopic capability. As the ASTER instrument is affected by clouds, areas of permanent cloud cover present gaps in the data.

Despite the earlier specifications, due to errors in the data the practical horizontal resolution for the data is estimated to be of 2.4 arc-seconds and the absolute vertical errors at the 95% confidence interval have been estimated at 17 m (Tachikawa et al. 2011b).

The ASTER GDEM version 2 elevation estimates for the Zambezi River basin area are depicted in Figure 4.3.
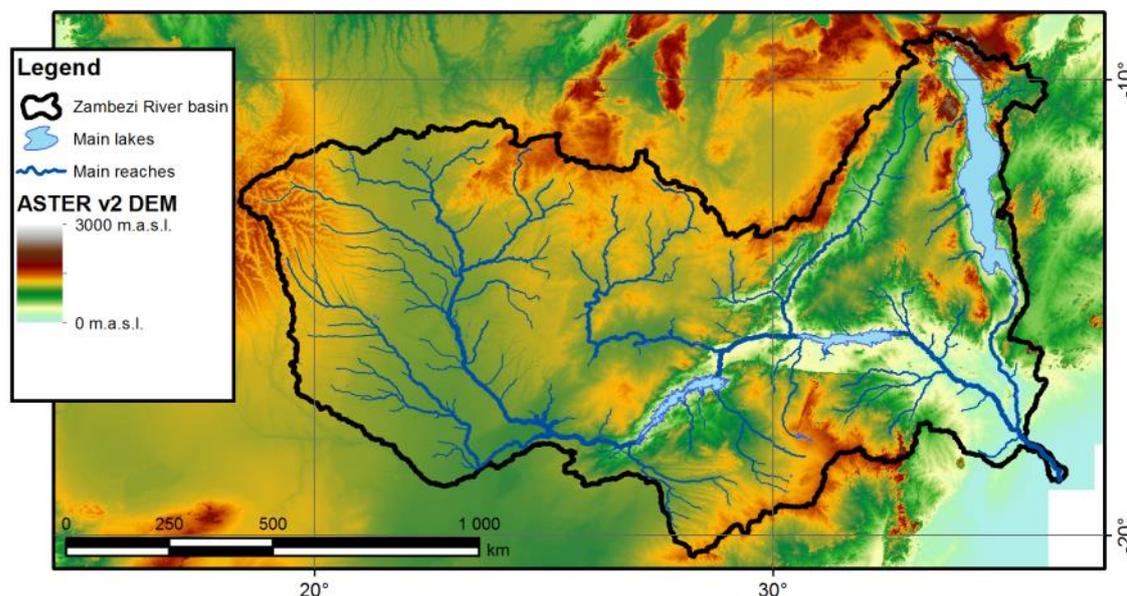
Figure 4.3. ASTER GDEM version 2 elevation estimates for the Zambezi River basin area.

Comparison

Numerous works have been carried out on the comparison of different GDEMs (e.g. Marschalk et al. 2004, Chang et al. 2010, Hirt et al. 2010), with comparisons being generally favorable to SRTM data amongst the freely available models.

Regarding hydrology applications, the use of DEMs and their associated uncertainty raises several concerns (Wechsler 2007) and the quality of derived hydrological parameters cannot be directly associated with reported mean absolute or relative errors[4].

In the scope of the present work and in order to derive the proposed hydrological model, all the aforementioned DEMs were subjectively tested regarding their potential to derive an accurate model of the Zambezi. This is justified due to the disagreements between datasets.

As an example of such disagreements, the difference between ASTER GDEM version 2 and SRTM version 4.1 data over the area of interest is illustrated in Figure 4.4. ASTER was the newest dataset, having also the best horizontal resolution, and therefore held large promise for the purposes of the present work. This was not verified. Clearly, substantial deviations whose width is equal to ASTER's swath are visible (which implies that the problem lies in the ASTER data). Understandably, even after due processing of the ASTER GDEM prior to watershed derivation, these abrupt differences effectively constitute barriers to the flow and lead to gross errors.

Other enlightening comparison can be made between the HydroSHEDS conditioned DEM and SRTM version 4.1. The computed difference between both products is shown in Figure 4.5, which is marked by the lower elevations of the main river streams and water bodies of the HydroSHEDS conditioned DEM.

---

[4] This is easily understandable as, under particular conditions, a very small number of erroneously estimated pixel elevations is enough to deviate a river stream and, consequently, misplace the contributions of arbitrarily large watersheds.
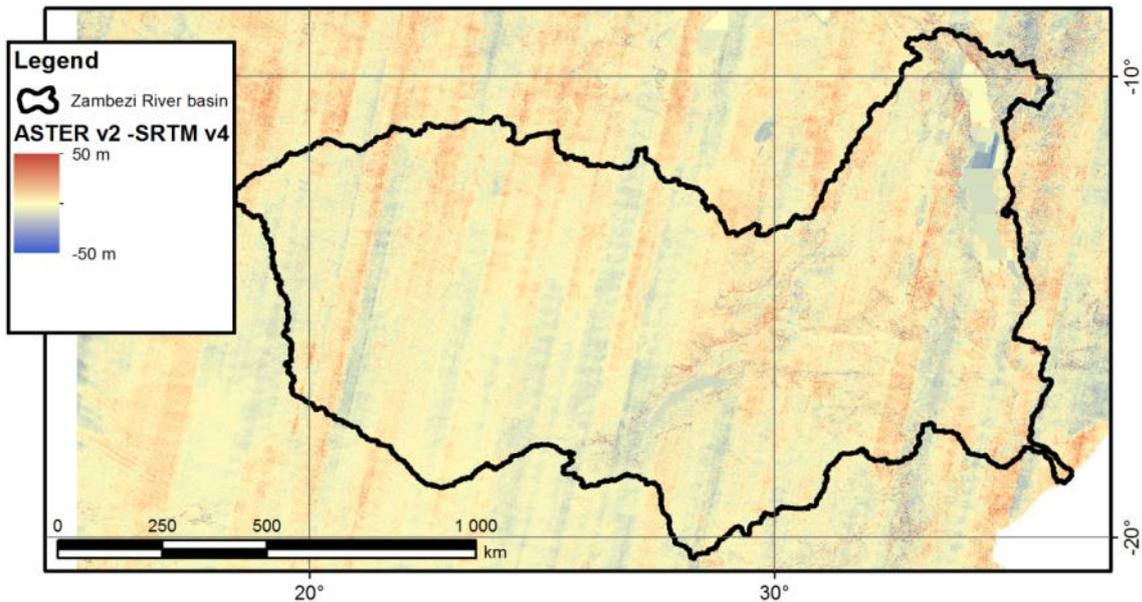
Figure 4.4. Difference between the ASTER GDEM version 2 and the SRTM version 4.1 data over the Zambezi River basin area.

While the HydroSHEDS conditioned DEM provides excellent results when used to derive the river network[5], the consistently lower elevations at certain points raise questions as to the validity of topographic parameters for hydrologic analyses such as slopes.
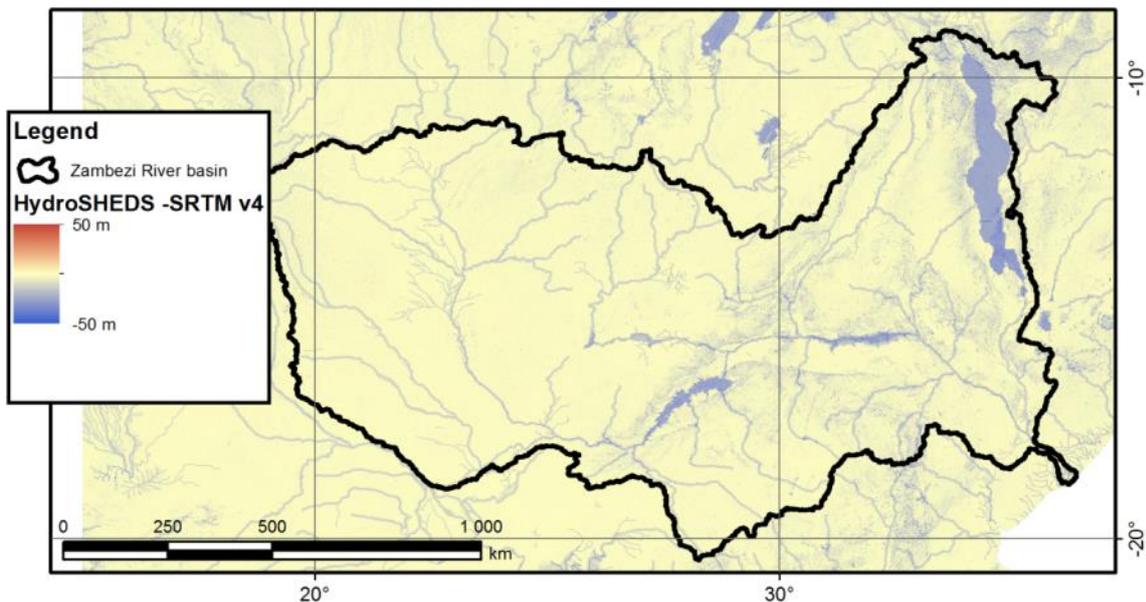
Figure 4.5. Difference between the HydroSHEDS conditioned DEM and the SRTM version 4.1 data over the Zambezi River basin area.

Another plot overlays the river networks derived by HydroSHEDS and HYDRO1k (Figure 4.6). If at first glance and due to the large scale, both results appear to have a good agreement, the areas that are drained to wrong watersheds are very relevant, which compromises the application of HYDRO1k.

---

[5] The (subjective) comparison was made by overlaying the derived river networks with Google Earth satellite imagery.
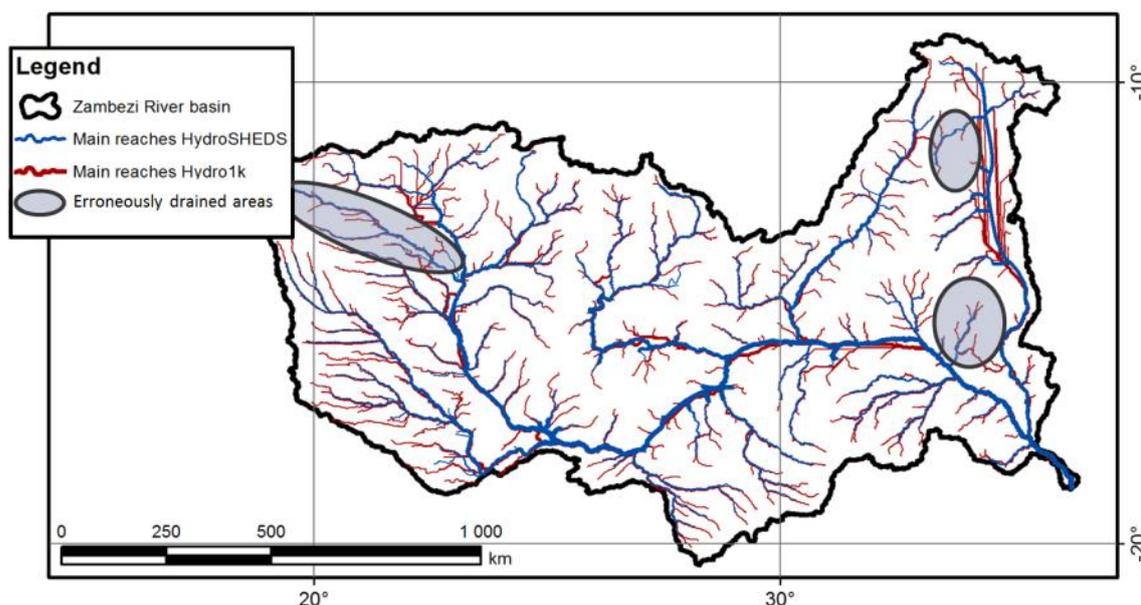
Figure 4.6. HydroSHEDS- and HYDRO1k-derived river networks overlaid within the Zambezi River basin. Most relevant erroneously drained areas highlighted in gray.

The main reach of the Zambezi between the Kariba and Cahora Bassa dams was analyzed in further detail. Nearly 90 cross-sections of the Zambezi Valley were evaluated regarding the data contained in the HydroSHEDS, SRTM version 4.1, ASTER GDEM version 2, and the Google Earth (GE) DEM.

The GE DEM is used to provide elevations to the globe featured in the equally named software. It is mostly based on SRTM data but contains contributions for several other sources in poorly characterized areas. Software was written in the Java language in the scope of the present work specifically for the purpose of retrieving these GE elevations.

The location of the evaluated cross-sections is presented in Figure 4.7 along with information on the locations of three cross-sections chosen to be represented below.
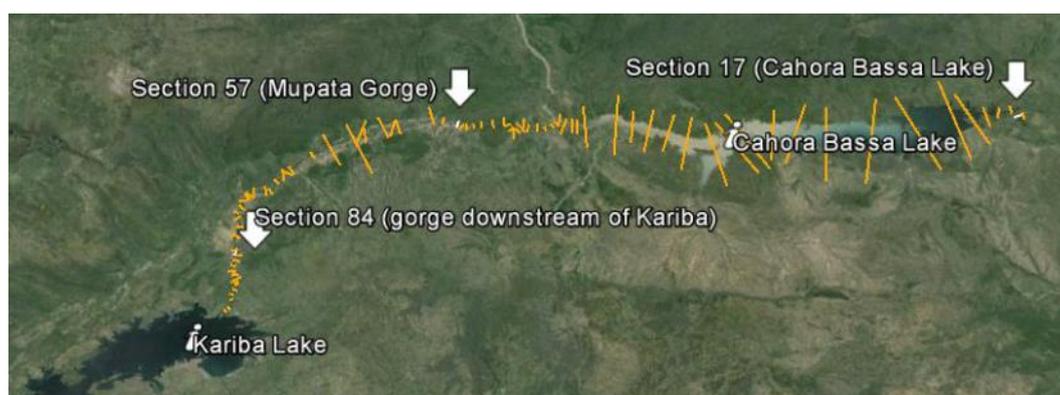
Figure 4.7. Scheme of the evaluated cross-sections of the Zambezi Valley between the Kariba and Cahora Bassa dams.

Data for a cross-section in the gorge downstream of the Kariba dam, along with another at the entrance of the Mupata Gorge, are plotted in Figure 4.8 and Figure 4.9, respectively. While SRTM version 4.1 and GE data tend to agree, ASTER GDEM version 2 actually seems to exaggerate the elevation of the actual streams – a behavior which is obviously undesirable for hydrologic applications. On the contrary – but as expected – HydroSHEDS presents underestimated elevations.
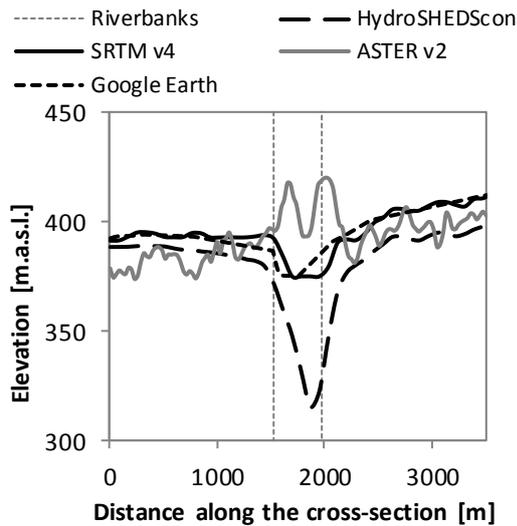
Figure 4.8. Cross-section of the Zambezi Valley in the gorge downstream of the Kariba dam (section 84).
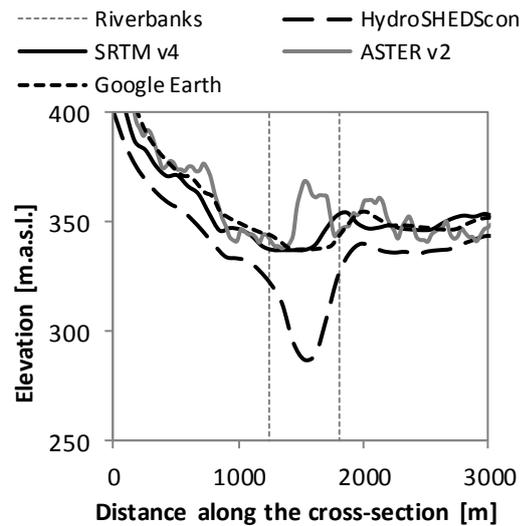
Figure 4.9. Cross-section of the Zambezi Valley at the entrance of the Mupata Gorge (section 57).

Another cross-section, within the Cahora Bassa Lake, was assessed (Figure 4.10). Here, an additional resource was available in the form of a topographic map from the Portuguese *Instituto Geográfico do Exército* (IGeoE) covering the area of the valley prior the completion of the Cahora Bassa dam and the rise of the water level[6] (Figure 4.11). All models perform reasonably well in this cross-section given their nature and the limitation imposed by the lake's surface. Despite this, SRTM version 4.1 appears to provide the best fit.
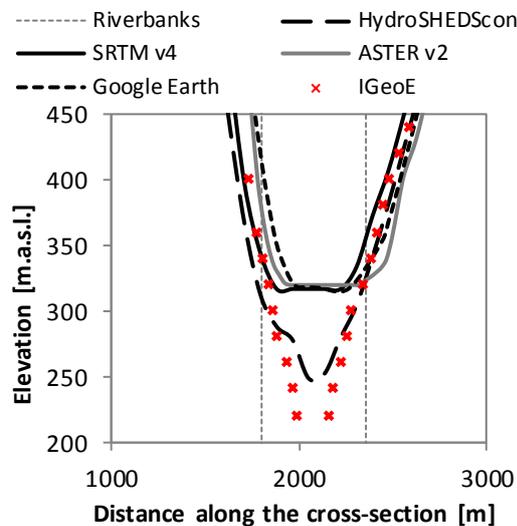
Figure 4.10. Cross-section of the Zambezi Valley in the Cahora Bassa Lake close to the dam (section 17).
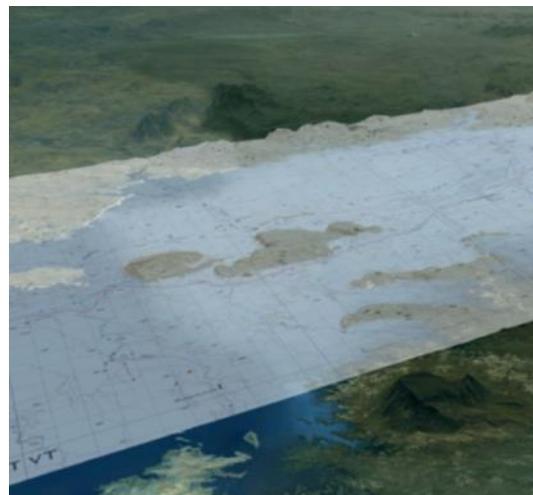
Figure 4.11. Topographic map of the Cahora Bassa Lake Valley prior to the dam completion.

Being established that the choice of an appropriate DEM is relevant for hydrological applications and that resolution or mean vertical errors are not necessarily linked to its reliability when used for such purposes, after the described analysis the HydroSHEDS conditioned DEM was found to be the best alternative in order to derive the river network and watersheds, while SRTM version 4.1 is recommended for the calculation of topographic parameters.

---

[6] The information contained in this map was crucial to the establishment of the volume-elevation curve of the Cahora Bassa Lake which was used in the proposed hydrological model.

4.3 Rainfall estimates

Rainfall is extremely important hydrological variable (e.g. Yapo et al. 1996, Andréassian et al. 2001, Fekete et al. 2004, Meskele et al. 2009). In the ZRB, difficult access to historical series of gauged records well distributed in space practically forces large scale modelers to resort to remote sensed rainfall. This was, in effect, observed in most, if not all, recent studies (for reference see Section 3.1).

Today, several products provide daily to sub-daily rainfall estimates that cover most of the World (Figure 4.12). Although they can differ widely in regard to their estimates, the basic principles of rainfall estimation from space platforms are shared among them.

Figure 4.12. Mean monthly rainfall (mm) in the ±60° latitude band from 1998 to 2011 (TMPA 3B42 v7a data).

Due to the high spatial and temporal variability of rainfall fields, achieving accurate remote sensing rainfall estimates presents an extremely difficult challenge. Three main kinds of instrument are used in order to achieve *quasi*-Global rainfall estimates:

- The backbone of satellite rainfall estimates are passive Microwave (MW) sensors mounted on satellites flying relatively close to the Earth in polar orbits. MW information is interesting as it penetrates the clouds, actually detecting precipitation particles beneath them. In these orbits, the satellites' altitude is relatively low and they move fast in relation to the Earth's surface. This, along with limitations on the sensors' swath, means that a given part of the World is only observed by a given polar-orbiting satellite from time to time, typically twice each day or with an interval of up to a few days (depending on the orbit and the sensor's swath).
  Due to the variability of the rainfall field, several measurements from passive MW sensors on a constellation of satellites are advised to properly characterize it.
- Passive MW coverage is not permanent and can be complemented with less informative Infrared (IR) and Visible (VIS) sensors. These can be mounted on polar-orbiting satellites or on distant platforms placed on geostationary orbit (such as the Meteosat series). The cloud-top brightness temperatures derived from IR can be correlated to cloud height and rainfall rates. VIS data provides complementary information about cloud thickness and possibility of producing precipitation.
- Finally, rainfall can be best estimated by using active MW sensors. These sensors, however, consume more power than passive ones and have smaller swaths, which translates into longer revisit times. So far, however, the sole orbiting active MW sensor adapted to estimate rainfall is the Precipitation Radar (PR) (Kozu et al. 2001) onboard the Tropical Rainfall Measuring Mission (TRMM) satellite (Simpson et al. 1996).

The IR and VIS images are only indirectly related to the surface precipitation rate. Also MWs do not "measure" surface precipitation rate, being more prone to capture a weighted vertical integration of rain and/or ice water content. Additionally, both region and season influence the observed signals over land areas (Petty and Krajewski 1996).

Historically, numerous satellite-derived rainfall estimates have been proposed (e.g. Petty and Krajewski 1996). Below, some of the arguably most promising datasets (TMPA 3B42, RFE2.0,

CMORPH, PERSIANN, and GSMaP MVK) working at fine spatial scales and reporting at sub-daily to daily time steps are characterized.

In the near future, advances are expected in the field of satellite rainfall products with the Global Precipitation Measurement (GPM) mission whose bold goal is to provide the next-generation of rain and snow worldwide estimates each three hours (Kidd and Huffman 2011). In fact, in the first months of 2014 GPM's core observatory satellite is being placed in orbit with a specialized 13-channel microwave imager and a dual-frequency precipitation radar in its payload.

TMPA 3B42

TRMM Multisatellite Precipitation Analysis (TMPA) provides a calibration-based sequential scheme for combining precipitation estimates from multiple satellites as well as gauge analyses (where feasible) (Huffman et al. 2007).

The TMPA 3B42 algorithm covers the globe between the 50°S and 50°N latitudes and is published with a 0.25°×0.25° spatial resolution encompassing periods of three hours. With a data series from 1998 to present, the algorithm has undergone several modifications. In late 2012 TMPA 3B42 version 7a was released.

In the derivation of the research grade TMPA data, contributions from several passive MW sensors (TMI, SSM/I, AMSR-E and AMSU-B), rainfall estimates from IR sensors, combined estimates from TRMM's Microwave Imager (TMI) and PR (TRMM Combined Instrument, TCI), and interpolated ground rainfall measurements are used.

Towards the final product, the best MW estimates from the different sensors, calibrated based on TCI and TMI observations, are merged resulting in a high quality MW precipitation measurement. The next step is to calibrate IR brightness temperatures using the aforementioned high quality measurement. Calibrated MW and IR data are then merged, being MW's taken "as is" where available, promoting local accuracy at the expense of a more homogeneous dataset. The final step is to scale the MW-IR estimates resorting to rain gauge data, which is done based on monthly aggregations (Huffman et al. 2007).
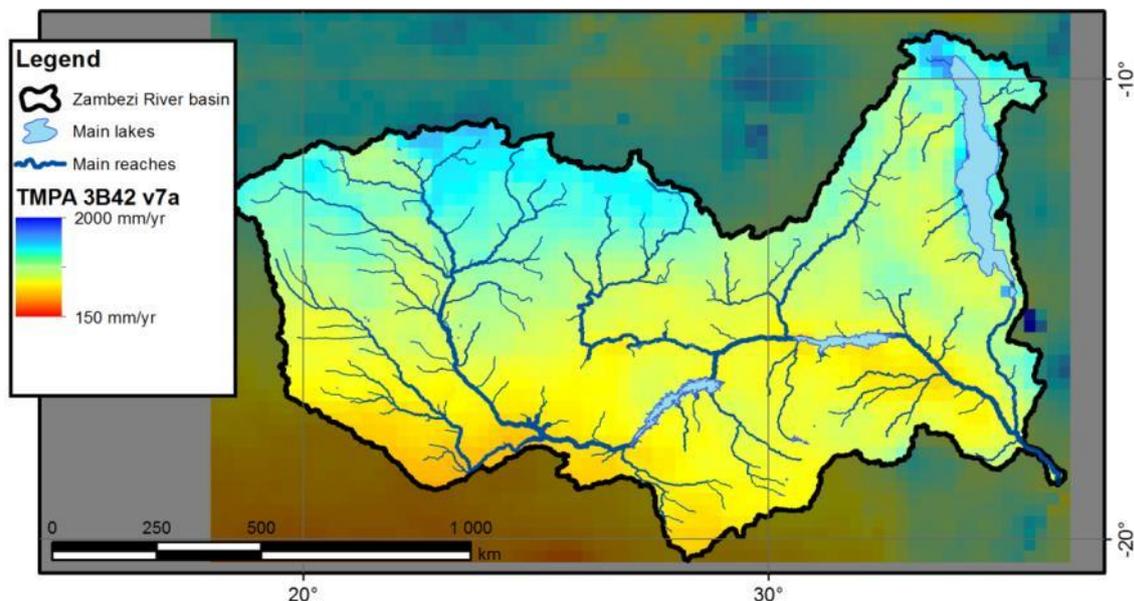
Figure 4.13. Mean annual rainfall estimations from the TMPA 3B42 version 7a product over the ZRB area (January 1998 to December 2011).

NOAA/FEWS RFE2.0

The National Oceanic and Atmospheric Administration (NOAA)/ Famine Early Warning System (FEWS) Rainfall Estimate 2.0 (RFE2.0) is produced daily for the African Continent[7], covering the period from January 2001 to present with a spatial resolution of 0.1°×0.1°.

The algorithm used to produce the RFE2.0 estimates is based on the work by Xie and Arkin (1996). Four data sources are used: daily rain Global Telecommunications Station gauge data and the following precipitation estimates: AMSU-B MW, SSM/I MW, and IR cloud-top temperature-based.

Due to the fact that, separately, each input source is incomplete in spatial coverage and contains non-negligible random error and systematic bias (NOAA Climate Prediction Center 2001), a two-step merging process is carried out. After the individual sources of data are processed, the IR and two MW precipitation estimates are linearly combined according to a maximum likelihood method based on their estimated random error. In a second step, the resulting estimate is compared and combined with rain gauge data. Here, the "shape" of the rainfall field is obtained from the satellite observations, while the amplitude of the rainfall is defined by the observations at rain gauges (where a sufficient number is available) (Reynolds 1988, Xie and Arkin 1996).

The mean annual RFE2.0 rainfall estimation for the ZRB area is presented in Figure 4.14.
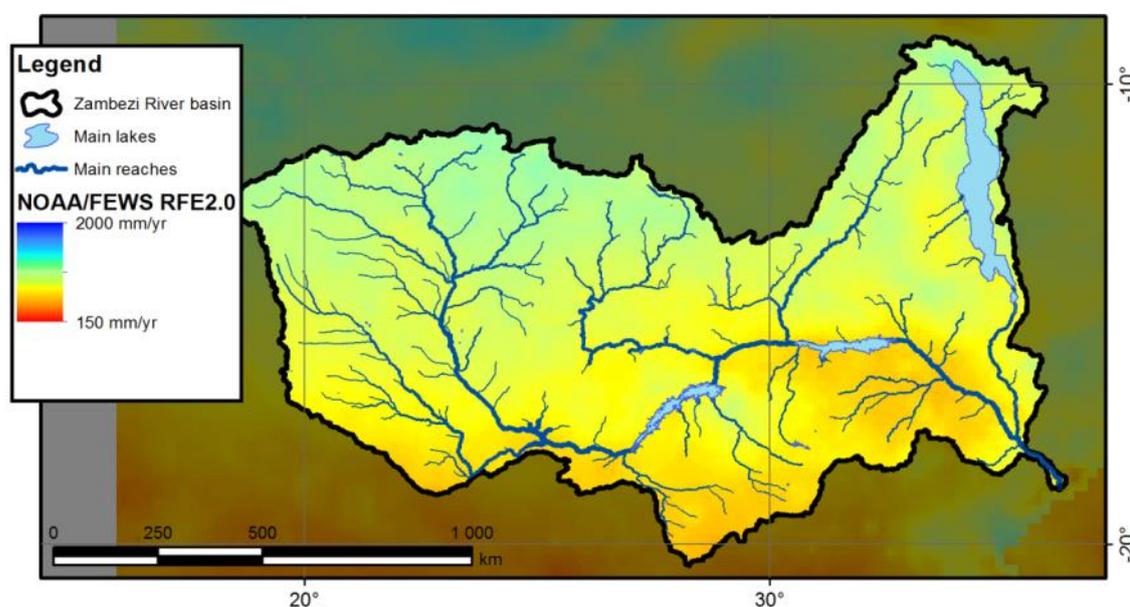
Figure 4.14. Mean annual rainfall estimations from the NOAA/FEWS RFE2.0 product over the ZRB area (January 2001 to December 2009).

CMORPH

The Climate Prediction Center Morphing Method (CMORPH) presents yet another approach at rainfall estimation. Being a morphing technique, it does not simply "merge" data from several sources, but uses motion vectors from IR observations to propagate the relatively high quality precipitation derived from MW data (Joyce et al. 2004).

The basic idea underlying the morphing concept is that, as IR precipitation estimates are affected by a number of problems, using only estimates derived from passive MW sensors is advantageous. Unlike other algorithms, which employ IR precipitation in areas where no MW coverage is available, CMORPH uses IR indirectly, solely as a guide to propagate the MW features between valid observations.

---

[7] The data is produced overland for the 40°S-40°N / 20°W-55°E window.

The CMORPH dataset produces estimates every 30 minutes in the ±60° latitude band, with a spatial resolution of 0.0727°×0.0727°. MW rainfall from the TMI, SSM/I, AMSU-B and AMSR-E sensors is used in its derivation.

The main steps of the algorithm are: passive MW rainfall ordering and mapping in a fine-resolution grid; surface snow and ice scanning, as it affects passive MW sensors; normalization of the AMSU-B estimates; derivation of half-hourly cloud system advection vectors from IR data and application of bias corrections; forward and backward propagation of MW rainfall estimates using the derived cloud system; and morphing the rainfall fields by a weighted combination of the forward and backward propagated fields.

The mean annual CMORPH rainfall estimation for the ZRB area is presented in Figure 4.14. Here, the data is displayed at a 0.25°×0.25° spatial resolution, which was adopted in this work. Also, it should be noticed that the data that was used covers the period starting from December 2002 to present and was the product of an evolving algorithm designated CMORPH version 0.x. A newly released CMORPH version 1.0, computed with a stabilized algorithm, covers data from 1998 to present.
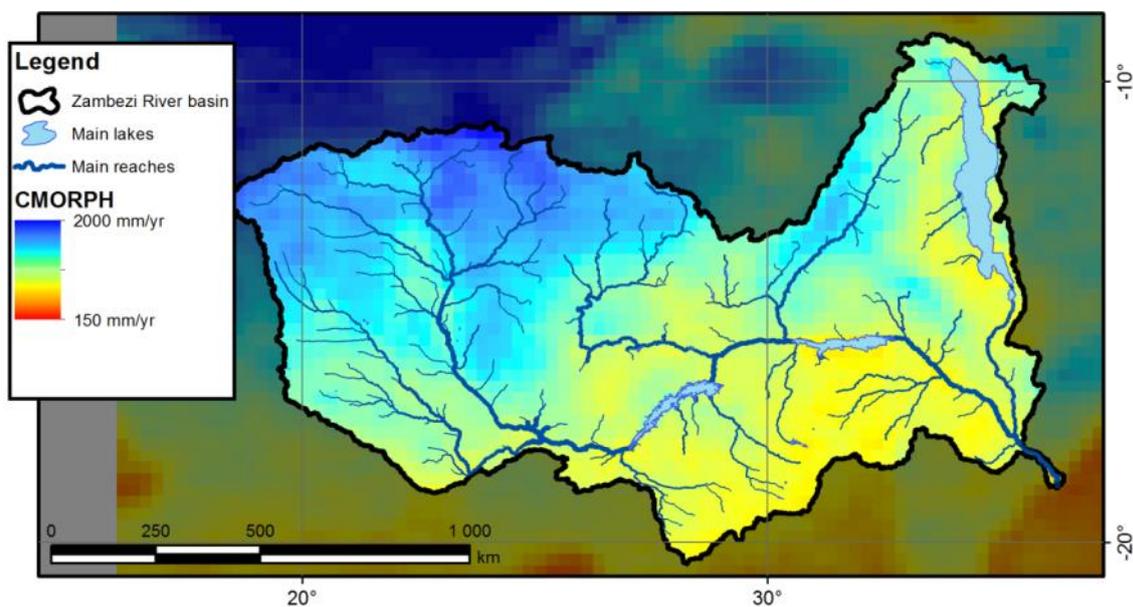
Figure 4.15. Mean annual rainfall estimations from the CMORPH version 0.x product over the ZRB area (January 2003 to December 2009; the data is shown at the 0.25°×0.25° spatial resolution).

PERSIANN

The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) is designed to be capable of extracting and combining information from data of various types including, for example, IR and microwave satellite imagery, rain gauge and ground-based radar data, and ground-surface topographic information (Hsu et al. 1997).

Various contributions describing PERSIANN enhancements have been published and retrieving exact details about the characteristics of the published dataset[8] was not possible. In general, the PERSIANN algorithm starts by the application of a SOM[9] ANN in order to classify inputs relevant to rainfall estimation (such as topography, IR-derived and MW-derived precipitations in the vicinity of the location of interest, radar data, variability of these fields, etc.). The activated node of the SOM layer is then linearly transformed in order to produce estimates, being that the coefficients of this transformation are adaptively adjusted in order to calibrate the model (Hsu et al. 1997, Sorooshian et al. 2000).

---

[8] The used PERSIANN dataset can be found at http://chrs.web.uci.edu/persiann/data.html
[9] Self-Organizing Map. An example of a SOM at work is described in Section 3.4.1.

This set-up grants PERSIANN great flexibility, an example of which are the sub-products developed to date, such as PERSIANN-GT (Sorooshian et al. 2000), PERSIANN-CCS and PERSIANN-GCCS (Hong et al. 2007).

PERSIANN data is available from March 2000 to present and covers the ±60° latitude band. In the scope of this thesis, a 3-hourly aggregation with a spatial resolution of 0.25°×0.25° was employed. PERSIANN estimates for the ZRB area are depicted in Figure 4.16.
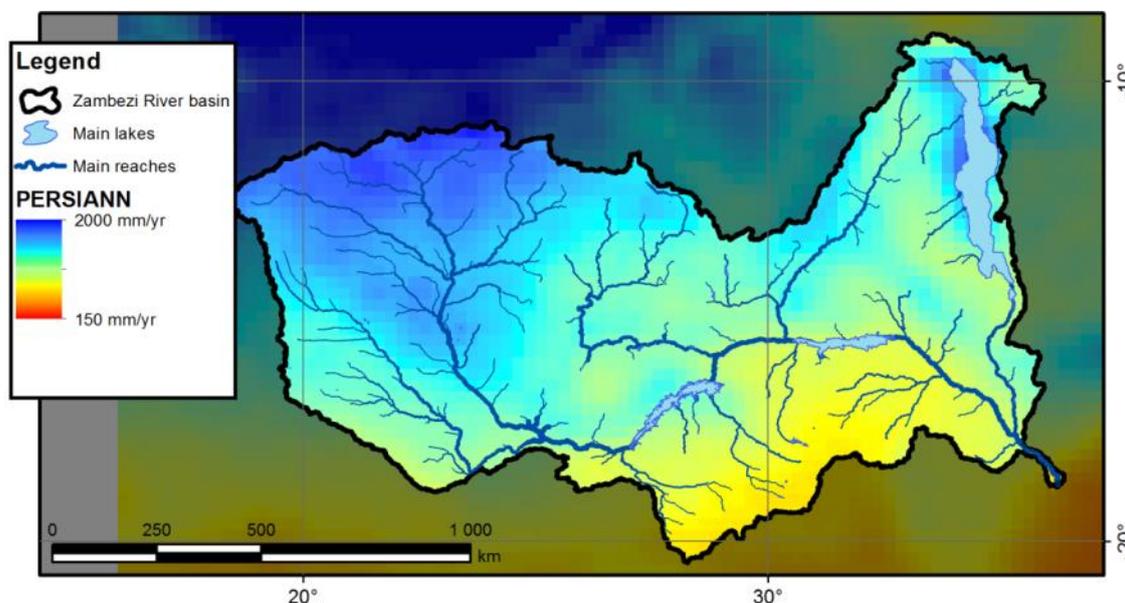
Figure 4.16. Mean annual rainfall estimations from the PERSIANN product over the ZRB area (January 2003 to December 2009).

GSMaP MVK

The most recent satellite rainfall product at this time is the Global Satellite Mapping of Precipitation (GSMaP) (Kubota et al. 2007).

GSMaP makes use of hydrometeor profiles derived from the PR onboard TRMM, statistical rain/no-rain classification and combined scattering-based surface rainfalls computed depending on rainfall intensities. In short, the model builds upon data retrieved by the TRMM mission in order to enhance the rainfall estimates of passive MW sensors (GSMaP MWR).

The GSMaP Moving Vector with Kalman filter method (MVK) product processes GSMaP MWR in a similar fashion to CMORPH's. In GSMaP MVK, however, a Kalman filter is used in order to refine the precipitation rates of the propagated fields (Ushio et al. 2009).

The GSMaP MVK dataset is available for every hour at a 0.1°×0.1° spatial resolution between the 60°S and 60°N latitudes. Similarly to CMORPH, GSMaP MVK makes use of the TMI, SSM/I and AMSR-E MW sensors.

In this work GSMaP MVK version 5.222 was used. Figure 4.17 contains its estimates over the ZRB.

Comparison

To date, numerous comparisons have been carried out regarding the previously mentioned rainfall estimates. Although facing a number of issues (e.g. Petty and Krajewski 1996, Joyce et al. 2004), benchmarking against gauged surface records and land-based precipitation radars is useful in order to understand the capabilities and limitations of each product.
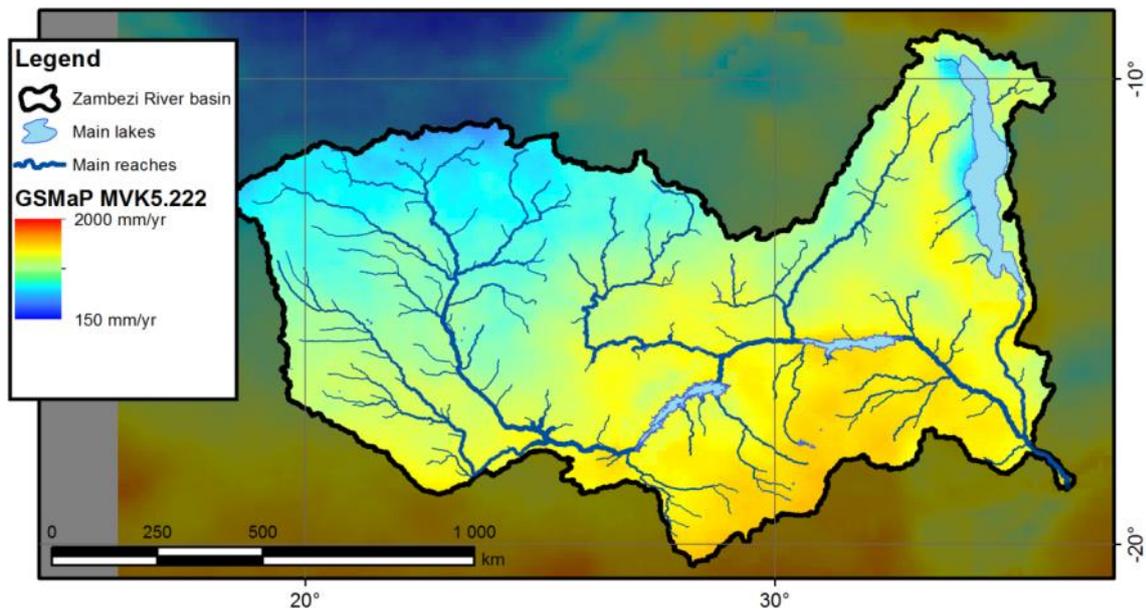
Figure 4.17. Mean annual rainfall estimations from the GSMaP MVK version 5.222 product over the ZRB area (March 2001 to December 2009).

Besides the validation assessments that are commonly published accompanying newly released products, broad comparisons are particularly informative. Nearly all major efforts have targeted specific areas such as Japan, the USA, Europe or Australia, where access to surface validation data is easier (e.g. Ebert et al. 2007).

The perhaps more relevant comparisons focused on the African Continent are not so numerous or broad-reaching. Jobard et al. (2011) provide a solid overview of the performance of 10 rainfall products at a 10-day time step over the West African monsoon region. Their results highlight the usefulness of incorporating gauge measurements into the estimates. Another relevant finding was that local products (such as NOAA/FEWS RFE2.0) perform best. Dinku et al. (2007) published somewhat conflicting results in a 10-daily comparison for East Africa, where NOAA/FEWS RFE2.0 led to poor estimates under certain criteria. These two examples highlight the need for site-specific validation.

Focused on the ZRB, Cohen Liechti et al. (2011) provided a summary of satellite rainfall products' evaluations over Africa and compared the TMPA 3B42 version 6, NOAA/FEWS RFE2.0 and CMORPH products to ground measurements and the Global Precipitation Climatology Centre's (GPCC) full data reanalysis product version 4 by looking at several time steps (daily, 10-daily and monthly).

Along with other publications (e.g. Huffman et al. 2007), the study shows a large scattering of the tested products for the daily time step (Figure 4.18) which tends to be smoothed out for longer aggregations.
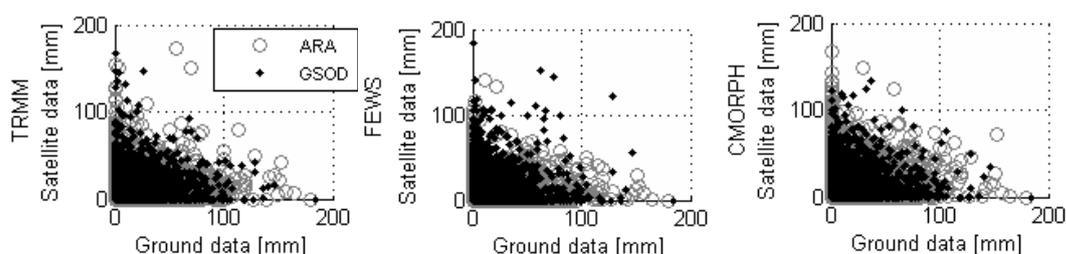
Figure 4.18. Scatter plots of TMPA 3B42 version 6 (TRMM), NOAA/FEWS RFE2.0 (FEWS) and CMORPH daily estimates and gauged records (from ARA-Zambeze and GSOD, NOAA's Global Surface Summary of the Day). Adapted from Cohen Liechti et al. (2011).

Of the three data sets, and in line with the results from Jobard et al. (2011), the CMORPH estimate provided the worst results. TMPA 3B42 version 6 and NOAA/FEWS RFE2.0 performed similarly. Such findings are not surprising as both TMPA 3B42 and RFE2.0 benefit from ground information.

Daily scatter plots depict somewhat bleak prospects regarding the application of such satellite rainfall estimates for detailed hydrological modeling. The spatial scales involved in the ZRB, however, have the potential to average out some of the estimate's random errors.

Notwithstanding, the average rainfalls valued by these products also show worrying discrepancies. From Figure 4.19 to Figure 4.23 differences in terms of their mean annual rainfall estimates over the ZRB are shown for all the described products.
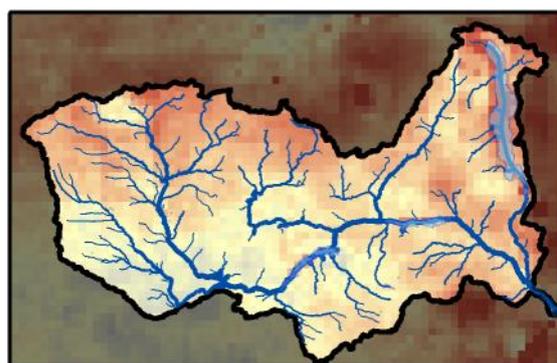
Figure 4.19. NOAA/FEWS RFE2.0 - TMPA 3B42 v7a mean annual rainfall estimates from 2001 to 2009.
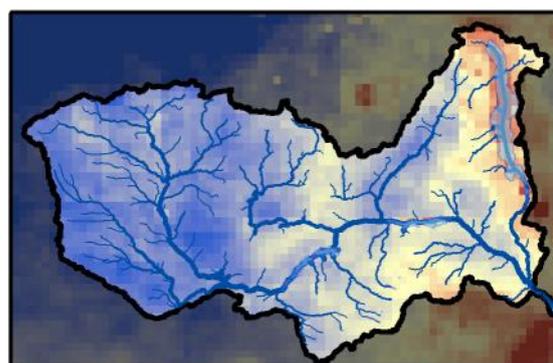
Figure 4.20. CMORPH v0.x - TMPA 3B42 v7a mean annual rainfall estimates from 2003 to 2009.
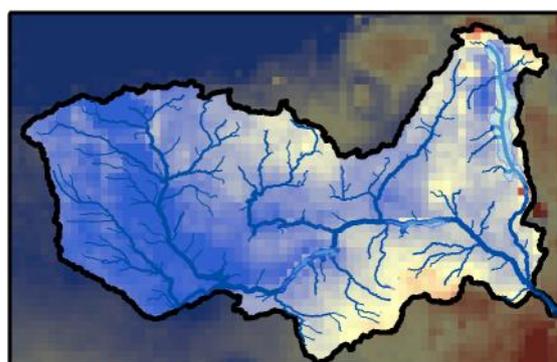
Figure 4.21. PERSIANN - TMPA 3B42 v7a mean annual rainfall estimates from 2003 to 2009.
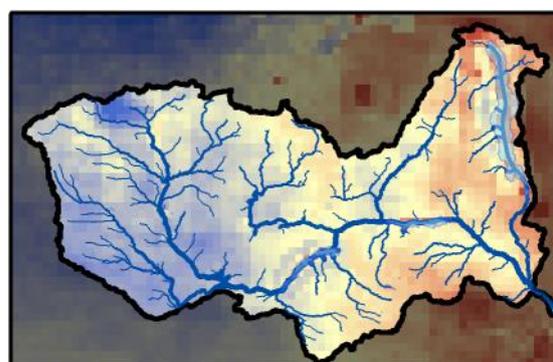
Figure 4.22. GSMaP MVK 5.222 - TMPA 3B42 v7a mean annual rainfall estimates from 2001 to 2009.

Some of the observed differences, of more than 500 mm/yr, can be put into perspective by considering the mean annual rainfall value over the Zambezi, which is according to the best estimates of approximately 1000 mm/yr. CMORPH and PERSIANN in particular show an acknowledgeable tendency towards overestimation.

Similarly worthy of notice is the difference between TMPA 3B42's versions 6 and 7a (Figure 4.23), also reported in other parts of the World (Zulkafli et al. 2013). In spite of new evaluations of TMPA 3B42

version 7a's product performance over the ZRB being lacking, Cohen Liechti (2013) reported substantial gains by using this data as forcing to a hydrological model of the basin.

In a temporal analysis of the three best adjusted products (TMPA 3B42 7a, NOAA/FEWS RFE2.0 and GSMaP MVK 5.222) it can be seen that besides the magnitude of the rainfall, also its annual distribution is captured differently from product to product. As an example, average monthly rainfall aggregations for the Kafue, Shire River and Chobe subbasins (see Figure 4.24 for reference) are depicted along with 5, 25, 75 and 95% quantiles from Figure 4.25 to Figure 4.31.
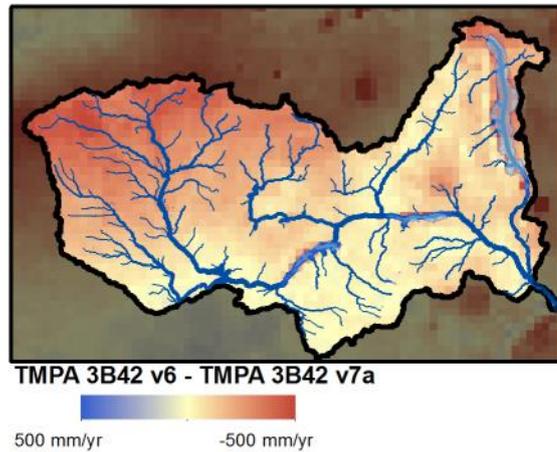
Figure 4.23. TMPA 3B42 v6 - TMPA 3B42 v7a mean annual rainfall estimates from 1998 to 2009.

Figure 4.24. Main subbasins of the ZRB.
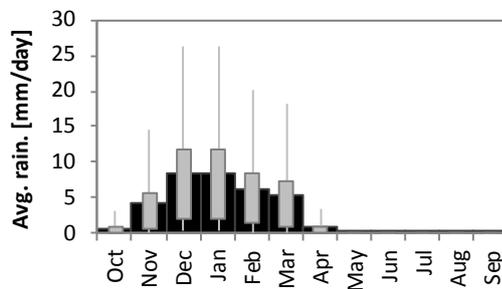
Figure 4.25. Boxplot of monthly TMPA 3B42 v7a rainfall from 2003 to 2009 over the Kafue subbasin.
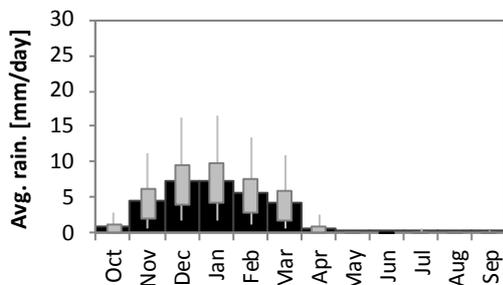
Figure 4.26. Boxplot of monthly NOAA/FEWS RFE2.0 rainfall from 2003 to 2009 over the Kafue subbasin.
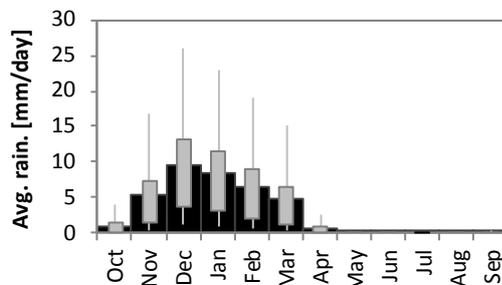
Figure 4.27. Boxplot of monthly GSMaP MVK rainfall from 2003 to 2009 over the Kafue subbasin.
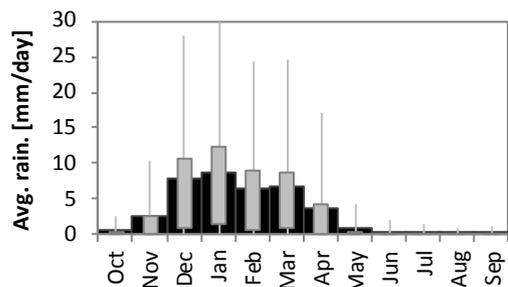
Figure 4.28. Boxplot of monthly TMPA 3B42 v7a rainfall from 2003 to 2009 over the Shire River subbasin.
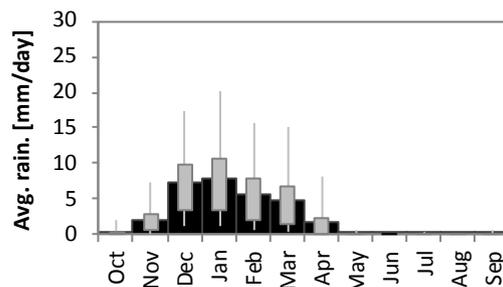
Figure 4.29. Boxplot of monthly NOAA/FEWS RFE2.0 rainfall from 2003 to 2009 over the Shire River subbasin.
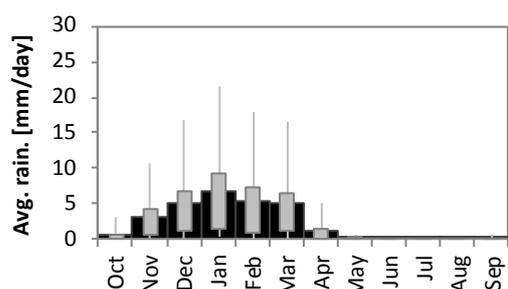
Figure 4.30. Boxplot of monthly TMPA 3B42 v7a rainfall from 2003 to 2009 over the Chobe subbasin.
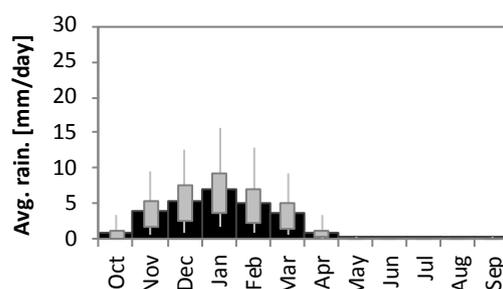
Figure 4.31. Boxplot of monthly NOAA/FEWS RFE2.0 rainfall from 2003 to 2009 over the Chobe subbasin.

Resulting from this assessment, it is believed that at the moment TMPA 3B42 v7a data constitutes the most advantageous alternative in the light of the thesis' goals. This is due to several factors: good performance in comparison to alternatives; longer temporal coverage than either NOAA/FEWS RFE2.0 or GSMaP MVK; the availability of a related near real-time product (TMPA 3B42-RT); and satisfactory results for similar hydrological applications.

4.4   Discharge and water level gauge measurements

Discharge and water level measurements are essential for the calibration of hydrological models. Unlike most of the modeling data described in the present chapter, which can be estimated from remote sensing sources, discharge series come almost exclusively from ground measurements, being the same true for water levels except for some localized cases where altimetry can be employed (see Section 4.5).

Due to a large effort by ADAPT researchers in order to retrieve data from stakeholders, a wealth of information was available for model forcing and/or calibration. The main data discharge and water level data used in the scope of this work is displayed in Appendix I.C. Its main sources have been: ARA-Zambeze; the Department of Water Affairs of Zambia; the Global Runoff Data Centre, 56068 Koblenz, Germany; Hidroeléctrica de Cahora Bassa; the Zambezi River Authority; and ZESCO.

Going much further than Appendix I.C, the ADAPT online database, described in Appendix II.A, contains detailed data and metadata about the series. It can be found at http://zambezi.epfl.ch.

4.5   Altimetry water level measurements

Unlike for other studies focused in the ZRB area (e.g. Michailovsky 2013), altimetry water level measurements were only marginally employed in the present work.

The water level data collected within ADAPT included long water level time series at the most relevant reservoirs (Kariba, Cahora Bassa, Itezhi-Tezhi, and Kafue Gorge) as well as at the Malawi Lake.

It was remarked that the ground data for the Malawi Lake, whose last entry made available to ADAPT is dated of October 2002, could be elongated by resorting to satellite altimetry data. To that end, smoothed Lake Malawi height variations were downloaded from the Global Reservoir and Lake Elevation Database[10] covering a period spanning from late 1992 to present.

The series is comprised by data from three sources: TOPEX/Poseidon historical archive (1993 to 2002), Jason-1 Interim GDR 20Hz altimetry (2002 to 2009), and OSTM Interim GDR 20Hz altimetry (ice mode) (2008 to present). For a discussion on the available continental surface water's altimetry products and associated issues refer to Calmant et al. (2008).

Owing to a very good agreement between ground data and height variations in the overlapping period from 1992 to 2002, the altimetry estimates were appended to the original series after a simple linear transformation whose coefficients minimized the discrepancies observed from 1992 to 2002.

4.6   Land use and soil type maps

In order to correctly characterize the HRUs used in the SWAT model, information on land use and soil type is of paramount importance.

Following the work of Cohen Liechti (2013), itself based on Schuol et al. (2008a, 2008b), soil type was characterized by the Digital Soil Map of the World and Derived Soil Properties version 3.5, produced by the Food and Agriculture Organization (FAO) of the United Nations at a resolution of 10 km (Food and Agriculture Organization of the United Nations (FAO) 1995). In it, about 5000 soil types are differentiated and some soil properties for two layers (0-30 cm and 30-100 cm) are provided. Further soil characteristics used by the SWAT model (e.g. particle-size distribution, bulk density, organic carbon content, available water capacity, saturated hydraulic conductivity, etc.) were provided by Reynolds et al. (2000) or calculated by using pedotransfer functions implemented in the model Rosetta (Schuol et al. 2008a, Schuol et al. 2008b). The soil map of the basin is illustrated in Figure 4.32 along with the basic soil types (the complete soil map breaks into much finer detail, which is not practical to illustrate).
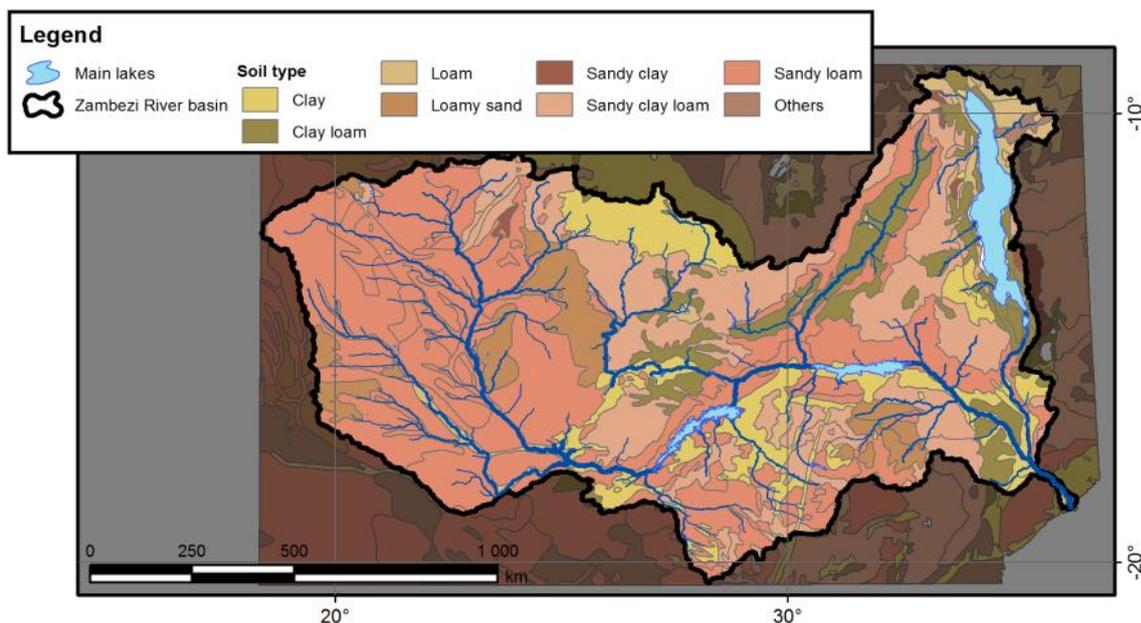
Figure 4.32. Soil type map for the ZRB.

Similarly, land use information was also taken from the work of Schuol et al. (2008a, 2008b) who employed a land use map originated in the Global Land Cover Characterization version 2

---

[10] Data retrieved from http://www.pecad.fas.usda.gov/cropexplorer/global_reservoir/.

(Loveland et al. 2000), which has a 1 km spatial resolution and 24 classes of land use representation. The parameterization of the land use classes (e.g. leaf area index, maximum stomatal conductance, maximum root depth, optimal and minimum temperature for plant growth) was based on the characteristics of the classes defined in the SWAT original database and on literature research (Schuol et al. 2008a, Schuol et al. 2008b). The land use map is illustrated in Figure 4.33.
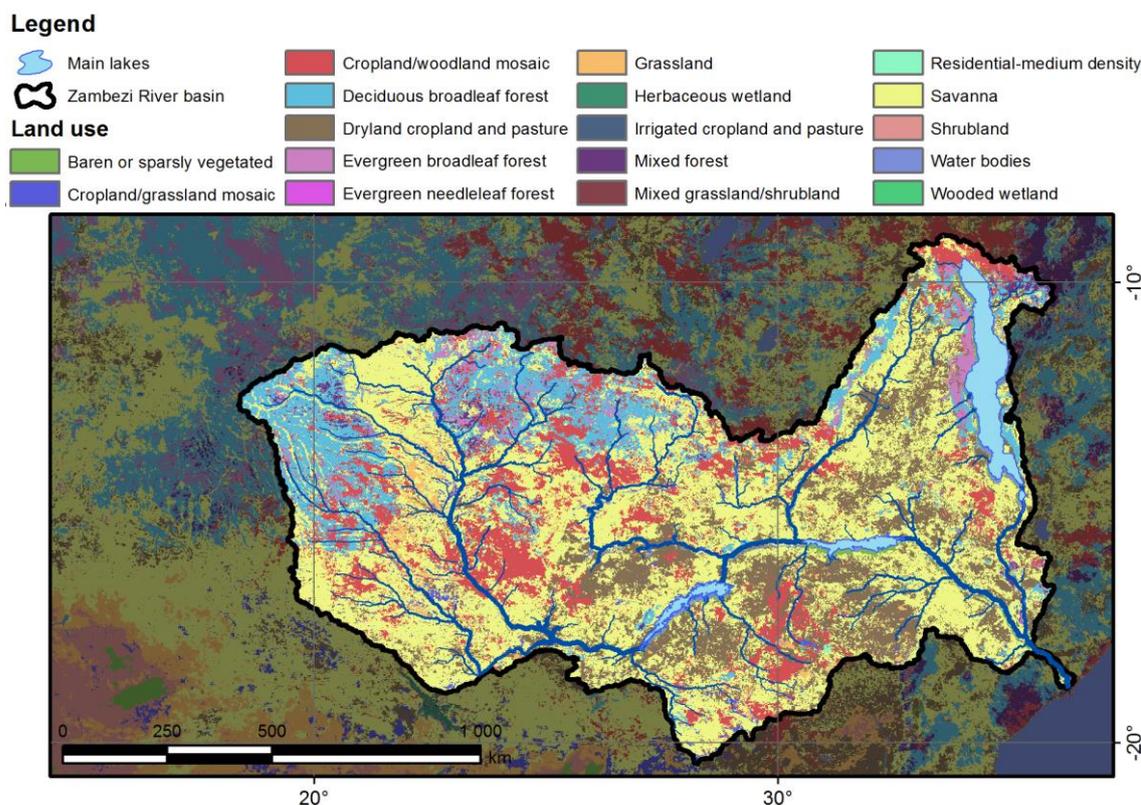
Figure 4.33. Land use map for the ZRB.

Other – more recent – sources of soil map and land use could have been adopted. For example, a new version of the soil map, the Harmonized World Soil Database (Nachtergaele et al. 2008), is now available. While obviously having the potential to condition the model's response and, therefore, being interesting to evaluate, the derivation of either soil type or land use data is complex and well beyond the scope of this work. Additionally, even if new characterizations are taken "as is", all their information must be translated into SWAT parameters prior to HRU calculation, which is a specialized task in its own right.

4.7   Data from climate reanalysis models

In order to estimate Potential Evapotranspiration (PET) as accurately as possible, additional types of data are needed, the actual needs being dependent on the equation used for this calculation.

After testing independent PET estimates and the Penman-Monteith equation in a preliminary SWAT model, Cohen Liechti (2013) observed that the use of the Hargreaves method (Hargreaves and Samani 1985), dependent only on air temperature as an external input, led to the best model responses.

A convenient means of obtaining climate data is to resort to one of many existing climate reanalysis models. Covering very large periods at sub-daily time steps, such models provide a wealth of data. Examples are the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim, whose data was used by Michailovsky (2013) as a forcing for a ZRB SWAT

model, or the NOAA/National Weather Service National Centers for Environmental Prediction's (NCEP) two reanalysis models.

In the present case temperature data was obtained from the NCEP/Department of Energy (DOE) Reanalysis 2 model (Kanamitsu et al. 2002) provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their website at http://www.esrl.noaa.gov/psd/. Data coverage is nearly World-wide and spanning from 1979 to 2013. Values are produced for each six hours. In the case of surface air temperature, the data is organized in a global T62 Gaussian grid. At the latitudes of the ZRB this grid can be converted to regular square pixels of roughly 1.875°×1.875° with small distortions. The average surface air temperatures for the ZRB, from 1979 to present, are illustrated in Figure 4.34.

Due to the relatively coarse grid and the fact that most values are not supported in physical measurement, and despite of the relative homogeneity of the air temperature field, testing a finer alternative would be worthwhile.
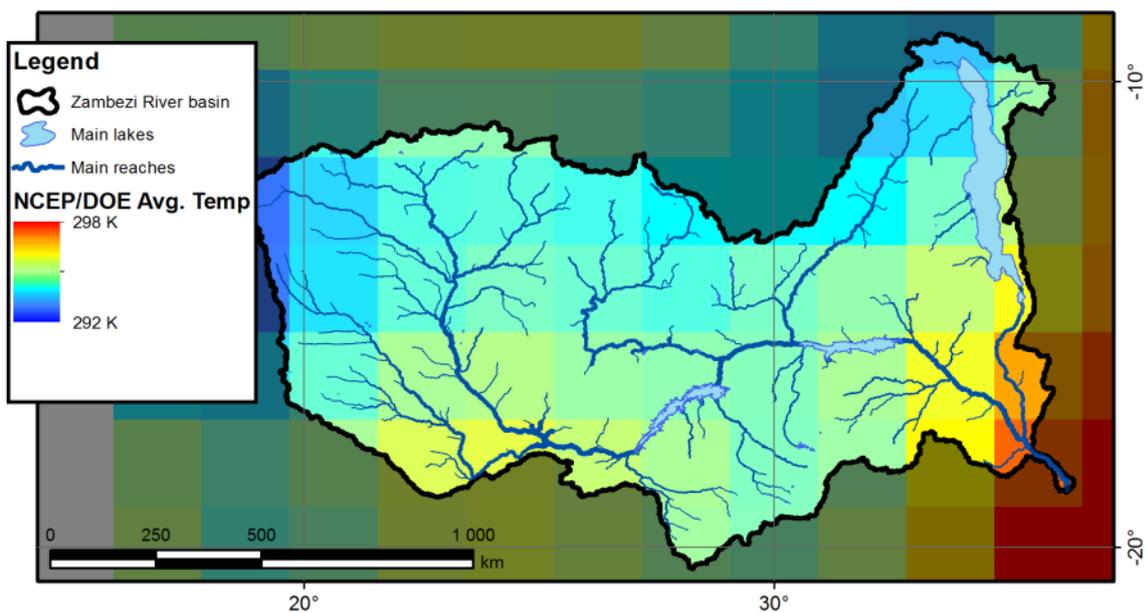
Figure 4.34. Average surface air temperature taken from the NCEP/DOE Reanalysis 2 dataset (1979 to present).

5 EXTENDING MODELING DATA: THE PATTERN-ORIENTED MEMORY INTERPOLATION TECHNIQUE

5.1 Motivation and objectives

Hydrological modeling over the whole ZRB is constrained by the availability of rainfall data. Aiming at daily or sub-daily time steps, this translates into a practical limitation in the modeled periods as:

- no daily, detailed, reliable and homogeneous rainfall products cover periods prior to 1998;
- rain gauge records in the ZRB are too sparse and incomplete in order to produce reliable areal rainfall maps using common interpolation tools;
- even if there was access to adequate gauge records, there would be a marked difference in the support of the rainfall fields before and after 1998 (interpolated and remotely sensed, respectively), making it ill-advised to use such an artificially heterogeneous combined field as forcing for a calibrated hydrological model[1].

Acknowledgeable climate variations with periods of some decades are associated with the ZRB. As illustrations of the magnitude of such variations, observing a long discharge time series at Victoria Falls (Figure 5.1) or the fact that, completely stopping its outflow to the Shire River, the Malawi Lake dropped remarkably low levels at the beginning of the 20th century should suffice (Delvaux 1995).
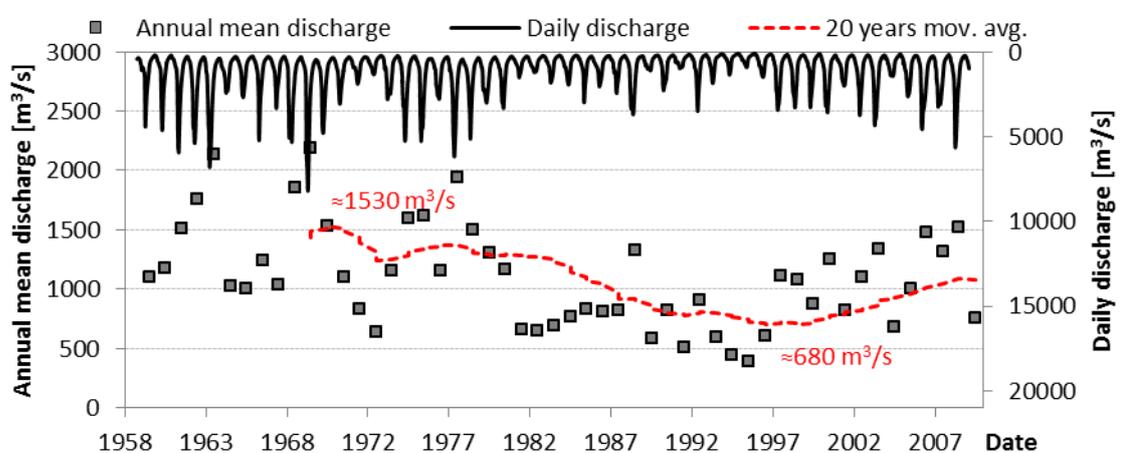
Figure 5.1. Historical variability of discharges at Victoria Falls.

---

[1] This point is thouroughly discussed in Chapter 8.

This constitutes one of two reasons why a long model calibration period in the ZRB is particularly important. The other is that, for some of the calibration points, much more data was collected about the period prior to 1998 than from 1998 to present. Without a long calibration period, sizeable subbasins such as the Luangwa are extremely hard to calibrate by direct means.

With this in mind, a novel methodology to interpolate historical rainfall from sparse records with increased accuracy, while guaranteeing a homogeneous support of the rainfall field was developed.

5.2   Introduction

The Pattern-Oriented Memory (POM) interpolation technique, described below, was published in two papers from which must of the chapter is drawn (Matos et al. 2013a, Matos et al. 2014). Due to the novelty of the method, its performance, the relevance to other aspects of the research, and the share of the work devoted to its development, the author took the risk of allowing this chapter to go into greater depth than the remainder.

The ability to reproduce areal rainfall maps from point observations is a demanding challenge and, quite often, one of vital importance in hydrological studies. In the past decades the high variability of rainfall over space and time and how to mathematically cope with it has been a subject of major concern of hydrologists (Lanza et al. 2001).

The task of producing interpolated rainfall maps is a far-reaching problem whose solution must take into account, firstly, the spatial and temporal scales at which the rainfall estimates are available and, secondly, the location and amount of the existing data. In this multivariate problem, the performance of different interpolation methods varies greatly.

Rainfall areal distributions are a valuable input for hydrological models. The calibration and validation of these models become particularly demanding tasks which usually require long series of data (rainfall, discharge, evapotranspiration, etc.). As seen previously, satellite derived rainfall products can provide distributed rainfall estimates that cover the equivalent of a dense measurement grid, in some cases, much denser than the one comprised by the existing rain gauge networks. In such areas, where the traditional rain gauge networks are sparse, it is convenient to resort to satellite derived rainfall estimates directly as input data or, as an alternative, to a combination of satellite, gauge, and eventually, radar data (Pegram et al. 2004, Sinclair and Pegram 2005).

When satellite derived rainfall estimates or the associated combined products are used as rainfall input data the performance and robustness of the hydrological models are constrained by the length of the satellite derived rainfall datasets (near-worldwide reliable estimates have become available from 1998 onwards, with the Tropical Rainfall Measuring Mission, TRMM). This all too often implies that the modeler will not be able to utilize precious historical discharge data that is useful for calibration and validation purposes. The ability to produce accurate areal distributions of historical or "older" rainfall events (denoting the events that pre-date reliable satellite derived rainfall estimates) would greatly contribute to alleviate this situation.

To date, several interpolation methods have been proposed, applied and compared in the task of rainfall interpolation. Thiessen polygons (Thiessen 1911) and Inverse Distance Weighting (IDW) (Shepard 1968) are techniques of straightforward application. IDW in particular combines a good comparative performance with simplicity, having a single free parameter. More recently, alternatives have been developed. Among these Kriging (Cressie 1990) outstands for its popularity and performance (e.g. Vicente-Serrano et al. 2003, Lloyd 2005), but models such as the geographically weighted regression (Brunsdon et al. 2001, Fotheringham et al. 2002), splines (Hancock and Hutchinson 2006), Radial Basis Function Networks (RBFN) (Broomhead and Lowe 1988), and Improved RBFNs (IRBFN) (Lin and Chen 2004) have also been reported to yield good results. Additionally, other techniques like kernel regression (Takeda et al. 2007), Artificial Neural Networks (ANN) (Sivapragasam et al. 2010) and Bayesian interpolation (Woodbury and Ulrych 2000)

have been proposed to enhance or execute interpolation tasks. Most of these methods, however, perform the interpolation for a given period in time, generally disregarding the full set of available records or, at best, doing so implicitly, and yield suboptimal results when distances between point measurements are large and short time steps are used. Explicitly using the full set of available records to perform the rainfall interpolation task could lead to improved performances, particularly so when considering the large amount of information that is contained in satellite derived rainfall estimates.

In the context of this research, POM refers to the ability to recall (identify and reproduce) complex non-linear patterns from related historical rainfall events; models that can display pattern-oriented memory are, therefore, an interesting approach as they can use the full set of satellite derived rainfall estimates in order to perform each interpolation task. This newly proposed method employs asynchronous information derived from areal rainfall estimates in addition to information from the data points that serve as the basis for the interpolation, applying regression techniques to produce its estimates. Its potential to achieve better performances than established interpolation methods is mainly due to its feature of explicitly making use of more data, the non-linear behavior and the capacity to recall recorded observations.

POM interpolation (Matos et al. 2013a) explicitly takes into account the time dimension of the rainfall in order to interpolate historical areal rainfall maps. It does so in two main phases, depicted in Figure 5.2.
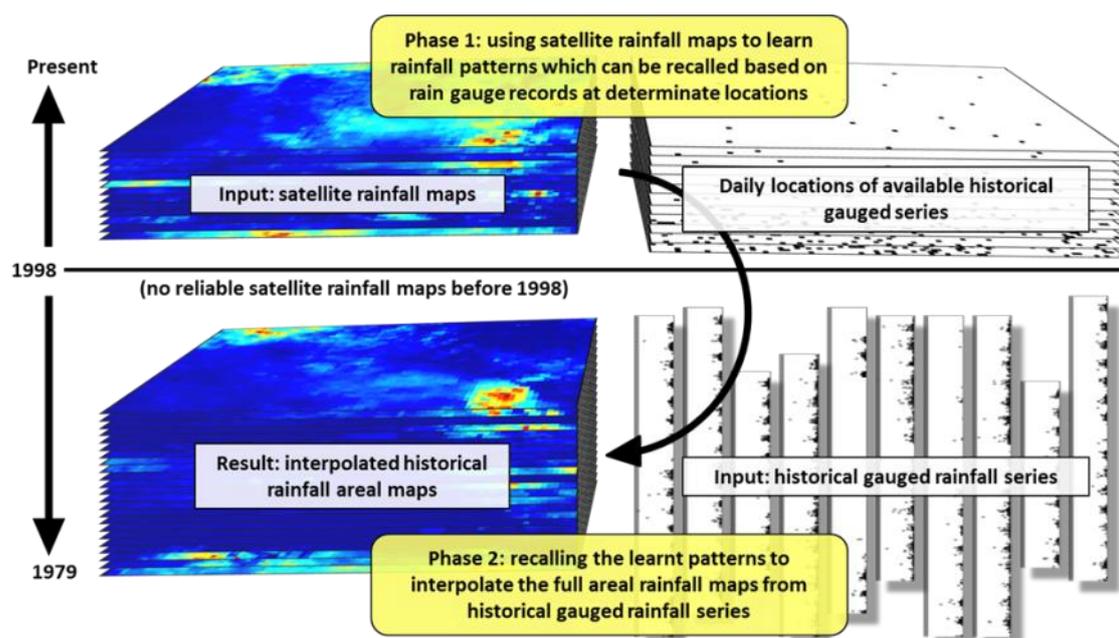
Figure 5.2. Scheme of the POM interpolation procedure applied at a daily time step for the period from 1979 to 1997 (Matos et al. 2014).

Below, the results illustrating the comparative capabilities of POM are presented. This is done by following two distinct approaches. The first approach, outlined in Section 5.4.1, provides a comprehensive analysis of POM in comparison to other interpolation techniques under different conditions of spatial data availability, number of patterns used for training, and time step. Because such a comparison would not be possible using rain gauge data over the ZRB given the size and distribution of the available historical series, proxy satellite data was employed. In order to rule out unforeseen effects stemming from the application of a single product, TMPA 3B42 version 6, NOAA/FEWS RFE2.0, PERSIANN and CMORPH datasets were used (see Section 4.3 for details on these datasets). The second approach, detailed in Section 5.4.2, involves the application of POM to real historical rain gauge series. A comparison to other interpolation techniques was also made,

now comparing interpolated results to a hold-out sample of historical records using a Monte-Carlo cross-validation procedure.

5.3   Rainfall interpolation methods

The POM interpolation method was compared to Thiessen polygons, IDW, Kriging and IRBFNs. Below, the principles underlying these methods are outlined, with an emphasis on POM. Specific details on the application of each method are not given here due to space considerations, but it should be highlighted that an effort was made to employ state-of-the-art models in the comparison. Obviously, neither Thiessen polygons nor IDW fall into this category, but "enhanced" versions of Kriging (Multiple-Realization Ordinary Kriging, MROK) and IRBFNs constitute substantial improvements over standard interpolators.

Thiessen polygons

The Thiessen polygons method applied to interpolation purposes is simply a nearest neighbor algorithm, expressed in equation (5.106), where $\hat{\mathbf{y}}_p$ is the unknown series at point p, $\mathbf{y}_i$ or $\mathbf{y}_k$ the known series at points i or k, and $d_{k,p}$ represents the distance from point k to the point being evaluated.

$$\begin{cases} \hat{\mathbf{y}}_p = \mathbf{y}_i \\ i = \arg\min_k \left\{ d_{k,p} \right\} \end{cases} \tag{5.106}$$

Inverse distance weighting

IDW is a widely known method of spatial interpolation which weights observations by a simple non-linear relationship of their distance to the point of interest. Proposed by Shepard (1968), it can be described by equation (5.107), where $\hat{\mathbf{y}}_p$ is the interpolated value at point p, $\mathbf{y}_i$ represents the observed value at a generic point i, N is the number of points with observations, $d_{i,p}$ has the value of the distance between points p and i, and β is the model's free parameter, often assumed equal to 2.

$$\hat{\mathbf{y}}_p = \frac{\sum_{i=1}^{N} \mathbf{y}_i \cdot d_{i,p}^{-\beta}}{\sum_{i=1}^{N} d_{i,p}^{-\beta}} \tag{5.107}$$

Kriging

Kriging is also a spatial interpolation method where each observation is weighted according to a function of its distance to the point of interest. The function used to weight the observations can, however, be more complex than that of IDW, being dependent on the relationship between the observation's variance and their distance. There are several Kriging models. Simple Kriging, treating the unknown values as a random variable, assumes that its expected value is known and that the underlying field is stationary. In Ordinary Kriging (OK), the expected value is considered unknown. Using co-Kriging it is possible to associate the field's variation to additional variables (Myers 1982). Kriging with external drift (e.g. Tobin et al. 2011), universal Kriging (Matheron 1969) or regression Kriging are examples of models where the assumption of field stationarity is relaxed by considering linear, polynomial or other classes of trend characterizing the mean value of the field. Indicator Kriging, disjunctive Kriging and block Kriging (Grimes et al. 1999, Haberlandt 2007) provide further examples of Kriging models with non-linear behavior.

OK interpolation, used in this contribution, is accomplished through a linear combination of observations and can be described based on equation (5.108):

$$\hat{\mathbf{y}}_p = \hat{Z}\left(c_p\right) = \sum_{i=1}^{N} w_i \cdot Z\left(c_i\right) = \sum_{i=1}^{N} w_i \cdot y_i \qquad (5.108)$$

where $\hat{\mathbf{y}}_p$ is the estimated value at point p; $\mathbf{y}_i$ is a known value at point i; c_p and c_i are the coordinates of points p and i; w_i are the Kriging weights; and $Z(\cdot)$ and $\hat{Z}(\cdot)$ represent the random field and its Kriging estimate at given points.

It stems from equation (5.108) that the Kriging weights characterize the interpolation results. According to the assumptions of each Kriging model, and given that the variogram quantifies the degree of spatial dependence of a random field by assuming that its variation between any pair of points is a function of their distance, the weights are computed in order to minimize the variance of the prediction error for a chosen theoretical variogram. Because the variogram of the random field is often unknown, the theoretical variogram used to estimate the optimal Kriging weights must be adjusted to an experimental variogram obtained from the observed data.

When daily rainfall fields are used to calculate experimental variograms results differ greatly from day to day. By adopting a Single-Realization Ordinary Kriging (SROK) approach, where independently fitted theoretical variograms are used at each time step, a degradation of the interpolation performance was observed. Alternatively, "multi-realization" experimental variograms, consisting of an average of normalized variograms, were used in a Multiple-Realization Ordinary Kriging (MROK) set-up. Details on the application of this methodology can be found in Bacchi and Kottegoda (1995) and Matos et al. (2013a)[2].

Radial basis function networks
RBFNs, described in Section 3.4.3, are a type of ANN particularly suitable for interpolation tasks.

RBFNs can be regarded as general regression or interpolation models. In rainfall interpolation tasks, it is usual to consider the latter. In that case, the model is fitted and applied to a single rainfall realization, being the input variables point coordinates. The fitting consists in determining the unique set of weights, $\mathbf{W}$, and intercepts, $\mathbf{b}$, that reduce the estimate error in the measured input points to 0.

Here, a modified version of the base model – improved RBFNs (IRBFN) (Lin and Chen 2004) –was applied using information derived from the fitting of an exponential model to a multiple-realization experimental variogram.

Pattern-Oriented Memory
POM interpolation is a performing method based on the idea that rainfall patterns exist and can be identified over a certain area. In its essence, the proposed interpolation method performs a non-linear regression from points with observed data to the points of interest in order to produce estimates. In general terms, the regression for point p can be expressed by equation (5.109), where $\mathbf{y}^{obs}$ represents a vector of point observations in the same time frame and in the vicinity of p, vector $\mathbf{w}_p$ characterizes the regression at that point, and $\hat{\mathbf{y}}_p$ has the meaning previously introduced:

$$\hat{y}_p = f\left(\mathbf{y}^{obs}, \mathbf{w}_p\right) \qquad (5.109)$$

---

[2] Alternative formulations of Kriging were tested with worse performances, namely universal Kriging with polynomials up to the 3rd degree. The anisotropy of the rainfall field was also ruled out.

From equation (5.109) it is evident that POM interpolation depends on the definition of $\mathbf{w}_p$, which is unknown *a priori*. In order to define it, a set of input-output patterns ($\mathbf{y}^{obs}$ and $\hat{\mathbf{y}}_p$) must be used. Because, in recent years, satellite rainfall estimates have become available for most of the globe, the problem of collecting the input-output patterns necessary to characterize the regression can be easily solved under the assumption that these satellite estimates behave similarly to the true rainfall in relative terms. As such, although POM uses historical rain gauge data as a basis for interpolation ($\mathbf{y}^{obs}$), the possibility that an adequate $\mathbf{w}_p$ can be estimated based on recent satellite data is tested. Taking into account the formulation of POM, the assumed similarity between the satellite estimates and true rainfall does not need to be expressed in absolute rainfall terms, but mainly be translated in a qualitative agreement between the areas displaying high and low intensity rainfall at each time step and their relative magnitudes.

The estimation of $\mathbf{w}_p$, $\hat{\mathbf{w}}_p$, can be accomplished through equation (5.110) where $g(\cdot)$ represents a generic function whose inputs are: $\mathbf{y}_p^{sat}$, which is a series of satellite estimates at point p ; $\mathbf{Y}^{sat}$, a matrix containing series of satellite estimates over the gauged points in the vicinity of p ; and $\mathbf{h}_p$, a set of hyperparameters that control how $\hat{\mathbf{w}}_p$ is computed:

$$\hat{\mathbf{w}}_p = g\left(\mathbf{y}_p^{sat}, \mathbf{Y}^{sat}, \mathbf{h}_p\right) \tag{5.110}$$

For any given point and date, the POM interpolation method functions, therefore, in two independent (yet similar) fields. Firstly, the $\hat{\mathbf{w}}_p$ estimate is calculated solely using recent satellite estimates. Secondly, $\hat{\mathbf{w}}_p$ is employed to characterize a regression (5.111) which is undertaken based on the gauged local rainfall field described by $\mathbf{y}^{obs}$.

$$\hat{y}_p = f\left(\mathbf{y}^{obs}, \hat{\mathbf{w}}_p\right) \tag{5.111}$$

While any regression model could be used to perform the described task (from multiple linear regression to ANNs), not every regression model guarantees the identification of patterns which is the main strength of POM. Regularized radial basis methods, such as SVRs or LS-SVRs (see Section 3.4 for details on the models) are most fit to perform the regression tasks associated with POM. In fact, it is the application of such models with a Gaussian kernel that grants the "memory" property to the interpolation. With it, each new pattern is projected onto the feature space according to its similarity with particular entries of the training set, "stored" in $\hat{\mathbf{w}}_p$. When the regression is undertaken for a given point and time step, the trained LS-SVR computes a measure of the proximity between the gauged rainfall, $\mathbf{y}^{obs}$, and the training patterns present in $\hat{\mathbf{w}}_p$, a process of effectively recalling past rainfall events. Then, these events are combined according to their proximity in order to produce the desired interpolation.

Unlike IDW or Kriging, which perform interpolations directly on the gauged rainfall field, POM is subject to the additional requirement that satellite and gauged estimates share similar properties. When the match between both types of estimates is perfect, no potential loss in the performance is observed. However, satellite estimates have been shown to disagree to some extent with rain gauge data (e.g. Cohen Liechti et al. 2011). Measures to mitigate this effect were not studied in this

work, but it is shown that the advantages of POM clearly outweigh the losses of performance that may result from it.

5.4   Methodology

5.4.1   Broad comparison with other interpolation models using satellite data

The first part of the work attempted to establish in what circumstances the POM interpolation method would outperform competitors under a broad range of conditions (spatial data availability, time step and number of training patterns).

For clarity, the problem of matching satellite rainfall and rain gauge data was not addressed at this stage. Instead, several historical satellite rainfall estimates were split into training/validation and test subsets, the latter being used as proxies for historical rain gauge data. The rainfall patterns were "learnt" for a restricted set of input pixels, being that afterwards, for the test period, the full reconstructed rainfall maps were compared with true satellite rainfall estimates. The process is schematized in Figure 5.3.
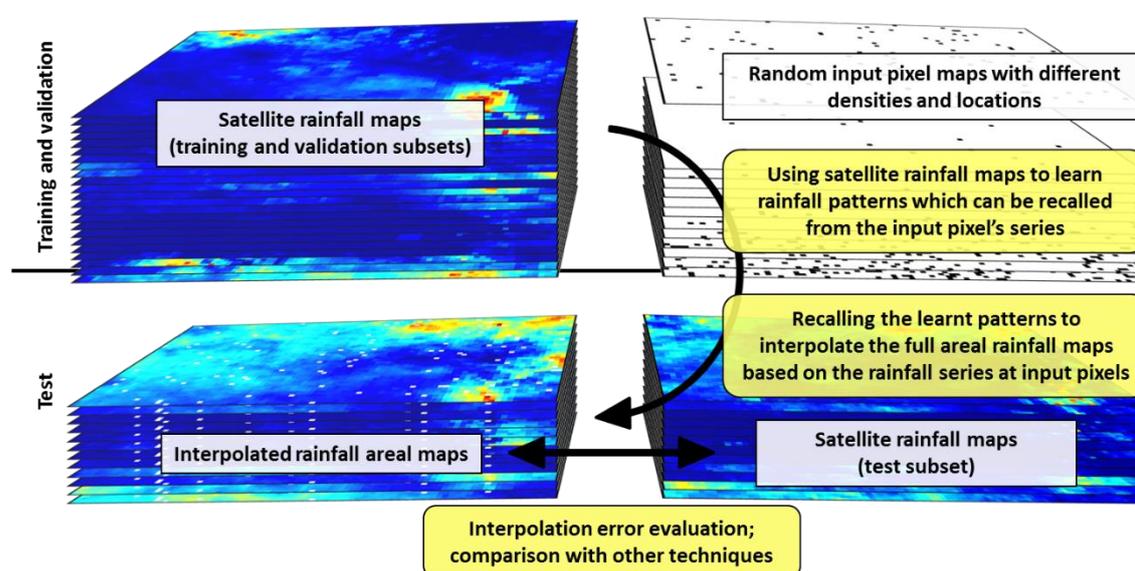
Figure 5.3. Scheme of the comparison to other interpolation methods using satellite data as a proxy for true rainfall.

The employment of satellite rainfall estimates as a proxy for gauged data has several advantages. Firstly, use of satellite data at this stage helps to separate issues associated with the learning and recalling of rainfall patterns and the matching of rain gauge and satellite data. Secondly, the use of areal satellite estimates can substitute the otherwise recommendable cross-validation procedure using rain gauge data – particularly hazardous due to sparse availability of data. Finally, evaluating the methodology over several distinct satellite rainfall estimates is a superior alternative to the employment of synthetically generated rainfall maps, which could also be resorted to.

According to Cohen Liechti et al (2011) and as illustrated before (Section 4.3) satellite derived rainfall estimates disagree substantially over the study area. Additionally, the satellite estimates result from data filtering techniques, being that the possibility that a good adjustment of a given interpolation model tested on this kind of data can be more associated with the applied filter than with the underlying rainfall phenomenon cannot be discarded. Resulting from the prior considerations, carrying out the analysis for several satellite estimates was considered essential.

A cross-validation scheme was employed to train, validate and test the different models. Varying periods within the ranges of January 1998 to February 2005 and April 2007 to December 2009 were used for training and validation; the training and validation subset lengths varied according to the

specific dataset, from 4 years in the case of CMORPH and PERSIANN to 6 years in the case of NOAA/FEWS RFE2.0 and 10 years in the case of TMPA 3B42 v6. The test period was always from March 2005 to February 2007.

The analysis was undertaken for six different spatial data availabilities, roughly equivalent to having available historical rainfall series data from 0.5, 1, 2.5, 5, 10 and 20% of all the pixels in the area. Although the percentage of data is not a measure of data availability recommended for direct application to analogous procedures, these values can alternatively be interpreted in terms of the expected distance to the nearest measured pixel from a non-measured location. Using this metric, it can be stated that the considered data availabilities correspond to average expected distances of 1.82, 1.45, 0.87, 0.63, 0.45 and 0.34° from any pixel center to the nearest measured source of data (roughly 218, 174, 104, 76, 54 and 40 km at the considered latitudes). This metric will be referred to as Average Expected Distance to the Nearest Measured Pixel (AEDNMP), being used to characterize the data sparseness.

In order to discard the possible influence of a particular configuration of data source locations, three different randomly selected subsets (in space and time) were considered for each combination of dataset, AEDNMP and temporal resolution. For each of the considered spatial data availabilities (one for each row of Figure 5.4) three alternative random subsets of "input" pixels were generated (seen column-wise in Figure 5.4). On the temporal plane, random samplings of the training and validation subsets were made for each input set.

Aggregating the temporal and spatial facets of the analysis, the same combinations of input pixels and training and validation subsets were kept for all the rainfall products and time scales, so that the models could be compared under identical conditions.

5.4.2 Interpolating from historical gauged rainfall series

With the comparative performance of POM well established, the methodology was applied to historical gauged rainfall series, from 1979 to 1997. In order to accomplish this task, the rainfall records extracted from the Global Surface Summary of the Day (GSOD) archive and produced by the NOAA's National Climatic Data Center (NCDC) were interpolated using TMPA 3B42 v7a satellite rainfall estimates gathered from 1998 to 2011. The locations of the used rain gauges are indicated in Figure 5.5.

In a sense, the analysis performed with gauged data is necessarily more limited than one resorting to satellite rainfall estimations as gauge locations are fixed and the constraining lack of access to data of the ZRB does not allow a full evaluation of the effect of different degrees of its spatial availability. Conversely, because the change of support (satellite rainfall and gauged rainfall do not necessarily behave exactly alike) which is inherent to POM but not to the other methods, an evaluation with rain gauge data is important. Again, a comparison to alternative methods was undertaken, although restrained to the best performing ones (IDW and MROK).

Applying POM to historical gauged series also entails additional challenges, namely:

- POM interpolation is a computationally expensive method, whose demands grow non-linearly with the number of training patterns. The improvements in performance that should be expected by using a longer training period have to be weighed against a dramatic increase in computation time.
- As a consequence of the application of the LS-SVR model (the best performing one), the average of the estimates tends to equal the average of the output training values, which can be undesirable, leading to overestimations during the dry seasons and underestimations during the wet periods.

Contributing to the solution of both issues, when, instead of a single regression model, valid yearlong, seasonal models are applied, the number of training patterns per season will be reduced and the driver for under and overestimations weakened.

In spite of their advantages, the loss of information accompanying the usage seasonal regressions must be accounted for. For example, by not having a sufficiently long series of high intensity rainfalls used for training in the dry season, substantial estimation errors can result if high intensity rainfall events do occur in that period. To prevent this, a "summary" of the rainfall events that can take place throughout the year at the region of interest should be added to each season's training sets. Here, such a summary was made by clustering the satellite rainfall data using Self-Organizing Maps (SOM).
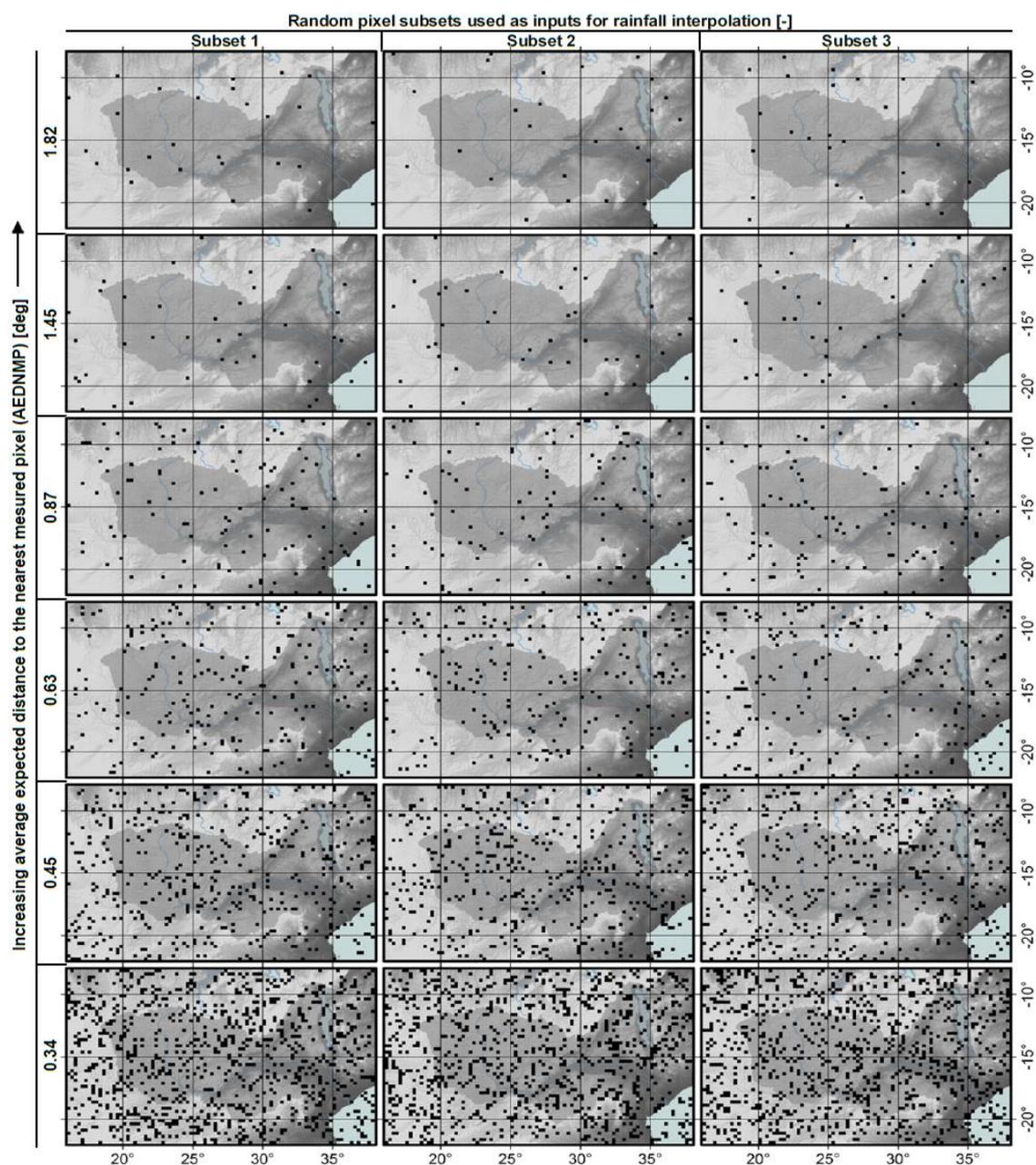
Figure 5.4. Subsets of random input rainfall maps for posterior interpolation. From top to bottom, the expected average distance to the nearest measured pixel is: 1.82, 1.45, 0.87, 0.64, 0.45 and 0.34°. From left to right: three random input pixel subsets used for model training.
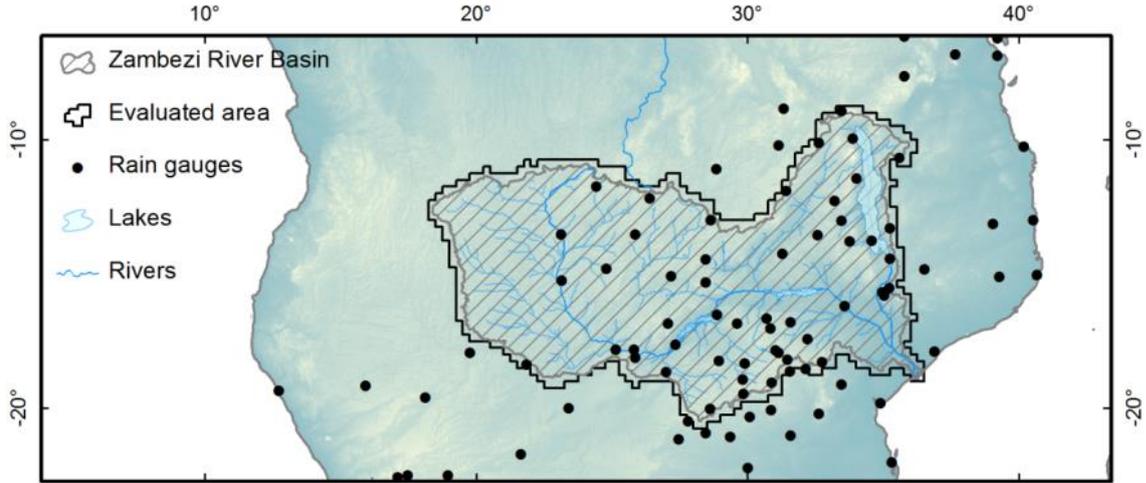
Figure 5.5. Location of the used rain gauges and delineation of the evaluated area.

The SOM (Kohonen 1982) is an unsupervised ANN with the ability to map high-dimensional spaces into a regular, usually two-dimensional grid (Kohonen and Honkela 2007). Also, it can be used as a clustering algorithm (e.g. Mangiameli et al. 1996).

SOMs are usually comprised of two layers. One input layer with one node for each input dimension and one output layer that assumes the form of an interconnected grid. Each of the output nodes is defined in a space with the same dimensions of the input patterns. During its unsupervised learning phase, the SOM is iteratively presented with the training patterns. For every pattern that is presented, the output node which is closer to it according to a criterion of Euclidean distance is activated (in a winner-take-all strategy). Following this, the winning node as well as those in its vicinity move closer to the coordinates of the input, causing the network to adapt. Once the positions of the nodes in the output layer stabilize (usually after all training patterns have been presented several times), the training stops. The clustering interpretation of the SOM can be that each output node's coordinates constitute the center of a cluster and the cluster members are inputs that activate it (see Section 3.4.1 for a brief example of this ANN's capabilities).

The comparison to IDW and Kriging was made comparing interpolated rainfall to a hold-out sample of nearly 180 000 daily gauged records from 1979 to 1997 corresponding to randomly distributed locations within the basin

5.4.3 Error measures
Five distinct error measures were evaluated: the Normalized Root Mean Square Error, NRMSE (5.112), the overall Relative Bias, RB (5.113), the Variance Ratio, VAR (5.114), the Mean Absolute Error, MAE (5.115) and the Pearson correlation coefficient, r.

$$NRMSE_{\mathbf{y},\hat{\mathbf{y}}} = \sqrt{\frac{1}{N \cdot Var(\mathbf{y})} \sum_{i=1}^{N} \left(y_i - \hat{y}_i \right)^2} \qquad (5.112)$$

$$RB_{\mathbf{y},\hat{\mathbf{y}}} = \frac{\sum_{i=1}^{N} \left(y_i - \hat{y}_i \right)}{\sum_{i=1}^{N} y_i} \qquad (5.113)$$

$$VAR_{\mathbf{y},\hat{\mathbf{y}}} = \frac{\sum_{i=1}^{N} \left(\hat{y}_i - \overline{\hat{y}} \right)^2}{\sum_{i=1}^{N} \left(y_i - \overline{y} \right)^2} \qquad (5.114)$$

$$MAE_{\mathbf{y},\hat{\mathbf{y}}} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right|$$
(5.115)

The NRMSE is an error measure similar to the widely used Root Mean Square Error (RMSE), normalized with the variance of the observations. While a NRMSE of 0 ascertains a perfect fitting, when it equals 1 the estimation's performance is equivalent to that of the mean value of the true series. The RB measure, desirably equal to 0, is employed to account for the bias of the estimates, usually very relevant in hydrology. The variability of the rainfall field, which tends to decrease with the use of interpolation procedures, is taken into account through the VAR measure which has a target value of 1. In spite of their analytical value, it must be emphasized that RB and VAR should be interpreted in the light of alternative measures due to the fact that, on occasion, inadequate estimates can result in relatively good values for both. The MAE measure was adopted due to its robustness to outliers; a perfect fitting corresponding to a MAE of 0.

5.5 Broad comparison with other interpolation models using satellite data

5.5.1 Number of training patterns

POM interpolation models were optimized after a training procedure using several years of records (at the daily, weekly and monthly time steps). However, given that these models can learn, it is reasonable to argue that, as a general trend, the lengthier the time series used for training, the better the model's fitness becomes.

This hypothesis was confirmed for all cases. As an example, from Figure 5.6 to Figure 5.8, three error measures (NRMSE, RB and VAR) for a particular daily time step LS-SVR (TMPA 3B42 v6, 0.87° AEDNMP, and random input pixel subset 2 – see Figure 5.4) are plotted against the number of days used for training. Although in the figures a logarithmic trend provides a good fitting to the experimental points, it must be stressed that while for the studied ranges the fitting is satisfactory, it cannot hold for an indefinitely long training period and, therefore, extrapolations are not valid.
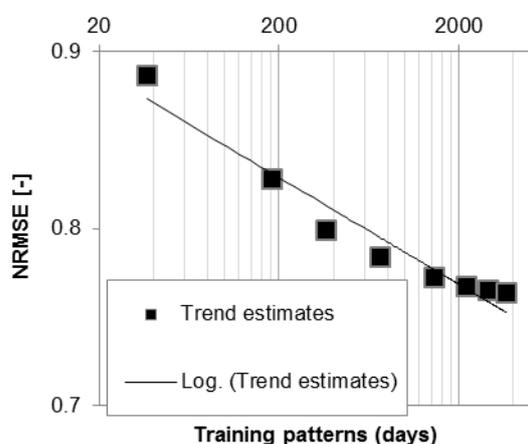
Figure 5.6. NRMSE evolution with training time series length. Example for TMPA 3B42 v6 daily data with AEDNMP of 0.87°. Random input pixel subset 2.
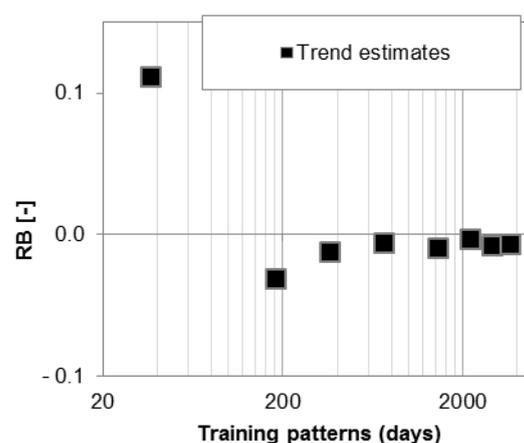
Figure 5.7. RB evolution with training time series length. Example for TMPA 3B42 v6 daily data with AEDNMP of 0.87°. Random input pixel subset 2.

5.5.2 Spatial data availability and time step

The effects of spatial data availability and interpolation time step on the relative performance of the different methods were evaluated according to the set-up of AEDNMP's and random input pixel subsets presented in Figure 5.4.

Summary results for the daily time step are shown in Table 5.1. NRMSE, RB and VAR values for each satellite rainfall dataset, interpolation method and ADNMP are presented. The presented values are averages from the three random input pixel sets tested for each of these combinations. Values for the three best performing interpolation methods (LS-SVR POM, MROK and IDW) are highlighted in bold font. The best results for each dataset/AEDNMP combination are additionally shown in red color.
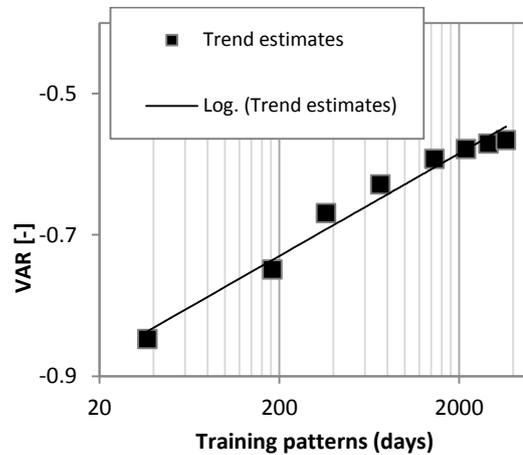
Figure 5.8. VAR evolution with training time series length. Example for TMPA 3B42 v6 daily data with AEDNMP of 0.87°. Random input pixel subset 2.

As can be clearly seen, LS-SVR POM performs best in every case for the NRMSE measure. The RB, albeit assuming generally low values for every model, is also well captured by LS-SVR POM for low spatial data availabilities (high AEDNMP), but generally surpassed by Thiessen polygons and IDW for higher spatial data availabilities. This is accomplished by the latter models, however, at the cost of higher NRMSE. When compared to the second best performing model, MROK, LS-SVR POM generally shows better RB. Similarly, the best VAR performances are attained by Thiessen polygons at the cost of worse NRMSE. Among the best performing models the differences for this measure are generally small.

Results for the monthly time step, presented in Table 5.2, differ from the daily ones being that POM is only comparatively better for the lower spatial data availabilities and when more data is used for training (for the TMPA 3B42 v6 dataset 10 years were used, accompanied by 6 years for NOAA/FEWS RFE2.0 and 4 years for PERSIANN and CMORPH). These results are particularly evident by looking at the NRMSE measure, as regarding RB POM remains quite competitive.

It is believed that POM is particularly advantageous when the input information about the rainfall field is lowest. This is not surprising, as POM is complementing "observed" input data with a long series of satellite records but does not share the theoretical formalism of models such as Kriging. For the almost erratic daily rainfall field, where the spatial correlation quickly diminishes with distance, POM led to the best results even for the highest spatial data availabilities tested. On much more stable monthly rainfall fields, Kriging only seems to be surpassed when spatial data availability is very low. In the intermediate weekly data, not shown here, these observations are corroborated.

5.5.3   Hyperparameter trends

As previously clarified, POM makes use of machine learning non-linear regression models such as SVR or LS-SVR. The behavior of such models is itself dependent on the definition of hyperparameters. In the case of LS-SVR, C and σ (refer to Sections 3.4.4 and 3.4.5, where, $\gamma = 1/\sigma$ for a brief discussion of their effects).

The optimization of these parameters is a major issue regarding POM's practical application as it is a very time-consuming process. Values for C and σ were optimized for each combination of dataset, AEDNMP and random input pixel subset using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) model (Auger and Hansen 2005, Hansen 2006, García et al. 2009, Hansen 2010), a random search evolutionary algorithm adequate to real-parameter optimization of non-linear, non-convex functions, in which the candidate solutions are sampled according to a multivariate normal distribution. Results for C and σ are presented in Figure 5.9 and Figure 5.10, respectively.

LS-SVR is a general regression model and the CMA-ES algorithm optimized its hyperparameters with no consideration for the nature of the rainfall interpolation task being performed, namely in terms of dependence with distance or variograms, for example. In this sense, the existence of a strong relationship between the optimized hyperparameter's values and AEDNMP is a remarkable finding and provides evidence that spatial data availability, more than input pixel distribution or the period selected for training and validation, is a key factor to be taken into account.

Table 5.1. Error measures for the daily time step. For each satellite dataset, interpolation method and AEDNMP the averaged results for the three corresponding random input data subsets of Figure 5.4 are presented.

		Average expected distance to the nearest measured pixel (AEDNMP) [deg]																	
		NRMSE [-]						RB [-]						VAR [-]					
		1.82	1.45	0.87	0.63	0.45	0.34	1.82	1.45	0.87	0.63	0.45	0.34	1.82	1.45	0.87	0.63	0.45	0.34
CMORPH	**Thiessen**	1.07	1.02	0.89	0.78	0.67	0.55	0.03	0.03	0.00	0.00	0.00	0.00	1.06	1.03	0.97	0.94	0.93	0.90
	IDW	0.86	0.83	0.74	0.68	0.62	0.56	0.04	0.02	0.00	0.01	0.00	0.00	0.51	0.51	0.51	0.50	0.52	0.55
	SROK	0.88	0.86	0.84	0.80	0.71	0.64	0.07	0.06	0.18	0.18	0.06	0.02	0.32	0.32	0.37	0.41	0.44	0.48
	MROK	0.85	0.82	0.75	0.69	0.62	0.55	0.04	0.04	0.02	0.02	0.01	0.01	0.38	0.39	0.42	0.46	0.51	0.57
	IRBFN	0.96	1.02	0.80	0.70	0.63	0.52	0.07	0.10	0.04	0.03	0.02	0.02	0.73	0.94	0.72	0.71	0.78	0.80
	LS-SVR POM	0.83	0.81	0.74	0.68	0.60	0.50	0.01	0.01	0.01	0.01	0.02	0.02	0.32	0.35	0.47	0.55	0.65	0.76
NOAA FEWS RFE2.0	**Thiessen**	0.85	0.87	0.68	0.60	0.51	0.42	0.03	0.02	0.00	0.00	0.00	0.00	1.01	1.14	0.97	0.97	0.96	0.94
	IDW	0.72	0.70	0.59	0.54	0.48	0.43	0.03	0.01	0.00	0.01	0.00	0.00	0.60	0.64	0.63	0.64	0.68	0.71
	SROK	0.77	0.74	0.63	0.57	0.51	0.46	0.05	0.04	0.02	0.02	0.02	0.01	0.46	0.50	0.54	0.60	0.66	0.70
	MROK	0.72	0.69	0.59	0.53	0.46	0.41	0.02	0.02	0.01	0.01	0.01	0.01	0.52	0.56	0.59	0.65	0.71	0.75
	IRBFN	0.79	0.78	0.60	0.53	0.61	0.39	0.02	0.04	0.01	0.02	0.02	0.01	0.74	0.84	0.77	0.82	1.06	0.88
	LS-SVR POM	0.70	0.66	0.57	0.51	0.46	0.39	0.00	0.01	0.01	0.01	0.01	0.01	0.50	0.54	0.67	0.72	0.80	0.86
PERSIANN	**Thiessen**	0.78	0.74	0.58	0.47	0.37	0.28	0.01	0.02	0.00	0.00	0.00	0.00	0.98	1.03	0.98	0.97	0.97	0.97
	IDW	0.67	0.63	0.52	0.46	0.40	0.34	0.02	0.02	0.00	0.00	0.00	0.00	0.60	0.66	0.68	0.68	0.72	0.77
	SROK	0.73	0.70	0.57	0.50	0.41	0.34	0.04	0.04	0.05	0.08	0.02	0.01	0.46	0.52	0.61	0.68	0.72	0.78
	MROK	0.66	0.62	0.51	0.43	0.35	0.28	0.01	0.03	0.01	0.01	0.01	0.01	0.55	0.61	0.67	0.72	0.78	0.84
	IRBFN	0.72	0.70	0.51	0.41	0.32	0.25	-0.01	0.05	0.01	0.01	0.01	0.01	0.71	0.83	0.81	0.84	0.88	0.92
	LS-SVR POM	0.65	0.61	0.49	0.41	0.32	0.23	-0.01	0.00	0.01	0.01	0.01	0.01	0.56	0.62	0.76	0.83	0.89	0.95
TMPA 3B42	**Thiessen**	1.08	1.08	0.93	0.84	0.74	0.64	0.02	0.02	0.00	0.00	-0.01	0.00	1.01	1.07	0.96	0.93	0.91	0.90
	IDW	0.89	0.87	0.77	0.72	0.67	0.61	0.01	-0.01	-0.01	0.01	-0.01	0.00	0.48	0.49	0.48	0.48	0.48	0.52
	SROK	0.90	0.89	0.81	0.77	0.72	0.67	0.05	0.03	0.04	0.04	0.03	0.02	0.29	0.31	0.33	0.37	0.41	0.45
	MROK	0.87	0.85	0.77	0.72	0.68	0.61	0.03	0.01	0.02	0.02	0.02	0.02	0.32	0.34	0.37	0.41	0.46	0.52
	IRBFN	1.00	1.05	0.84	0.77	0.72	0.67	0.09	0.15	0.05	0.05	0.04	0.04	0.73	0.95	0.72	0.73	0.78	0.87
	LS-SVR POM	0.85	0.83	0.76	0.71	0.65	0.58	0.00	-0.01	-0.01	0.01	0.01	0.02	0.29	0.32	0.44	0.52	0.59	0.69

Table 5.2. Error measures for the monthly time step. For each satellite dataset, interpolation method and AEDNMP the averaged results for the three corresponding random input data subsets of Figure 5.4 are presented.

		Average expected distance to the nearest measured pixel (AEDNMP) [deg]																	
		NRMSE [-]						RB [-]						VAR [-]					
		1.82	1.45	0.87	0.63	0.45	0.34	1.82	1.45	0.87	0.63	0.45	0.34	1.82	1.45	0.87	0.63	0.45	0.34
CMORPH	Thiessen	0.50	0.48	0.39	0.34	0.28	0.23	0.03	0.03	0.00	0.00	0.00	0.00	1.06	1.07	1.01	0.99	0.99	0.98
	IDW	0.43	0.42	0.36	0.33	0.30	0.29	0.04	0.01	-0.01	0.01	0.00	0.00	0.89	0.85	0.80	0.82	0.82	0.81
	SROK	0.43	0.40	0.33	0.30	0.26	0.23	0.03	0.03	0.02	0.01	0.01	0.01	0.83	0.87	0.87	0.89	0.90	0.92
	MROK	0.40	0.38	0.32	0.29	0.25	0.22	0.03	0.03	0.02	0.01	0.01	0.01	0.89	0.92	0.90	0.90	0.91	0.93
	IRBFN	0.45	0.45	0.36	0.30	0.28	0.22	0.02	0.02	0.01	0.00	0.00	0.00	0.91	0.95	0.95	0.94	0.97	0.96
	LS-SVR POM	0.41	0.39	0.35	0.33	0.30	0.26	0.00	0.00	0.01	0.00	0.00	0.00	0.86	0.86	0.89	0.90	0.92	0.94
NOAA FEWS RFE2.0	Thiessen	0.39	0.40	0.29	0.25	0.21	0.17	0.03	0.02	0.00	0.00	0.00	0.00	1.02	1.05	1.00	0.99	1.00	0.99
	IDW	0.35	0.35	0.29	0.26	0.24	0.23	0.03	0.01	-0.01	0.01	0.01	0.00	0.90	0.89	0.86	0.86	0.88	0.86
	SROK	0.35	0.32	0.24	0.21	0.18	0.16	0.02	0.02	0.01	0.01	0.01	0.01	0.86	0.91	0.90	0.93	0.95	0.96
	MROK	0.32	0.31	0.24	0.21	0.18	0.15	0.02	0.02	0.00	0.01	0.01	0.00	0.90	0.92	0.92	0.94	0.95	0.96
	IRBFN	0.35	0.36	0.25	0.22	0.26	0.15	-0.01	0.01	-0.01	0.01	0.01	0.01	0.86	0.94	0.93	0.96	1.01	0.98
	LS-SVR POM	0.33	0.31	0.26	0.24	0.21	0.18	0.00	0.00	0.00	0.01	0.01	0.00	0.88	0.89	0.93	0.96	0.97	0.97
PERSIANN	Thiessen	0.35	0.33	0.24	0.19	0.14	0.11	0.01	0.02	0.00	0.00	0.00	0.00	0.99	1.05	1.01	0.99	0.99	0.99
	IDW	0.33	0.32	0.26	0.23	0.21	0.19	0.02	0.02	0.00	0.01	0.01	0.00	0.88	0.90	0.85	0.86	0.87	0.87
	SROK	0.31	0.28	0.19	0.16	0.12	0.09	0.01	0.03	0.01	0.00	0.01	0.01	0.85	0.94	0.93	0.94	0.96	0.98
	MROK	0.27	0.26	0.18	0.15	0.12	0.09	0.01	0.03	0.01	0.00	0.00	0.00	0.90	0.97	0.96	0.96	0.97	0.98
	IRBFN	0.31	0.31	0.20	0.16	0.12	0.09	-0.04	0.01	0.00	0.00	0.01	0.00	0.82	0.95	0.95	0.96	0.98	0.99
	LS-SVR POM	0.31	0.28	0.22	0.19	0.16	0.14	-0.02	0.00	0.00	0.00	0.00	0.00	0.89	0.92	0.94	0.96	0.98	0.99
TMPA 3B42	Thiessen	0.53	0.53	0.42	0.37	0.32	0.27	0.02	0.02	0.00	0.00	-0.01	0.00	1.05	1.08	1.01	0.98	0.97	0.98
	IDW	0.45	0.45	0.36	0.33	0.30	0.27	0.01	-0.01	-0.01	0.01	-0.01	0.00	0.82	0.86	0.84	0.85	0.85	0.88
	SROK	0.45	0.44	0.35	0.31	0.28	0.25	0.04	0.02	0.01	0.01	0.01	0.01	0.79	0.85	0.85	0.87	0.88	0.91
	MROK	0.43	0.42	0.34	0.31	0.27	0.24	0.03	0.01	0.01	0.01	0.01	0.01	0.83	0.86	0.86	0.88	0.89	0.92
	IRBFN	0.50	0.51	0.38	0.33	0.30	0.28	0.04	0.04	0.00	0.01	0.00	0.01	0.90	0.98	0.92	0.93	0.94	0.97
	LS-SVR POM	0.43	0.41	0.36	0.33	0.30	0.27	0.01	0.01	0.00	0.01	0.00	0.00	0.83	0.85	0.88	0.91	0.91	0.95
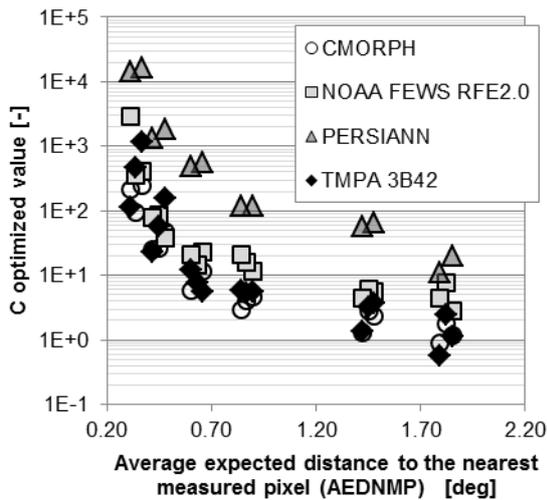
Figure 5.9. Evolution of the optimal C hyperparameter (LS-SVR) with spatial data availability. For each AEDNMP three results corresponding to the three different subsets are presented (outliers for PERSIANN's subset 2 are omitted).
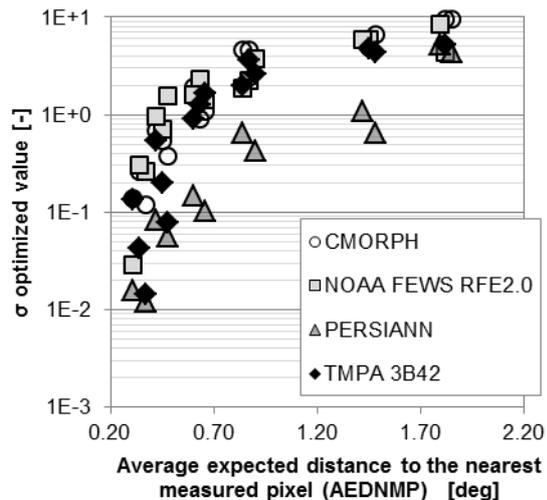
Figure 5.10. Evolution of the optimal σ hyperparameter (LS-SVR) with spatial data availability. For each AEDNMP three results corresponding to the three different subsets are presented (outliers for PERSIANN's subset 2 are omitted).

Moreover, it is interesting to observe that the TMPA 3B42 v6, NOAA/FEWS RFE2.0 and CMORPH datasets lead to strikingly similar results, which seems to indicate that the optimal hyperparameter's values depend more on the underlying nature of the rainfall process than on specific datasets used in the interpolation.

In order to produce the results published here, calculations were conducted in parallel in three high-end desktops and two calculation servers for a period close to one year. The derivation of simple laws relating AEDNMP (or equivalent spatial data availability evaluation measures) to adequate hyperparameter's values would greatly alleviate such a computational burden.

5.6 Interpolating from historical gauged rainfall series

5.6.1 Reduction of input's dimensionality and noise

The earlier assessment, conducted for proxy satellite data, relied on a limited number of training patterns (a maximum of 10 years of data in the case of the TMPA 3B42 v6 dataset). Additionally, the performance evaluation was not carried out exhaustively, but for a random number of points corresponding to a fraction of the total pixels.

In this more practical analysis, using directly gauged rainfall data, it is beneficial to use the maximum amount of training data possible. In this case 14 years (1998 to 2011) of daily TMPA 3B42 v7a satellite rainfall estimates were used in order to interpolate pixels corresponding to the whole ZRB. Their POM computation can be extremely heavy. Partly to address this issue, partly as a means to reduce seasonal estimation bias, the POM interpolation procedure was split into several seasonal models. Four seasons were identified based on mean TMPA 3B42 v7a estimates: a dry season from June to August, an ascending season from September to November, a wet season from December to March, and a descending season from April to May. Doing so, however, entailed the risk of impairing the interpolation of abnormal wet days during the dry season and, conversely, of dry days occurring in typically wet months. As such, following the implementation of a seasonal POM methodology, it was considered important to include some information about the full range of rainfall patterns observed throughout the year in the training data associated with each season. To do so, SOMs were adopted in order to classify 14 years of TMPA 3B42 data into a limited number of clusters.

The vectors used to characterize the data covering the ZRB area have 4928 entries (56x88 pixels) which constitute an appreciably high dimensionality. If directly introduced in a SOM, they would lead to both an unnecessary computational effort and the likely inclusion of noise in the resulting clusters. Alleviating both problems, Principal Components Analysis (PCA) was used to translate the original 4928 variables into 511 principal components prior to training the SOM. The determination of the number of principal components was made monitoring the percentage of total variance kept (87%) while balancing the number of components and the quality of the reconstructed rainfall maps. An example of the effect of the dimensionality reduction on one daily storm event is illustrated in Figure 5.11.

A 19x19 SOM output grid was adopted, corresponding to 361 clusters that approximately contain the main range of observed rainfalls over the area of interest and, therefore "summarize" the possible rainfall events. A subset of 9 of the total 361 characteristic daily rainfall events identified by the SOM is shown in Figure 5.12.

The choices of the number of principal components and size of the SOM were done after limited testing. Although a thorough study of these values would probably contribute to an improvement in the overall POM method performance, the choices described above sufficed to attain acknowlegeable results, as will be seen in the next section.
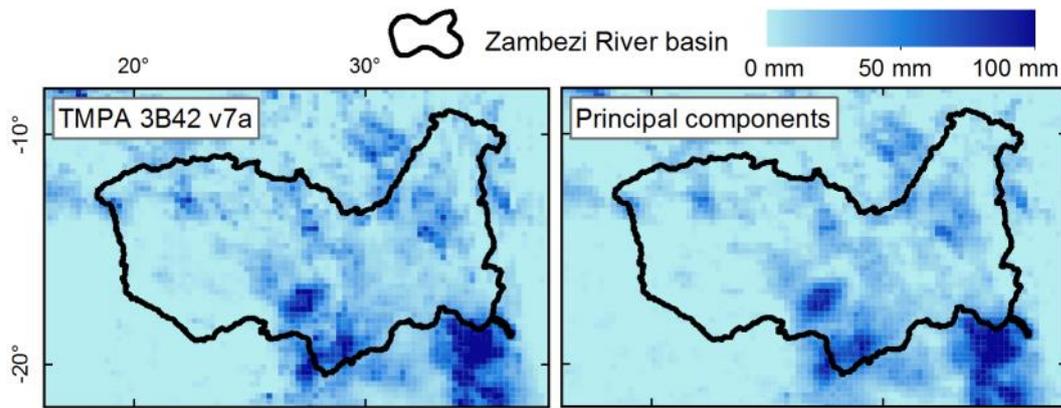
Figure 5.11. Example of the smoothing effect accomplished by the application of PCA (original data vs. data reconstructed from the 511 principal components).
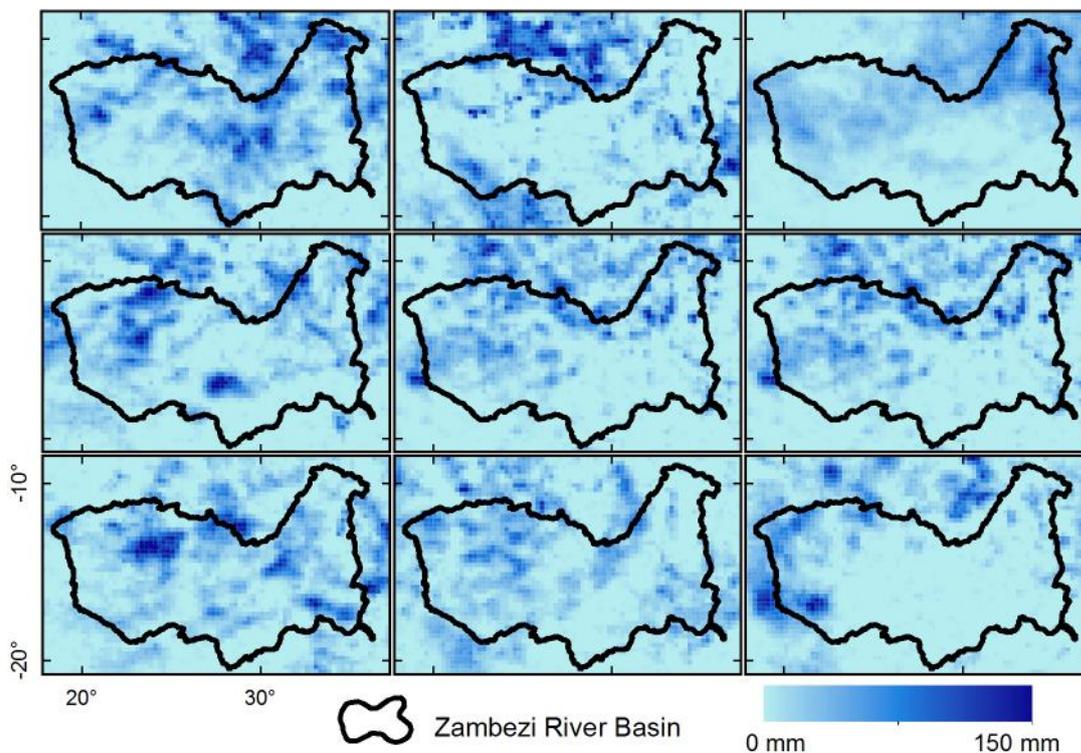
Figure 5.12. Illustration of a 3x3 subset of the characteristic rainfall patterns contained in the 19x19 trained SOM.

5.6.2   Cross-validation results from 1979 to 1997

Three methods, IDW, MROK and POM, were compared using a Monte Carlo cross-validation procedure. The results presented in this section were computed for interpolations from 1979 to 1997 using nearly 180 000 points, as mentioned previously.

LS-SVR hyperparameters were optimized for 500 points using the CMA-ES algorithm. In order to estimate adequate hyperparameters for all pixels and time steps of the analysis, a power law inspired in the relationships observed in Figure 5.9 and Figure 5.10 was then fitted to these 500 points using the distance to the nearest gauging station with valid measurements as an independent variable.

Daily results are summarized in Table 5.3. From the three interpolation methods, IDW led to the worst results, except in terms of VAR, the ratio of the interpolated series' variance over the gauged

rainfall series'. POM, obtaining the best results in the remaining measures, clearly outstands regarding RB, MAE, and r.

Putting into perspective these results, one can resort to earlier work by Cohen Liechti et al. (2011) on the match between satellite products and rain gauge measurements over the ZRB area in the period spanning from 2003 to 2009. In that contribution, the direct Pearson correlation found between TMPA 3B42 v6 data and GSOD rain gauges was of 28% with a bias of -0.11 mm/day. Although the reported values are not directly comparable with the r of 27% or a bias of -0.08 mm/day (corresponding to the RB of -4.4%) obtained here with POM (the time frames are different and Cohen Liechti et al. (2011) did not implement any interpolation method, having compared satellite and gauge data directly), they certainly highlight the quality of the interpolations achieved by the method, hinting that, at least in these measures, the POM interpolated values seem to be equivalent to TMPA 3B42 v6 data.

Table 5.3. Cross-validation performance measures for the three tested models at the daily scale. Measures computed for the period from 1979 to 1997. Best values highlighted in bold red font.

	NRMSE	RB	VAR	MAE [mm]	r
IDW	1.032	-8.9%	**31.0%**	2.703	22.1%
MROK	0.985	-9.5%	20.0%	2.713	25.7%
POM	**0.983**	**-4.4%**	22.0%	**2.615**	**27.0%**

For 30-day aggregations the results improve (as expected) but the relative comparison remains unchanged with POM outstanding as the best method (Table 5.4). Scatter plots of the cross-validation aggregated data points are illustrated in Figure 5.13

Table 5.4. Cross-validation performance measures for the three tested models. Aggregation over 30 days. Measures computed for the period from 1979 to 1997. Best values highlighted in bold red font.

	NRMSE	RB	VAR	MAE [mm]	r
IDW	0.719	-8.8%	**62.9%**	34.483	70.3%
MROK	0.691	-9.4%	57.1%	34.141	72.7%
POM	**0.684**	**-4.4%**	60.7%	**32.564**	**73.2%**
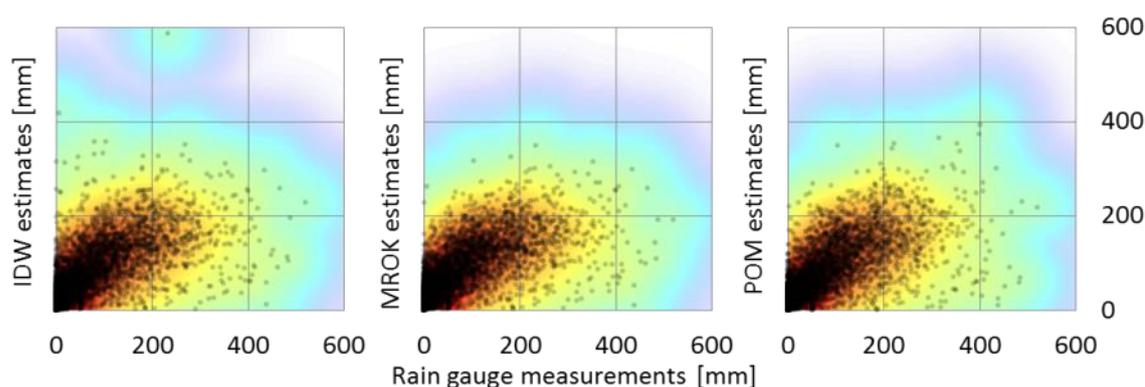
Figure 5.13. Scatter plots of interpolation estimates vs. gauged computed within the cross-validation scheme. 30 days' aggregated rainfall values. The normalized point density is shaded from red to white in the background.

Again putting the results into perspective, for monthly data Cohen Liechti et al. (2011) reported an r of 83% and a bias of 9.04 mm/month. While the aggregated POM results performed worse in terms of r (73%), they showed an equivalent bias of -2.5 mm/month, less than one third of the value calculated for TMPA 3B42 v6 monthly data. As is the case for daily data, these observations are only indicative as a direct comparison between values is not possible.

A geographical comparison and interpretation of the Pearson correlation coefficients calculated for MROK and POM interpolations of daily rainfall was also carried out and is illustrated in Figure 5.14 (IDW was left out due to its marginally worse results and in the interest of figure readability). In it, over every cross-validation location, the changes in r when POM is applied instead of MROK are presented. As can be seen, in the majority of the cases r was higher for POM than MROK, being that in the most isolated locations the differences are generally more acknowledgeable. This is in line with the observation that POM is particularly advantageous where spatial data is scarce.
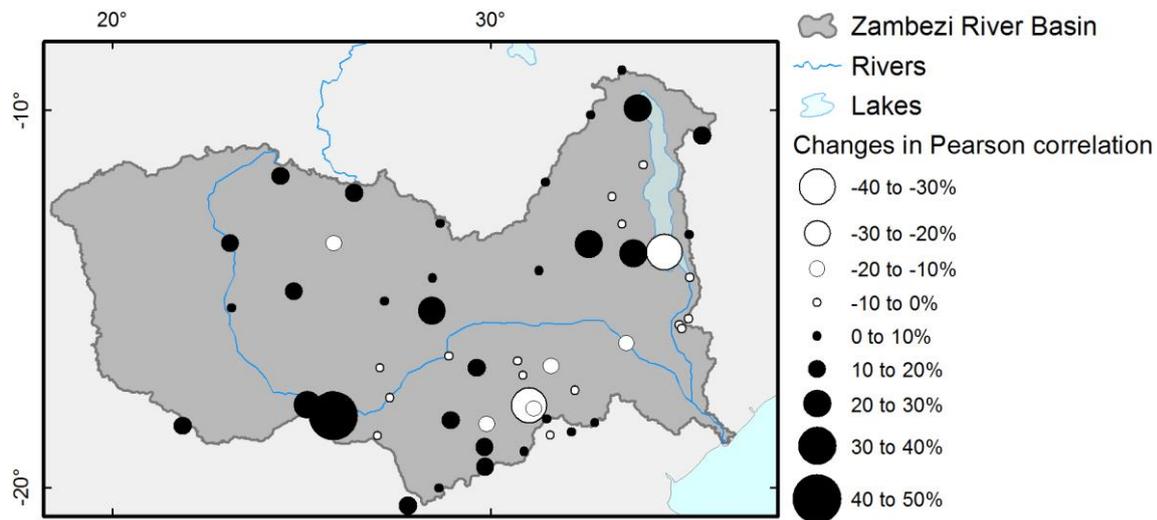
Figure 5.14. Changes in terms of local Pearson correlation coefficients when the POM method is applied instead of MROK.

Finally, comparing proxy data and real rain gauge measurements' results, it is relevant that POM seems to perform comparatively better for the latter. At first glance this would be unexpected due to the potential problems related to the change of support (the POM regression is being performed in the "space" of satellite rainfall estimates and not directly on gauged records). Apparently, those problems are outweighed by the increased variability that gauged rainfall fields usually display over satellite interpolated ones.

5.7   Potential and limitations

The POM interpolation method holds great promise as a consequence of the results presented above. Notwithstanding, despite historical rainfall interpolation performances toppling those of IDW or Kriging and seeming to be on par with that of recent satellite rainfall estimates, POM has shortcomings worth referencing.

The proposed interpolation method relies on the undertaking of a substantial number of non-linear regressions which involve heavy computational efforts. Namely, when implemented using the LS-SVR model, it can take orders of magnitude longer than Kriging and, specially, IDW. If hyperparameter optimization is taken into account, the gap between computation times widens even further.

Another issue is related to the complexity of the method and the associated hazard of generating poor interpolations. Namely, the user should be familiar with the principles underlying SVRs and clustering (if a seasonal implementation is made).

In order to perform the regressions target values must be supplied. Although recent satellite data can conveniently serve for training purposes over most of the World, regions where these data do not exist or perform badly can pose a challenge.

As stated previously, it stems from the nature of POM interpolation that the calculations are carried out in a field different from that of the inputs and outputs. This is done when, for example, rain

gauge input data is used to generate point interpolations, but the regressions are made based entirely on satellite data. Because point measurements and satellite data do not match perfectly, the implicit assumption that they are similar might contribute to some loss of performance or, ultimately, the inadequacy of the interpolation if both fields share no relation. Accounting for the difference explicitly through transformation functions could, consequently, be beneficial. Also, if the similarity between satellite data and the true rainfall fields is not demonstrated *a priori*, the quality of the interpolations should be corroborated by a process of cross-validation.

Finally, the results presented above can likely be enhanced by more thoroughly optimizing hyperparameters, defining a more sensible seasonal distribution of rainfall events, or finding a more adequate set of clusters to represent their range of possible spatial distributions.

5.8   Conclusions

Tested under a broad set of cases, the POM methodology has proven itself useful in order to interpolate sparse rainfall data, particularly so when the underlying field is marked by high variability. Under such conditions, POM was superior to any of the methods compared against it. POM was also shown to increase interpolation performance as more training data is available, which certainly highlights the potential of the method for future use.

In terms of its agreement with rain gauge data, namely regarding Pearson correlation coefficient and bias, POM's interpolation results are even comparable with direct estimates from recent satellite rainfall products.

Having disadvantages in its complexity and computationally demanding implementation, POM has great potential for improvement in several different aspects, the most relevant likely being: the enhancement of satellite rainfall estimates, the lengthening of the series used for training the regression models, and a better definition of the method's hyperparameters.

Still, POM has another valuable property: its interpolations derive many features from the areal rainfall field that is used as a basis for the interpolation. In line with the motivation that led to this work, when POM interpolated rainfall is applied to hydrologic models it effectively opens up new possibilities related to extended calibration and the simulation of historical events which would otherwise be difficult to exploit. This is discussed in detail in Chapter 8.

6 TOWARDS A MACHINE LEARNING LONG-RANGE DAILY STREAMFLOW FORECAST MODEL FOR THE UPPER ZAMBEZI RIVER BASIN

6.1 Motivation and objectives

Developing applications of physically based or semi-conceptual hydrological models such as SWAT can be an enormous effort (for more on this claim refer to Chapters 7 and 9). Even when the large amounts of data needed to set up and run the models are promptly available and in the case that a good knowledge of the basin is held, a long amount of time must usually be spent in preparation. To this, all the intricacies of a computationally demanding calibration process must be added.

In return, such models can constitute reliable assets to be used in order to better understand the basins' responses and offer the opportunity to capitalize on the gained knowledge in future decisions. Examples of their utility are the assessment of climate change effects, the planning of major hydroelectric schemes (e.g. Cohen Liechti 2013), the definition of environmental flows, the evaluation of land use practices, the calculation of water availabilities, the estimation of extreme flows, etc.

Notwithstanding their great usefulness, when physically based or semi-conceptual hydrological models models do not comprise continuous data acquisition (the case for SWAT and many others of the same type), they are seldom the best choice in order to perform tasks such as flood forecasting or the support of operational decisions.

Data-driven models, whether inspired in time-series analysis theory or machine learning, can be more flexible and, ultimately, a good compromise between performance and development requirements. If, on the negative side, data-driven models provide very limited understanding of the hydrologic processes and can only be used to retrieve information in determined points of the basin (those which were used for calibration), on the positive side they can be set up and calibrated in a relatively short period of time and easily adapted for forecasting purposes.

Stemming from this, the objectives of the work described in this chapter were:

- to understand the key aspects to take into account when implementing data-driven discharge forecasting models in a large-scale river basin;
- to assess which are the best performing algorithms to do so; and
- to contribute towards the improvement of existing data-driven discharge forecasting methodologies.

6.2   Introduction

River discharge and water stage forecasting have for long been a topic of scientific interest. In the literature, an abundant and growing list of contributions addressing them can be found. Historically, three main approaches have been employed in the task of river discharge forecasting: physically based, conceptual, and data-driven. This chapter focuses on the latter. Traditionally, the most popular data-driven models applied to discharge and water stage forecasting are the Autoregressive Moving Average (ARMA) and derivations (e.g. Abrahart and See 2000, Mohammadi et al. 2006, Anderson et al. 2013) and ANNs (e.g. Abrahart and See 2000, Toth and Brath 2007, Piotrowski and Napiorkowski 2011, Sivapragasam et al. 2013), but alternative methodologies have been proposed such as SVR (e.g. Lin et al. 2006, Wang et al. 2009), functional networks (Bruen and Yang 2005), recurrent fuzzy systems (Evsukoff et al. 2011), recurrent neural networks (Kumar et al. 2004), genetic programming (Wang et al. 2009, Londhe and Charhate 2010), model trees (Londhe and Charhate 2010), or k-nearest neighbors (Sivakumar et al. 2002, Wu and Chau 2010). Although the performances reported in literature vary greatly (Wu and Chau 2010), authors typically acknowledge the newly proposed methods as promising substitutes when compared to traditional time-series tools such as ARMA or its integrated and/or seasonal counterparts (commonly addressed as SARIMA).

Here, a flow forecast system is proposed and evaluated for the Zambezi River at Victoria Falls. That section drains the over 360 000 km^2 Upper Zambezi subbasin. Not far downstream the section lays Kariba, whose operation greatly influences the Zambezi basin's economy and ecosystem. Characterized by its large size, pronounced seasonal changes in its hydrological response, covering a range of climates from tropical to semi-arid, and being influenced by large and complex wetlands, the Upper Zambezi presents an interesting challenge with regard to flow forecasting. The location and main features of the catchment are presented in Figure 6..
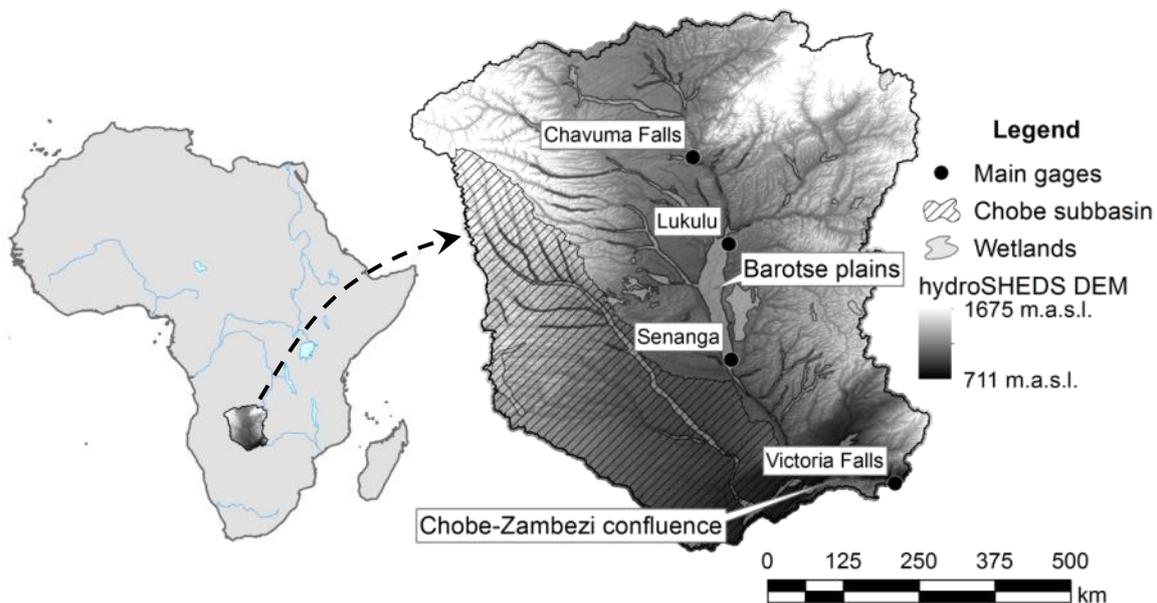
Figure 6.1. Location and main features of the Upper Zambezi River basin.

Forecasts for lead times of 7, 30 and 60 days, based on discharge gauge measurements and satellite rainfall estimates, are produced according to different models – ARMA, MLP, SVR and LS-SVR – which provided promising results in early studies (Matos et al. 2013b).

Data-driven models are particularly sensitive to the number, relevance and independence between inputs (Akhtar et al. 2009). Due to this, the data pre-processing strategies leading up to the definition of model inputs can be as important as the choice of the model itself. Among other aspects, parsimony in the number of inputs is of paramount importance. When using areal rainfall observations as exogenous variables – as is the case here – an adequate rainfall aggregation

scheme is useful. Extending earlier works (Akhtar et al. 2009, Matos et al. 2013b), a rainfall aggregation scheme based on hydraulic considerations was developed.

The performance of the various models is evaluated under the scope of five error measures and discussed in detail. Namely, the possible benefits and risks of applying non-linear models are examined. Emphasis is also put on the properties of forecasts obtained with SVR and LS-SVR, which have been reported in literature to perform well (e.g. Liong and Sivapragasam 2002, Asefa et al. 2006, Lin et al. 2006) but, as will be shown, entail certain risks to their application.

Another issue commonly associated with forecasts is their uncertainty (Beven and Binley 1992, Krzysztofowicz 2002, Montanari 2007, Moradkhani and Sorooshian 2008, McMillan et al. 2012). A great number of publications have been devoted to the subject, which has originated a lively debate within the hydrological community (e.g. Mantovan and Todini 2006, Beven et al. 2007, Mantovan et al. 2007), but quantifying the uncertainty of river discharge forecasts remains a topic of major interest. In particular, there are great advantages in associating occurrence probabilities to discharge forecasts (Cloke and Pappenberger 2009). Several ways of achieving this goal have been proposed, from the combination of ensemble weather forecasts with deterministic hydrological models to informal and formal Bayesian frameworks. In general, such models are computationally expensive, rely on complex theory, and include underlying assumptions that may not be verified in practice. More importantly, in many cases they fail to achieve accurate representations of the forecasts' probabilities. The last part of this chapter is dedicated to the introduction of a conceptually simple, assumption free, computationally parsimonious and highly flexible methodology of obtaining empirical representations of forecast probability distributions.

Although an attempt has been made to write the chapter in a coherent and sequential manner, Sections 6.4, 6.5 and 6.6 in particular correspond to diverse topics and should be, to some extent, regarded in that light. Section 6.4 compares several deterministic forecasting models and tests a new scheme of rainfall aggregation based on hydraulic considerations. Section 6.5 focuses on characteristics of the SVR and LS-SVR models and provides evidence that, despite their good comparative performance and under certain conditions they can produce particularly faulty forecasts. Finally, results from a novel empirical uncertainty post-processor with great potential are presented in Section 6.6.

6.3   Data and methods

6.3.1   Data
The data used in the works described in this chapter is comprised of two DEMs (HydroSHEDS and SRTM v4.1, see Section 4.2 for details), historical discharge data from the Chavuma, Senanga and Victoria Falls gauging stations (Figure 6.1, see Section 4.4 for details) and TMPA 3B42 v7a rainfall satellite estimates (described in Section 4.3).

For the purpose of rainfall aggregation, a map of wetland areas (Wetlands version 1.0 of the global polygon dataset compiled by UNEP World Conservation Monitoring Centre, 1993) and a number of river widths (retrieved manually from Google Earth imagery) were also used.

6.3.2   Models
In an early work, not presented here, a total of 7 data-driven models were tested regarding their accuracy when performing 30 and 60 days' lead time forecasts for discharge at Victoria Falls (Matos et al. 2013b). Both feedforward and recurrent alternatives (capable of reproducing dynamical responses) were tested, namely:

- MLPs, RBFNs, and LS-SVR, as feedforward models;
- ARMA, layer-recurrent ANNs, Non-linear Autoregressive with Exogenous Inputs (NARX) MLPs, and NARX LS-SVR, as recurrent models.

From the results then achieved it was evident that the substantially higher computational demands of non-linear recurrent models (the latter named above except ARMA, which is linear) do not appear to lead to substantially improved forecasts. Quite the opposite: they seem to promote a higher variability in the results. On the side of feedforward alternatives, RBFN displayed unsatisfactory performance. Resulting from the reported findings, the presented analysis focused on ARMA, MLPs, SVR and LS-SVR.

There are several ways in which to pose the forecasting problem, particularly considering the machine learning models. The ARMA model (in this case better named as ARMAX, as it uses exogenous inputs such as rainfall or upstream discharge series) followed a standard implantation in which one step-ahead forecasts are made. This implementation is schematized in Figure 6.2, where $\hat{y}_i$ is the desired forecast, x^* is the input vector and $f\left(x^*, W\right)$ represents the ARMAX model[1].

Feedforward models have the ability to directly provide forecasts for any chosen lead time. Figure 6.3, which illustrates the adopted feedforward model's operation, mainly differs from Figure 6.2 as there is no feedback loop. Due to this, the lead time, like the choice of past observations, is used as input and needs to be defined externally[2].
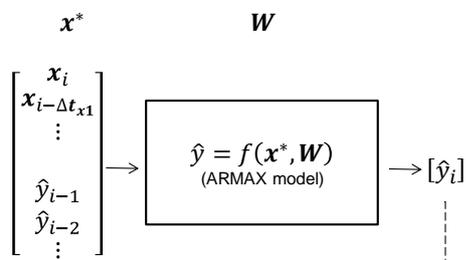
Figure 6.2. Scheme of an ARMAX model operation.
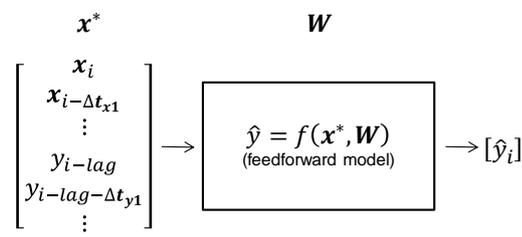
Figure 6.3. Scheme of a feedforward model operation.

Autoregressive with moving average

ARMA are mathematical models of the persistence, or autocorrelation in a time series (Wu and Chau 2010). Widely used in hydrology, an $\mathrm{ARMA}\left(p, q\right)$ model is characterized by p, the autoregressive order, and q, the moving-average order. More complex versions of ARMA can include a polynomial trend in the underlying time series or account for seasonal effects. If both are considered, a seasonal autoregressive integrated moving-average model, described in detail in several publications (e.g. Brockwell and Davis 2002) and represented as SARIMA $\left(p, o, q\right) \times \left(P, O, Q\right)_s$ can be defined by equation (6.116):

$$\Omega\left(B^s\right)\omega\left(B\right)\left(1-B^s\right)^O\left(1-B\right)^o X_t = \alpha + \Psi\left(B^s\right)\psi\left(B\right)Z_t \tag{6.116}$$

where B is a delay operator $\left(B^i X_t = X_{t-i}\right)$, Z represents a white noise process, and Ω, ω, Ψ, and ψ operators that act as described in the set of equations (6.117).

---

[1] x^* is comprised of exogenous inputs x_z and past forecasts $\hat{y}_z$. The index z is defined by the current time step (i) and a lag. While for $\hat{y}_z$ the lags are incremental, for x_z they are defined outside the model, hence the use of Δt_{x1}.

[2] The lead time is now explicitly defined as the *lag* index. The series of past observations does not need to be continuous as for the ARMAX model. The chosen past observations to be used are defined by the Δt_{y1} index.

$$\begin{cases} \omega(B) = 1 - \omega_1 B - \omega_2 B^2 - \ldots - \omega_p B^p \\ \psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots + \psi_q B^q \\ \Omega(B^s) = 1 - \Omega_1 B^s - \Omega_2 B^{2s} - \ldots - \Omega_P B^{Ps} \\ \Psi(B^s) = 1 + \Psi_1 B^s + \Psi_2 B^{2s} + \ldots + \Psi_Q B^{Qs} \end{cases}$$

(6.117)

The first set of parameters of the $\text{SARIMA}(p,o,q) \times (P,O,Q)_s$ model is associated with the autoregressive component (p), the degree of the underlying polynomial trend (o), and the moving-average window size (q). The seasonal component is captured by the analogous second set, defined by the seasonality period (s), associated autoregressive (P), trend (O), and moving-average (Q) parameters.

The SARIMA model, presented above for the univariate case, is readily extended for multiple variables. In the present work, example daily series without evident underlying trends are used. This implies that the integration components are not needed and that the seasonality period is expected to be of approximately 365 days. Unfortunately, using such a large seasonality period has very heavy computational costs. In order to avoid these costs while capturing the seasonal behavior, a Fourier series with eight terms was fitted to a training series and presented to an ARMA model with exogenous inputs: $\text{ARMAX}(p,q)$.

Multilayer perceptron

Numerous successful applications of MLPs have been reported in hydrology and, particularly, in attempts to forecast discharges (e.g. Baratti et al. 2003, Kumar et al. 2004). The model's description, done in Section 3.4.2, will not be repeated here. However, a particularly relevant aspect of the MLP should be emphasized. As previously stated, MLP training is not deterministic, being affected, among other things, by initial conditions and the choice of training and validation subsets. As such, it is basically not viable to reproduce a trained MLP's response with an independently trained model, even if similar in its architecture[3]. Owing to the same reasons, if conditions are altered the MLP's response will be different, and thus its performance.

Stemming from such considerations, it makes little sense to define or compare the performance of a single instance of a chosen MLP architecture with that of alternative models as the results are not generalizable.

In order to overcome this, an ensemble of MLPs should be used in order to characterize the performance of a given architecture. In an application to discharge forecasting, it is also reasonable to use, instead of a single MLP, an ensemble of selected trained models.

Support vector regression least-squares support vector regression

SVR and LS-SVR models have been already introduced in Sections 3.4.4 and 3.4.5, not being their description reproduced here. Like for the MLP, however, particularities of SVR and LS-SVR in regard to flow forecasting are discussed.

The application of SVR as a data-driven discharge and water stage forecasting tool is more recent than that of ARMA or MLP models. *Liong and Sivapragasam* (2002) compared SVR with an ANN model for 1 and 7 days lead time forecasts, stating that SVR offers better generalization performance, particularly when the available training set is limited, but remarking that such SVR implementations were recent and should be further researched. *Lin et al.* (2006) compared SVR to ARMA and ANN models for long-term discharge prediction finding that SVR performed better and advocating it as a strong candidate for the prediction of long-term discharges. Likewise, the

---

[3] This is "only" a practical limitation. If all the relevant training and validation parameters are reproduced the response from different MLP models will be the same. In spite of this, the full set of relevant information is never provided in practice.

capability of SVR to forecast short and long term streamflows was highlighted by *Asefa et al.* (2006). Additional results supporting good comparative performances by SVR models have been published by *Yu et al.* (2006) and *Lin et al.* (2009) regarding sub-daily forecasts. Elaborate implementations, such as the crisp distribution SVR, which uses several independent SVR models adapted to clustered input data to perform forecasts, surpassed competing methods (Wu et al. 2009). LS-SVR has been compared to ARMA and ANN models also with promising results (Samsudin et al. 2011, Matos et al. 2013b).

Due to the structural risk minimization principle they follow, SVR and LS-SVR tend towards "reasonable" predictions when inputs stretch outside the training domain. Such "reasonable" predictions are typically within the range of values found during training. This behavior, revealed with non-linear RBF kernels, can be greatly advantageous in many regression problems and is not present in other non-linear regressors such as most ANNs (unlike SVR or LS-SVR these become unpredictable outside of the training domain). Figure 6.4 illustrates the principle in operation for two simple problems. There, an input variable x is used to predict y. In both examples the fitting in the region used for training is excellent, but quickly, as x values depart from the training domain, predictions for y converge to the aforementioned "reasonable" values.
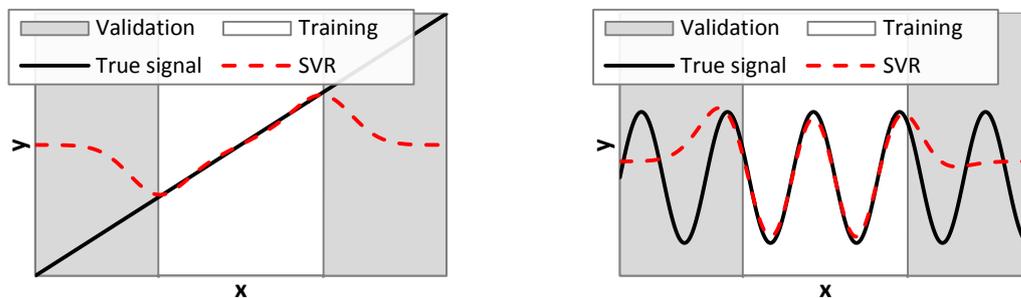
Figure 6.4. Simple examples of the structural risk minimization principle at work for SVR with RBF kernels.

Two other examples can demonstrate part of the influence of the ε-insensitive loss function implemented in SVR. Firstly, a set of linear observations contaminated by white noise and the presence of outliers is approximated by SVR and LS-SVR linear models in Figure 6.5 While the LS-SVR predictions are clearly affected by the three outlier observations (as would most models relying on the minimization of the Mean Square Error, MSE), SVR remains robust owing to its loss function (based on the Mean Absolute Error, MAE). From another angle, depicted in Figure 6.6, a sine wave is approximated by linear kernel SVR and LS-SVR models. Due to the symmetry of the input wave and the limitations of the regressors, after training, constant (horizontal) estimates are obtained. What is noteworthy here is that, while the LS-SVR model's estimate coincides with the average of the input wave, as the majority of the observations fall below this average, so do the SVR's predictions. In summary, while the ε-insensitive loss function renders SVR robust to outliers, it also can lead to volume estimation failures.
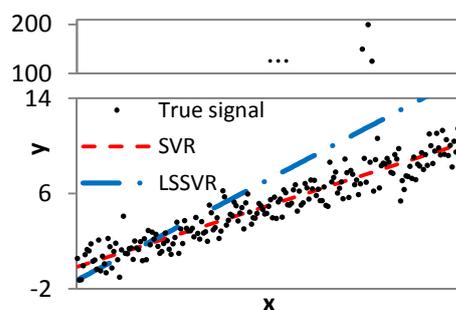
Figure 6.5. How the ε-insensitive loss function affects SVR in the presence of outliers.
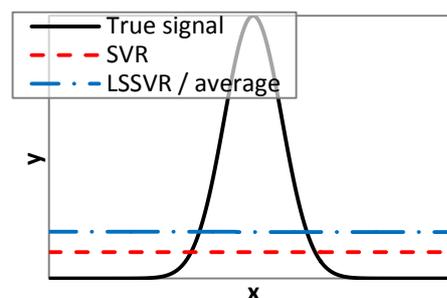
Figure 6.6. How the ε-insensitive loss function can lead SVR to volume errors.

6.3.3 Performance evaluation

A set of five commonly used error measures were considered in the analysis, namely the Nash-Sutcliffe efficiency (NE) coefficient (Nash and Sutcliffe 1970), the Root MSE (RMSE), the MAE, the Peak Ratio (PKR), and the Volume Ratio (VOR).

The first four measures are defined for convenience in equations (6.118) to (6.121). Being commonly used in hydrology, it suffices to clarify that, in the equations, N represents the length of the observed series, Q_t the observed discharge at time step t, $\hat{Q}_t$ the corresponding simulated discharge, and $\bar{Q}_t$ the average observed discharge.

$$\text{NE} = 1 - \frac{\sum_{t=1}^{N}\left(Q_t - \hat{Q}_t\right)^2}{\sum_{t=1}^{N}\left(Q_t - \bar{Q}_t\right)^2} \tag{6.118}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}\left(Q_t - \hat{Q}_t\right)^2}{N}} \tag{6.119}$$

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|Q_t - \hat{Q}_t\right| \tag{6.120}$$

$$\text{VR} = \sum_{t=1}^{N}Q_t \left/ \sum_{t=1}^{N}\hat{Q}_t \right. \tag{6.121}$$

The definition of the considered peaks is subjective. In the present work, only maximum yearly discharge observations were taken as peaks and the value for PKR was computed according to equation (6.122).

$$\text{PKR} = \frac{1}{\Pi}\sum_{\pi=1}^{\Pi}\frac{Q_{t(\pi)}}{\hat{Q}_{t(\pi)}} \tag{6.122}$$

where Π represents the number of years with valid peaks, π stands for the year of interest, $t(\pi)$ is an index identifying the day in which the maximum discharge observation in year ξ is observed and, finally, $Q_.$ and $\hat{Q}_.$ represent observed and estimated discharges, respectively.

6.3.4 Rainfall aggregation scheme

The substantial time flows take to travel along the Upper Zambezi added to the regularization effect of major wetlands such as the Barotse Plains and the Chobe-Zambezi confluence make the accuracy of relatively distant forecasts adequate for a number of purposes, particularly if information from relevant gauging stations upstream or rainfall is considered. In order to conduct the present analysis areal rainfall estimates resulting from Pattern-Oriented Memory (POM, see Chapter 5) interpolation and TMPA 3B42 v7a data were joined, covering the period from 1982 to 2009.

All the tested forecasting models are sensitive to the number of inputs, being that the amount of information contained in each series and the detrimental effects of adding noise to the model must be weighed in every case. None of the tested methods is able to efficiently cope with the large amount of data contained in pixel-wise rainfall series, which would amount to hundreds of inputs. As such, a method of aggregation that is able to reduce this load while preserving the maximum of useful information is needed. Desirably, the method should operate in the space and time dimensions.

A straightforward way of achieving such an aggregation, implemented by Matos et al. (2013b), could start by computing distances to the section of interest. With these distances, knowing the average time of travel of the floods between a number of control points, a map could be produced with the travel times to the section of interest (Figure 6.7). Depending on the desired forecast horizon, relevant rainfall bands could be defined, the pixel data within each band be aggregated, and a smoothing algorithm applied on the time dimension (e.g. moving windows) in order to successfully pre-process lower-dimension rainfall inputs.
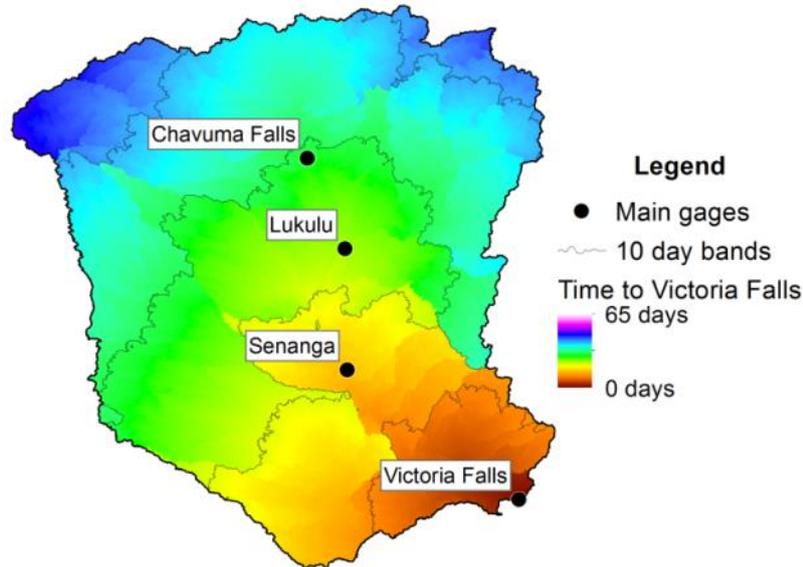
Figure 6.7. Map of travel times to Victoria Falls considering a calibrated constant flow velocity.

Such an approach does, however, implicitly contain major simplifications (e.g. a constant flow travel time) which seldom hold in practice. In the past, Akhtar et al. (2009) have addressed this issue by assuming different velocities for overland and channel flows, which albeit constituting an improvement, is still a substantially simplified approach.

Here, a more precise methodology for obtaining a map of travel times was devised. Assuming the validity of Manning's equation, the mean channel flow velocity can be calculated by equation (6.123):

$$u = \frac{1}{n_{ch}} \cdot R_{rch}^{2/3} \cdot s_{ch}^{1/2} \tag{6.123}$$

where u is the mean flow velocity, n_{ch} represents the roughness coefficient, R_{rch} stands for the hydraulic radius and s_{ch} is the channel slope.

Additionally assuming that $R_{rch} \approx h$ – a condition valid when the channel width, w, is appreciably larger than h – the water height can be explicitly derived from (6.123) for any given flow, Q, using equation (6.124). Under the same conditions, and through simple manipulation of equations (6.123) and (6.124), equation (6.125) ensues, providing a practical means for estimating the mean channel velocity.

$$h = \left(\frac{Q}{1/n_{ch} \cdot s_{ch}^{1/2} \cdot w} \right)^{3/5} \tag{6.124}$$

$$u = n_{ch}^{-3/5} \cdot s_{ch}^{3/10} \cdot Q^{2/5} \cdot w^{-2/5} \tag{6.125}$$

More relevantly, from (6.125) stems that the ratio of the mean flow velocities in any two channels can be obtained from the ratios of their roughness, slope, flow and width, as shown in equation (6.126):

$$\frac{u_1}{u_2} = \left(\frac{n_{ch,1}}{n_{ch,2}}\right)^{-3/5} \cdot \left(\frac{s_{ch,1}}{s_{ch,2}}\right)^{3/10} \cdot \left(\frac{Q_1}{Q_2}\right)^{2/5} \cdot \left(\frac{w_1}{w_2}\right)^{-2/5} \tag{6.126}$$

According to Rodríguez-Iturbe and Rinaldo (2001), citing Leopold et al. (1964), reference flows, Q, can be described by a $Q \propto A_{sb}^{\delta}$ relation, where A_{sb} is the contributing area and δ a constant exponent. In the case of a recurrence interval of 1.5 years $\delta = 0.75$, while for the mean annual discharge $\delta = 1$. In the scope of the present contribution a value of $\delta = 0.75$ was taken. Furthermore, a rainfall term was added to the relation, as shown in (6.127), where $\bar{P}$ represents the average annual rainfall in the contributing area.

$$Q \propto A_{sb}^{\delta} \cdot \bar{P} \tag{6.127}$$

Because only the ratio of the reference flows is relevant in equation (6.126), the determination of net rainfalls is not important as long as similar precipitation losses can be admitted. The channel slope term can be very easily obtained from a DEM. Channel width can be acquired from aerial surveys. Based on the fact that there is a relationship between channel width, w, and discharge given by (6.128) (Rodríguez-Iturbe and Rinaldo 2001, citing Leopold and Maddock, 1953), fitted values for α and β were obtained for several river cross-sections of the Upper Zambezi (Figure 6.8).

$$w = \alpha \cdot Q^{\beta} \tag{6.128}$$
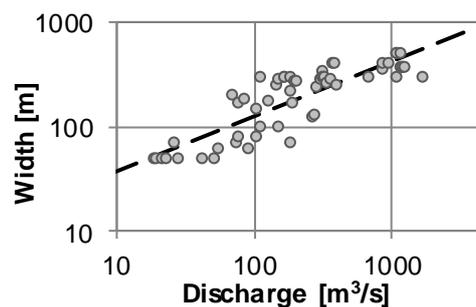
Figure 6.8. Fitted relationship of discharges and channel width for control points in the Upper Zambezi basin.

Under such a methodology for calculating channel velocities, roughness coefficients are the only parameter relevant which is difficult to obtain without undertaking direct surveys.

Combining (6.126) and (6.128), a simple equation that relates the mean velocities in two distinct channels can be obtained:

$$\ln\left(u_1\right) = \ln\left(u_2\right) - \frac{3}{5}\log\left(\frac{n_{ch,1}}{n_{ch,2}}\right) + \frac{3}{10}\log\left(\frac{s_{ch,1}}{s_{ch,2}}\right) + \frac{2}{5}\left[\delta\log\left(\frac{A_{sb1}}{A_{sb2}}\right) + \log\left(\frac{\bar{P}_1}{\bar{P}}\right)\right]\left(1-\beta\right) \quad (6.129)$$

Based on equation (6.129), a map of relative flow velocities was computed for the Upper Zambezi. Flow ratios were calculated resorting to a calibrated equation (6.127), channel slopes derived according to a DEM, and two admitted classes of channel roughness where the n_{ch} in wetlands was considered 4 times larger than in other locations.

From the relative flow velocities an adimensional travel time map was computed. Using the mean flow travel times from Chavuma (28 days), Lukulu (25 days), and Senanga (9 days) to Victoria Falls (Matos et al. 2013b), the travel times were calibrated. Overland flow velocities, assumed to be dominant down to 800 m from the watershed divides and admitted to be of 1 cm/s, were finally used to complete the travel time map for the whole catchment. The map, presented in Figure 6.9, diverges greatly from the one based on flow lengths alone (Figure 6.7) and, unlike it, evidences the expected fractal imprint due to the acceleration of the flow in the main channels.
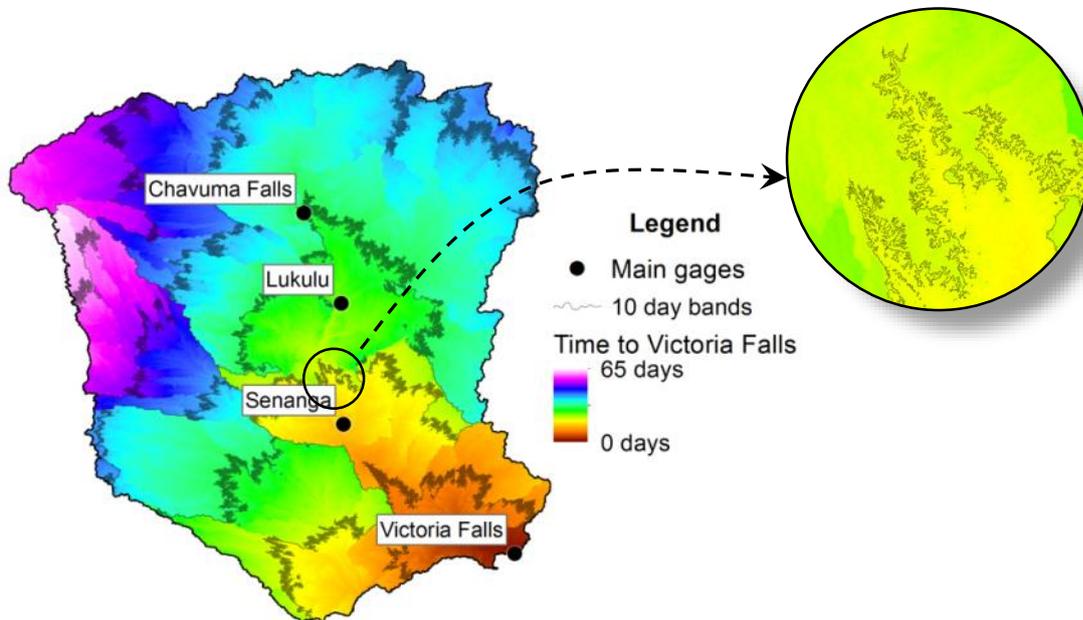
Figure 6.9. Map of travel times to Victoria Falls calculated using equation (6.129). Detail of the fractal imprint due to the acceleration of the flow in the main channels.

For each of the three analyzed forecast horizons (7, 30 and 60 days) and based on the map presented in Figure 6.9, the daily rainfall values were aggregated into relevant bands. Because of the topography of the Chobe-Zambezi confluence, marked by small slopes and large wetland areas, the Chobe subbasin (see Figure 6.1) does not always contribute to flows in the Zambezi's main reach. This knowledge was taken into account in the definition of the bands. For the 7-day forecasts three bands were defined within the zone estimated to contribute more to the discharge (from 4 to 10 days' time to Victoria Falls). The aggregated rainfall for each band was then rescaled to discharge units and bands contributing less than 4 m³/s on average were aggregated. The resulting two bands can are illustrated in Figure 6.10 (left).

A similar process was implemented for the 30 and 60 days forecast horizons. In the former, the main contribution zone was assumed to span from 24 to 36 days' time to Victoria Falls and an average threshold of 30 m³/s was considered for band aggregation (Figure 6.10, middle). Regarding the latter, the main contribution zone was enlarged due to higher uncertainty and taken between 26

and 72 days' time to Victoria Falls, being the average threshold assumed to be of 60 m³/s (Figure 6.10, right).
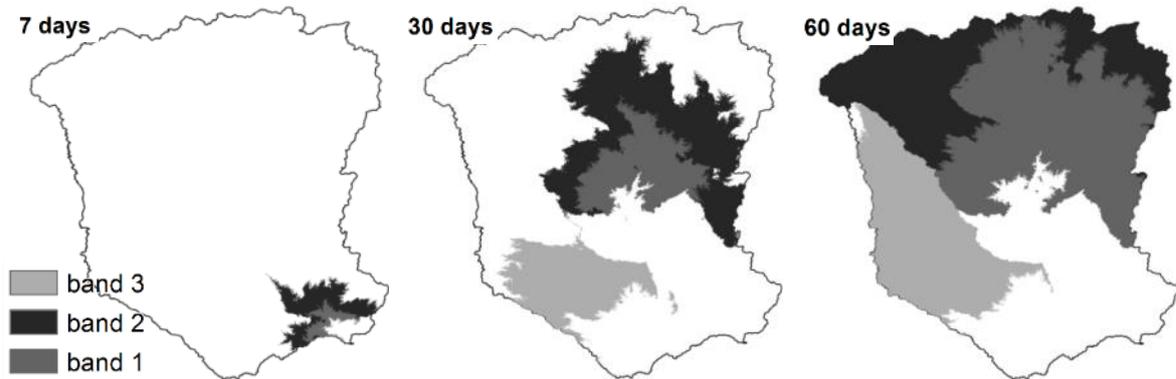
Figure 6.10. Adopted rainfall aggregation bands for different forecast lead times.

With the final definition of the contribution bands, the spatial reduction of dimensionality of rainfall inputs was complete. Time-wise, a simple leaky integrator or fading-memory filter was applied (Zarchan and Musoff 2009) according to equation (6.130).

$$\hat{x}_t = \hat{x}_{t-1} + (1 - \varphi)(x_t - \hat{x}_{t-1}) \tag{6.130}$$

where $\hat{x}_t$ is the filtered estimate for time step t, $\hat{x}_{t-1}$ is the previous estimate, φ represents the fading constant and x_t stands for the actual measurement or estimate at time step t.

According to the definition of the fading-memory filter, the contribution, χ, of each past observation is given, asymptotically, by an exponential function (6.131).

$$\chi_{t-i} = \delta \cdot e^{i \cdot \varphi} \tag{6.131}$$

where i is the order of the observation, and δ and φ are constants[4].

With this knowledge, it is possible to define φ such that 90% of the contributions come from observations within the past k days by considering the integral in equation (6.132).

$$\sum_{i=0}^{k-1} \chi_{t-i} \approx \int_{0.5}^{k+0.5} \alpha \cdot e^{z \cdot \varphi} dz \tag{6.132}$$

In the scope of the present work two values for k were tested (10 and 20 days, corresponding to φ's of 0.79 and 0.89, respectively).

6.3.5   A novel, simple and non-parametric way of evaluating forecast uncertainty
The interest of the hydrological community in characterizing hydrological uncertainty and the probability distribution of forecasts was already introduced, as well as some of the difficulties associated with this process.

---

[4] The δ and φ constants are used here exclusively in the scope of the fading-memory filter and should not be confused with earlier variables (e.g. from equations (6.127) and (6.129)).

Cloke and Pappenberger (2009) have done an extensive review on ensemble flow forecasting. Concluding, they highlight that forecasts based on ensemble prediction systems do not result in true probabilities of flooding, as uncertainties are not treated fully and the assumptions of some of the approaches are violated.

Most informal and formal approaches rely on strong assumptions regarding model errors such as normality, independence or homoscedasticity; assumptions that are manifestly not valid regarding streamflow forecasts. Through the introduction of more complex schemes, however, these assumptions can be relaxed and replaced by less constraining ones. Examples are the assumptions of autocorrelation or of a linear relationship between the standard deviation of the errors and flow (e.g. Schoups and Vrugt 2010). Albeit constituting powerful tools, models of such complexity are hard to set up and demand substantial computing power, being ultimately limited by the assumptions that underlie them in all cases.

Other approaches take forecast ensembles that do not represent the true probability distribution of streamflow magnitudes and correct them by applying methodologies such as wavelet transformation (Bogner and Pappenberger 2011). Again, methodologies that rely on prior assumptions and are appreciably complex.

On a conceptual plane, focusing on the probability of the forecast coming true and not on the probability of a given event taking place has some advantages: firstly, because the probability of a natural event taking place is an abstraction impossible to verify directly and only subject to be inferred (being that the validity of such an inference can be argued upon) and, secondly, because focusing on the probability of the forecast coming true allows for an empirical evaluation of the forecasting capabilities.

In the developed research the idea of producing realistic empirical probability distributions of streamflow forecasts is exploited in a simple, non-parametric way.

In order to do so, one can start with a deterministic regression model such as a MLP, which can provide a forecast, $\hat{y}$, given a set of inputs, $\mathbf{X}$, and internal parameters, $\mathbf{W}$.

$$\hat{y} = f\left(\mathbf{X}, \mathbf{W}\right) \tag{6.133}$$

The internal parameters obviously define the characteristics of the forecast. Usually, for MLPs a backpropagation algorithm is applied during training. As previously mentioned, differences in initial conditions and/or training and validation subsets can influence the result and, by training several MLPs, an ensemble of forecasts with some variability is obtained. This variability, however, is related to the nature of the MLPs and not the phenomenon being modeled because every member of the ensemble simply tries to approximate the "true" response. Consequently, a MLP ensemble resulting from several backpropagated networks can be expected to contain little information about the uncertainty of its forecasts.

If the whole ensemble is trained in parallel, however, it can be used to empirically represent the probability distribution of its forecasts coming true. This can be achieved by optimizing internal parameters ($\mathbf{W}$) with a multi-objective optimization scheme (see Section 3.3.3 for a discussion on this type of optimization models) that approximates the Pareto surface spanned by at least two conflicting objectives of a particular sort.

The first objective aims at making each member of the ensemble as accurate as possible[5]. For example, the MSE or the MAE can be minimized. Here the MSE was used.

---

[5] This is where backpropagation and other common methods to train MLPs stop.

The second objective aims at minimizing the proportion of the observations which exceed the member's forecast. This can be achieved by a measure such as the Exceedance Fraction (EF), calculated according to equations (6.134) and (6.135):

$$EF = \frac{\sum_{t=1}^{N} \vartheta_t}{N}$$ (6.134)

$$\begin{cases} \vartheta_t = 1 & \text{if } Q_t > \hat{Q}_t \\ \vartheta_t = 0 & \text{otherwise} \end{cases}$$ (6.135)

where N is the number of training observations, Q_t is the observed flow at time step t, and $\hat{Q}_t$ is the corresponding forecast.

Obviously, MSE and EF are conflicting objectives. While decreasing the MSE will tend towards an EF of 0.5[6], an EF of 0 means that a member's forecasts always exceed the observations. When a Pareto surface mapping the optimal trade-offs between EF and MSE is obtained, one can simply slice that surface at a given EF band in order to obtain the best forecasts (according to MSE) with a desired probability of non-exceedance. By evaluating selected slices of the surface, one can achieve the characterization of the probability of each coming true[7].

Finally, an optional regularization criterion aiming to train and select more parsimonious ensemble members can be included in the set of objective functions. While this can improve the quality of the results, it does not alter in any way the concepts introduced above.

6.4 Model comparison

A comparison of deterministic data-driven models for forecasting streamflows at Victoria Falls was made. The comparison included ARMAX, MLP, SVR and LS-SVR models operating at lead times of 7, 30 and 60 days.

Producing daily forecasts can be particularly challenging due to the high period associated with the yearly high-low flow cycles. For example, models such as seasonal ARMA are hardly computationally feasible when periods of 365 observations are needed. In order to circumvent this limitation, seasonal models were not applied, but a fitted Fourier series was included as an external input.

For every model and lead time, three sets of inputs were evaluated:

- past streamflows observed at Victoria Falls and corresponding Fourier series approximation – F input set;
- F plus aggregated observed rainfall – FR input set;
- FR plus observed discharge at upstream stations (Chavuma Falls and Senanga) – FRD input set.

In Figure 6.11 the historical series at Victoria Falls, Senanga and Chavuma falls are presented. Also, model training and evaluation periods are indicated.

In order to promote a fair comparison, all combinations of model, lead time and input set were evaluated for the same time frames. Due to gaps in the Chavuma Falls and Senanga discharge

---

[6] Only approximately. In reality more could be said on the subject as for a non-simetric probability distribution this is not the case.
[7] A Pareto surface generated according to the above description will span between an EF of 0 and 0.5 (approximately), which corresponds to the upper half (50%) of the probability distribution associated with the forecasting errors. Because the problem is symmetric, in order to obtain the lower half, it suffices to minimize 1-EF instead of EF.

data, the evaluation periods were subject to some constraints. The effective observations used for model evaluation are presented in Figure 6.12.
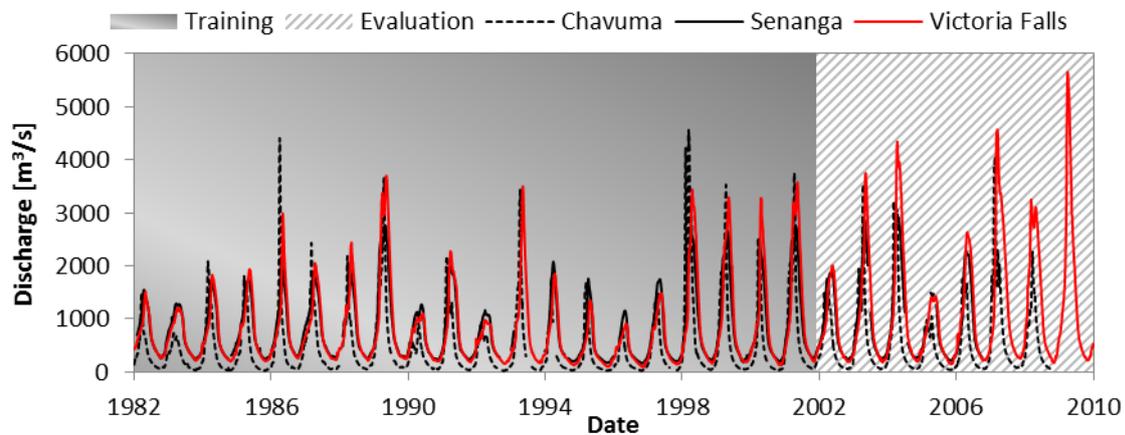
Figure 6.11. Discharge data series retrieved for the Upper Zambezi River along with training and evaluation periods.

Also presented in Figure 6.12 is the adjustment of a Fourier series with four terms to the observed discharges at Victoria Falls (the adjustment was undetaken for the training period, from 1982 to 2001).
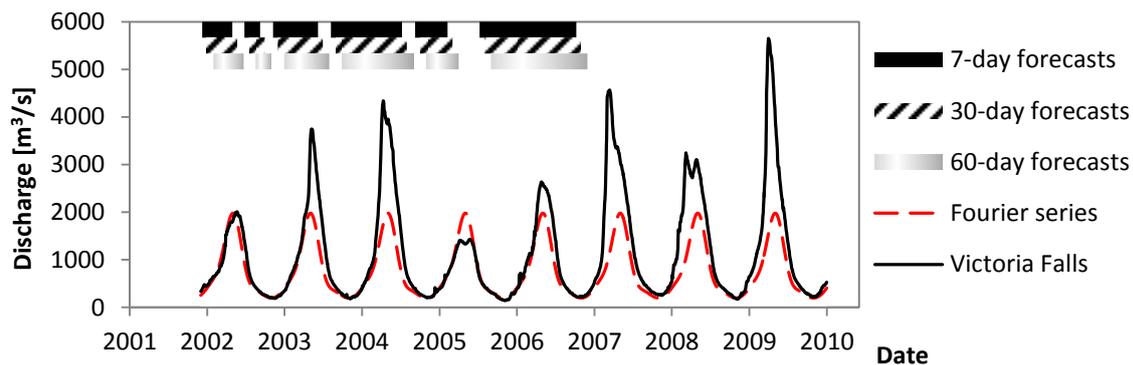
Figure 6.12. Time frames used for model comparison (indicated on top) and Fourier series adjustment to the observed steam flows at Victoria Falls.

Each alternative method has different characteristics and requirements. Regarding ARMAX, the selection of the p and q parameters was carried out by trial and error, choosing the model with the least Akaike information criterion value. Computations were accomplished using the *R* software *forecast* package (Hyndman et al. 2013).

MLP models (tested with 2 and 10 hidden nodes: MLP_2 and MLP_{10}, respectively) were evaluated through the mean value of ensembles with 600 members after removal of the worst candidates (candidates producing negative streamflow forecasts and very poorly adjusted to the Fourier series).

In SVR and LS-SVR the choice of kernels is subjective and hyperparameters must be optimized. Linear and Radial Basis Function (RBF) kernels were tested. Optimizations were carried out using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) model (Auger and Hansen 2005, Hansen 2006, García et al. 2009, Hansen 2010), a random search evolutionary algorithm adequate for real-parameter optimization of non-linear, non-convex functions, in which the candidate solutions are sampled according to a multivariate normal distribution.

Results are presented from Table 6.1 to Table 6.3, organized in respect to forecast lead time. The best three results regarding each error measure are highlighted in bold red font.

As a general trend, it can be seen that there is a degradation of all measures when lead time increases. While this would be expected, this degradation is quite limited in light of the long forecasting lead times, which attests to the usefulness of the techniques but, more importantly, reflects that the hydrographs generated in the Upper Zambezi are very smooth; a consequence of the basin's large area and the presence of large wetlands.

For a lead time of 7 days, forecasts are nearly perfect, with NE's over 0.99. Although the FRD input set appears to produce the best performances, comparisons between the different models are of little significance given the proximity among results.

For 30 days lead time differences are more relevant (Table 6.2). Again, the FRD input is beneficial. In fact, the choice of input set seems to be more important than the choice of model. In parallel, the simpler models (ARMAX, LS-SVR with linear kernel and MLP_2) achieve the best forecasts. A final observation is important regarding this lead time: simpler models do not appear to make use of the rainfall information as, when results for F and FR input sets are analyzed, it can be seen that differences are more noticeable for MLP_{10}, SVR RBF and LS-SVR RBF than for ARMAX, MLP_2, SVR linear or LS-SVR linear.

Table 6.1. Evaluation of the model's performances for a lead time of 7 days during the evaluation period. F, FR and FRD inputs considered.

		NE	PKR	VR	RMSE	MAE
ARMAX	(F)	0.991	**1.001**	0.989	93.7	44.4
	(FR)	0.991	1.008	0.989	93.5	44.3
	(FRD)	0.991	0.995	0.989	94.0	41.8
MLP_2	(F)	0.985	0.971	0.980	121.2	56.0
	(FR)	0.985	0.976	0.981	120.0	55.9
	(FRD)	**0.993**	**0.999**	0.990	**80.0**	44.6
MLP_{10}	(F)	0.977	0.947	0.975	146.0	62.8
	(FR)	0.979	0.957	0.978	140.8	62.2
	(FRD)	0.991	1.009	**0.993**	90.2	45.1
SVR RBF	(F)	0.982	0.947	0.983	131.8	55.8
	(FR)	0.977	0.939	0.975	151.1	61.2
	(FRD)	**0.994**	1.007	**0.997**	**73.8**	**38.2**
SVR Linear	(F)	0.987	0.995	0.980	113.3	52.3
	(FR)	0.987	0.998	0.982	112.7	51.9
	(FRD)	0.992	0.985	0.984	88.1	**40.6**
LS-SVR RBF	(F)	0.988	0.976	0.980	107.9	48.3
	(FR)	0.989	0.980	0.985	106.0	46.4
	(FRD)	0.990	1.002	0.984	98.5	44.3
LS-SVR Linear	(F)	0.988	**1.002**	0.988	108.8	49.3
	(FR)	0.988	1.002	0.988	108.6	49.0
	(FRD)	**0.994**	1.004	**0.994**	**79.7**	**40.9**

Table 6.2. Evaluation of the model's performances for a lead time of 30 days during the evaluation period. F, FR and FRD inputs considered.

		NE	PKR	VR	RMSE	MAE
ARMAX	(F)	0.881	0.785	0.916	346.0	158.8
	(FR)	0.881	0.741	0.928	346.0	159.3
	(FRD)	**0.932**	0.846	0.920	**261.5**	**136.0**
MLP_2	(F)	0.892	0.821	0.934	328.6	165.5
	(FR)	0.898	0.845	**0.938**	318.4	163.9
	(FRD)	**0.933**	0.886	0.920	**258.7**	143.6
MLP_{10}	(F)	0.878	0.804	0.923	349.2	175.7
	(FR)	0.892	0.837	0.936	327.6	170.1
	(FRD)	0.930	**0.910**	0.925	264.4	150.2
SVR RBF	(F)	0.733	0.712	0.854	519.6	241.4
	(FR)	0.834	0.715	0.868	409.8	182.6
	(FRD)	0.917	0.826	0.898	289.9	145.4
SVR Linear	(F)	0.847	0.726	0.879	393.1	182.6
	(FR)	0.839	0.721	0.872	403.6	184.9
	(FRD)	0.922	0.839	0.903	280.5	148.5
LS-SVR RBF	(F)	0.886	0.811	0.897	339.7	155.5
	(FR)	0.914	0.853	**0.938**	294.6	**142.7**
	(FRD)	0.931	0.868	0.911	263.6	142.7
LS-SVR Linear	(F)	0.879	0.780	0.932	350.1	159.1
	(FR)	0.886	0.809	**0.939**	338.7	160.2
	(FRD)	**0.933**	0.868	0.917	**260.9**	141.3

The error measures for the lead time of 60 days are presented in Table 6.3. Results are remarkable as LS-SVR RBF provides clearly the best alternative from every angle. Furthermore, unlike for shorter lead times, it is not the FRD input set that leads to best fitting models, but FR. It is believed that this is related to the physical properties of the Upper Zambezi. From Figure 6.9 it can be seen that water precipitating on the upstream parts of the catchment is expected to take more than 60 days to reach Victoria Falls. On the contrary, flows passing at Chavuma Falls and Senanga have a time of travel down to Victoria Falls of less than 30 days and, therefore, contain very little information that can be used 60 days into the future. Because parsimony affects model performance, the FRD set compares negatively to FR, which is less likely to contain "noise" in this case.

As before, not all models took advantage of the added information in FR. In fact, rainfall information seems to be better processed by the non-linear implementations (MLP_2, MLP_{10} and LS-SVR RBF, with the exception of SVR RBF).

Table 6.3. Evaluation of the model's performances for a lead time of 60 days during the evaluation period. F, FR and FRD inputs considered.

		NE	PKR	VR	RMSE	MAE
ARMAX	(F)	0.748	0.799	0.835	501.5	247.6
	(FR)	0.747	0.756	0.835	502.7	248.0
	(FRD)	0.674	0.665	0.831	570.2	314.0
MLP_2	(F)	0.745	0.779	0.840	501.9	267.9
	(FR)	0.760	0.788	0.854	487.4	255.6
	(FRD)	0.760	0.761	0.857	487.1	251.6
MLP_{10}	(F)	0.784	0.825	0.856	462.0	248.7
	(FR)	0.791	**0.837**	**0.893**	454.7	250.3
	(FRD)	0.751	0.784	0.873	496.7	267.9
SVR RBF	(F)	0.736	0.720	0.807	513.6	255.9
	(FR)	0.715	0.799	0.833	533.1	255.4
	(FRD)	0.715	0.749	0.839	533.0	271.3
SVR Linear	(F)	0.643	0.678	0.759	596.5	303.4
	(FR)	0.644	0.669	0.769	595.7	301.6
	(FRD)	0.676	0.631	0.769	568.3	276.7
LS-SVR RBF	(F)	**0.800**	**0.850**	0.856	**446.4**	233.3
	(FR)	**0.818**	0.841	**0.916**	**426.3**	**225.4**
	(FRD)	**0.797**	0.802	**0.893**	**450.5**	**233.1**
LS-SVR Linear	(F)	0.739	0.777	0.844	510.7	270.4
	(FR)	0.738	0.754	0.829	510.9	265.6
	(FRD)	0.776	0.750	0.857	472.3	**227.0**

As discussed earlier, regardless of their average performance, MLP ensembles obtained by backpropagation do not represent the forecast probability distribution. An example is given in Figure 6.13, where the ensemble of MLP_{10} models (represented within its 0.05 – 0.95 quantiles can be compared with observed discharges.
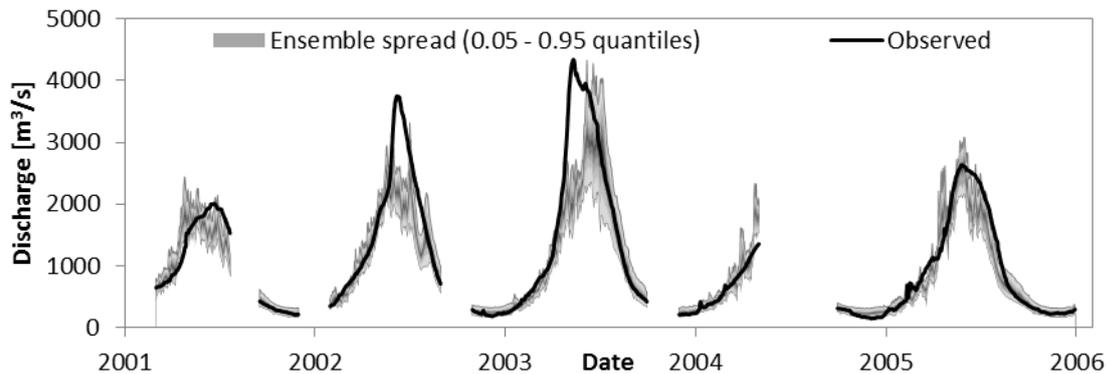
Figure 6.13. Illustration of the dispersion of a MLP_{10} ensemble for the 60-day lead time forecast.

In a more quantitative assessment (Figure 6.14), the inability of the ensemble to capture high discharges is evidenced when observed empirical exceedance probability is compared with simulated empirical exceedance probability (a comparison analogous to that of a Q-Q plot).

A global interpretation of the results can be that, as lead times increase and the forecasting problem becomes more complex, non-linear models gain advantage over linear ones. Also, rainfall information from the upper parts of the catchment gains importance as a predictive variable.

Finally, in works on deterministic forecast it is common that the models' capabilities are assessed through an absolute analysis of the error measures. As examples, one could argue that a positive NE coefficient or, alternatively, that a NE superior to that of a naïve model (which would take the last observation as the best forecast) would indicate some forecasting skill.
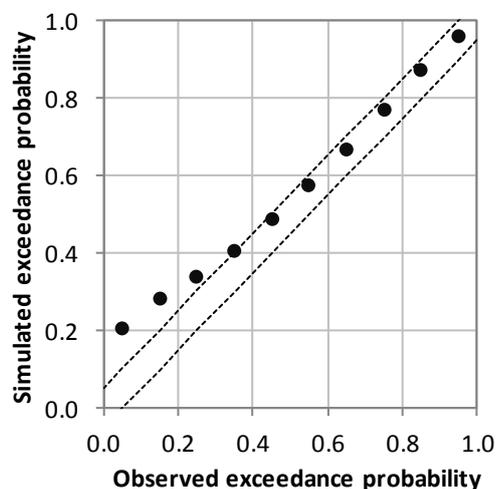
Figure 6.14. Comparison between observed and simulated exceedance probabilities for an ensemble of MLP$_{10}$ models trained by backpropagation. Ensemble for 60-day lead time forecasts. The dashed lines represent the bounds of the rejection region of the Kolmogorov-Smirnov test at a 0.95 CI.

It is expected that events sufficiently afar in time are mutually independent (this meaning that for such lead times any forecasting effort is irrelevant). During the execution of the simulations which led to the earlier results, lead times of 90 and 120 were also tested (not shown here). For these longer lead times, it was observed that performances tended to decrease, approaching the values obtained by using the fitted Fourier series as a predictor. In fact, because the Fourier series is periodical, it is completely independent of lead time – it provides a "no forecast" baseline.

Because a Fourier series represents a skill threshold below which there is no advantage in applying forecasting models, in future works concerning long lead times there would be an advantage if such models were benchmarked against it. As an example, the performance of forecasts made based on the fit of a Fourier series at Victoria Falls are presented in Table 6.4. If stationarity is assumed, an impressive minimum NE coefficient can still be guaranteed in predictions over the evaluation period, regardless of the chosen lead time.

Table 6.4. Fourier series benchmark.

	NE	PKR	VR	RMSE	MAE
Fourier	0.656	0.690	0.743	589.5	297.0

6.5 Limitations of support vector machine regression discharge forecasting

From earlier works (enumerated in Section 6.3.2) and based on the previous results, it can be inferred that SVR and LS-SVR models show remarkable aptitude regarding the deterministic forecasting of streamflow.

Despite these earlier findings, in the light of SVR theory, the reliability of such models for forecasting extreme flood events raises concerns. In Section 6.3.2 it has been shown how SVR models, by virtue of their ε-insensitive loss function, can underestimate streamflow volumes. Additionally, the results of the combination of the structural risk minimization principle implemented in SVR and LS-SVR

models and RBF kernels have been illustrated for simple examples; when used to forecast extreme flood events this combination can have profound impacts.

The analysis presented below goes over these impacts. For simplicity, only the lead time of 30 days, the ARMAX, SVR and LS-SVR models, and the F input set (comprised of discharge observations at Victoria Falls and corresponding Fourier series) were evaluated. As discharges from upstream are no longer included in the analysis, the earlier problem with data gaps is not present. As such, the training and test/evaluation periods were changed in order to encompass all the period from 1982 to 2010.

To assess the forecast capabilities of SVR and LS-SVR during extreme flood events, an additional input set was prepared. This constitutes a "reduced discharge" scenario in which all the observations of flows above 2500 m³/s were omitted from the training input set, guaranteeing that some events in the evaluation period were out of the training domain (Figure 6.15).
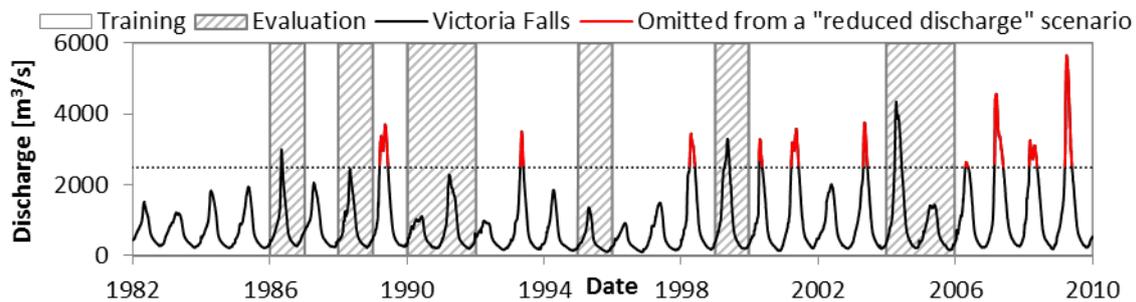
Figure 6.15. Historical daily discharges at Victoria Falls. Definition of training and evaluation periods and identification of observations omitted from the "reduced discharge" scenario.

Results for the base scenario, from which no data was omitted, agree reasonably well with the ones presented in Section 6.4 (Table 6.5). Again, the best results for each error measure are highlighted in bold red font.

Table 6.5. Results of a standard model comparison for a 30-day lead time forecast at Victoria Falls. The F input set was considered.

	Training				Evaluation			
	NE	**VR**	**RMSE**	**MAE**	**NE**	**VR**	**RMSE**	**MAE**
ARMAX	0.841	0.992	361.8	172.8	0.827	0.926	326.5	165.1
SVR Lin.	0.825	0.951	379.6	170.9	0.854	**0.980**	299.8	156.6
SVR RBF	**0.881**	0.963	**313.4**	**132.2**	**0.870**	0.973	**282.7**	**144.9**
LS-SVR Lin.	0.833	**1.000**	370.8	178.0	0.855	1.035	299.0	167.3
LS-SVR RBF	**0.897**	**1.000**	**291.4**	**142.1**	**0.879**	1.012	**273.1**	**151.7**

When the reduced dataset is used, however, results for the evaluation period change radically, as can be seen from an inspection of Table 6.6. Now, SVR RBF and, particularly, LS-SVR RBF show acknowledgedly low skills.

Table 6.6. Results of a "reduced discharge" scenario for a 30-day lead time forecast at Victoria Falls (evaluation period). The F input set was considered.

	NE	**VR**	**RMSE**	**MAE**
ARMAX	0.832	0.936	321.9	**162.8**
SVR Lin.	**0.843**	**0.951**	**310.6**	**159.2**
SVR RBF	0.775	0.931	372.2	170.6
LS-SVR Lin.	**0.839**	**0.980**	**314.8**	165.9
LS-SVR RBF	0.685	0.882	440.5	192.1

The flood volume underestimation effect expected in SVR forecasts is difficult to recognize for RBF kernels due to the strong influence of the structural risk minimization principle. They are, however, identifiable for the linear kernels, being that the VR measure for the linear SVR's forecasts is considerably lower than for the linear LS-SVR's.

Looking at the produced 30-day lead time forecasts, one of the hazards of the structural risk minimization principle at work in discharge forecasting applications is evidenced. Considering SVR RBF results, reproduced in Figure 6.16, and particularly in the years of 1999 and 2004, predictions drop abruptly as lagged observations fed into the model exceed 2 500 m³/s.
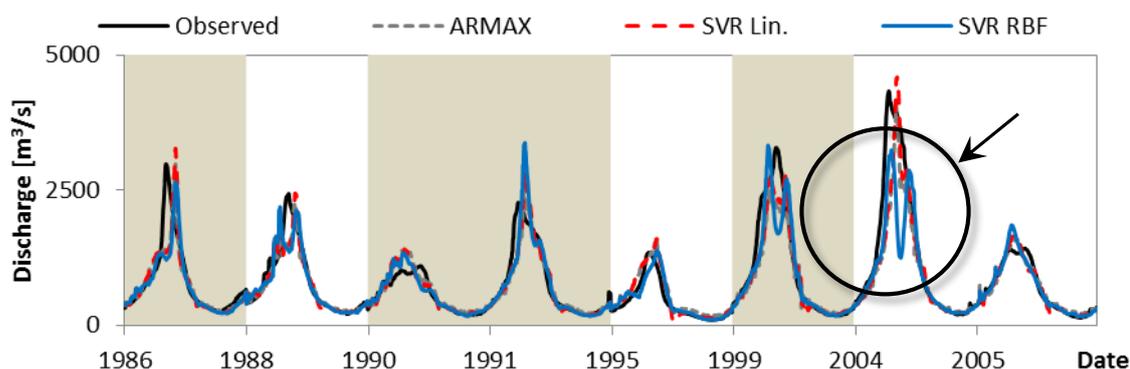
Figure 6.16. Observed and forecasted hydrographs for a 30-day lead time at Victoria Falls using a "reduced discharge" subset for training. SVR linear and SVR RBF models (discontinuous date axis).

Depicted in Figure 6.17, the effect is even worsened for LS-SVR RBF whose predictions for the peak flood of 2004 drop well below expected yearly flood maxima. As hinted by Table 6.6's results, linear models are not affected by similar effects.
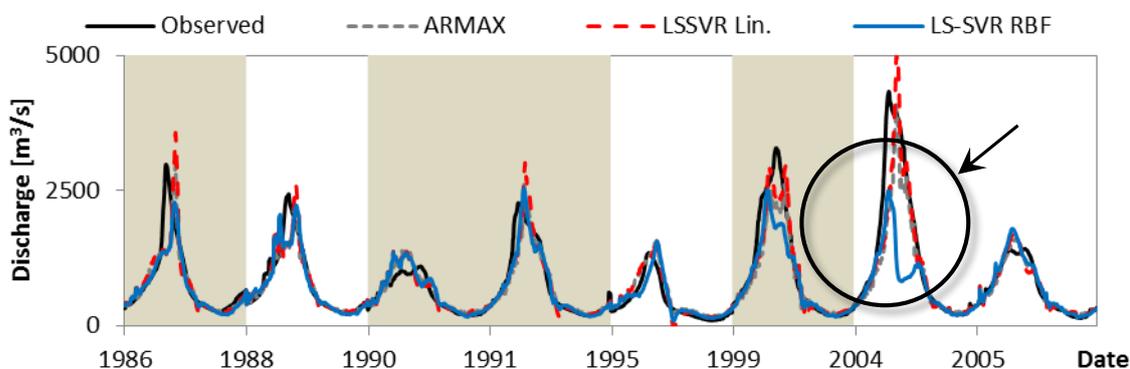
Figure 6.17. Observed and forecasted hydrographs for a 30-day lead time at Victoria Falls using a "reduced discharge" subset for training. LS-SVR linear and LS-SVR RBF models (discontinuous date axis).

The present analysis puts forward that due to their inbuilt structural risk minimization principle, RBF SVR and LS-SVR are theoretically prone to predict unexceptional amplitude discharges in the face of rare peak floods. The effect yields obvious and serious risks and should be clearly addressed in future implementations. Related to this claim, it can be argued that no ANN or SVR model should be employed beyond its training domain. This is, naturally, a sensible preposition. The case is, however, that discharge forecasting problems are posed in multi-dimensional input domains (in the present case with lagged discharges and Fourier series approximations, but easily augmented by including additional inputs such as rainfall and evaporation estimates or upstream lagged discharges). Finding whether one is within the "safe" training domain is simply not straightforward and, in practice, hardly ever done. Additionally, because high discharges are rare and often the ones more interesting to forecast, there is a strong driver towards extending the valid domain of the models. Models such as most ANNs become unpredictable under these conditions – which is

desirable from a certain point of view as any analysis of an ensemble of trained ANNs will probably show a marked increase in prediction dispersion, providing an implicit warning of the likely errors (Figure 6.13 is an example of this, with a wider ensemble in corresponding to higher discharges). RBF SVR and LS-SVR's responses are, on the contrary and as shown, quite predictable – on the underestimation side – and thus deserve particular caution.

The ε-insensitive loss function minimized in SVR (not LS-SVR) models was also analyzed. Being related to the MAE, the ε-insensitive loss function is a distinctive feature of SVR models as it grants them both robustness to outliers and increased computational performance. Under the safe assumption that the regression is not perfect, while minimizing the MSE tends to produce models whose simulations' mean values match the mean values of the observations, minimizing the MAE does not. In the case of yearly discharge series, the high likelihood of lower-than-average observations frequently means that MAE minimization will result in forecasting models whose integrated discharges underestimate true flood volumes. In many discharge forecast applications, such as reservoir management, a tendency towards systematically underestimating flood volumes can also entail clear risks.

6.6 Capturing uncertainty: the development of empirical forecast ensembles

All the previous methods provide deterministic forecasts without any information about the uncertainty associated with the forecast. Even the MLP backpropagated ensembles, which come closest to a "soft" characterization of forecast uncertainty, seem to be ultimately unable to accomplish a satisfactory result (Figure 6.13 and Figure 6.14).

Implementing the ideas laid out in Section 6.3.5, one can optimize a whole ensemble of MLPs simultaneously in order to obtain a correct distribution of the probability of forecasts being accurate. The touchstone of the approach is to select appropriate conflicting objectives and characterizing the Pareto surface spanned by their realizations. Three objectives were applied:

- a measure of member deterministic accuracy – $\mathrm{Log}_{10}\left(\mathrm{MSE}\right)$ [8];
- a measure of the empirical exceedance probability associated with each member's forecasts – EF;
- and a measure of the complexity of each member model (the standard deviation of the connection weights)[9].

As indicated previously (see this chapter's footnote 7, in page 115), two separate optimizations must be carried out: one for exceedance probabilities lower than 50% and the other for exceedance probabilities of over 50%. Additionally, separate ensembles were trained for 12 windows of three months so as to avoid that exceedances during the high flow months could be cancelled out by non-exceedances during the low flow months.

A Multi-Algorithm Genetically Adaptive Multi-objective method (AMALGAM) (Vrugt and Robinson 2007, Vrugt et al. 2009a) was employed in order to characterize the desired Pareto surface by optimizing MLP connection weights[10]. AMALGAM is a meta-optimization algorithm in the sense that it oversees the execution of four distinct multi-optimization algorithms: non-dominated sorted genetic algorithm-II (NSGA-II) (Deb et al. 2002b), particle swarm optimization (PSO) (Kennedy and Eberhart 1995), adaptive Metropolis search (AMS) (Haario et al. 2001), and differential evolution (DE) (Storn and Price 1997). Because AMALGAM can harness the strengths of any of the four algorithms that run in parallel, it compares favorably to other multi-objective optimization algorithms (e.g. Zhang et al. 2010, Zhang et al. 2011).

---

[8] The logarithm was simply used in order to make the problem more tractable for the optimization algorithm.
[9] Other measures of model complexity could be used, having this one been chosen for its simplicity.
[10] Although AMALGAM was chosen due to its reported performance, many other multi-objective optimization algorithms could be applied.

An example of a Pareto surface obtained for the 3-D space spanned by the objective functions is shown in Figure 6.18. On the top right corner, a 3-D view of the ensemble members' results is illustrated. Underneath, a matrix plot unveils 2-D projections of the surface where the trade-offs between $\mathrm{Log}_{10}(\mathrm{MSE})$, EF, and the regularization measure are evident.
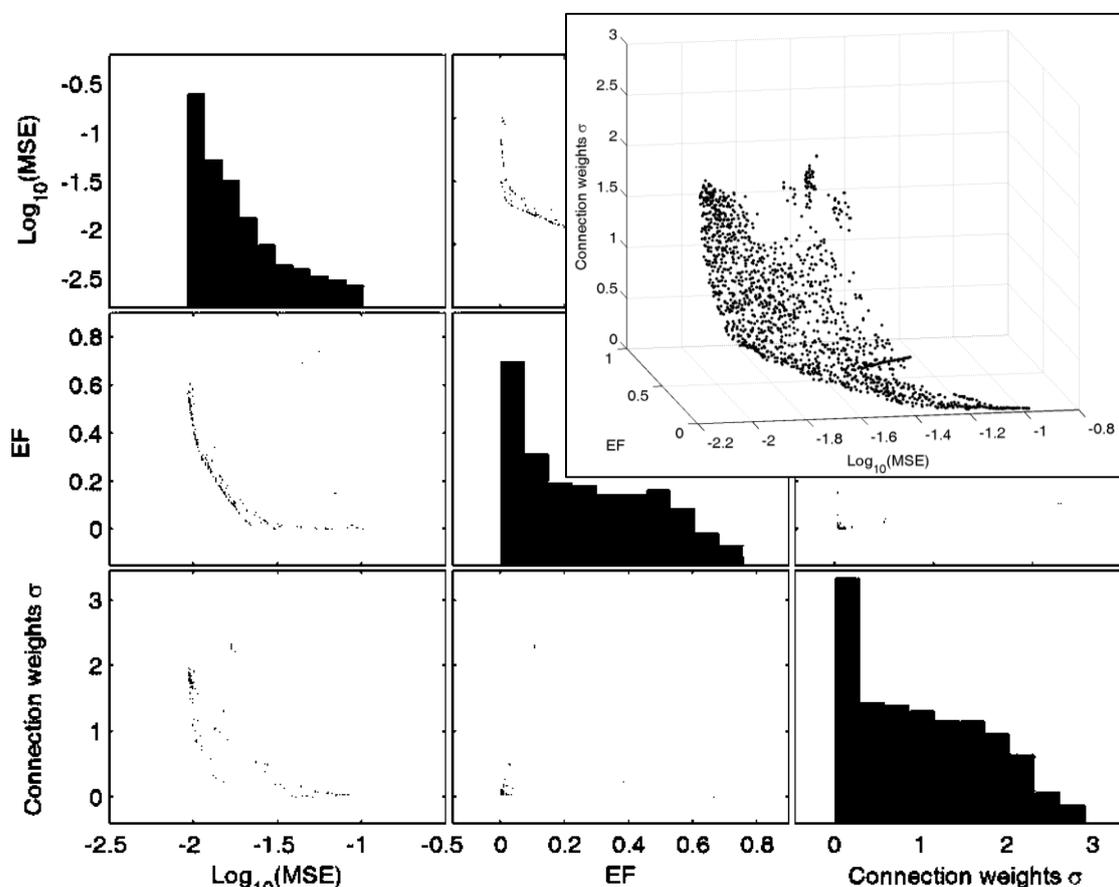
Figure 6.18. Representation of a Pareto surface obtained for exceedance probabilities of 0 to 50% during the months of April, May and June.

Because the aim of the multi-objective model is to approximate the Pareto surface, it is not the trained ensemble itself that characterizes forecast probabilities. In order to obtain such probabilities, the Pareto surface must be sliced in EF bands which correspond to several probability thresholds. Then, the observations falling within each band can be averaged so as to get a forecast associated with any desired probability.

The concept was applied to a 30-day s lead time forecast for Victoria Falls using the F input set. The training and evaluation periods are identified in Figure 6.15. The forecasts for the training period are illustrated in Figure 6.19. The forecast probability is characterized by 11 non-exceedance thresholds (2.5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 97.5%). Because the date axis is discontinuous, light brown shaded areas on the background identify the contiguous dates.

Observing the adjustment of the forecast, it can be seen that, as desirable, the lowest exceedance probability bands are hardly ever crossed. Furthermore, it is interesting to notice that the ensemble spread varies greatly according to flow magnitude and period of the year.

Putting the results into perspective, Figure 6.20 compares observed and simulated exceedance probabilities. Optimally, results should follow a straight line from (0,0) to (1,1), which is the case even for the very hard to model low exceedance probabilities. Kolmogorov-Smirnov test error bounds were added to the plot, indicating that, at a 95% confidence interval, the hypothesis that the simulated probabilities and the observed ones share the same distribution cannot be rejected.
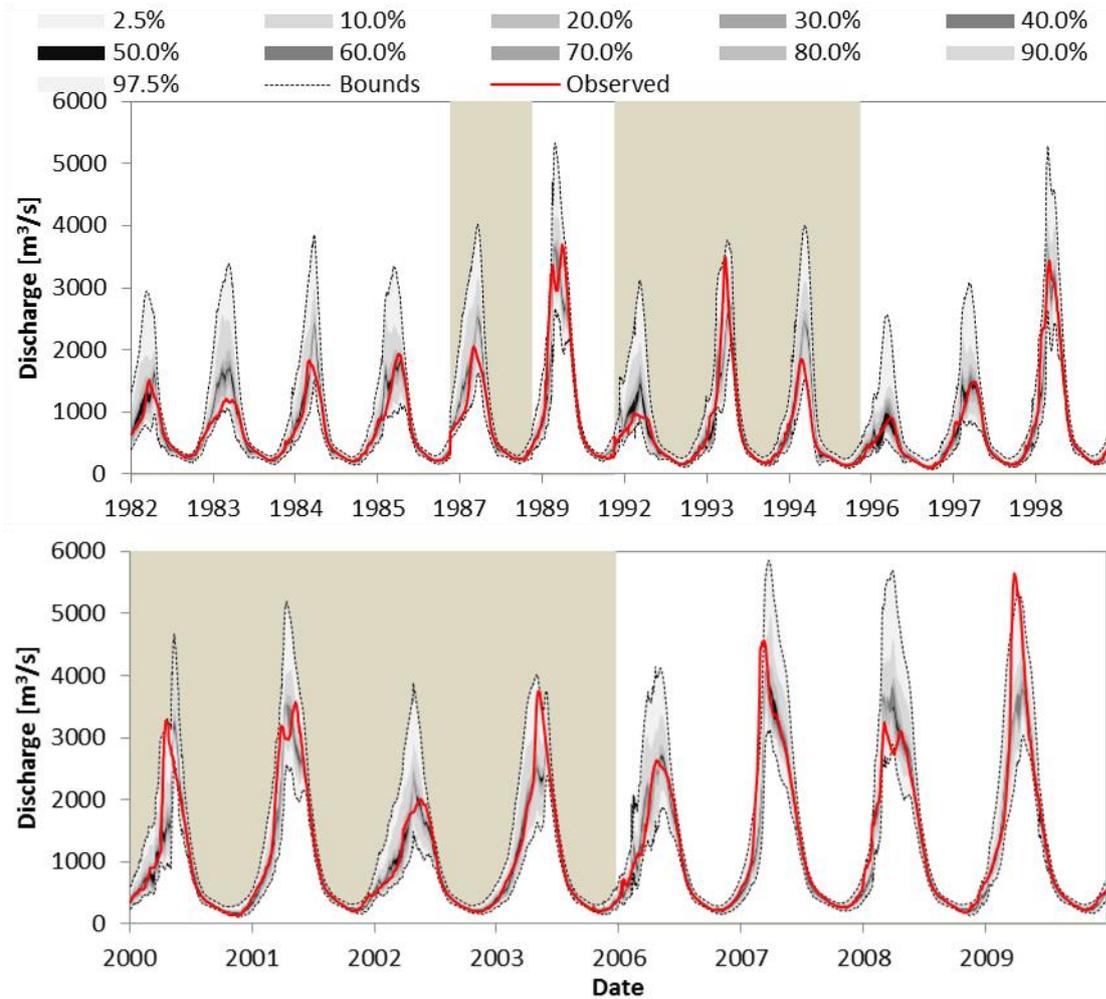
Figure 6.19. Probabilistic forecasts for a 30-day lead time at Victoria Falls. Training period.

Dealing with machine learning models and MLPs in particular it is highly advisable to evaluate results in an independent evaluation period. In that line, the adjustment of the forecast for the test period is presented in Figure 6.21. Again, the probability bands seem to agree well with observations.
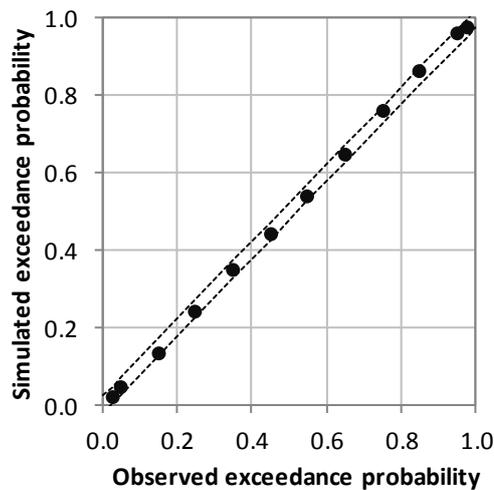
Figure 6.20. Comparison between observed and simulated exceedance probabilities for a probabilistically correct MLP ensemble. 30-day lead time forecasts. Training period. The dashed lines represent the bounds of the rejection region of the Kolmogorov-Smirnov test at a 0.95 CI.
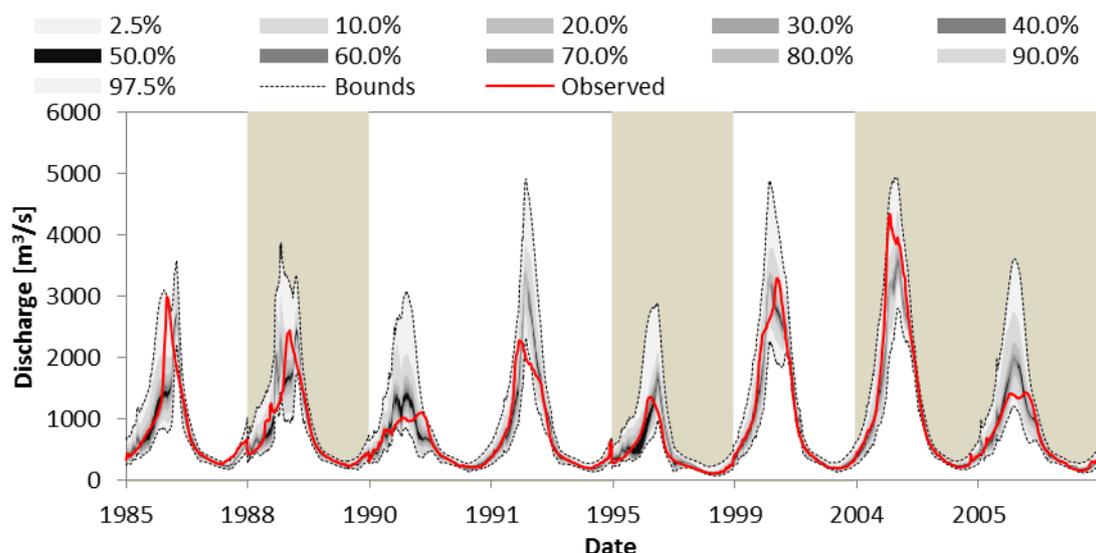
Figure 6.21. Probabilistic forecasts for a 30-day lead time at Victoria Falls. Evaluation period.

The evaluation period observed-simulated exceedance plot (Figure 6.22) further supports the validity of the methodology. Again, the hypothesis that both distributions are the same cannot be discarded at a confidence interval (CI) of 95%, according to the Kolmogorov-Smirnov test.
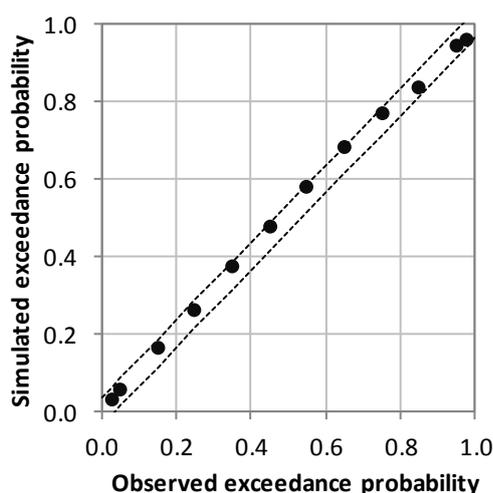
Figure 6.22. Comparison between observed and simulated exceedance probabilities for a probabilistically correct MLP ensemble. 30-day lead time forecasts. Evaluation period. The dashed lines represent the bounds of the rejection region of the Kolmogorov-Smirnov test at a 0.95 CI.

Finally, as one of the advantages of this method is the absence of theoretical assumptions, it is interesting to observe how the ensemble spread evolves. In fact, the ensemble spread can be associated with the heteroscedasticity of the error – a condition that must be explicitly addressed when using formal Bayesian inference methods to characterize forecast uncertainty.

In Figure 6.23 the ensemble spread is plotted against the magnitude of the observed discharge for rising and receding flow periods. As can be seen, although there is a general trend of higher spreads for higher observed discharges, spreads in the receding flow periods are generally smaller than during rising flows – something overlooked even in complex BI frameworks (e.g. Schoups and Vrugt 2010).
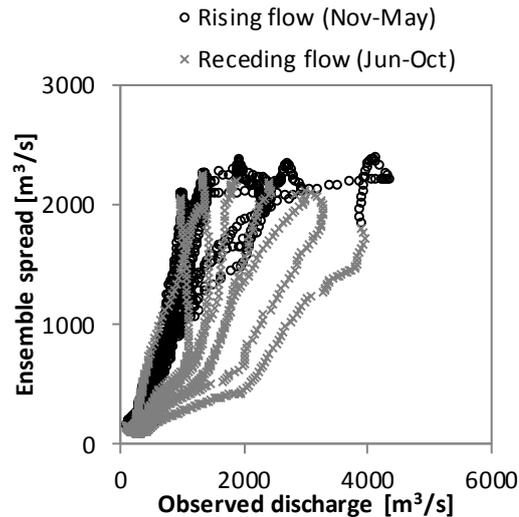
Figure 6.23. Ensemble spread vs. observed discharge for rising and receding flow periods.

6.7 Conclusions

There are several conclusions worth summarizing in the wake of the previous analyses. In first place, the newly proposed scheme for rainfall aggregation, based on hydraulic considerations and easily implemented resorting to remote sensing data was able to enhance forecasting results, particularly for a 60-day lead time. In fact, it has been observed that the choice of an appropriate input set can have more weight on performance than the choice of the forecasting model itself.

Addressing the comparison of alternative deterministic forecast models, it was remarked that, as lead time increases, so do the advantages of applying more complex algorithms. In parallel, the non-linear approaches appear to benefit from rainfall information while linear models such as ARMAX do not.

RBF SVR and LS-SVR models outstood in the comparative analysis, especially regarding the 60-day lead time. In a closer assessment of both models, it has been shown that they can lead to gross underestimations of flood peaks and volumes. Despite this, and although apparently less performing under non-extreme conditions, linear SVR and LS-SVR can also surpass traditional ARMAX models while not yielding the same risks of expected peak flood underestimation, thus possibly consisting on a safer and promising alternative.

A novel method for developing empirical forecast probability distributions was developed. Being of simple and fast application, requiring no prior assumptions and performing remarkably well, this method holds great promise. In its proposed implementation MLPs were used. While this grants flexibility to the whole framework – optimization is fast, being that lead times and inputs can be easily changed – other types of ensemble members could be considered such as RBFNs or even SWAT models.

Because the probabilities depicted by the ensemble are not the probabilities of a certain occurrence taking place, but rather the probabilities of a certain forecast being observed in practice, they are comprised of the aleatoric uncertainty associated with the natural phenomena and the epistemic uncertainty which can be linked to imperfections of the forecast. By this token, a better forecasting system would produce better resolution empirical probability distribution functions. Using this criterion, different input sets and ensemble member types can be easily compared.

7 PREPARATION OF THE HYDROLOGICAL MODEL AND DEVELOPMENT OF THE CALIBRATION INTERFACE

7.1 Introduction

This is the first of three chapters (besides the present one, Chapters 8 and 9) which are specially related to the Soil and Water Assessment Tool (SWAT) hydrological model. SWAT is semi-distributed and most of its equations are based on considerations about the physical nature of the modeled processes. It constitutes a very broad-reaching model with a strong water quality component which was not explored in this work. Details on SWAT's features, theory and applications can be found in Chapter 3 and will not be reproduced here.

Unlike the data-driven models developed and tested in Chapter 6, SWAT's application and, it can be argued, the use of traditional physically-based or even conceptual hydrological models in general, entails burdensome phases of data pre-processing and model preparation, calibration and validation.

Owing to its research character, the present work went beyond the normal application of the SWAT model, namely by:

- analyzing broad parts of SWAT's source code;
- changing that code where needed;
- developing an external highly flexible calibration interface; and
- investing in enhanced calculation performance.

The background work described below consumed the lion's share of the author's time and albeit the following discussion does not constitute a coherent and driven scientific contribution, it provides a base for better understanding the conducted research in its goals and limitations. Most importantly, it is hoped that it can be a guide to future SWAT users and provide some contribution to its developers.

The contents are distributed through four sections. Section 7.2 focuses on the data pre-processing strategy and on model preparation using the ArcSWAT interface (Winchell et al. 2010)[1]. A brief summary of input data preparation is given in Section 7.3. The main changes to SWAT's source code are discussed in Section 7.4. In Section 7.5 the development of a flexible calibration interface in Matlab is described. Finally, recommendations for future improvements are made in Section 7.6.

---

[1] ArcSWAT is an input interface for SWAT, integrating it with the ArcVIEW GIS software and greatly facilitating model preparation. It can be found at http://swat.tamu.edu/software/arcswat/.

7.2   Model preparation

Model preparation was done through ArcSWAT following a standard procedure which is, for the most part, not described here. The data used during model preparation (refer to Chapter 4 for further information) was:

- the HydroSHEDS DEM (for the derivation of the river network) (Lehner et al. 2008);
- the SRTM DEM version 4.1 (for the calculation of the subbasin's physical parameters) (Jarvis et al. 2008);
- the Global Land Cover Characterization version 2 (Loveland et al. 2000); and
- the Digital Soil Map of the World and Derived Soil Properties version 3.5 (Food and Agriculture Organization of the United Nations (FAO) 1995).

Two distinct SWAT models were set up. One as parsimonious as possible, with a small number of computational subbasins and Hydrologic Response Units (HRU); another, more detailed, constraining the subbasin's area to less than 5000 km^2 and adopting a number of HRUs "adequate" to represent the main land cover and soil types. The synthesis numbers characterizing both models are indicated in Table 7.1:

Table 7.1. Main numbers characterizing the parsimonious and detailed SWAT models.

	Parsimonious model	Detailed model
Number of subbasins	53	325
Number of HRUs	145	1198
Slope classes[2]	[0, 0.5[; [0.5, 2[; [2, inf[	[0, 0.5[; [0.5, 2[; [2, inf[
Soil type threshold[3]	35%	25%
Land use threshold[3]	30%	25%
Slope threshold[3]	20%	25%

The distribution of computational subbasins in both models is illustrated in Figure 7.1, for the parsimonious model, and Figure 7.2, for the detailed one. Two larger illustrations, including information about calibration points and considered waterbodies, can be found in Appendix I.D and in Appendix I.E.
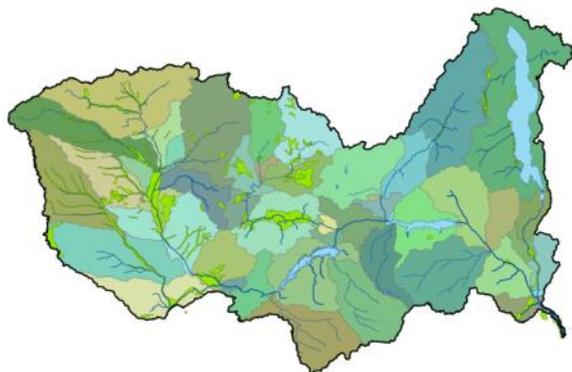
Figure 7.1. Subbasins considered in the parsimonious model.
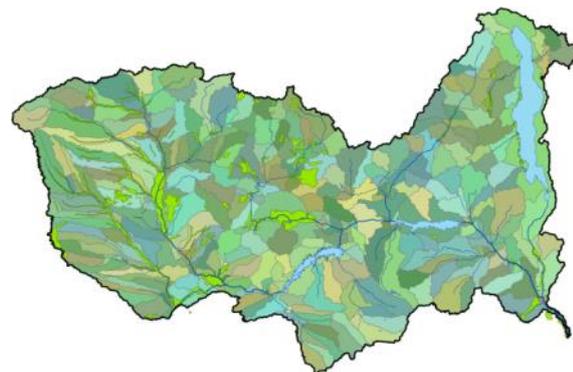
Figure 7.2. Subbasins considered in the detailed model.

As mentioned, the set up process was quite standard, one particularity being the way in which waterbodies were represented. During model set up, ArcSWAT assumes that there is a maximum of one waterbody, one reach and one junction per subbasin. Additionally, waterbodies are overlaid

---

[2] Slope is calculated for every pixel in the basin. The slope classes are used to aggregate the pixels into a restricted number of representative groups.

[3] HRUs characterize areas thought to have similar hydrologic response; in the case of ArcSWAT according to a combination of soil type, land use and slope class. To improve the model's computational performance, thresholds can be chosen in order to dismiss HRUs in minority within each subbasin.

to the subbasin where they stand, being that area doubled. This poses a problem when representing large waterbodies.

Indeed, it is normal that several streams with a relevant drainage area converge into a large waterbody. Because of the limitation of one reach and one junction per subbasin, however, in order to simulate the streams that converge towards the waterbody, subbasins are created within that area (see the example of Figure 7.3, where multiple subbasins intersect the Kariba reservoir). The heart of the matter is that such waterbodies, spreading in reality over several subbasins, can only be associated with one in SWAT and, therefore, their contributions to "runoff" cannot ever be fully deduced from those computed for the land phase. In order to avoid this, all sizeable reaches contributing to the simulated waterbodies were linked to the waterbody's outlet (in the example of Figure 7.4, the Kariba reservoir is now totally contained in one large subbasin).
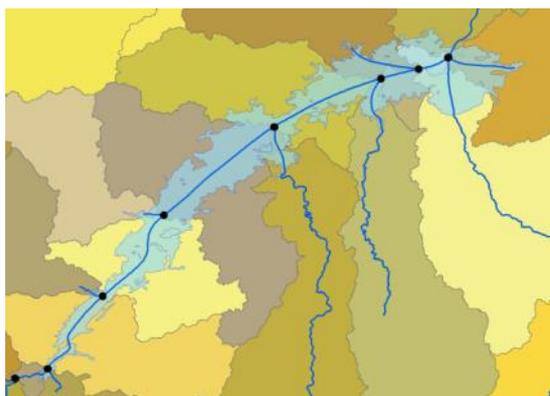
Figure 7.3. ArcSWAT default representation of the Kariba reservoir.
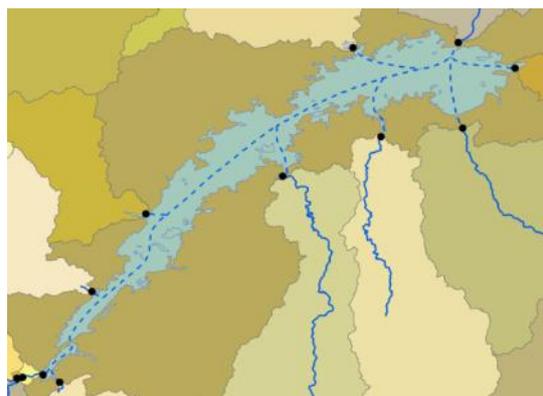
Figure 7.4. Modified representation of the Kariba reservoir (instantaneous routing in the dashed lines).

Cohen Liechti (2013) simulated the ZRB with a SWAT setup similar to the "detailed" model tested here. Following her work, without a specific representation within SWAT, wetlands located in the main reaches were modeled as uncontrolled reservoirs (Cohen Liechti et al. 2014). The equations used to simulate them, added to SWAT, are described in Section 3.2.4.

Perhaps due to the use of the conditioned HydroSHEDS DEM, but also due to necessary assumptions of the ArcSWAT software, the derived channel widths for most of the ZRB were grossly exaggerated. This led to extremely high transmission losses that, for example, dried the river between Kariba and Cahora Bassa, where observed flows average nearly 2000 m^3/s. In order to overcome this and similar undesirable effects, most channel widths had to be manually corrected resorting to satellite imagery.

7.3 Input data preparation

Input data was already introduced in Chapter 4. Here, however, it is relevant to clarify how this data is passed on to the SWAT models.

Daily areal maps of rainfall and temperature covering the period from 1979 to 2009 were used. While the temperature data was retrieved solely from the NCEP/DOE Reanalysis 2 model (Kanamitsu et al. 2002) (see Section 4.7 for details), the rainfall dataset was built upon two distinct sources. From 1979 to 1997 areal rainfall resulting from the application of the Pattern-Oriented Memory (POM) interpolation technique (described in Chapter 5) was used. From 1998 to 2009, the TMPA 3B42 version 7a satellite estimates (described in Section 4.3) were chosen.

SWAT is only prepared to make use of one daily estimate of precipitation and temperature per subbasin. As such, a bilinear interpolation scheme was employed in order to aggregate the daily data contained in the aforementioned areal maps to the shape of each subbasin.

7.4   Alterations to the Soil and Water Assessment Tool source code

As previously mentioned, SWAT's domain of applicability goes far and wide. Only a fraction of the working principles of SWAT was overviewed in Section 3.2. Despite its complexity and the fact that SWAT was been applied several times to African watersheds and, particularly, to the ZRB, some issues related to its source code had to be addressed prior to a successful output.

The main modifications to the source code are justified below. One of the latest versions of the hydrological model was used as a basis for these alterations (SWAT 2012, revision 585)[4].

The relatively large waterbodies present in the ZRB have a great impact on streamflows, either by reducing peak flows and increasing low flows, in the case of wetlands, or by arbitrarily conditioning discharges downstream, in the case of the larger dams. Without proper simulation of these waterbodies, a daily modeling of the ZRB is a sterile endeavor. In effect, most of the changes to the source code addressed below are related to waterbodies.

Definition of outflows

Uncontrolled outflows were modeled according to the formulation by Cohen Liechti et al. (2014), described in Section 3.2.4, where it is defined from equation (3.70) to (3.72). For controlled reservoirs, also discussed in Section 3.2.4, SWAT makes two main options available:

- a time series of reservoir outflows is provided; or
- outflows are estimated according to a set of simplified parameters such as emergency and principal spillway stored volumes, the definition of a flood season, monthly target reservoir storage, etc.

When complete outflow series are available no problems arise. SWAT, however, cannot cope with missing outflow values. One of the changes made to the source code was precisely to blend the two aforementioned options so that outflow time-series are used when available but the model is capable of autonomously estimating outflows when such data is missing. This estimation is based on monthly target volumes and a free parameter (see equation (3.74), page 40) characterizing the reservoir's reaction time to deviations from it.

Correction of the waterbody surface area calculation

The combined surface of the major waterbodies in the basin is such that the evaporation they promote carries substantial impact on the runoff at the Delta (around 15-20% of the annual total (Beck 2010)).

In the original SWAT source code waterbody surface is defined according to equation (7.136):

$$A_t^{rsv} = \alpha \left(V_t^{rsv} \right)^{\beta} \tag{7.136}$$

where A_t^{rsv} S_{rsv} is the water surface at day t, V_t^{rsv} stands for the water volume, and α and β are free parameters.

Their values are calculated based on two surface-volume pairs: β according to equation (7.137) and α according to equation (7.138):

$$\beta = \frac{\log_{10}\left(A_t^{rsv,1}\right) - \log_{10}\left(A_t^{rsv,2}\right)}{\log_{10}\left(V_t^{rsv,1}\right) - \log_{10}\left(V_t^{rsv,2}\right)} \tag{7.137}$$

---

[4] As this chapter is being written SWAT 2012 revision 614 is the most recent version made available to the public.

$$\alpha = \left(\frac{A_t^{rsv}}{V_t^{rsv}} \right)^{\beta}$$

(7.138)

where any of the two points can be used in the latter[5].

It is easy to confirm that equation (7.138) is not accurate and should be substituted by equation (7.139).

$$\alpha = \frac{A_t^{rsv}}{V_t^{rsv\,\beta}}$$

(7.139)

In large reservoirs, such as Cahora Bassa, the miscalculation of the α coefficient amounted to substantial deviations in surface estimates and, consequently, to direct rainfall and evaporation contributions to the stored volume. The correction was communicated by the author to SWAT's developers and is fixed since Revision 612.

Dynamic correction of subbasin area based on waterbody surface
It has already been mentioned that SWAT doubles the waterbodies' surfaces due to their overlap with subbasins. The effect is fixed only indirectly, through the definition of "water" HRUs, whose contributions to runoff are (in principle) not taken.

HRUs are defined during model set up. Their areas are therefore constant throughout the simulations. If this approximation can be considered acceptable for reservoirs and lakes when their surface is maintained approximately constant, for wetlands, which are usually shallow areas displaying marked seasonal variations in inundated surface, it is argued that it might be inadequate.

In order to overcome this issue, code was written so that, for each time step, the waterbody's surface is correctly deduced from the "active" area of its subbasin.

Influence of "water HRUs" on subbasin runoff
Before, it was stated that the contributions of "water" HRUs were (in principle) not taken for runoff computation as this is done directly at the waterbodies. This wording was chosen because it is not so and an indirect influence of such HRUs on subbasin runoff exists.

Simulating the land phase, SWAT begins by computing specific runoff for each HRU. Specific runoffs for every subbasin are then computed according to a weighted average proportional to the area of each HRU; contributions of water HRUs are naturally neglected.

In order to obtain runoff volumes, the code proceeds by multiplying the averaged specific runoff by the subbasin's area. Firstly, however, the averaged specific runoff is divided by the fraction of the subbasin's surface which is not covered in water; in other words, even if most of the subbasin is water, its runoff volume will be erroneously computed as if the small surface which is "dry" was covering all the area.

Because this approximately doubled the contributions from waterbodies and can, in special cases, induce large errors, the source code was corrected.

Introduction of volume-area curves in waterbodies
Equation (7.136), discussed above, is used by SWAT to estimate the surface of waterbodies at each time step. Assuming linearity and being based on only two points, it was considered admissible to replace it by a more solid method.

---

[5] Although this choice could be arbitrary, SWAT uses the points associated with principal and emergency spillways – designations which do not necessarily correspond to their physical counterparts.

In the proposed version of the code, a waterbody's surface can alternatively be interpolated from a user-supplied volume-area curve.

Introduction of volume-elevation curves in waterbodies

The original process of estimating water level from stored volumes is quite similar to the previous one (for surface), also relying on the assumption of linearity and a calibration using two points.

Again, this was optionally substituted by a user-defined volume-elevation curve.

Introduction of elevation-maximum discharge curves in controlled reservoirs

SWAT provides a means of specifying maximum monthly discharges from reservoirs. This limit is, however, independent from water level, which is not common in real cases.

As before, the option to take advantage of a user-defined elevation-maximum discharge curve for controlled reservoirs was introduced.

Creation of local surface runoff lag parameters

Finally, while the vast majority of SWAT's parameters are locally defined at the subbasin or, in most cases, HRU level, the surface runoff lag parameter can only be defined at the basin's scope.

This parameter, introduced in Section 3.2.3, and applied in equation (3.13), controls the amount of time that the generated surface runoff takes to reach the river network. As will be shown in Section 9.5.3 (abbreviated as SURLAG), this is one of the parameters that most influence the subbasins' hydrological responses during the rainy season and, given the size and heterogeneity of the ZRB, it is unlikely that a single value adapted to all subbasins exists.

Following this reasoning, an option of using specific surface runoff lag parameters for every subbasin was developed.

7.5 Development of a calibration interface

7.5.1 Enhancing computational performance

Computational performance is a major issue regarding semi-distributed models such as SWAT, particularly during the calibration phase. For the ZRB, after taking all possible options to speed up computations, a single simulation of the detailed model on a daily time step (comprising 31 years) takes a little over 12 minutes to be completed on a high-end desktop computer[6] running on Windows 7.

This equates to about 120 simulations per day. Nonetheless, the time spent to automatically read and evaluate the model's performance must also be taken into account. In fact, it adds substantially to the simulation time. Due to this, a more realistic estimate is that the number of simulations per day amounts to nearly 100. Putting this number into perspective, numerous single-objective optimizations of the ZRB SWAT model needed at least 20 000 simulations in order to converge to satisfying results[7]. Conservatively, this means that a single calibration run takes at best 20 days to complete. Yet, if Multi-Objective (MO) or Bayesian Inference (BI) calibration algorithms are used, the number of needed model runs is very likely to increase considerably.

SWAT does not make use of parallel computing. This, of course, means that there is a potential to decrease calibration times by running simulations in separate processor threads. The first step towards enhancing computational performance was to develop a multithreaded calibration code. It so happens that SWAT relies on intensive read/write operations to disk. With the more than 10 000 distinct files that the detailed model needs to run, disk access is a bottleneck. In fact, it was remarked that, on a Windows operating system, the use of two processors would not yield any

---

[6] Intel i7-2600 CPU @ 3.40 GHz.
[7] Considering about 14 calibration parameters and using the performing Covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm.

tangible performance increase; using more processors entailed even more concurrent read/write operations leading eventually worse performances than those attained by using a single processor.

The problem was dealt with in three ways: SWAT's source code was modified so that text outputs were substituted by binary ones; the parallel model runs were distributed among several independent disks; and, most importantly, the operating system was changed to Linux[8]. By doing so, major performance gains were attained. Instead of 12 minutes, the developed calibration interface could now, on average, change parameters, run the model, read its results, and evaluate performances at multiple locations in 40 seconds: a performance gain of more than 20 times. Under the same conditions a parsimonious model of the whole ZRB can, on average, be evaluated in less than 10 seconds.

7.5.2   Promoting flexibility

The calibration interface was developed in Matlab. The base for the code, which was mostly re-written and considerably expanded, was kindly provided by Chetan Maringanti and Théodora Cohen Liechti.

The calibration interface undertakes the necessary pre-processing of the calibration runs. Additionally, it serves as a wrapper to the SWAT model, changing its parameters as well as reading and evaluating results. Within it, specialized optimization algorithms are called. For Single-Objective (SO) calibration runs the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm (Auger and Hansen 2005, Hansen 2006, García et al. 2009, Hansen 2010) was used. For MO calibrations A Multi-Algorithm Genetically Adaptive Multi-objective method (AMALGAM) (Vrugt and Robinson 2007, Vrugt et al. 2009a) was employed[9].

Care was taken to develop a robust and flexible calibration interface that would have the ability to go further than the scope of the present research. Some traits of the code are introduced below.

Seamless control over the calibration data
An online public platform for data collection, storage and sharing was developed in the scope of this research. Appendix II.A describes this resource. Here, it is important to remark that existing discharge or water level time-series within the ZRB can be easily downloaded and, without any change to the files, automatically used by the interface for model calibration.

Calibration based on observed discharges, water levels, and volumes
Because data scarcity is a problem in the ZRB and, in particular, there is few data that can be used for model calibration in some parts of the basin, the interface makes use, not only of observed discharges, but also of water level measurements and volumes stored in waterbodies.

In order to be used, the water level measurements need to be converted to discharges. Because the rating curves are not known, they are optimized at every model run using an evolutionary strategy. Typically, the resulting information is less reliable than actual discharges. As such, their weight on the global optimization changes from one model run to another and is proportional to the quality of the adjustment achieved for the rating curve[10].

Limitless choice of fitness measures
The choice of fitness measures is quite limitless, without any need to change the interface's code. All that is needed is that the function's file, whose arguments are the observed and simulated series, is placed in a given folder. An example of code for the Nash-Sutcliffe fitness measure is illustrated:

---

[8] Due to a different approach regarding file management, SWAT appears to perform appreciably better on Linux systems. Also, the SWAT source code was compiled with the performing Intel Fortran Compiler, made available by EPFL, using optimal performance options.
[9] Both optimization algorithms are briefly described in Chapter 9.
[10] Chapter 9 contais more information on this topic.

```
function performance=nashSutcliffe(observed, simulated)
    performance=1-sum((observed-simulated).^2)/sum((observed-mean(observed)).^2);
end
```

The actual fitness values to be passed on to the optimization algorithms can be composites of the aforementioned fitness measures. Specification is done through text which is parsed by the interface, virtually enabling any transformation of the base results. An example is shown below[11]:

```
reachFitnessMetrics={'(1-nashSutcliffe)','mae*0.8+rmse*0.2'};
```

Calibration aggregating multiple points

The fitness values, calculated as described above for a given reach or reservoir, can then be aggregated over distinct calibration points. Because the relative importance of each calibration point might differ, as can the number of available records[12], a weighting scheme was adopted. Taking Ξ as the aggregated fitness value, the proposed weighting scheme can be described through equations (7.140) and (7.141):

$$\Xi = \frac{\sum_{n=1}^{N} w_n \cdot f_n(\mathbf{y}, \hat{\mathbf{y}})}{\sum_{n=1}^{N} w_n} \tag{7.140}$$

$$w_n = \left(\alpha \cdot D_{(n,\text{magnitude})} \times \beta \cdot D_{(n,\text{length})} \right)^{\gamma} \tag{7.141}$$

where N is the number of calibration points; $\mathbf{y}$ is the observed series; $\hat{\mathbf{y}}$ is the simulated series; w_n represents the weight of the fitness value computed at point n; $D_{(n,\text{magnitude})}$ is the observed series' average value; $D_{(n,\text{length})}$ is the observed series' length; and α, β, and γ are adjustable parameters controlled by the user.

Calibration of specific, conflicting, local points
Additionally, it is possible not only to use aggregated values, but also to consider the minimization of particular fitness values for chosen locations in a MO optimization problem. This approach can be combined with the aggregation of multiple points.

Time-weighted fitness measures
Another option included in the interface foresees that time might play a role in the importance of the calibration data. Using this option, fitness measures are computed on a yearly basis and combined according to a weighted average. Weights can be arbitrarily specified, but will normally grant more importance to the more recent observations.

Cascading calibration
The heterogeneity of the ZRB and the computational challenges that were faced prompted the development of an optimal cascading approach to calibration.

When this option is chosen, the interface analyzes the river network and calibrates selected subbasins separately; proceeding downstream and using calibrated simulation results from the upstream areas as cascading inputs. As only deterministic outputs should be passed downstream, the cascading calibration was only made viable within a SO optimizations.

---

[11] In this example two fitness measures are defined. Because optimization algorithms normally perform minimizations, the Nash-Sutcliffe efficiency coefficient is not used directly on the first. On the second, the MAE and the RMSE are weighted.

[12] The relative importance between calibration points can change, for example, between a point located downstream, in the main channel of the Zambezi and one placed at a minor upstream tributary. Similarly, series with many missing records should be considered to contain less information than complete ones and, therefore, desirably have less weigth on the calibration.

As the interface "crops" the SWAT simulation on-the-fly, for each step of the cascade the number of simulated subbasins is small, reducing the computation time manifold.

Regional calibration

Finally, the heterogeneity of the ZRB renders the existence of basin-wide optimal parameters unlikely. Motivated by this view, another calibration approach was developed. The approach, applicable to SO or MO frameworks, assumes that chosen parameters are only regionally homogeneous, differing among different parts of the basin. The regions of "homogeneity" can be defined independently for particular subbasin configurations.

7.6   Recommendations for future improvements

The SWAT model constitutes a gargantuan effort and, to some extent, a work in progress with many contributors. Coded in FORTRAN, the sheer number of variables and subroutines that are employed renders it extremely hard to assimilate.

During model set up, perhaps the issue with the most impact is the definition of channel widths. In spite of its large influence on simulations, among the myriad of input files, all laden with parameters, channel width is easily overlooked; it should be addressed with care.

The issues related to waterbodies were particularly emphasized in the light of this research's goals. Whenever they are expected to exert noticeable influence on a basin's behavior, however, the modifications described above are highly recommended.

Additionally, there are issues that, although detected, were not worthwhile addressing given the objectives and constraints of this work. One example is the process by which reach transmission losses are computed, which firstly calculates daily contribution to bank storage, then evaporation from the stream, and finally bank storage contribution to streamflow. Seemingly innocuous, this ordering can mask, for example, the evaporation of all the available streamflow in a given reach as, at the end of the day, outflow is still produced by water flowing back into the reach from bank storage[13].

Finally – and quite unfortunately – in the final stages of the research it was unveiled that, when the Soil Conservation Service's (SCS) equations (see Section 3.2.3) are used in order to compute surface runoff, the canopy storage can be grossly overestimated by SWAT. In fact, SCS's equations take an aggregate estimate of the initial abstractions that includes the canopy storage. SWAT, however, seems to explicitly subtract the canopy storage component from the precipitation prior to the application of the SCS's equations. This conceptually means the abstraction is considered twice. Still, when SCS's equations are used, an exception within the code resets the water contents of the canopy at the beginning of each time step. Obviously, if the canopy is parameterized such that its storage volume is large, considerable errors will ensue. From this point of view, considering a maximum canopy storage of 0 whenever SCS's equations are in use is recommended[14].

Regarding calibration, recommendations have to be that investing in computational performance can have great returns – in the present case translated into a 20-fold enhancement.

The calibration interface whose main features have been described above, is publicly shared in the ADAPT online database, at http://zambezi.epfl.ch.

---

[13] A solution to this problem was not implemented. In fact, such a solution might not be straighforwad at all as bank storage and evaporation are computed sequentially and within different routines by SWAT.

[14] At the time this error was identified, it was impractical to re-run the model's optimizations. The maximum canopy storage, however, was considered as a free parameter in the procedure. As such, although conceptually lacking, it is believed that the simulations will not be extremely affected.

After further inspection, the finding will be communicated to the SWAT community.

8   ADVANTAGES OF THE PATTERN-ORIENTED MEMORY RAINFALL INTERPOLATION METHOD FOR HYDROLOGICAL MODELING

8.1   Introduction

The Pattern-Oriented Memory (POM) rainfall interpolation method (Matos et al. 2013a, Matos et al. 2014) was introduced in Chapter 5. Making use of state-of-the-art satellite rainfall estimations and explicitly accounting for time during the interpolation process, the method seems to outperform competing approaches when spatial data is scarce and the areal rainfall field displays high variability. Beyond the performance gains which have been demonstrated, POM interpolations are particularly interesting due to additional properties, object of this chapter.

Even taking the adequacy of satellite daily rainfall estimates' values as forcing for hydrological models as granted, some constraints arise from their use. These constraints are based on:

1. the limited availability of satellite estimates, which do not encompass periods prior to 1998;
2. the fact that the nature, or support[1], of satellite rainfall estimates is different from that of gauged series; and
3. arguably, the existing series of satellite rainfall estimates are not suitable to represent the variability of rainfall-runoff responses that typically occur within Southern Africa (Hughes 2006).

In practice, if a model is prepared to run based on satellite rainfall estimates, it is uncertain if it can produce reliable results when forced by rainfall estimates with different support. Focusing on an inverse problem, Artan et al. (2007) calibrated a model using Inverse Distance Weighting (IDW) and tested daily satellite rainfall as forcing, obtaining poor results. Conclusions were that a small bias in satellite rainfall estimates was amplified in streamflow simulations and that satellite data should not be used on models calibrated with gauged rainfall data.

In an in-depth study about the combination of satellite and gauged rainfall data over Southern Africa[2], Hughes (2006) highlights the incompatibilities between gauged and satellite rainfall data when used by a calibrated hydrological model. In fact, if the model is calibrated on gauged rainfall data, its parameters will inevitably reflect any errors or inadequacies of those series (such as

---

[1] The terminology "change of support" is commonly used in geostatistics (e.g. Gelfand et al. 2001) to address issues related to the mismatch between point and block or areal data. Here, the terminology is used more broadly to encompass also other features that contribute to diffenreces between rainfall datasets. An illustrative example of such features is the mismatch in quantile-quantile plots between satellite and gauged rainfall estimates.

[2] Including the Kafue River.

a lack of adequate spatial representation) or resulting from the way in which the gauge data were used to obtain subbasin average rainfall inputs. In the ZRB, due to the low spatial availability of gauge records, models are preferentially calibrated resorting satellite rainfall estimates. If possible, the use of gauge data would, however, still be desirable.

With POM, the main spatial patterns observed within the satellite rainfall field used as a basis for the interpolation are transmitted to the interpolated surface being, to some degree, the support maintained. Stemming from this, it is hypothesized that POM can have, beyond performance, valuable advantages over traditional rainfall interpolation methods in regard to hydrological modeling. These can be summarized as:

- allowing model calibration at points which would otherwise be difficult to consider;
- carrying out historical simulations with increased accuracy;
- improving model reliability by adopting longer calibration periods.

The first advantage is self-evident in the case of discontinued discharge series. In the ZRB there are plenty of "old" discharge series which have been discontinued for decades and which, without adequate contemporary forcing data, cannot be used for calibration purposes. One particularly illustrative example within the ZRB can be found on the Luangwa catchment (see Appendix I.A, Figure AI.6 for the location). The available discharge series for the Mfuwe hydrometric station, gauging a considerable area of the Upper Luangwa catchment, is presented in Figure 8.1.
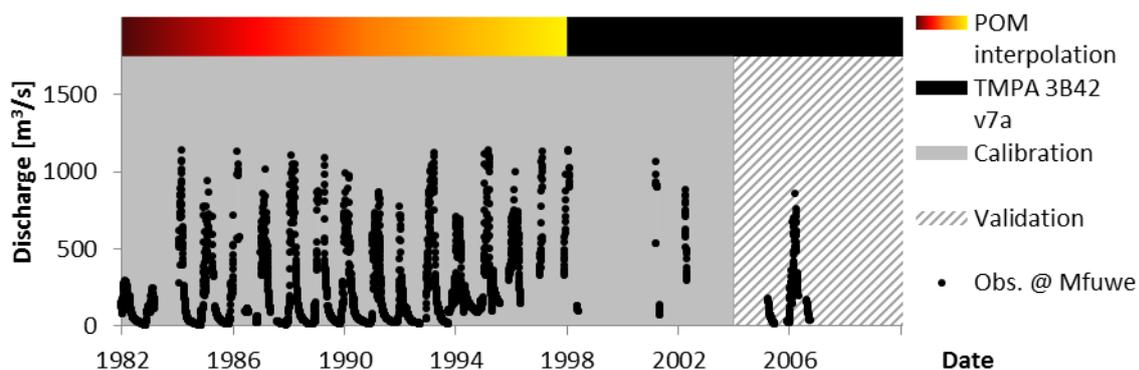
Figure 8.1. Illustration of the effects of longer forcing series on simulation periods; particularly the ones devoted to calibration. Example of observed flow series at Mfuwe (see Appendix I.A, Figure AI.6 for the location).

As can be seen, the very few observations that overlap the period for which satellite rainfall estimates are available are largely insufficient for successful calibration efforts. Older records, however, are plentiful. By using POM interpolations in addition to satellite rainfall estimates, the overall span of possible simulations more than doubles while the time frame that can be devoted to calibration more than triples[3].

The other two advantages, related to the change of support regarding rainfall inputs, require a deeper analysis in order to be proven. The devised methodology, results and discussion are presented in the remainder of the chapter.

---

[3] These numbers assume an original calibration, using solely satellite estimates as the rainfall forcing, spanning between 1998 and 2003 and a validation period from 2004 onwards, which are periods used in the scope of Chapter 9 and in the work of Cohen Liechti (2013)

8.2   Methodology

8.2.1   Increased accuracy in historical simulations

In order to evaluate the earlier hypotheses it must first be established that, when combined with satellite data, the use of historical areal rainfall maps interpolated using POM leads to better simulations than the use of rainfall maps derived by alternative techniques.

To do so, the detailed Soil and Water Assessment Tool (SWAT) model of the ZRB, described in Chapter 7, was employed. The parameter calibration (of which a detailed discussion is deferred to Chapter 9) was undertaken for the period from 1998 to 2003[4] using a cascading approach.

The calibrated model was employed to simulate a longer time frame, from 1982 to 2009. While for the calibration only TRMM 3B42 v7a estimates were taken as rainfall forcing, in order to prepare inputs to be used in the wider simulation (comprising an additional period from 1982 to 1997), interpolations were needed.

Three sets of interpolated areal rainfall maps were computed using POM, IDW, and Multiple-Realization Ordinary Kriging (MROK)[5], being model forcing based on them. Simulation results referring to the period of 1982 to 1997 were then compared.

8.2.2   Advantages of a longer calibration period

The question of whether POM interpolations can lead to more reliable hydrological simulations goes deeper. Even after establishing that rainfall forcing derived from POM is better than the competing ones, two additional conditions must be verified:

1. POM's historical rainfall must be sufficiently accurate and similar to satellite estimates so that an actual model calibration can be performed on them; and
2. There should be relevant information in historical series prior to 1998 that is not assimilated by the hydrological model when shorter, more recent calibration periods are adopted.

In order to test these conditions, two SWAT models, resorting to extended or standard calibrations, are evaluated over an independent test period. The extended calibration makes use of POM and TMPA 3B42 v7a data and goes from 1982 to 2003, c(82-03). The standard calibration employs TMPA 3B42 v7a data exclusively and spans from 1998 to 2003, c(98-03). The test period is defined from 2004 to 2009.

8.2.3   Evaluated locations

The evaluation of results was limited by data constraints and issues associated with model calibration; easier in unregulated catchments, it was undertaken for 7 points of the ZRB, illustrated in Figure 8.2. As simulations cascade, the results of downstream points encompass upstream information.

8.2.4   Error measures

Four error measures were considered in the analysis, namely the Nash-Sutcliffe efficiency (NE) coefficient (Nash and Sutcliffe 1970), the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE), and the Volume Ratio (VR).

These measures have already been introduced in the scope of Chapter 6 and will not be reproduced where. Their mathematical description can be found in Section 6.3.3.

---

[4] While the methodology is the same, this period should not to be confused with that of the calibrations conducted in Chapter 9, which comprises the years from 1982 to 2003.
[5] As IDW and MROK were the most competing established interpolation models according to the findings of Chapter 5. A brief introduction of the two techniques is provided in the same chapter.
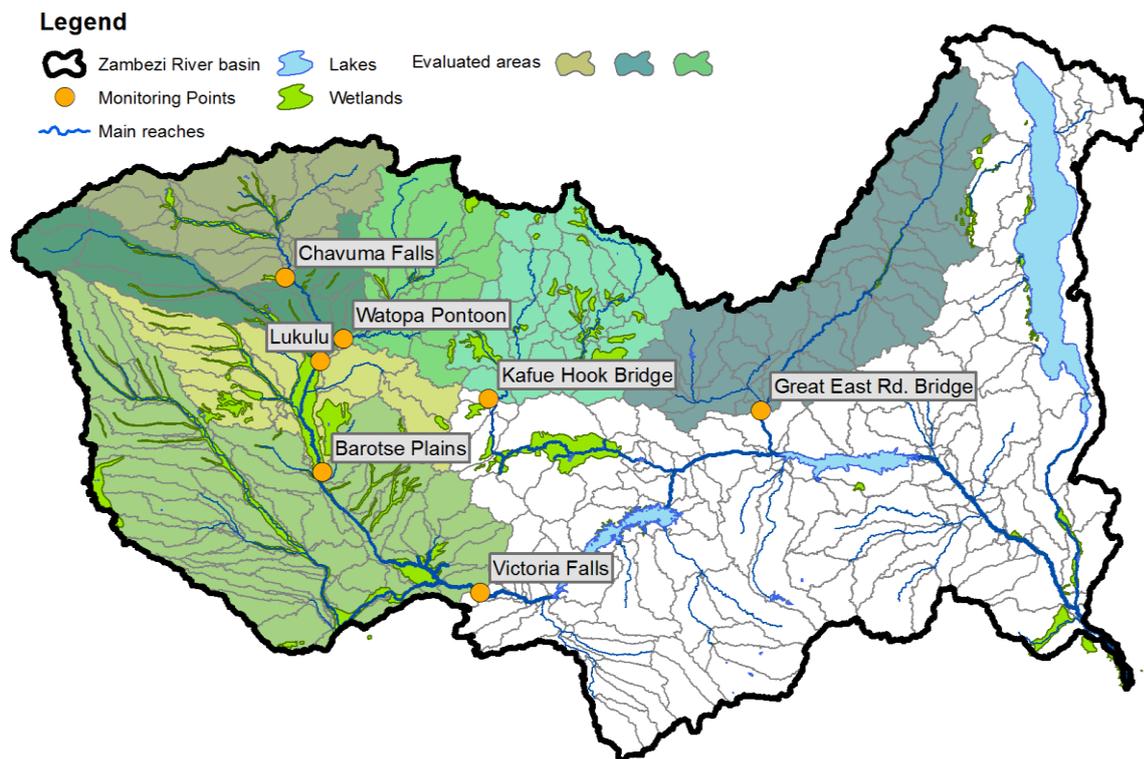
Figure 8.2. Areas evaluated in order to assess the potential advantages of the POM interpolation.

8.3   Results and discussion

8.3.1   Increased accuracy in historical simulations

The accuracy of historical simulations was assessed for 7 distinct points of the ZRB from 1982 to 1997, which is the time frame where interpolated rainfall maps are used as model forcing.

Observed and simulated hydrograph sets are shown below for two points of the basin. The first, referring to Kafue Hook Bridge, are presented in Figure 8.3.

The Kafue Hook Bridge location was chosen as it presents the worst case among the studied ones. Even in this worse case, however, it can be noticed that POM's interpolated rainfall provides a much better forcing for the calibrated SWAT model than does IDW's or MROK's.

The second example illustrates the hydrographs for Victoria Falls, a location chosen because it encapsulates the majority of the analyzed area and as its hydrographs are very similar to the remainder. As can be seen in Figure 8.4, while IDW and MROK interpolations lead to nearly constant responses, incurring in large volume underestimations, the application of POM equates to much more appropriate simulations. In fact, POM's simulations seem to be remarkably close to the ones which benefit from satellite rainfall estimates as forcing (1998 onwards).

The quantitative evaluation of the different forcings is translated in Table 8.1, where the best results for each case are highlighted in bold red font. Albeit not being perfect, simulations that use POM are superior to others on every account, be it in relation to error measure or location.

Straying away from an assessment of the absolute quality of the results, it can be clearly seen that, if historical interpolations are needed, the POM interpolation method presents considerable advantages over the tested alternatives.
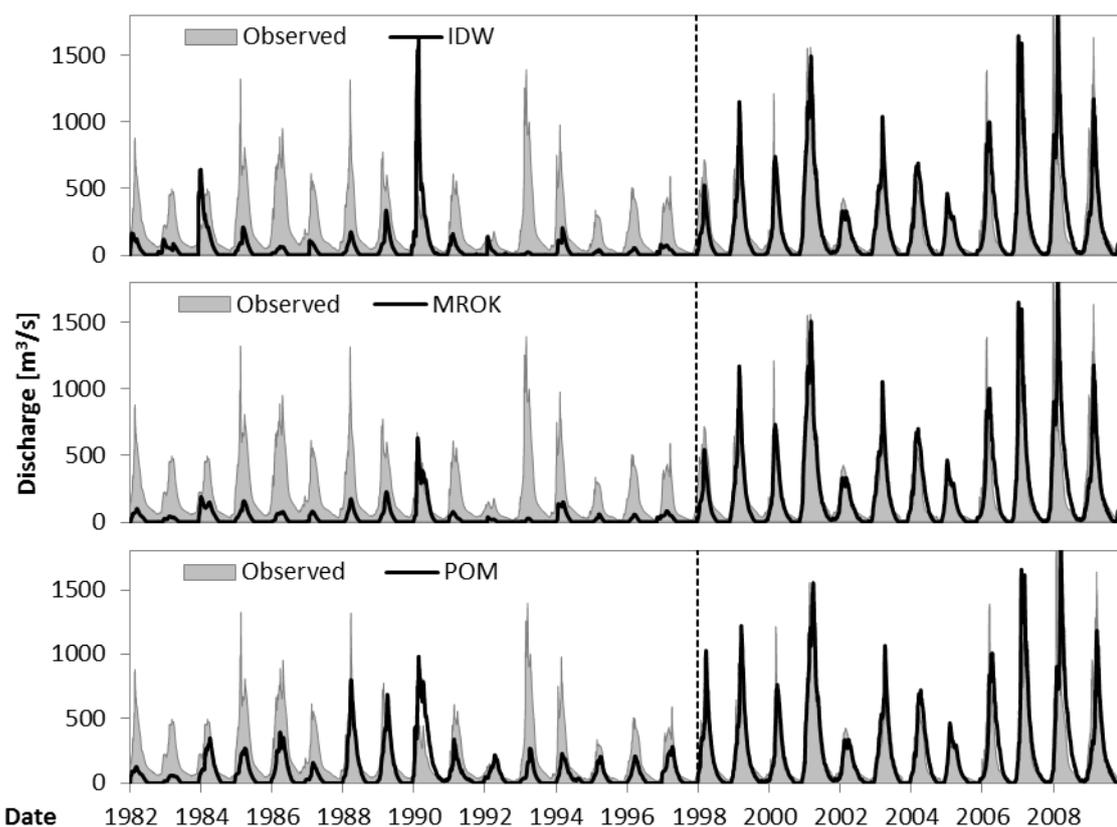
Figure 8.3. Simulated and observed hydrographs for Kafue Hook Bridge. From top to bottom, IDW, MROK and POM rainfall interpolations are used as forcing from 1982 to 1997.
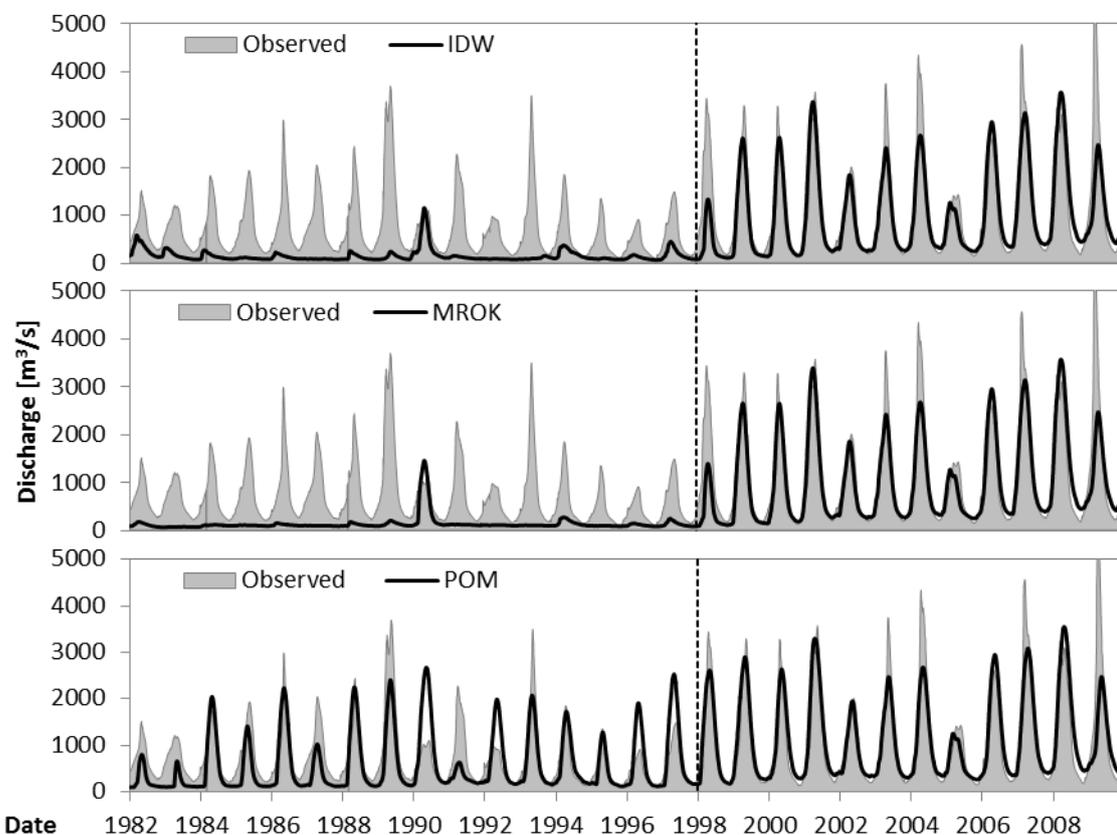
Figure 8.4. Simulated and observed hydrographs for Victoria Falls. From top to bottom, IDW, MROK and POM rainfall interpolations are used as forcing from 1982 to 1997.

Table 8.1. Overall comparison of simulation performance measures from 1982 to 1997.

|		POM	MROK	IDW			POM	MROK	IDW
	NE	**0.13**	-0.37	-0.42		NE	**0.40**	-0.86	-0.77
Chavuma	VR	**1.31**	0.15	0.09	Victoria	VR	**0.82**	0.19	0.21
Falls	RMSE	**517.9**	649.8	660.8	Falls	RMSE	**481.5**	844.4	822.9
	MAE	**285.5**	377.9	390.7		MAE	**331.1**	586.0	562.1
	NE	**0.20**	-0.65	-0.55		NE	**0.09**	-0.46	-0.50
Watopa	VR	**0.48**	0.26	0.32	Kafue	VR	**0.43**	0.15	0.23
Pontoon	RMSE	**130.3**	186.6	181.0	Hook	RMSE	**204.9**	259.8	263.2
	MAE	**98.6**	140.1	133.5	Bridge	MAE	**136.6**	176.9	176.4
	NE	**-0.50**	-1.43	-1.38		NE	**0.21**	-0.01	-0.04
	VR	**0.82**	0.08	0.07	Great East	VR	**0.37**	0.18	0.17
Lukulu	RMSE	**561.2**	714.9	708.6	Rd. Bridge	RMSE	**1176.5**	1329.3	1349.4
	MAE	**416.1**	570.8	565.5		MAE	**473.0**	579.7	588.5
Barotse	NE	**0.33**	-1.34	-1.41					
Plains	VR	**0.78**	0.17	0.14					
(Senanga)	RMSE	**444.3**	827.6	840.1					
	MAE	**328.3**	627.0	646.3					

8.3.2 Advantages of a longer calibration period

Being the superior quality of POM historical simulations established, the potential advantages of a longer calibration period sustained by their use are analyzed below. Such an analysis is naturally associated with the calibration and validation procedures themselves.

The calibration of hydrological models is a wide and complex topic. An introduction to it was provided in Section 3.3 and a deeper discussion of the thematic ensues in Chapter 9. At this stage, however, it should suffice to keep in mind that a successful calibration can only be confirmed in the light of a representative independent validation period. An example is provided in Figure 8.5.

Addressing the main question of this section, results from two calibration runs of the detailed SWAT model are evaluated at Chavuma Falls. In Figure 8.5 the evolution of both calibrations can be evaluated according to the best fitted model at each iteration in the light of the NE error measure[6]. For every iteration, results are evaluated over the calibration and validation periods. For the extended calibration, c(82-03), the period from 1982 to 2003 is considered, while for the standard calibration, c(98-03), observations between 1998 and 2003 are used. The validation reports to the time frame between 2004 and 2009.
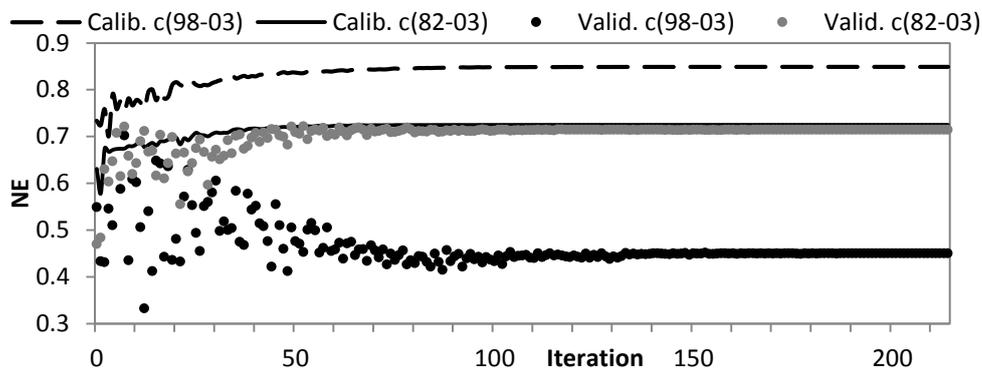
Figure 8.5. Evolution of two calibrations of the detailed SWAT model at Chavuma Falls.

For Chavuma Falls and looking at the calibration periods alone, both c(98-03) and c(82-03) show improvements as the iterations progress, roughly stabilizing after 60 or 70 iterations are completed. Also, the standard calibration, c(98-03), achieves NE coefficients of approximately 0.85, while the

---

[6] The application of the NE error measure is justified in Chapter 9.

extended calibration, c(82-03), converges to a more modest NE coefficient of 0.70. Looking at the validation period – the one that matters most –, the results change markedly. c(98-03) leads to a NE coefficient below 0.50, while for c(82-03) the 0.70 NE is kept.

In this illustrative example, the use of the extended calibration was clearly beneficial. The same did not occur at every location. In Table 8.2 the results are summarized. For c(98-03) and c(82-03), error measures corresponding to the calibration and validation periods are presented. The best validation results in each case are highlighted in bold red font.

Table 8.2. Error measures for calibration and validation periods. Standard (c(98-03)) and extended (c(82-03)) calibrations considered.

		Calibration		Validation				Calibration		Validation	
		c(98-03)	c(82-03)	c(98-03)	c(82-03)			c(98-03)	c(82-03)	c(98-03)	c(82-03)
Chavuma Falls	**NE**	0.85	0.72	0.45	**0.71**	**Victoria Falls**	**NE**	0.91	0.76	0.72	**0.73**
	VR	1.06	1.05	1.29	**1.14**		**VR**	1.07	0.98	**0.99**	0.99
	RMSE	341.4	351.8	555.8	**400.7**		**RMSE**	295.4	360.8	628.0	**617.6**
	MAE	187.8	206.6	285.2	**220.1**		**MAE**	215.5	238.6	374.7	**361.6**
Watopa Pontoon	**NE**	0.88	0.61	0.69	**0.71**	**Kafue Hook Bridge**	**NE**	0.89	0.60	**0.60**	0.60
	VR	0.90	0.81	0.81	**1.03**		**VR**	0.96	0.87	**1.09**	1.39
	RMSE	67.2	101.8	127.9	**125.5**		**RMSE**	96.3	150.8	**196.5**	198.0
	MAE	58.1	75.7	101.2	**89.6**		**MAE**	64.1	91.0	**115.5**	132.3
Lukulu	**NE**	0.76	0.63	0.54	**0.77**	**Great East Rd. Brige**	**NE**	0.59	0.43	**0.55**	0.05
	VR	0.91	0.96	0.91	**1.04**		**VR**	0.76	0.96	**0.94**	1.84
	RMSE	366.0	332.7	518.8	**367.2**		**RMSE**	573.3	950.6	**626.0**	908.4
	MAE	297.1	234.6	357.4	**246.4**		**MAE**	318.2	408.8	**335.1**	522.8
Barotse Plains (Senanga)	**NE**	0.95	0.81	0.83	**0.84**						
	VR	1.04	0.98	**1.03**	1.07						
	RMSE	168.7	275.2	328.8	**322.4**						
	MAE	128.5	195.5	243.4	**225.2**						

In the majority of the points the use of an extended calibration seems to have improved the reliability of the hydrological model. This is the case of Chavuma Falls, Watopa Pontoon, Lukulu, Senanga and Victoria Falls.

On Kafue Hook Bridge the standard calibration performed better – but only marginally – being the only remarkable difference observable through the VR. Great East Rd. Bridge's results evidence clear benefits in the use of a standard calibration. Here, however, a hasty interpretation can be challenged on grounds of the data's representativity, as will be seen in the following.

In order to further analyze the differences between standard and extended calibrations, hydrographs are plotted for three locations: Chavuma Falls, where the benefits of the latter seem to have been greatest; Kafue Hook Bridge, where the extended calibration did not surpass the standard; and Great East Rd. Bridge, whose results hint a clear disadvantage of the extended calibration.

In Figure 8.6 the observed and simulated hydrographs are displayed for Chavuma Falls. The validation period can be seen on top, being the calibration time frame depicted below. In general terms, it can be observed that the standard calibration leads to an overestimation of high flows both during validation and from 1982 to 1997.

In agreement with the numerical results, the hydrographs for Kafue Hook Bridge, illustrated in Figure 8.7, are harder to classify. During the validation period it appears that the extended calibration leads to poorer representations of the recession of the flood while performing slightly better in the simulation of some of the peaks. From 1982 to 1997, there is a clear advantage of the extended calibration over the standard one. This, however, could be expected and is of limited significance

as extended calibrations make use of observations in that same period and standard calibrations do not. Also, it is perhaps worth referencing that it was for Kafue Hook Bridge that POM's interpolated rainfall forcing led to worst results in the first part of the analysis (Figure 8.3); something that is likely to have affected negatively the extended calibration.
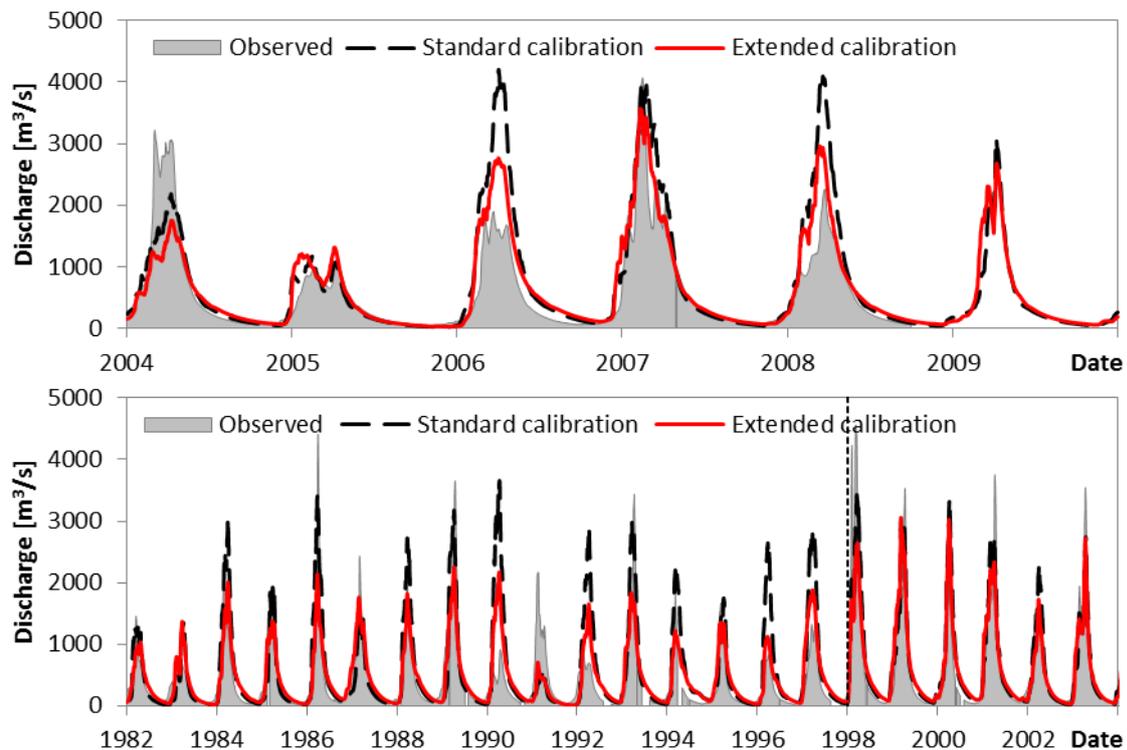

Figure 8.6. Simulated and observed hydrographs for Chavuma Falls. Validation (top) and calibration (bottom) periods are depicted. Year of 1998 marked with a dotted line.
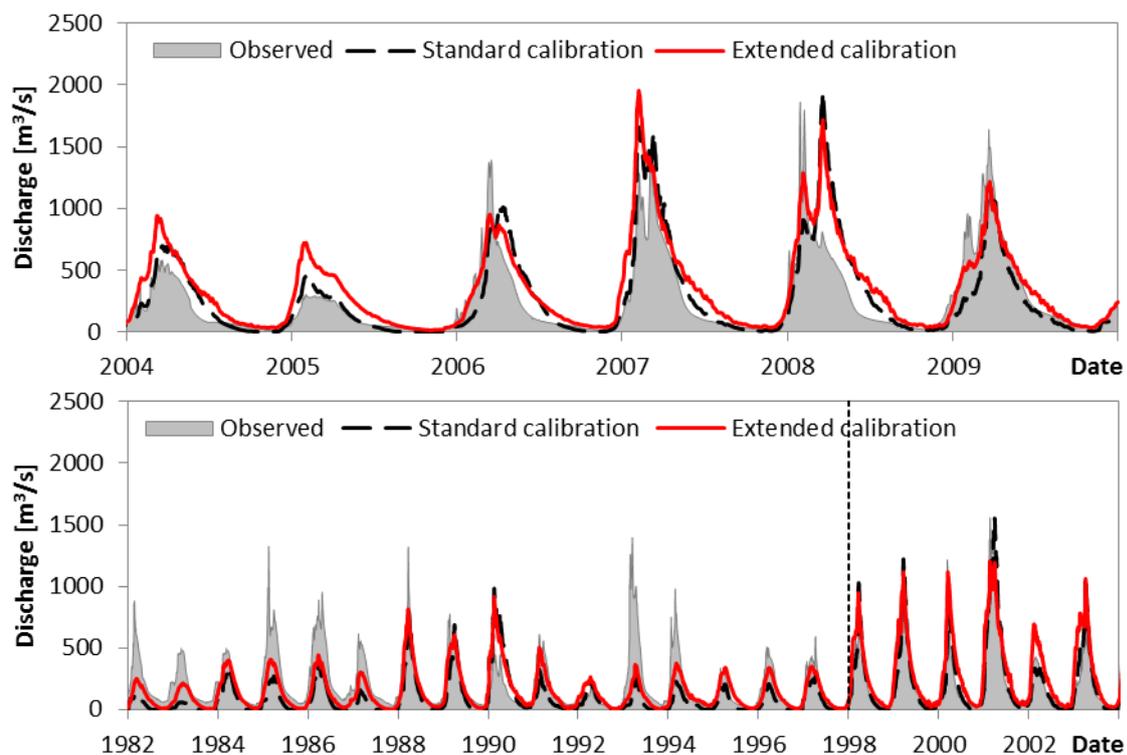

Figure 8.7. Simulated and observed hydrographs for Kafue Hook Bridge. Validation (top) and calibration (bottom) periods are depicted. Year of 1998 marked with a dotted line.

Finally, Figure 8.8 presents hydrographs for Great East Rd. Bridge. Looking at the bottom plot it can immediately be seen that the Luangwa subbasin has a very fast hydrological response, with abrupt changes in the daily discharge[7]. While this might have played a role in the numerical results shown above, it is likely that the reason underlying the markedly worse validation performance of the extended calibration is the lack of data from 2003 to 2009.
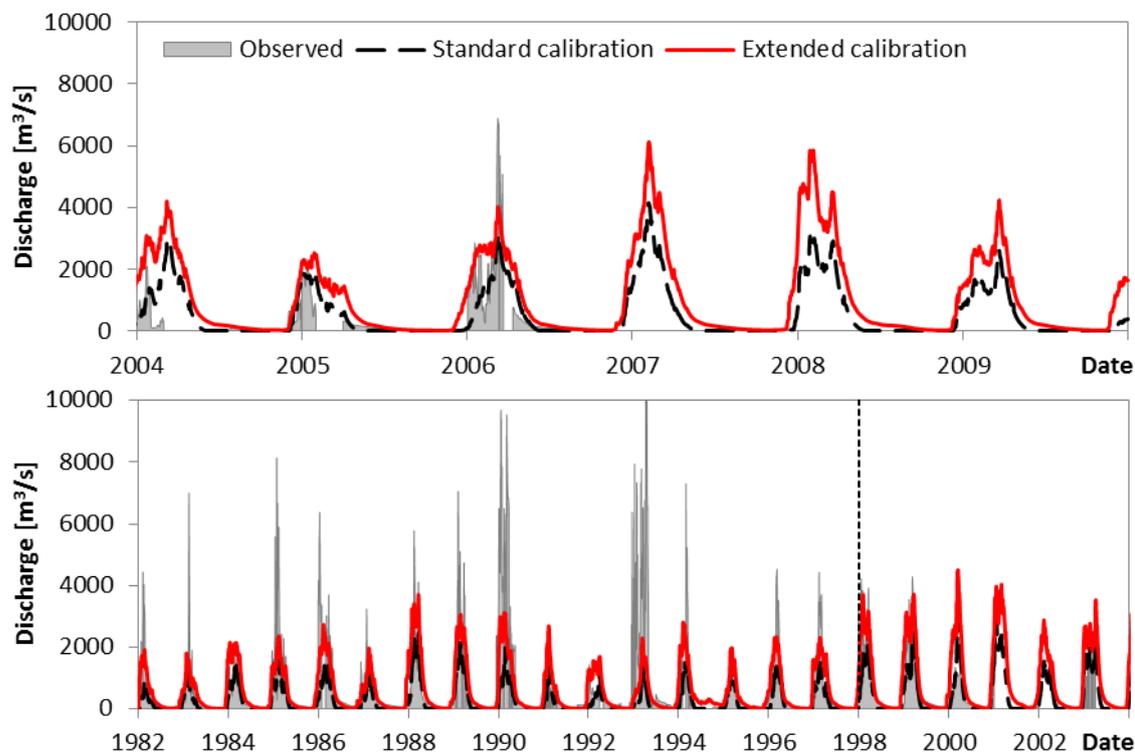
Figure 8.8. Simulated and observed hydrographs for Great East Rd. Bridge. Validation (top) and calibration (bottom) periods are depicted. Year of 1998 marked with a dotted line.

8.4 Conclusions

Often, satellite rainfall estimates are the best option as inputs to hydrological models. When this is the case, daily simulations are usually constrained to a time frame starting in 1998. If the interpolation of gauged records prior to 1998 is undertaken in order to generate historical areal rainfall maps to be used in a satellite-calibrated model, a change of support problem might emerge due to the intrinsically different nature of the generated interpolations and the satellite estimates.

Owing to its working principle, POM interpolation should alleviate the aforementioned change of support problem. In fact, because POM uses satellite rainfall estimates as a basis for the gauge records' interpolation, the main patterns observed in that data are propagated to the interpolations.

In the present chapter properties of POM interpolated rainfall when applied as a forcing of hydrological models were analyzed. In this scope three main hypotheses were put into focus:

- POM can allow model calibration at points which would otherwise be difficult to consider;

---

[7] In fact, it seems plausible that the series is affected by some measurement errors. Notwithstanding, there is no practical way of detecting eventual errors and the recorded discharges at Great East Rd. Bridge are valuable. In the Luangwa, only Mfuwe has another useful discharge series but, unfortunately, it displays many gaps (Figure 8.1) and integrates less than half of the subbasin's area.

- historical simulations can be performed with increased accuracy when POM interpolations are used;
- such historical simulations can be used in extended calibrations enhancing model reliability.

Results from seven locations within the ZRB seem to confirm such hypotheses. The first is self-evident when available discharge series that can be used for model calibration pre-date 1998 (particularly if the other two hypotheses are confirmed). The second was clearly demonstrated, being that POM-derived rainfall forcing led to the best historical results according to all tested error measures and for all locations. Finally, it was shown that, under adequate conditions, an extended calibration using POM interpolations can be advantageous or on par with standard calibrations using only satellite data.

It ensues from the findings of Chapter 5 that the performance of POM rainfall interpolations changes with spatial data availability and the variability of the underlying rainfall field. As such, under different conditions the use of similar historical rainfall interpolations might not be as advantageous as it appears to be for the ZRB. Also, the usefulness of an extended calibration period does not depend only on the quality of the historical rainfall interpolations and their similarity to the used satellite rainfall estimates, it hinges on the amount of information that can be added to the standard calibration as well.

9 HYDROLOGICAL MODEL CALIBRATION

9.1 Introduction

The present chapter appears as this document comes to a close. In large measure, a great deal of the developed work converges here.

It is very hard – and arguably pointless – to fully convey the complexity that can be associated with an accurate and encompassing calibration of a hydrological model for the ZRB. This complexity stems perhaps from the fact that the objective of achieving a performing and reliable hydrological model of such a large area is on the one hand, a much applied goal and, on the other hand, an extremely far-reaching one.

The applied character of the goal dictates that the problem cannot easily be posed in a way that is comfortable to the researcher or allows that clear conclusions are drawn. The goal is far-reaching because a plethora of aspects have to be balanced and accounted for at all times.

Among the aspects to be taken into account comes, firstly, the notorious trade-off between calibration performance and model reliability. In second come practical issues such as computational performance, scarcity of information, or the choice of the optimization algorithm. Still, other issues gain particular relevance in the ZRB. As examples:

How to address the optimization problem?
- Consider a Single-Objective (SO) optimization?
- Consider a Multiple-Objective (MO) optimization where the different objectives are defined by different error measures?
- Consider a MO optimization in which the objectives are defined by the same error measure at different locations?
- Consider a Bayesian Inspired (BI) optimization at the expense of higher computational demands?

How to properly define objectives?
- How to find the right balance between diversity and number of objectives, knowing that added dimensions hinder the optimization effort?
- How to combine error measures computed at different locations into one coherent and representative objective?
- How to combine different error measures computed at the same location, overcoming scale issues?
- How to achieve objectives that incorporate information derived from different types of data (e.g. discharges, volumes and water levels)?

How to improve the model?
- How to better describe hydrological and hydraulic processes?
- What optimization algorithm to choose?
- What parameters to change in order to obtain good phenomena representations while promoting parsimoniousness?
- How to set an appropriate scope for the selected parameters (e.g. global, regional, local...)?

How to make the best use of the scarce information available?
- How far back in time to extend the calibration?
- How to properly address the challenge posed by the spatial heterogeneity of the basin?
- How to reliably calibrate ungauged areas resorting to information retrieved from nearby zones?
- How to cope with the likely event that not all of the observed data used for calibration is accurate?
- How to integrate simulations that were generated using different forcing supports (see Chapter 8 for a discussion on the subject in regard to rainfall)?

These are but a fraction of the relevant (and still unsolved) questions that can be posed. Evidently, it is far beyond the scope of a document of this nature to present sound answers to them.

Owing to the practical nature of hydrological modeling, whether these and other problems are explicitly dealt with or not, they cannot simply be cast aside. In this work, a choice was made to attempt overviewing the calibration problem in its full complexity. Obviously, this choice was costly, not least in terms of the share of the workload it demanded, but also in the meandering direction of the research and on the tractability of the methodology.

In the following sections, having the previous questions in mind, a concise (and partially rearranged) description of the efforts devoted to model calibration and validation is unfolded.

9.2   Methodology

9.2.1   Overall strategy
In order to address the main questions surrounding the calibration of hydrological models (with an emphasis on SWAT) two models have been prepared. With the goal of gaining insight into what is an adequate degree of complexity to model the ZRB given the available resources (mainly in terms of the quantity, quality and distribution of calibration and forcing data), a parsimonious and a detailed representation of the ZRB in SWAT have been evaluated. Their preparation, along with details on the calibration framework, is described in Chapter 7.

Past hydrological modeling attempts within the ZRB, briefly described in Section 3.1, have either been focused on particular areas of the basin or consisted of broad overviews of its hydrology working at monthly or yearly time steps. One notable exception is the work by Cohen Liechti (2013) – to the author's knowledge the first to endeavor the modeling of the whole ZRB on a daily scale.

Although large gains stand to be made from accurate models of the basin, modelers shun away from these for good reasons. Chief among them are the ZRB's large size and heterogeneity, which allied to the lack of ground hydrological data (both for model calibration and forcing), pose an enormous challenge. While areas upstream of the Kariba and Itezhi-Tezhi reservoirs are characterized by a fair amount of information, downstream of the confluence of the Kafue and the Zambezi rivers' information lacks and modeling becomes extremely difficult. In particular, very little recent discharge or water level data about the southern bank of the Zambezi River was available in the scope of this study and, to the North, the Luangwa and Shire subbasins[1] pose interesting challenges. The poorly monitored Luangwa is marked by a fast hydrological response

---

[1] Refer to Appendix I.A for these locations.

and can generate very high flood peaks. The Shire subbasin's discharges, in spite of being driven by the outflow of the Malawi Lake, reflect water resources management changes in recent times (e.g. the indirect control of levels on the lake); due to this, the extent to which historical data is accurate within the subbasin is uncertain. Finally, as one approaches the delta, only water level series are available. Because lack of information about river cross sections and changing profiles due to high sediment transport rates, the establishment of rating curves is difficult and their applicability short-lived. Calibrating precise hydrological models on such grounds is hard and even the very recent Cohen Liechti's (2013) model was not validated downstream of Cahora Bassa.

The chosen calibration approach should be able to cope with the issues above and yet retain tractable complexity and manageable computational demands. Firstly, one must decide whether to implement SO, MO or BI based schemes.

The general issues related to the calibration of hydrological models have been discussed in Section 3.3. An important aspect of model calibration that was not addressed before is that MO optimization tools and BI approaches solve far more complex problems than the one posed by SO optimization. In fact, while SO optimization searches for an optimal (single) point in the solution space, both MO and BI approaches search for an ensemble of solutions in order to characterize the Pareto set or the posterior probability distribution function of the parameters.

While MO and BI approaches, and particularly the latter, hold great promise, for some complex systems it is extremely challenging to find even a single satisfying solution, let alone the much larger numbers required to characterize the simulation's probability distribution function or the objective's Pareto surface.

Due to BI's high computational demands (e.g. Laloy and Vrugt 2012, Chu et al. 2014, Vrugt and Laloy 2014), which are perhaps not well-matched to a SWAT model of the ZRB[2], the approach was not applied and the work focused on SO and, mostly MO frameworks.

The definition of objectives with the goal of calibrating the models has to take into account several factors:

- Each model's performance is not defined by the quality of the simulations at one particular site, but rather at several locations throughout the ZRB.
- The model's performance should reflect not only the similarity between simulated and observed discharges, but also include information pertaining to stored volumes and water levels.
- The quality of a simulated series can seldom be completely characterized by a particular error measure.
- Not all the calibration information is necessarily reliable and/or coherent with the model's structure.
- The equifinality concept (Beven 1993) (see Section 3.1).

Objectives can be based on error measures such as the Mean Squared Error (MSE) or the Nash-Sutcliffe efficiency (NE) coefficient, but numerous alternatives exist. Depending on definition, some objectives may be adimensional, while others are relative to the units of the series being evaluated.

Defining an encompassing objective for a basin-wide SO calibration is difficult at best. The reason is that one has to account for different types of series, not necessarily easy to relate (e.g. discharge and volume), and to do it simultaneously for several distinct locations. Still, among the locations the series do not necessarily share a common number of records or mean value – which should reflect on their overall representation in the objective.

---

[2] Vrugt and Laloy (2014) consider optimization problems which might take up to much more than 10^6 (actually closer to 10^7) model evaluations in order coverge. Altough these problems comprise more than 200 parameters, being more complex than the ones solved in this document, 10^6 evaluations of the parsimonious SWAT ZRB model would take four months to complete in a high-end desktop computer with 8 models running in parallel – this, with no guanrantee of a satisfactory end result.

Illustrating difficulties, if one were to choose the minimization of the MSE as an objective and then average the result among different locations, the objective would be biased towards granting more importance to the ones with larger deviations from the observations in absolute terms. With The Nash-Sutcliffe efficiency (NE) coefficient this would not happen. Being an adimensional alternative, NE, does not present the same problems. It, however, varies in the range $]-\infty, 1]$, and the NE coefficient for poor adjustments between simulation and observations plummets fast through the negative scale. When attempting to combine NE's for different locations, even if many display near perfect fittings (NE near to 1), it is not uncommon that a single discordant series conditions the averaged NE for all locations. A scenario in which a single location can completely dominate the objective is naturally undesirable.

After objectives are properly defined, there should be a concern about the generalization capabilities of the produced model or, in other words, its capability to perform reliably not only when forced by "known" data, used for calibration, but also in the presence of "unknown" data. Here, this is tested using a validation period. In order to promote reliability (and as discussed in Chapter 3), the model's complexity – partly represented by the number of free parameters – should be as small as possible. Conversely, it would be expected that a certain degree of complexity is needed in order to reproduce also complex phenomena (such as river discharge). An adequate model will have, therefore, as much complexity as needed in order to perform well and as little complexity as demanded in order to be reliable.

Heterogeneity plays an important role in the Zambezi's hydrology. While it can be represented by using local parameters, not all subbasins may have independent parameters under penalty that the model's complexity will be such that reliability drops to undesired levels. Here, efforts were made in order to explore the role of heterogeneity and to develop models which account for it while keeping the degradation of reliability in check.

The first calibration attempt to be made was a MO one, handling reaches and reservoirs separately in two independent objectives and basing them on measures that do not produce bias when aggregated for multiple series. The scope of this calibration was global, being that heterogeneity was not considered. Both the parsimonious and the detailed model were tested. The approach was designated global calibration.

Following, a SO calibration was attempted. In order to overcome difficulties related to the combination of information about several locations and different types of series in one objective, while accounting for heterogeneity, this calibration was conducted separately in distinct areas of the ZRB. Calibrations progressed from upstream to downstream and information flowed as in a cascade. This approach was designated as cascading calibration.

A third attempt combined a MO optimization with heterogeneity. To do so, the most relevant parameters to the ZRB SWAT model were selected through a Fourier Amplitude Sensitivity Test (FAST), being allowed to vary regionally within the calibration framework. In order to control the model's complexity, an additional objective which minimized heterogeneity was included. The resulting methodology was named regional-regularized calibration.

Finally, a MO calibration scheme which combined the results of the SO cascading calibration and the regional-regularization capabilities of the third attempt was devised. Running with the results of the cascading calibration as a base model, the MO regional-regularized framework was used in order to further calibrate solely the most relevant parameters in a so-called cascading-regularized approach.

9.2.2 The Covariance Matrix Adaptation Evolution Strategy optimization algorithm

The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) optimization algorithm (Auger and Hansen 2005, Hansen 2006, García et al. 2009, Hansen 2010) was used in the scope of the SO cascading calibrations. It is a random search evolutionary algorithm adequate to real-parameter

optimization of non-linear, non-convex functions, in which the candidate solutions are sampled according to a multivariate normal distribution.

CMA-ES is a general purpose optimization tool. Its performance is well documented in comparative analyses (Hansen 2006, García et al. 2009). The main domains of application of the technique are non-separable functions that are ill-conditioned or rugged and an interesting feature of the CMA-ES is its quasi parameter-free nature. Also, CMA-ES is reported to overcome several problems often associated with evolutionary algorithms.

CMA-ES was chosen mostly because it is reported in specialized literature to perform well without prior assumptions or knowledge of the problem – which is, in practice, the case.

9.2.3   A Multi-Algorithm Genetically Adaptive Multi-objective method
The (A) Multi-Algorithm Genetically Adaptive Multi-objective Method (AMALGAM) optimization algorithm (Vrugt and Robinson 2007, Vrugt et al. 2009a) was employed in all MO calibrations.

Having been applied successfully along with the SWAT model in the past (Zhang et al. 2010, Cohen Liechti 2013), it is a meta heuristic optimization algorithm that draws its performance from several distinct MO optimization concepts and can approximate the Pareto surface of non-dominated solutions spanned by an arbitrary set of objectives. As a common evolutionary algorithm, it evaluates a population of possible solutions at each iteration. Owing to its meta heuristic nature, however, it produces new populations according to four independent optimization models:

- Non-dominated Sorted Genetic Algorithm-II (NSGA-II) (Deb et al. 2002b);
- Particle Swarm Optimization (PSO) (Kennedy and Eberhart 1995);
- adaptive metropolis search (Haario et al. 2001); and
- differential evolution (Storn and Price 1997).

In order to maintain population size, a selection has to be made among the offspring produced by each model. In order to do so, AMALGM accounts for the past success of the models in finding non-dominated solutions, selecting more offspring from the most promising ones.

9.2.4   The Fourier Amplitude Sensitivity Test algorithm
The FAST (e.g. Cukier et al. 1973, Schaibly and Shuler 1973, Saltelli et al. 1999, Cannavó 2012) is an ANOVA (analysis of variance) type global first order sensitivity test. It consists in the comparison between the variances of observations related to a given parameter and the overall variance.

In its essence, FAST is an elegant way of sampling the parameter space through the establishment of a one-dimensional search path which regularly covers it regardless of the number of parameters. Also, it provides a computationally efficient way of computing first order sensitivity indices based on Fourier coefficients (motivating its designation). It should be remarked that the FAST is limited to first order sensitivities and its accuracy is dependent on the number of points used to sample the parameter space.

The description of FAST, not justified in the scope of this work, can be found in numerous sources (e.g. Cukier et al. 1973). Finally, there are several sensitivity tests which extend the capabilities of the FAST. These, however, generally resort to increases in the number of model evaluations. In light of the objectives of this work, being the evaluation of the ZRB models a computationally demanding task, it was considered that what stood to be gained by a more complex analysis would not justify the increase in computation time.

9.2.5   Optimized parameters
Careful consideration of the theoretical principles implemented in SWAT (see Section 3.2), experience gathered through the work of Cohen Liechti (2013), and a lengthy process of trial and error led to a list of SWAT parameters, presented in Table 9.1, which was selected for close inspection. Indication of optimization bounds and whether the log transform was applied to their dimension during optimization is given in an extended table, presented in Appendix I.F.

Table 9.1. Synthesis of the SWAT parameters selected for optimization.

Code	Designation	Reference	Scale
ALPHA_BF	Baseflow alpha factor	α^{gw} , eq. (3.44) , page 33	absolute
Ave	Available water capacity of the soil layer	SW_i^{exc} , eq. (3.36) , page 32	**relative**
CANMX	Maximum canopy storage	See Section 7.6	absolute
CH_K2	Effective hydraulic conductivity in main channel alluvium	K^{rch} , eq. (3.59) , page 37	absolute
CN2	Initial SCS runoff curve number for moisture condition II	CN_2 , eq. (3.4) , page 25	**relative**
Depth	Depth from soil surface to bottom of layer	$z_i^{upper} - z_i^{lower}$, see page 31	**relative**
ESCO	Soil evaporation compensation factor	$esco$, eq. (3.28), page 30	absolute
EVRSV	Lake evaporation coefficient	η , eq. (3.68) , page 39	absolute
GW_DELAY	Groundwater delay time	δ_{gw} , eq. (3.40) , page 32	absolute
GW_REVAP	Groundwater "revap" coefficient	β^{rev} , eq. (3.42) , page 33	absolute
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur	$W^{sh,revap,thr}$, eq. (3.43) , page 33	absolute
SURLAG	Surface runoff lag coefficient	$surlag$, eq. (3.13), page 27	absolute
ALPHA	Natural reservoir alpha coefficient	α^{rsv} , eq. (3.72) , page 40	absolute
BETA	Natural reservoir beta coefficient	β^{rsv} , eq. (3.72) , page 40	absolute
K	Natural reservoir k coefficient	k^{rsv} , eq. (3.71) , page 39	absolute
H0	Natural reservoir h0 coefficient	$d^{rsv,thr}$, eq. (3.71) , page 39	absolute

The last four parameters in Table 9.1 are related to the uncontrolled outflow of major waterbodies in the basin, namely the Barotse Plains, the Chobe-Zambezi Confluence, the Kafue Flats and the Malawi Lake.

During a long phase of preliminary calibration efforts, in which these four parameters, taking different values for each waterbody, were adjusted along with the remainder, it was remarked that the attained performances were severely degraded due to bottleneck effects. As such, it was decided to independently calibrate those parameters directly with observed data before the main calibration run, where their value would be fixed. The hydrographs and volume series resulting from their adjustment are illustrated in Appendix I.G.

9.2.6 Error measures and objective definition

Building upon extensive testing and relying again on earlier work by Cohen Liechti (2013), six error measures were selected to serve as a basis for evaluating the quality of the produced simulations. Being that there is no limitation to the number of measures presented here, due to dimensionality concerns only few should actually be used simultaneously to calibrate the hydrological model. These are specified in the subsequent sections for each particular case.

Notwithstanding that some of these measures have already been introduced in the previous chapters, they are displayed again for convenience. Among the ones introduced earlier (e.g. Section 6.3.3) are the Mean Absolute Error (MAE), the Root MSE (RMSE), the NE coefficient, and the Volume Ratio (VR). Additionally, relative versions of the MAE and the RMSE, based on the observed variation interval, were introduced: the Average Interval-Relative Absolute Deviation (AIRAD) and the Interval-Relative RMSE (IRRMSE).

The measures are formally defined from equation (9.142) to equation (9.147). Among them N represents the length of the observed series, Q_t the observed discharge at time step t , $\hat{Q}_t$ the

corresponding simulated discharge, $\bar{Q}_t$ the average observed discharge, and $\underset{max}{\Delta} Q$ the range of the observed discharges.

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left| Q_t - \hat{Q}_t \right| \tag{9.142}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}\left(Q_t - \hat{Q}_t\right)^2}{N}} \tag{9.143}$$

$$\mathrm{NE} = 1 - \frac{\sum_{t=1}^{N}\left(Q_t - \hat{Q}_t\right)^2}{\sum_{t=1}^{N}\left(Q_t - \bar{Q}_t\right)^2} \tag{9.144}$$

$$\mathrm{VR} = \sum_{t=1}^{N}Q_t \bigg/ \sum_{t=1}^{N}\hat{Q}_t \tag{9.145}$$

$$\mathrm{AIRAD} = \frac{\mathrm{MAE}}{\underset{max}{\Delta} Q} \tag{9.146}$$

$$\mathrm{IRRMSE} = \frac{\mathrm{RMSE}}{\underset{max}{\Delta} Q} \tag{9.147}$$

As put forward above, the model calibration was not limited to the evaluation of discharge series. Where appropriate, also reservoir (or lake) volumes and channel water levels were used.

Reservoir volumes are directly computed by SWAT. In their case, the same error measures can be used with a seamless replacement of the discharge by volume series[3].

In regard to water levels, however, added difficulties arise. In fact, while it is possible to extract water level information from SWAT (albeit not straightforward), estimating water levels is not a main concern of the model and, thus, those values are regarded as potentially lacking accuracy. Additional work was undertaken, therefore, in order to match observed water levels and simulated discharges.

A simple rating curve of the type presented in equation (9.148) is adopted as to relate discharges Q [m³/s] to water levels h [m]. For each location of interest the three free parameters – α, β, and h_0 [m] – are unknown.

$$Q = \alpha\left(h - h_0\right)^{\beta} \tag{9.148}$$

Given this limitation, for each simulation a nested optimization procedure, based on the CMA-ES algorithm, is used to find the best combination of rating curve parameters according to the NE measure.

---

[3] The question of whether the same error measures are fit for discharge and volume is not relevant at this stage. As emphasized in Section 7.5.1, the model user has full control over which error measures (or even functions of error measures) are applied to what type of data, location, etc.

Once an acceptable calibration is found for the rating curve, the observed water levels are converted into proxy observed discharges and subsequently handled in a similar way to their "real" counterparts.

One issue remains, however: the reliability of the proxy discharges is expected to be less than that of "real" data. As a consequence, it is undesirable that they have similar weight in the overall model calibration. In order to materialize this concern, a coefficient was devised to selectively reduce the weight of water level-derived information on the calibration results.

An arbitrary minimum weight was set at 0.1. Such a minimum is reached when the NE of the calibrated rating curve approaches 0.5. When NE grows towards 1, there is evidently excellent agreement between water levels and discharges and, accordingly, the weight rises to 1. Conversely, as optimal NE values degrade, one can be fairly confident that the simulated discharges diverge from observed water levels. In this latter case, the water level information is also quite useful. Due to this, weights of "poorly" adjusted series were defined to arbitrarily rise to 0.4. This is translated by the convex weight function in Figure 9.1.
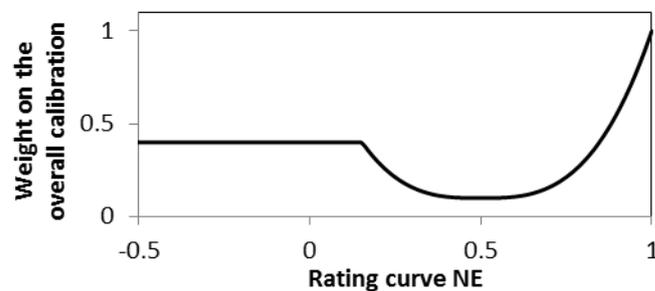

Figure 9.1. Weight function applied to water level observations. The weight is dependent on the optimized rating curve's NE coefficient.

As introduced in Section 7.5.2, the calibration interface can also associate weights to error measures according to their date. As the POM interpolated rainfall, used from 1979 to 1997, can be less reliable as a model input than satellite estimates (as discussed in detail in Chapter 8), and also because the basin's state (e.g. regarding land cover) is likely to have changed to some degree during the simulation period, an exponential weighting scheme was employed. In order to detail the yearly weights for each error measure, it was arbitrarily defined that, under the exponential scheme, observations retrieved from 1998 to 2003 should transmit the same amount of information as those concerning the period from 1982 to 1997. The weights are illustrated in Figure 9.2.
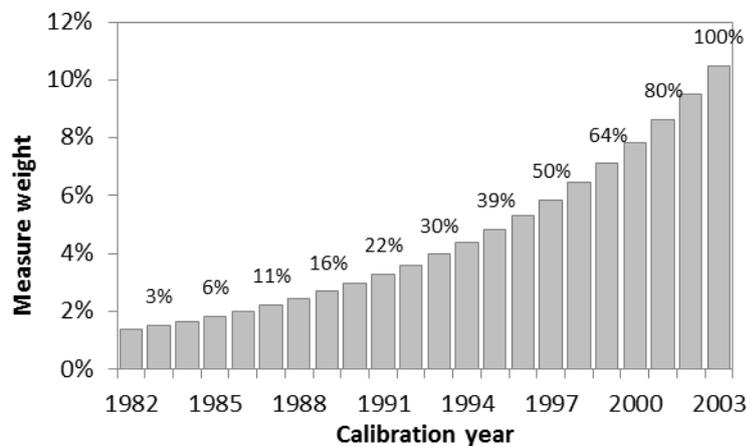

Figure 9.2. Exponential weighting scheme prioritizing the more recent observations. Cumulative weight valued on top of every other column.

The time-weighting of error measures, however, should be applied carefully as it does not work necessarily well with all error measures. For example, a time weighted NE coefficient can lead to a substantially distorted evaluation of the simulation's fitness, depending on the shape of the yearly hydrographs.

The objective for the SO optimization was the maximization of the NE error measure. As highlighted before, there are some difficulties when the NE coefficient from several series is aggregated due to a disproportionate weight of poor performances. Due to this, it was not used in the MO calibrations – which coincidently were carried out for the whole basin.

The MO calibrations were based on two main objectives. Of these, one represents the fitness of reach discharge and water level series, while the other reflects the fitness of the series of stored volumes at the main reservoirs.

For the reaches, the defined objective has been defined as the minimization of the weighted RMSE. The weighting accounted for the amount of available records in the observed series and included the aforementioned exponential scheme for prioritizing more recent observations. In reservoirs, due to the large differences in the stored volumes between the Malawi Lake and the Kafue Gorge Dam's reservoir, the minimization of an adimensional error measure was chosen: the IRRMSE.

Additionally, a third objective, used in two of the calibration set-ups, aimed at the minimization of a parameter heterogeneity measure. This measure is detailed in Section 9.5.2, below.

9.3   Global calibration

9.3.1   Endeavoring a multi-objective global calibration

When a global parameter calibration is undertaken the concept of equifinality (Beven 1993), discussed in Section 3.3, should be kept in mind, particularly due to the heterogeneity of the ZRB.

In fact, it is unlikely that, considering the capabilities of current hydrological models, a single parameter set can reproduce the entire range of hydrological responses of the ZRB. Due to this, no global SO optimization was carried out. Instead, a MO optimization similar to the one pursued by Cohen Liechti (2013) was done.

The optimization was carried out for the parsimonious and the detailed models. Most of the parameters in Table 9.1 were allowed to change, except for the last four, related to the uncontrolled outflow of the major waterbodies.

The AMALGAM algorithm was employed aiming at the convergence of an ensemble of 40 parameter sets to a Pareto surface spanned by two objectives. These objectives were the minimization of the weighted RMSE in reaches and the minimization of the weighted IRRMSE in the monitored waterbodies. Simulations were run for both models over the course of several days past the completion of 150 iterations.

The topology of the parsimonious and detailed model's, along with the location of observed time-series can be seen in Appendix I.D and Appendix I.E.

9.3.2   Results and discussion

The first element in the analysis should be the representation of the objective space. For both models, the values of reaches' RMSE and reservoirs' IRRMSE are plotted in Figure 9.3 and in Figure 9.4, respectively. These refer to the calibration period, from 1982 to 2003.
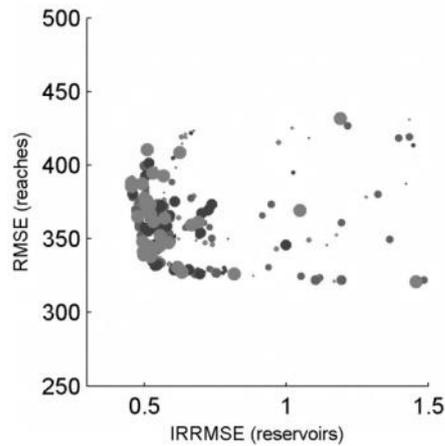
Figure 9.3. Final objective space for the global calibration of the parsimonious model. Results based on the 10 last iterations; the latest results are displayed above and with larger radius
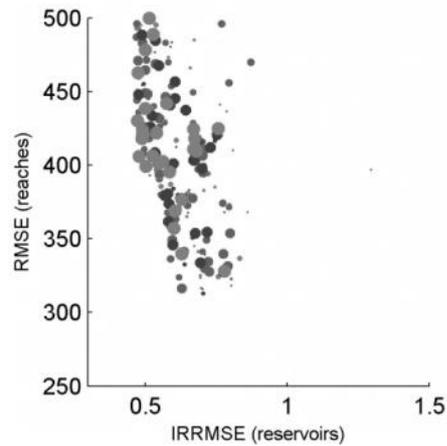

Figure 9.4. Final objective space for the global calibration of the detailed model. Results based on the 10 last iterations; the latest results are displayed above and with larger radius.

Several aspects are worthy of notice regarding the illustrated objective spaces. Firstly, a substantial scattering of the solutions can be seen. This is undesirable and could possibly be avoided by: 1) employing larger ensembles and 2) allowing the optimization to run for longer. Particularly the first option could be worth pursuing if computational demands were to be disregarded. Despite this limitation, valuable information can be derived from the figures.

Most interestingly, it can be clearly seen that for both models – with emphasis on the parsimonious one – there is a trade-off between the performance at the reservoirs and at the reaches.

Because the models' parameters are global, trade-offs can also be observed between some calibration locations. In Figure 9.5, the RMSE obtained for several global parameter sets of two close calibration points (Senanga and Victoria Falls) can be inspected. It appears that low RMSE's at Senanga are a necessary condition for low RMSE's at Victoria Falls. Furthermore, there is a positive correlation among most of the depicted parameter sets between the two locations. This is not surprising as the locations placed nearby on the main reach of the Zambezi. An example of a clear trade-off if given in Figure 9.6. There, the RMSE's at Kafue Gorge and Victoria Falls are shown. While Victoria Falls is on the Upper Zambezi, Kafue Gorge lays on the Kafue subbasin. It appears that the parameter sets that produce the best RMSE's at Victoria Falls are not the ones that perform best for Kafue Gorge.
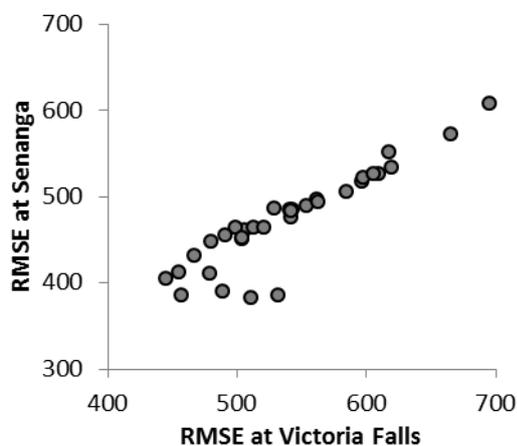

Figure 9.5. Illustration of a positive correlation among calibration points. Error measures from the global calibration of the parsimonious model.
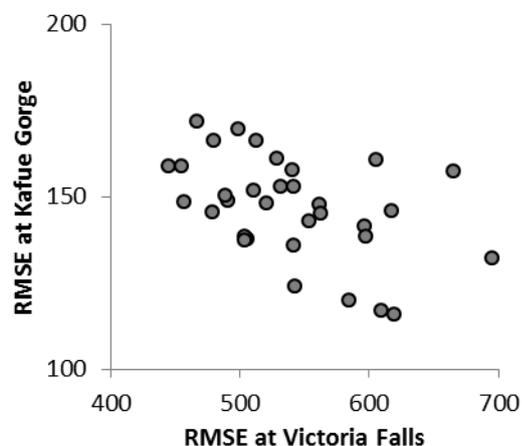

Figure 9.6. Illustration of a negative correlation among calibration points. Error measures from the global calibration of the parsimonious model.

Focusing on the differences between parsimonious and detailed models, it can be seen that:

- There is a larger scattering of the detailed model's objectives. The explanation for this effect is not evident. The "walk" of the optimization algorithm is partly random and, simply, the scattering might have been dictated by chance in a less successful calibration run. Additionally, because the detailed model includes a finer representation of forcing variables and subbasin heterogeneity, it is possible that the solution space spanned by it is rougher than that of the parsimonious model. If so this would contribute to a solution space harder to search by the optimization algorithm.
- The best reach RMSE is smaller for the detailed model. This hints that the application of detailed models can be worth pursuing.
- The reservoir's objective, IRRMSE, appears to perform better for the parsimonious model. Somewhat conflicting with the previous remark, this can be attributed to a less accomplished optimization run of the detailed model or, alternatively on topological or input inadequacies.
- The bounding Pareto surface for both models differs appreciably. The distinct Pareto surfaces appear to indicate that the models have different strong and weak points. Notwithstanding, in order to support this claim it would be desirable to run several more simulations, if possible, for longer and using larger ensembles.

Another approach to the analysis of the calibration results is the inspection of error measures at different locations. In Table 9.2 and in Table 9.3 the selected error measures are presented for six discharge, one water level and three volume series. Unlike the previous, the following results were computed concerning the validation period (2004 to 2009).

Table 9.2. Summarized validation results of the parsimonious model's global calibration. Best results among the ensemble. For VR the variation is presented.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	316.3	504.6	0.079	0.125	0.55	0.98 - 2.26
Senanga	236.6	305.2	0.087	0.112	0.85	0.62 - 1.37
Kafue Hook Bridge	96.1	168.9	0.052	0.091	0.71	0.04 - 0.85
Great East Rd. Bridge	383.9	819.8	0.056	0.119	0.23	0.04 - 0.39
Kafue Gorge	86.9	128.2	0.082	0.121	0.55	0.15 - 1.2
Victoria Falls	345.6	523.4	0.063	0.095	0.80	0.61 - 1.05
Marromeu (level)	0.592	0.831	0.107	0.151	0.11	0.98 - 1.05
Kariba	4.52E+09	5.55E+09	0.108	0.133	0.73	0.86 - 1.2
Cahora Bassa	2.76E+09	3.49E+09	0.224	0.283	-0.41	0.6 - 1.17
Malawi Lake	6.02E+09	7.24E+09	0.125	0.151	0.65	1 - 1

Table 9.3. Summarized validation results of the detailed model's global calibration. Best results among the ensemble. For VR the variation is presented.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	229.3	416.6	0.057	0.104	0.69	0.68 - 1.26
Senanga	288.0	411.4	0.106	0.151	0.73	0.47 - 1
Kafue Hook Bridge	120.2	211.9	0.065	0.115	0.54	0.26 - 0.63
Great East Rd. Bridge	351.4	669.4	0.051	0.097	0.49	1.05 - 1.99
Kafue Gorge	102.6	147.2	0.097	0.139	0.41	0.46 - 0.93
Victoria Falls	393.4	651.7	0.071	0.118	0.70	0.48 - 1.04
Marromeu (level)	0.542	0.777	0.098	0.141	0.22	1 - 1.13
Kariba	3.24E+09	4.02E+09	0.077	0.096	0.86	0.79 - 1.11
Cahora Bassa	1.55E+10	1.76E+10	1.254	1.428	-34.98	1.3 - 1.57
Malawi Lake	5.70E+09	6.87E+09	0.119	0.143	0.68	1 - 1

Of the different measures, perhaps the adimensional ones (particularly AIRAD, IRRMSE and NE) are of easier interpretation. From AIRAD and IRRMSE values below one it can be seen that the simulation errors are inferior to the observed series' variation.

Neither of the models dominates the other. Notwithstanding, despite being worse in some respects, the parsimonious model can overall be considered as better.

A more thorough evaluation can be achieved by directly assessing the simulated series. Notwithstanding, too many of them have been produced in order to allow for an individual inspection (40 series at each calibration location for both models). Below, solely examples for Victoria Falls and the Kariba dam are illustrated. In Appendix I.H and Appendix I.I additional information can be found.

Figure 9.7 refers to simulations produced for Victoria Falls by the parsimonious model. Observed historical flows can be compared to the simulations which attained the best reaches' objective (minimum weighted RMSE) and the best reservoirs' objective (minimum weighted IRRMSE) among the ensemble[4]. Also, a best local simulation according to a NE coefficient criterion is depicted. Finally, on the background, values falling between the 25 and 75% quantiles of the ensemble are shadowed. Two vertical dashed lines mark the transition between POM interpolated and TMPA 3B42 v7a rainfall inputs (1998) and between calibration and validation periods (2004). Figure 9.8 shows an equivalent representation of the detailed model's results.
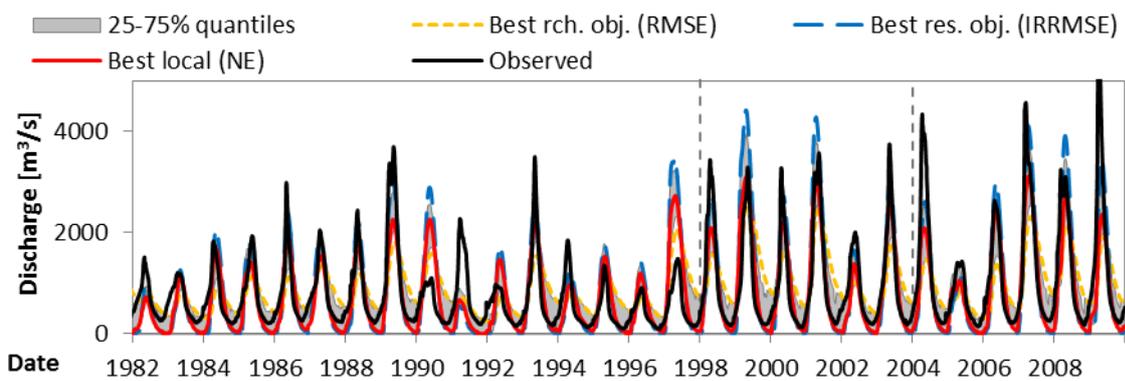

Figure 9.7. Synthesis hydrographs of the global calibration of the parsimonious model at Victoria Falls.
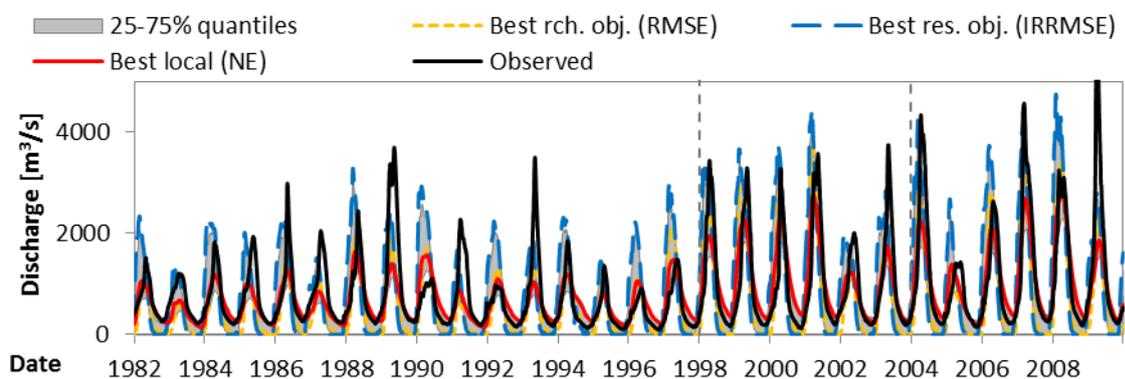

Figure 9.8. Synthesis hydrographs of the global calibration of the detailed model at Victoria Falls.

While both ensembles seem adequate – particularly so for the most recent years – performance is still far from perfect. The simulations corresponding to the best objectives for reaches and reservoirs are particularly revealing. For the parsimonious model, the optimization of the reaches' objective led to the overestimation of low flows and a substantial departure from the observations. For the

---

[4] Because these objectives are computed aggregating all calibration series, they do not necessarily correspond to satisfactory simulation at particular locations.

detailed model the, optimizations of both objectives also compromised the accuracy of the corresponding simulations, albeit in a different way.

Victoria Falls is upstream of the Kariba Dam. Due to this, it is interesting to inspect results at that location. In Figure 9.9 they are presented for the case of the parsimonious model.
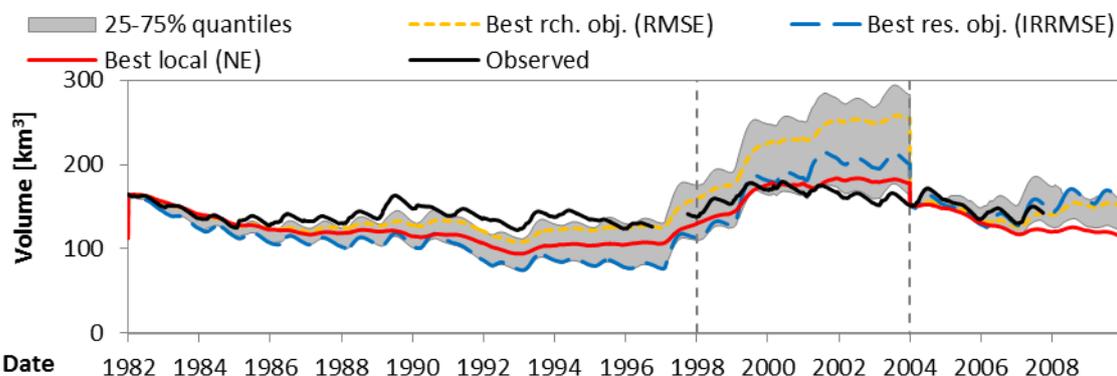

Figure 9.9. Synthesis volume series of the global calibration of the parsimonious model at the Kariba Dam.

At first glance, the simulations of the volume stored in the reservoir are worse than their discharge counterparts at Victoria Falls, evidencing larger relative deviations. Such an interpretation should, however, account for the fact that the volume series is representing an accumulated discharge error. As a consequence, obtaining accurate volume series is extremely difficult over long periods – as is the case – being that even small bias in the inflow simulations are translated into substantial volume differences.

During simulation, the volumes of the monitored waterbodies have been reset at the beginning at 2004, coinciding with the beginning of the validation period, in order for it to be independent from errors accumulated in the preceding years.

Overall, the results of the parsimonious model are very encouraging. Looking at the validation period in particular, performances are excellent. The results obtained for the detailed model, presented in Figure 9.10, differ from the latter substantially (notice that the vertical axis' scale is different between figures). Unlike before, there seems to be a tendency to underestimate volumes. Although the simulations corresponding to optimized reservoir objective and the local best NE depict seasonality better than their parsimonious counterparts, the simulation corresponding to the overall best reaches objective shows an alarming departure from the observations. This, in fact, hints at another case of a marked trade-off.
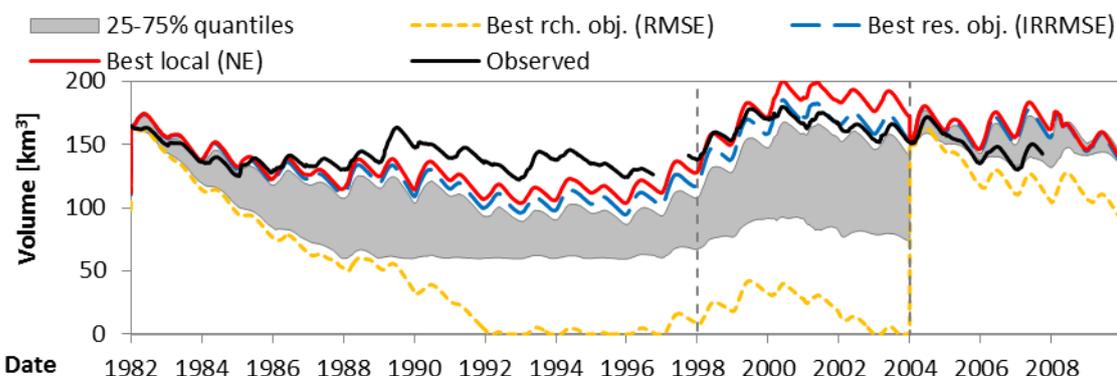

Figure 9.10. Synthesis volume series of the global calibration of the detailed model at the Kariba Dam.

Responsible for the diversity among simulations are, of course, SWAT's parameter values. An overview of their influence over the results can be gained by inspecting their distributions on the final optimization ensembles. Also, it is interesting to address how these distributions differ from the parsimonious to the detailed model.

Two examples are particularly useful for the discussion. The CN2 parameter is, as will be shown in Section 9.5.3, one with great impact on the model's results. Its distribution in the final ensembles is presented in Figure 9.11, in the case of the parsimonious model, and in Figure 9.12, in the case of the detailed one. In both cases there is a fairly spread distribution of the parameter around 0.8 / 0.9[5]. There seems, therefore, to be some coherence between the parsimonious and the detailed model: the parameters assume similar values and, furthermore, their spread is likely to be needed to approximate the desired Pareto surface.
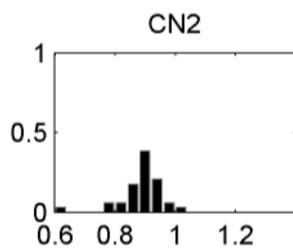

Figure 9.11. Distribution of CN2 values among the ensemble for the last iteration of the global calibration of the parsimonious model. Histogram normalized according to ensemble size (40).
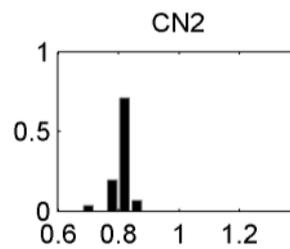
Figure 9.12. Distribution of CN2 values among the ensemble for the last iteration of the global calibration of the detailed model. Histogram normalized according to ensemble size (40).

Looking at the CANMX parameter (related to the maximum canopy storage), in Figure 9.13 and Figure 9.14, one can see that there is no agreement between ensembles. While for the parsimonious model there is little spread and values are high, for the detailed model the obtained CANMX values are distributed with a majority of low values[6].
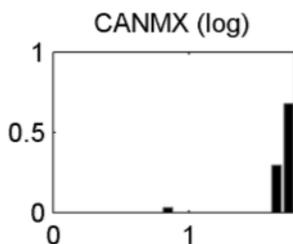

Figure 9.13. Distribution of CANMX values among the ensemble for the last iteration of the global calibration of the parsimonious model. Histogram normalized according to ensemble size (40).
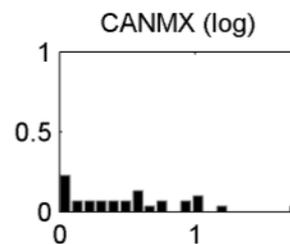
Figure 9.14. Distribution of CANMX values among the ensemble for the last iteration of the global calibration of the detailed model. Histogram normalized according to ensemble size (40).

The disparity between the representations of CANMX values in the two ensembles can be attributed to several factors. The differences inherent to the parsimonious and detailed models might have a significant contribution. Also, the random component of the optimization can be responsible for a "settlement" of the ensemble in different regions of the solution space. In addition to this, if the impact of CANMX on the objectives is low, the random behavior dominates. Finally,

---

[5] Notice that, as specified in Table 9.1, CN2's scale is relative. The original curve number values – defined for particular HRU units by ArcSWAT – are simply adjusted by a factor equal to the optimized parameter.
[6] The differences are even more marked than the figures hint, as the CANMX values there were applied the log transform.

CANMX appears to be poorly implemented in SWAT (this was discussed in Section 7.6), which might have led to this apparent incoherence.

A lot more could be stated about the global calibration run. In this sense, detailed results for the global calibrations of both models are presented in Appendix I.H and Appendix I.I and are well worth inspecting. Within them, detailed numerical calibration and validation results can be found. Also, synthesis series for an extended set of calibration points are illustrated and histograms containing the final distribution of the optimized parameters are shown.

9.4 Cascading calibration

9.4.1 Driver for a cascading calibration

In basin as large as the ZRB, the assumption of global parameters is likely to be one of the key drivers of model errors. A practical way of assessing this is to perform separate calibrations for different regions of the basin. Because downstream regions are not independent from the hydrological processes taking place upstream, a cascading calibration procedure was devised.

According to the proposed procedure, the calibration is first attempted for the upstream areas. Once satisfying parameter sets have been found, a series of simulated outflows is produced. Simulated outflows from upstream are then passed to downstream areas so that they, too, can be properly calibrated.

Each local calibration addresses not only the errors inherent to the representation of the hydrological processes within it, but also the errors eventually associated with the simulation of the upstream areas. This means that information is effectively being cascaded downstream.

There are several advantages to such an approach. Firstly, only the area under study needs to be simulated at each step of the cascade, which reduces substantially the computational demands. Also, no major changes need to be made to the optimization set-up[7].

Disadvantages can also be highlighted. For one, the consideration of a cascade renders the application of a MO optimization difficult. In fact, if it is acknowledged that no single best parameter set exists, but that a continuous Pareto surface should rather be considered, there is no single simulation that can be passed downstream, but an infinite number of optimal simulations. Along the cascade, each of those optimal simulations would span an equivalent Pareto surface, rapidly adding to the problem's complexity. Another disadvantage is that, since there is no regularization, very different parameters can be found for similar areas; this does not promote parsimony and is, therefore, undesirable.

9.4.2 Details of the cascading calibration

As before, both parsimonious and detailed models were optimized. The areas of the cascade were chosen in order to maximize the use of calibration data. This calibration, however, was not pursued down to the delta due to the poor quality of the simulations at Cahora Bassa and at the downstream part of the Shire subbasin. The cascade areas for both models are identified in Figure 9.15 and in Figure 9.16.

Mostly due to difficulties associated with a representative aggregation of the NE coefficient computed at different locations, the use of this measure was avoided in the MO basin-wide optimizations. Here, however, a cascade of SO optimizations using data from one or few locations is performed and, therefore, the same limitations to the use of the NE coefficient do not hold. The maximization of the NE was thus chosen as the objective to be pursued.

---

[7] In fact, all parameters can be regarded as "global" for each of the submodels contained in the cascade.
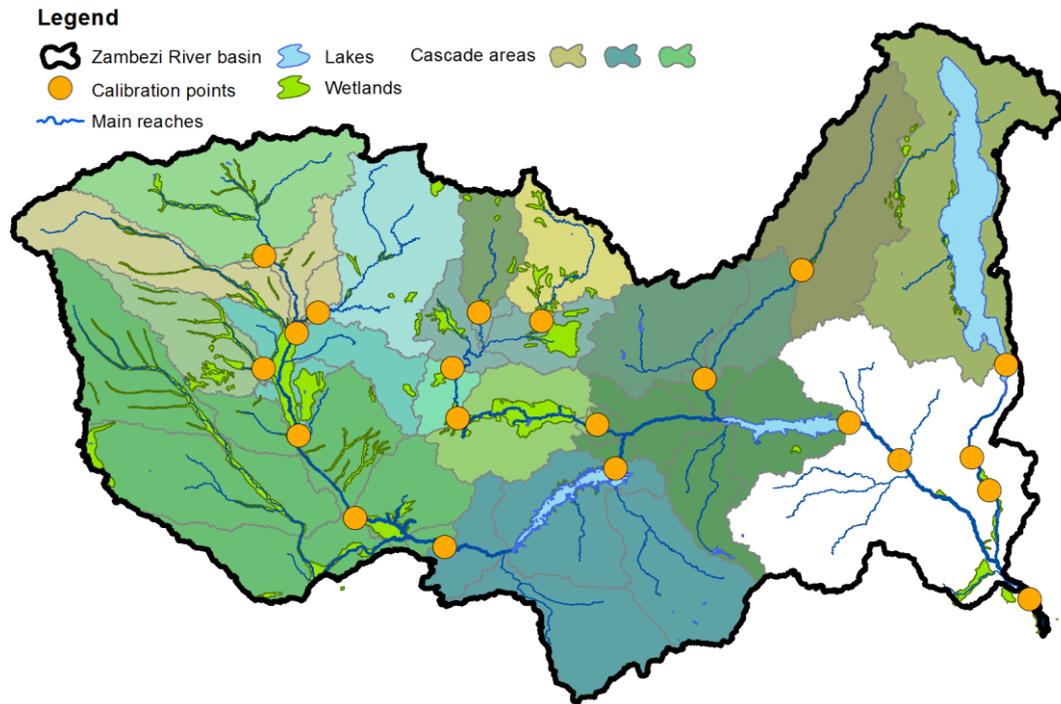
Figure 9.15. Areas considered in the cascading calibration for the parsimonious model.
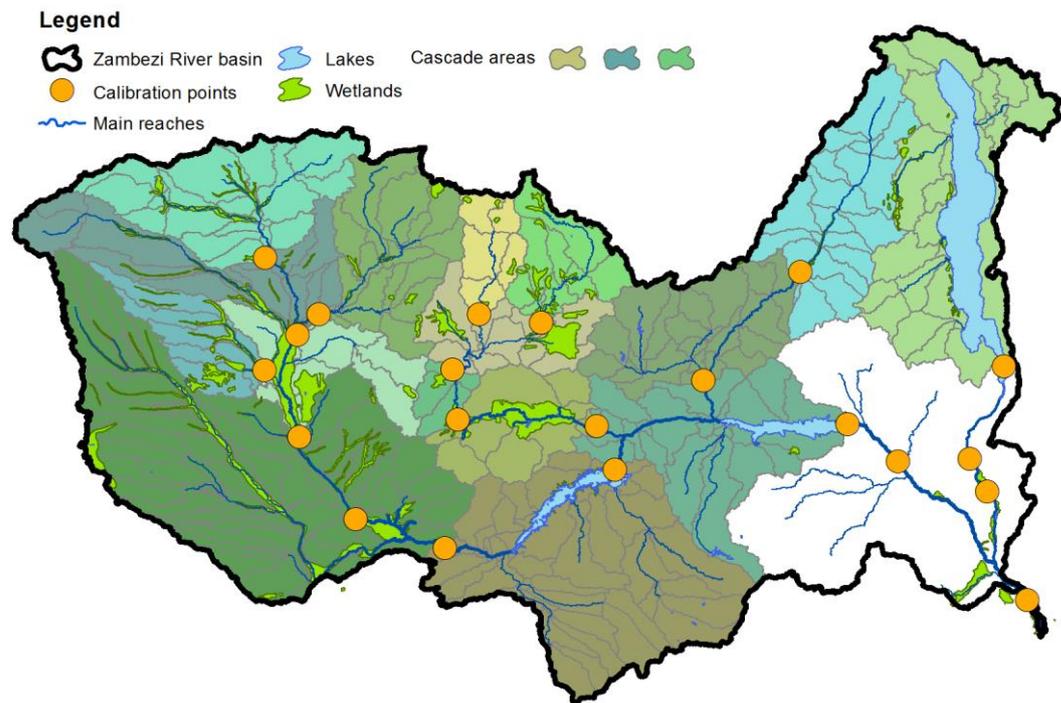

Figure 9.16. Areas considered in the cascading calibration for the detailed model.

Not being implemented at the time of the computations and neither being compatible with the NE measure, the time-weighting of calibration years was not employed here. Also, at this stage the observed outflows of Kafue Gorge were not used for model calibration, what made the calibration of the Kafue Flats dependent on the evaluation of the series of stored volumes at the Kafue Gorge Dam's relatively small reservoir. This proved to be very difficult for the optimization algorithm to cope with.

Finally, during the execution of the cascading calibration, the uncontrolled outflow parameters were being included in the optimization parameter set (obviously only in the basins where

uncontrolled reservoirs existed). Because of the bottlenecks introduced by poor calibration of these parameters, the optimization runs faced particularly challenging difficulties.

The process was carried out using the CMA-ES algorithm until the results stabilized and stopping criteria were met, usually between 150 and 400 iterations. The populations were, in most of the cases, fixed at 60 parameter sets.

9.4.3 Results and discussion

Due to the fact that an ensemble is not produced, an easier assessment of the cascading calibration performances is possible. Starting with the comparison of the validation error measures obtained at key points for the parsimonious (Table 9.4) and detailed (Table 9.5) models, one can immediately remark that: there are points showing very low performances (e.g. Kafue Gorge); and, again, the parsimonious model tends to be superior to the detailed one.

Table 9.4. Summarized validation results of the parsimonious model's cascading calibration.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	251.1	407.2	0.062	0.101	0.70	0.98
Senanga	170.8	289.1	0.063	0.106	0.87	0.91
Kafue Hook Bridge	100.1	149.6	0.054	0.081	0.77	1.17
Great East Rd. Bridge	614.8	1028.1	0.089	0.150	-0.21	2.00
Kafue Gorge	151.8	259.2	0.143	0.245	-0.83	1.41
Victoria Falls	365.6	641.9	0.066	0.117	0.70	0.93
Marromeu (level)	0.563	0.770	0.102	0.139	0.24	0.95
Kariba	3.57E+09	4.12E+09	0.085	0.098	0.85	1.00
Cahora Bassa	9.79E+09	1.16E+10	0.793	0.940	-14.62	1.09
Malawi Lake	2.40E+10	2.98E+10	0.500	0.620	-4.96	1.00

Table 9.5. Summarized validation results of the detailed model's cascading calibration.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	231.2	421.7	0.057	0.105	0.68	1.15
Senanga	252.5	347.3	0.093	0.128	0.81	1.12
Kafue Hook Bridge	132.3	198.0	0.072	0.107	0.60	1.39
Great East Rd. Bridge	522.9	908.3	0.076	0.132	0.05	1.84
Kafue Gorge	386.8	444.5	0.366	0.420	-4.39	2.40
Victoria Falls	367.6	606.1	0.067	0.110	0.74	1.03
Marromeu (level)	0.680	0.898	0.123	0.163	-0.03	0.92
Kariba	1.25E+10	1.57E+10	0.300	0.374	-1.13	1.07
Cahora Bassa	7.29E+10	8.16E+10	5.902	6.611	-770.78	2.45
Malawi Lake	6.96E+10	7.47E+10	1.448	1.555	-36.52	1.01

The detailed model, however, produced better results at some locations such as Victoria Falls (Figure 9.17), which are quite satisfactory.

Although not generally, in particular cases the cascade-calibrated parsimonious model obtained results which compare favorably to those of the global calibrations. This is mainly the example of upstream catchments such as Chavuma Falls, but also of the Kariba Dam, which is modeled with great accuracy in the last years of the simulation (Figure 9.18).

Particularly for the detailed model, the calibration was not successful for all subbasins. For example, the bottleneck problem induced by the inclusion of the uncontrolled reservoir parameters in the calibration led to an extremely poor representation of the Malawi Lake, as shown in Figure 9.19.
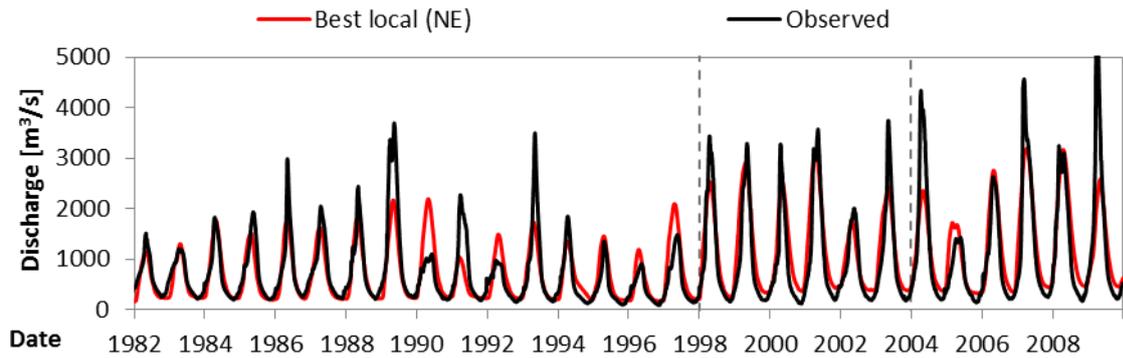
Figure 9.17. Observed and simulated hydrographs after the cascading calibration of the detailed model at Victoria Falls.
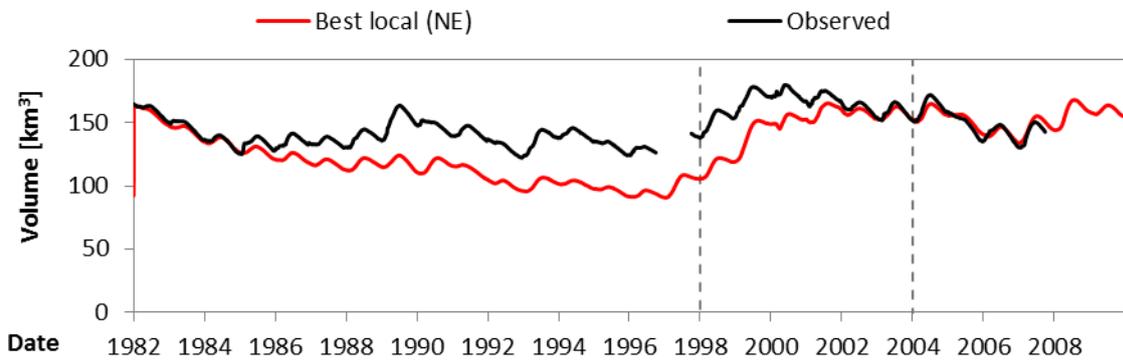

Figure 9.18. Observed and simulated volume series after the cascading calibration of the parsimonious model at the Kariba Dam.
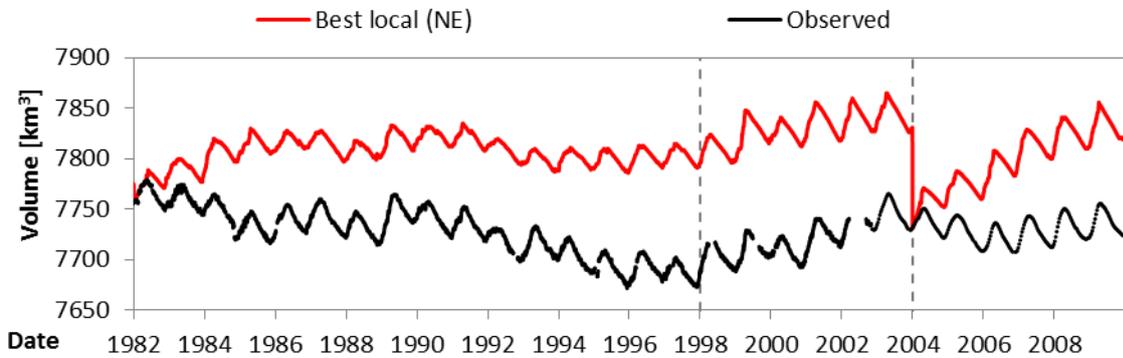

Figure 9.19. Observed and simulated stored volume series after the cascading calibration of the detailed model at the Malawi Lake.

Overall, it cannot be said that a cascading calibration such as the one proposed is superior to the global MO set-up. Notwithstanding, their comparison also highlights that, as expected, the global calibration is not optimal for all calibration points. Detailed results for the cascading calibration can be found in Appendix I.J and Appendix I.K

9.5 Regional-regularized calibration

9.5.1 Driver for the regional-regularized calibration

The regional-regularized calibration is yet another attempt at addressing the shortcomings of the global calibration.

The cascading calibrations might owe their disappointing performances to an effective large number of free parameters (one "independent" set for each cascade area). Also, within the

cascade framework information flows only downstream, being that each area must cope alone with the simulation errors inherited from the preceding areas.

A regional-regularized calibration might alleviate both problems while still allowing for the parameters to represent the basin's heterogeneity.

9.5.2   Details of the regional-regularized calibration

The idea behind the regional-regularized calibration is to promote a MO optimization of parameters very similar to the one used for the global calibration, but where the more relevant parameters are allowed to change among pre-defined regions.

A regularization scheme was introduced in order to restrict the free variation of the regional parameters. By doing so, if equivalent performances can be obtained using similar parameter values for all regions, the algorithm will choose the most homogeneous solution.

As for the global calibration, the objectives of minimizing the weighted RMSE at the reaches and the weighted IRRMSE at the reservoirs were adopted. In addition, a third objective was introduced in order to regularize the ensemble. This objective – minimizing the sum of the coefficients of variation for each regional parameter – acts as a measure of the parameter set's heterogeneity.

By incorporating an objective of minimizing heterogeneity one can, on the one hand and as implied, enforce the regularization of the ensemble. Also, the global calibration conduced previously constitutes a particular case of the regional-regularized calibration where heterogeneity is equal to zero. In perfect conditions, therefore, by evaluating RMSE or IRRMSE against heterogeneity values one should be able to assess the benefits of the regional-regularized approach. Finally, augmenting the number of free parameters increases the chances of "overfitting" of the model. In general, increasing the number of free parameters holds the potential to improve calibration error measures; notwithstanding, past a certain point this might not be reflected in validation error measures, which may eventually be degraded. Controlling the heterogeneity of the solutions could be useful in order to balance calibration and validation performances even when the number of free parameters is great.

On the downside, the inclusion of a minimum heterogeneity objective adds one extra dimension to the Pareto surface being approximated and can, therefore, demand more computational resources (such as increasing ensemble size and number of iterations) in order for the optimization run to be successful.

Again, the AMALGAM optimization algorithm was employed. The number of iterations was over 150 and the number of ensemble members fixed at 160 parameter sets.

The 10 selected regions[8] are illustrated in Figure 9.20, for the parsimonious model, and in Figure 9.21, for the detailed model.

9.5.3   Promoting parsimoniousness: selecting the most relevant parameters

With 10 regions to be considered, it is not feasible to calibrate either model with all the selected parameters taking regional values. As such, it is paramount to select parameters to be regionalized by defining those which the models are more sensitive to.

Selecting the most relevant parameters is not straightforward. A decision based on subjective judgment would be hard to justify due to the complex interpretation of the model and the difficulty of distinguishing the effects attributable to each parameter. Effectively, the parameters' influence on the results has been shown to vary according to subbasin and even season (Reusser et al. 2011, Reusser and Zehe 2011, Cohen Liechti et al. 2014). By the other hand, the distributions of parameters obtained for different areas of the cascade at the end of the calibration (not

---

[8] The selection was based on the 13 main catchments of the ZRB presented in Chapter 2. The smallest subbasins were aggregated to larger ones in order to promote model parsomoniousness.

presented in this document) were of difficult interpretation and provided little insight; they varied greatly between models and among subbasins.
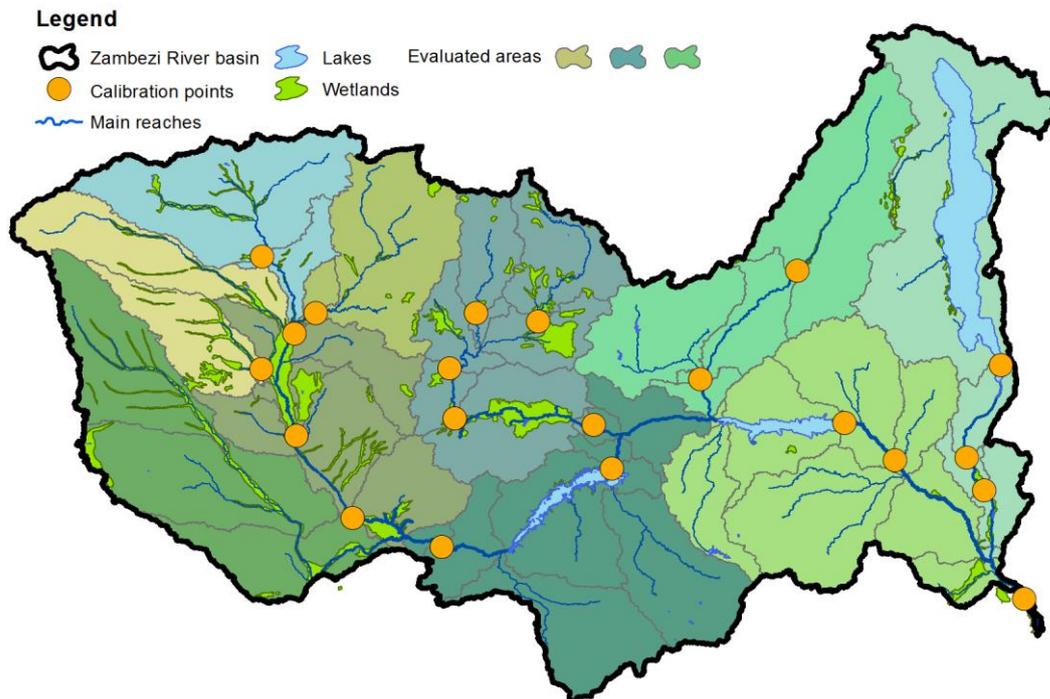

**Figure 9.20. Calibration regions for the parsimonious model.**
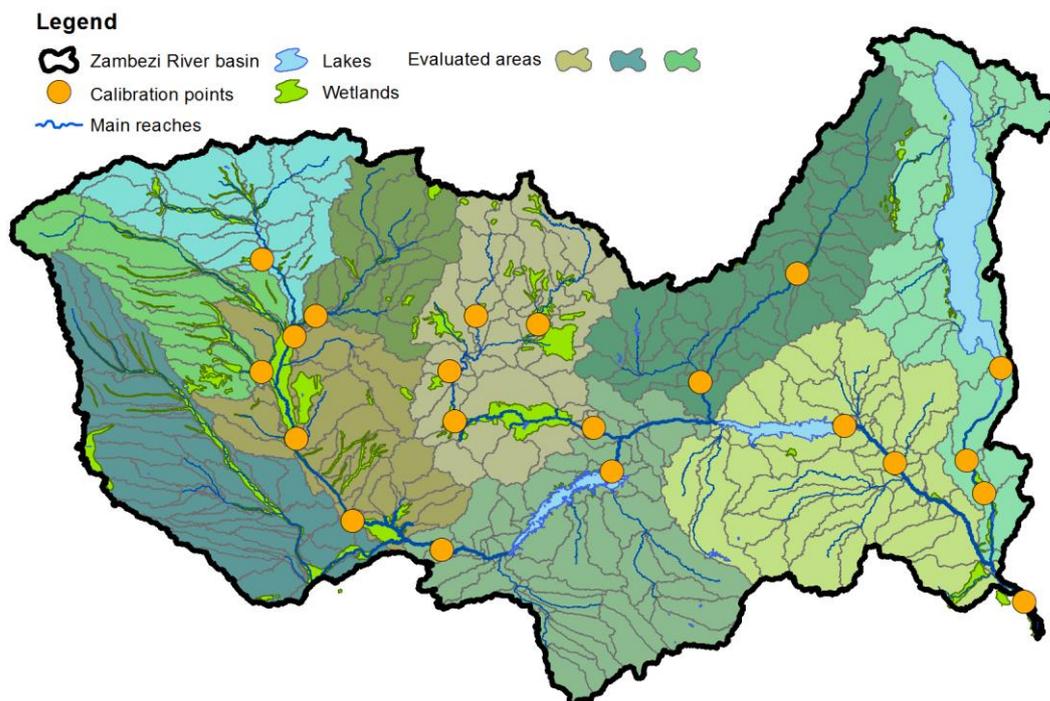
Figure 9.21. Calibration regions for the detailed model.

In order to make an informed decision about the most relevant parameters, a FAST was carried out for four selected locations with distinct hydrologic behaviors. These were Chavuma Falls, Kafue Hook Bridge, Great East Road Bridge and the Malawi Lake.

The FAST can be conducted in regard to specific outputs of the simulations. One common application is associating different combinations of model parameters to an objective or error measure (in the present case the NE coefficient). This analysis was undertaken.

Objectives or error measures transform a complex output into a single number. While this is extremely useful, it also leads to the loss of a considerable amount of information. This is naturally reflected in the FAST results.

Following earlier work (Reusser et al. 2011, Reusser and Zehe 2011, Cohen Liechti et al. 2014), a deeper insight into the sensitivity of each parameter was obtained by using FAST in order to evaluate directly model outputs at each time step. To this end, FAST was directly applied to daily simulated discharges encompassing the period from 1982 to 2003. Yearly results were then averaged in order to filter out noise.

The time averaged sensitivities obtained for Chavuma Falls are presented in Figure 9.22. As can be seen, the model's sensitivity to each parameter changes clearly throughout the year, being related to the discharge.
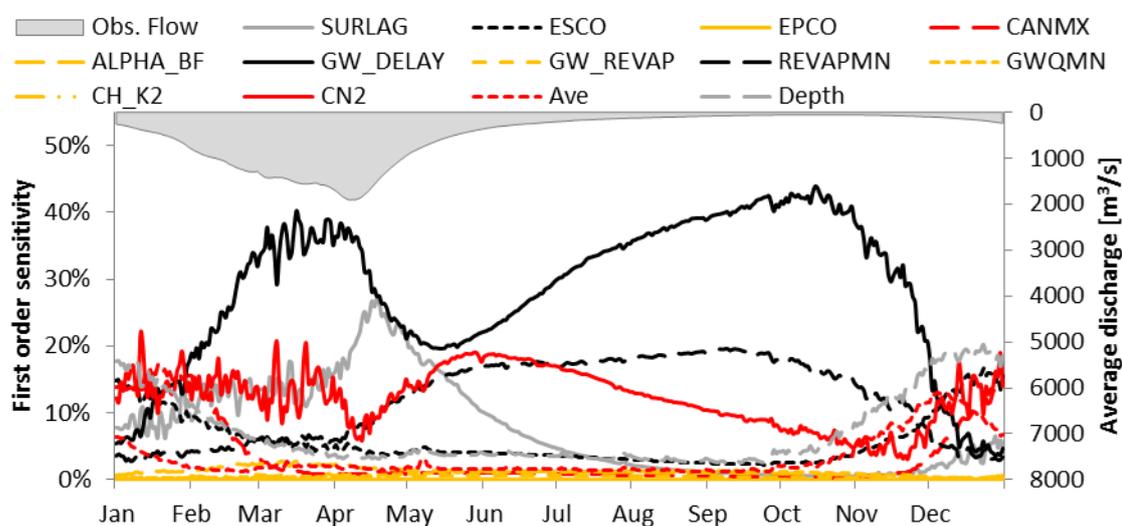

Figure 9.22. Results of the FAST applied directly to simulation results. Results of the parsimonious model for Chavuma Falls.

Such results are valuable in order to understand the hydrological model's behavior. For example, SURLAG (related to surface runoff) is only important in the wake of the rainy season and while surface flow is being produced. Conversely, during low flows the model displays more sensitivity to REVAPMN (which acts as a sink for groundwater).

Another interpretation can be made by normalizing the first order sensitivity index obtained by FAST by the average discharge. As can be observed in Figure 9.23, the time normalized sensitivity reveals that parameters which are predominantly reflected during low flows can ultimately have very little impact on the overall simulations.

The normalized results obtained for Kafue Hook Bridge and Great East Road Bridge are presented in Figure 9.24 and Figure 9.25, respectively. The results, similar to those of Chavuma Falls, differ mainly in terms of the sensitivity to SURLAG and ALPHA_BF (the groundwater recession constant).

In Appendix I.L the non-normalized results of the FAST for Kafue Hook Bridge and Great East Road Bridge can be found.

FAST results should, however, still be interpreted with care as they provide a "global" sensitivity test. The solution spaces spanned by hydrological models are well-known for their non-linearity. Non-linearity means that the global sensitivity to a given parameter might be totally different from local

sensitivity and, when calibrating a hydrological model, one is precisely settling in local optimal points.
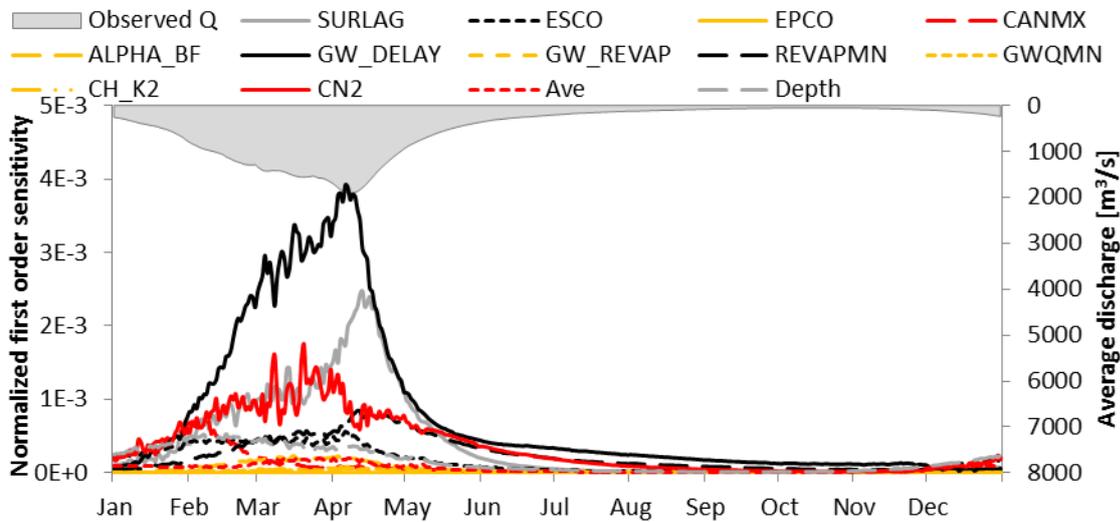

Figure 9.23. Results of the FAST applied directly to simulation results. Normalization of the first order sensitivity index by the average discharge. Results of the parsimonious model for Chavuma Falls.
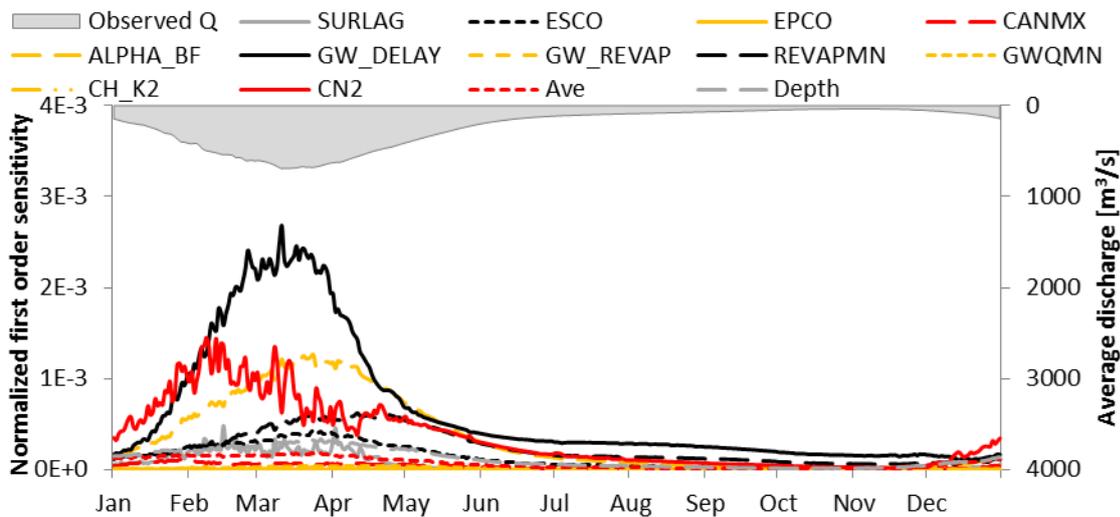

Figure 9.24. Results of the FAST applied directly to simulation results. Normalization of the first order sensitivity index by the average discharge. Results of the parsimonious model for Kafue Hook Bridge.

Because, in fact, most of the solution space might lead to inadequate simulations, modelers are usually interested in restricted regions of the solution space. In such cases, the sensitivity of the model to each parameter can change radically. In order to test and illustrate this, a FAST was run on a constrained region of the solution space in the vicinity of the optimal parameter set obtained for Chavuma Falls during the cascading calibration. The results, presented in Figure 9.26, show that local sensitivity differs completely from the global one – at least in this case.

Aggregated "global" findings using the three described FAST "modes" are presented from Figure 9.27 to Figure 9.29 for Chavuma Falls, Kafue Hook Bridge and Great East Road Bridge. Figure 9.30 represents the local values obtained for Chavuma Falls.
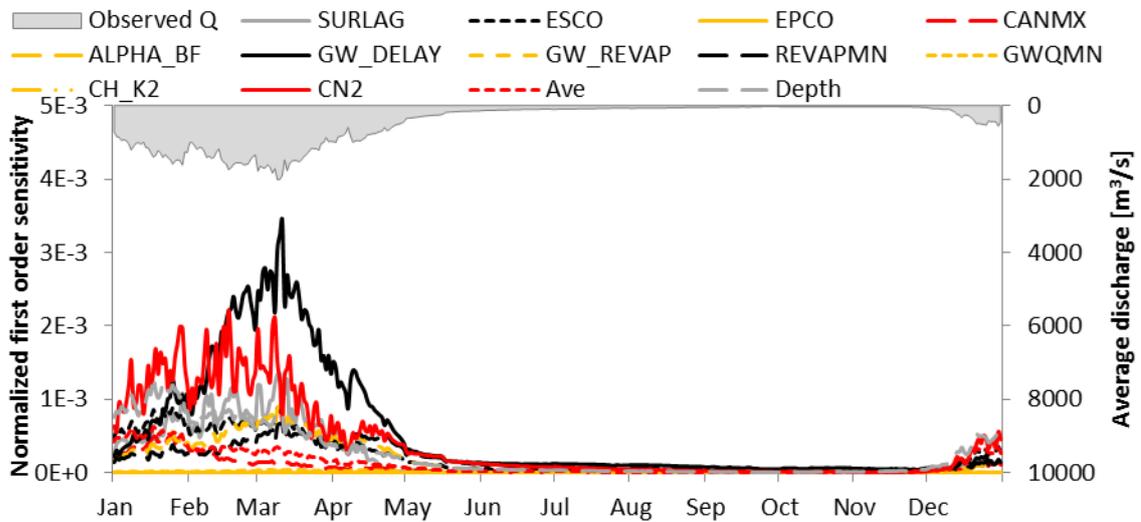
Figure 9.25. Results of the FAST applied directly to simulation results. Normalization of the first order sensitivity index by the average discharge. Results of the parsimonious model for Great East Rd. Bridge.
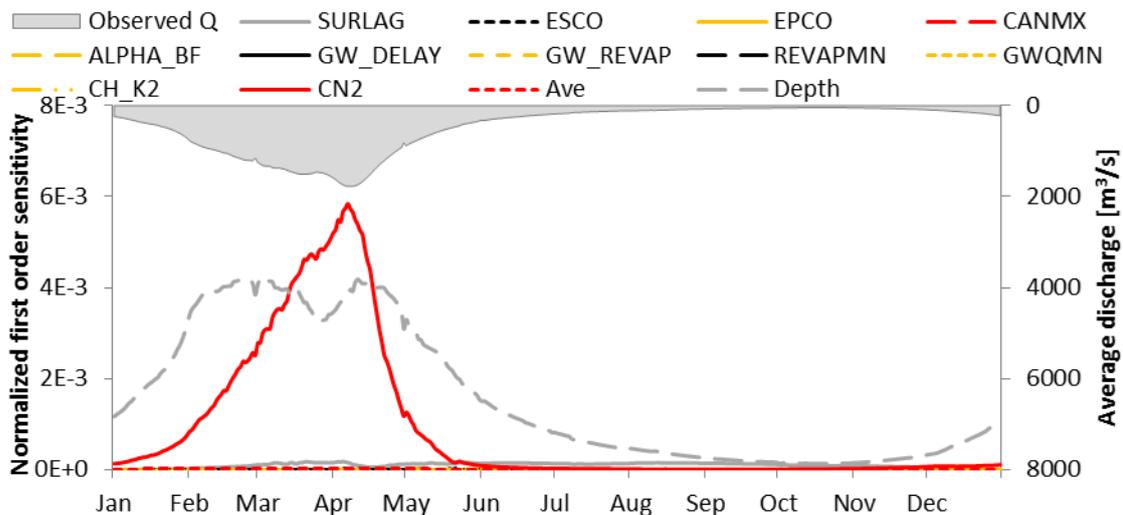

Figure 9.26. Results of the FAST applied directly to simulation results using parameters in the vicinity of an optimal solution. Normalization of the first order sensitivity index by the average discharge. Results of the parsimonious model for Chavuma Falls.

At the Malawi Lake a stored volume series is produced. Because such a series does not necessarily follow similar yearly cycles, the first order sensitivity is difficult to aggregate into a yearly basis. As a consequence, only the NE-based FAST was employed at that location. The results evidence a large sensitivity to CN2 and some of the reservoir parameters (ALPHA and K).

According to the global results, the most important parameters appear to be GW_DELAY, CN2, SURLAG and ALPHA_BF. As there are 10 regions, however, it is important to restrict as much as possible the number of regional parameters. As such, based on the obtained results and preliminary testing, only GW_DELAY, CN2 and SURLAG were selected.
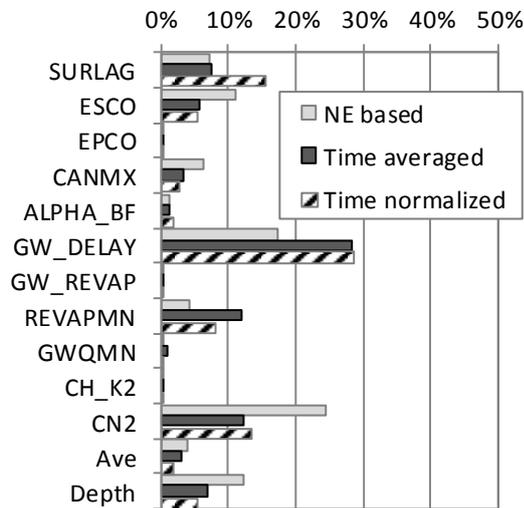
First order sensitivity index

Figure 9.27. Aggregated "global" FAST results for Chavuma Falls.
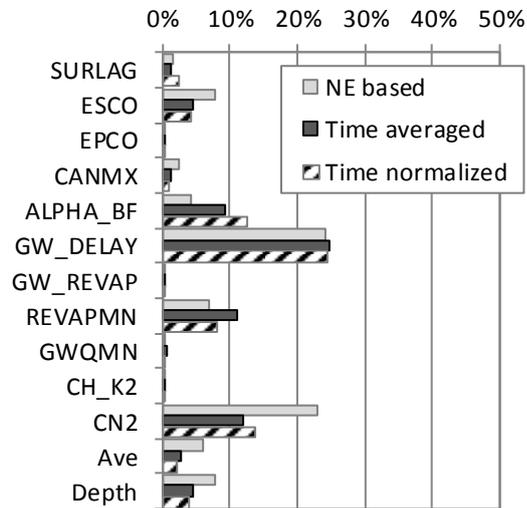
First order sensitivity index

Figure 9.28. Aggregated "global" FAST results for Kafue Hook Bridge.
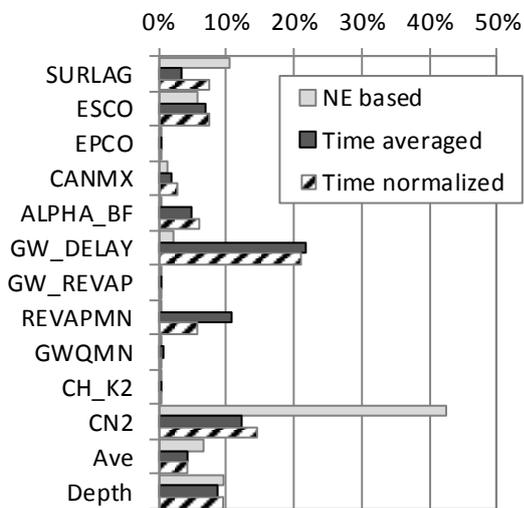
First order sensitivity index

Figure 9.29. Aggregated "global" FAST results for Great East Road Bridge.
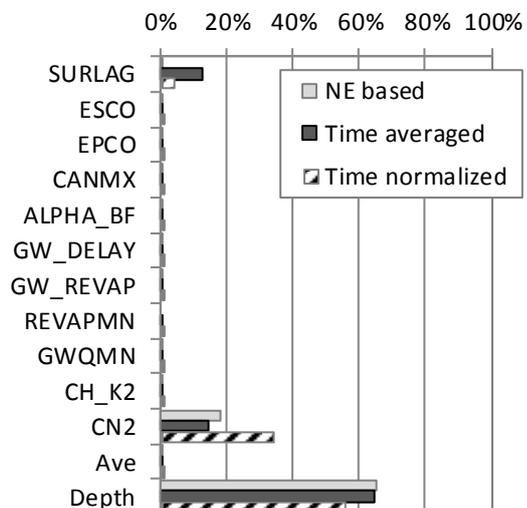
First order sensitivity index

**Figure 9.30. Aggregated local FAST results for Chavuma Falls..**
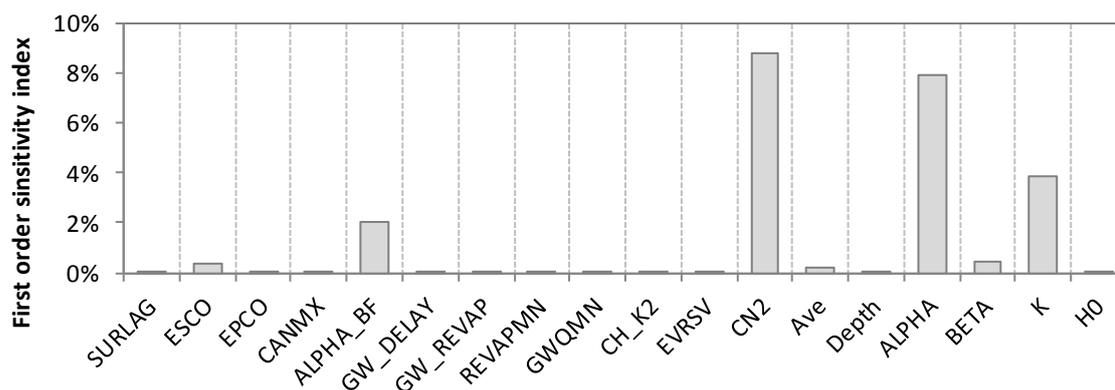
Figure 9.31. FAST results for the Malawi Lake (Based on the NE metric).

9.5.4   Results and discussion

The inspection of the optimized Pareto surfaces provides a first insight into the potential of the regional-regularized calibration. As now there are three objectives being optimized, the 3-D Pareto surface is depicted resorting to three 2-D projections.

The results obtained for the parsimonious model, displayed in Figure 9.32, can be compared with those of Figure 9.3 (page 156). Through such a comparison it can be seen that, most relevantly, the reservoir's objective is much better met using the regional approach and that the trade-off between reservoirs and reaches is not so evident (hinting that the trade-off is a result of the heterogeneity inherent to the basin, not reproduced by the global calibration).
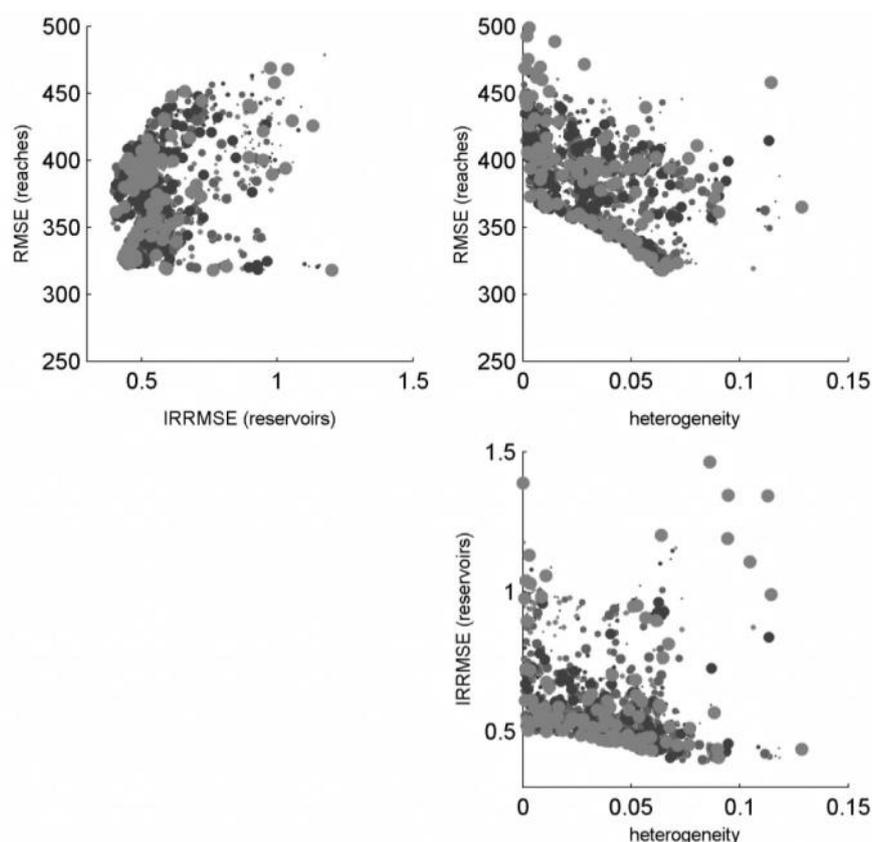

Figure 9.32. Final objective space for the regional-regularized calibration of the parsimonious model. Results based on the 10 last iterations; the latest results are displayed above and with larger radius

Accounting for heterogeneity it can be seen that both objectives show a tendency towards improvement as heterogeneity increases, which provides yet another hint at the value of a regionalized calibration for the ZRB.

Conclusions about the success of the AMALGAM optimization run can also be drawn. In fact, it has fallen short of its potential. This can be inferred resorting again to a comparison with the global calibration's results. In light of the previous discussion, were the present optimization to meet its full potential, some RMSE values corresponding to no heterogeneity would be less or equal to 320 (the best value obtained for the global calibration).

Similar conclusions can be made for the detailed model's results, presented in Figure 9.33. Here, however, the comparison to the global calibration's Pareto surface (Figure 9.4, page 156) is even more advantageous. Also worthy of notice is the fact that the weighted RMSE computed for reaches is best for the detailed model.
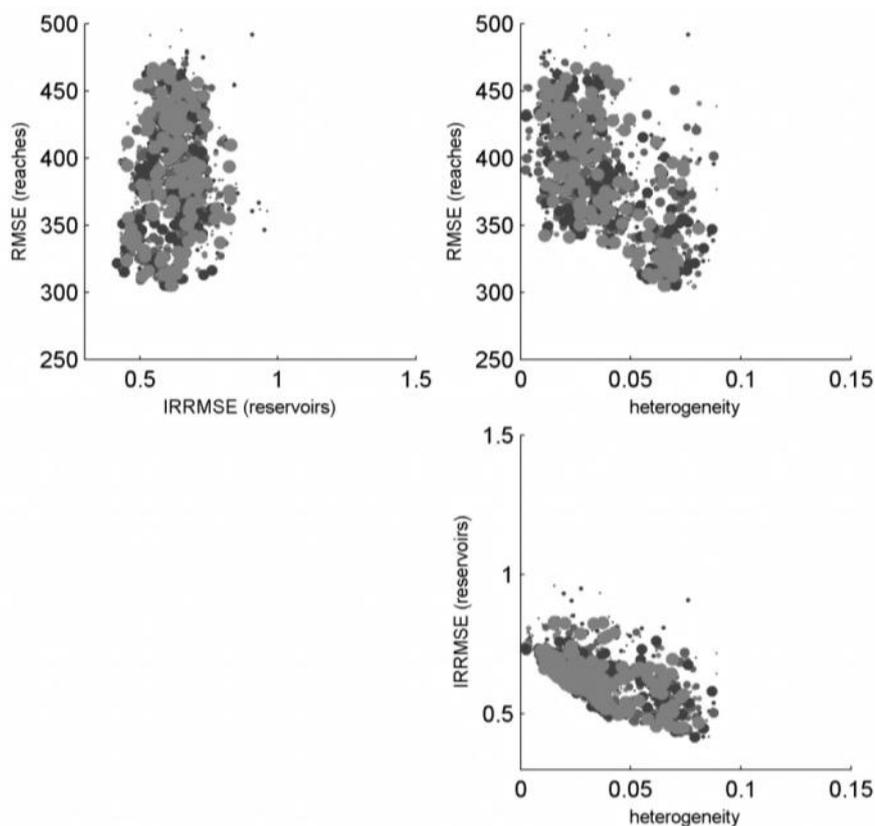
Figure 9.33. Final objective space for the regional-regularized calibration of the detailed model. Results based on the 10 last iterations; the latest results are displayed above and with larger radius

The top validation error measures at key calibration points present a different but revealing perspective. Again, comparing the parsimonious model results (Table 9.6) with those of the global calibration (Table 9.2, page 157), it can be seen that the global calibration's measures are vastly superior. Calibration error measures, presented in Appendix I.M, depict a similar relationship. In light of the obtained Pareto surfaces this is surprising. However counterintuitive, the observations can be easily explained.

Table 9.6. Summarized validation results of the parsimonious model's regional-regularized calibration. Best results among the ensemble. For VR the variation is presented.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	342.8	558.3	0.085	0.139	0.45	1.17 - 3.15
Senanga	338.8	436.5	0.124	0.160	0.70	0.64 - 3.19
Kafue Hook Bridge	113.8	187.5	0.062	0.101	0.64	0.03 - 8.94
Great East Rd. Bridge	331.8	669.2	0.048	0.097	0.49	0.03 - 6.28
Kafue Gorge	93.2	145.0	0.088	0.137	0.43	0.13 - 3.3
Victoria Falls	429.0	662.9	0.078	0.120	0.68	0.56 - 4.36
Marromeu (level)	0.609	0.809	0.110	0.146	0.16	0.88 - 1.21
Kariba	5.05E+09	5.88E+09	0.121	0.141	0.70	0.85 - 2.93
Cahora Bassa	3.61E+09	4.76E+09	0.293	0.385	-1.62	0.56 - 3.56
Malawi Lake	6.98E+09	8.70E+09	0.145	0.181	0.49	1 - 1.01

Firstly, Table 9.6 contains the summary of the best measures obtained for each calibration point while the corresponding Pareto surface is spanned by weighted error measures, which account for all the parameters at each point. What is being seen is then derivate from two factors implicitly acknowledged before: 1) the regional-regularized calibration fell short of its potential; and 2) the regional parameters increase the coherence of the model.

The first remark is supported by the fact that, being the global optimization a particular case of the regional-regularized one, the latter has the potential to produce equal or better results than the former within the ensemble.

The second remark is justified by the fact that, if in spite of showing worse performances locally, the regional-regularized calibration leads to better overall measures (the objectives are better met), it is because the models are more balanced among different calibration points. It is now evident that the global calibration produces good local results at the expense of the quality of the simulations at other parts of the basin.

The detailed model shows improvements to the parsimonious one in some of the calibration points. Exceptions are the Kafue region and the Cahora Bassa reservoir. Values are presented in Table 9.7.

Table 9.7. Summarized validation results of the detailed model's regional-regularized calibration. Best results among the ensemble. For VR the variation is presented.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	260.3	432.3	0.065	0.107	0.67	0.97 - 1.9
Senanga	265.1	418.6	0.097	0.154	0.72	0.61 - 1.31
Kafue Hook Bridge	193.9	304.6	0.105	0.165	0.05	0.07 - 0.27
Great East Rd. Bridge	339.2	626.7	0.049	0.091	0.55	0.93 - 1.92
Kafue Gorge	167.1	222.7	0.158	0.210	-0.35	0.18 - 0.48
Victoria Falls	330.4	609.1	0.060	0.111	0.73	0.52 - 1.33
Marromeu (level)	0.546	0.758	0.099	0.137	0.26	0.93 - 1.11
Kariba	2.43E+09	3.16E+09	0.058	0.076	0.91	0.82 - 1.19
Cahora Bassa	1.08E+10	1.21E+10	0.875	0.980	-15.96	1.21 - 1.54
Malawi Lake	5.85E+09	7.13E+09	0.122	0.148	0.66	1 - 1

Below, the synthesis of the optimized ensembles can be evaluated graphically for three selected sites. From the hydrographs of the parsimonious model at Victoria Falls (Figure 9.34), it can be seen that, although a wide ensemble is produced, covering well the observations, the local best hydrograph according to NE departs noticeably from the observed series during low flows. The ensemble spread, much wider than that of the global calibrations, is likely a consequence of the employed regularization.
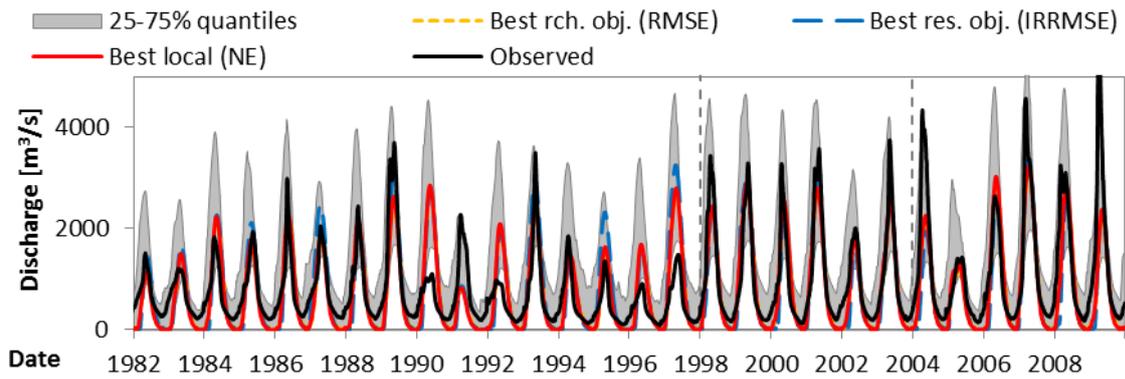
Figure 9.34. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Victoria Falls.

It has been noticed that the detailed model's calibration was not successful at all locations. As an example, the generated hydrographs at Kafue Gorge, depicted in Figure 9.35, could be improved.
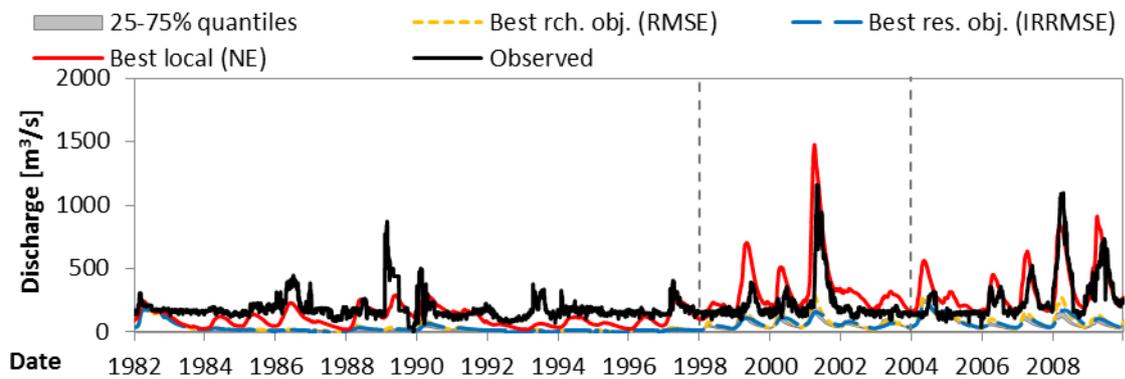
Figure 9.35. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Kafue Gorge.

Finally, looking at the stored volumes at Kariba, shown in Figure 9.36 and Figure 9.37 for the parsimonious and detailed models, it can be observed that the trade-off between the minimization of the reaches' weighted RMSE and the simulations performance at the reservoirs still remains for the detailed model. For the parsimonious model this is not evidently seen.
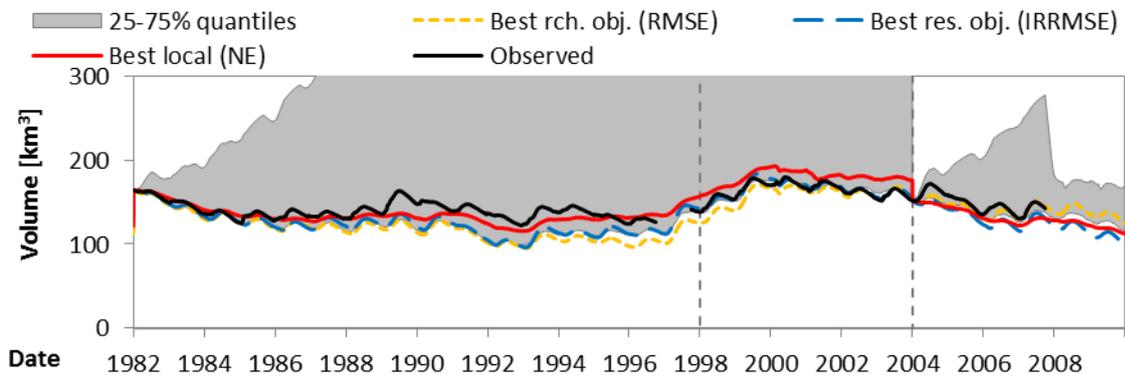
Figure 9.36. Synthesis stored volume series of the regional-regularized calibration of the parsimonious model at the Kariba Dam.
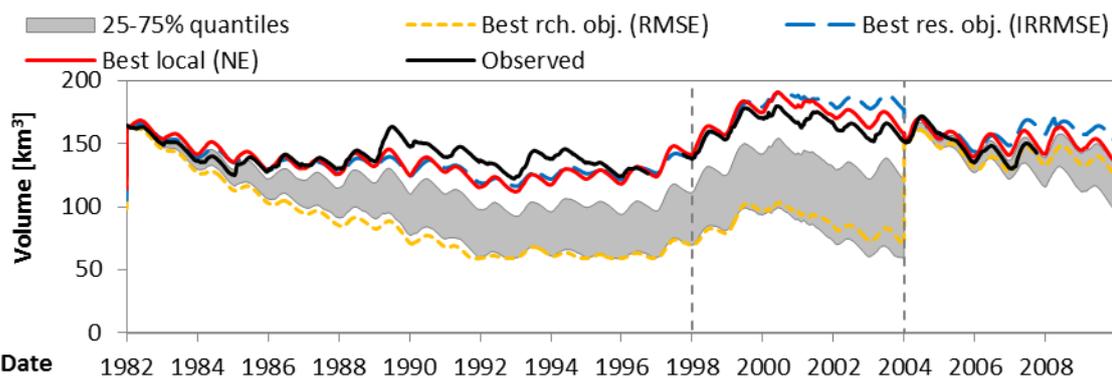
Figure 9.37. Synthesis stored volume series of the regional-regularized calibration of the detailed model at the Kariba Dam.

Regarding the parameter distributions within the final ensemble, there is an acknowledgeable similarity between the parsimonious and detailed models. The regionalized ones in particular (CN2, GW_DELAY and SURLAG) follow very close distributions. Parameters such as "Depth" (representing the depth from the surface to the bottom of the soil layer) and CANMX evidence differences. The distributions of "Depth" are reproduced in Figure 9.38 and Figure 9.39.
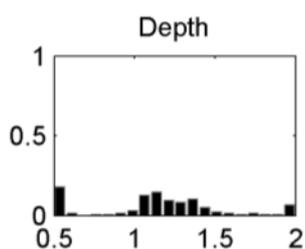

Figure 9.38. Distribution of Depth values among the ensemble. Last iteration of the regional-regularized calibration of the parsimonious model. Histogram normalized according to ensemble size (160).
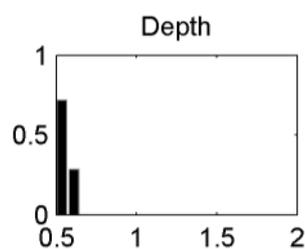

Figure 9.39. Distribution of Depth values among the ensemble. Last iteration of the regional-regularized calibration of the detailed model. Histogram normalized according to ensemble size (160).

Detailed results for the regional-regularized calibrations of both models are presented in Appendix I.M and Appendix I.N. Again, detailed numerical calibration and validation results can be found in the appendices along with synthesis series for an extended set of calibration points. Histograms illustrating the final distribution of the optimized parameters are also included.

9.6 Combining approaches with a cascading-regularized calibration

9.6.1 Driver for a combined approach

Until now all calibration set-ups displayed strong and weak points. Among them, the cascading calibration was particularly underexploited. The reasons for this have already been laid out and are related to the downstream sense of the information flow, sub-optimal solutions due to the inclusion of uncontrolled reservoir parameters in the set to be optimized and the "independence" between parameters in varying areas of the cascade. The last proposed calibration intends to address such issues through the implementation of a MO optimization procedure which, like the regional-regularized one, contains a regularization scheme.

9.6.2 Details of the regional-regularized calibration

The cascading-regularized calibration consisted of running a regional calibration over a ZRB model, whose base parameters were the ones inherited from the cascading calibration.

The idea is to avoid the use of global parameters altogether by: 1) considering a restricted set of regional parameters to be calibrated and regularized (CN2, GW_DELAY, and SURLAG); and 2) using the local cascading calibration's results in order to by-pass the simultaneous global calibration of the remaining.

If working desirably, such a scheme would benefit from locally defined parameters while maintaining coherence between the simulations at different calibration points and, through the regularization scheme, still avoid "overfitting".

The three selected parameters were calibrated using the AMALGAM optimization algorithm for the minimization of the three objectives defined for the regional-regularized calibration. The number of iterations was also over 150 and the number of ensemble members kept at 160 parameter sets. The regions were left unchanged.

9.6.3 Results and discussion

The computed Pareto surfaces for the parsimonious and detailed models are presented in Figure 9.40 and Figure 9.41.

The results from the parsimonious model are satisfying. Similarly to what has been evidenced in the previous calibrations, clear trade-offs are exhibited between objectives – with heterogeneity in particular.
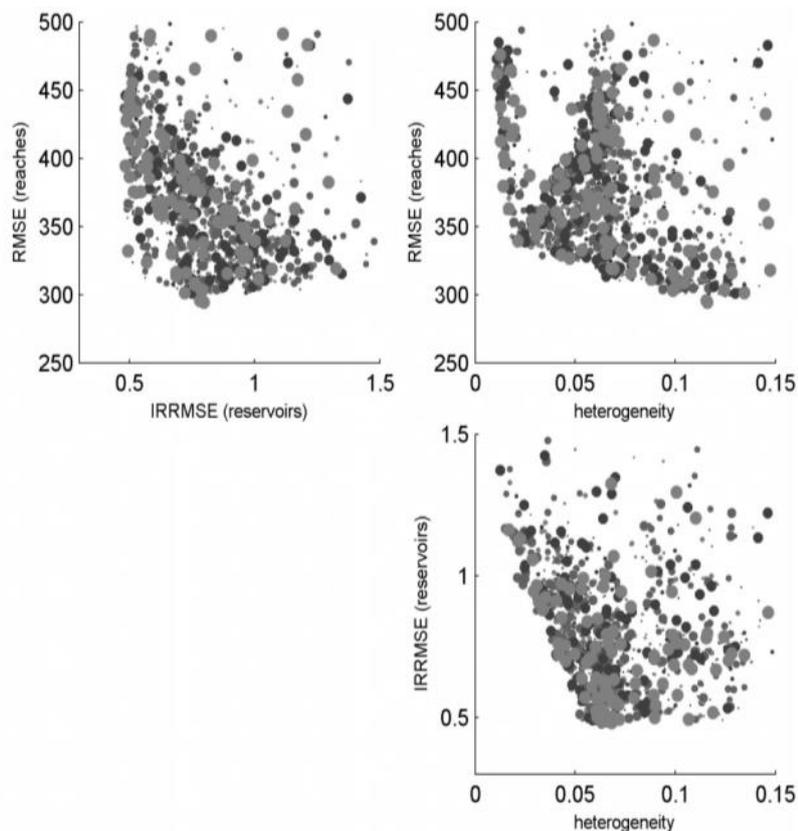
Figure 9.40. Final objective space for the cascading-regularized calibration of the parsimonious model. Results based on the 10 last iterations; the latest results are displayed above and with larger radius.

One interesting fact is that the present calibration approach yielded the best overall weighted RMSE for the reaches. Also worthy of notice is the relationship between the latter parameter and heterogeneity, whose trade-off displays two slopes. It is possible that this reflects the presence of

areas with markedly different behaviors and of areas that, albeit distinct, do not translate their heterogeneity in RMSE results so strongly.

Based on the information contained in Figure 9.41 it can be inferred that the detailed model's calibration failed, being the values obtained for reach and reservoir's objectives unexpectedly high.

The error measures presented in Table 9.8, referring to the validation period of the parsimonious model, reveal that it does not perform equally well across the basin. While the theoretical advantages of the cascading calibration have been put to good use at locations such as Chavuma Falls, error measures for the Kafue Hook Bridge or the Cahora Bassa dam are unsatisfying.

Table 9.9 is particularly revealing of why the detailed model's calibration performs poorly. In fact, the calibration does not appear to have failed throughout the whole ZRB. In a claim further supported by the full results contained in Appendix I.P, what caused the soaring climb of the objective functions were poor performances at particular points, namely the Malawi Lake[9] and the Kafue subbasin. The Malawi Lake in particular influences five downstream stations, three of which to a great extent: Mangochi, Chiromo and Chikwawa. All three evidence very poor performances. In parallel, volumes are greatly overestimated at Cahora Bassa. This is likely due excess inflows from the Luangwa subbasin (Great East Road. Bridge).
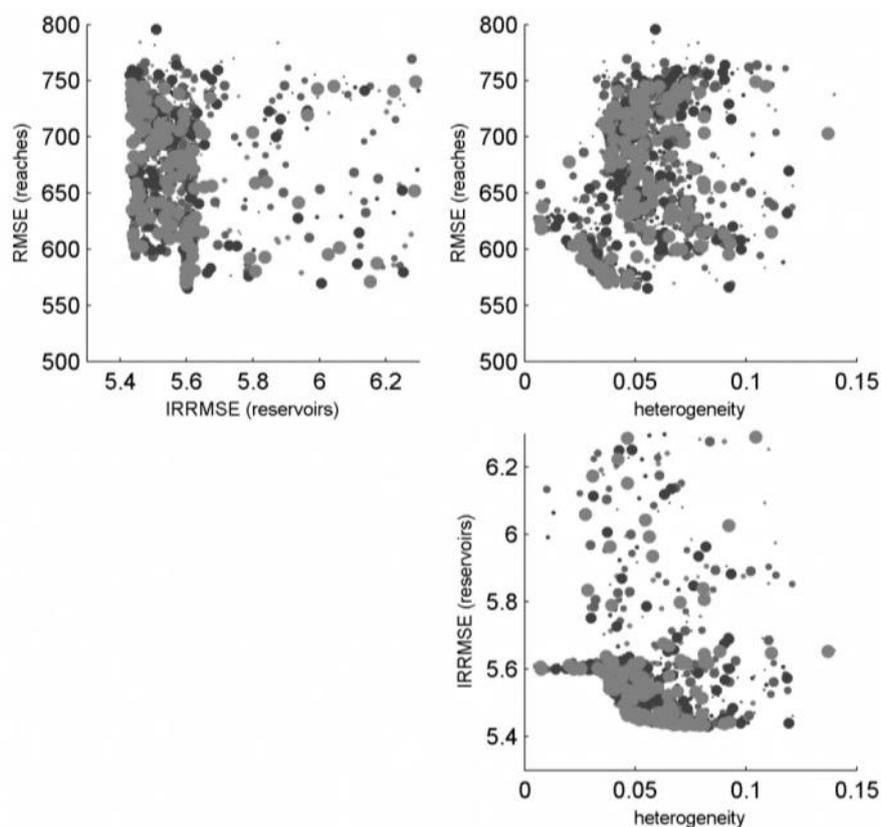

Figure 9.41. Final objective space for the cascading-regularized calibration of the detailed model. Results based on the 10 last iterations; the latest results are displayed above and with larger radius.

---

[9] The deviations at the Malawi Lake are partly owed to the decision of using the externally calibrated uncontrolled reservoir outflow parameters and not directly the ones computed within the cascading calibration. This can easily corrected in future implementations or, better yet, a cascading calibration that uses the fixed outflow parameters can be endevored.

Table 9.8. Summarized validation results of the parsimonious model's cascading-regularized calibration. Best results among the ensemble. For VR the variation is presented.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	237.9	410.1	0.059	0.102	0.70	0.13 - 1.59
Senanga	207.5	306.5	0.076	0.113	0.85	0.36 - 1.54
Kafue Hook Bridge	163.4	272.0	0.088	0.147	0.24	0.35 - 2.69
Great East Rd. Bridge	536.5	766.6	0.078	0.112	0.33	1.12 - 4.36
Kafue Gorge	100.4	144.6	0.095	0.137	0.43	0.77 - 1.3
Victoria Falls	314.7	571.3	0.057	0.104	0.77	0.39 - 1.89
Marromeu (level)	0.560	0.712	0.101	0.129	0.35	0.78 - 1.11
Kariba	2.93E+09	3.44E+09	0.070	0.082	0.90	0.73 - 1.4
Cahora Bassa	9.41E+09	1.14E+10	0.762	0.923	-14.03	0.22 - 1.54
Malawi Lake	1.53E+10	1.98E+10	0.319	0.411	-1.62	1 - 1

Two selected hydrograph synthesis can illustrate observations about the cascading-regularized calibrations. As an example of good agreement to the observations, the results of the parsimonious model at Victoria Falls are presented in Figure 9.42.

In Figure 9.43, regarding the outputs of the detailed model at Kafue Hook Bridge, a case is presented where simulations are not satisfactory. The stored volume deviations at the Malawi Lake are mostly inherited from the cascading calibration (refer to Figure 9.19, page 164).

Table 9.9. Summarized validation results of the detailed model's cascading-regularized calibration. Best results among the ensemble. For VR the variation is presented.

|	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	260.8	464.8	0.065	0.116	0.62	0.52 - 1.13
Senanga	216.9	314.5	0.080	0.115	0.84	0.56 - 1.23
Kafue Hook Bridge	148.6	258.8	0.080	0.140	0.31	0.57 - 0.69
Great East Rd. Bridge	421.5	695.8	0.061	0.101	0.44	1.05 - 2.09
Kafue Gorge	307.2	334.6	0.290	0.316	-2.05	2.02 - 2.14
Victoria Falls	367.4	593.6	0.067	0.108	0.75	0.48 - 1.29
Marromeu (level)	0.526	0.777	0.095	0.141	0.23	0.9 - 1.03
Kariba	4.59E+09	5.15E+09	0.110	0.123	0.77	0.79 - 1.25
Cahora Bassa	4.91E+10	5.65E+10	3.978	4.573	-368.23	1.98 - 2.26
Malawi Lake	5.92E+10	6.43E+10	1.233	1.338	-26.78	1.01 - 1.01
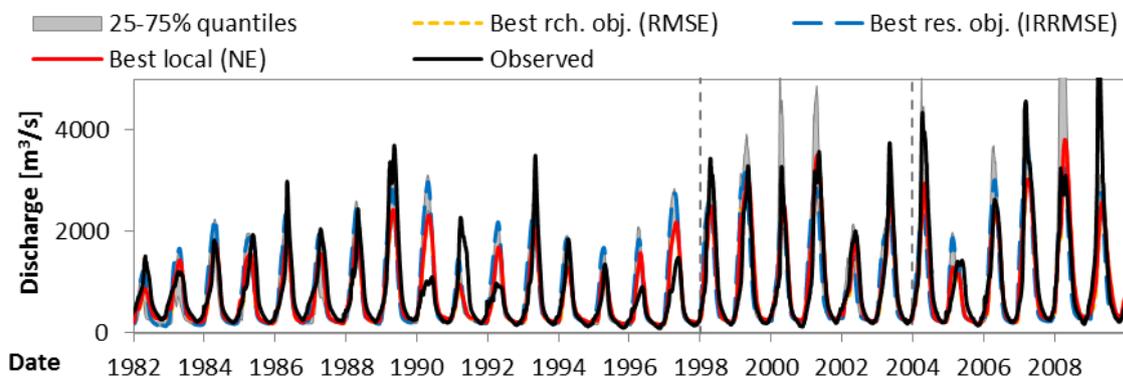
Figure 9.42. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Victoria Falls.
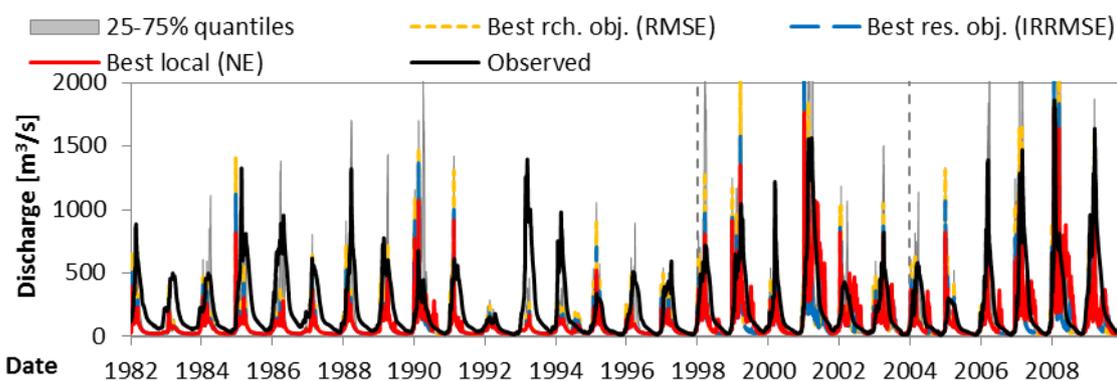
Figure 9.43. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Kafue Hook Bridge.

Comparing the optimized parameter's distributions, displayed in Figure 9.44 and Figure 9.45, it seems that, despite its poor objective values, the detailed model's parameters are more smoothly distributed amidst the ensemble.
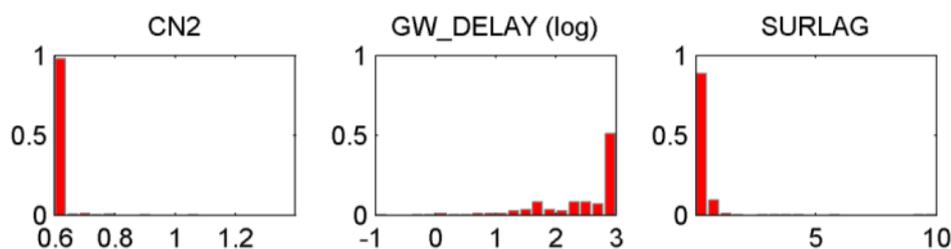

Figure 9.44. Parameter distribution among the ensemble for the last iteration of the cascading-regularized calibration of the parsimonious model. Histogram normalized according to ensemble size (160).
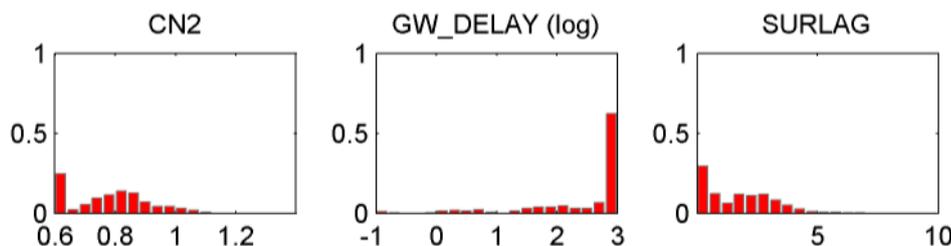
Figure 9.45. Parameter distribution among the ensemble for the last iteration of the cascading-regularized calibration of the detailed model. Histogram normalized according to ensemble size (160).

Detailed results for the cascading-regularized calibrations of both models are presented in Appendix I.P and Appendix I.L. Again, detailed numerical calibration and validation results can be found in the appendices along with synthesis series for an extended set of calibration points.

9.7   Comparing calibration strategies

Drawing near to the end of the chapter, an effort to choose the "best" model – and calibration procedure – is worth making. As can, however, be hinted from the previous discussions, this is no easy task.

Due to several factors, from chance to modeling decisions, the detailed model did not clearly outperform the parsimonious one. In fact, on several occasions it seemed to perform worse. Adding

this to the fact that the price of higher detail is a four-fold increase in overall computation time[10], there are no strong arguments in favor of using the detailed model.

This does not mean that the detailed model does not hold the potential to surpass the detailed one but, in the present work, this potential was not harnessed.

Among calibration approaches, the cascading calibration displayed several weaknesses. For example, it is hard to pursue it down to the ZRB's delta and obtaining good results at the reservoirs is very difficult due to the one-sense flow of information. Furthermore, because the cascading calibration roughly constitutes a particular case of the cascading-regularized one, when heterogeneities are potentially very marked, it is believed that the original cascading calibration can only be improved by the addition of the regularization scheme.

From these considerations, the best solution was sought among the results of three candidate calibration approaches (global, regional-regularized, and cascading-regularized) applied at the parsimonious model. Among these, there is no clear "winner".

A good understanding of the different parameter's performances according to each of the calibration schemes can be achieved by focusing separately on locations along the main reach of the Zambezi, along the tributaries, and for the monitored reservoirs.

Detailed evaluations were made using the AIRAD, IRRMSE and NE error measures as these are unitless and, consequently, allow for an easier comparison among calibration points. The comparison of AIRAD results for stations on the main reach of the Zambezi River is portrayed in Figure 9.46 (calibration period) and Figure 9.47 (validation period).

Being the desirable AIRAD equal to zero at every calibration point, the calibration approaches can be easily ordered by their performance. The cascading-regularized calibration comes first, followed by the global one.
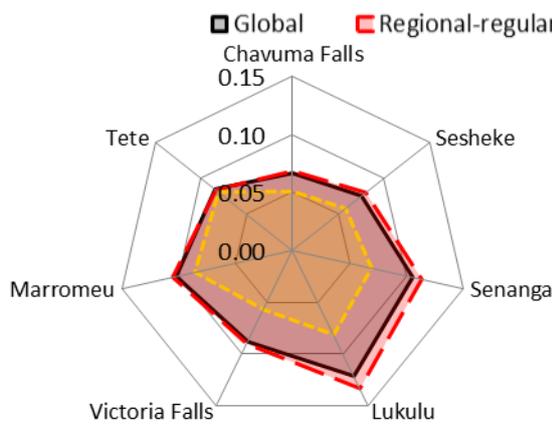

Figure 9.46. Best calibration AIRAD of the parsimonious global, regional-regularized, and cascading-regularized models at points along the main reach of the Zambezi River.
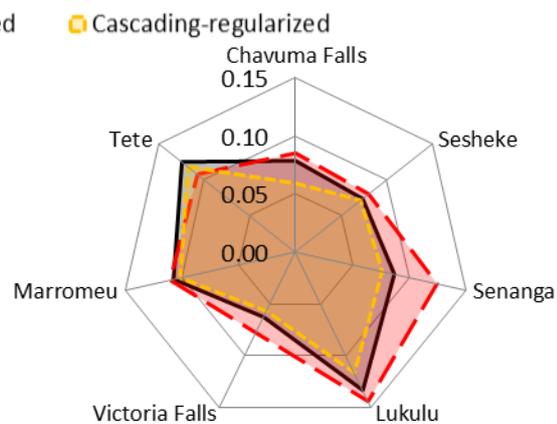
Figure 9.47. Best validation AIRAD of the parsimonious global, regional-regularized, and cascading-regularized models at points along the main reach of the Zambezi River.

Looking at tributaries, whose comparison is depicted in Figure 9.48 and in Figure 9.49, the situation is inversed. Now, the most advantageous alternative seems to be the regional-regularized calibration, followed by the global one.

---

[10] This includes result processing, which takes roughly the same amount of time independently from model complexity. In practice, the detailed model runs, therefore, more than four times slower than the parsimonious one. Additionally, due to very large number of files being handled during the simulations, only computers with several independent hard drives can cope with the parallel processing of the detailed model; it is not the case for the parsimonious one.
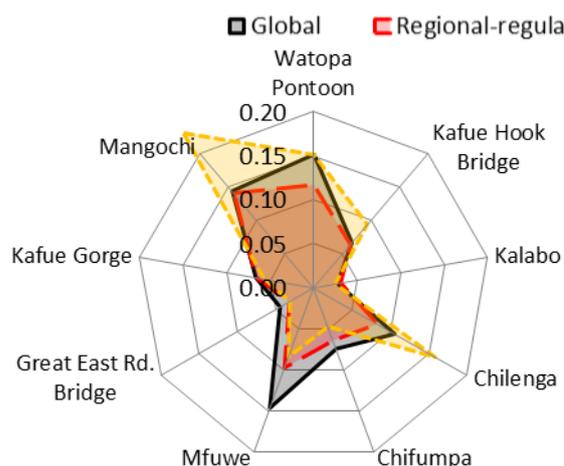
Figure 9.48. Best calibration AIRAD of the parsimonious global, regional-regularized, and cascading-regularized models at points along the main tributaries.
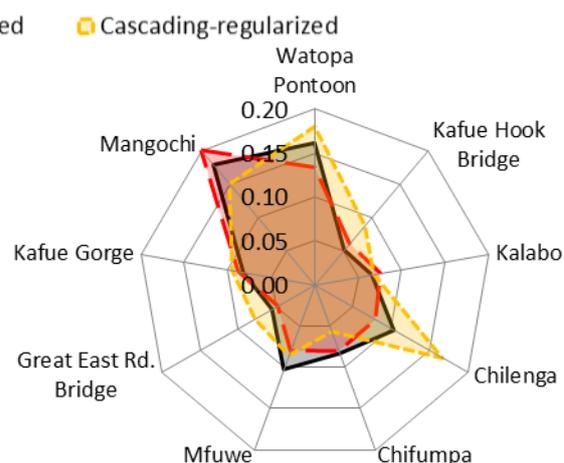

Figure 9.49. Best validation AIRAD of the parsimonious global, regional-regularized, and cascading-regularized models at points along the main tributaries.

Finally, Figure 9.50 and Figure 9.51 contain results associated with the monitored reservoirs. Through their inspection, it can be seen that the regional-regularized calibration appears to have performed best, again followed by the global approach.
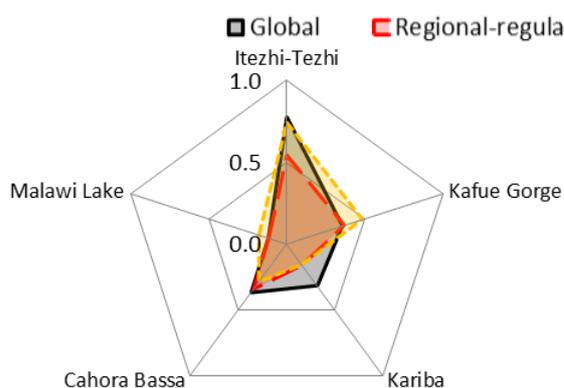
Figure 9.50. Best calibration AIRAD of the parsimonious global, regional-regularized, and cascading-regularized models at the monitored reservoirs.
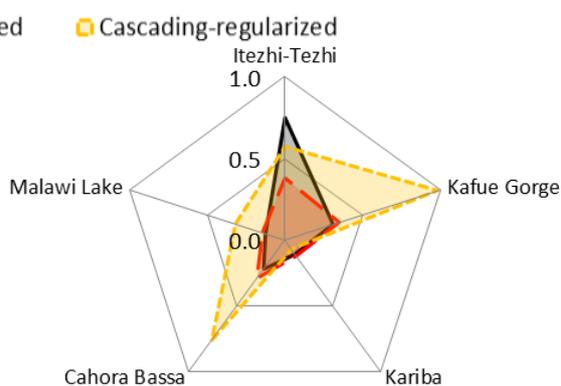
Figure 9.51. Best validation AIRAD of the parsimonious global, regional-regularized, and cascading-regularized models at the monitored reservoirs.

A similar interpretation arises from the analysis of the IRRMSE and NE. Equivalent plots for these measures can be found in Appendix I.Q.

Based on these findings, the regional-regularized calibration of the parsimonious model is proposed for general aims as it produced coherent simulations (as discussed above) and obtained the best results for tributaries and reservoirs at the expense of only a little degradation of main reach measures.

In a study addressing the main reach of the Zambezi River – reservoir volumes being neglected – the cascading-regularized parsimonious model would have to be recommended.

Finally, the globally calibrated parsimonious model constitutes a sound all-round model which could be used if tributaries are closely inspected.

9.8   Conclusions

The questions raised at the onset of this chapter were many and cover a wide range of topics related to the calibration of hydrological models. In order to address them, millions of SWAT model runs were performed during the course of two years in three high-end desktop computers.

Model reliability was a central concern of the research. In particular, the issue of how to address the concomitant heterogeneity of the ZRB and scarcity of information has underpinned the whole model development and calibration phases.

While scarcity of information was partly tackled by the work described in Chapters 5 and 8, with the development of the Pattern-Oriented Memory (POM) interpolation technique and the consequent possibility of extending calibration times significantly, heterogeneity still had to be addressed during model calibration.

Along with findings related to the FAST, the calibration results clearly demonstrated the importance of considering a heterogeneous set of parameters. In practice this gave strength to the case for a MO calibration in detriment of a SO one. Also, it prompted the development of ways to "control" reliability in models of appreciable complexity. To this end, a regional-regularization scheme was developed. Additionally, because there are considerable regions of the ZRB which remain greatly ungauged or have only recently become monitored, a way of transposing parameters from neighboring areas to them is needed. If working properly, a regional-regularization calibration can be used to that end.

Albeit the proposed regional-regularization was successful in calibrating heterogeneous models while keeping validation errors in check, the resulting models did not meet their full potential. Also, the heterogeneity minimization objective was not observed to be directly associated with validation performance. While future trials with larger ensemble sizes and longer calibration runs are recommended in order to fully access the regional-regularized calibration methodology, the "easier" problem of calibrating a heterogeneous model without regularization should also be looked upon.

While most published researches on the subject emphasize the merits of certain optimization algorithms, little attention is granted to their implementation in real, complex cases (where computation time matters) and to the fact that poor parameter spaces can exist where even the best performing optimization models face difficulties. In the case of the ZRB, the uncontrolled reservoir outflow parameters constituted bottlenecks in the optimization procedure and severely hindered the performance of the optimization models. It is recommended that parameters which can completely and abruptly alter the model's response are, if possible, not included in the general optimization or at least handled with particular caution.

The comparison of the parsimonious and the detailed models tended towards the former due to failed calibrations and the greater computational demands of the latter; arguments only lightly contradicted by some isolated performance gains. In the future, an increase in ensemble size and iterations, along with the removal of the heterogeneity minimization objective, might contribute to decisive argument for one of them.

Global, regional-regularized and cascading-regularized calibration approaches were evaluated. While the global calibration provided solid results, it did not account for the basin's heterogeneity, which led to poorer performances at the tributaries. Because the cascading-regularized calibration suffered some degradation from modeling decisions and less accomplished local optimizations, the regional-regularized approach is recommended as the calibration strategy for the SWAT ZRB models. In spite of not leading to the best results on every account – most notably performance along the main reach of the Zambezi River – it achieved the best calibrations of tributary subbasins and reservoirs with little concessions on other accounts. Additionally, it was shown to produce more coherent and balanced simulations. Finally, the calibration results did not reflect the approach's potential, which translates into the possibility of future improvements.

The produced models are considered fit for supporting water resources assessments and planning. If coupled with the novel and very promising methodology described in Section 6.6 or a more complex BI framework, uncertainty could be associated with the model's outputs, further increasing its potential.

In order to extend the calibrated model's capabilities to include forecast, coupling then with a filtering scheme (e.g. an Unscented Kalman filter) would be interesting, as new observations could be continuously used in order to improve simulations.

The best performing models have been made available online, either to be downloaded or to be interactively used. To facilitate this, a new platform, developed within the present research, will be employed. The description of the platform – the ADAPT online database – is the object of Appendix II.A.

Finally, there are prospects to contribute to the European Space Agency's (ESA) TIGER initiative[11] with the calibrated SWAT models of the ZRB.

---

[11] http://www.tiger.esa.int/.

10 Main conclusions and perspectives on future work

10.1 Overview

The ultimate aim of the research was to prepare a hydraulic-hydrological model of the Zambezi River Basin (ZRB) at a daily time scale. It should have potential to be applied in parallel and subsequent researches being developed under the African Dams Project (ADAPT) and directly by stakeholders. Owing to the dire needs of Southern African populations and to a true scope for improvement regarding water resources management practices and the capacitation of professionals and researchers within the sector, every effort was made to direct the research herein reported into producing tangible and applicable results.

A wide range of topics was studied in the course of the work; from the careful assessment of modeling data, passing by the development of a novel rainfall interpolation technique, diverting into the field of machine learning models applied to discharge forecasts, and culminating in the thorough study of calibration strategies for semi-distributed models of the ZRB. Striving to avoid the all too often alluring trap of conducting research for the sake of developing something original, the guideline directing the work has been to identify the main challenges to be addressed prior to meeting the established goals and devise ways of overcoming them. Being naturally impossible to tackle all challenges on the horizon, subjective choices had to be made in order to select those which were both useful to address and realistically within reach.

Resorting to remote sensing data is, in the case of the ZRB, a real necessity. The investment in the careful evaluation of its sources arose from the realization that substantial modeling errors could stem from the use of inadequate products.

In a scenario of scarce ground data, the belief that gains were to be made from the use of extended rainfall series (something later confirmed, as described in Chapter 8), for which no suitable models had been found in literature, prompted the experimentation with machine learning algorithms in order to produce an entirely new rainfall interpolation method. This held great risks but was ultimately beneficial.

The study of machine learning models for discharge forecast was encouraged by the realization that, without a proper data assimilation technique, models such as the Soil and Water Assessment Tool (SWAT) are unfit for such a task. More so, even coupled with adequate assimilation techniques, traditional hydrological models' forecasts depend on estimates of future forcings and entail a burdensome development phase. The proposed machine learning models, having the virtues of being flexible and extremely easy to apply, attained very promising results.

Great emphasis was placed on the improvement and calibration of SWAT. Knowing that adequate calibrations of complex models usually demand an excessively large amount of runs, an effort was made to shorten the time a SWAT simulation of the ZRB took to complete. Without the 20-fold improvements that were attained, the analyses of the proposed calibration strategies would have certainly not been possible. Along with the efficiency concerns, the belief in the major role of heterogeneity within the basin led to the revision of the SWAT source code, which potentiated the identification and, at times, correction of relevant issues. Among these were the inaccurate calculation of reservoir surfaces, the independence between land and reservoir areas, and the overestimation of canopy rainfall interception.

10.2 Introduction of the Pattern-Oriented Memory interpolation technique

The novel Pattern-Oriented Memory (POM) rainfall interpolation technique is based on non-linear regression machine learning models and makes use of satellite rainfall estimates in order to interpolate historical ground records. It was shown to be superior to popular techniques such as Kriging in scenarios where data is spatially scarce (the case of the ZRB). Not only does this methodology have the potential to perform rainfall interpolations with fewer errors than other methods, it also displays the capability to improve as more and better satellite data becomes available. Furthermore, it has been shown to be particularly fit to extend recent satellite rainfall estimates due to the fact that it effectively transposes their main features to the historical series, greatly reducing the usually constraining "change of support" problems. Resorting to POM, daily hydrologic simulations of the ZRB, previously limited to a period starting in 1998, can now be carried out from as far back as 1979.

10.3 Machine learning models applied to streamflow forecast

The research related to the use of machine learning models for streamflow forecast was conducted in four fronts: the comparison of alternative models (e.g. Autoregressive Moving-Average (ARMA), Artificial Neural Networks (ANN) and Support-Vector Regression (SVR)); the enhancement of rainfall aggregation techniques; the study of limitations inherent to SVR forecasting models; and the development of a non-parametric empirical uncertainty post-processor.

Going beyond the development of deterministic forecasting models with promising accuracies, even for long lead times of up to 60 days at Victoria Falls, the conducted research motivated a reevaluation of previous findings reported in literature by showing that SVR models are particularly hazardous when used for discharge forecasting purposes; a conclusion based on their underlying theoretical principles and easily observable in practice.

10.4 A novel non-parametric post-processor for empirical uncertainty

In what is regarded as being the most promising outcome of the thesis, a novel non-parametric empirical uncertainty post-processor was developed. By carrying out the simultaneous training of a large-sized ensemble of ANNs under a particular set of conflicting objectives, the proposed methodology is able to effectively generate probabilistically correct uncertainties of detrended series given a representative set of training patterns. It is an (informal) technique that, unlike Bayesian methods (formal), does not require the definition of likelihood functions nor an external "conceptual" model of the phenomenon being modeled. Being extremely versatile and straightforward to set up, it can be easily adapted to incorporate new information and, depending on the problem at hand, can run orders of magnitude faster than Bayesian methods.

Having been successfully used in order to forecast discharges at Victoria Falls and Kafue Hook Bridge, preliminary results place its performance on par with that of state-of-the-art methodologies employed to characterize discharge forecast uncertainty. It is believed that, due to its non-

parametric nature, it does not share the limitations brought about by the need to specify approximate theoretical likelihood distributions of formal uncertainty estimators.

The potential range of applicability of the methodology goes well beyond discharge forecasting and even hydrology. Being a non-parametric empirical uncertainty post-processor, it can theoretically be used to characterize uncertainty of estimates made in fields as diverse as economy, chemistry or physics.

10.5   Revising the Soil and Water Assessment Tool's adequacy to the Zambezi River basin.

The analysis of the SWAT model in its application to the ZRB has evidenced inadequacies in the code. Not putting into question the overall value of the model – which is great – the issues that have been identified affect results, particularly when relatively large water bodies are present; in that case, validation should be addressed with care. If SWAT's calibration is regarded as a general optimization problem and proper validation is carried out (as was the case here), its structural shortcomings, while potentially contributing to a degradation of the model's potential performance, should not be constraining. Conversely, the alternative approaches of setting the model's main parameters according to theoretical considerations, field measurements or attending to findings reported in literature are regarded as ill-advised.

The inadequacies signaled in SWAT code were found in the midst of an effort to understand the model and explain observed results. Being the checking of SWAT's code a task well outside the scope of the present research, the question of whether more inadequacies would be found as a result of a thorough evaluation procedure cannot but to be asked. Probably the answer would be yes.

10.6   Proposal of new model calibration procedures

A large effort to propose adequate calibration procedures for the ZRB's hydrologic models was made. These models (SWAT-based), while not competing with machine learning alternatives for the purpose of streamflow forecast, present enormous potential given their ability to accurately simulate, not only water quantity related processes, but water quality as well, being almost ideally suited to inform medium to long-term water resources management decisions.

In total, four different approaches, encompassing Single- (SO) and Multi-Objective (MO) optimizations, were evaluated over two SWAT models. Resorting to millions of simulations, the full calibration of daily hydrological models covering the whole basin, from the Upper Zambezi to a few kilometers upstream from the Delta (Marromeu) was accomplished – it is believed – for the first time. To that end, effective methods for an effective combination of discharge, water level and stored volume information acquired at multiple locations had to be considered.

Heterogeneity was shown to play a noticeable role in the basin's hydrology and it is recommended that the calibration of future models allows for the definition of regional parameters. The tested calibration approaches accounted for the ZRB's heterogeneity to different degrees: from a global model which does not allow parameter heterogeneity to a cascading-regularized model which considers varying sets of parameters at each region. Best results were obtained using a regional-regularized calibration approach due to its capacity of approximating contributions, not only of subbasins along the main reach of the Zambezi, but also along its tributaries.

From the two tested models, the simpler one (labeled "parsimonious") was found to be advantageous to a more complex ("detailed") alternative. In fact, having four times less computational demands and producing similar or better results than its detailed counterpart, the use of the parsimonious model is encouraged. Although such findings are intrinsically related to SWAT, it is believed that in light of the complexity-data-performance-reliability trade-off, the scarce amount of information available for the calibration of daily rainfall-runoff models over the ZRB favors the use of simpler models.

10.7 Sharing results and future work

Along with original data and works from fellow ADAPT researchers, the most important outcomes of the research will be conveyed to stakeholders through the ongoing ADAPT online database project. In the future, it is expected that the database becomes useful for anyone interested in ZRB's water resources and that it is given continuity by local partners.

In the wake of the proposed methodologies and presented findings there are several topics worthwhile pursuing in the future. One is undoubtedly the undertaking of an extended regional-regularized calibration of the parsimonious SWAT model of the ZRB using a larger ensemble and taking advantage of more iterations. Also of great use would be to couple the calibrated model with a data assimilation procedure – such as an unscented Kalman filter – so that it could provide a sound alternative for river discharge forecast.

Another promising avenue of research is related to the POM interpolation method. In particular, it would be interesting to apply it to regions with a high variability of the rainfall field and for shorter time steps – possibly replacing satellite by radar data as the "source" of rainfall patterns. Additionally, profiting from the advantages of POM regarding hydrological modeling, further testing on to what extent POM reduces the "change of support" problem would be interesting.

Finally, it has certainly not slipped the reader's attention that the developed non-parametric empirical uncertainty post-processor (yet to be found a proper designation) is regarded with the utmost interest. Notwithstanding its relatively small share of the document – consequence of a late development – it is believed that such interest is justified due to the promising results and to a very wide range of potential applications. Future work will certainly pass by the application of the method to streamflow forecast within the Zambezi and to the characterization of the uncertainty of calibrated SWAT ZRB models. Also, it is believed that benchmarking the methodology in well-known areas, such as the Rhone River basin, and comparing it with formal uncertainty estimation methods will yield motivating results.

Overall, it is expected that the outcomes of this thesis can be profitably used in practice, namely in tasks such as dam operation optimization and synchronization, environmental impact assessments, evaluation of future scenarios (predicting responses to climate change and increased demands) and a broad range of other studies related to themes such as wildlife, water chemistry, sediment transport, and integrated water management focusing on the ZRB. More importantly, it is hoped that the reported findings can be built upon by others in order to potentiate new, promising, and unforeseen methodologies.

REFERENCES

Abbaspour, K. C. 2012. SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs - a user manual. .

Abbaspour, K. C., J. Yang, I. Maximov, R. Siber, K. Bogner, J. Mieleitner, J. Zobrist, and R. Srinivasan. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. Journal of Hydrology **333**(2):413-430.

Abrahart, R. J. and L. See. 2000. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrological Processes **14**(11-12):2157-2172.

Aduah, M. S. 2007. Multitemporal remote sensing for mapping and monitoring floods, an approach towards validation of the KAFRIBA model, Kafue Flats, Zambia. Master. International Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands.

Aerts, J., M. Kriek, and M. Schepel. 1999. STREAM (Spatial tools for river basins and environment and analysis of management options): 'set up and requirements'. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere **24**(6):591-595.

Akhtar, M., G. Corzo, S. v. Andel, and A. Jonoski. 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. Hydrology and Earth System Sciences **13**(9):1607-1618.

Anderson, P. L., M. M. Meerschaert, and K. Zhang. 2013. Forecasting with prediction intervals for periodic autoregressive moving average models. Journal of Time Series Analysis **34**(2):187-193.

Andréassian, V., C. Perrin, C. Michel, I. Usart-Sanchez, and J. Lavabre. 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. Journal of Hydrology **250**(1-4):206-223.

Arnold, J., D. Moriasi, P. Gassman, K. Abbaspour, M. White, R. Srinivasan, C. Santhi, R. Harmel, A. Van Griensven, and M. Van Liew. 2012. SWAT: Model use, calibration, and validation. Transactions of the ASABE **55**(4):1491-1508.

Artan, G., H. Gadain, J. Smith, K. Asante, C. Bandaragoda, and J. Verdin. 2007. Adequacy of satellite derived rainfall data for stream flow modeling. Natural Hazards **43**(2):167-185.

Asefa, T., M. Kemblowski, M. McKee, and A. Khalil. 2006. Multi-time scale stream flow predictions: The support vector machines approach. Journal of Hydrology **318**(1):7-16.

Auger, A. and N. Hansen. 2005. A restart CMA evolution strategy with increasing population size. Pages 1769-1776 *in* IEEE Congress on Evolutionary Computation. IEEE, Edinburgh, Scotland.

Bacchi, B. and N. T. Kottegoda. 1995. Identification and calibration of spatial correlation patterns of rainfall. Journal of Hydrology **165**(1-4):311-348.

Bannari, A., D. Morin, F. Bonn, and A. R. Huete. 1995. A review of vegetation indices. Remote Sensing Reviews **13**(1-2):95-120.

Baratti, R., B. Cannas, A. Fanni, M. Pintus, G. M. Sechi, and N. Toreno. 2003. River flow forecast for reservoir management through neural networks. Neurocomputing **55**(3):421-437.

Bartsch, A., W. Wagner, K. Scipal, C. Pathe, D. Sabel, and P. Wolski. 2009. Global monitoring of wetlands – the value of ENVISAT ASAR Global mode. Journal of environmental management **90**(7):2226-2233.

Beck, L. 2010. Transboundary water allocation in the Zambezi River Basin. Ph.D. Swiss Federal Institute of Technology Zurich (ETHZ), Zurich.

Beilfuss, R. 2001. Prescribed flooding and restoration potential in the Zambezi Delta, Mozambique. Working paper #3. International Crane Foundation, USA.

Beilfuss, R. and D. Dos Santos. 2001. Patterns of hydrological change in the Zambezi Delta, Mozambique. Working Paper #2. International Crane Foundation, USA.

Bekele, E. G. and J. W. Nicklow. 2007. Multi-objective automatic calibration of SWAT using NSGA-II. Journal of Hydrology **341**(3–4):165-176.

Berry, P., R. Smith, J. Freeman, and J. Benveniste. 2008. Towards a new global digital elevation model. Pages 431-435 Observing our changing Earth. Springer Berlin Heidelberg.

Berry, P. A. M., J. D. Garlick, J. A. Freeman, and E. L. Mathers. 2005. Global inland water monitoring from multi-mission altimetry. Geophysical Research Letters **32**(16):L16401.

Berry, P. A. M., R. G. Smith, and J. Benveniste. 2010. ACE2: the new global digital elevation model. Pages 231-237 *in* S. P. Mertikas, editor. Gravity, Geoid and Earth Observation. Springer Berlin Heidelberg.

Beven, K. 1989. Changing ideas in hydrology — The case of physically-based models. Journal of Hydrology **105**(1–2):157-172.

Beven, K. 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. Advances in Water Resources **16**(1):41-51.

Beven, K. 2009. Comment on "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" by Jasper A. Vrugt, Cajo J. F. ter Braak, Hoshin V. Gupta and Bruce A. Robinson. Stochastic environmental research and risk assessment **23**(7):1059-1060.

Beven, K. and A. Binley. 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrological Processes **6**(3):279-298.

Beven, K., P. Smith, and J. Freer. 2007. Comment on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology" by Pietro Mantovan and Ezio Todini. Journal of Hydrology **338**(3):315-318.

Birkett, C., C. Reynolds, B. Beckley, and B. Doorn. 2011. From research to operations: The USDA Global Reservoir and Lake Monitor. Pages 19-50 *in* S. Vignudelli, A. G. Kostianoy, P. Cipollini, and J. Benveniste, editors. Coastal Altimetry. Springer Berlin Heidelberg.

Birkett, C. M. 1994. Radar altimetry: A new concept in monitoring lake level changes. EOS, Transactions American Geophysical Union **75**(24):273-275.

Bogner, K. and F. Pappenberger. 2011. Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system. Water Resources Research **47**(7):W07524.

Brakenridge, G. R., S. V. Nghiem, E. Anderson, and S. Chien. 2005. Space-based measurement of river runoff. EOS, Transactions American Geophysical Union **86**(19):185-188.

Brakenridge, R. and E. Anderson. 2006. MODIS-based flood detection, mapping and measurement: the potential for operational hydrological applications. Pages 1-12 *in* J. Marsalek, G. Stancalie, and G. Balint, editors. Transboundary Floods: Reducing Risks Through Flood Management. Springer Netherlands.

Breiman, L. 2001. Random Forests. Machine learning **45**(1):5-32.

Brockwell, P. J. and R. A. Davis. 2002. Introduction to time series and forecasting. 2nd edition. Springer-Verlag, New York.

Broomhead, D. and D. Lowe. 1988. Radial basis functions, multi-variable functional interpolation and adaptive networks. Royal signals and radar establishment Malvern, England.

Brown, M. E. and V. M. Escobar. 2014. Assessment of Soil Moisture Data Requirements by the Potential SMAP Data User Community: Review of SMAP Mission User Community. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of **7**(1):277-283.

Broyden, C. G. 1970. The convergence of a class of double-rank minimization algorithms 1. general considerations. IMA Journal of Applied Mathematics **6**(1):76-90.

Bruen, M. and J. Yang. 2005. Functional networks in real-time flood forecasting - a novel application. Advances in Water Resources **28**(9):899-909.

Brunsdon, C., J. McClatchey, and D. Unwin. 2001. Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression. International journal of climatology **21**(4):455-466.

Brutsaert, W. 2005. Hydrology: an introduction. Cambridge University Press, Cambridge, U.K.

Calmant, S., F. Seyler, and J. Cretaux. 2008. Monitoring Continental Surface Waters by Satellite Altimetry. Surveys in Geophysics **29**(4-5):247-269.

Cannavó, F. 2012. Sensitivity analysis for volcanic source modeling quality assessment and model selection. Computers & Geosciences **44**(0):52-59.

Chahine, M. T. 1992. The hydrological cycle and its influence on climate. Nature **359**:373-380.

Chang, C. C. and C. J. Lin. 2011. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) **2**(3).

Chang, H., X. Li, and L. Ge. 2010. Assessment of SRTM, ACE2 and ASTER-GDEM using RTK-GPS. 15th Australasian Remote Sensing & Photogrammetry Conference, Australia.

Chen, J. 2010. Flooded with water deficits : optimal timing of an environmental flow release, using flow forecasts in the operational management of a reservoir. Master. University of Twente, Twente.

Chen, J. and B. Adams. 2006. Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. Journal of Hydrology **318**(1-4):232-249.

Chu, W., T. Yang, and X. Gao. 2014. Comment on "High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing" by Eric Laloy and Jasper A. Vrugt. Water Resources Research.

Cloke, H. L. and F. Pappenberger. 2009. Ensemble flood forecasting: A review. Journal of Hydrology **375**(3–4):613-626.

Cohen Liechti, T. 2013. Influence of dam operation on water resources management under different scenarios in the Zambezi River Basin considering environmental objectives and hydropower. Ph.D. thesis 6012 (2013). École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Cohen Liechti, T., J. P. Matos, J.-L. Boillat, and A. J. Schleiss. 2011. Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi River Basin. Hydrology and Earth System Sciences **8**:8173-8201.

Cohen Liechti, T., J. P. Matos, D. Ferràs Segura, J.-L. Boillat, and A. J. Schleiss. 2014. Hydrological modelling of the Zambezi River basin taking into account floodplain behaviour by a modified reservoir approach. International Journal of River Basin Management:1-13.

Cortes, C. and V. Vapnik. 1995. Support-vector networks. Machine learning **20**(3):273-297.

CRC Catchment Hydrology. 2005. Series on Model Choice: 1. General approaches to modelling and practical issues of model choice. Cooperative Research Centre for Catchment Hydrology.

Cressie, N. 1990. The origins of kriging. Mathematical Geology **22**(3):239-252.

Cukier, R. I., C. M. Fortuin, K. E. Shuler, A. G. Petschek, and J. H. Schaibly. 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. The Journal of Chemical Physics **59**(8):3873-3878.

Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems **2**(4):303-314.

Danielson, J. J. and D. B. Gesch. 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010). U. S. Geological Survey, Reston, Virginia, USA.

Dawdy, D. R. and T. O'Donnell. 1965. Mathematical models of catchment behavior. Journal of the Hydraulics Division, ASCE **91**(HY4):113-137.

De Brabanter, K., P. Karsmakers, F. Ojeda, C. Alzate, J. De Brabanter, K. Pelckmans, B. De Moor, J. Vandewalle, and J. Suykens. 2010. LS-SVMlab Toolbox User's Guide version 1.8. ESAT-SISTA, K.U.Leuven, Leuven.

De Vos, N. and T. Rientjes. 2005. Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation. Hydrology and Earth System Sciences **9**(1-2):111-126.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002a. A fast and elitist multiobjective genetic algorithm: NSGA-II. Evolutionary Computation, IEEE Transactions on **6**(2):182-197.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002b. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation **6**(2):182-197.

Defries, R. S. and J. R. G. Townshend. 1994. NDVI-derived land cover classifications at a global scale. International Journal of Remote Sensing **15**(17):3567-3586.

Delvaux, D. 1995. Age of Lake Malawi (Nyasa) and water level fluctuations. Annual report 1993 & 1994. Royal Museum of Central Africa, Department of Geology and Mineralogy, Tervuren, Belgium.

Denconsult. 1998. Sector studies under ZACPLAN.

DHV. 1980. Kafue Flats hydrological studies.

DHV. 2004. Final report, Integrated water resources management project for the Kafue Flats Zambia, Components 1, 2, 3.

Dinku, T., P. Ceccato, E. Grover-Kopec, M. Lemma, S. J. Connor, and C. F. Ropelewski. 2007. Validation of satellite rainfall products over East Africa's complex topography. International Journal of Remote Sensing **28**(7):1503-1526.

Diskin, M. H. and E. Simon. 1977. A procedure for the selection of objective functions for hydrologic simulation models. Journal of Hydrology **34**(1–2):129-149.

Doherty, J. 2005. User's manual for PEST, version 10. Watermark Numerical Computing, Brisbane, Australia.

Drucker, H., C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik. 1996. Support vector regression machines. Pages 155-161 *in* Advances in neural information processing systems. MIT Press, Denver, Colorado, USA.

Duan, Q., S. Sorooshian, and V. Gupta. 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research **28**(4):1015-1031.

Eberhart, R. and J. Kennedy. 1995. A new optimizer using particle swarm theory. Pages 39-43 *in* Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on. IEEE.

Ebert, E. E., J. E. Janowiak, and C. Kidd. 2007. Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bulletin of the American Meteorological Society **88**(1):47-64.

Efstratiadis, A. and D. Koutsoyiannis. 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. Hydrological Sciences Journal–Journal des Sciences Hydrologiques **55**(1):58-78.

Euroconsult Mott MacDonald. 2007. Integrated water resources management stategy for the Zambezi River Basin - Rapid assessment report.

Evsukoff, A. G., B. S. L. P. de Lima, and N. F. F. Ebecken. 2011. Long-term runoff modeling using rainfall forecasts with application to the Iguaçu river basin. Water resources management **25**(3):963-985.

Farr, T. G., P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf. 2007. The Shuttle Radar Topography Mission. Reviews of Geophysics **45**(2):RG2004.

Fekete, B., C. Vörösmarty, J. Roads, and C. Willmott. 2004. Uncertainties in precipitation and their impacts on runoff estimates. Journal of Climate **17**:294-304.

Fletcher, R. 1970. A new approach to variable metric algorithms. The computer journal **13**(3):317-322.

Food and Agriculture Organization of the United Nations (FAO). 1995. Digital soil map of the World and derived soil properties (version 3.5) [CD-ROM]. FAO.

Forster, M. R. 2005. Notice: No Free Lunches for Anyone, Bayesians Included. Department of Philosophy, University of Wisconsin–Madison Madison, USA.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons Inc.

Fujisada, H., G. B. Bailey, G. G. Kelly, S. Hara, and M. J. Abrams. 2005. ASTER DEM performance. Geoscience and Remote Sensing, IEEE Transactions on **43**(12):2707-2714.

Gao, B.-C. 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment **58**(3):257-266.

Garcia Hernandez, J., F. Jordan, J. Dubois, J.-L. Boillat, and A. J. Schleiss. 2007. Routing System II - Modélisation d'écoulements dans des systèmes hydrauliques. 1661-1179, Laboratoire de Constructions Hydrauliques, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

García, S., D. Molina, M. Lozano, and F. Herrera. 2009. A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the CEC'2005 special session on real parameter optimization. Journal of Heuristics **15**(6):617-644.

Gelfand, A. E., L. Zhu, and B. P. Carlin. 2001. On the change of support problem for spatio-temporal data. Biostatistics **2**(1):31-45.

Gerrits, A. M. J. 2005. Hydrological modelling of the Zambezi catchment for gravity measurements. Master. Techinical University Delft, Delft, The Netherlands.

Gesch, D. B., K. L. Verdin, and S. K. Greenlee. 1999. New land surface digital elevation model covers the Earth. EOS, Transactions American Geophysical Union **80**(6):69-70.

Goldberg, D. and J. Holland. 1988. Genetic Algorithms and Machine Learning. Machine learning **3**(2-3):95-99.

Goldfarb, D. 1970. A family of variable-metric methods derived by variational means. Mathematics of computation **24**(109):23-26.

Govender, M. and C. Everson. 2005. Modelling streamflow from two small South African experimental catchments using the SWAT model. Hydrological Processes **19**(3):683-692.

Govindaraju, R. S. 2000a. Artificial neural networks in hydrology. I: Preliminary concepts. Journal of Hydrologic Engineering **5**(2):115-123.

Govindaraju, R. S. 2000b. Artificial neural networks in hydrology. II: hydrologic applications. Journal of Hydrologic Engineering **5**(2):124-137.

Griensven, A. v. and T. Meixner. 2006. Methods to quantify and identify the sources of uncertainty for river basin water quality models. Water Science & Technology **53**(1):51-59.

Grimes, D., E. Pardo-Iguzquiza, and R. Bonifacio. 1999. Optimal areal rainfall estimation using raingauges and satellite data. Journal of Hydrology **222**(1-4):93-108.

Gupta, H. V., S. Sorooshian, and P. O. Yapo. 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resources Research **34**(4):751-763.

Gupta, H. V., S. Sorooshian, and P. O. Yapo. 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. Journal of Hydrologic Engineering **4**(2):135-143.

Haario, H., E. Saksman, and J. Tamminen. 2001. An adaptive Metropolis algorithm. Bernoulli **7**(2):223-242.

Haberlandt, U. 2007. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. Journal of Hydrology **332**(1-2):144-157.

Hagan, M. T. and M. B. Menhaj. 1994. Training feedforward networks with the Marquardt algorithm. IEEE transactions on neural networks **5**(6):989-993.

Hancock, P. and M. Hutchinson. 2006. Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines. Environmental Modelling & Software **21**(12):1684-1694.

Hansen, N. 2006. The CMA evolution strategy: a comparing review. Pages 75-102 *in* J. Lozano, P. Larrañaga, I. Inza, and E. Bengoetxea, editors. Towards a new evolutionary computation. Springer Berlin Heidelberg.

Hansen, N. 2010. The CMA evolution strategy: A tutorial.

Hargreaves, G. H. and Z. A. Samani. 1985. Reference crop evapotranspiration from ambient air temperature. American Society of Agricultural Engineers.

Haykin, S. 1994. Neural networks: a comprehensive foundation. Prentice Hall, Upper Saddle River, NJ, USA.

Hirt, C., M. S. Filmer, and W. E. Featherstone. 2010. Comparison and validation of the recent freely available ASTER-GDEM ver1, SRTM ver4.1 and GEODATA DEM-9S ver3 digital elevation models over Australia. Australian Journal of Earth Sciences **57**(3):337-347.

Hong, Y., D. Gochis, J.-t. Cheng, K.-l. Hsu, and S. Sorooshian. 2007. Evaluation of PERSIANN-CCS Rainfall Measurement Using the NAME Event Rain Gauge Network. Journal of Hydrometeorology **8**(3):469-482.

Hsu, K.-l., X. Gao, S. Sorooshian, and H. V. Gupta. 1997. Precipitation estimation from remotely sensed Information using artificial neural networks. Journal of Applied Meteorology **36**(9):1176-1190.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker. 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology **8**(1):38-55.

Hughes, D. A. 2006. Comparison of satellite rainfall data with observations from gauging station networks. Journal of Hydrology **327**(3–4):399-410.

Hyndman, R. J., G. Athanasopoulos, S. Razbash, D. Schmidt, Z. Zhou, Y. Khan, and C. Bergmeir. 2013. forecast: forecasting functions for time series and linear models. R package version 4.8. http://CRAN.R-project.org/package=forecast.

Ibbit, R. P. 1970. Systematic parameter fitting for conceptual models of catchment hydrology. Ph.D. thesis. Imperial College London, University of London, London, UK.

Jarvis, A., H. Reuter, A. Nelson, and E. Guevara. 2008. Hole-filled SRTM for the globe Version 4. CGIAR-CSI.

Jobard, I., F. Chopin, J. C. Berges, and R. Roca. 2011. An intercomparison of 10-day satellite precipitation products during West African monsoon. International Journal of Remote Sensing **32**(9):2353-2376.

Jordan, F. and A. J. Schleiss. 2007. Modèle de prévision et de gestion des crues - optimisation des opérations des aménagements hydroélectriques à accumulation pour la réduction des débits de crue. 1661-1179, Laboratoire de Constructions Hydrauliques, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie. 2004. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. Journal of Hydrometeorology **5**(3):487-503.

Juízo, D., J. P. Matos, L. Katiyo, S. Grass, F. Nyatsanza, and H. Makurira. 2012. Recommendations on modelling for the implementation of environmental flow project in the Zambezi River Basin. Zambezi River Authority, Hidroeléctrica de Cahora Bassa, ARA-Zambeze, ZESCO, Worldwide Wildlife Fund, International Crane Foundation, The University of Zambia, Universidade Eduardo Mondlane and UNESCO-IHE.

Jury, M. R. and M. E. Gwazantini. 2002. Climate variability in Malawi, part 2: sensitivity and prediction of lake levels. International journal of climatology **22**(11):1303-1312.

Kalma, J., T. McVicar, and M. McCabe. 2008. Estimating Land Surface Evaporation: A Review of Methods Using Remotely Sensed Surface Temperature Data. Surveys in Geophysics **29**(4-5):421-469.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter. 2002. NCEP-DOE AMIP-II reanalysis (R-2). Bulletin of the American Meteorological Society **83**(11):1631-1644.

Karlsson, J. M. and W. Arnberg. 2011. Quality analysis of SRTM and HYDRO1K: a case study of flood inundation in Mozambique. International Journal of Remote Sensing **32**(1):267-285.

Kasabov, N. 1996. Foundations of neural networks, fuzzy systems, and knowledge engineering. The MIT Press.

Katiyo, L. 2012. Overview of modelling studies in the Zambezi River Basin. Power2Flow project.

Kavetski, D., G. Kuczera, and S. W. Franks. 2006. Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. Journal of Hydrology **320**(1–2):173-186.

Kellndorfer, J., W. Walker, L. Pierce, C. Dobson, J. A. Fites, C. Hunsaker, J. Vona, and M. Clutter. 2004. Vegetation height estimation from Shuttle Radar Topography Mission and National Elevation Datasets. Remote Sensing of Environment **93**(3):339-358.

Kennedy, J. and R. Eberhart. 1995. Particle swarm optimization. Pages 1942-1948 *in* Neural Networks, 1995. Proceedings., IEEE International Conference on. IEEE.

Kerr, Y., P. Waldteufel, J. Wigneron, J. Martinuzzi, J. Font, and M. Berger. 2001. Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Transactions on Geoscience and Remote Sensing **39**(8):1729.

Kidd, C. and G. Huffman. 2011. Global precipitation measurement. Meteorological Applications **18**(3):334-353.

King, J. 2012. Establishing the process for developing and implementing environmental flows for the Zambezi River Basin., WWF, Cape Town, South Africa.

Klees, R., X. Liu, T. Wittwer, B. Gunter, E. Revtova, R. Tenzer, P. Ditmar, H. Winsemius, and H. Savenije. 2008. A comparison of global and regional GRACE models for land hydrology. Surveys in Geophysics **29**(4):335-359.

Klees, R., E. Zapreeva, H. Winsemius, and H. Savenije. 2006. The bias in GRACE estimates of continental water storage variations. Hydrology and Earth System Sciences Discussions **3**(6):3557-3594.

Köck, F., P. Meier, W. Kinzelbach, and I. Hajnsek. 2011. Modeling of the Kafue Flats Wetland, Zambia: Use of a TanDEM-X DEM and flood maps to improve the hydraulic model of overland flow and water retention in the floodplains.*in* TanDEM-X Science Team Meeting, German Aerospace Center, Oberpfaffenhofen, Germany.

Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. Biological cybernetics **43**(1):59-69.

Kohonen, T. and T. Honkela. 2007. Kohonen network. Scholarpedia **2**(1):1568.

Kozu, T., T. Kawanishi, H. Kuroiwa, M. Kojima, K. Oikawa, H. Kumagai, K. Okamoto, M. Okumura, H. Nakatsuka, and K. Nishikawa. 2001. Development of precipitation radar

onboard the Tropical Rainfall Measuring Mission (TRMM) satellite. IEEE Transactions on Geoscience and Remote Sensing **39**(1):102-116.

Krieger, G., A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink. 2007. TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry. Geoscience and Remote Sensing, IEEE Transactions on **45**(11):3317-3341.

Krzysztofowicz, R. 2002. Bayesian system for probabilistic river stage forecasting. Journal of Hydrology **268**(1–4):16-40.

Kubota, T., S. Shige, H. Hashizume, K. Aonashi, N. Takahashi, S. Seto, Y. Takayabu, T. Ushio, K. Nakagawa, and K. Iwanami. 2007. Global precipitation map using satellite-borne microwave radiometers by the GSMaP Project: Production and validation. Geoscience and Remote Sensing, IEEE Transactions on **45**(7):2259-2275.

Kumar, D. N., K. S. Raju, and T. Sathish. 2004. River flow forecasting using recurrent neural networks. Water resources management **18**(2):143-161.

Kustas, W. P. and J. M. Norman. 1996. Use of remote sensing for evapotranspiration monitoring over land surfaces. Hydrological Sciences Journal **41**(4):495-516.

Laloy, E. and J. A. Vrugt. 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing. Water Resources Research **48**(1):W01526.

Landert, J. 2008. Modeling Biogeochemistry using the Soil and Water Assessment Tool in the Zambezi River Basin. Swiss Federal Institute of Technology Zurich.

Lanza, L., J. Ramírez, and E. Todini. 2001. Stochastic rainfall interpolation and downscaling. Hydrology and Earth System Sciences **5**(2):139-143.

Lehner, B., K. Verdin, and A. Jarvis. 2006. HydroSHEDS technical documentation (available at

http://hydrosheds.cr.usgs.gov). World Wildlife Fund, Washington, DC, USA.

Lehner, B., K. Verdin, and A. Jarvis. 2008. New global hydrography derived from spaceborne elevation data. Eos Trans. AGU **89**(10):93-94.

Leopold, L., M. Wolman, and J. Miller. 1964. Fluvial process in geomorphology.

Lin, G.-F., G.-R. Chen, P.-Y. Huang, and Y.-C. Chou. 2009. Support vector machine-based models for hourly reservoir inflow forecasting during typhoon-warning periods. Journal of Hydrology **372**(1):17-29.

Lin, G.-F. and L.-H. Chen. 2004. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. Journal of Hydrology **288**(3-4):288-298.

Lin, J.-Y., C.-T. Cheng, and K.-W. Chau. 2006. Using support vector machines for long-term discharge prediction. Hydrological Sciences Journal **51**(4):599-612.

Liong, S.-Y. and C. Sivapragasam. 2002. Flood stage forecasting with support vector machines. Journal of the American Water Resources Association **38**(1):173-186.

Lloyd, C. 2005. Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. Journal of Hydrology **308**(1-4):128-150.

Londhe, S. and S. Charhate. 2010. Comparison of data-driven modelling techniques for river flow forecasting. Hydrological Sciences Journal **55**(7):1163-1174.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. W. M. J. Yang, and J. W. Merchant. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing **21**(6-7):1303-1330.

MacKay, D. J. 1992a. A practical Bayesian framework for backpropagation networks. Neural computation **4**(3):448-472.

MacKay, D. J. C. 1992b. Bayesian interpolation. Neural computation **4**(3):415-447.

Mangiameli, P., S. K. Chen, and D. West. 1996. A comparison of SOM neural network and hierarchical clustering methods. European Journal of Operational Research **93**(2):402-417.

Mantovan, P. and E. Todini. 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. Journal of Hydrology **330**(1):368-381.

Mantovan, P., E. Todini, and M. L. V. Martina. 2007. Reply to comment by Keith Beven, Paul Smith and Jim Freer on "Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology". Journal of Hydrology **338**(3–4):319-324.

Marschalk, U., A. Roth, M. Eineder, and S. Suchandt. 2004. Comparison of DEMs derived from SRTM/X- and C-band. Pages 4531-4534 vol.4537 *in* Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International.

Martone, M., P. Rizzoli, B. Bräutigam, and G. Krieger. 2013. First 2 years of TanDEM-X mission: Interferometric performance overview. Radio Science **48**(5):617-627.

Matheron, G. 1969. Le krigeage universel. École nationale supérieure des mines de Paris.

Matondo, J. and P. Mortensen. 1998. Water resource assessment for the Zambezi river basin. Water International **23**(4):256-262.

Matos, J. P. 2008. Using artificial neural networks to model the degradation of water supply systems (Utilização de redes neuronais artificiais para a modelação da degradação de sistemas de distribuição de água). Master's thesis. Technical University of Lisbon, Instituto Superior Técnico, Lisbon, Portugal.

Matos, J. P., T. Cohen Liechti, D. Juízo, M. M. Portela, and A. J. Schleiss. 2013a. Can satellite based pattern-oriented memory improve the interpolation of sparse historical rainfall records? Journal of Hydrology **492**(7):102-116.

Matos, J. P., T. Cohen Liechti, M. M. Portela, and A. J. Schleiss. 2013b. Coupling satellite rainfall estimates and machine learning techniques for flow forecast: application to a large catchment in southern africa.*in* Proceedings of 35th IAHR World Congress. Tsinghua University Press, Chengdu, China.

Matos, J. P., T. Cohen Liechti, M. M. Portela, and A. J. Schleiss. 2014. Pattern-oriented memory interpolation of sparse historical rainfall records. Journal of Hydrology **510**:493-503.

Matos, J. P., T. Cohen, J.-L. Boillat, A. J. Schleiss, and M. M. Portela. 2010. Analysis of flow regime changes due to operation of large reservoirs on the Zambezi River. 6th Internationational Symposyium on Environmental Hydraulics, Athens, Greece.

Mayaux, P., E. Bartholomé, S. Fritz, and A. Belward. 2004. A new land-cover map of Africa for the year 2000. Journal of Biogeography **31**(6):861-877.

Mazvimavi, D. and P. Wolski. 2006. Long-term variations of annual flows of the Okavango and Zambezi Rivers. Physics and Chemistry of the Earth **31**:944-951.

McMillan, H., T. Krueger, and J. Freer. 2012. Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. Hydrological Processes **26**(26):4078-4111.

Meier, P. 2012. Real-time hydrologic modelling and floodplain modelling in the Kafue river basin, Zambia Ph.D. Swiss Federal Institute of Technology, Zurich, Zurich, Switzerland.

Meier, P., A. Frömelt, and W. Kinzelbach. 2011. Hydrological real-time modelling in the Zambezi river basin using satellite-based soil moisture and rainfall data. Hydrology and Earth System Sciences **15**(3):999-1008.

Meier, P. and W. Kinzelbach. 2010. Modeling flooding patterns in the Kafue Flats, Zambia.*in* EGU General Assembly 2010, Vienna, Austria.

Mekanik, F., M. Imteaz, S. Gato-Trinidad, and A. Elmahdi. 2013. Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology **503**:11-21.

Mertens, J. 2013. African Dams Project: an integrated water resources managament study. Final stakeholder report. ETH Zurich, Zurich.

Meskele, T., H. Moradkhani, and M. PE. 2009. Impacts of Different Rainfall Estimates on Hydrological Simulation and Satellite Rainfall Retrieval Error Propagation. ASCE.

Michailovsky, C. I. 2008. Comparing GRACE water storage observations and regional-scale hydrological models for Southern Africa. Technical University of Denmark.

Michailovsky, C. I. 2013. Updating river basin models with radar altimetry. Ph.D. thesis. Technical University of Denmark, Lyngby, Denmark.

Michailovsky, C. I. and P. Bauer-Gottwein. 2013. Operational reservoir inflow forecasting with radar altimetry: the Zambezi case study. Hydrol. Earth Syst. Sci. Discuss. **10**(7):9615-9644.

Michailovsky, C. I., S. McEnnis, P. A. M. Berry, R. Smith, and P. Bauer-Gottwein. 2012. River monitoring from satellite radar altimetry in the Zambezi River Basin. Hydrol. Earth Syst. Sci. Discuss. **9**(3):3203-3235.

Mohammadi, K., H. R. Eslami, and R. Kahawita. 2006. Parameter estimation of an ARMA model for river flow forecasting using goal programming. Journal of Hydrology **331**(1):293-299.

Montanari, A. 2007. What do we mean by 'uncertainty'? The need for a consistent wording about uncertainty assessment in hydrology. Hydrological Processes **21**(6):841-845.

Moore, A. E., F. P. D. Cotterill, M. P. L. Main, and H. B. Williams. 2008. The Zambezi River. Pages 311-332 Large rivers: geomorphology and management. John Wiley & Sons, Ltd, Chichester, UK.

Moradkhani, H. and S. Sorooshian. 2008. General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis. Pages 1-24 *in* S. Sorooshian, K.-L. Hsu, E. Coppola, B. Tomassetti, M. Verdecchia, and G. Visconti, editors. Hydrological modelling and the water cycle. Springer Berlin Heidelberg.

Moreira, A., G. Krieger, I. Hajnsek, D. Hounam, M. Werner, S. Riegger, and E. Settelmeyer. 2004. TanDEM-X: a TerraSAR-X add-on satellite for single-pass SAR interferometry. Pages 1000-1003 *in* Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International. IEEE.

Moriasi, D., B. Wilson, K. Douglas-Mankin, J. Arnold, and P. Gowda. 2012. Hydrologic and water quality models: use, calibration, and validation. Transactions of the ASABE **55**(4):1241-1247.

Mwelwa, E. M. 2004. The application of the monthly time step Pitman rainfall-runoff model to the Kafue River Basin of Zambia. Master. Rhodes University, Grahamstown, South Africa.

Myers, D. E. 1982. Matrix formulation of co-kriging. Mathematical Geology **14**(3):249-257.

Nachtergaele, F., H. van Velthuizen, L. Verelst, N. Batjes, K. Dijkshoorn, V. van Engelen, G. Fischer, A. Jones, L. Montanarella, M. Petri, S. Prieler, E. Teixeira, D. Wilberg, and X. Shi. 2008. Harmonized World soil database version 1.0. FAO/IIASA/ISRIC/ISSCAS/JRC, Rome, Italy and Laxenburg, Austria.

Nash, J. and J. Sutcliffe. 1970. River flow forecasting through conceptual models part I - A discussion of principles. Journal of Hydrology **10**(3):282-290.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams. 2011. Soil and water assessment tool theoretical documentation, version 2009. Grassland, Soil and Water Research Laboratory, Agricultural Research Service; Brackland Research Center, Texas Agricultural Experiment Station, Temple, Texas, USA.

NOAA Climate Prediction Center. 2001. The NOAA Climate Prediction Center african rainfall estimation algorithm version 2. 0.

Pegram, G., I. Deyzel, S. Sinclair, P. Visser, D. Terblanche, and G. Green. 2004. Daily mapping of 24 hr rainfall at pixel scale over South Africa using satellite, radar and raingauge data.

Perrin, M. 2013. Gestion intégrée des ressources en eau du bassin du Zambèze. Master's thesis. École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Petty, G. W. and W. F. Krajewski. 1996. Satellite estimation of precipitation over land. Hydrological Sciences Journal **41**(4):433-451.

Pilgrim, D., T. Chapman, and D. Doran. 1988. Problems of rainfall-runoff modelling in arid and semiarid regions/Problèmes de la mise au point de modèles pluie-écoulement dans les régions arides et semi-arides. Hydrological Sciences Journal **33**(4):379-400.

Piotrowski, A. P. and J. J. Napiorkowski. 2011. Optimizing neural networks for river flow forecasting – Evolutionary Computation methods versus the Levenberg–Marquardt approach. Journal of Hydrology **407**(1):12-27.

Rahman, K., C. Maringanti, M. Beniston, F. Widmer, K. Abbaspour, and A. Lehmann. 2013. Streamflow Modeling in a Highly Managed Mountainous Glacier Watershed Using SWAT: The Upper Rhone River Watershed Case in Switzerland. Water resources management **27**(2):323-339.

Refsgaard, J. C. 1997. Parameterisation, calibration and validation of distributed hydrological models. Journal of Hydrology **198**(1–4):69-97.

Reusser, D. E., W. Buytaert, and E. Zehe. 2011. Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test. Water Resources Research **47**(7):W07551.

Reusser, D. E. and E. Zehe. 2011. Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity. Water Resources Research **47**(7):W07550.

Reuter, H. I., A. Nelson, and A. Jarvis. 2007. An evaluation of void-filling interpolation methods for SRTM data. International Journal of Geographical Information Science **21**(9):983-1008.

Reynolds, C. A., T. J. Jackson, and W. J. Rawls. 2000. Estimating soil water-holding capacities by linking the Food and Agriculture Organization Soil map of the world with global pedon databases and continuous pedotransfer functions. Water Resources Research **36**(12):3653-3662.

Reynolds, R. W. 1988. A Real-Time Global Sea Surface Temperature Analysis. Journal of Climate **1**(1):75-87.

Rodríguez-Iturbe, I. and A. Rinaldo. 2001. Fractal river basins: chance and self-organization. Cambridge University Press.

Rodriguez, E., C. S. Morris, J. E. Belz, E. C. Chapin, J. M. Martin, W. Daffer, and S. Hensley. 2005. An assessment of the SRTM topographic products. NASA Jet Propulsion Laboratory.

Roerink, G. J., Z. Su, and M. Menenti. 2000. S-SEBI: A simple remote sensing algorithm to estimate the surface energy balance. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere **25**(2):147-157.

Ronco, P. 2008. Sediment budget of unsurveyed rivers at watershed scale: the case of Lower Zambezi. Ph.D. thesis. Università Degli Studi di Padova, Padova.

Ronco, P., G. Fasolato, M. Nones, and G. Di Silvio. 2010. Morphological effects of damming on lower Zambezi River. Geomorphology **115**(1):43-55.

Rosenblatt, F. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review **65**(6):386.

Rouhani, H., D. Gorissen, P. Willems, and J. Feyen. 2007. Improved rainfall-runoff modeling combining a semi-distributed model with artificial neural networks.*in* Proceedings of the SWAT Workshops and Conference.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1985. Learning internal representations by error propagation. DTIC Document.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning representations by back-propagating errors. Nature **323**.

Saltelli, A., S. Tarantola, and K. P. S. Chan. 1999. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. Technometrics **41**(1):39-56.

Samsudin, R., P. Saad, and A. Shabri. 2011. River flow time series using least squares support vector machines. Hydrology and Earth System Sciences **15**(6):1835-1852.

Schaibly, J. H. and K. E. Shuler. 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II Applications. The Journal of Chemical Physics **59**(8):3879-3888.

Schölkopf, B., A. J. Smola, R. C. Williamson, and P. L. Bartlett. 2000. New support vector algorithms. Neural computation **12**(5):1207-1245.

Schoups, G. and J. A. Vrugt. 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. Water Resources Research **46**(10):W10531.

Schuol, J., K. C. Abbaspour, R. Srinivasan, and H. Yang. 2008a. Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. Journal of Hydrology **352**(1–2):30-49.

Schuol, J., K. C. Abbaspour, H. Yang, R. Srinivasan, and A. J. B. Zehnder. 2008b. Modeling blue and green water availability in Africa. Water Resour. Res **44**:W07406.

Scipal, K., C. Scheffler, and W. Wagner. 2005. Soil moisture-runoff relation at the catchment scale as observed with coarse resolution microwave remote sensing. Hydrol. Earth Syst. Sci. **9**(3):173-183.

Shamseldin, A. Y. and K. M. O'Connor. 2001. A non-linear neural network technique for updating of river flow forecasts. Hydrology and Earth System Sciences **5**(4):577-597.

Shanno, D. F. 1970. Conditioning of quasi-Newton methods for function minimization. Mathematics of computation **24**(111):647-656.

Shepard, D. 1968. A two-dimensional interpolation function for irregularly-spaced data. Pages 517-524 *in* Proceedings of the 1968 23rd ACM national conference

Simpson, J., C. Kummerow, W. Tao, and R. Adler. 1996. On the tropical rainfall measuring mission (TRMM). Meteorology and Atmospheric physics **60**(1):19-36.

Sinclair, S. and G. Pegram. 2005. Combining radar and rain gauge rainfall estimates using conditional merging. Atmospheric Science Letters **6**(1):19-22.

Sivakumar, B., A. W. Jayawardena, and T. M. K. G. Fernando. 2002. River flow forecasting: use of phase-space reconstruction and artificial neural networks approaches. Journal of Hydrology **265**(1-4):225-245.

Sivapragasam, C., V. Arun, and D. Giridhar. 2010. A simple approach for improving spatial interpolation of rainfall using ANN. Meteorology and Atmospheric physics:1-7.

Sivapragasam, C., S. Vanitha, N. Muttil, K. Suganya, S. Suji, M. T. Selvi, R. Selvi, and S. J. Sudha. 2013. Monthly flow forecast for Mississippi River basin using artificial neural networks. Neural Computing and Applications.

Smith, R. G. and P. A. M. Berry. 2010. ACE2: global digital elevation model. EAPRS Laboratory, De Montfort University, Leicester, UK.

Sorooshian, S. and V. K. Gupta. 1983. Automatic calibration of conceptual rainfall-runoff models: The question of parameter observability and uniqueness. Water Resources Research **19**(1):260-268.

Sorooshian, S., K. Hsu, X. Gao, H. Gupta, B. Imam, and D. Braithwaite. 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. Bulletin of the American Meteorological Society **81**(9):2035-2046.

Srinivas, N. and K. Deb. 1994. Muiltiobjective optimization using nondominated sorting in genetic algorithms. Evolutionary computation **2**(3):221-248.

Srinivasan, R., X. Zhang, and J. Arnold. 2010. SWAT ungauged: hydrological budget and crop yield predictions in the Upper Mississippi River Basin. Transactions of the ASABE **53**(5):1533-1546.

Storn, R. and K. Price. 1997. Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization **11**(4):341-359.

Suykens, J. A. K., T. V. Gestel, J. D. Brabanter, B. D. Moor, and J. Vandewalle. 2002. Least squares support vector machines. World Scientific Pub. Co. Inc., Singapore.

Suykens, J. A. K. and J. Vandewalle. 1999. Least squares support vector machine classifiers. Neural processing letters **9**(3):293-300.

Swinnen, E. and F. Veroustraete. 2008. Extending the SPOT-VEGETATION NDVI time series (1998-2006) back in time with NOAA-AVHRR data (1985-1998) for Southern Africa. Geoscience and Remote Sensing, IEEE Transactions on **46**(2):558-572.

SWRSD Zambezi Basin Joint Venture. 2011. Dam synchronization and flood releases in the Zambezi River Basin project., Consultancy report for the German Federal Ministry for Economic Cooperation and Development (GTZ) and the UK Department for International Development (DIFD).

Tachikawa, T., M. Hato, M. Kaku, and A. Iwasaki. 2011a. Characteristics of ASTER GDEM version 2. Pages 3657-3660 *in* Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International.

Tachikawa, T., M. Kaku, A. Iwasaki, D. Gesch, M. Oimoen, Z. Zhang, J. Danielson, T. Krieger, B. Curtis, and J. Haase. 2011b. ASTER global digital elevation model version 2 - summary of validation results. ASTER GDEM Validation Team.

Takeda, H., S. Farsiu, and P. Milanfar. 2007. Kernel regression for image processing and reconstruction. Image Processing, IEEE Transactions on **16**(2):349-366.

Tapiador, F. J., C. Kidd, V. Levizzani, and F. S. Marzano. 2010. A neural networks–based fusion technique to estimate half-hourly rainfall estimates at 0.1 resolution from satellite passive microwave and infrared data. Journal of Applied Meteorology **43**(4).

The World Bank. 2010. The Zambezi River basin. A multi-sector investment opportunities analysis. The World Bank, Washington DC, USA.

Thiessen, A. H. 1911. Precipitation averages for large areas. Monthly weather review **39**(7):1082-1089.

Thirumalaiah, K. and M. Deo. 1998. River stage forecasting using artificial neural networks. Journal of Hydrologic Engineering **3**(1):26-32.

Tilmant, A., W. Kinzelbach, L. Beevers, and D. Juizo. 2010. Optimal water allocation in the Zambezi Basin.*in* 2010 International Congress on Environmental Modelling and Software, Ottawa, Canada.

Tobin, C., L. Nicotina, M. B. Parlange, A. Berne, and A. Rinaldo. 2011. Improved interpolation of meteorological forcings for hydrologic applications in a Swiss Alpine region. Journal of Hydrology **401**(1-2):77-89.

Toth, E. and A. Brath. 2007. Multistep ahead streamflow forecasting: Role of calibration data in conceptual and neural network modeling. Water Resources Research **43**(11):W11405.

Tucker, C. J., J. E. Pinzon, M. E. Brown, D. A. Slayback, E. W. Pak, R. Mahoney, E. F. Vermote, and N. El Saleous. 2005. An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. International Journal of Remote Sensing **26**(20):4485-4498.

Tucker, C. J., J. R. G. Townshend, and T. E. Goff. 1985. African Land-Cover Classification Using Satellite Data. Science **227**(4685):369-375.

Tyson, P. D., G. R. J. Cooper, and T. S. McCarthy. 2002. Millennial to multi-decadal variability in the climate of southern Africa. International Journal of Climatology **22**(9):1105-1117.

U.S. Geological Survey. 2001. HYDRO1k elevation derivative database. U.S. Geological Survey (USGS), Washington, DC, USA.

Ushio, T., K. Sasashige, T. Kubota, S. Shige, K. i. Okamoto, K. Aonashi, T. Inoue, N. Takahashi, T. Iguchi, M. Kachi, R. Oki, T. Morimoto, and Z.-I. Kawasaki. 2009. A Kalman Filter Approach to the Global Satellite Mapping of Precipitation (GSMaP) from Combined Passive Microwave and Infrared Radiometric Data. Journal of the Meteorological Society of Japan. Ser. II **87A**:137-151.

Valyon, J. and G. Horváth. 2007. A Robust LS-SVM Regression. International Journal of Computer, Information Science and Engineering **1**(7):692-697.

van Griensven, A. and W. Bauwens. 2003. Multiobjective autocalibration for semidistributed water quality models. Water Resources Research **39**(12):1348.

van Leeuwen, W. J. D., A. R. Huete, and T. W. Laing. 1999. MODIS Vegetation Index Compositing Approach: A Prototype with AVHRR Data. Remote Sensing of Environment **69**(3):264-280.

van Liew, M. W., J. Arnold, and D. Bosch. 2005. Problems and potential of autocalibrating a hydrologic model. Transactions of the ASAE **48**(3):1025-1040.

van Zyl, J. J. 2001. The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography. Acta Astronautica **48**(5-12):559-565.

Vicente-Serrano, S. M., M. A. Saz Sánchez, and J. M. Cuadrat. 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): Application to annual precipitation and temperature. Climate research **24**(2):161-180.

Vörösmarty, C. J. and B. Moore. 1991. Modelling basin-scale hydrology in support of physical climate and global biogeochemical studies: An example using the Zambezi River. Surveys in Geophysics **12**(1-3):271-311.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian. 2003. Effective and efficient algorithm for multiobjective optimization of hydrologic models. Water Resources Research **39**(8):1214.

Vrugt, J. A. and E. Laloy. 2014. Reply to comment by Chu et al. on "High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing". Water Resources Research:n/a-n/a.

Vrugt, J. A. and B. A. Robinson. 2007. Improved evolutionary optimization from genetically adaptive multimethod search. Proceedings of the National Academy of Sciences **104**(3):708.

Vrugt, J. A., B. A. Robinson, and J. M. Hyman. 2009a. Self-adaptive multimethod search for global optimization in real-parameter spaces. Evolutionary Computation, IEEE Transactions on **13**(2):243-259.

Vrugt, J. A., C. J. Ter Braak, M. P. Clark, J. M. Hyman, and B. A. Robinson. 2008. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. Water Resources Research **44**(12).

Vrugt, J. A., C. J. Ter Braak, H. V. Gupta, and B. A. Robinson. 2009b. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic environmental research and risk assessment **23**(7):1011-1026.

Wagner, W., G. Bloschl, P. Pampaloni, J.-C. Calvet, B. Bizzarri, J.-P. Wigneron, and Y. Kerr. 2007. Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. Nordic hydrology **38**(1):1-20.

Wahr, J., M. Molenaar, and F. Bryan. 1998. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical Research: Solid Earth **103**(B12):30205-30229.

Wang, W.-C., K.-W. Chau, C.-T. Cheng, and L. Qiu. 2009. A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of Hydrology **374**(3-4):294-306.

Wechsler, S. P. 2007. Uncertainties associated with digital elevation models for hydrologic applications: a review. Hydrol. Earth Syst. Sci. **11**(4):1481-1500.

Winchell, M., R. Srinivasan, M. Di Luzio, and J. G. Arnold. 2010. ArcSWAT interface for SWAT2009. User's guide. Brackland Research and Extension Center, Texas AgriLife Research; Grassland, Soil and Water Research Laboratory, USDA Agricultural Research Service, Temple, Texas, USA.

Winsemius, H., B. Schaefli, A. Montanari, and H. Savenije. 2009. On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. Water Resour. Res **45**(W12422):W12422.

Winsemius, H. C., H. H. G. Savenije, A. M. J. Gerrits, E. A. Zapreeva, and R. Klees. 2006a. Comparison of two model approaches in the Zambezi river basin with regard to model reliability and identifiability. Hydrol. Earth Syst. Sci. **10**(3):339-352.

Winsemius, H. C., H. H. G. Savenije, N. C. van de Giesen, B. J. J. M. van den Hurk, E. A. Zapreeva, and R. Klees. 2006b. Assessment of Gravity Recovery and Climate Experiment (GRACE) temporal signature over the upper Zambezi. Water Resources Research **42**(12):W12201.

Wolpert, D. H. 1996a. The Existence of A Priori Distinctions Between Learning Algorithms. Neural computation **8**(7):1391-1420.

Wolpert, D. H. 1996b. The Lack of A Priori Distinctions Between Learning Algorithms. Neural computation **8**(7):1341-1390.

Wolpert, D. H. and W. G. Macready. 1997. No free lunch theorems for optimization. Evolutionary Computation, IEEE Transactions on **1**(1):67-82.

Woodbury, A. D. and T. J. Ulrych. 2000. A full-Bayesian approach to the groundwater inverse problem for steady state flow. Water Resources Research **36**(8):2081-2093.

Wu, C. L. and K. W. Chau. 2010. Data-driven models for monthly streamflow time series prediction. Engineering Applications of Artificial Intelligence **23**(8):1350-1367.

Wu, C. L., K. W. Chau, and Y. S. Li. 2009. Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. Water Resources Research **45**(8):W08432.

Xie, P. and P. Arkin. 1996. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. Journal of Climate **9**(4):840-858.

Yang, J., P. Reichert, K. Abbaspour, J. Xia, and H. Yang. 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. Journal of Hydrology **358**(1):1-23.

Yapo, P. O., H. V. Gupta, and S. Sorooshian. 1996. Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. Journal of Hydrology **181**(1–4):23-48.

Yapo, P. O., H. V. Gupta, and S. Sorooshian. 1998. Multi-objective global optimization for hydrologic models. Journal of Hydrology **204**(1–4):83-97.

Yew Gan, T., E. M. Dlamini, and G. F. Biftu. 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. Journal of Hydrology **192**(1–4):81-103.

Yu, P.-S., S.-T. Chen, and I.-F. Chang. 2006. Support vector regression for real-time flood stage forecasting. Journal of Hydrology **328**(3):704-716.

Zarchan, P. and H. Musoff. 2009. Fundamentals of Kalman filtering: a practical approach. 3 edition. American Institute of Aeronautics and Astronautics, Reston, USA.

Zhang, X., R. Srinivasan, J. Arnold, R. C. Izaurralde, and D. Bosch. 2011. Simultaneous calibration of surface flow and baseflow simulations: A revisit of the SWAT model calibration framework. Hydrological Processes **25**(14):2313-2320.

Zhang, X., R. Srinivasan, and M. V. Liew. 2010. On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model. Hydrological Processes **24**(8):955-969.

Zhang, X., R. Srinivasan, K. Zhao, and M. V. Liew. 2009. Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. Hydrological Processes **23**(3):430-441.

Zulkafli, Z., W. Buytaert, C. Onof, B. Manz, E. Tarnavsky, W. Lavado, and J.-L. Guyot. 2013. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean-Amazon river basins. Journal of Hydrometeorology.

APPENDIX I

APPENDIX I.A. MAIN SUBBASINS

In this appendix the location and main features of the main subbasins of the Zambezi are identified and located.
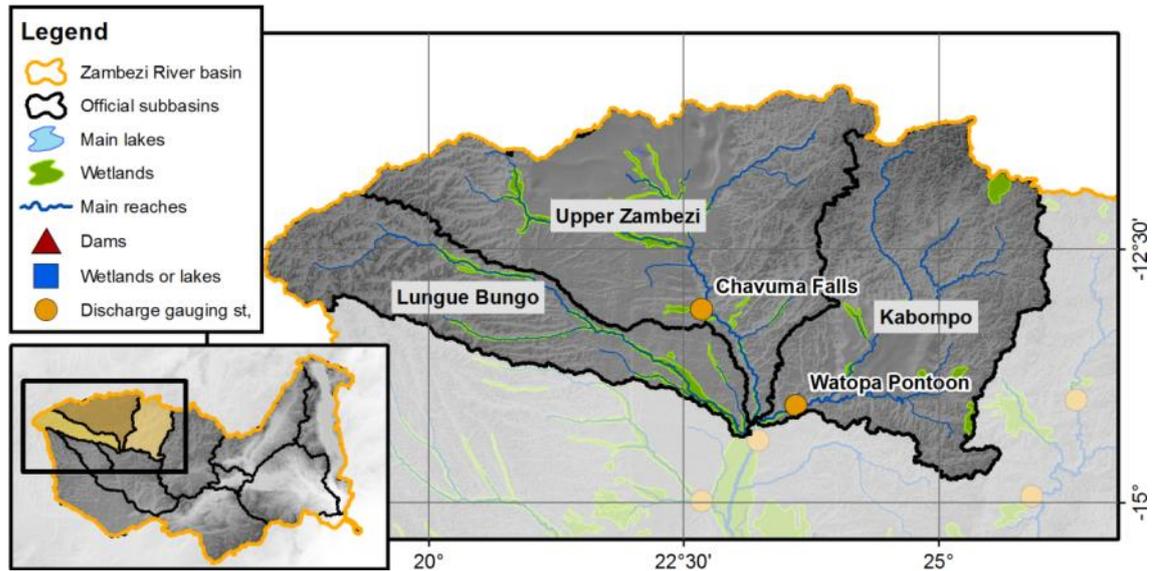

**Figure AI.1. Location and main features of the Upper Zambezi, Lungue Bungo and Kabompo subbasins.**
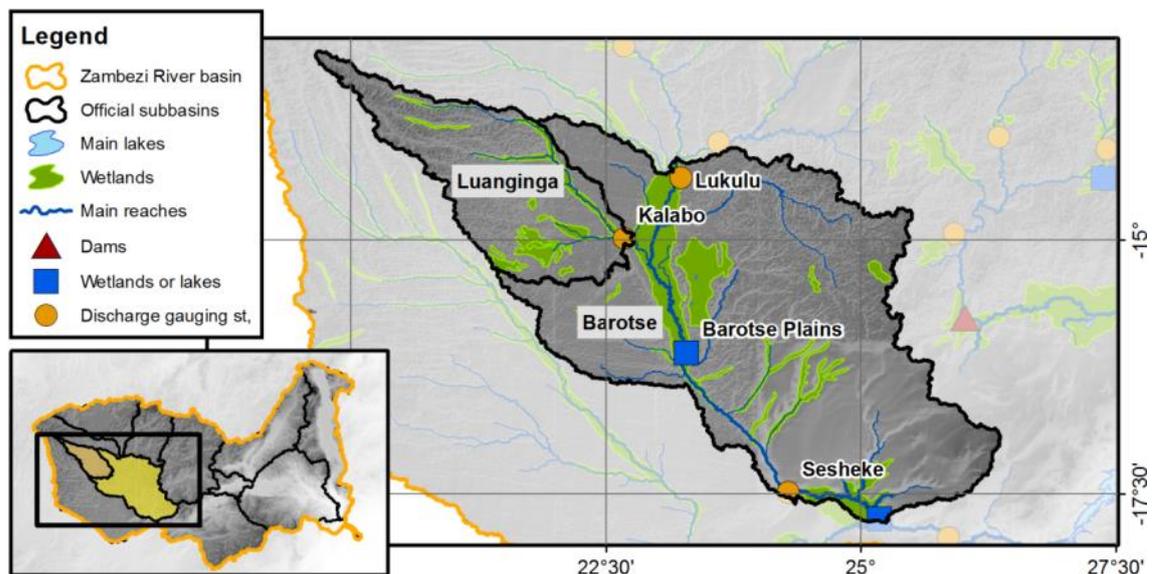

**Figure AI.2. Location and main features of the Barotse and Luanginga subbasins.**
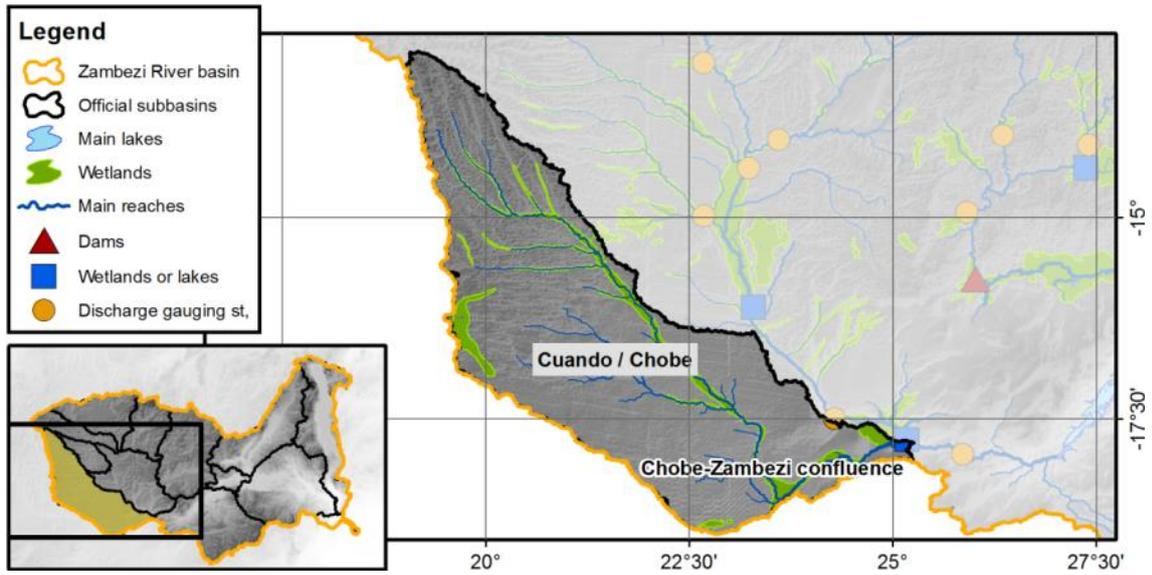
**Figure AI.3. Location and main features of the Cuando/Chobe subbasin.**
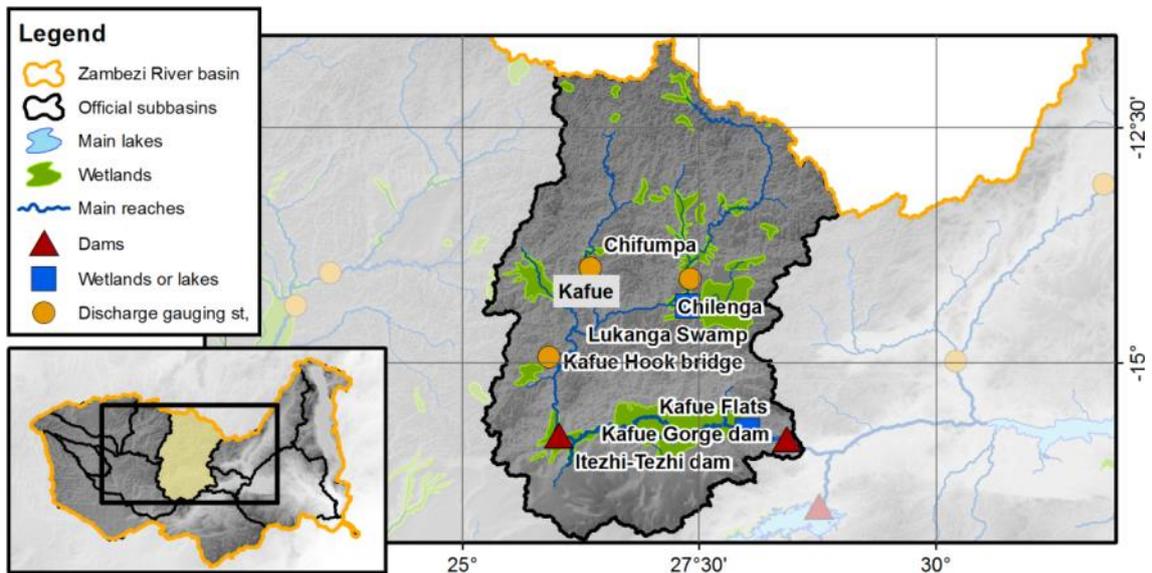
Figure AI.4. Location and main features of the Kafue subbasin.
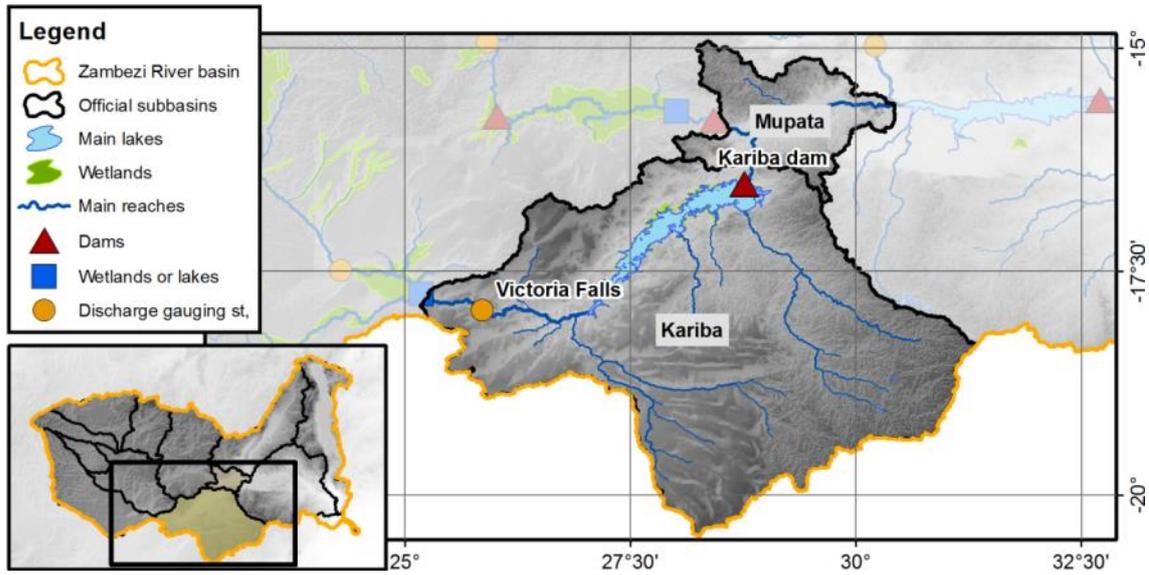
Figure AI.5. Location and main features of the Kariba and Mupata subbasins.
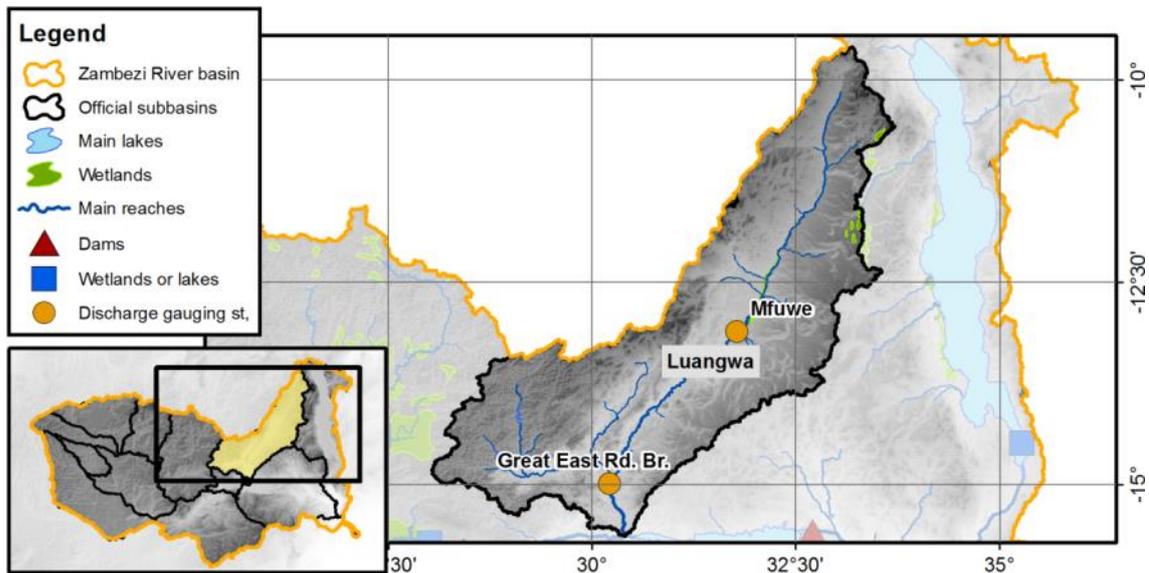
Figure AI.6. Location and main features of the Luangwa subbasin.
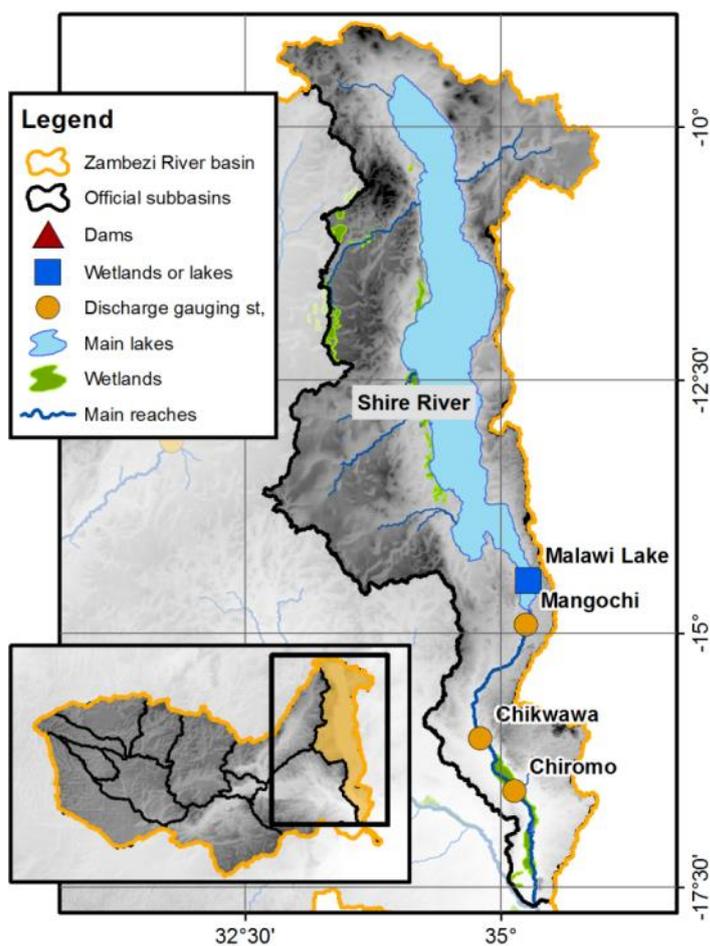
Figure AI.7. Location and main features of the Shire River subbasin.
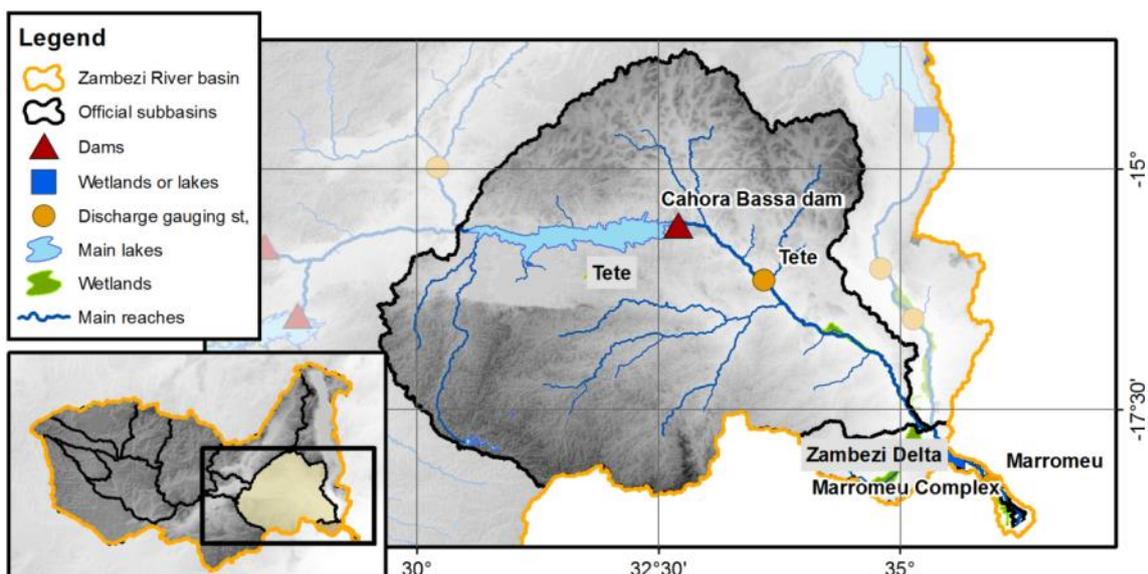
Figure AI.8. Location and main features of the Tete and Zambezi Delta subbasins.

APPENDIX I.B. MAIN CHARACTERISTICS OF LAKES, RESERVOIRS AND WETLANDS
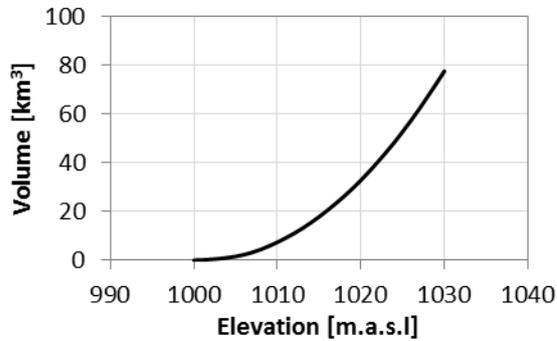
Figure AI.9. Elevation-volume curve adopted for the Barotse Plains.
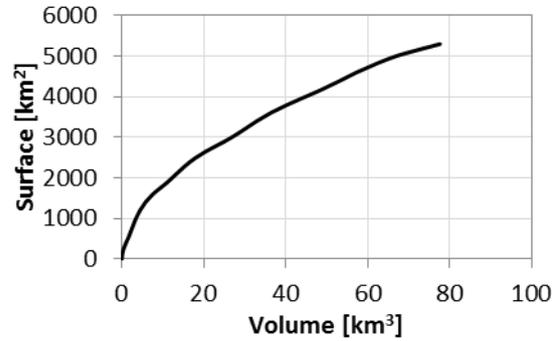
Figure AI.10. Volume-surface curve adopted for the Barotse Plains.
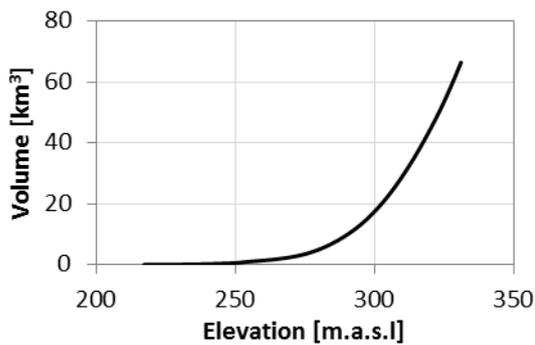
Figure AI.11. Elevation-volume curve adopted for the Cahora Bassa reservoir.
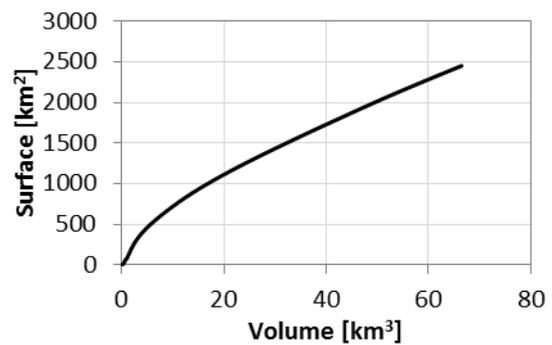
Figure AI.12. Volume-surface curve adopted for the Cahora Bassa reservoir.
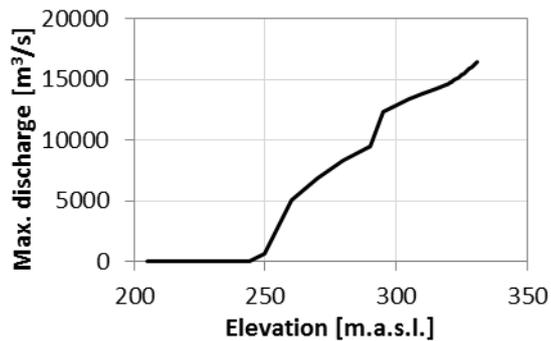
Figure AI.13. Maximum discharge curve adopted for the Cahora Bassa dam.
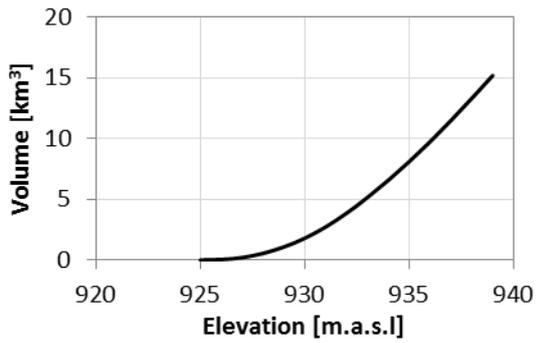
**Figure AI.14. Elevation-volume curve adopted for the Chobe-Zambezi Confluence.**
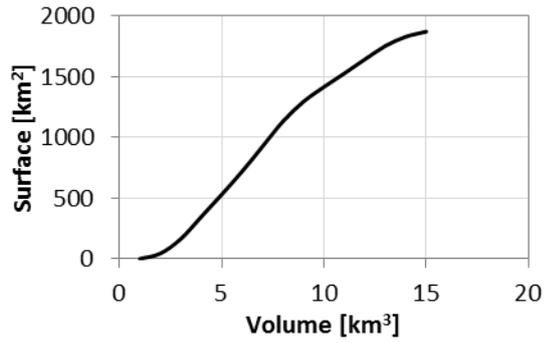
Figure AI.15. Volume-surface curve adopted for the Chobe-Zambezi Confluence.
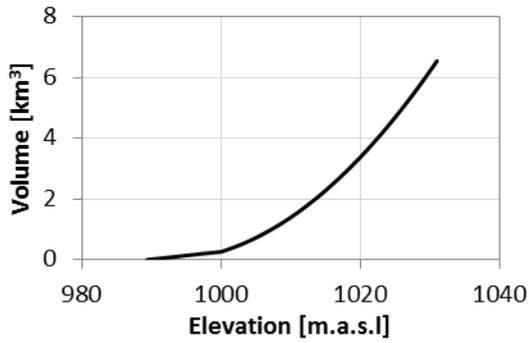

**Figure AI.16. Elevation-volume curve adopted for the Itezhi-Tezhi reservoir.**
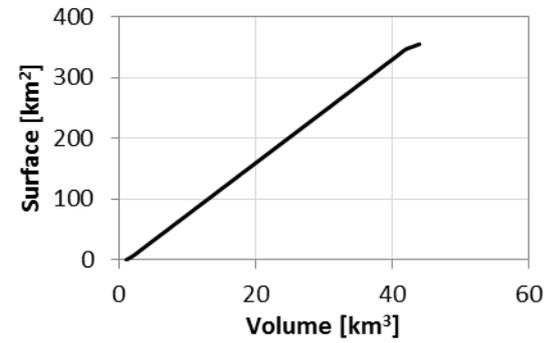
Figure AI.17. Volume-surface curve adopted for the Itezhi-Tezhi reservoir.
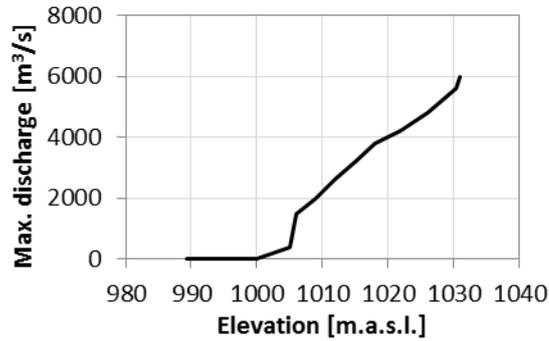

**Figure AI.18. Maximum discharge curve adopted for the Itezhi-Tezhi dam.**
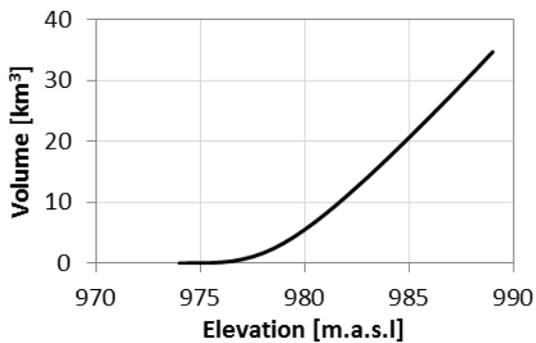
Figure AI.19. Elevation-volume curve adopted for the Kafue Flats.
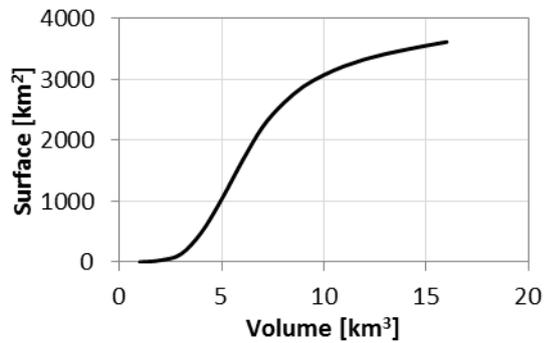
Figure AI.20. Volume-surface curve adopted for the Kafue Flats.
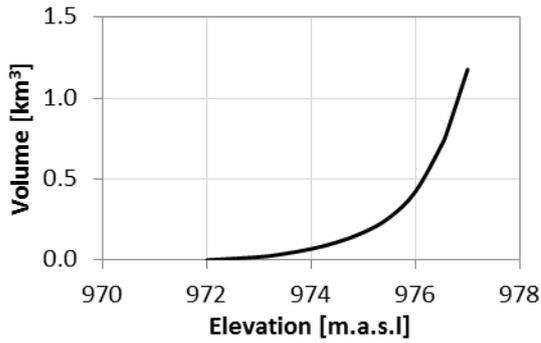
Figure AI.21. Elevation-volume curve adopted for the Kafue Gorge reservoir.
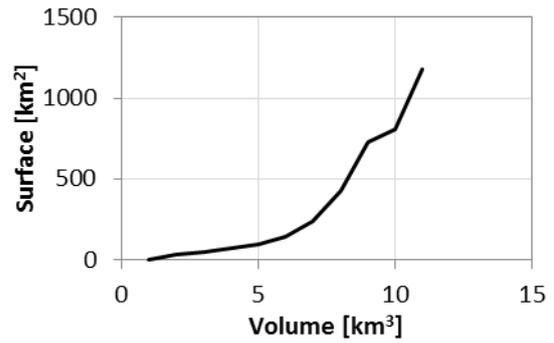
Figure AI.22. Volume-surface curve adopted for the Kafue Gorge reservoir.
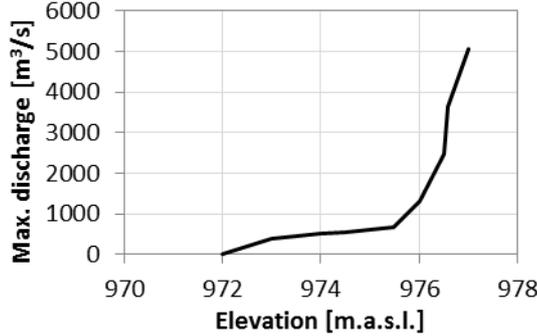
Figure AI.23. Maximum discharge curve adopted for the Kafue Gorge dam.
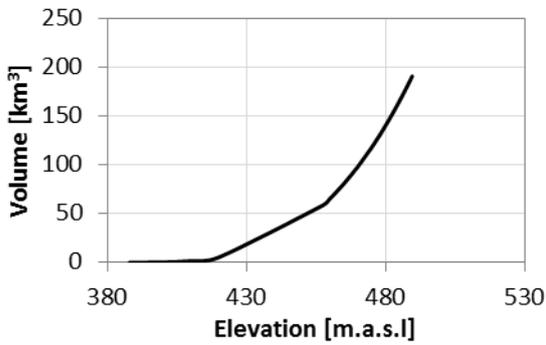
Figure AI.24. Elevation-volume curve adopted for the Kafue Gorge reservoir.
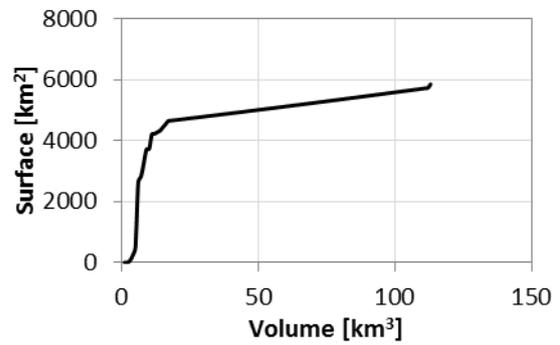
Figure AI.25. Volume-surface curve adopted for the Kafue Gorge reservoir.
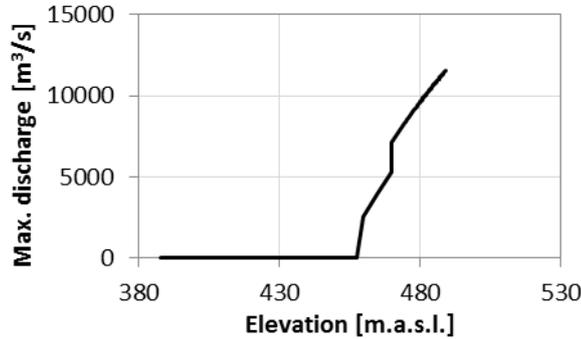
Figure AI.26. Maximum discharge curve adopted for the Kafue Gorge dam.
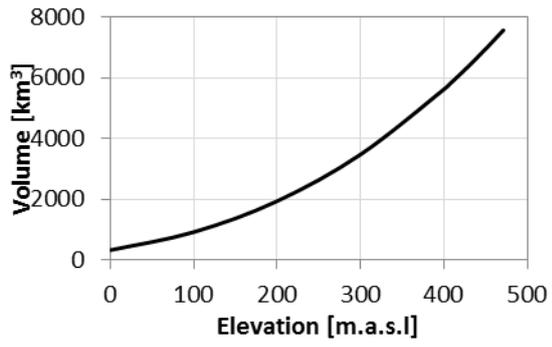
**Figure AI.27. Elevation-volume curve adopted for the Kafue Flats.**
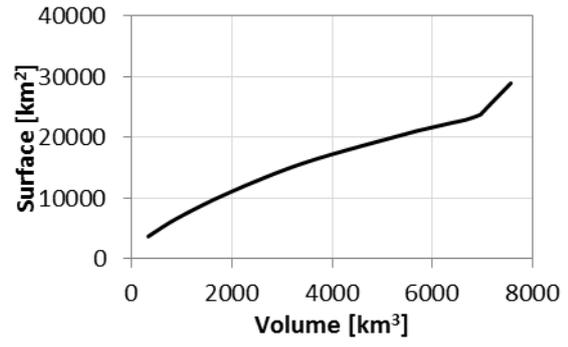
Figure AI.28. Volume-surface curve adopted for the Kafue Flats.

APPENDIX I.C. MAIN DISCHARGE AND WATER LEVEL SERIES USED IN THE SCOPE OF THE WORK
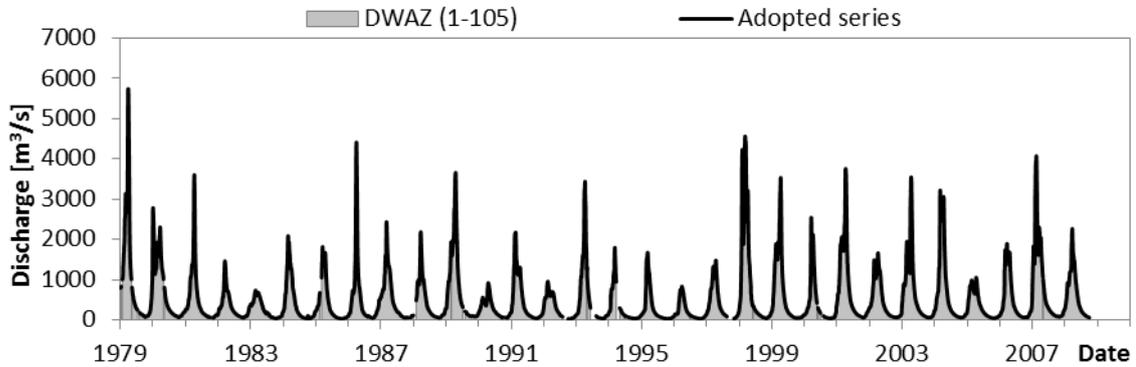
Figure AI.29. Discharge data for Chavuma Falls.
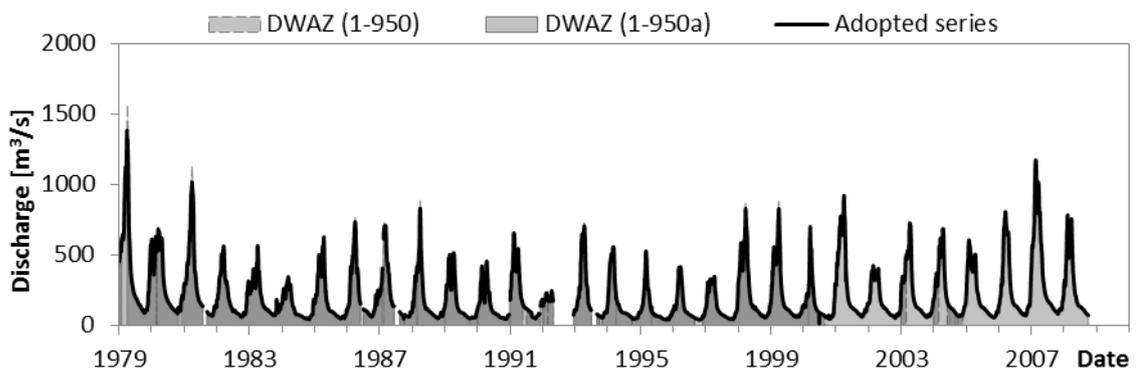
Figure AI.30. Discharge data for Watopa Pontoon.
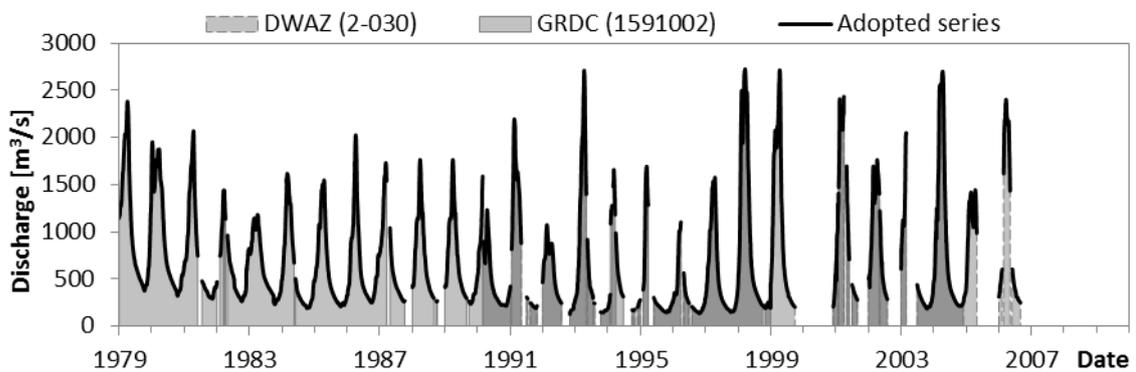
Figure AI.31. Discharge data for Lukulu.
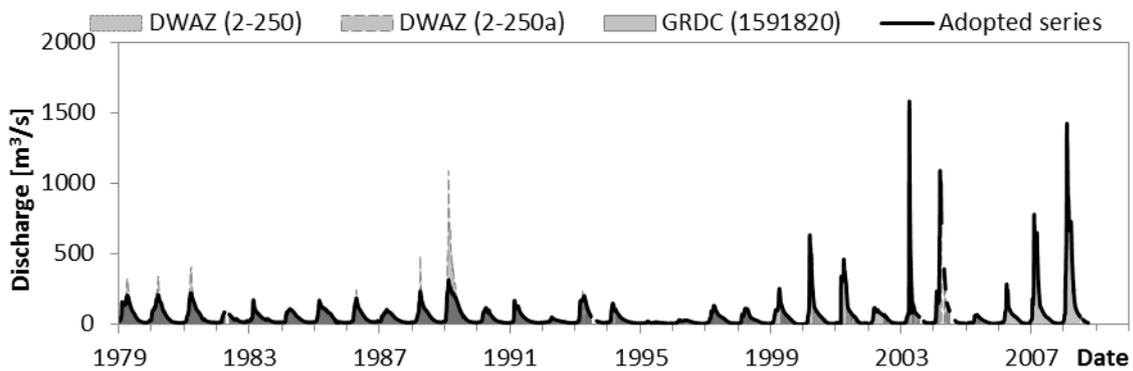
Figure AI.32. Discharge data for Kalabo.
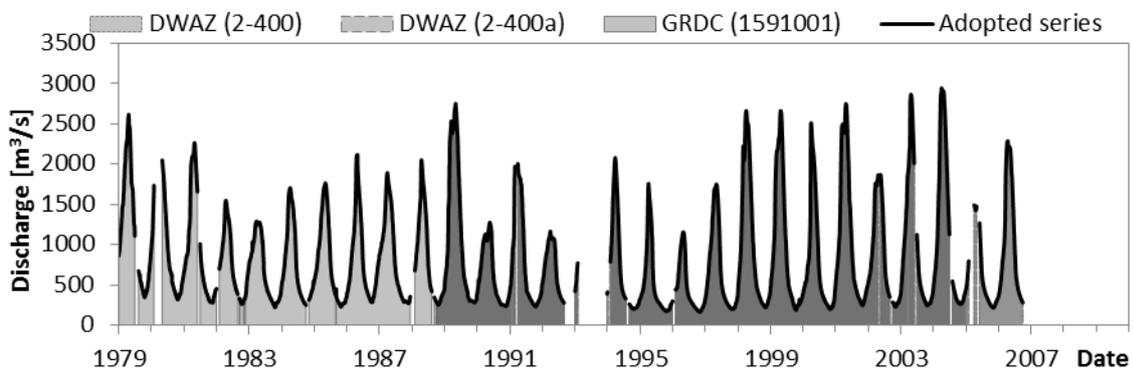
Figure AI.33. Discharge data for Senanga.
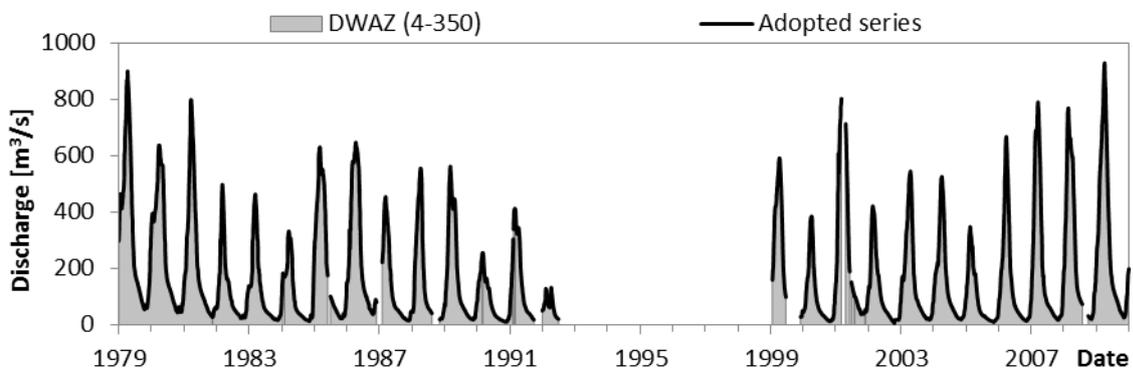
Figure AI.34. Discharge data for Chilenga.
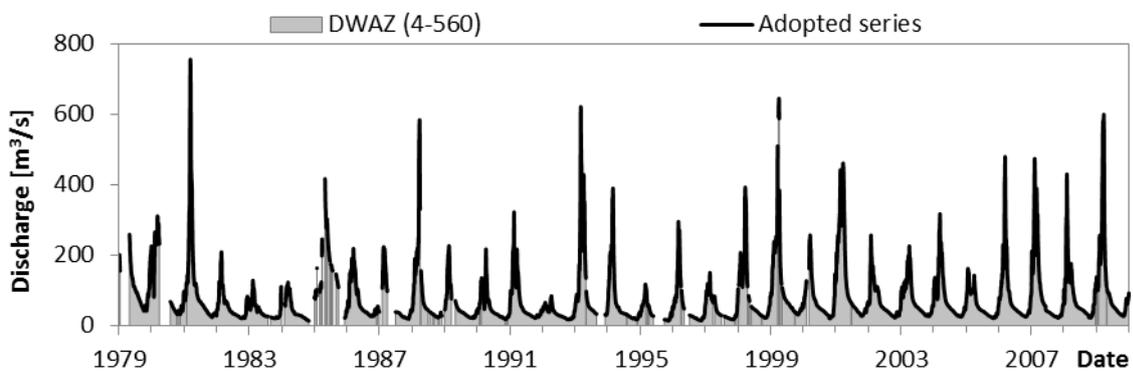
Figure AI.35. Discharge data for Chifumpa.
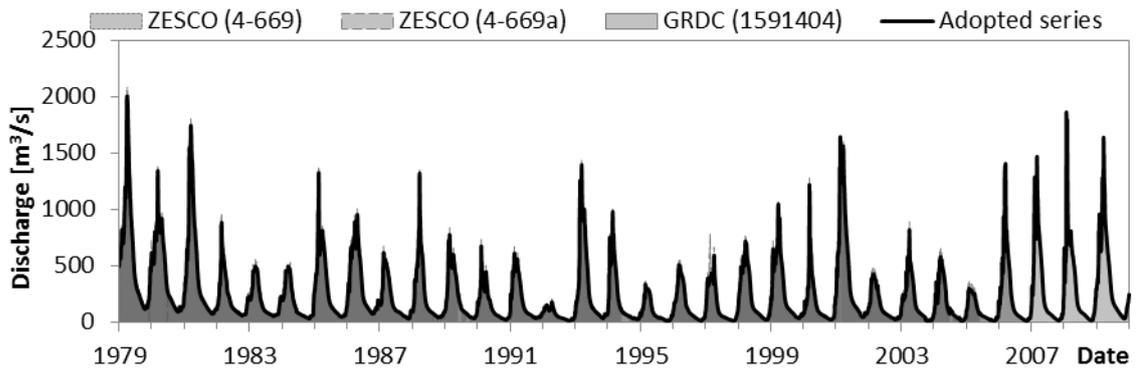
Figure AI.36. Discharge data for Kafue Hook Bridge.
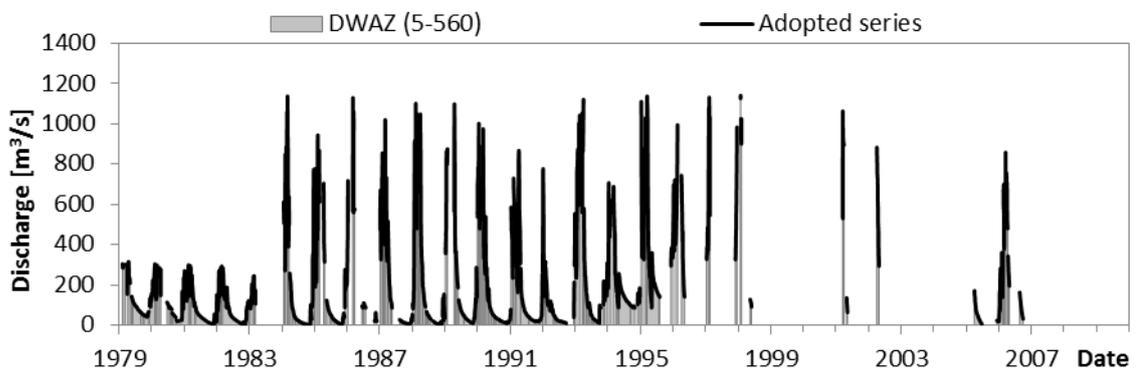
Figure AI.37. Discharge data for Mfuwe.
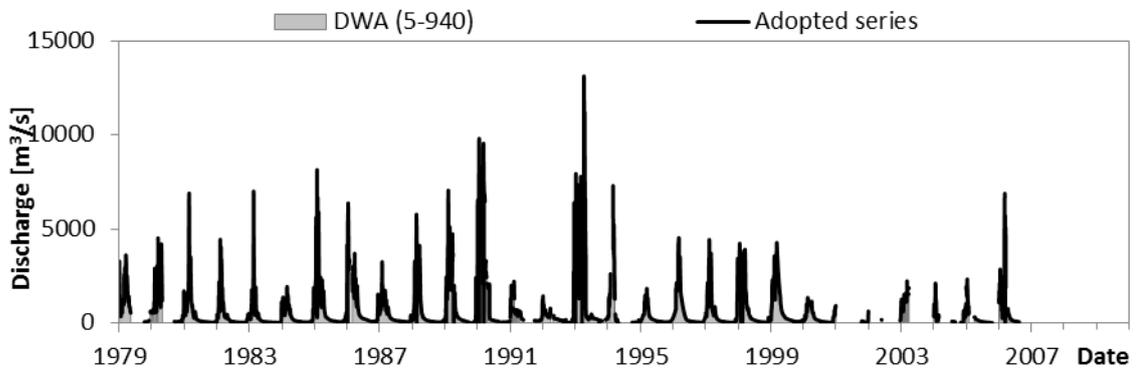
Figure AI.38. Discharge data for Great East Rd. Bridge.
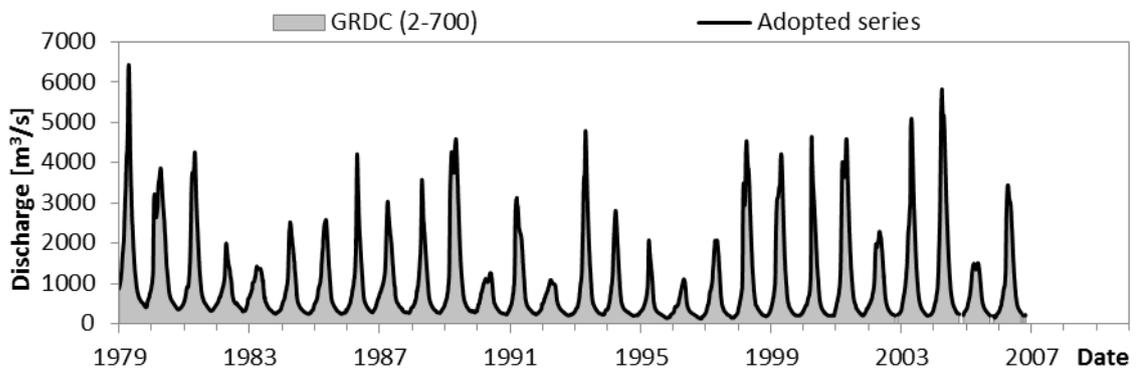
Figure AI.39. Discharge data for Sesheke.
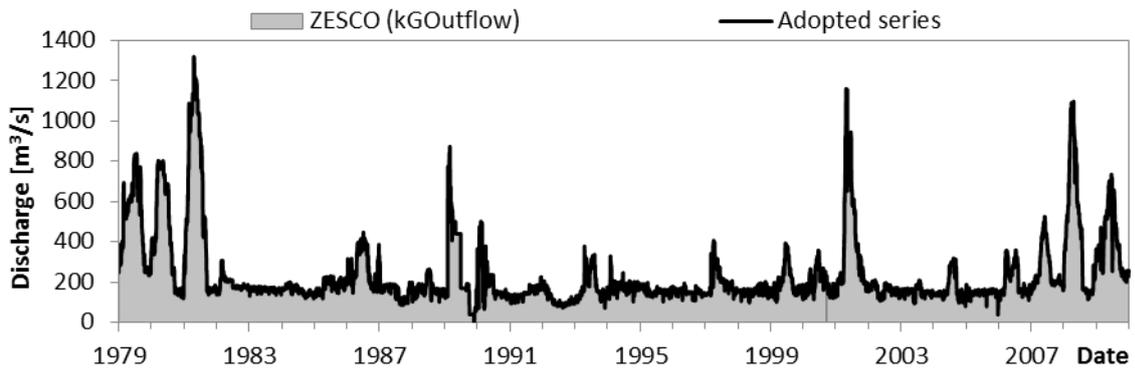
Figure AI.40. Discharge data for Kafue Flats (Kafue Gorge dam outflow).
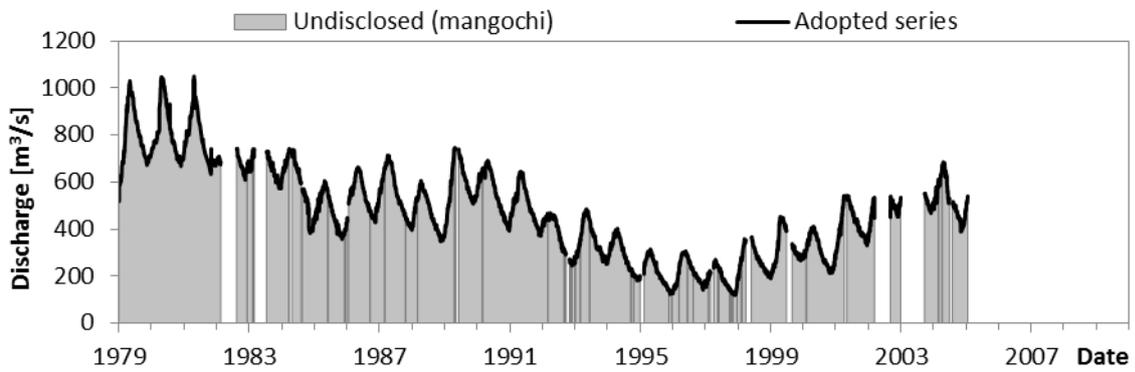
Figure AI.41. Discharge data for Mangochi.
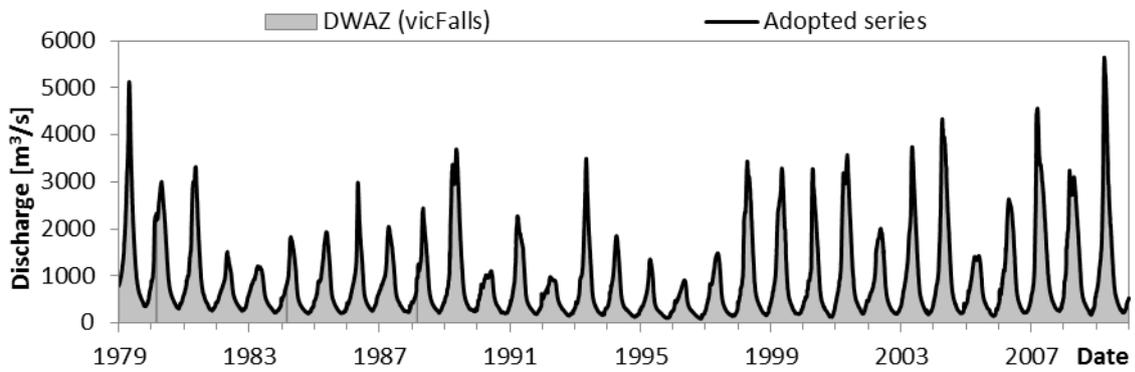
Figure AI.42. Discharge data for Victoria Falls.
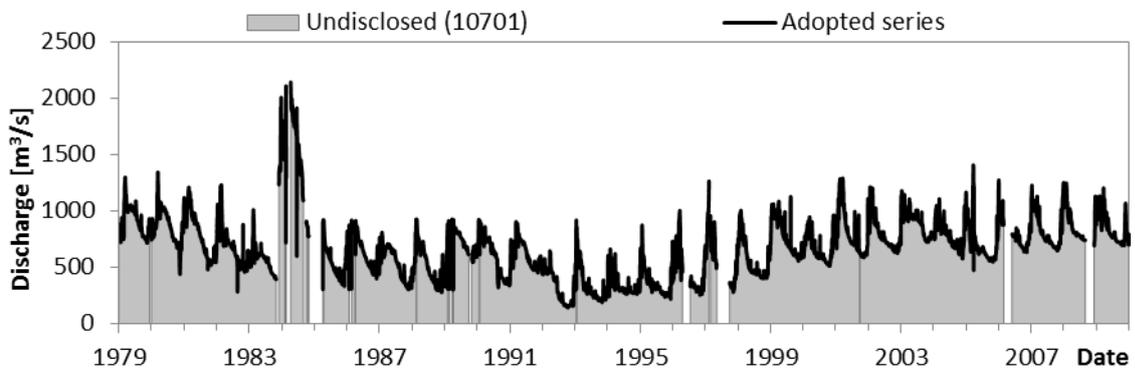
Figure AI.43. Discharge data for Chiromo.
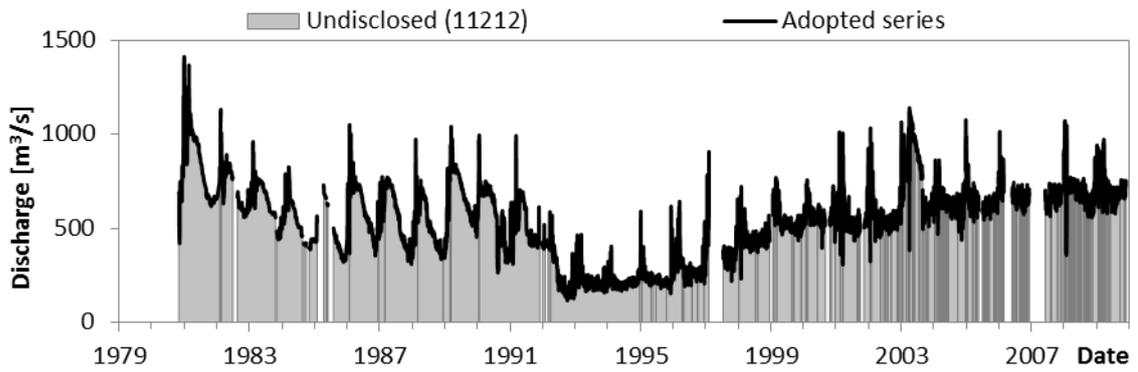
Figure AI.44. Discharge data for Chikwawa.
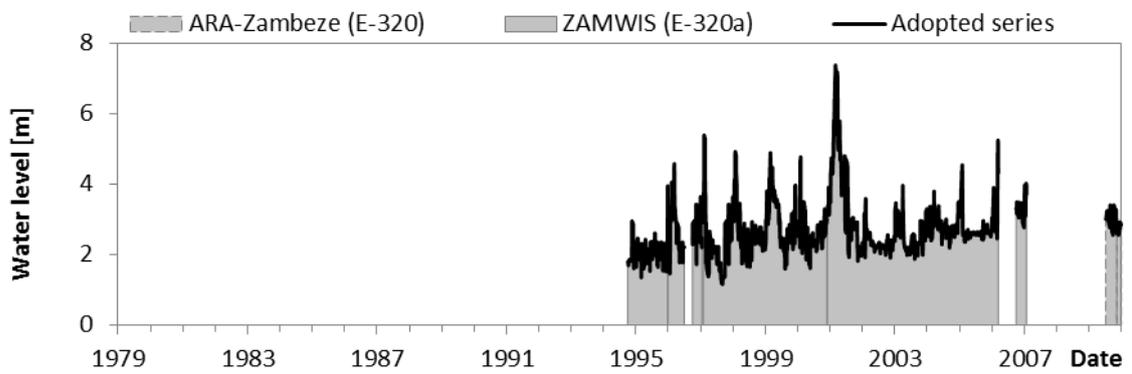
Figure AI.45. Water level data for Tete.
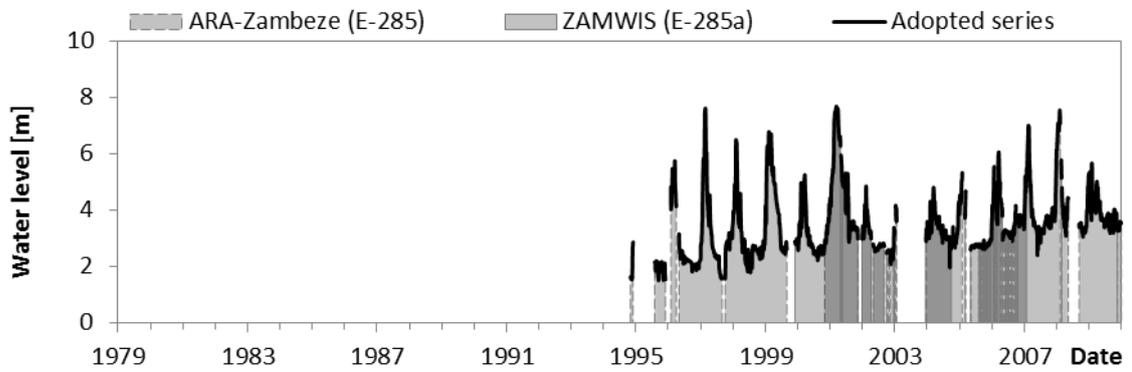
Figure AI.46. Water level data for Marromeu (Sena Sugar).
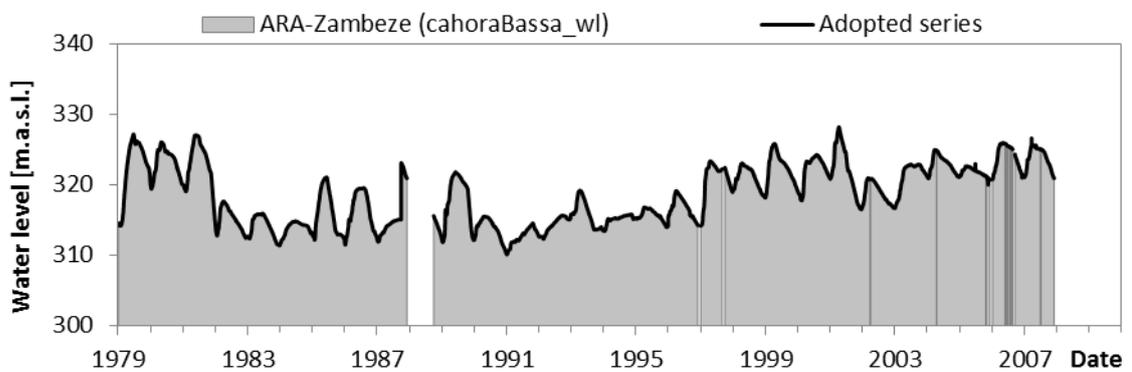
Figure AI.47. Stored volume data for the Cahora Bassa reservoir.
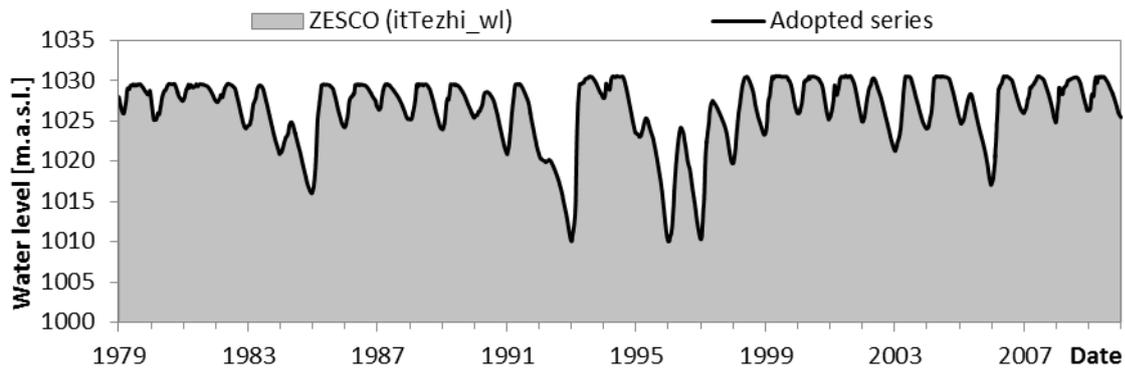
Figure AI.48. Stored volume data for the Itezhi-Tezhi reservoir.
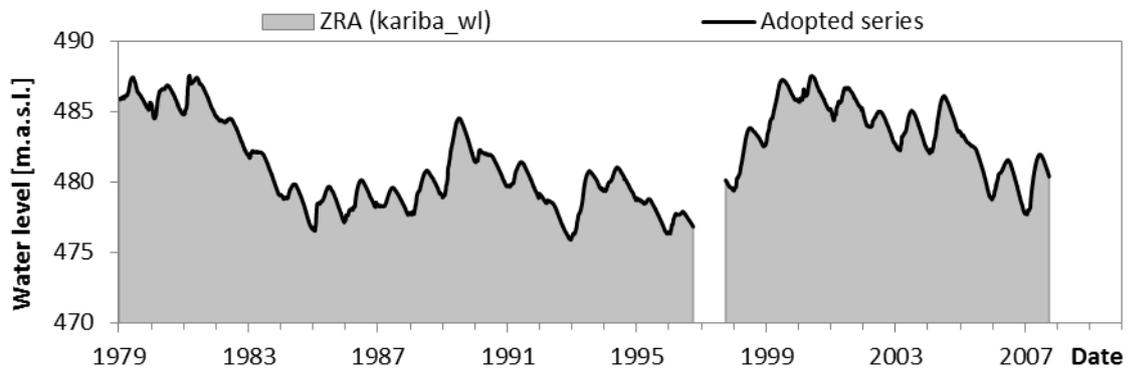
Figure AI.49. Stored volume data for the Kariba reservoir.
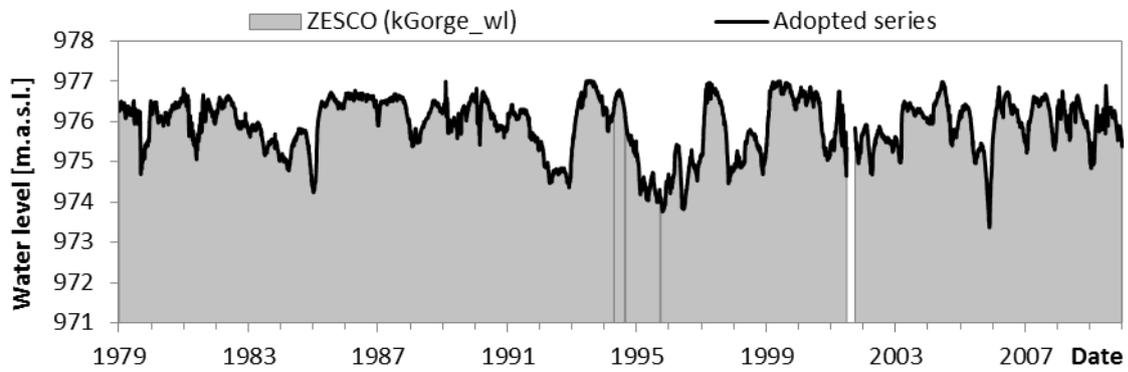
Figure AI.50. Stored volume data for the Kafue Gorge reservoir.
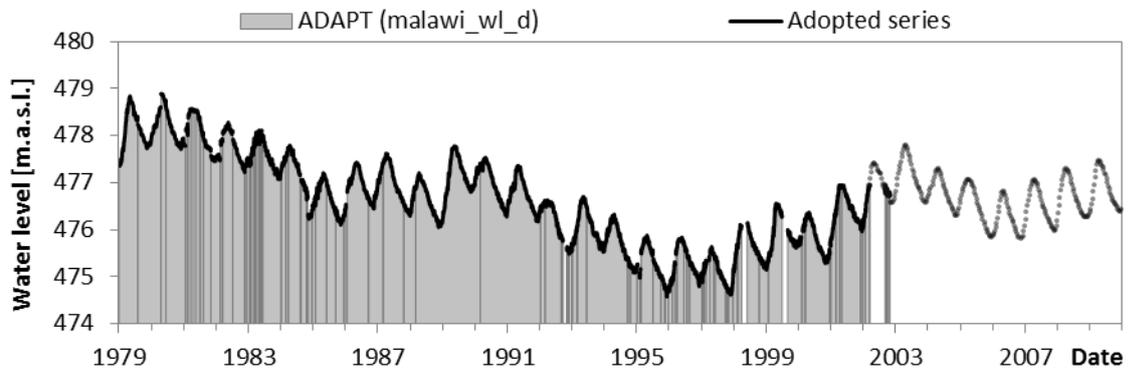
Figure AI.51. Stored volume data for the Malawi Lake.
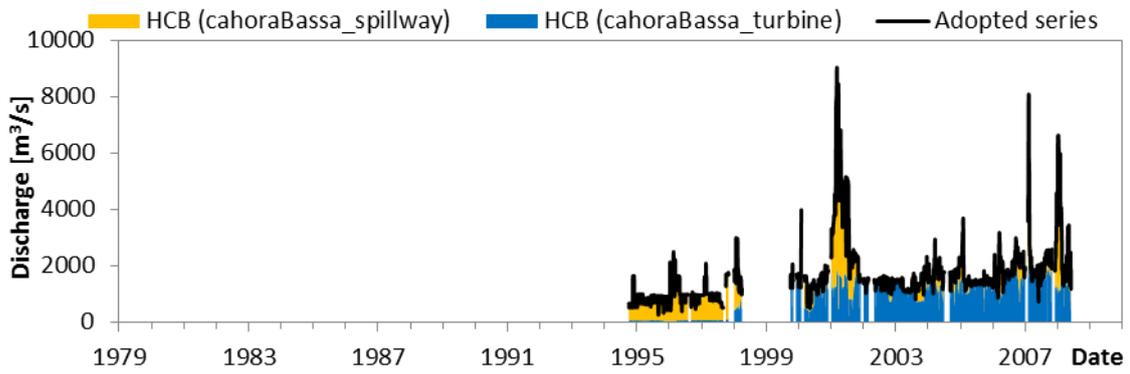
Figure AI.52. Outflow data for the Cahora Bassa dam.
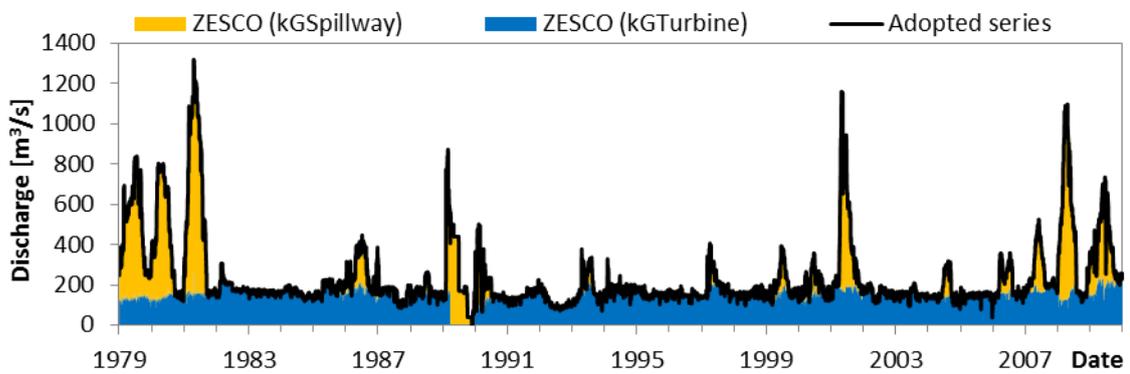
Figure AI.53. Outflow data for the Kafue Gorge dam.
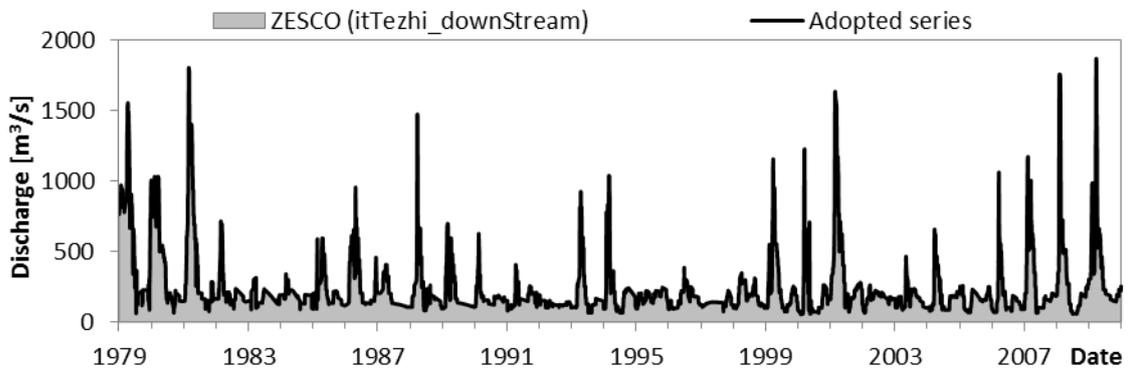
Figure AI.54. Outflow data for the Itezhi-Tezhi dam.
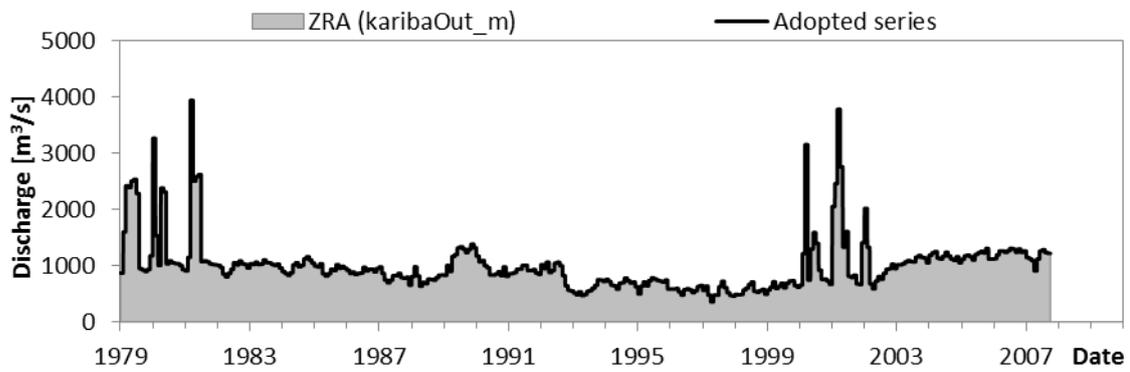
Figure AI.55. Outflow data for the Kariba dam (monthly).

APPENDIX I.D. TOPOLOGY OF THE PARSIMONIOUS SWAT MODEL
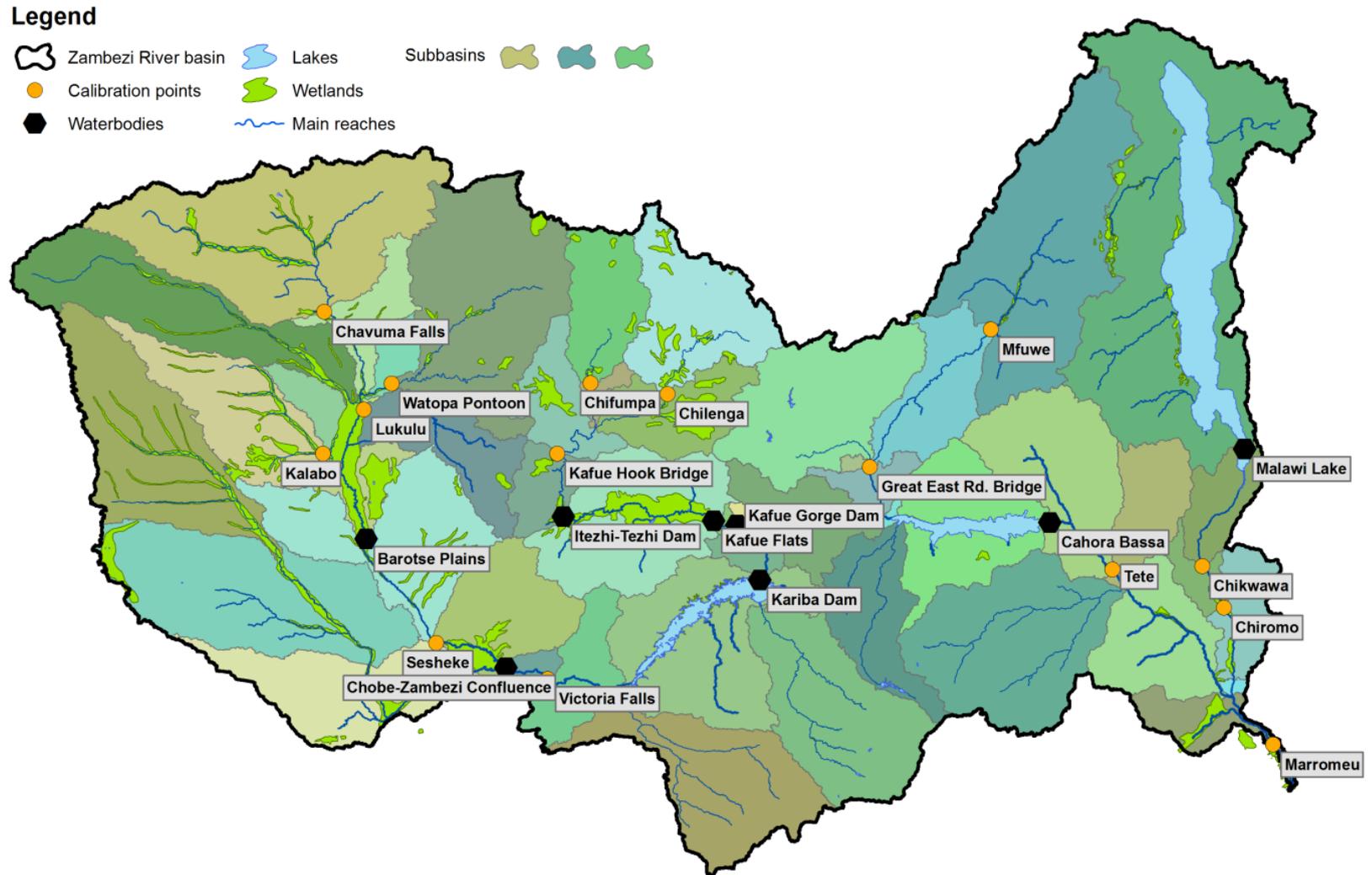
Figure AI.56. Topology of the parsimonious SWAT model.

APPENDIX I.E. TOPOLOGY OF THE DETAILED SWAT MODEL
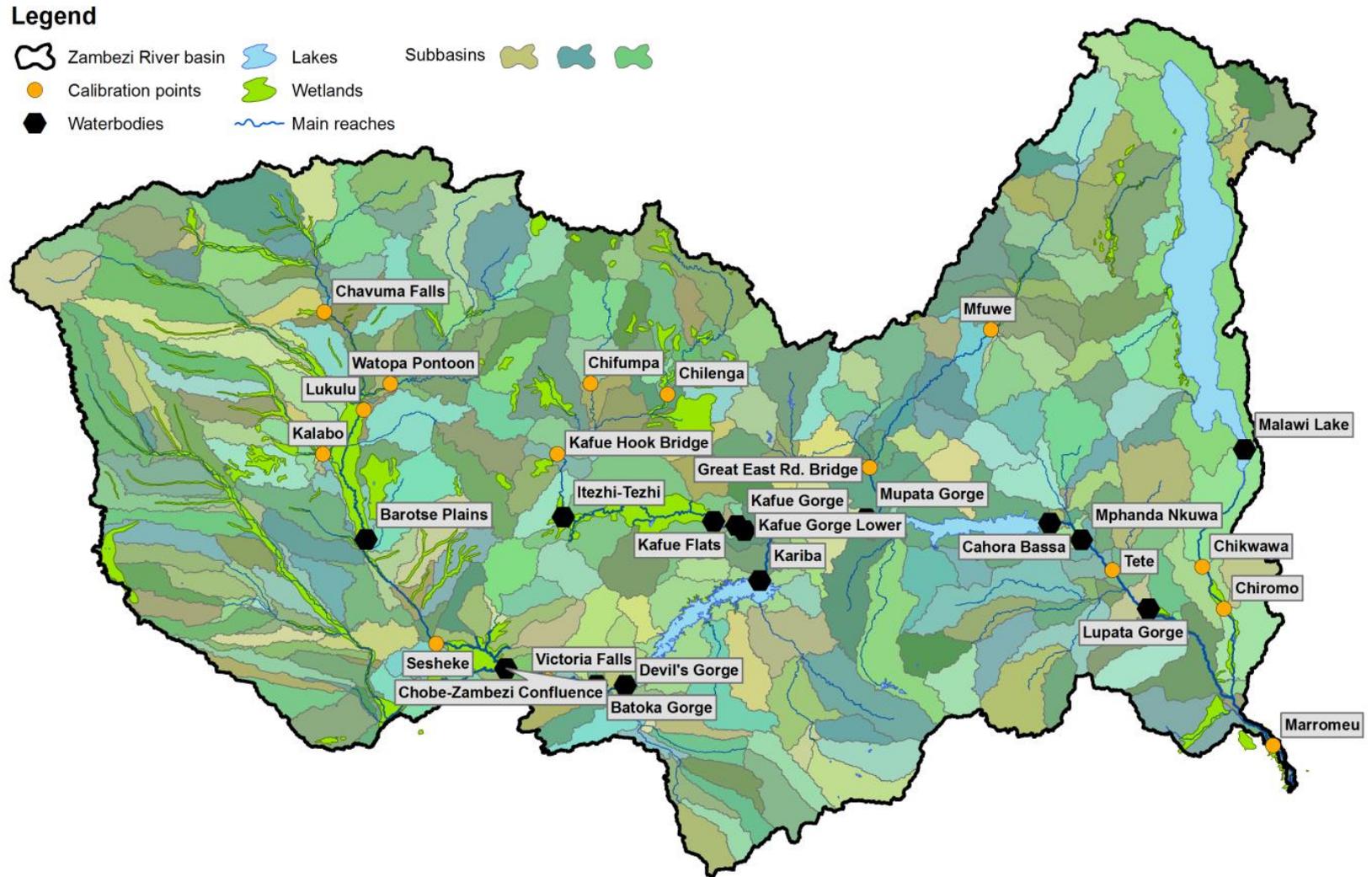
Figure AI.57. Topology of the detailed SWAT model.

APPENDIX I.F. SWAT PARAMETERS CONSIDERED FOR OPTIMIZATION

The appendix contains a list of the parameters considered for optimization. CH_K2, EPCO and GWQMN were only employed in the cascading calibration. Due to evidences of their little impact on the results, they were not used in the other calibration set-ups.

ALPHA, BETA, K, and H0 – uncontrolled reservoir parameters – were only calibrated along with the remaining parameters in the cascading calibration. For the alternative calibrations, they were fixed based on the relevant available observations. Results from the procedure are shown in Appendix I.G.

Table AI.1. Synthesis of the SWAT parameters selected for optimization. Indication of references and optimization assumptions.

Code	Designation	Reference	Overall lower bound	Overall upper bound	Scale	Log
ALPHA_BF	Baseflow alpha factor	α^{gw} , eq. (3.44) , page 33	0	1	absolute	no
Ave	Available water capacity of the soil layer	SW_i^{exc} , eq. (3.36) , page 32	0.5	2	**relative**	no
CANMX	Maximum canopy storage	See Section 7.6	1	60	absolute	no
CN2	Initial SCS runoff curve number for moisture condition II	CN_2 , eq. (3.4) , page 25	0.6	1.4	**relative**	no
Depth	Depth from soil surface to bottom of layer	$z_i^{upper} - z_i^{lower}$, see page 31	0.5	2	**relative**	no
ESCO	Soil evaporation compensation factor	$esco$, eq. (3.28), page 30	0.001	1	absolute	**yes**
EVRSV	Lake evaporation coefficient	η , eq. (3.68) , page 39	0.4	1.6	absolute	no
GW_DELAY	Groundwater delay time	δ_{gw} , eq. (3.40) , page 32	0.1	1000	absolute	**yes**
GW_REVAP	Groundwater "revap" coefficient	β^{rev} , eq. (3.42) , page 33	0.02	0.2	absolute	no
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur	$W^{sh,revap,thr}$, eq. (3.43) , page 33	1	400	absolute	**yes**
SURLAG	Surface runoff lag coefficient	$surlag$, eq. (3.13), page 27	0.01	10	absolute	no
CH_K2	Effective hydraulic conductivity in main channel alluvium	K^{rch} , eq. (3.59) , page 37	0.1	50	absolute	**yes**
EPCO	Plant uptake compensation factor	Not described in the manuscript. Refer to Neitsch et al. (2011)	0.001	1	absolute	**yes**
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur	$W_t^{sh,thr}$, eq. (3.44) , page 33	1	400	absolute	**yes**
ALPHA	Natural reservoir alpha coefficient	α^{rsv} , eq. (3.72) , page 40	10	10000	absolute	**yes**
BETA	Natural reservoir beta coefficient	β^{rsv} , eq. (3.72) , page 40	0.1	5	absolute	no
K	Natural reservoir k coefficient	k^{rsv} , eq. (3.71) , page 39	1	10000	absolute	**yes**
H0	Natural reservoir h0 coefficient	$d^{rsv,thr}$, eq. (3.71) , page 39	0	20	absolute	no

APPENDIX I.G. CALIBRATION OF UNCONTROLLED OUTFLOW PARAMETERS
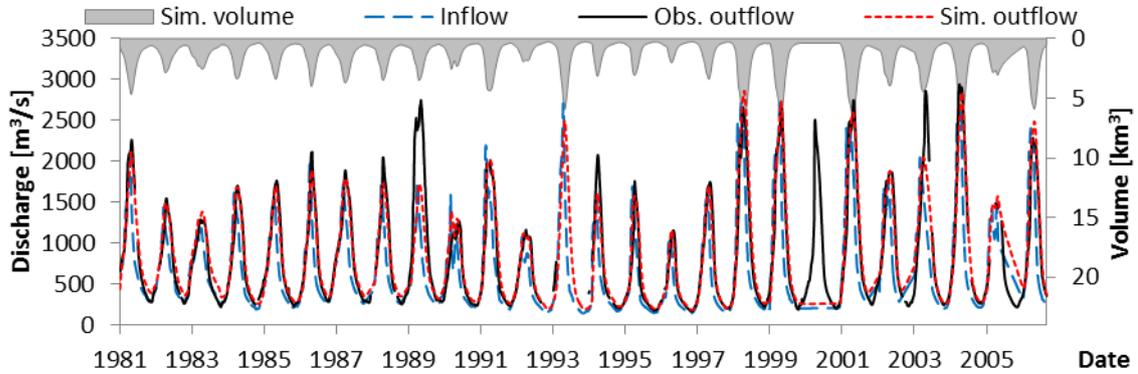
Figure AI.58. Parameterizing the Barotse Plains wetland outflows.
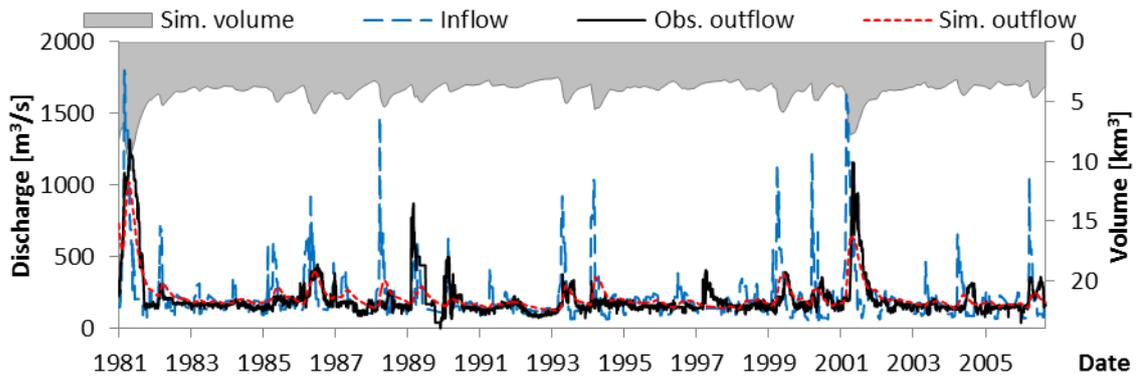
Figure AI.59. Parameterizing the Kafue Flats wetland outflows – convergence to an unrealistic set of parameters which leads to excess storage.
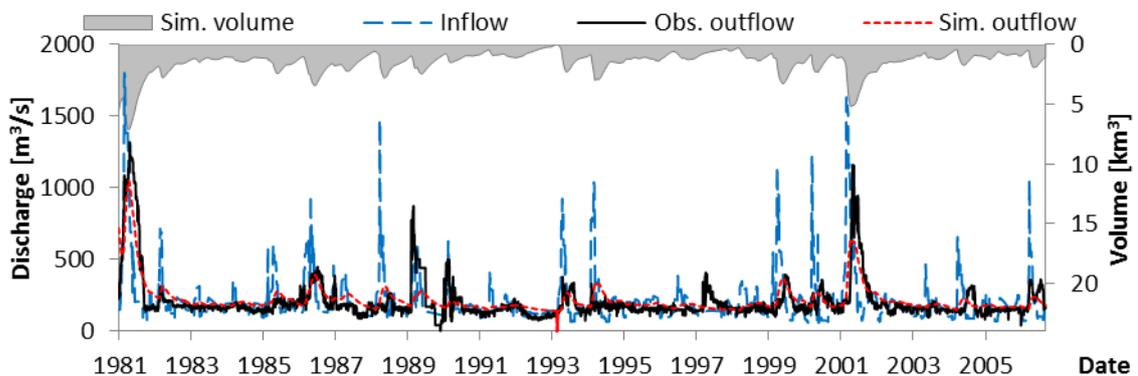
Figure AI.60. Parameterizing the Kafue Flats wetland outflows.
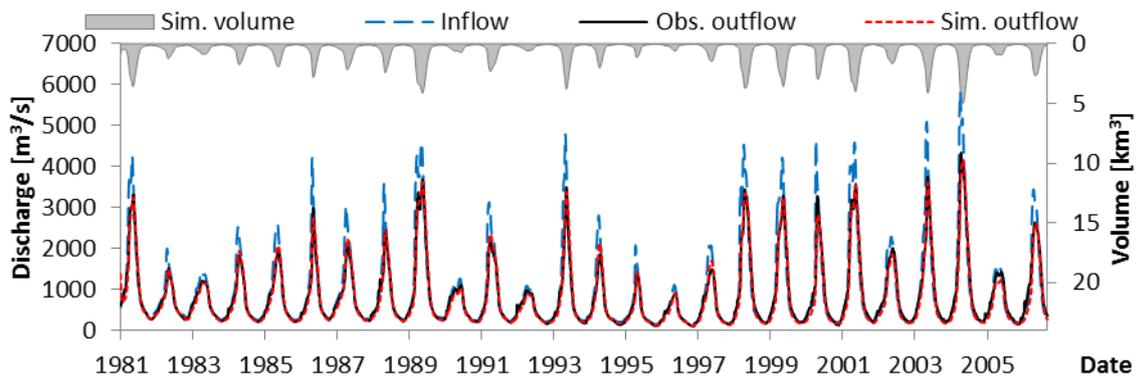
Figure AI.61. Parameterizing the Chobe-Zambezi Confluence wetland outflows.
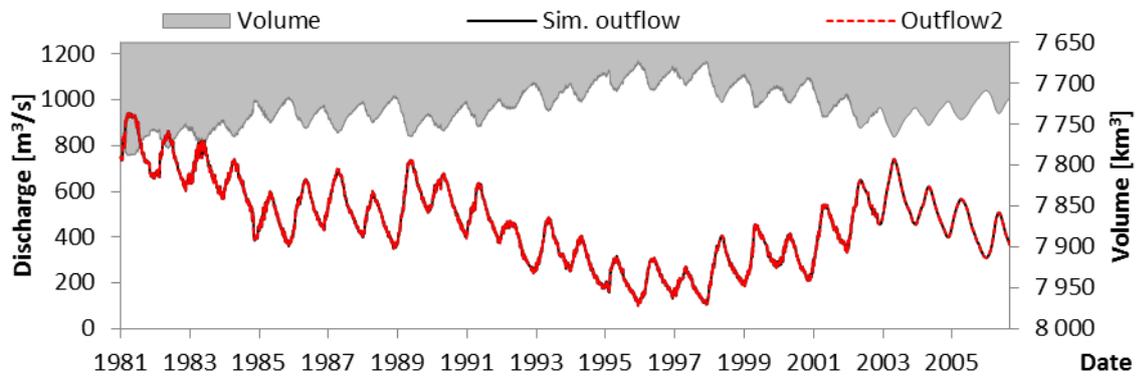
Figure AI.62. Parameterizing the Malawi Lake outflows.

APPENDIX I.H. RESULTS FOR THE GLOBAL CALIBRATION OF THE PARSIMONIOUS MODEL

Full ensemble results

The full ensemble results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Results from the 10% worst performing parameter sets regarding each calibration objective are not included.

Table AI.2. Full minimum ensemble results for the global calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	302.1	492.7	0.067	0.109	-1.23	1.14
Watopa Pontoon	134.3	173.1	0.152	0.196	-1.23	0.02
Chiromo	185.3	270.0	0.093	0.135	-0.45	0.63
Chikwawa	114.1	145.5	0.112	0.142	0.03	0.77
Sesheke	374.2	567.3	0.075	0.114	0.05	0.66
Senanga	284.5	382.0	0.105	0.142	-0.29	0.75
Lukulu	314.6	438.1	0.121	0.168	-2.08	0.91
Kafue Hook Bridge	104.0	176.6	0.067	0.114	-0.74	0.02
Kalabo	43.0	76.7	0.027	0.049	-0.42	0.01
Chilenga	84.0	137.9	0.106	0.173	-0.86	0.01
Chifumpa	46.8	72.7	0.074	0.115	-0.86	0.03
Mfuwe	166.4	266.3	0.146	0.234	-0.65	0.00
Great East Rd. Bridge	578.2	1281.3	0.044	0.098	-0.27	0.02
Kafue Gorge	77.4	116.0	0.067	0.100	-2.65	0.08
Mangochi	90.2	106.8	0.144	0.170	0.45	0.87
Victoria Falls	321.6	445.3	0.088	0.122	0.10	0.79
Marromeu (level)	0.630	0.783	0.102	0.127	0.24	1.00
Tete (level)	0.525	0.701	0.085	0.113	-2.29	0.87
Itezhi-Tezhi	3.90E+09	4.13E+09	0.776	0.821	-16.4	0.0
Kafue Gorge	3.96E+08	4.75E+08	0.356	0.426	-376.7	0.0
Kariba	1.83E+10	2.19E+10	0.318	0.380	-140.7	0.7
Cahora Bassa	9.70E+09	1.08E+10	0.371	0.413	-16.0	0.7
Malawi Lake	1.28E+10	1.53E+10	0.118	0.142	0.5	1.0

Table AI.3. Full minimum ensemble results for the global calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	316.3	504.6	0.079	0.125	-1.33	0.98
Watopa Pontoon	180.5	229.7	0.162	0.206	-1.29	0.02
Chiromo	291.1	314.8	0.312	0.338	-5.94	0.53
Chikwawa	181.6	202.1	0.252	0.281	-11.76	0.63
Sesheke	417.0	768.8	0.074	0.136	-0.03	0.52
Senanga	236.6	305.2	0.087	0.112	0.11	0.62
Lukulu	331.1	454.5	0.133	0.183	-0.57	0.57
Kafue Hook Bridge	96.1	168.9	0.052	0.091	-0.55	0.04
Kalabo	93.3	206.1	0.066	0.145	-0.26	0.01
Chilenga	95.3	162.5	0.104	0.177	-0.76	0.01
Chifumpa	48.1	72.4	0.083	0.125	-0.89	0.03
Mfuwe	87.3	129.1	0.103	0.152	-0.75	0.01
Great East Rd. Bridge	383.9	819.8	0.056	0.119	-0.24	0.04
Kafue Gorge	86.9	128.2	0.082	0.121	-1.59	0.15
Mangochi	53.0	65.6	0.179	0.221	0.09	0.88
Victoria Falls	345.6	523.4	0.063	0.095	0.07	0.61
Marromeu (level)	0.592	0.831	0.107	0.151	-0.97	0.98
Tete (level)	0.377	0.632	0.125	0.209	-16.88	0.92
Itezhi-Tezhi	2.77E+09	3.22E+09	0.749	0.872	-38.42	0.05
Kafue Gorge	3.46E+08	3.93E+08	0.308	0.350	-1862.04	0.01
Kariba	4.52E+09	5.55E+09	0.108	0.133	-10.35	0.86
Cahora Bassa	2.76E+09	3.49E+09	0.224	0.283	-73.85	0.60
Malawi Lake	6.02E+09	7.24E+09	0.125	0.151	0.40	1.00

Table AI.4. Full median ensemble results for the global calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	426.4	588.3	0.094	0.130	0.23	1.31
Watopa Pontoon	161.0	214.7	0.182	0.243	-0.76	0.14
Chiromo	222.2	313.2	0.111	0.157	-0.25	0.68
Chikwawa	133.7	170.1	0.131	0.166	0.22	0.84
Sesheke	449.7	689.3	0.091	0.139	0.50	0.82
Senanga	365.0	479.4	0.135	0.178	0.42	0.91
Lukulu	391.5	500.4	0.151	0.192	0.16	1.09
Kafue Hook Bridge	156.3	234.7	0.101	0.151	0.02	0.30
Kalabo	46.4	84.1	0.029	0.053	-0.19	0.16
Chilenga	125.7	187.8	0.158	0.236	-0.30	0.20
Chifumpa	60.8	88.8	0.096	0.141	-0.40	0.25
Mfuwe	192.9	309.6	0.169	0.272	-0.56	0.05
Great East Rd. Bridge	643.8	1384.7	0.049	0.106	-0.22	0.06
Kafue Gorge	114.5	148.2	0.099	0.128	-0.85	0.43
Mangochi	92.6	110.6	0.148	0.176	0.51	0.95
Victoria Falls	391.7	536.7	0.107	0.147	0.48	0.98
Marromeu (level)	0.862	0.999	0.140	0.162	0.44	1.03
Tete (level)	0.732	0.944	0.118	0.152	-0.22	0.96
Itezhi-Tezhi	4.67E+09	4.84E+09	0.930	0.963	-14.9	0.1
Kafue Gorge	4.60E+08	5.35E+08	0.413	0.481	-2.8	0.1
Kariba	3.06E+10	3.58E+10	0.531	0.621	-5.0	0.9
Cahora Bassa	1.25E+10	1.40E+10	0.476	0.536	-5.0	0.9
Malawi Lake	1.31E+10	1.57E+10	0.122	0.145	0.6	1.0

Table AI.5. Full median ensemble results for the global calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	422.1	605.5	0.105	0.150	0.35	1.14
Watopa Pontoon	218.4	292.8	0.196	0.262	-0.60	0.20
Chiromo	345.3	369.3	0.370	0.396	-4.92	0.57
Chikwawa	225.6	242.1	0.313	0.336	-9.68	0.67
Sesheke	615.7	1052.9	0.109	0.186	0.33	0.61
Senanga	396.6	542.1	0.146	0.199	0.53	0.72
Lukulu	484.9	688.5	0.195	0.277	0.19	0.74
Kafue Hook Bridge	147.9	241.8	0.080	0.131	0.40	0.49
Kalabo	101.1	225.5	0.071	0.159	-0.13	0.14
Chilenga	143.9	222.6	0.157	0.242	-0.04	0.27
Chifumpa	65.2	93.7	0.113	0.162	-0.13	0.32
Mfuwe	140.8	213.3	0.165	0.251	-0.18	0.23
Great East Rd. Bridge	453.8	983.6	0.066	0.143	-0.11	0.14
Kafue Gorge	125.8	183.3	0.119	0.173	0.08	0.63
Mangochi	60.7	73.5	0.205	0.248	0.16	0.89
Victoria Falls	515.8	802.1	0.094	0.146	0.54	0.75
Marromeu (level)	0.887	1.060	0.161	0.192	-0.44	1.01
Tete (level)	0.645	0.971	0.213	0.321	-5.92	0.97
Itezhi-Tezhi	4.53E+09	4.70E+09	1.228	1.275	-29.04	0.18
Kafue Gorge	4.53E+08	5.03E+08	0.403	0.447	-4.14	0.37
Kariba	1.25E+10	1.36E+10	0.299	0.325	-0.60	0.95
Cahora Bassa	8.24E+09	1.05E+10	0.667	0.850	-11.75	0.85
Malawi Lake	6.45E+09	7.83E+09	0.134	0.163	0.59	1.00

Table AI.6. Full maximum ensemble results for the global calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	655.9	1002.3	0.145	0.221	0.46	2.28
Watopa Pontoon	182.9	241.9	0.207	0.273	-0.14	0.29
Chiromo	245.2	337.4	0.123	0.169	0.07	0.79
Chikwawa	148.8	189.9	0.145	0.186	0.43	0.95
Sesheke	627.4	944.7	0.126	0.190	0.66	1.21
Senanga	507.5	714.3	0.188	0.265	0.63	1.41
Lukulu	650.1	959.1	0.250	0.369	0.36	1.78
Kafue Hook Bridge	211.4	313.5	0.136	0.202	0.45	0.61
Kalabo	51.3	92.0	0.032	0.058	0.01	0.36
Chilenga	154.7	224.7	0.194	0.282	0.30	0.51
Chifumpa	71.1	102.5	0.113	0.162	0.06	0.52
Mfuwe	201.3	319.0	0.177	0.280	-0.15	0.18
Great East Rd. Bridge	672.1	1412.6	0.051	0.108	-0.04	0.17
Kafue Gorge	176.5	208.1	0.152	0.179	-0.13	0.87
Mangochi	97.2	117.0	0.155	0.186	0.54	1.08
Victoria Falls	510.8	703.6	0.140	0.193	0.64	1.38
Marromeu (level)	1.007	1.170	0.163	0.190	0.66	1.06
Tete (level)	1.145	1.549	0.185	0.250	0.33	1.00
Itezhi-Tezhi	4.89E+09	5.06E+09	0.974	1.007	-10.6	0.2
Kafue Gorge	2.77E+09	5.33E+09	2.489	4.788	-2.0	5.7
Kariba	1.43E+11	1.73E+11	2.485	3.008	-1.3	2.0
Cahora Bassa	1.86E+10	2.37E+10	0.709	0.904	-2.5	1.2
Malawi Lake	1.37E+10	1.69E+10	0.127	0.156	0.6	1.0

Table AI.7. Full maximum ensemble results for the global calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	774.0	1143.7	0.192	0.284	0.55	2.26
Watopa Pontoon	266.8	350.1	0.239	0.313	0.01	0.35
Chiromo	376.6	399.8	0.404	0.429	-3.30	0.64
Chikwawa	252.0	264.5	0.350	0.368	-6.45	0.74
Sesheke	828.9	1304.6	0.146	0.230	0.64	1.16
Senanga	596.9	750.9	0.219	0.276	0.85	1.37
Lukulu	703.8	960.6	0.283	0.386	0.65	1.68
Kafue Hook Bridge	244.9	387.8	0.132	0.210	0.71	0.85
Kalabo	112.3	238.4	0.079	0.168	0.05	0.28
Chilenga	193.3	289.5	0.210	0.315	0.45	0.57
Chifumpa	85.2	121.1	0.148	0.210	0.33	0.58
Mfuwe	175.2	259.9	0.206	0.305	0.57	0.54
Great East Rd. Bridge	500.9	1040.8	0.073	0.151	0.23	0.39
Kafue Gorge	234.9	308.1	0.222	0.291	0.55	1.20
Mangochi	64.2	76.6	0.216	0.258	0.33	0.91
Victoria Falls	818.8	1135.9	0.149	0.206	0.80	1.05
Marromeu (level)	1.088	1.239	0.197	0.224	0.11	1.05
Tete (level)	1.020	1.561	0.337	0.516	-1.93	1.00
Itezhi-Tezhi	5.25E+09	5.39E+09	1.422	1.460	-13.04	0.50
Kafue Gorge	8.34E+09	9.57E+09	7.428	8.522	-2.14	17.51
Kariba	3.04E+10	3.62E+10	0.727	0.865	0.73	1.20
Cahora Bassa	2.02E+10	2.54E+10	1.633	2.059	-0.41	1.17
Malawi Lake	7.57E+09	9.44E+09	0.158	0.197	0.65	1.00

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Their interpretation is introduced in the Section 9.3.2 of the main text.
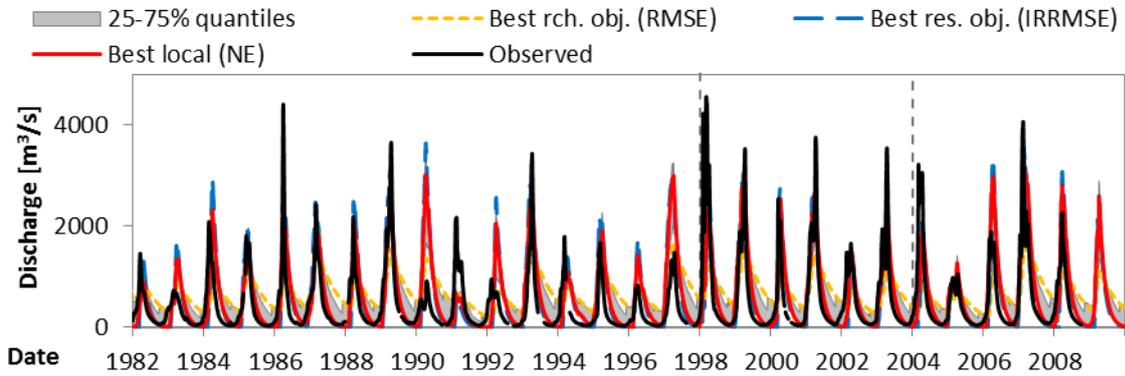
Figure AI.63. Synthesis hydrographs of the global calibration of the parsimonious model at Chavuma Falls.
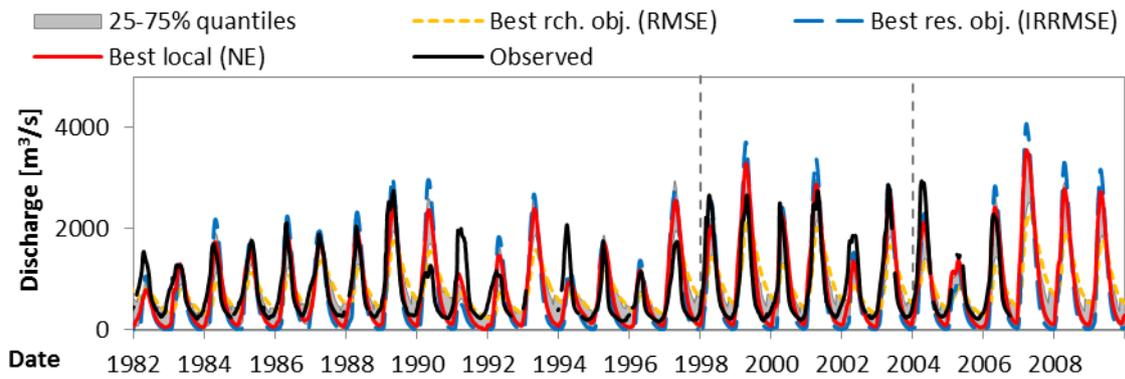
Figure AI.64. Synthesis hydrographs of the global calibration of the parsimonious model at Senanga.
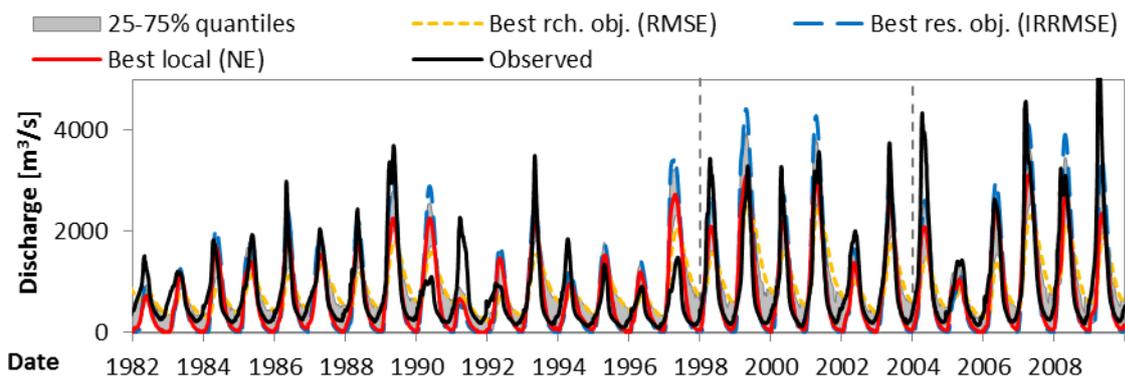
Figure AI.65. Synthesis hydrographs of the global calibration of the parsimonious model at Victoria Falls.
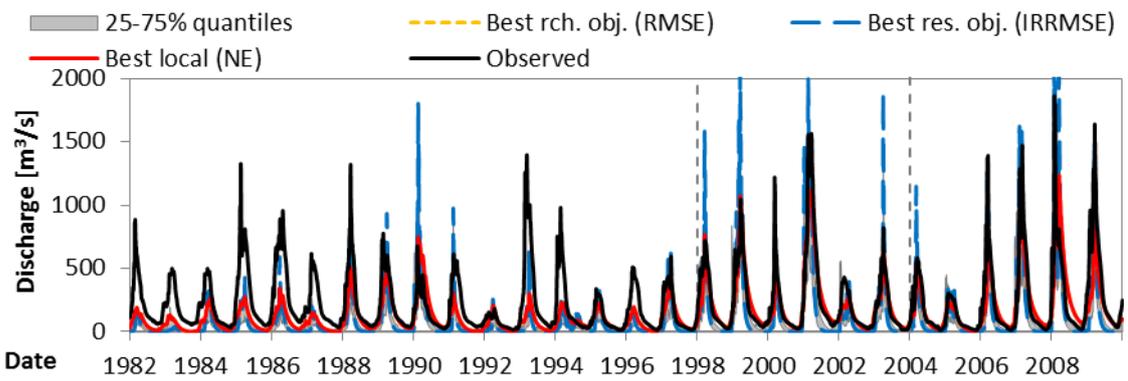
Figure AI.66. Synthesis hydrographs of the global calibration of the parsimonious model at Kafue Hook Bridge.
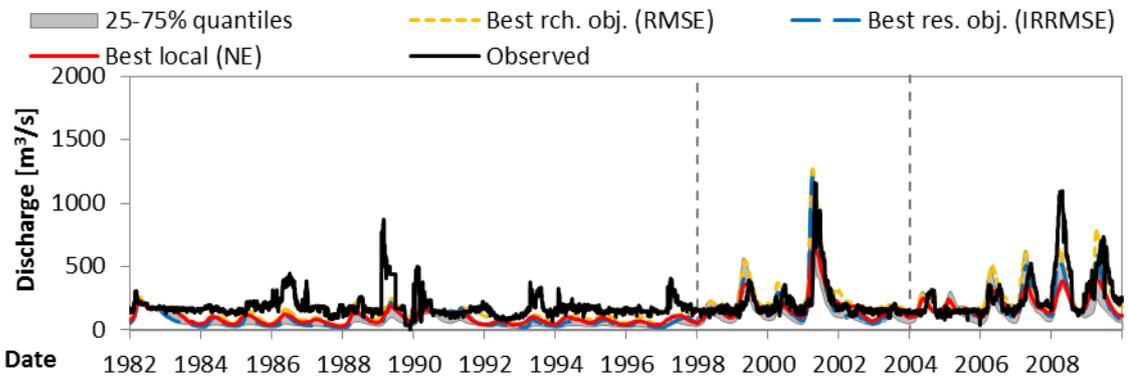
Figure AI.67. Synthesis hydrographs of the global calibration of the parsimonious model at Kafue Gorge.
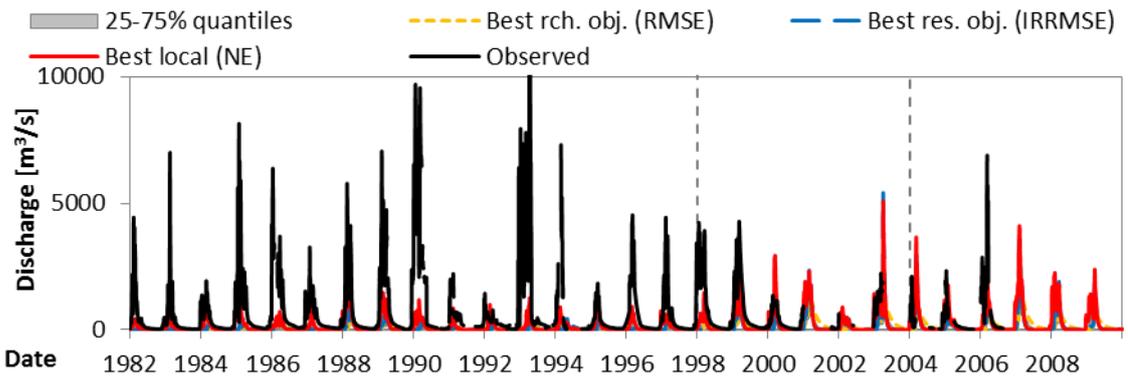
Figure AI.68. Synthesis hydrographs of the global calibration of the parsimonious model at Great East Road Bridge.
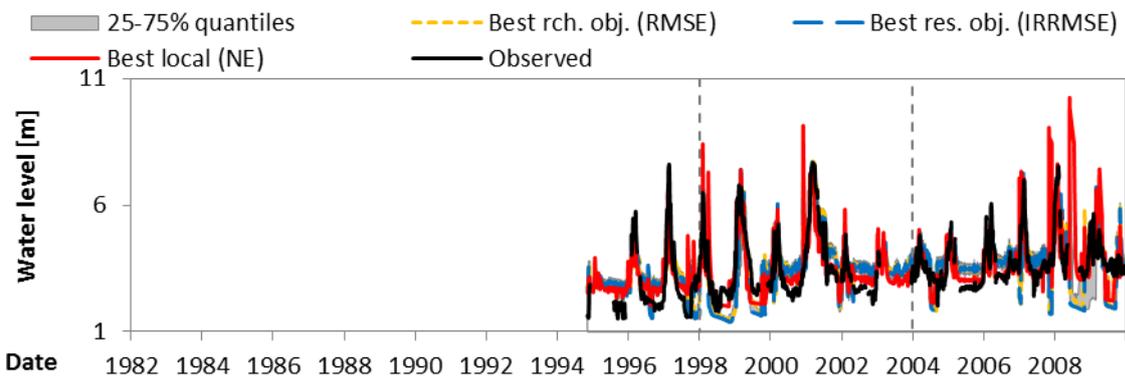
Figure AI.69. Synthesis water level series of the global calibration of the parsimonious model at Marromeu.
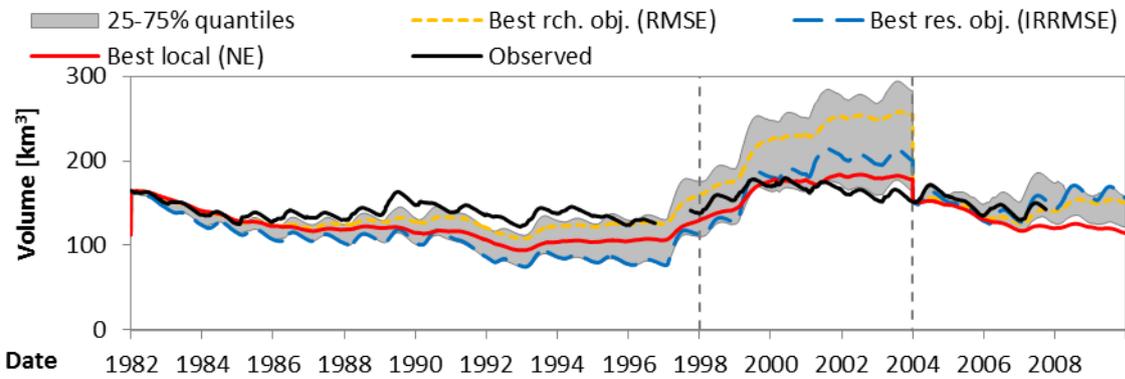
Figure AI.70. Synthesis volume series of the global calibration of the parsimonious model at the Kariba Dam.
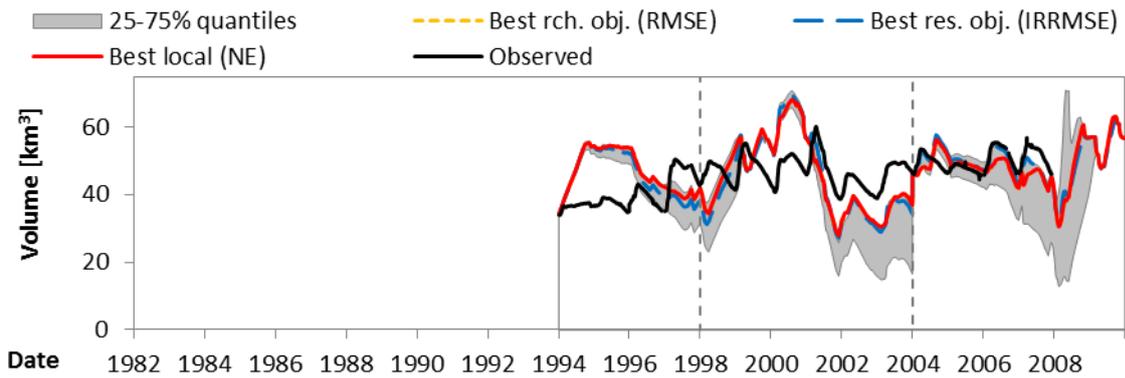
Figure AI.71. Synthesis volume series of the global calibration of the parsimonious model at the Cahora Bassa Dam.
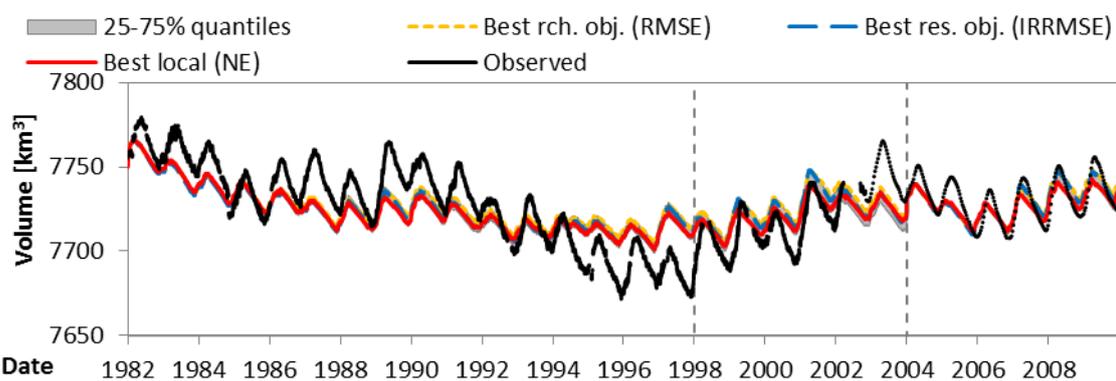
Figure AI.72. Synthesis volume series of the global calibration of the parsimonious model at the Malawi Lake.

Final parameter distribution
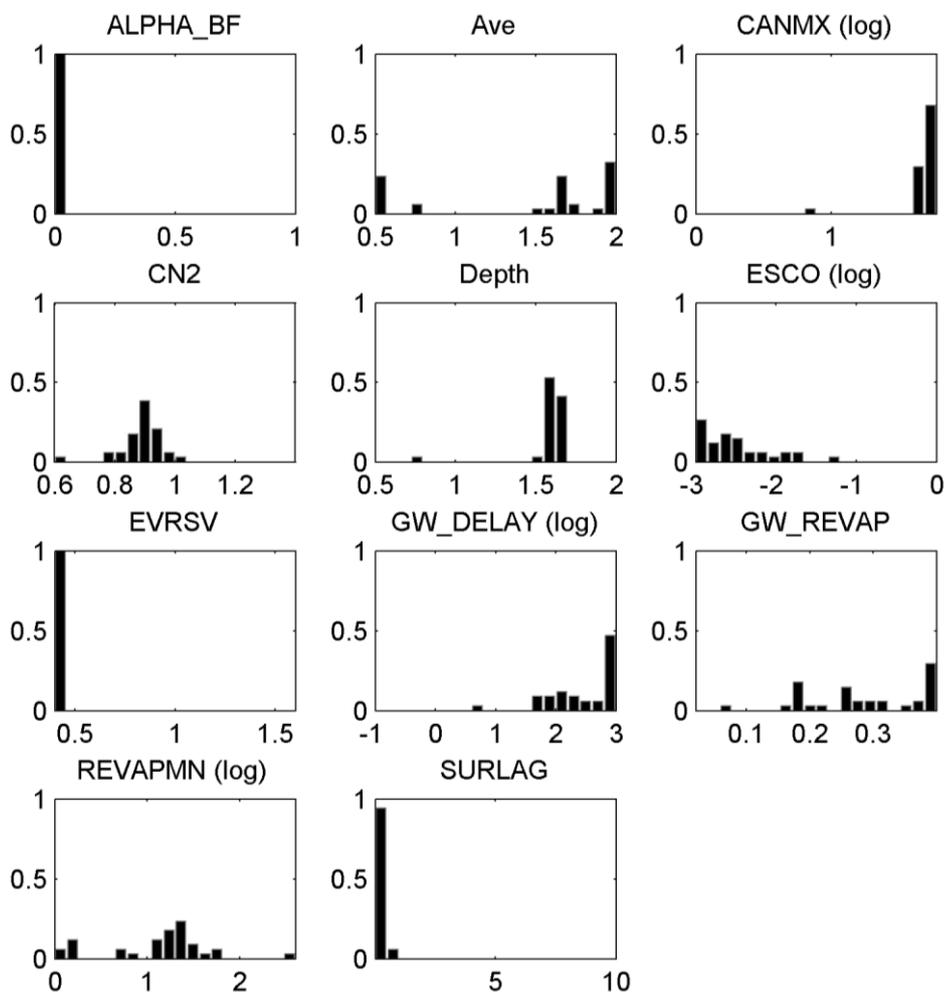
Figure AI.73. Parameter distribution among the ensemble for the last iteration of the global calibration of the parsimonious model. Histogram normalized according to ensemble size (40).

APPENDIX I.I. RESULTS FOR THE GLOBAL CALIBRATION OF THE DETAILED MODEL

Full ensemble results

The full ensemble results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Results from the 10% worst performing parameter sets regarding each calibration objective are not included.

Table AI.8. Full minimum ensemble results for the global calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	203.0	402.4	0.045	0.089	-0.03	0.59
Watopa Pontoon	134.9	178.0	0.152	0.201	-0.97	0.10
Chiromo	180.4	242.2	0.090	0.121	-0.16	0.72
Chikwawa	122.3	150.2	0.120	0.147	0.18	0.89
Sesheke	399.2	695.2	0.080	0.140	-0.27	0.36
Senanga	233.7	344.7	0.087	0.128	-0.45	0.41
Lukulu	198.5	288.1	0.076	0.111	-0.58	0.47
Kafue Hook Bridge	153.2	234.7	0.099	0.151	-0.57	0.11
Kalabo	39.9	74.2	0.025	0.047	-0.55	0.09
Chilenga	131.2	195.7	0.165	0.246	-0.77	0.05
Chifumpa	54.1	82.3	0.086	0.130	-0.72	0.15
Mfuwe	144.3	240.6	0.127	0.211	-2.53	0.90
Great East Rd. Bridge	397.8	959.1	0.030	0.073	0.02	0.52
Kafue Gorge	105.3	134.8	0.091	0.116	-1.62	0.26
Mangochi	68.7	86.8	0.110	0.138	0.39	0.98
Victoria Falls	304.5	482.0	0.083	0.132	-0.22	0.45
Marromeu (level)	0.738	0.934	0.120	0.151	0.33	1.03
Tete (level)	0.525	0.720	0.085	0.116	-0.32	0.94
Itezhi-Tezhi	4.54E+09	4.71E+09	0.905	0.938	-16.29	0.02
Kafue Gorge	4.01E+08	4.82E+08	0.360	0.433	-6.87	0.01
Kariba	1.61E+10	1.91E+10	0.279	0.331	-46.69	0.38
Cahora Bassa	8.84E+09	1.16E+10	0.337	0.444	-12.14	0.89
Malawi Lake	9.47E+09	1.19E+10	0.088	0.110	0.48	1.00

Table AI.9. Full minimum ensemble results for the global calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	229.3	416.6	0.057	0.104	-0.19	0.68
Watopa Pontoon	161.3	222.7	0.144	0.199	-0.81	0.18
Chiromo	242.0	267.3	0.260	0.287	-4.17	0.61
Chikwawa	141.8	164.0	0.197	0.228	-9.21	0.70
Sesheke	572.4	939.1	0.101	0.166	-0.26	0.39
Senanga	288.0	411.4	0.106	0.151	-0.27	0.47
Lukulu	308.4	407.2	0.124	0.164	-0.68	0.48
Kafue Hook Bridge	120.2	211.9	0.065	0.115	-0.24	0.26
Kalabo	77.8	169.2	0.055	0.119	-0.15	0.12
Chilenga	137.9	215.3	0.150	0.234	-0.56	0.12
Chifumpa	54.6	80.6	0.095	0.139	-0.54	0.18
Mfuwe	242.1	284.1	0.284	0.334	-20.40	2.07
Great East Rd. Bridge	351.4	669.4	0.051	0.097	-0.74	1.05
Kafue Gorge	102.6	147.2	0.097	0.139	-0.64	0.46
Mangochi	34.8	41.7	0.117	0.141	0.65	0.96
Victoria Falls	393.4	651.7	0.071	0.118	-0.13	0.48
Marromeu (level)	0.542	0.777	0.098	0.141	-0.32	1.00
Tete (level)	0.311	0.580	0.103	0.192	-4.90	0.99
Itezhi-Tezhi	3.70E+09	4.18E+09	1.004	1.133	-32.82	0.12
Kafue Gorge	5.21E+08	5.94E+08	0.464	0.529	-265.14	0.61
Kariba	3.24E+09	4.02E+09	0.077	0.096	-9.42	0.79
Cahora Bassa	1.55E+10	1.76E+10	1.254	1.428	-112.53	1.30
Malawi Lake	5.70E+09	6.87E+09	0.119	0.143	0.35	1.00

Table AI.10. Full median ensemble results for the global calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	339.6	552.1	0.075	0.122	0.32	1.15
Watopa Pontoon	145.8	194.5	0.165	0.220	-0.44	0.23
Chiromo	208.8	284.1	0.105	0.142	-0.03	0.77
Chikwawa	127.2	159.9	0.124	0.156	0.31	0.94
Sesheke	496.3	814.6	0.100	0.164	0.30	0.75
Senanga	346.5	474.0	0.128	0.176	0.43	0.85
Lukulu	307.0	413.5	0.118	0.159	0.43	1.01
Kafue Hook Bridge	168.7	253.8	0.109	0.164	-0.14	0.27
Kalabo	49.9	84.3	0.032	0.053	-0.19	0.35
Chilenga	136.0	202.6	0.171	0.255	-0.51	0.14
Chifumpa	61.0	89.7	0.097	0.142	-0.42	0.26
Mfuwe	211.6	350.4	0.186	0.308	-0.99	1.60
Great East Rd. Bridge	471.6	1061.5	0.036	0.081	0.28	0.87
Kafue Gorge	115.1	146.8	0.099	0.127	-0.82	0.48
Mangochi	71.2	88.7	0.113	0.141	0.69	1.04
Victoria Falls	432.6	617.1	0.118	0.169	0.31	0.98
Marromeu (level)	0.803	1.033	0.130	0.167	0.40	1.05
Tete (level)	0.606	0.811	0.098	0.131	0.10	1.02
Itezhi-Tezhi	4.68E+09	4.84E+09	0.931	0.964	-14.95	0.06
Kafue Gorge	4.20E+08	5.04E+08	0.377	0.453	-2.38	0.19
Kariba	3.06E+10	3.48E+10	0.531	0.605	-4.73	0.80
Cahora Bassa	1.64E+10	1.92E+10	0.626	0.732	-10.17	1.36
Malawi Lake	9.79E+09	1.23E+10	0.091	0.114	0.74	1.00

Table AI.11. Full median ensemble results for the global calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	365.6	584.4	0.091	0.145	0.39	1.19
Watopa Pontoon	192.6	255.1	0.172	0.228	-0.22	0.34
Chiromo	297.7	325.1	0.319	0.349	-3.59	0.63
Chikwawa	191.5	212.2	0.266	0.295	-7.21	0.73
Sesheke	669.8	1094.8	0.118	0.193	0.28	0.83
Senanga	430.6	583.7	0.158	0.214	0.46	0.94
Lukulu	458.3	595.1	0.184	0.239	0.40	1.10
Kafue Hook Bridge	170.8	262.6	0.092	0.142	0.29	0.52
Kalabo	94.1	187.2	0.066	0.132	0.22	0.42
Chilenga	149.3	231.8	0.163	0.252	-0.13	0.26
Chifumpa	64.9	91.4	0.112	0.158	-0.07	0.33
Mfuwe	492.0	757.0	0.578	0.890	-13.86	3.52
Great East Rd. Bridge	602.9	996.8	0.088	0.145	-0.14	1.80
Kafue Gorge	120.5	169.6	0.114	0.160	0.22	0.76
Mangochi	36.7	42.8	0.124	0.144	0.72	0.98
Victoria Falls	576.7	822.5	0.105	0.149	0.51	0.94
Marromeu (level)	0.579	0.802	0.105	0.145	0.18	1.01
Tete (level)	0.416	0.702	0.137	0.232	-2.61	1.00
Itezhi-Tezhi	4.24E+09	4.60E+09	1.149	1.246	-27.71	0.23
Kafue Gorge	1.17E+09	1.53E+09	1.039	1.363	-46.64	2.60
Kariba	9.83E+09	1.18E+10	0.235	0.282	-0.21	1.06
Cahora Bassa	2.54E+10	2.82E+10	2.060	2.282	-90.96	1.51
Malawi Lake	5.94E+09	7.19E+09	0.124	0.150	0.65	1.00

Table AI.12. Full maximum ensemble results for the global calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	442.4	680.7	0.098	0.150	0.64	1.24
Watopa Pontoon	168.2	227.6	0.190	0.257	-0.21	0.29
Chiromo	221.5	301.7	0.111	0.151	0.25	0.89
Chikwawa	134.2	173.8	0.131	0.170	0.39	1.09
Sesheke	662.4	1093.7	0.133	0.220	0.49	0.80
Senanga	514.0	756.9	0.190	0.280	0.70	0.90
Lukulu	521.5	687.6	0.201	0.264	0.72	1.07
Kafue Hook Bridge	192.4	297.7	0.124	0.192	0.02	0.34
Kalabo	55.6	96.2	0.035	0.061	0.08	0.46
Chilenga	148.7	219.5	0.187	0.276	-0.41	0.17
Chifumpa	65.0	98.5	0.103	0.156	-0.20	0.31
Mfuwe	250.5	466.5	0.220	0.409	0.06	1.75
Great East Rd. Bridge	614.3	1243.0	0.047	0.095	0.42	0.95
Kafue Gorge	143.0	176.4	0.123	0.152	-0.53	0.60
Mangochi	103.6	123.5	0.165	0.197	0.70	1.21
Victoria Falls	609.0	822.0	0.167	0.225	0.58	1.04
Marromeu (level)	0.881	1.096	0.143	0.178	0.51	1.08
Tete (level)	0.703	0.980	0.113	0.158	0.29	1.06
Itezhi-Tezhi	4.88E+09	5.04E+09	0.971	1.004	-14.10	0.09
Kafue Gorge	5.43E+08	7.69E+08	0.488	0.691	-2.09	0.68
Kariba	9.09E+10	1.00E+11	1.579	1.745	-0.72	0.93
Cahora Bassa	1.80E+10	2.08E+10	0.687	0.794	-3.10	1.40
Malawi Lake	1.42E+10	1.75E+10	0.132	0.162	0.76	1.00

Table AI.13. Full maximum ensemble results for the global calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	542.8	818.4	0.135	0.203	0.69	1.26
Watopa Pontoon	223.8	311.6	0.200	0.279	0.07	0.42
Chiromo	316.7	345.1	0.340	0.370	-2.10	0.70
Chikwawa	211.5	236.7	0.294	0.329	-3.90	0.80
Sesheke	928.9	1444.7	0.164	0.255	0.47	0.89
Senanga	674.0	896.0	0.248	0.329	0.73	1.00
Lukulu	744.2	991.8	0.299	0.398	0.72	1.16
Kafue Hook Bridge	196.3	346.8	0.106	0.188	0.54	0.63
Kalabo	106.5	227.8	0.075	0.161	0.36	0.54
Chilenga	172.9	272.4	0.188	0.297	0.03	0.32
Chifumpa	72.3	109.3	0.125	0.189	0.17	0.40
Mfuwe	544.3	908.4	0.640	1.067	-1.09	3.86
Great East Rd. Bridge	688.2	1232.9	0.100	0.179	0.49	1.99
Kafue Gorge	170.3	245.1	0.161	0.232	0.41	0.93
Mangochi	42.5	47.6	0.143	0.161	0.73	1.00
Victoria Falls	812.4	1256.3	0.148	0.228	0.70	1.04
Marromeu (level)	0.848	1.016	0.154	0.184	0.22	1.13
Tete (level)	0.554	0.897	0.183	0.296	-1.47	1.05
Itezhi-Tezhi	4.85E+09	4.99E+09	1.314	1.353	-22.75	0.34
Kafue Gorge	3.09E+09	3.62E+09	2.753	3.221	-6.17	6.93
Kariba	3.14E+10	3.47E+10	0.751	0.829	0.86	1.11
Cahora Bassa	2.83E+10	3.13E+10	2.295	2.536	-34.98	1.57
Malawi Lake	8.03E+09	9.83E+09	0.167	0.205	0.68	1.00

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Their interpretation is introduced in the Section 9.3.2 of the main text.
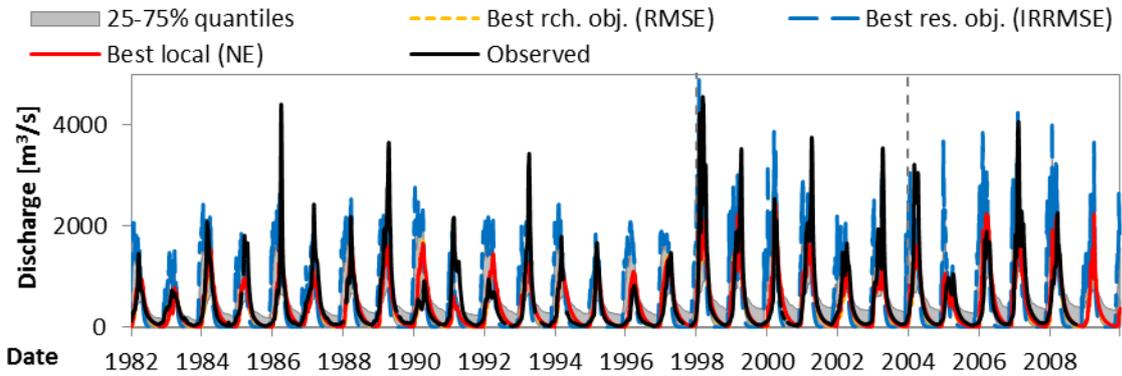
Figure AI.74. Synthesis hydrographs of the global calibration of the detailed model at Chavuma Falls.
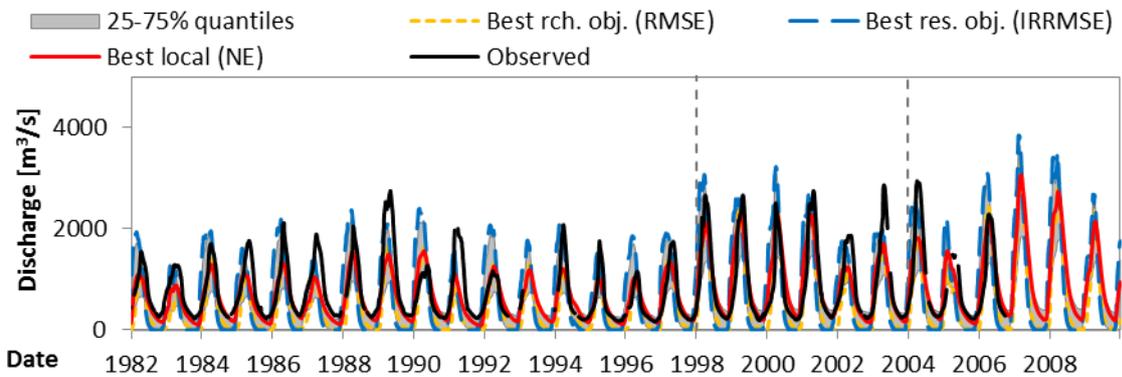
Figure AI.75. Synthesis hydrographs of the global calibration of the detailed model at Senanga.
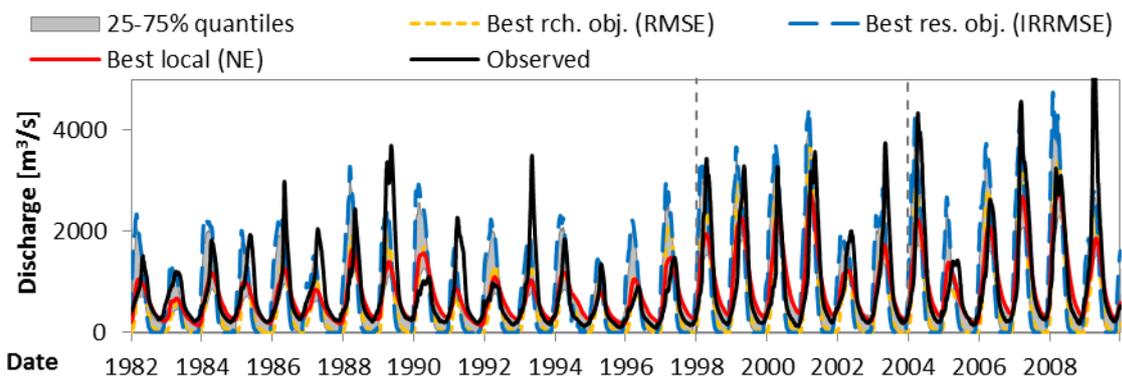
Figure AI.76. Synthesis hydrographs of the global calibration of the detailed model at Victoria Falls.
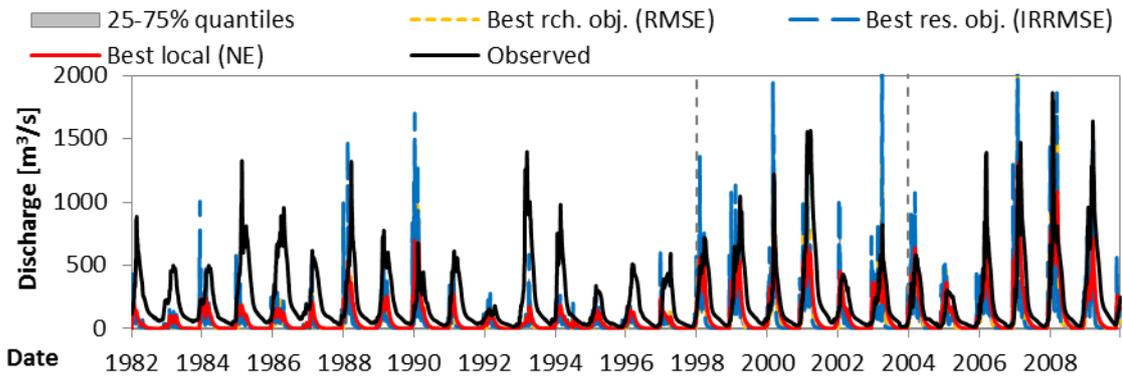
Figure AI.77. Synthesis hydrographs of the global calibration of the detailed model at Kafue Hook Bridge.
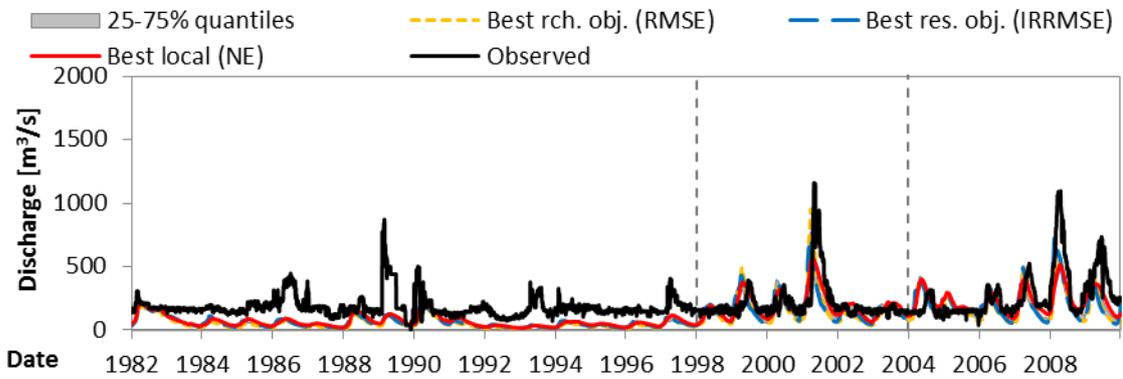
Figure AI.78. Synthesis hydrographs of the global calibration of the detailed model at Kafue Gorge.
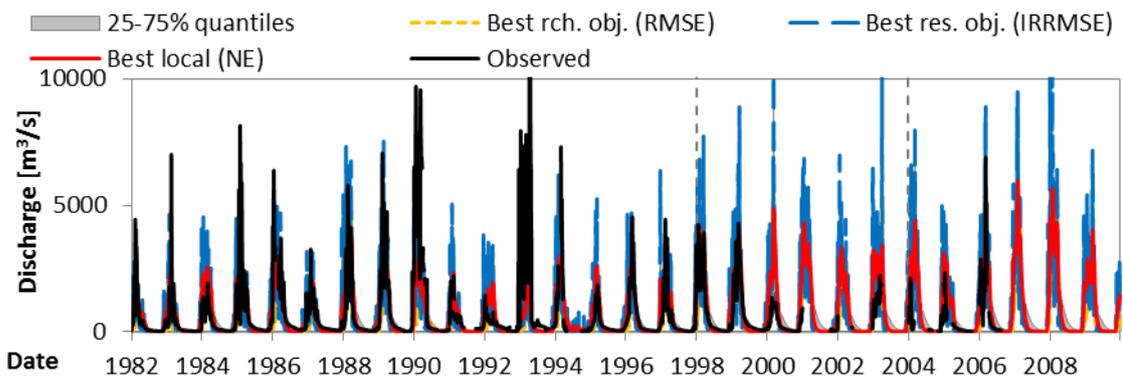
Figure AI.79. Synthesis hydrographs of the global calibration of the detailed model at Great East Road Bridge.
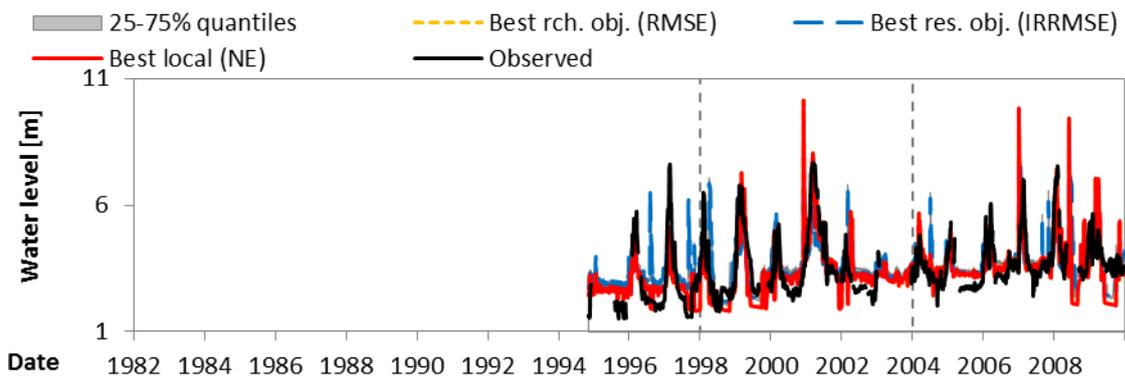
Figure AI.80. Synthesis water level series of the global calibration of the detailed model at Marromeu.
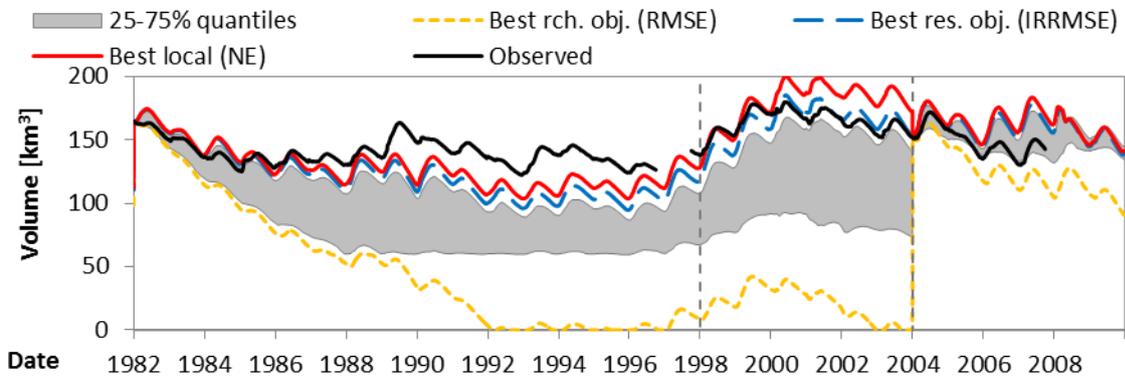
Figure AI.81. Synthesis volume series of the global calibration of the detailed model at the Kariba Dam.
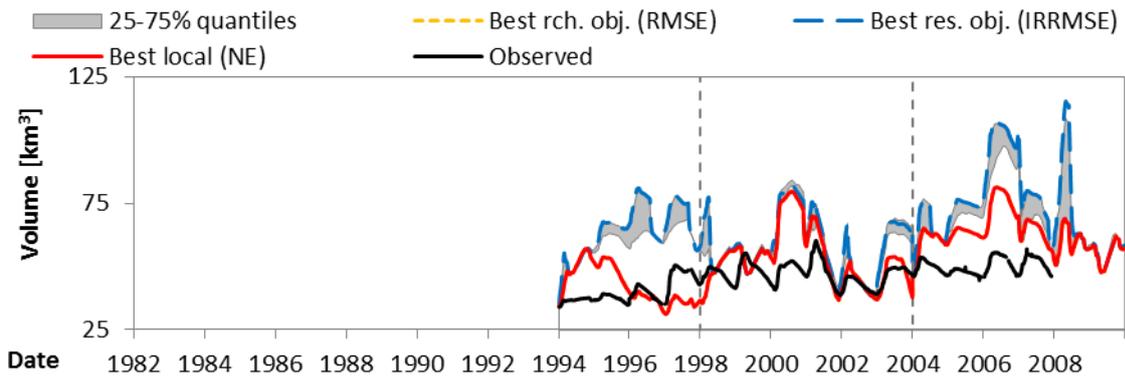
Figure AI.82. Synthesis volume series of the global calibration of the detailed model at the Cahora Bassa Dam.
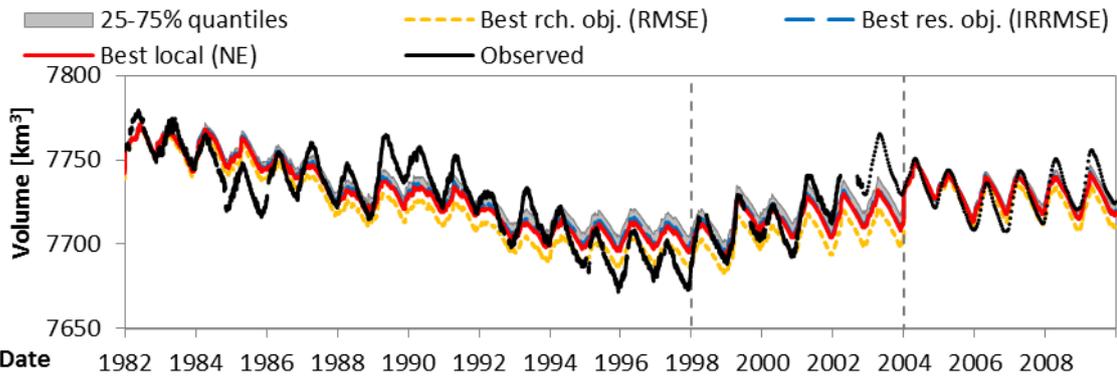
Figure AI.83. Synthesis volume series of the global calibration of the detailed model at the Malawi Lake.

Final parameter distribution
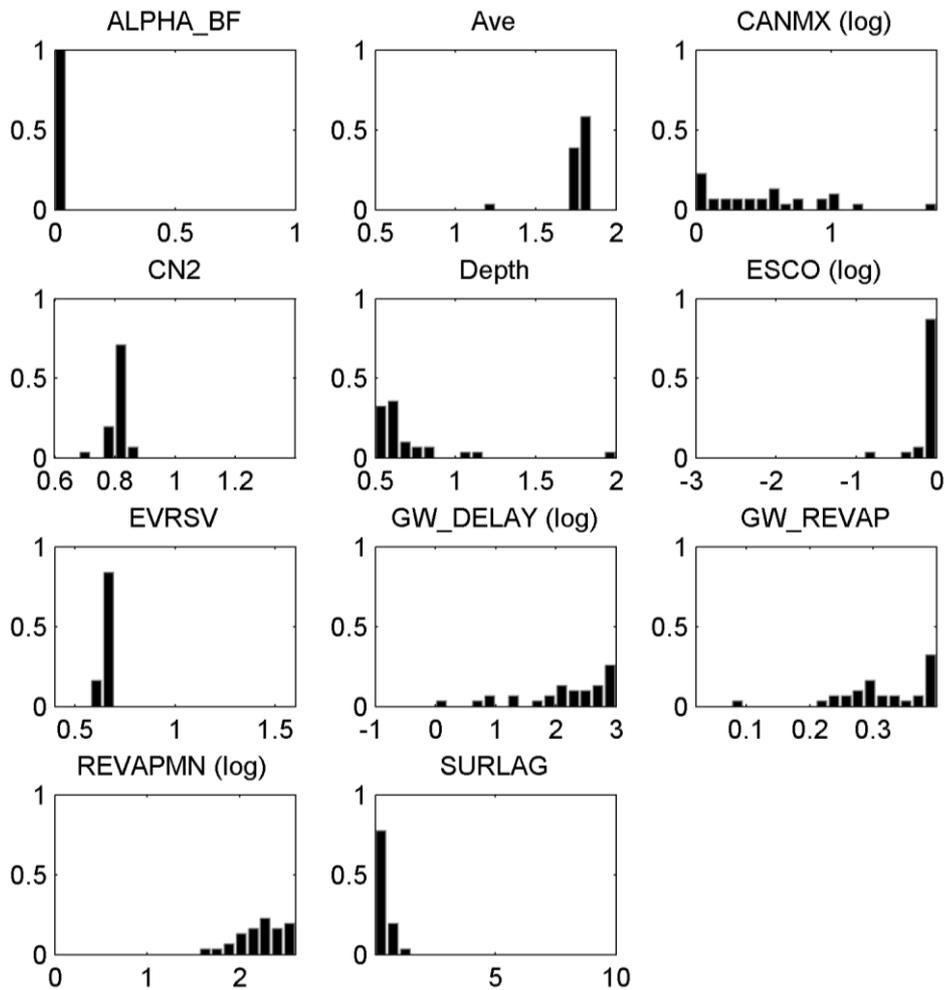
Figure AI.84. Parameter distribution among the ensemble for the last iteration of the global calibration of the detailed model. Histogram normalized according to ensemble size (40).

APPENDIX I.J. RESULTS FOR THE CASCADING CALIBRATION OF THE PARSIMONIOUS MODEL

Full results

The full results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Due to the fact that the cascading calibration does not make use of an ensemble, only two tables are presented.

Table AI.14. Full results for the cascading calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	223.2	379.3	0.049	0.084	0.68	0.92
Watopa Pontoon	60.6	91.6	0.068	0.103	0.68	0.95
Chiromo	312.0	415.7	0.156	0.208	-1.21	1.42
Chikwawa	302.5	383.7	0.296	0.375	-2.98	1.60
Sesheke	313.9	567.1	0.063	0.114	0.66	0.85
Senanga	178.2	272.9	0.066	0.101	0.81	0.98
Lukulu	187.0	295.4	0.072	0.114	0.71	1.01
Kafue Hook Bridge	79.2	136.7	0.051	0.088	0.67	1.00
Kalabo	29.3	63.9	0.019	0.040	0.31	1.01
Chilenga	59.6	91.6	0.075	0.115	0.69	1.02
Chifumpa	32.1	61.1	0.051	0.097	0.34	1.18
Mfuwe	95.9	155.6	0.084	0.137	0.61	0.92
Great East Rd. Bridge	425.8	932.4	0.033	0.071	0.45	0.99
Kafue Gorge	80.4	210.8	0.069	0.182	-2.75	1.30
Mangochi	200.7	240.5	0.320	0.383	-1.31	1.44
Victoria Falls	220.5	343.4	0.060	0.094	0.79	1.01
Marromeu (level)	0.660	0.837	0.107	0.136	0.61	1.05
Tete (level)	0.530	0.723	0.085	0.117	0.28	0.98
Itezhi-Tezhi	3.23E+09	4.06E+09	0.643	0.809	-10.23	1.12
Kafue Gorge	8.76E+09	1.41E+10	7.869	12.714	-2662.85	19.51
Kariba	2.07E+10	2.49E+10	0.360	0.432	-1.92	0.86
Cahora Bassa	9.35E+09	1.22E+10	0.357	0.466	-3.52	1.16
Malawi Lake	2.59E+10	3.18E+10	0.241	0.296	-0.74	1.00

Table AI.15. Full results for the cascading calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	251.1	407.2	0.062	0.101	0.70	0.98
Watopa Pontoon	91.9	137.0	0.082	0.123	0.65	0.98
Chiromo	275.6	401.9	0.296	0.431	-6.02	1.16
Chikwawa	259.7	363.7	0.361	0.505	-23.12	1.22
Sesheke	448.2	881.1	0.079	0.155	0.53	0.77
Senanga	170.8	289.1	0.063	0.106	0.87	0.91
Lukulu	257.2	432.1	0.103	0.174	0.68	0.81
Kafue Hook Bridge	100.1	149.6	0.054	0.081	0.77	1.17
Kalabo	85.4	199.0	0.060	0.140	0.12	0.62
Chilenga	67.2	96.2	0.073	0.105	0.81	1.08
Chifumpa	34.4	62.0	0.059	0.107	0.51	1.13
Mfuwe	176.3	228.3	0.207	0.268	-0.35	1.85
Great East Rd. Bridge	614.8	1028.1	0.089	0.150	-0.21	2.00
Kafue Gorge	151.8	259.2	0.143	0.245	-0.83	1.41
Mangochi	42.5	49.9	0.143	0.168	0.61	0.98
Victoria Falls	365.6	641.9	0.066	0.117	0.70	0.93
Marromeu (level)	0.563	0.770	0.102	0.139	0.24	0.95
Tete (level)	0.476	0.758	0.157	0.250	-3.22	0.90
Itezhi-Tezhi	4.42E+09	5.39E+09	1.199	1.460	-38.41	1.80
Kafue Gorge	1.23E+10	1.45E+10	10.947	12.914	-4277.30	25.33
Kariba	3.57E+09	4.12E+09	0.085	0.098	0.85	1.00
Cahora Bassa	9.79E+09	1.16E+10	0.793	0.940	-14.62	1.09
Malawi Lake	2.40E+10	2.98E+10	0.500	0.620	-4.96	1.00

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Observations are compared against the optimized cascading simulation, which maximized the Nash-Sutcliffe efficiency coefficient (NE).
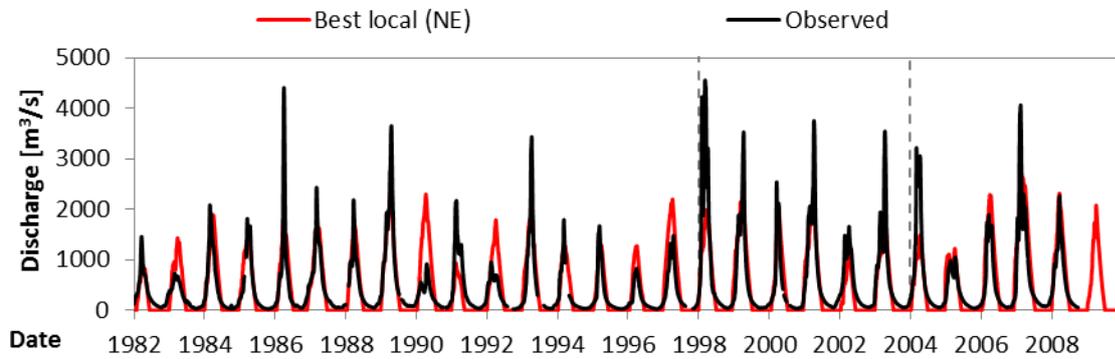
Figure AI.85. Observed and simulated hydrographs after the cascading calibration of the parsimonious model at Chavuma Falls.
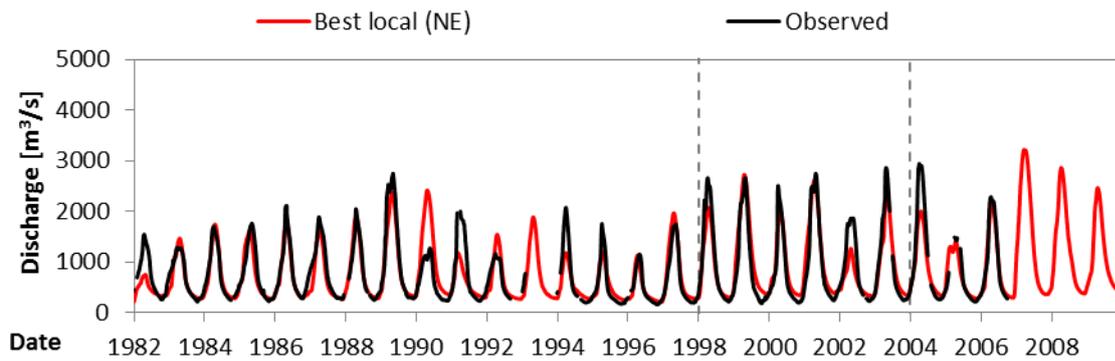
Figure AI.86. Observed and simulated hydrographs after the cascading calibration of the parsimonious model at Senanga.
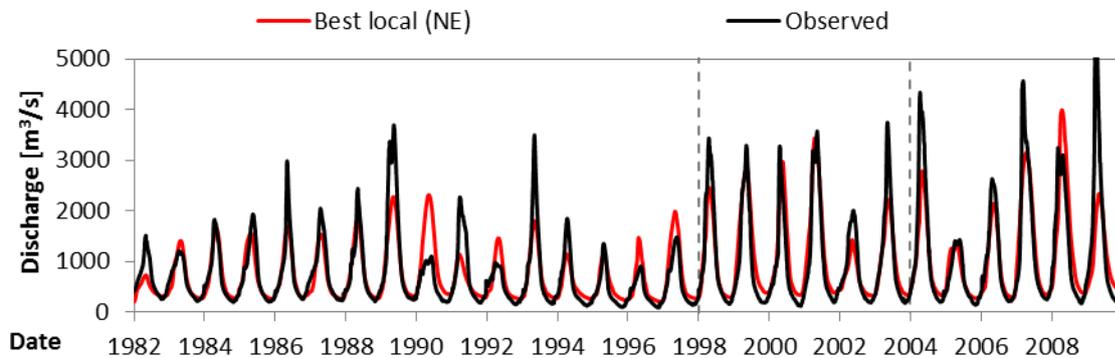
Figure AI.87. Observed and simulated hydrographs after the cascading calibration of the parsimonious model at Victoria Falls.
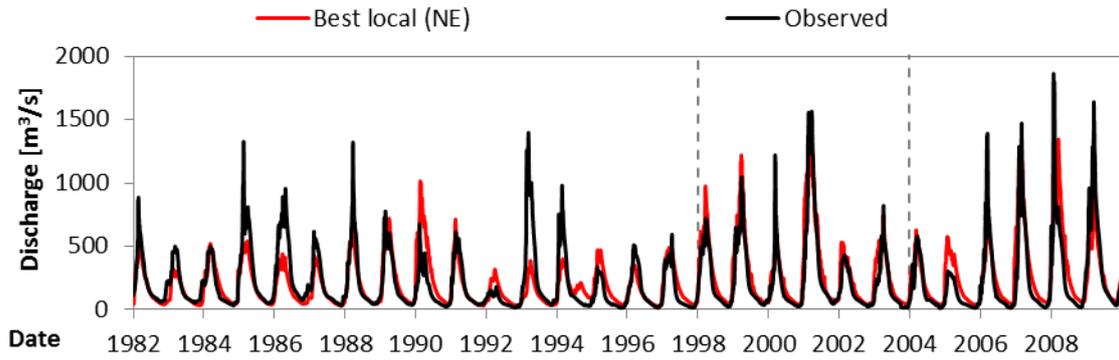
Figure AI.88. Observed and simulated hydrographs after the cascading calibration of the parsimonious model at Kafue Hook Bridge.
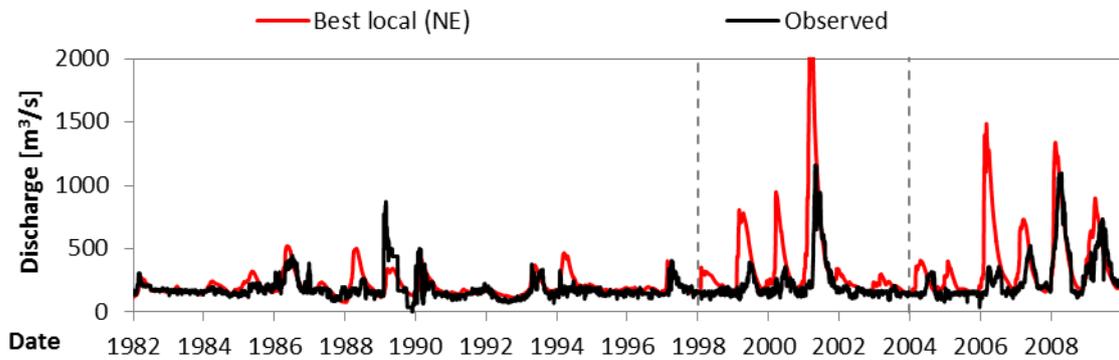
Figure AI.89. Observed and simulated hydrographs after the cascading calibration of the parsimonious model at Kafue Gorge.
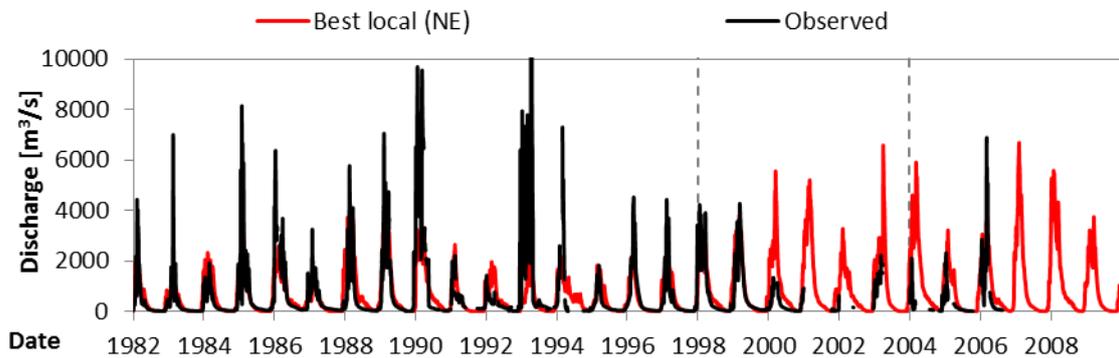
Figure AI.90. Observed and simulated hydrographs after the cascading calibration of the parsimonious model at Great East Road Bridge.
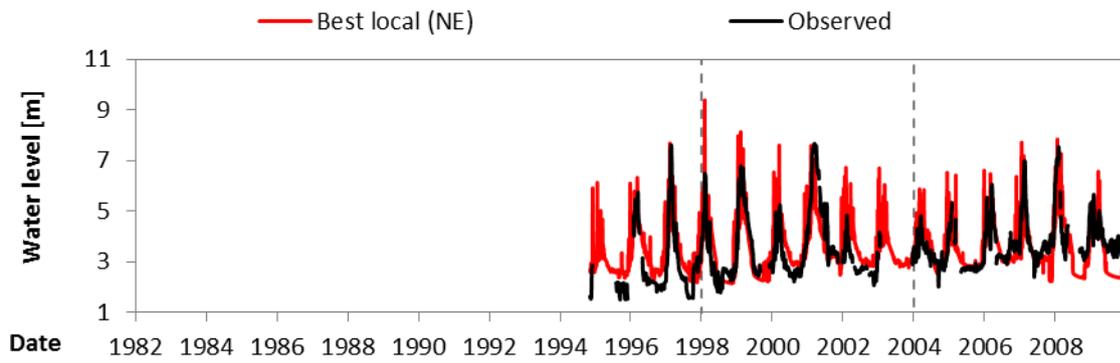
Figure AI.91. Observed and simulated water level series after the cascading calibration of the parsimonious model at Marromeu.
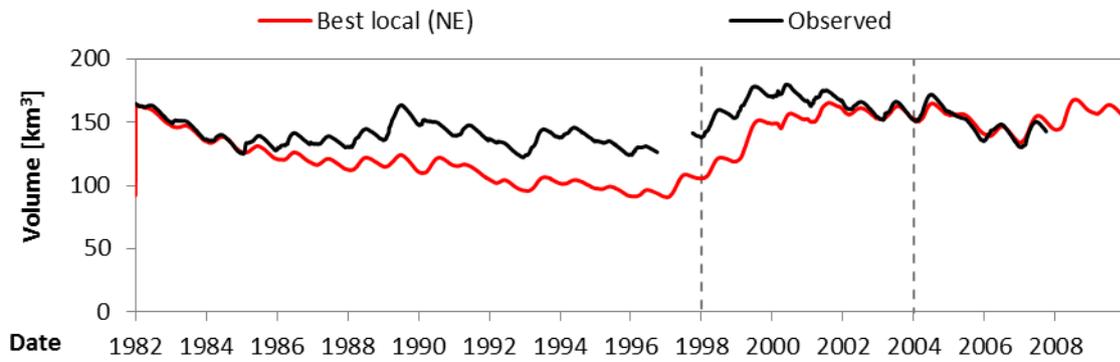
Figure AI.92. Observed and simulated volume series after the cascading calibration of the parsimonious model at the Kariba Dam.
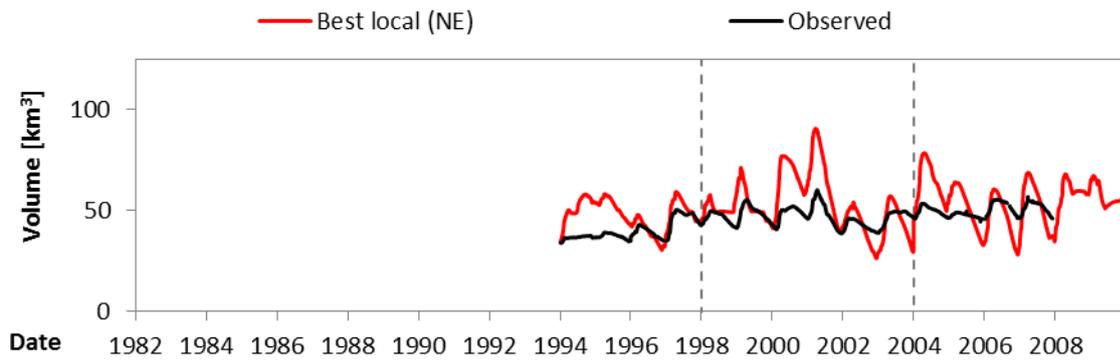
Figure AI.93. Observed and simulated volume series after the cascading calibration of the parsimonious model at the Cahora Bassa Dam.
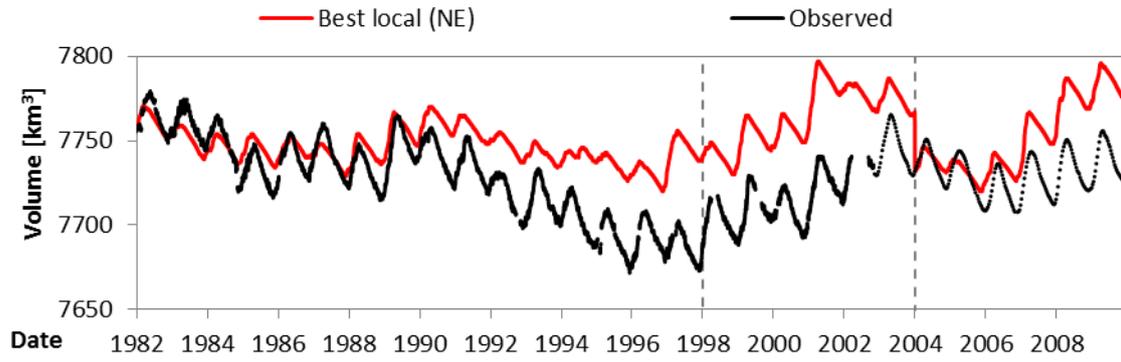
Figure AI.94. Observed and simulated volume series after the cascading calibration of the parsimonious model at the Malawi Lake.

APPENDIX I.K. RESULTS FOR THE CASCADING CALIBRATION OF THE DETAILED MODEL

Full results

The full results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Due to the fact that the cascading calibration does not make use of an ensemble, only two tables are presented.

Table AI.16. Full results for the cascading calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	204.2	361.2	0.045	0.080	0.71	1.07
Watopa Pontoon	73.8	105.0	0.083	0.119	0.58	0.85
Chiromo	835.3	929.2	0.418	0.465	-10.03	2.33
Chikwawa	859.6	943.1	0.840	0.922	-23.02	2.73
Sesheke	332.5	566.2	0.067	0.114	0.66	0.88
Senanga	197.8	282.6	0.073	0.105	0.80	1.02
Lukulu	243.1	352.9	0.094	0.136	0.58	0.99
Kafue Hook Bridge	91.0	150.8	0.059	0.097	0.60	0.87
Kalabo	27.8	59.5	0.018	0.038	0.41	0.93
Chilenga	66.1	102.5	0.083	0.129	0.61	0.91
Chifumpa	34.7	63.5	0.055	0.101	0.29	0.78
Mfuwe	108.5	164.0	0.095	0.144	0.56	0.83
Great East Rd. Bridge	408.8	950.6	0.031	0.073	0.43	0.96
Kafue Gorge	232.5	325.9	0.201	0.281	-7.95	2.18
Mangochi	747.2	770.8	1.191	1.228	-22.73	2.76
Victoria Falls	242.0	368.6	0.066	0.101	0.75	1.02
Marromeu (level)	0.714	1.035	0.116	0.168	0.40	0.99
Tete (level)	0.666	0.893	0.107	0.144	-0.09	1.12
Itezhi-Tezhi	3.58E+09	4.14E+09	0.714	0.823	-10.63	0.59
Kafue Gorge	5.21E+10	6.62E+10	46.812	59.468	-58276.15	111.5
Kariba	1.63E+10	1.95E+10	0.283	0.338	-0.79	0.93
Cahora Bassa	4.55E+10	5.20E+10	1.736	1.984	-81.00	2.03
Malawi Lake	8.49E+10	8.92E+10	0.788	0.828	-12.65	1.01

Table AI.17. Full results for the cascading calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	231.2	421.7	0.057	0.105	0.68	1.15
Watopa Pontoon	102.5	148.4	0.092	0.133	0.59	1.05
Chiromo	564.9	767.0	0.606	0.822	-24.55	1.69
Chikwawa	554.9	718.4	0.771	0.998	-93.11	1.82
Sesheke	486.2	837.4	0.086	0.148	0.58	0.96
Senanga	252.5	347.3	0.093	0.128	0.81	1.12
Lukulu	274.0	413.6	0.110	0.166	0.71	1.06
Kafue Hook Bridge	132.3	198.0	0.072	0.107	0.60	1.39
Kalabo	76.0	160.8	0.054	0.113	0.43	0.93
Chilenga	80.2	120.8	0.087	0.131	0.69	1.21
Chifumpa	38.8	59.7	0.067	0.103	0.54	1.01
Mfuwe	176.7	238.8	0.208	0.281	-0.48	1.77
Great East Rd. Bridge	522.9	908.3	0.076	0.132	0.05	1.84
Kafue Gorge	386.8	444.5	0.366	0.420	-4.39	2.40
Mangochi	156.4	179.6	0.527	0.606	-4.01	1.29
Victoria Falls	367.6	606.1	0.067	0.110	0.74	1.03
Marromeu (level)	0.680	0.898	0.123	0.163	-0.03	0.92
Tete (level)	0.403	0.540	0.133	0.178	-1.14	1.08
Itezhi-Tezhi	9.39E+09	1.09E+10	2.546	2.956	-160.53	2.70
Kafue Gorge	3.96E+10	4.44E+10	35.214	39.525	-4.01E+04	79.25
Kariba	1.25E+10	1.57E+10	0.300	0.374	-1.13	1.07
Cahora Bassa	7.29E+10	8.16E+10	5.902	6.611	-770.78	2.45
Malawi Lake	6.96E+10	7.47E+10	1.448	1.555	-36.52	1.01

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Observations are compared against the optimized cascading simulation, which maximized the Nash-Sutcliffe efficiency coefficient (NE).
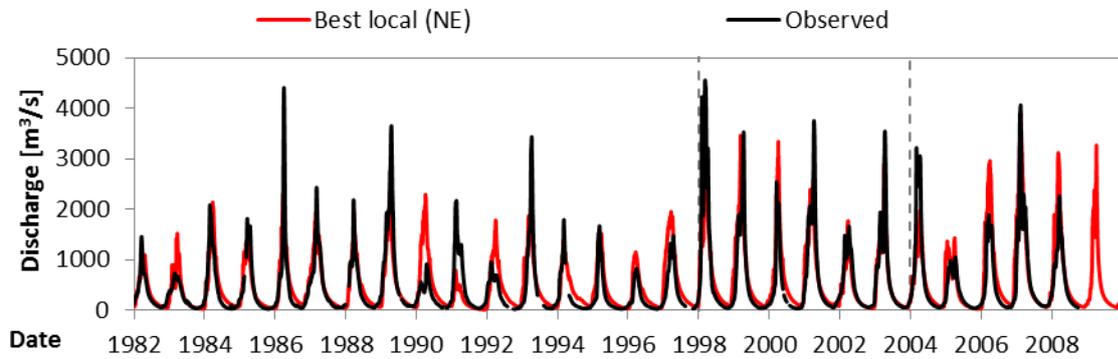

Figure AI.95. Observed and simulated hydrographs after the cascading calibration of the detailed model at Chavuma Falls.
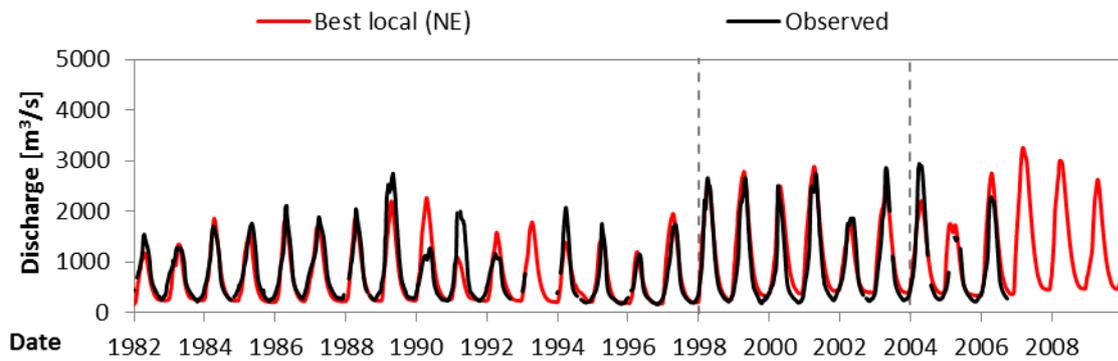

Figure AI.96. Observed and simulated hydrographs after the cascading calibration of the detailed model at Senanga.
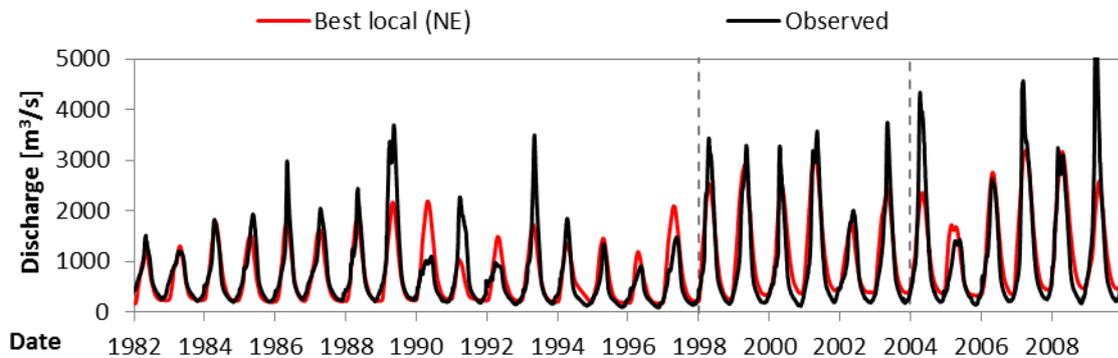
Figure AI.97. Observed and simulated hydrographs after the cascading calibration of the detailed model at Victoria Falls.
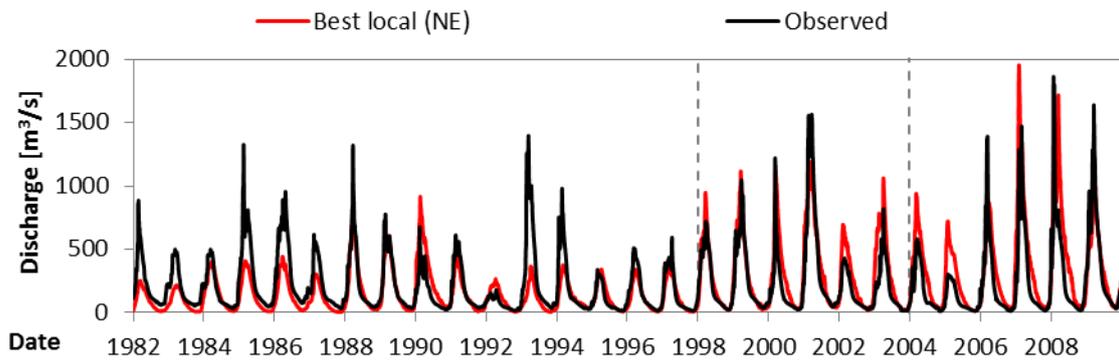
Figure AI.98. Observed and simulated hydrographs after the cascading calibration of the detailed model at Kafue Hook Bridge.
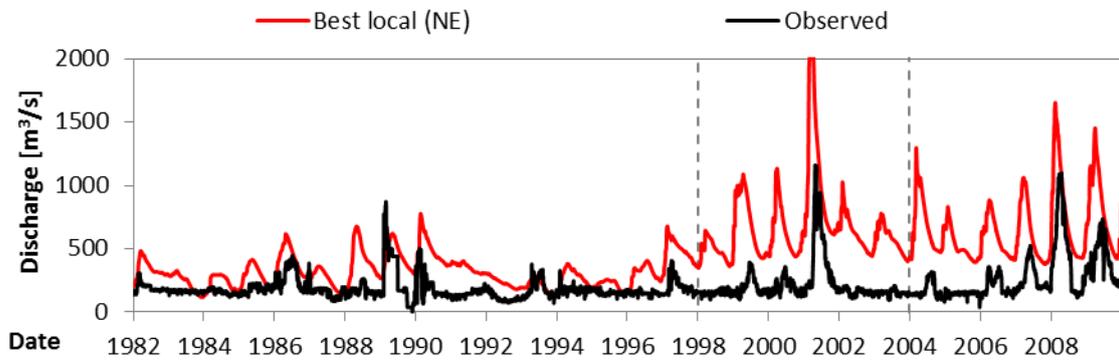
Figure AI.99. Observed and simulated hydrographs after the cascading calibration of the detailed model at Kafue Gorge.
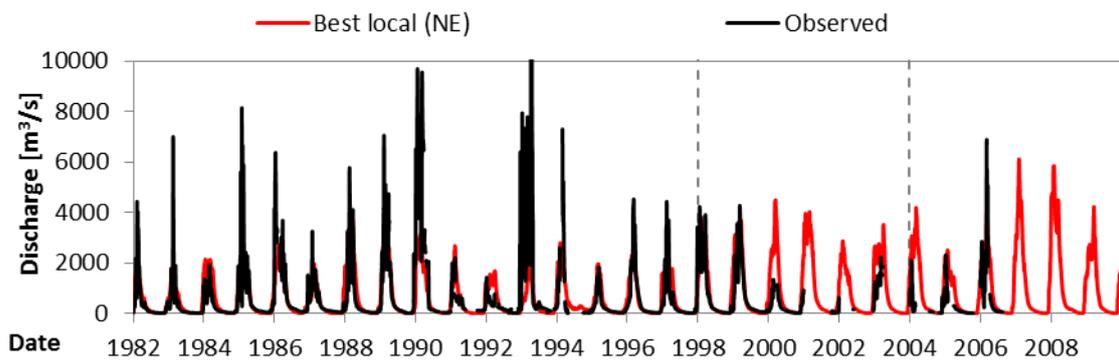
Figure AI.100. Observed and simulated hydrographs after the cascading calibration of the detailed model at Great East Road Bridge.
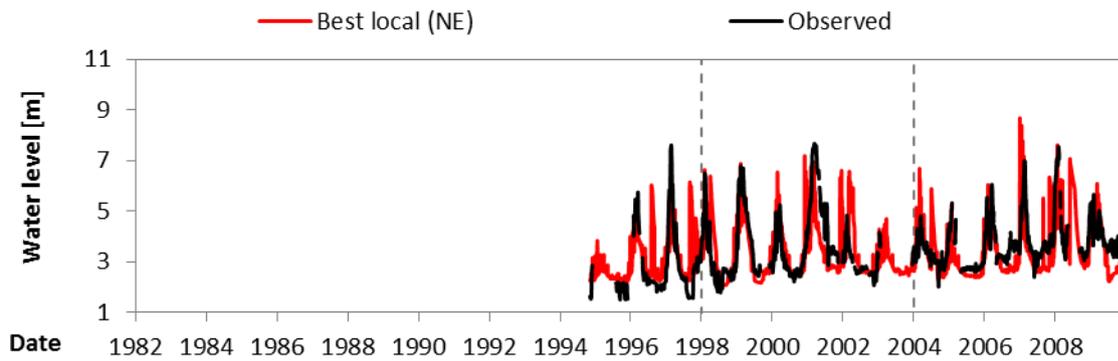
Figure AI.101. Observed and simulated water level series after the cascading calibration of the detailed model at Marromeu.
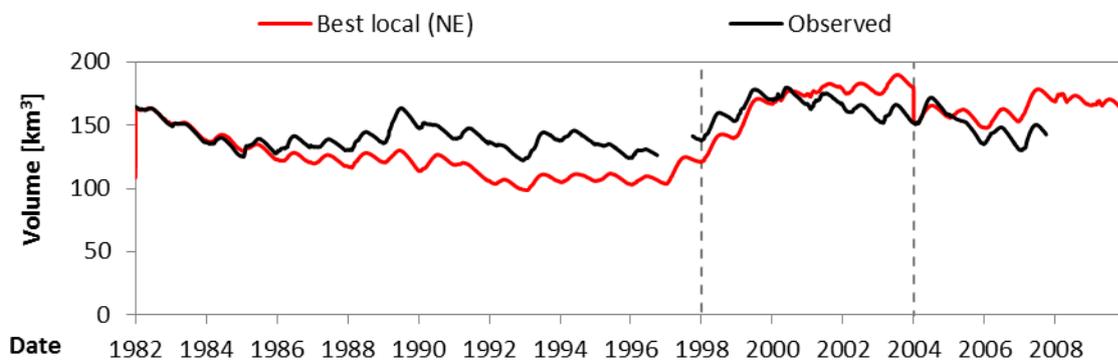
Figure AI.102. Observed and simulated stored volume series after the cascading calibration of the detailed model at the Kariba Dam.
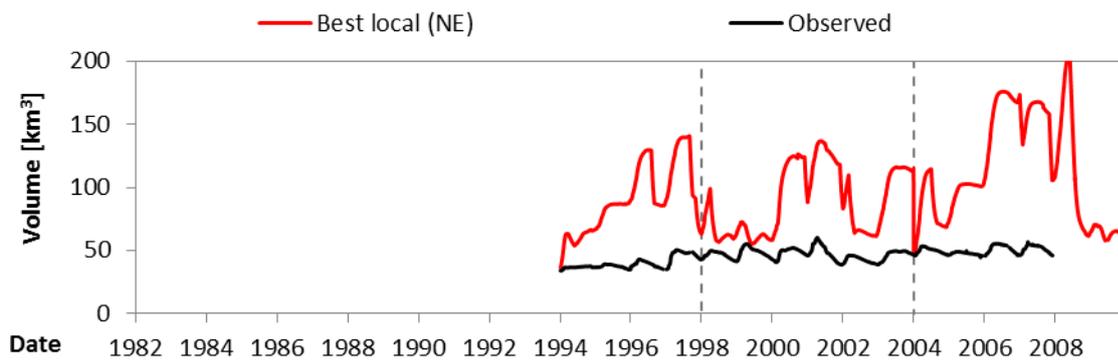
Figure AI.103. Observed and simulated stored volume series after the cascading calibration of the detailed model at the Cahora Bassa Dam.
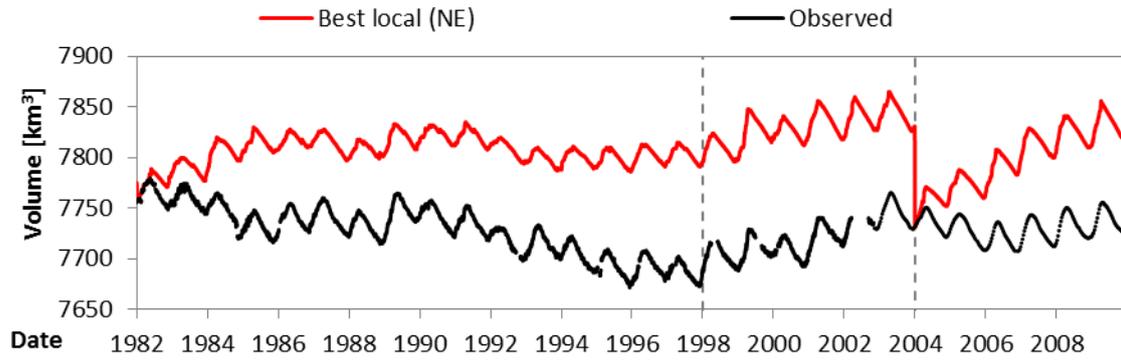
Figure AI.104. Observed and simulated stored volume series after the cascading calibration of the detailed model at the Malawi Lake.

APPENDIX I.L. ADDITIONAL RESULTS OF THE FAST
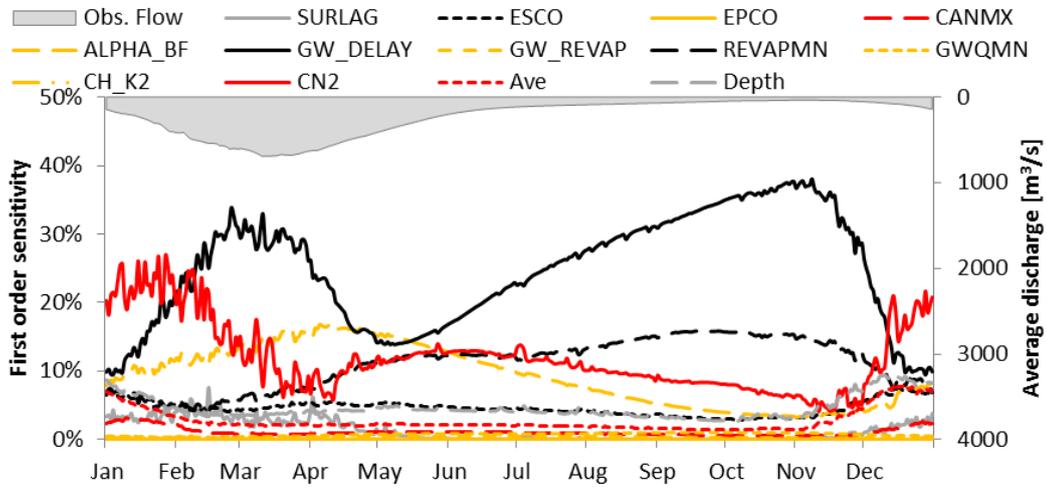
Figure AI.105. Results of the FAST applied directly to simulation results. Results of the parsimonious model for Kafue Hook Bridge.
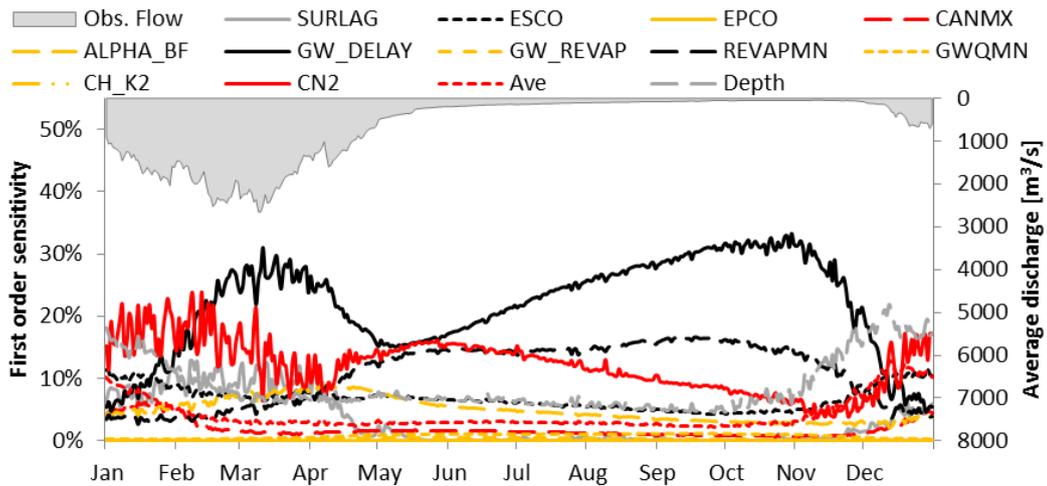
Figure AI.106. Results of the FAST applied directly to simulation results. Results of the parsimonious model for Great East Road Bridge.

APPENDIX I.M. RESULTS FOR THE REGIONAL-REGULARIZED CALIBRATION OF THE PARSIMONIOUS MODEL

Full ensemble results

The full ensemble results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Results from the 10% worst performing parameter sets regarding each calibration objective are not included.

Table AI.18. Full minimum ensemble results for the regional-regularized calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	310.6	520.5	0.069	0.115	-9.06	1.24
Watopa Pontoon	103.6	139.6	0.117	0.158	-49.55	0.02
Chiromo	193.9	280.8	0.097	0.141	-5.46	0.34
Chikwawa	119.3	154.0	0.117	0.150	-13.94	0.42
Sesheke	397.2	573.1	0.080	0.115	-7.89	0.67
Senanga	305.8	401.9	0.113	0.149	-17.29	0.76
Lukulu	345.5	457.0	0.133	0.176	-22.27	0.89
Kafue Hook Bridge	99.8	178.1	0.064	0.115	-146.58	0.01
Kalabo	48.8	88.0	0.031	0.056	-44.66	0.01
Chilenga	66.0	114.5	0.083	0.144	-45.53	0.01
Chifumpa	40.4	69.3	0.064	0.110	-115.88	0.02
Mfuwe	110.7	187.8	0.097	0.165	-44.62	0.00
Great East Rd. Bridge	416.8	953.2	0.032	0.073	-5.63	0.01
Kafue Gorge	76.9	111.7	0.066	0.096	-55.06	0.07
Mangochi	88.1	106.8	0.140	0.170	-18.61	0.46
Victoria Falls	329.2	444.8	0.090	0.122	-47.19	0.76
Marromeu (level)	0.649	0.884	0.105	0.143	-0.34	0.95
Tete (level)	0.520	0.694	0.084	0.112	-4.21	0.79
Itezhi-Tezhi	2.72E+09	3.28E+09	0.542	0.653	-3.49E+05	0.02
Kafue Gorge	4.09E+08	4.97E+08	0.367	0.446	-4.60E+05	0.00
Kariba	9.28E+09	1.19E+10	0.161	0.206	-1.07E+04	0.58
Cahora Bassa	8.99E+09	1.14E+10	0.343	0.436	-286.96	0.66
Malawi Lake	1.25E+10	1.49E+10	0.116	0.138	-10.41	1.00

Table AI.19. Full minimum ensemble results for the regional-regularized calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	342.8	558.3	0.085	0.139	-7.84	1.17
Watopa Pontoon	149.6	193.7	0.134	0.174	-41.74	0.02
Chiromo	285.2	317.4	0.306	0.340	-12.26	0.42
Chikwawa	193.6	222.8	0.269	0.310	-64.25	0.48
Sesheke	453.8	740.4	0.080	0.131	-4.68	0.55
Senanga	338.8	436.5	0.124	0.160	-11.93	0.64
Lukulu	357.3	497.5	0.144	0.200	-11.22	0.72
Kafue Hook Bridge	113.8	187.5	0.062	0.101	-122.55	0.03
Kalabo	107.4	225.0	0.076	0.159	-7.73	0.01
Chilenga	73.0	126.1	0.079	0.137	-33.50	0.01
Chifumpa	46.3	73.1	0.080	0.127	-107.37	0.03
Mfuwe	68.0	104.2	0.080	0.122	-135.70	0.01
Great East Rd. Bridge	331.8	669.2	0.048	0.097	-22.16	0.03
Kafue Gorge	93.2	145.0	0.088	0.137	-29.46	0.13
Mangochi	59.4	72.3	0.200	0.244	-2.15	0.83
Victoria Falls	429.0	662.9	0.078	0.120	-23.91	0.56
Marromeu (level)	0.609	0.809	0.110	0.146	-1.28	0.88
Tete (level)	0.325	0.489	0.107	0.161	-29.39	0.86
Itezhi-Tezhi	1.40E+09	2.02E+09	0.379	0.548	-7.97E+04	0.05
Kafue Gorge	3.91E+08	4.38E+08	0.348	0.390	-1.16E+05	0.01
Kariba	5.05E+09	5.88E+09	0.121	0.141	-936.13	0.85
Cahora Bassa	3.61E+09	4.76E+09	0.293	0.385	-2521.12	0.56
Malawi Lake	6.98E+09	8.70E+09	0.145	0.181	-32.95	1.00

Table AI.20. Full median ensemble results for the regional-regularized calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	533.4	666.3	0.118	0.147	0.01	1.55
Watopa Pontoon	182.4	241.3	0.206	0.273	-1.22	0.03
Chiromo	236.2	329.7	0.118	0.165	-0.39	0.66
Chikwawa	143.5	182.0	0.140	0.178	0.11	0.81
Sesheke	534.7	812.6	0.108	0.164	0.30	0.85
Senanga	418.1	522.7	0.155	0.194	0.31	0.97
Lukulu	426.3	576.9	0.164	0.222	-0.11	1.20
Kafue Hook Bridge	194.8	290.3	0.126	0.187	-0.50	0.17
Kalabo	51.2	91.8	0.032	0.058	-0.41	0.02
Chilenga	144.8	212.5	0.182	0.267	-0.66	0.11
Chifumpa	66.3	97.1	0.105	0.154	-0.67	0.18
Mfuwe	200.4	318.5	0.176	0.280	-0.65	0.01
Great East Rd. Bridge	670.3	1413.5	0.051	0.108	-0.27	0.03
Kafue Gorge	151.1	185.0	0.130	0.160	-1.88	0.33
Mangochi	90.4	109.4	0.144	0.174	0.52	0.92
Victoria Falls	463.5	620.7	0.127	0.170	0.30	0.97
Marromeu (level)	0.965	1.151	0.156	0.187	0.26	1.05
Tete (level)	0.941	1.215	0.152	0.196	-1.03	0.93
Itezhi-Tezhi	4.88E+09	5.04E+09	0.971	1.004	-16.28	0.03
Kafue Gorge	4.69E+08	5.44E+08	0.422	0.489	-2.93	0.02
Kariba	3.51E+10	4.11E+10	0.609	0.715	-7.00	1.00
Cahora Bassa	1.78E+10	2.25E+10	0.682	0.860	-14.42	0.83
Malawi Lake	1.28E+10	1.54E+10	0.119	0.143	0.59	1.00

Table AI.21. Full median ensemble results for the regional-regularized calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	561.0	715.3	0.139	0.178	0.09	1.56
Watopa Pontoon	266.1	349.3	0.238	0.313	-1.28	0.02
Chiromo	376.6	400.0	0.404	0.429	-5.95	0.53
Chikwawa	251.9	264.6	0.350	0.368	-11.76	0.63
Sesheke	685.6	1043.6	0.121	0.184	0.34	0.76
Senanga	489.1	689.3	0.180	0.253	0.25	0.89
Lukulu	552.8	723.3	0.222	0.291	0.11	1.10
Kafue Hook Bridge	219.7	354.7	0.119	0.192	-0.29	0.28
Kalabo	112.1	238.1	0.079	0.168	-0.26	0.02
Chilenga	178.1	270.6	0.194	0.295	-0.54	0.14
Chifumpa	78.4	112.5	0.136	0.195	-0.63	0.21
Mfuwe	174.2	260.3	0.205	0.306	-0.76	0.06
Great East Rd. Bridge	496.8	1043.0	0.072	0.152	-0.25	0.08
Kafue Gorge	198.9	268.7	0.188	0.254	-0.97	0.45
Mangochi	60.6	73.5	0.204	0.248	0.16	0.89
Victoria Falls	669.5	982.5	0.122	0.178	0.31	0.74
Marromeu (level)	0.907	1.096	0.164	0.198	-0.54	1.04
Tete (level)	0.908	1.294	0.300	0.428	-11.29	0.97
Itezhi-Tezhi	5.15E+09	5.30E+09	1.395	1.437	-37.15	0.11
Kafue Gorge	4.98E+08	5.46E+08	0.443	0.486	-5.06	0.18
Kariba	1.51E+10	1.69E+10	0.361	0.404	-1.48	0.99
Cahora Bassa	1.81E+10	2.33E+10	1.470	1.890	-62.05	0.70
Malawi Lake	7.56E+09	9.45E+09	0.157	0.197	0.40	1.00

Table AI.22. Full maximum ensemble results for the regional-regularized calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	1367.9	2127.3	0.302	0.469	0.40	3.61
Watopa Pontoon	647.7	1152.0	0.732	1.301	0.26	3.68
Chiromo	646.6	710.9	0.324	0.356	-0.01	1.99
Chikwawa	677.8	743.8	0.662	0.727	0.36	2.37
Sesheke	1971.3	2894.7	0.397	0.583	0.65	3.14
Senanga	1876.2	2685.0	0.695	0.995	0.59	3.31
Lukulu	1745.1	2636.4	0.671	1.014	0.30	3.43
Kafue Hook Bridge	1756.8	2884.2	1.132	1.859	0.44	8.86
Kalabo	249.2	521.6	0.158	0.330	-0.30	4.84
Chilenga	681.6	1125.2	0.857	1.414	0.52	5.04
Chifumpa	481.2	812.3	0.763	1.288	0.15	7.19
Mfuwe	1065.6	1676.7	0.935	1.472	0.43	6.18
Great East Rd. Bridge	2051.9	3230.5	0.157	0.247	0.42	3.83
Kafue Gorge	532.4	815.5	0.459	0.703	-0.05	3.76
Mangochi	648.5	700.9	1.033	1.117	0.54	2.53
Victoria Falls	3729.1	5160.7	1.021	1.412	0.64	5.65
Marromeu (level)	1.372	1.551	0.222	0.251	0.56	1.31
Tete (level)	1.268	1.947	0.204	0.314	0.34	1.19
Itezhi-Tezhi	6.18E+11	7.17E+11	123.131	142.716	-6.31	125.16
Kafue Gorge	1.59E+11	1.86E+11	142.987	167.045	-2.28	338.49
Kariba	1.29E+12	1.51E+12	22.370	26.181	0.34	9.77
Cahora Bassa	8.03E+10	9.74E+10	3.066	3.718	-2.97	2.81
Malawi Lake	7.45E+10	8.16E+10	0.692	0.757	0.62	1.01

Table AI.23. Full maximum ensemble results for the regional-regularized calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	1380.7	2229.0	0.343	0.554	0.45	3.15
Watopa Pontoon	850.5	1512.1	0.762	1.354	0.30	3.43
Chiromo	467.8	552.4	0.502	0.592	-3.38	1.43
Chikwawa	451.2	598.2	0.627	0.831	-8.05	1.63
Sesheke	2357.9	3069.6	0.416	0.541	0.67	2.96
Senanga	2068.8	2856.1	0.760	1.049	0.70	3.19
Lukulu	1915.4	2676.2	0.770	1.075	0.58	2.98
Kafue Hook Bridge	2080.9	3466.4	1.125	1.874	0.64	8.94
Kalabo	346.7	626.5	0.244	0.441	-0.13	3.56
Chilenga	772.2	1282.4	0.841	1.396	0.67	4.64
Chifumpa	543.4	918.0	0.941	1.589	0.31	6.84
Mfuwe	1677.0	2295.7	1.971	2.698	0.72	10.35
Great East Rd. Bridge	2805.7	4494.5	0.408	0.654	0.49	6.28
Kafue Gorge	646.5	1057.0	0.611	0.999	0.43	3.30
Mangochi	129.6	142.6	0.437	0.481	0.19	1.23
Victoria Falls	4119.7	5893.7	0.748	1.070	0.68	4.36
Marromeu (level)	1.182	1.334	0.214	0.242	0.16	1.21
Tete (level)	1.190	2.035	0.393	0.672	-0.75	1.14
Itezhi-Tezhi	2.08E+11	2.42E+11	56.352	65.667	-4.55	38.57
Kafue Gorge	6.60E+10	7.57E+10	58.718	67.364	-2.91	131.47
Kariba	2.88E+11	3.29E+11	6.889	7.860	0.70	2.93
Cahora Bassa	1.28E+11	1.48E+11	10.376	11.951	-1.62	3.56
Malawi Lake	6.19E+10	7.11E+10	1.287	1.479	0.49	1.01

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Their interpretation is introduced in the Section 9.3.2 of the main text.
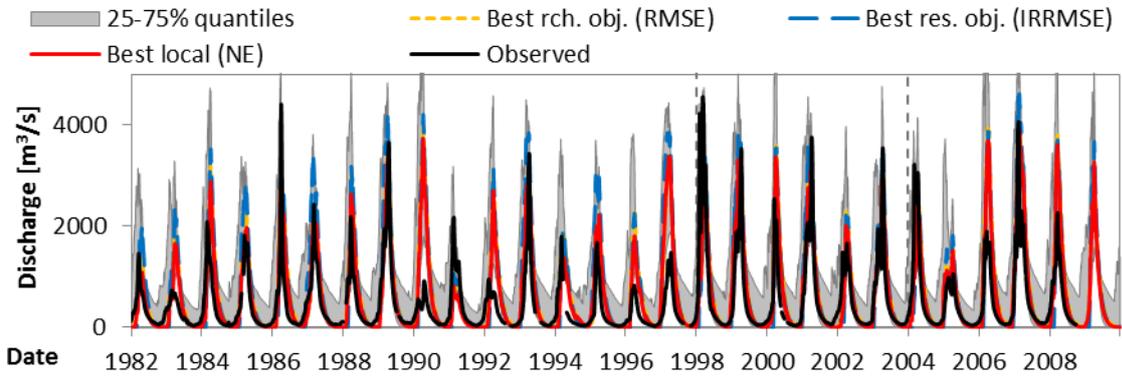
Figure AI.107. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Chavuma Falls.
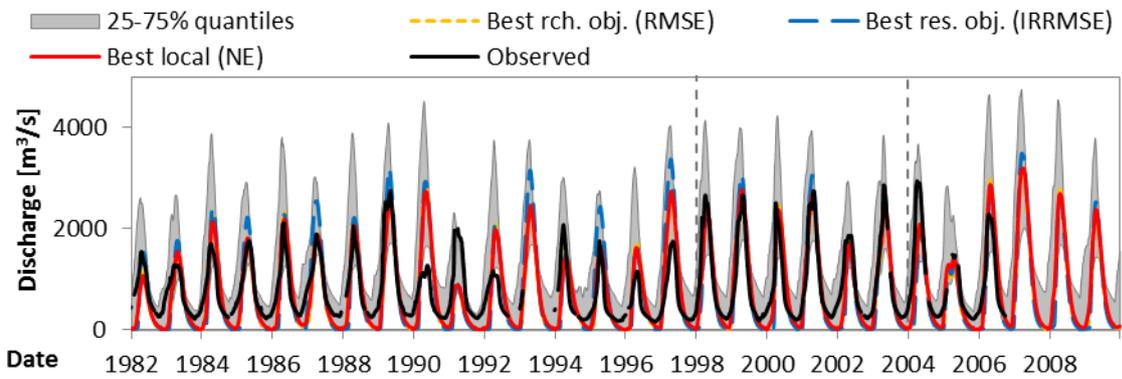
Figure AI.108. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Senanga.
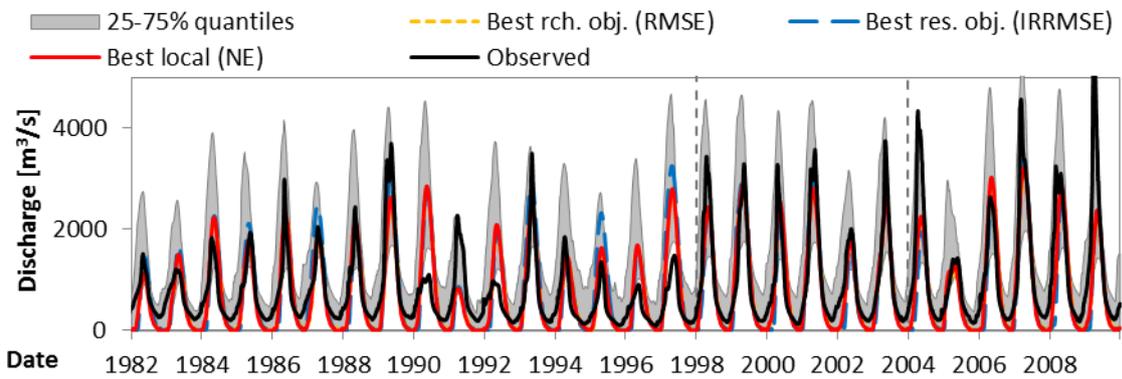
Figure AI.109. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Victoria Falls.
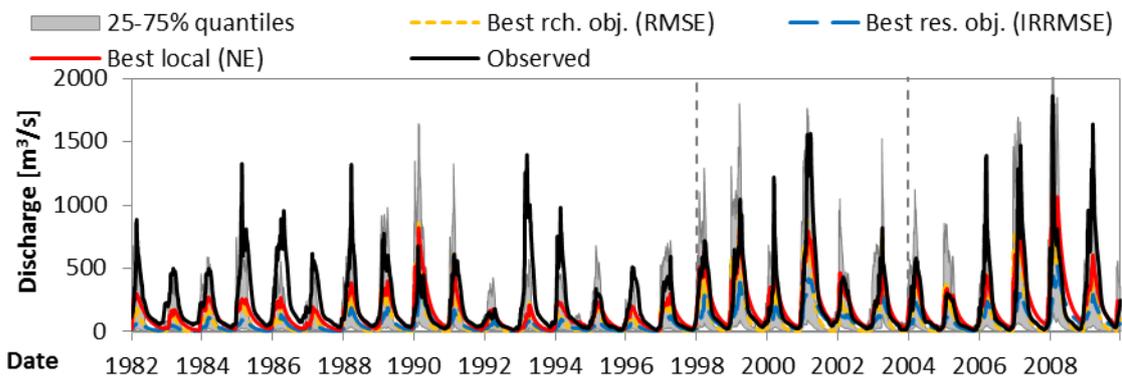
Figure AI.110. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Kafue Hook Bridge.
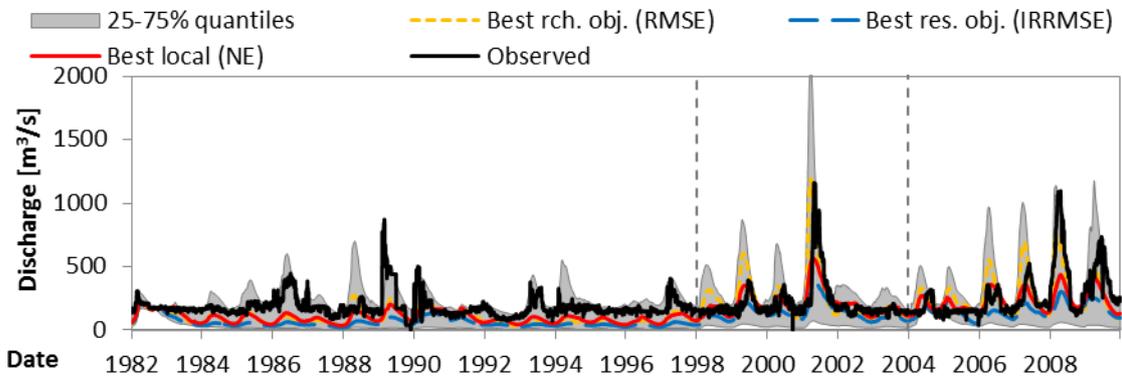
Figure AI.111. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Kafue Gorge.
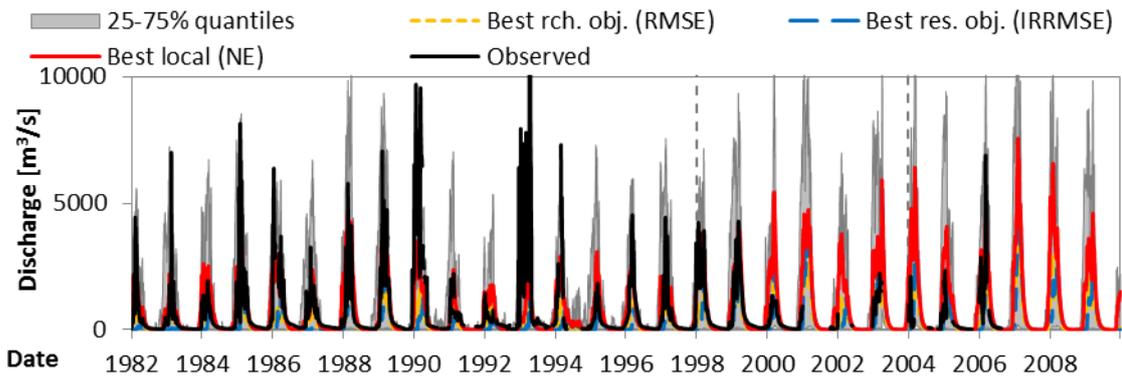
Figure AI.112. Synthesis hydrographs of the regional-regularized calibration of the parsimonious model at Great East Road Bridge.
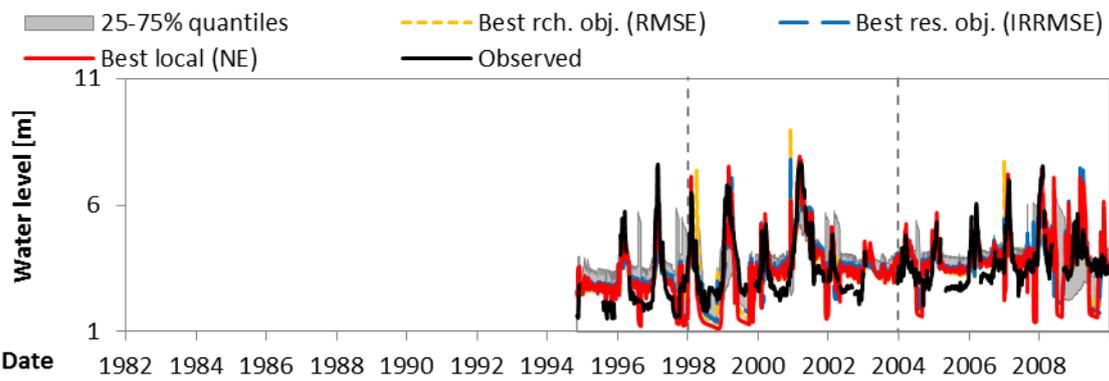
Figure AI.113. Synthesis water level series of the regional-regularized calibration of the parsimonious model at Marromeu.
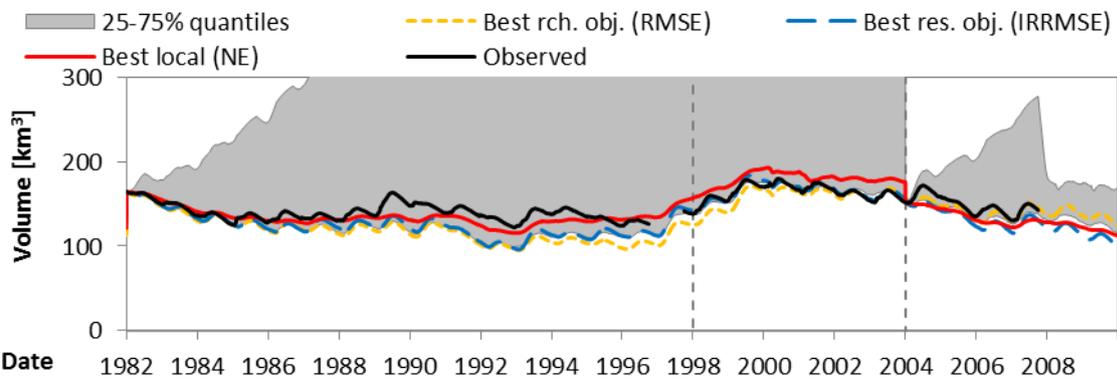
Figure AI.114. Synthesis volume series of the regional-regularized calibration of the parsimonious model at the Kariba Dam.
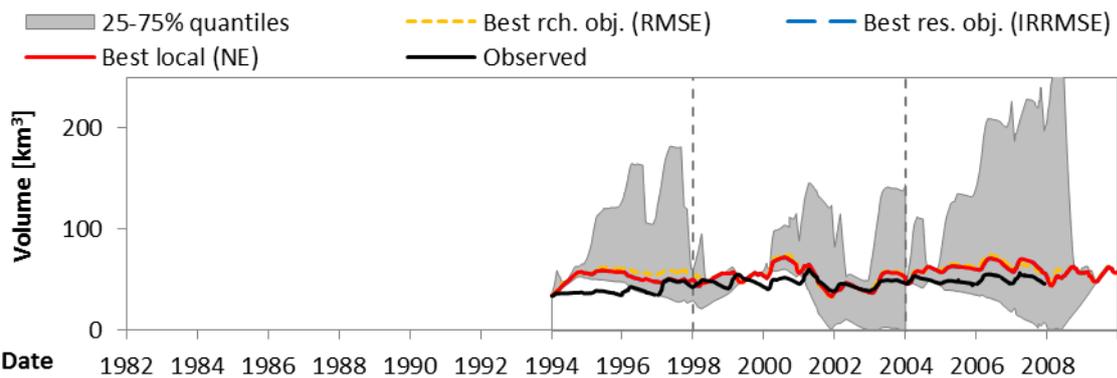
Figure AI.115. Synthesis volume series of the regional-regularized calibration of the parsimonious model at the Cahora Bassa Dam.
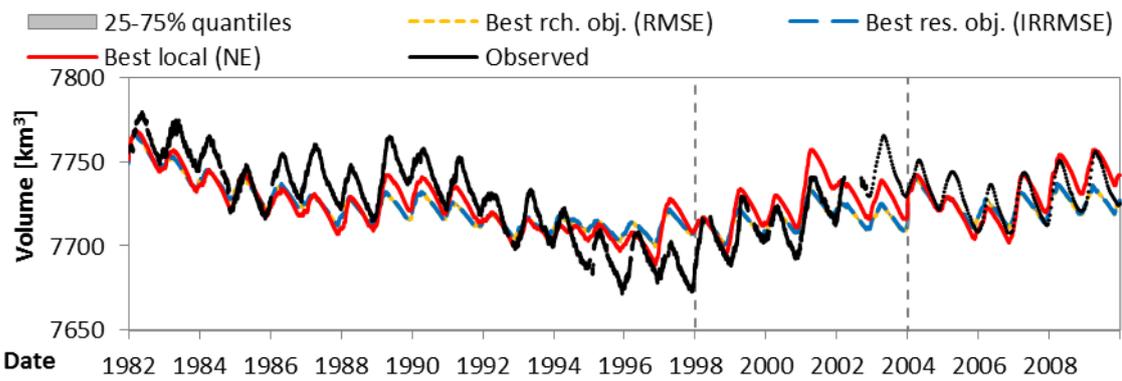
**Figure AI.116. Synthesis volume series of the regional-regularized calibration of the parsimonious model at the Malawi Lake.**

Final parameter distribution
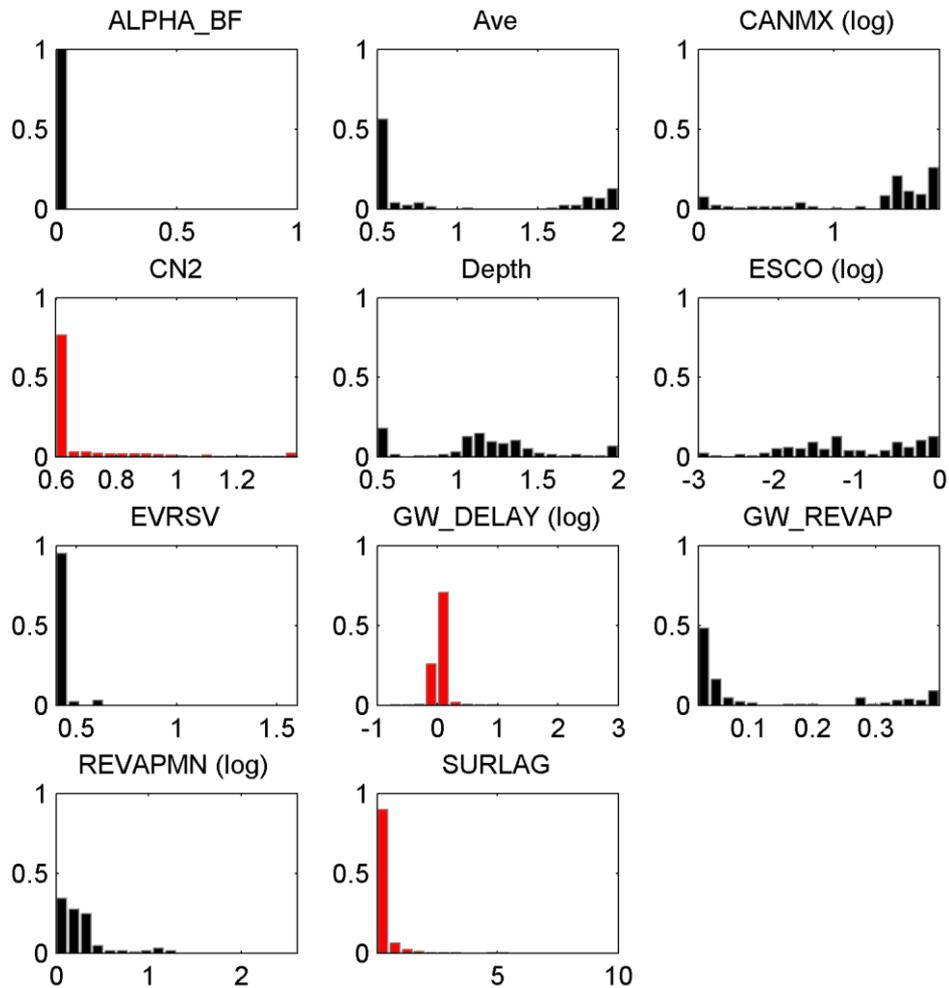
Figure AI.117. Parameter distribution among the ensemble for the last iteration of the regional-regularized calibration of the parsimonious model. Histogram normalized according to ensemble size (160). Regionalized parameters in red.

APPENDIX I.N. RESULTS FOR THE REGIONAL-REGULARIZED CALIBRATION OF THE DETAILED MODEL

Full ensemble results

The full ensemble results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Results from the 10% worst performing parameter sets regarding each calibration objective are not included.

Table AI.24. Full minimum ensemble results for the regional-regularized calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	228.2	407.8	0.050	0.090	-0.06	0.86
Watopa Pontoon	155.0	204.7	0.175	0.231	-1.18	0.04
Chiromo	189.3	246.6	0.095	0.123	-0.61	0.61
Chikwawa	130.0	158.7	0.127	0.155	-1.01	0.76
Sesheke	357.0	612.6	0.072	0.123	-0.16	0.47
Senanga	203.6	294.5	0.075	0.109	-0.24	0.54
Lukulu	185.4	285.9	0.071	0.110	-0.64	0.64
Kafue Hook Bridge	189.6	282.9	0.122	0.182	-0.75	0.03
Kalabo	44.4	81.6	0.028	0.052	-0.41	0.02
Chilenga	146.8	215.4	0.185	0.271	-0.86	0.01
Chifumpa	65.1	95.7	0.103	0.152	-0.84	0.05
Mfuwe	130.1	211.8	0.114	0.186	-2.72	0.89
Great East Rd. Bridge	396.0	943.0	0.030	0.072	0.03	0.48
Kafue Gorge	139.5	169.1	0.120	0.146	-2.44	0.11
Mangochi	70.8	90.0	0.113	0.143	-2.71	0.86
Victoria Falls	263.5	401.8	0.072	0.110	-0.10	0.55
Marromeu (level)	0.726	0.949	0.118	0.154	0.25	0.98
Tete (level)	0.531	0.722	0.086	0.116	-0.41	0.98
Itezhi-Tezhi	4.84E+09	5.00E+09	0.963	0.996	-16.39	0.02
Kafue Gorge	4.42E+08	5.15E+08	0.397	0.463	-2.93	0.00
Kariba	8.64E+09	1.04E+10	0.150	0.180	-46.77	0.39
Cahora Bassa	7.67E+09	9.86E+09	0.293	0.377	-14.08	1.01
Malawi Lake	9.72E+09	1.23E+10	0.090	0.114	-1.74	1.00

Table AI.25. Full minimum ensemble results for the regional-regularized calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	260.3	432.3	0.065	0.107	-0.74	0.97
Watopa Pontoon	198.9	263.1	0.178	0.236	-1.15	0.06
Chiromo	184.0	226.8	0.197	0.243	-6.23	0.53
Chikwawa	84.0	109.4	0.117	0.152	-13.71	0.61
Sesheke	544.2	919.5	0.096	0.162	-0.15	0.49
Senanga	265.1	418.6	0.097	0.154	-0.08	0.61
Lukulu	295.3	433.1	0.119	0.174	-0.78	0.63
Kafue Hook Bridge	193.9	304.6	0.105	0.165	-0.52	0.07
Kalabo	91.2	193.9	0.064	0.137	-0.25	0.02
Chilenga	170.4	255.8	0.185	0.278	-0.72	0.04
Chifumpa	72.9	102.3	0.126	0.177	-0.82	0.05
Mfuwe	219.4	268.4	0.258	0.315	-15.38	2.01
Great East Rd. Bridge	339.2	626.7	0.049	0.091	-0.34	0.93
Kafue Gorge	167.1	222.7	0.158	0.210	-1.43	0.18
Mangochi	35.1	44.9	0.118	0.151	0.49	0.94
Victoria Falls	330.4	609.1	0.060	0.111	-0.07	0.52
Marromeu (level)	0.546	0.758	0.099	0.137	-0.24	0.93
Tete (level)	0.322	0.585	0.106	0.193	-5.25	0.98
Itezhi-Tezhi	4.95E+09	5.09E+09	1.343	1.380	-38.27	0.05
Kafue Gorge	3.97E+08	4.63E+08	0.354	0.413	-6.56	0.03
Kariba	2.43E+09	3.16E+09	0.058	0.076	-9.57	0.82
Cahora Bassa	1.08E+10	1.21E+10	0.875	0.980	-103.84	1.21
Malawi Lake	5.85E+09	7.13E+09	0.122	0.148	-1.89	1.00

Table AI.26. Full median ensemble results for the regional-regularized calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	307.6	479.0	0.068	0.106	0.49	1.30
Watopa Pontoon	177.8	234.5	0.201	0.265	-1.09	0.05
Chiromo	207.0	268.4	0.104	0.134	0.08	0.91
Chikwawa	139.6	173.0	0.136	0.169	0.19	1.12
Sesheke	476.1	744.7	0.096	0.150	0.41	0.71
Senanga	318.0	415.6	0.118	0.154	0.56	0.82
Lukulu	294.7	399.0	0.113	0.153	0.47	1.03
Kafue Hook Bridge	204.4	304.5	0.132	0.196	-0.65	0.06
Kalabo	50.1	89.6	0.032	0.057	-0.35	0.06
Chilenga	152.8	223.0	0.192	0.280	-0.83	0.02
Chifumpa	69.5	100.7	0.110	0.160	-0.80	0.07
Mfuwe	245.2	317.4	0.215	0.279	-0.63	1.94
Great East Rd. Bridge	460.0	1007.6	0.035	0.077	0.36	0.95
Kafue Gorge	160.6	192.1	0.139	0.166	-2.11	0.17
Mangochi	128.7	150.7	0.205	0.240	0.09	1.29
Victoria Falls	372.0	534.0	0.102	0.146	0.48	0.88
Marromeu (level)	0.828	1.045	0.134	0.169	0.39	1.06
Tete (level)	0.594	0.804	0.096	0.130	0.11	1.01
Itezhi-Tezhi	4.89E+09	5.05E+09	0.973	1.006	-16.34	0.02
Kafue Gorge	4.68E+08	5.42E+08	0.421	0.487	-2.91	0.01
Kariba	3.65E+10	4.23E+10	0.633	0.735	-7.47	0.75
Cahora Bassa	1.66E+10	2.00E+10	0.634	0.764	-11.16	1.36
Malawi Lake	1.75E+10	2.11E+10	0.163	0.196	0.23	1.00

Table AI.27. Full median ensemble results for the regional-regularized calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	358.7	576.7	0.089	0.143	0.41	1.25
Watopa Pontoon	251.5	329.4	0.225	0.295	-1.03	0.08
Chiromo	284.0	314.9	0.305	0.338	-3.31	0.65
Chikwawa	165.9	185.5	0.231	0.258	-5.28	0.76
Sesheke	675.6	1052.3	0.119	0.186	0.33	0.70
Senanga	421.0	530.6	0.155	0.195	0.55	0.82
Lukulu	422.3	565.0	0.170	0.227	0.46	1.00
Kafue Hook Bridge	227.9	359.8	0.123	0.195	-0.33	0.11
Kalabo	106.0	226.3	0.075	0.159	-0.14	0.08
Chilenga	184.7	277.4	0.201	0.302	-0.61	0.06
Chifumpa	81.3	115.2	0.141	0.199	-0.71	0.07
Mfuwe	463.5	605.1	0.545	0.711	-8.50	3.48
Great East Rd. Bridge	546.4	878.3	0.079	0.128	0.12	1.65
Kafue Gorge	208.2	278.6	0.197	0.263	-1.12	0.27
Mangochi	40.3	46.6	0.136	0.157	0.66	0.98
Victoria Falls	566.1	832.2	0.103	0.151	0.50	0.74
Marromeu (level)	0.611	0.805	0.111	0.146	0.17	1.02
Tete (level)	0.412	0.694	0.136	0.229	-2.54	1.00
Itezhi-Tezhi	5.18E+09	5.33E+09	1.405	1.446	-37.64	0.06
Kafue Gorge	4.63E+08	5.04E+08	0.412	0.449	-4.16	0.08
Kariba	6.98E+09	7.91E+09	0.167	0.189	0.46	0.97
Cahora Bassa	2.28E+10	2.62E+10	1.848	2.123	-78.61	1.46
Malawi Lake	7.70E+09	9.78E+09	0.160	0.204	0.36	1.00

Table AI.28. Full maximum ensemble results for the regional-regularized calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	509.2	689.1	0.112	0.152	0.63	1.75
Watopa Pontoon	180.2	239.0	0.204	0.270	-0.60	0.18
Chiromo	267.9	355.1	0.134	0.178	0.22	1.18
Chikwawa	239.7	272.9	0.234	0.267	0.32	1.45
Sesheke	607.3	1046.7	0.122	0.211	0.60	1.04
Senanga	476.5	698.4	0.177	0.259	0.78	1.21
Lukulu	494.3	700.7	0.190	0.270	0.73	1.49
Kafue Hook Bridge	209.6	313.8	0.135	0.202	-0.42	0.13
Kalabo	51.1	91.7	0.032	0.058	-0.12	0.26
Chilenga	154.5	225.1	0.194	0.283	-0.70	0.06
Chifumpa	70.3	101.8	0.111	0.161	-0.62	0.15
Mfuwe	284.1	478.5	0.249	0.420	0.27	2.14
Great East Rd. Bridge	673.3	1239.0	0.051	0.095	0.44	1.05
Kafue Gorge	171.6	202.1	0.148	0.174	-1.41	0.31
Mangochi	288.2	304.8	0.459	0.486	0.68	1.68
Victoria Falls	540.5	778.8	0.148	0.213	0.71	1.28
Marromeu (level)	0.938	1.155	0.152	0.187	0.50	1.11
Tete (level)	0.707	1.012	0.114	0.163	0.28	1.04
Itezhi-Tezhi	4.89E+09	5.06E+09	0.974	1.007	-16.01	0.03
Kafue Gorge	4.69E+08	5.44E+08	0.422	0.489	-2.53	0.07
Kariba	8.92E+10	1.01E+11	1.548	1.746	0.49	1.20
Cahora Bassa	1.83E+10	2.23E+10	0.699	0.851	-1.95	1.41
Malawi Lake	3.67E+10	4.00E+10	0.340	0.371	0.74	1.00

Table AI.29. Full maximum ensemble results for the regional-regularized calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	680.1	988.9	0.169	0.246	0.67	1.90
Watopa Pontoon	256.8	338.9	0.230	0.303	-0.29	0.29
Chiromo	379.1	407.9	0.407	0.437	-1.24	0.77
Chikwawa	262.8	284.0	0.365	0.395	-1.18	0.90
Sesheke	828.9	1384.4	0.146	0.244	0.49	1.11
Senanga	584.3	825.7	0.215	0.303	0.72	1.31
Lukulu	702.5	1020.3	0.282	0.410	0.68	1.58
Kafue Hook Bridge	236.5	384.4	0.128	0.208	0.05	0.27
Kalabo	111.4	236.6	0.078	0.167	0.16	0.33
Chilenga	189.3	286.6	0.206	0.312	-0.37	0.14
Chifumpa	83.1	119.1	0.144	0.206	-0.35	0.18
Mfuwe	539.9	794.6	0.634	0.934	-0.87	3.89
Great East Rd. Bridge	746.0	1082.3	0.109	0.157	0.55	1.92
Kafue Gorge	228.6	298.8	0.216	0.282	-0.35	0.48
Mangochi	54.0	57.2	0.182	0.193	0.69	1.03
Victoria Falls	885.2	1223.0	0.161	0.222	0.73	1.33
Marromeu (level)	0.808	0.983	0.146	0.178	0.26	1.11
Tete (level)	0.502	0.923	0.166	0.305	-1.51	1.03
Itezhi-Tezhi	5.24E+09	5.38E+09	1.419	1.458	-34.21	0.10
Kafue Gorge	5.33E+08	6.10E+08	0.475	0.543	-3.37	0.65
Kariba	2.91E+10	3.49E+10	0.696	0.835	0.91	1.19
Cahora Bassa	2.71E+10	3.01E+10	2.192	2.437	-15.96	1.54
Malawi Lake	1.88E+10	2.07E+10	0.391	0.431	0.66	1.00

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Their interpretation is introduced in the Section 9.3.2 of the main text.
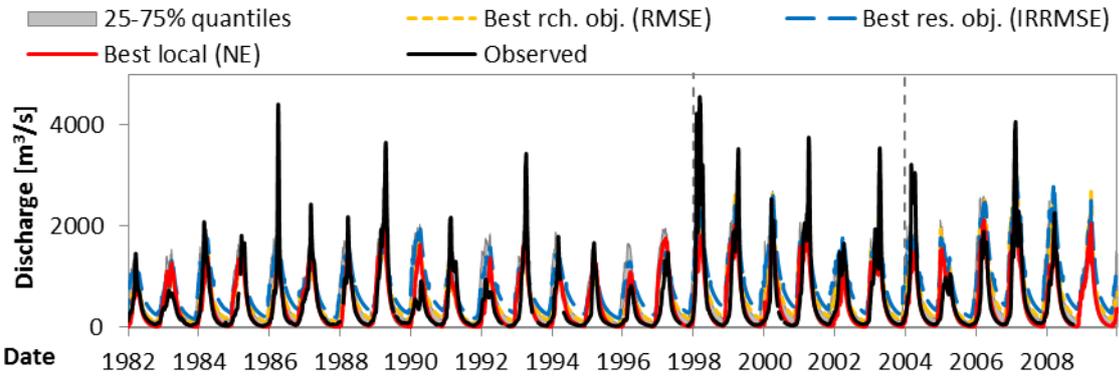

**Figure AI.118. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Chavuma Falls.**
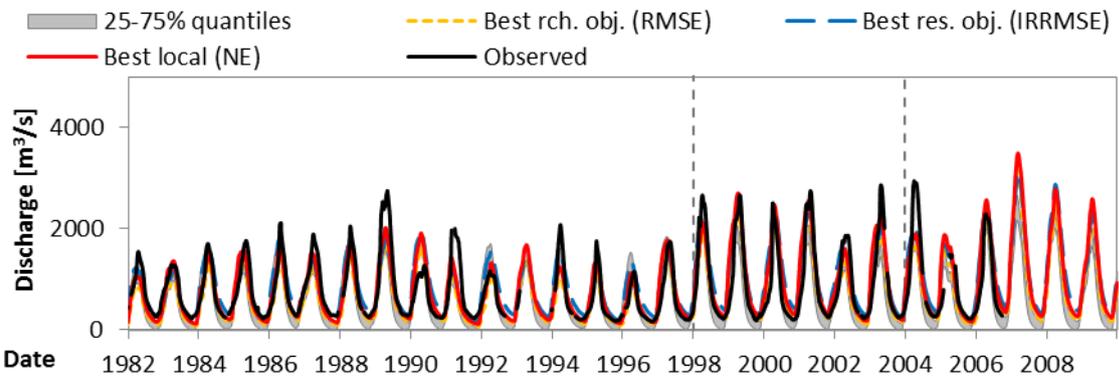
Figure AI.119. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Senanga.
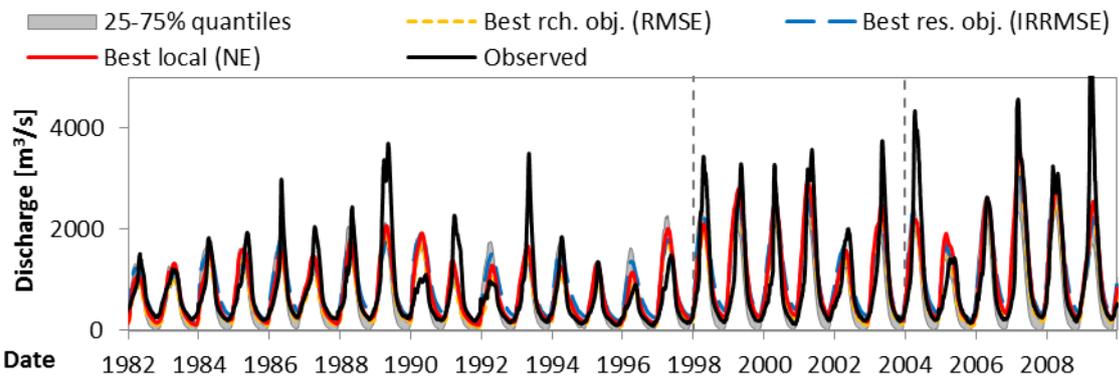

Figure AI.120. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Victoria Falls.
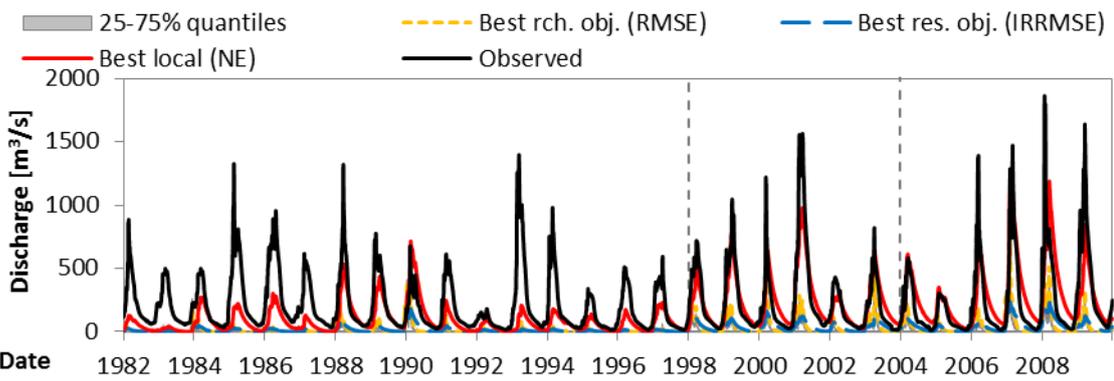
Figure AI.121. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Kafue Hook Bridge.
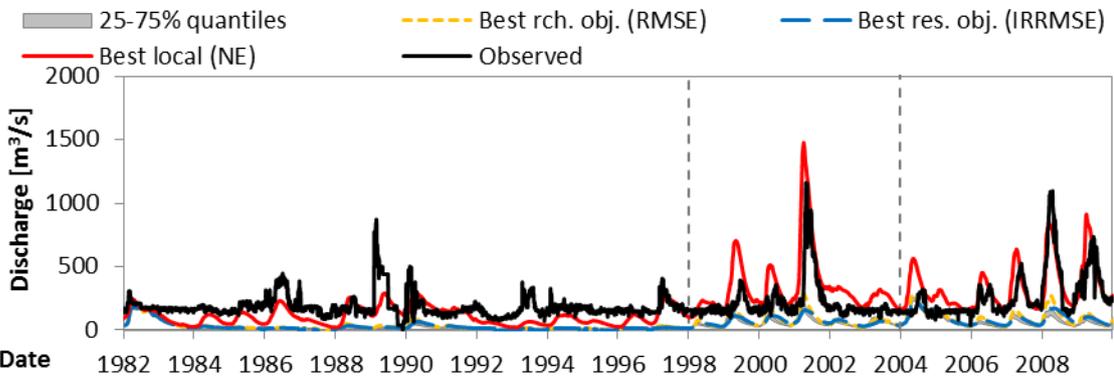

Figure AI.122. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Kafue Gorge.
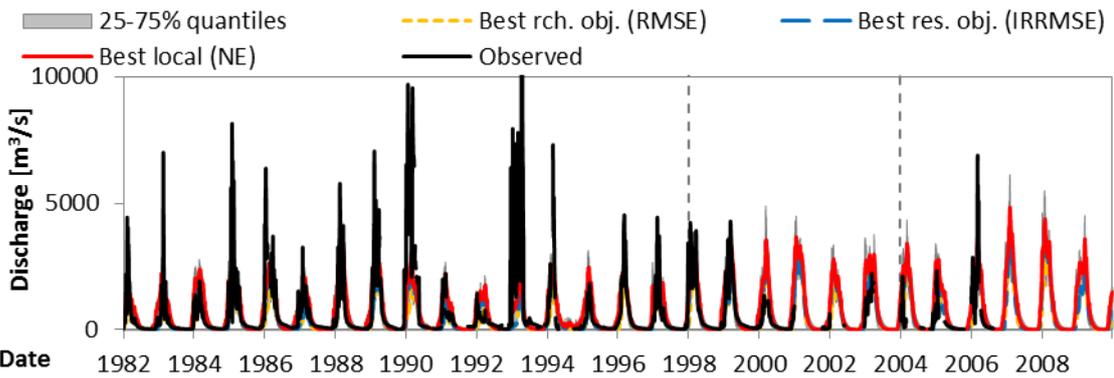
Figure AI.123. Synthesis hydrographs of the regional-regularized calibration of the detailed model at Great East Road Bridge.
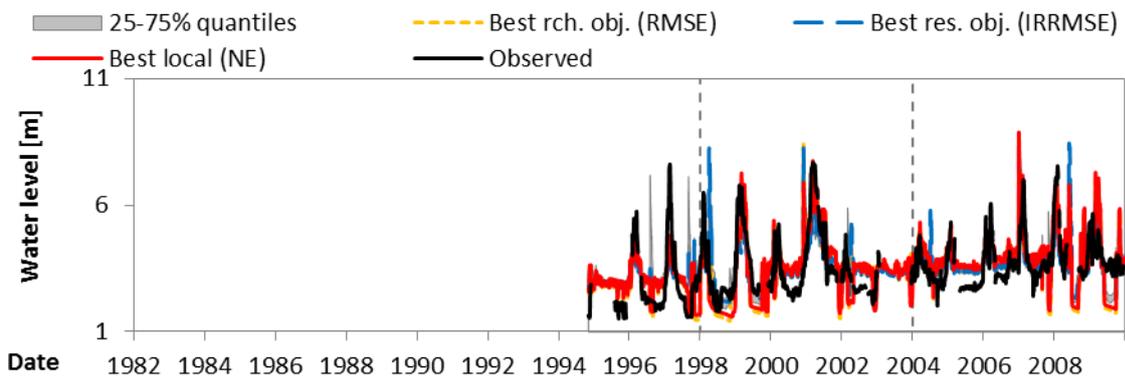
Figure AI.124. Synthesis water level series of the regional-regularized calibration of the detailed model at Marromeu.
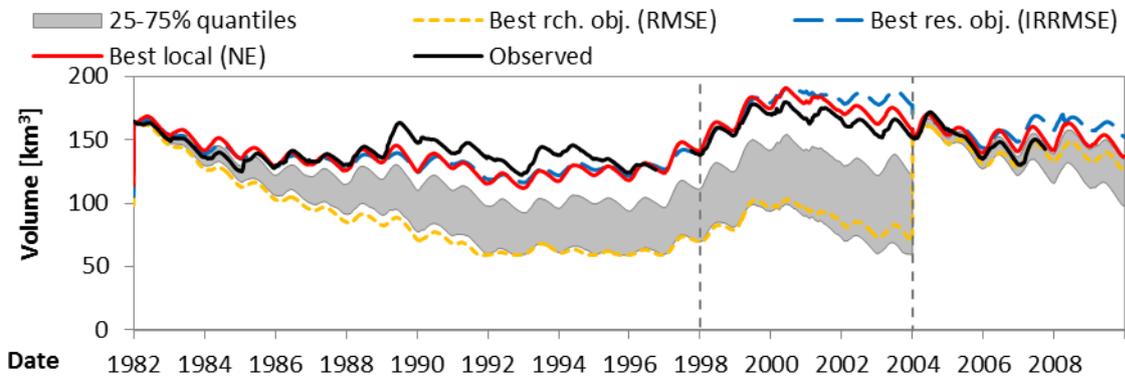
Figure AI.125. Synthesis volume series of the regional-regularized calibration of the detailed model at the Kariba Dam.
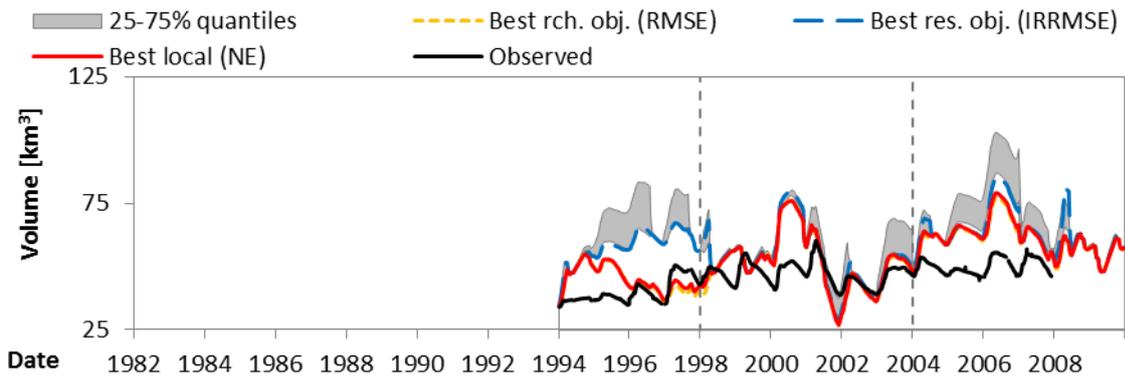
Figure AI.126. Synthesis volume series of the regional-regularized calibration of the detailed model at the Cahora Bassa Dam.
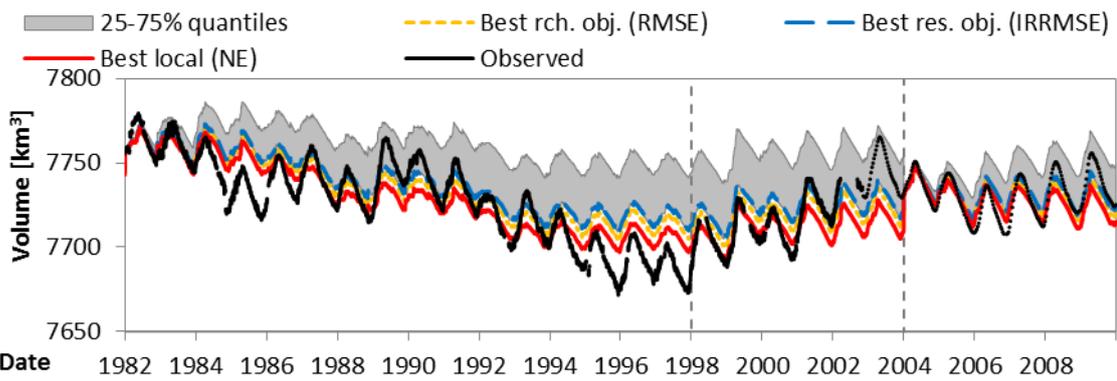
**Figure AI.127. Synthesis volume series of the regional-regularized calibration of the detailed model at the Malawi Lake.**

Final parameter distribution
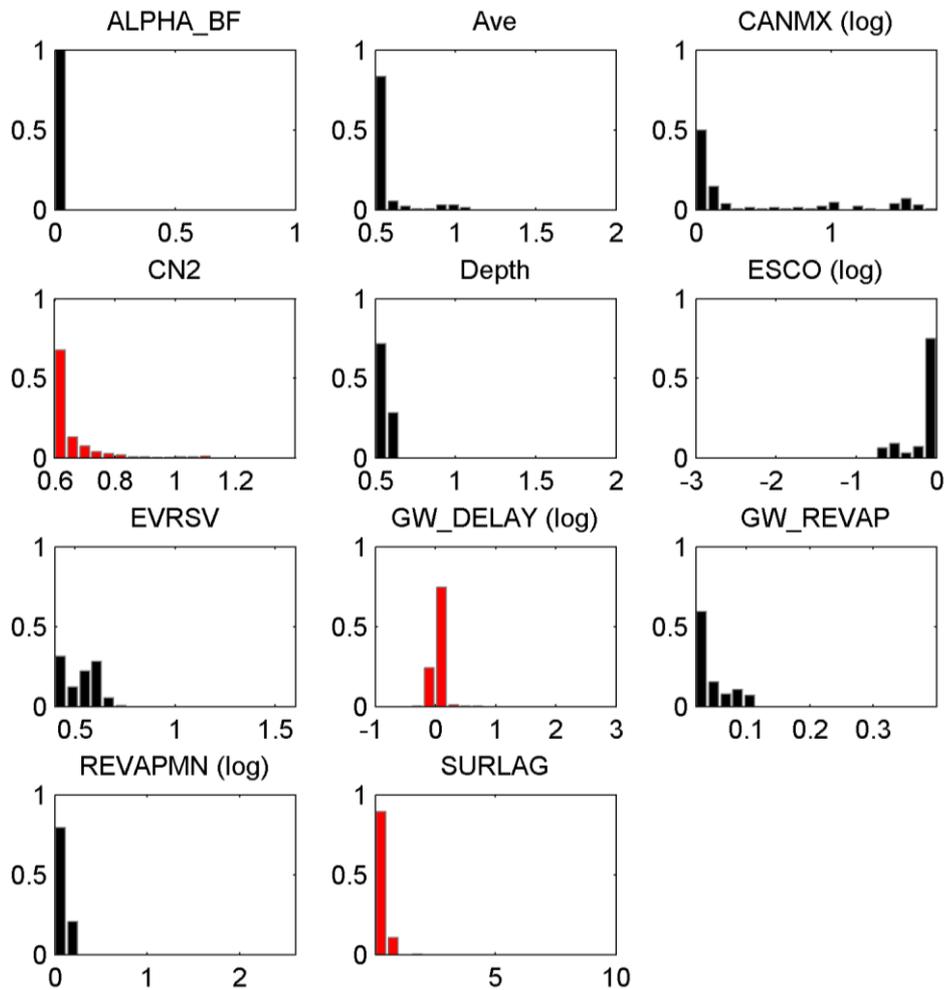
Figure AI.128. Parameter distribution among the ensemble for the last iteration of the regional-regularized calibration of the detailed model. Histogram normalized according to ensemble size (160). Regionalized parameters in red.

APPENDIX I.O. RESULTS FOR THE CASCADING-REGULARIZED CALIBRATION OF THE PARSIMONIOUS MODEL

Full ensemble results

The full ensemble results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Results from the 10% worst performing parameter sets regarding each calibration objective are not included.

Table AI.30. Full minimum ensemble results for the cascading-regularized calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	230.2	385.6	0.051	0.085	-0.89	0.02
Watopa Pontoon	134.0	181.7	0.151	0.205	-41.90	0.07
Chiromo	253.2	310.7	0.127	0.156	-1.00	1.29
Chikwawa	236.7	284.4	0.231	0.278	-2.22	1.45
Sesheke	287.5	498.8	0.058	0.100	-0.21	0.25
Senanga	187.0	289.0	0.069	0.107	-1.16	0.29
Lukulu	211.3	327.9	0.081	0.126	-2.38	0.11
Kafue Hook Bridge	145.7	237.9	0.094	0.153	-15.47	0.25
Kalabo	39.0	80.6	0.025	0.051	-20.95	0.07
Chilenga	126.2	190.3	0.159	0.239	-15.41	0.11
Chifumpa	29.8	58.8	0.047	0.093	-46.16	0.81
Mfuwe	95.6	157.3	0.084	0.138	-10.84	0.00
Great East Rd. Bridge	411.5	937.6	0.031	0.072	-2.12	0.49
Kafue Gorge	66.3	123.6	0.057	0.107	-2.00	0.44
Mangochi	144.2	184.0	0.230	0.293	-0.96	1.27
Victoria Falls	208.7	342.0	0.057	0.094	-2.37	0.28
Marromeu (level)	0.530	0.665	0.086	0.108	0.26	0.87
Tete (level)	0.502	0.639	0.081	0.103	-1.43	0.88
Itezhi-Tezhi	3.73E+09	4.03E+09	0.743	0.802	-8463.07	0.09
Kafue Gorge	5.37E+08	6.86E+08	0.483	0.616	-1271.61	0.48
Kariba	9.08E+09	1.11E+10	0.158	0.194	-118.91	0.18
Cahora Bassa	7.55E+09	9.46E+09	0.288	0.361	-48.04	0.17
Malawi Lake	1.92E+10	2.48E+10	0.178	0.230	-0.53	1.00

Table AI.31. Full minimum ensemble results for the cascading-regularized calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	237.9	410.1	0.059	0.102	-1.14	0.13
Watopa Pontoon	201.3	268.8	0.180	0.241	-49.69	0.13
Chiromo	140.8	172.3	0.151	0.185	-5.48	1.04
Chikwawa	143.2	181.3	0.199	0.252	-21.34	1.15
Sesheke	404.3	651.7	0.071	0.115	-0.14	0.29
Senanga	207.5	306.5	0.076	0.113	-0.50	0.36
Lukulu	289.4	430.2	0.116	0.173	-1.26	0.11
Kafue Hook Bridge	163.4	272.0	0.088	0.147	-15.09	0.35
Kalabo	96.7	204.4	0.068	0.144	-2.99	0.05
Chilenga	153.5	236.5	0.167	0.257	-13.46	0.22
Chifumpa	32.8	65.1	0.057	0.113	-49.87	0.73
Mfuwe	73.0	111.0	0.086	0.130	-43.08	0.01
Great East Rd. Bridge	536.5	766.6	0.078	0.112	-10.46	1.12
Kafue Gorge	100.4	144.6	0.095	0.137	-0.57	0.77
Mangochi	44.5	52.5	0.150	0.177	0.30	0.94
Victoria Falls	314.7	571.3	0.057	0.104	-3.02	0.39
Marromeu (level)	0.560	0.712	0.101	0.129	-1.29	0.78
Tete (level)	0.357	0.534	0.118	0.176	-14.31	0.83
Itezhi-Tezhi	2.11E+09	3.02E+09	0.572	0.818	-4780.56	0.16
Kafue Gorge	1.10E+09	1.33E+09	0.980	1.184	-2409.86	2.66
Kariba	2.93E+09	3.44E+09	0.070	0.082	-39.64	0.73
Cahora Bassa	9.41E+09	1.14E+10	0.762	0.923	-205.95	0.22
Malawi Lake	1.53E+10	1.98E+10	0.319	0.411	-4.45	1.00

Table AI.32. Full median ensemble results for the cascading-regularized calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	408.0	660.2	0.090	0.146	0.03	1.30
Watopa Pontoon	169.2	224.7	0.191	0.254	-0.92	0.14
Chiromo	260.1	316.2	0.130	0.158	-0.28	1.31
Chikwawa	244.2	292.1	0.239	0.285	-1.30	1.47
Sesheke	408.2	647.3	0.082	0.130	0.56	0.83
Senanga	309.4	459.8	0.115	0.170	0.46	0.97
Lukulu	369.1	559.3	0.142	0.215	-0.05	1.10
Kafue Hook Bridge	153.1	250.8	0.099	0.162	-0.12	0.41
Kalabo	47.7	87.9	0.030	0.056	-0.30	0.42
Chilenga	134.8	203.6	0.169	0.256	-0.52	0.31
Chifumpa	32.4	71.1	0.051	0.113	0.10	0.88
Mfuwe	194.1	317.6	0.170	0.279	-0.64	0.61
Great East Rd. Bridge	569.0	1170.4	0.043	0.089	0.13	0.84
Kafue Gorge	101.4	138.5	0.088	0.119	-0.62	0.62
Mangochi	163.1	196.8	0.260	0.314	-0.55	1.32
Victoria Falls	327.7	499.0	0.090	0.137	0.55	0.96
Marromeu (level)	0.710	0.911	0.115	0.148	0.54	1.03
Tete (level)	0.665	0.802	0.107	0.129	0.12	1.01
Itezhi-Tezhi	4.46E+09	4.65E+09	0.888	0.926	-13.69	0.12
Kafue Gorge	6.96E+08	1.07E+09	0.625	0.962	-14.25	0.81
Kariba	3.45E+10	4.40E+10	0.599	0.764	-8.15	0.80
Cahora Bassa	2.78E+10	3.16E+10	1.061	1.208	-29.40	0.41
Malawi Lake	2.17E+10	2.70E+10	0.201	0.251	-0.25	1.00

Table AI.33. Full median ensemble results for the cascading-regularized calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	479.4	764.8	0.119	0.190	-0.04	1.31
Watopa Pontoon	248.0	324.8	0.222	0.291	-0.97	0.23
Chiromo	151.2	191.7	0.162	0.206	-0.60	1.10
Chikwawa	160.7	210.0	0.223	0.292	-7.04	1.20
Sesheke	557.8	932.7	0.098	0.164	0.48	0.76
Senanga	349.2	492.7	0.128	0.181	0.62	0.92
Lukulu	459.4	668.1	0.185	0.268	0.24	0.95
Kafue Hook Bridge	177.0	298.7	0.096	0.162	0.08	0.48
Kalabo	106.0	226.2	0.075	0.159	-0.14	0.26
Chilenga	169.1	262.6	0.184	0.286	-0.45	0.35
Chifumpa	35.7	78.2	0.062	0.135	0.21	0.80
Mfuwe	176.0	271.0	0.207	0.318	-0.90	1.30
Great East Rd. Bridge	648.7	1089.6	0.094	0.159	-0.36	1.87
Kafue Gorge	117.9	167.7	0.111	0.158	0.23	0.91
Mangochi	51.3	61.6	0.173	0.208	0.41	0.98
Victoria Falls	417.2	675.7	0.076	0.123	0.67	0.84
Marromeu (level)	0.684	0.849	0.124	0.154	0.08	0.94
Tete (level)	0.730	0.809	0.241	0.267	-3.80	0.98
Itezhi-Tezhi	4.60E+09	4.77E+09	1.248	1.294	-29.95	0.18
Kafue Gorge	2.54E+09	3.16E+09	2.262	2.816	-202.49	5.68
Kariba	7.10E+09	8.01E+09	0.170	0.191	0.44	0.97
Cahora Bassa	2.64E+10	3.19E+10	2.142	2.585	-116.95	0.51
Malawi Lake	2.05E+10	2.46E+10	0.426	0.512	-3.07	1.00

Table AI.34. Full maximum ensemble results for the cascading-regularized calibration (calibration phase: 1982-2003) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	580.8	923.0	0.128	0.204	0.67	1.63
Watopa Pontoon	447.1	1061.2	0.505	1.199	-0.26	2.25
Chiromo	290.4	395.3	0.145	0.198	-0.23	1.38
Chikwawa	278.8	345.2	0.272	0.337	-1.18	1.56
Sesheke	707.9	1066.3	0.143	0.215	0.74	1.26
Senanga	576.9	921.8	0.214	0.342	0.79	1.43
Lukulu	588.7	1004.7	0.226	0.386	0.64	1.61
Kafue Hook Bridge	458.2	963.6	0.295	0.621	0.00	2.30
Kalabo	132.4	361.6	0.084	0.229	-0.09	2.78
Chilenga	330.9	668.3	0.416	0.840	-0.33	2.25
Chifumpa	199.0	516.0	0.315	0.818	0.39	2.83
Mfuwe	414.2	854.3	0.363	0.750	0.60	2.59
Great East Rd. Bridge	1211.5	2217.5	0.093	0.169	0.44	2.50
Kafue Gorge	126.1	188.6	0.109	0.163	-0.29	1.23
Mangochi	182.3	221.3	0.290	0.353	-0.35	1.39
Victoria Falls	667.8	1364.5	0.183	0.373	0.79	1.64
Marromeu (level)	1.026	1.150	0.166	0.186	0.75	1.20
Tete (level)	1.104	1.329	0.178	0.214	0.44	1.09
Itezhi-Tezhi	9.15E+10	1.12E+11	18.213	22.218	-10.03	19.37
Kafue Gorge	5.42E+09	9.78E+09	4.868	8.788	-5.26	12.43
Kariba	1.20E+11	1.59E+11	2.082	2.767	0.41	1.69
Cahora Bassa	3.71E+10	4.02E+10	1.416	1.534	-1.72	1.40
Malawi Lake	2.40E+10	2.98E+10	0.223	0.277	-0.05	1.00

Table AI.35. Full maximum ensemble results for the cascading-regularized calibration (validation phase: 2004-2009) of the parsimonious model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	687.6	1096.1	0.171	0.272	0.70	1.59
Watopa Pontoon	695.5	1646.7	0.623	1.475	-0.35	2.43
Chiromo	276.7	386.2	0.297	0.414	-0.29	1.15
Chikwawa	257.1	350.0	0.357	0.486	-5.00	1.25
Sesheke	879.2	1375.0	0.155	0.242	0.74	1.28
Senanga	616.5	972.0	0.226	0.357	0.85	1.54
Lukulu	853.5	1150.9	0.343	0.462	0.68	1.37
Kafue Hook Bridge	590.0	1250.8	0.319	0.676	0.24	2.69
Kalabo	190.8	423.5	0.134	0.298	0.07	2.07
Chilenga	408.5	830.1	0.445	0.904	-0.17	2.29
Chifumpa	249.7	628.9	0.432	1.089	0.45	2.99
Mfuwe	698.2	1303.6	0.821	1.532	0.68	4.76
Great East Rd. Bridge	1785.9	3161.8	0.260	0.460	0.33	4.36
Kafue Gorge	126.2	240.1	0.119	0.227	0.43	1.30
Mangochi	53.2	67.3	0.179	0.227	0.57	0.99
Victoria Falls	1144.8	2368.1	0.208	0.430	0.77	1.89
Marromeu (level)	1.089	1.337	0.197	0.242	0.35	1.11
Tete (level)	1.215	1.444	0.401	0.477	-1.09	1.12
Itezhi-Tezhi	5.04E+10	5.93E+10	13.658	16.083	-11.37	10.11
Kafue Gorge	9.09E+09	1.09E+10	8.095	9.694	-34.95	18.99
Kariba	5.96E+10	6.85E+10	1.423	1.637	0.90	1.40
Cahora Bassa	3.93E+10	4.23E+10	3.186	3.423	-14.03	1.54
Malawi Lake	2.35E+10	2.85E+10	0.490	0.593	-1.62	1.00

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Their interpretation is introduced in the Section 9.3.2 of the main text.
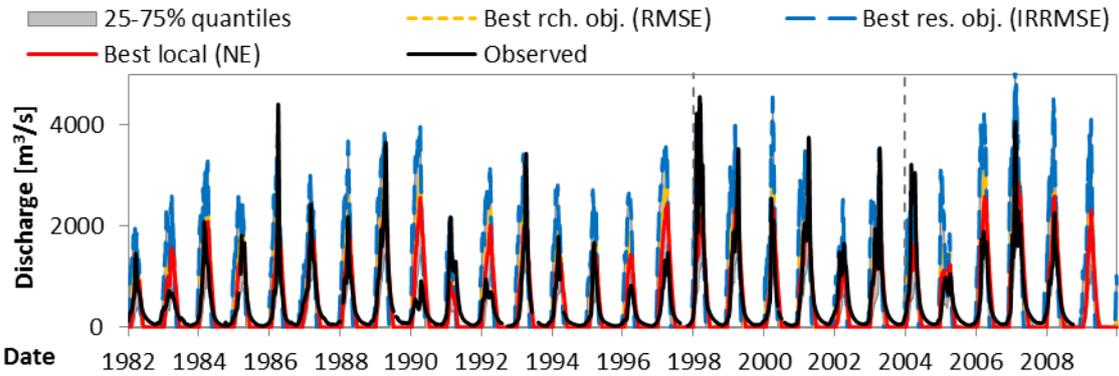

Figure AI.129. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Chavuma Falls.
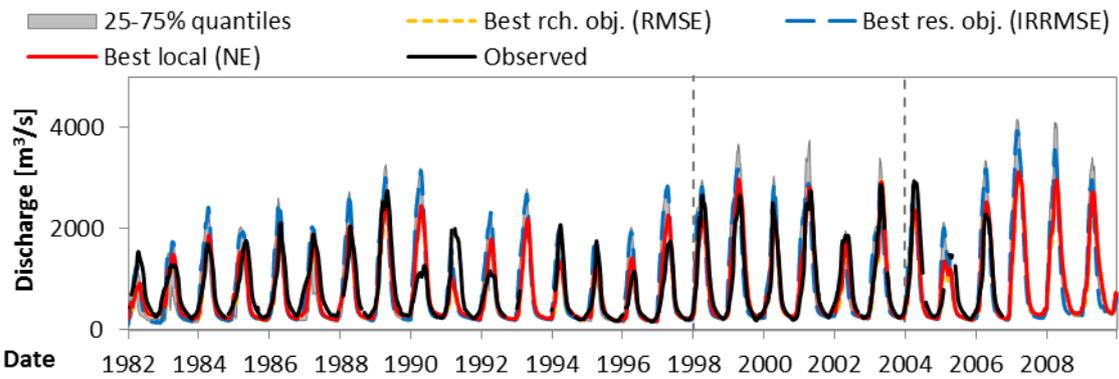

Figure AI.130. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Senanga.
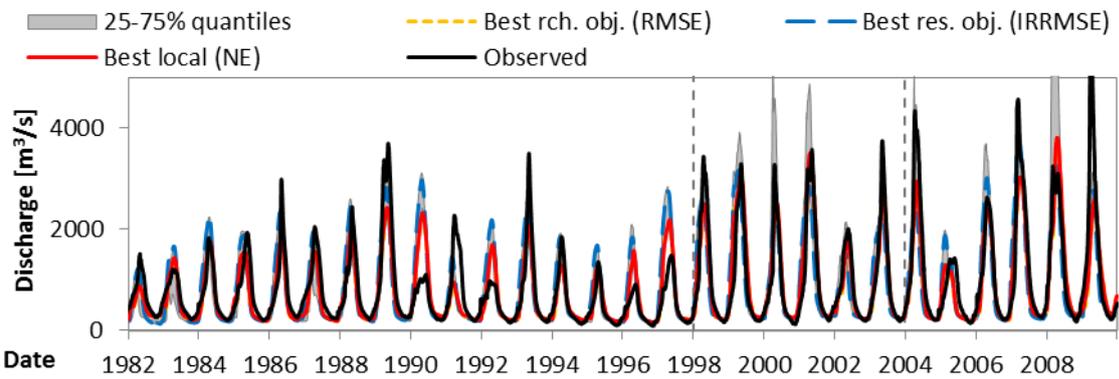

Figure AI.131. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Victoria Falls.
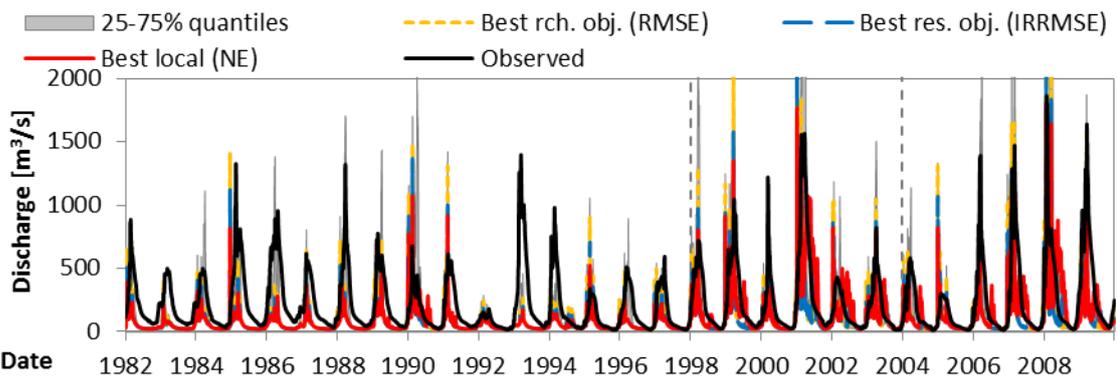
Figure AI.132. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Kafue Hook Bridge.
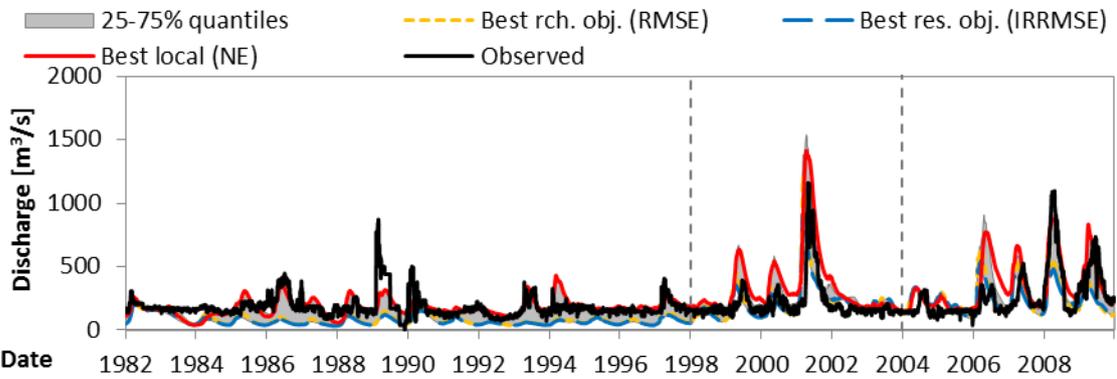
Figure AI.133. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Kafue Gorge.
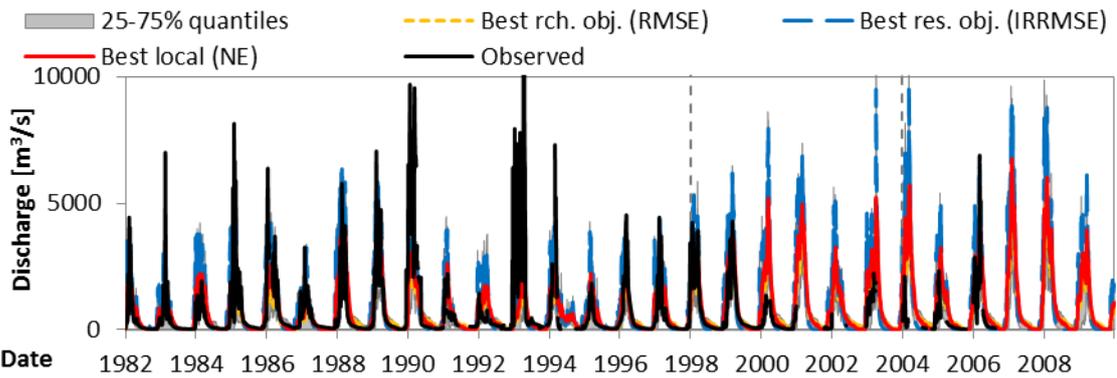

Figure AI.134. Synthesis hydrographs of the cascading-regularized calibration of the parsimonious model at Great East Road Bridge.
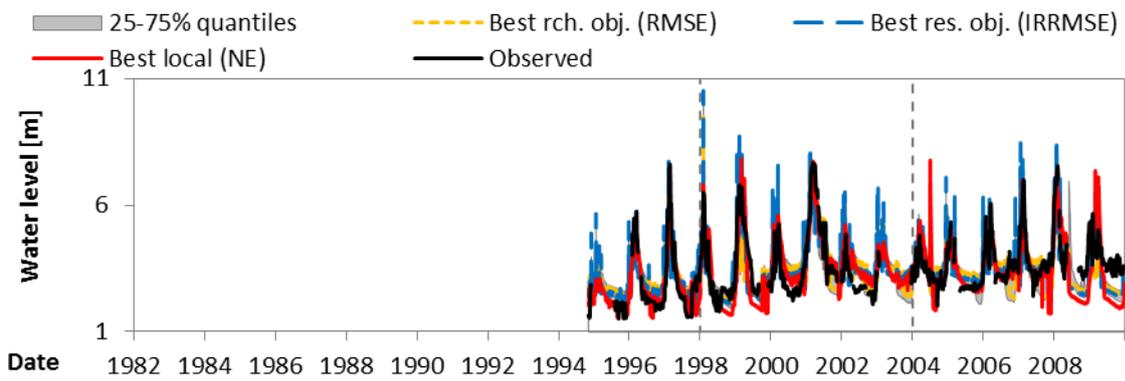
Figure AI.135. Synthesis water level series of the cascading-regularized calibration of the parsimonious model at Marromeu.
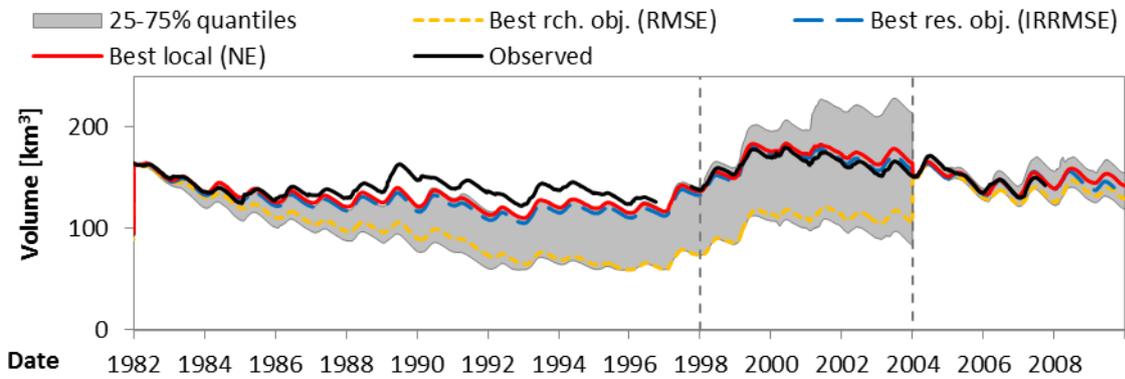

Figure AI.136. Synthesis volume series of the cascading-regularized calibration of the parsimonious model at the Kariba Dam.
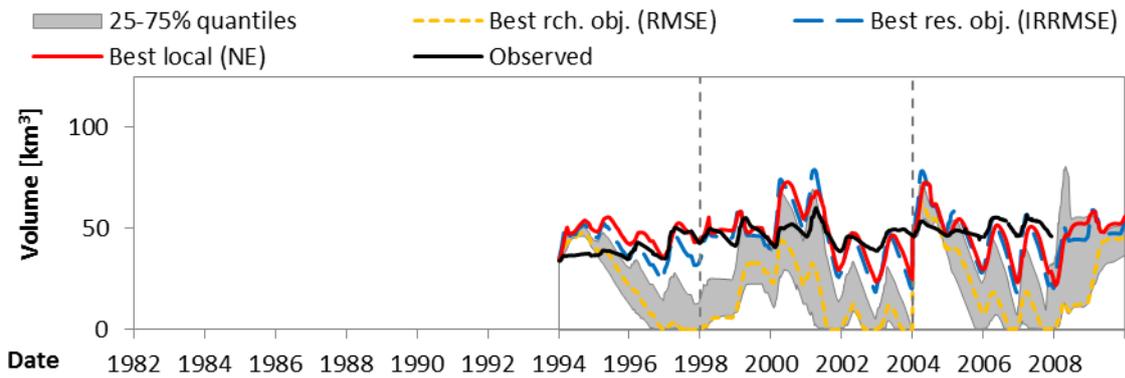

Figure AI.137. Synthesis volume series of the cascading-regularized calibration of the parsimonious model at the Cahora Bassa Dam.
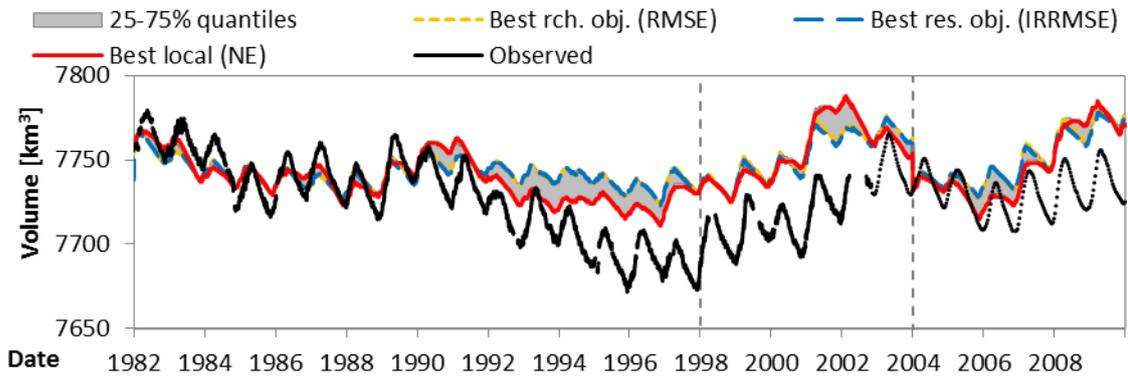
Figure AI.138. Synthesis volume series of the cascading-regularized calibration of the parsimonious model at the Malawi Lake.

APPENDIX I.P. RESULTS FOR THE CASCADING-REGULARIZED CALIBRATION OF THE DETAILED MODEL

Full ensemble results

The full ensemble results, below, refer to the minimum, median, and maximum calculated measures at each calibration location. Results from the 10% worst performing parameter sets regarding each calibration objective are not included.

Table AI.36. Full minimum ensemble results for the cascading-regularized calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	239.3	426.0	0.053	0.094	-0.09	0.53
Watopa Pontoon	141.4	190.3	0.160	0.215	-29.09	0.22
Chiromo	683.3	710.3	0.342	0.355	-7.43	2.06
Chikwawa	696.2	728.0	0.680	0.711	-17.82	2.40
Sesheke	331.0	534.3	0.067	0.108	-0.20	0.42
Senanga	235.6	343.7	0.087	0.127	-0.30	0.49
Lukulu	287.3	426.7	0.111	0.164	-1.91	0.37
Kafue Hook Bridge	135.8	232.0	0.088	0.150	-0.02	0.37
Kalabo	43.1	79.4	0.027	0.050	-11.45	0.09
Chilenga	107.4	174.0	0.135	0.219	-0.23	0.27
Chifumpa	37.3	66.8	0.059	0.106	0.20	0.63
Mfuwe	134.5	206.4	0.118	0.181	-0.11	0.35
Great East Rd. Bridge	424.4	1006.8	0.032	0.077	-0.02	0.45
Kafue Gorge	134.8	183.5	0.116	0.158	-2.51	1.57
Mangochi	605.2	631.6	0.964	1.006	-18.06	2.42
Victoria Falls	266.8	399.1	0.073	0.109	-0.15	0.48
Marromeu (level)	0.635	0.932	0.103	0.151	0.14	0.95
Tete (level)	0.533	0.768	0.086	0.124	-0.07	1.02
Itezhi-Tezhi	4.72E+09	4.88E+09	0.939	0.973	-15.44	0.05
Kafue Gorge	1.69E+10	2.36E+10	15.147	21.203	-1.08E+04	36.75
Kariba	1.94E+10	2.31E+10	0.337	0.402	-51.31	0.48
Cahora Bassa	2.21E+10	2.56E+10	0.846	0.977	-60.29	1.50
Malawi Lake	7.03E+10	7.53E+10	0.653	0.698	-10.30	1.01

Table AI.37. Full minimum ensemble results for the cascading-regularized calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	260.8	464.8	0.065	0.116	-0.23	0.52
Watopa Pontoon	200.6	269.0	0.180	0.241	-32.01	0.25
Chiromo	452.1	508.4	0.485	0.545	-18.54	1.55
Chikwawa	497.2	552.7	0.691	0.768	-68.73	1.74
Sesheke	394.7	727.4	0.070	0.128	-0.19	0.45
Senanga	216.9	314.5	0.080	0.115	-0.13	0.56
Lukulu	333.2	434.1	0.134	0.174	-0.76	0.33
Kafue Hook Bridge	148.6	258.8	0.080	0.140	0.21	0.57
Kalabo	86.5	197.4	0.061	0.139	-3.03	0.08
Chilenga	122.2	208.1	0.133	0.227	-0.07	0.35
Chifumpa	38.6	67.0	0.067	0.116	0.38	0.85
Mfuwe	130.5	164.2	0.153	0.193	-2.68	0.86
Great East Rd. Bridge	421.5	695.8	0.061	0.101	-0.27	1.05
Kafue Gorge	307.2	334.6	0.290	0.316	-2.45	2.02
Mangochi	118.1	144.4	0.398	0.487	-5.34	1.19
Victoria Falls	367.4	593.6	0.067	0.108	-0.13	0.48
Marromeu (level)	0.526	0.777	0.095	0.141	-0.16	0.90
Tete (level)	0.304	0.469	0.100	0.155	-1.29	1.02
Itezhi-Tezhi	3.92E+09	4.19E+09	1.061	1.135	-28.44	0.20
Kafue Gorge	3.21E+10	3.53E+10	28.574	31.459	-2.99E+04	64.49
Kariba	4.59E+09	5.15E+09	0.110	0.123	-15.44	0.79
Cahora Bassa	4.91E+10	5.65E+10	3.978	4.573	-608.05	1.98
Malawi Lake	5.92E+10	6.43E+10	1.233	1.338	-31.56	1.01

Table AI.38. Full median ensemble results for the cascading-regularized calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	254.3	474.1	0.056	0.105	0.50	0.89
Watopa Pontoon	147.0	198.6	0.166	0.224	-0.50	0.42
Chiromo	684.6	710.7	0.343	0.356	-5.45	2.06
Chikwawa	697.5	729.1	0.682	0.712	-13.35	2.40
Sesheke	393.3	652.7	0.079	0.131	0.55	0.72
Senanga	270.1	389.7	0.100	0.144	0.61	0.85
Lukulu	386.1	554.7	0.149	0.213	-0.03	0.85
Kafue Hook Bridge	142.7	240.0	0.092	0.155	-0.02	0.37
Kalabo	46.2	86.5	0.029	0.055	-0.26	0.13
Chilenga	115.9	182.7	0.146	0.230	-0.23	0.27
Chifumpa	37.5	67.0	0.059	0.106	0.20	0.63
Mfuwe	152.6	261.4	0.134	0.229	-0.11	0.35
Great East Rd. Bridge	560.2	1258.6	0.043	0.096	-0.01	0.45
Kafue Gorge	134.8	183.5	0.116	0.158	-1.84	1.57
Mangochi	606.1	632.5	0.966	1.008	-14.97	2.43
Victoria Falls	297.7	453.7	0.081	0.124	0.63	0.83
Marromeu (level)	0.861	1.130	0.139	0.183	0.29	1.05
Tete (level)	0.615	0.844	0.099	0.136	0.02	1.11
Itezhi-Tezhi	4.75E+09	4.92E+09	0.946	0.979	-15.44	0.05
Kafue Gorge	1.69E+10	2.36E+10	15.149	21.206	-7409.12	36.75
Kariba	4.61E+10	5.16E+10	0.801	0.896	-11.58	0.69
Cahora Bassa	2.96E+10	3.43E+10	1.129	1.308	-34.63	1.66
Malawi Lake	7.05E+10	7.54E+10	0.654	0.699	-8.74	1.01

Table AI.39. Full median ensemble results for the cascading-regularized calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	296.6	527.0	0.074	0.131	0.51	0.87
Watopa Pontoon	206.1	281.2	0.185	0.252	-0.48	0.60
Chiromo	453.8	509.2	0.487	0.546	-10.26	1.55
Chikwawa	498.3	553.1	0.693	0.769	-54.79	1.74
Sesheke	493.9	896.0	0.087	0.158	0.52	0.69
Senanga	270.9	397.9	0.099	0.146	0.75	0.86
Lukulu	497.6	657.6	0.200	0.264	0.26	0.78
Kafue Hook Bridge	152.2	276.9	0.082	0.150	0.21	0.57
Kalabo	108.1	231.1	0.076	0.163	-0.19	0.21
Chilenga	133.3	226.0	0.145	0.246	-0.07	0.35
Chifumpa	38.6	67.3	0.067	0.116	0.42	0.85
Mfuwe	130.5	168.6	0.153	0.198	0.26	0.86
Great East Rd. Bridge	627.9	919.6	0.091	0.134	0.03	1.08
Kafue Gorge	307.3	334.6	0.290	0.316	-2.05	2.02
Mangochi	120.3	146.6	0.406	0.494	-2.34	1.20
Victoria Falls	446.0	771.3	0.081	0.140	0.57	0.75
Marromeu (level)	0.572	0.852	0.104	0.154	0.07	0.98
Tete (level)	0.341	0.483	0.113	0.160	-0.71	1.06
Itezhi-Tezhi	4.44E+09	4.66E+09	1.204	1.262	-28.44	0.20
Kafue Gorge	3.21E+10	3.53E+10	28.577	31.463	-2.54E+04	64.50
Kariba	9.06E+09	1.23E+10	0.217	0.294	-0.31	0.99
Cahora Bassa	5.06E+10	5.80E+10	4.099	4.697	-388.56	2.01
Malawi Lake	5.95E+10	6.45E+10	1.238	1.342	-26.93	1.01

Table AI.40. Full maximum ensemble results for the cascading-regularized calibration (calibration phase: 1982-2003) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	398.5	702.0	0.088	0.155	0.60	1.16
Watopa Pontoon	401.7	888.8	0.454	1.004	-0.38	2.23
Chiromo	746.4	812.3	0.374	0.407	-5.45	2.18
Chikwawa	777.0	834.9	0.759	0.816	-13.31	2.56
Sesheke	612.4	1063.0	0.123	0.214	0.70	1.02
Senanga	483.5	715.4	0.179	0.265	0.70	1.20
Lukulu	564.6	933.1	0.217	0.359	0.39	1.28
Kafue Hook Bridge	142.7	240.1	0.092	0.155	0.04	0.43
Kalabo	107.7	272.4	0.068	0.172	-0.06	2.49
Chilenga	115.9	182.7	0.146	0.230	-0.11	0.34
Chifumpa	38.5	67.3	0.061	0.107	0.21	0.69
Mfuwe	160.2	261.4	0.141	0.229	0.31	0.82
Great East Rd. Bridge	599.2	1266.2	0.046	0.097	0.36	0.95
Kafue Gorge	149.8	204.2	0.129	0.176	-1.84	1.66
Mangochi	668.7	691.0	1.066	1.101	-14.93	2.57
Victoria Falls	493.0	796.6	0.135	0.218	0.71	1.22
Marromeu (level)	1.038	1.241	0.168	0.201	0.52	1.10
Tete (level)	0.666	0.884	0.107	0.142	0.19	1.14
Itezhi-Tezhi	4.75E+09	4.92E+09	0.946	0.979	-15.22	0.05
Kafue Gorge	2.12E+10	2.85E+10	19.024	25.573	-7406.99	45.90
Kariba	7.66E+10	1.05E+11	1.331	1.827	-1.53	1.51
Cahora Bassa	3.78E+10	4.49E+10	1.442	1.716	-18.89	1.85
Malawi Lake	7.72E+10	8.12E+10	0.717	0.753	-8.71	1.01

Table AI.41. Full maximum ensemble results for the cascading-regularized calibration (validation phase: 2004-2009) of the detailed model.

	MAE	RMSE	AIRAD	IRRMSE	NE	VR
Chavuma Falls	470.5	832.0	0.117	0.207	0.62	1.13
Watopa Pontoon	632.8	1328.8	0.567	1.190	-0.35	2.46
Chiromo	523.1	670.8	0.561	0.719	-10.23	1.64
Chikwawa	511.3	618.4	0.711	0.859	-54.71	1.76
Sesheke	808.9	1406.0	0.143	0.248	0.68	1.04
Senanga	545.9	843.1	0.200	0.310	0.84	1.23
Lukulu	664.2	1014.5	0.267	0.408	0.68	1.21
Kafue Hook Bridge	155.1	276.9	0.084	0.150	0.31	0.69
Kalabo	188.4	425.8	0.133	0.300	0.13	2.46
Chilenga	133.7	226.0	0.145	0.246	0.09	0.48
Chifumpa	42.3	69.6	0.073	0.121	0.42	0.90
Mfuwe	223.0	376.8	0.262	0.443	0.30	1.68
Great East Rd. Bridge	689.2	1053.3	0.100	0.153	0.44	2.09
Kafue Gorge	332.1	355.5	0.314	0.336	-2.05	2.14
Mangochi	182.9	202.2	0.617	0.682	-2.23	1.34
Victoria Falls	787.0	1254.7	0.143	0.228	0.75	1.29
Marromeu (level)	0.696	0.949	0.126	0.172	0.23	1.03
Tete (level)	0.403	0.559	0.133	0.184	-0.62	1.09
Itezhi-Tezhi	4.44E+09	4.66E+09	1.204	1.262	-22.83	0.29
Kafue Gorge	3.45E+10	3.83E+10	30.749	34.120	-2.54E+04	69.33
Kariba	3.66E+10	4.36E+10	0.874	1.041	0.77	1.25
Cahora Bassa	6.31E+10	7.25E+10	5.113	5.873	-368.23	2.26
Malawi Lake	6.53E+10	6.96E+10	1.359	1.448	-26.78	1.01

Synthesis hydrographs

Below, synthesis hydrographs are presented for selected calibration locations throughout the basin. Their interpretation is introduced in the Section 9.3.2 of the main text.
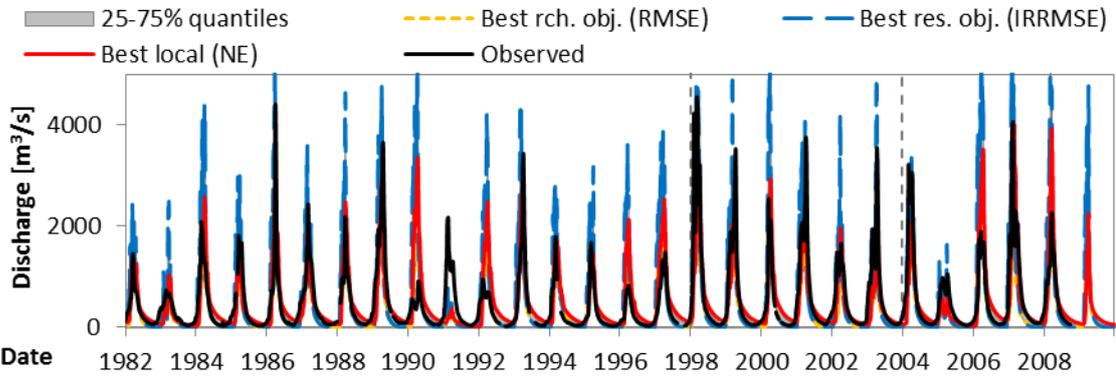

Figure AI.139. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Chavuma Falls.
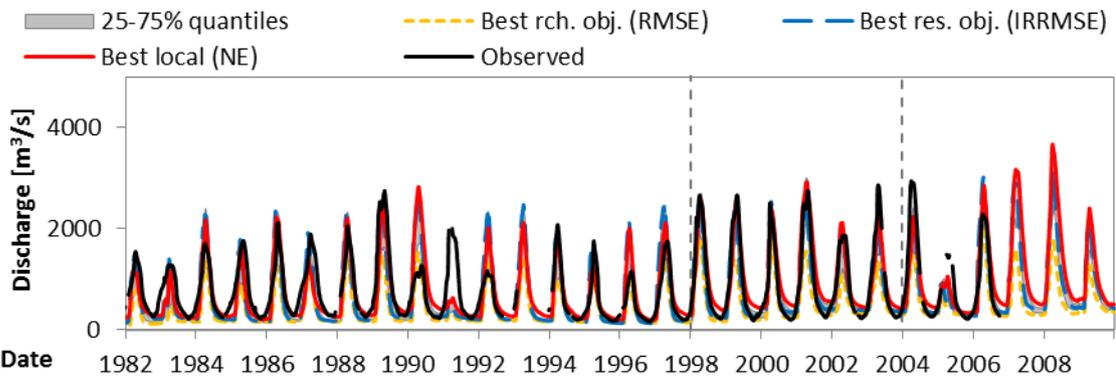

Figure AI.140. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Senanga.
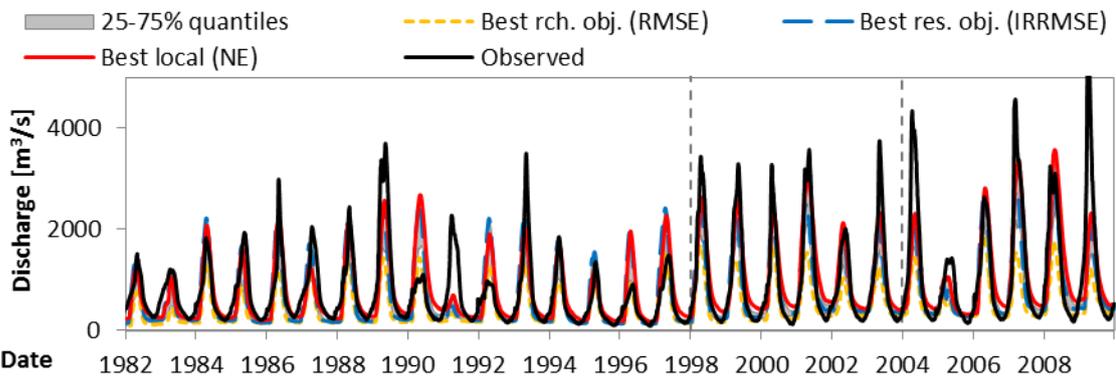

Figure AI.141. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Victoria Falls.
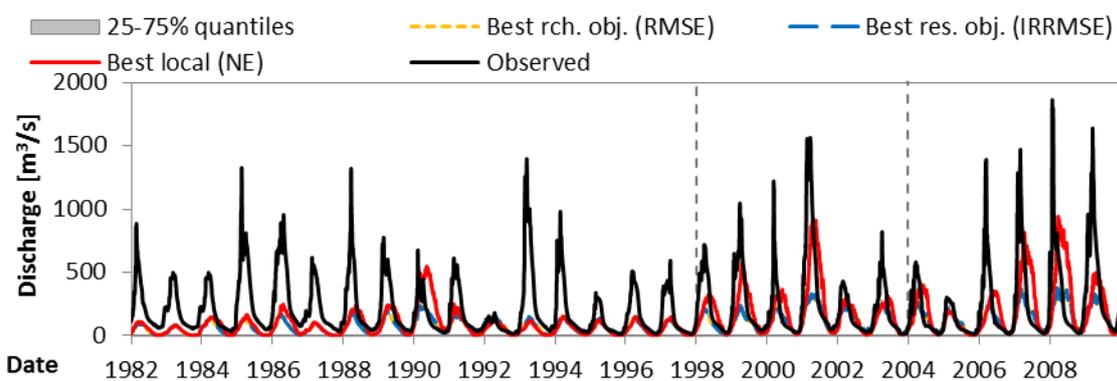
Figure AI.142. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Kafue Hook Bridge.
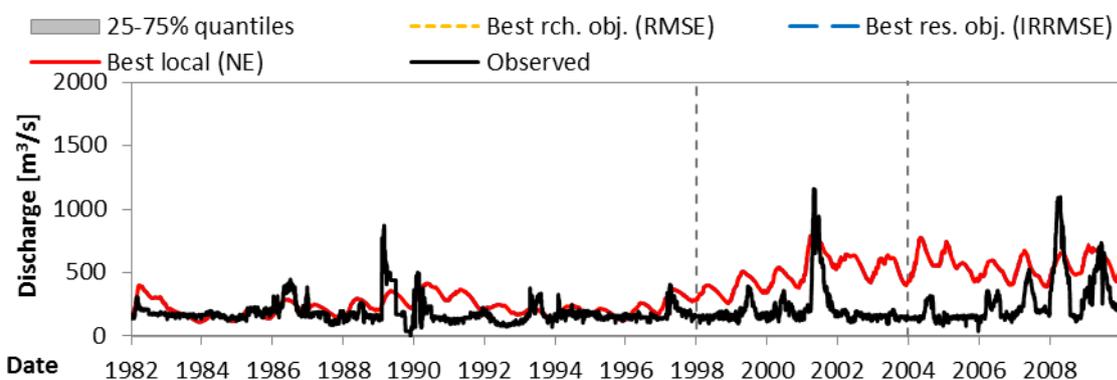
Figure AI.143. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Kafue Gorge.
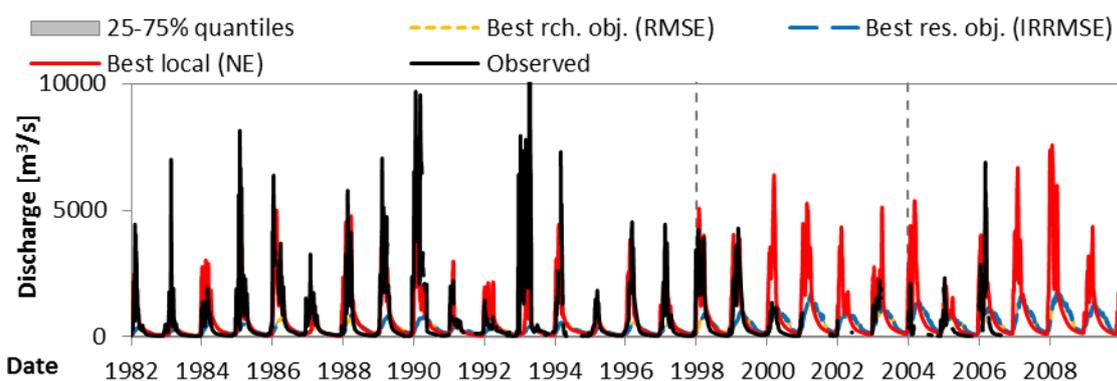
Figure AI.144. Synthesis hydrographs of the cascading-regularized calibration of the detailed model at Great East Road Bridge.
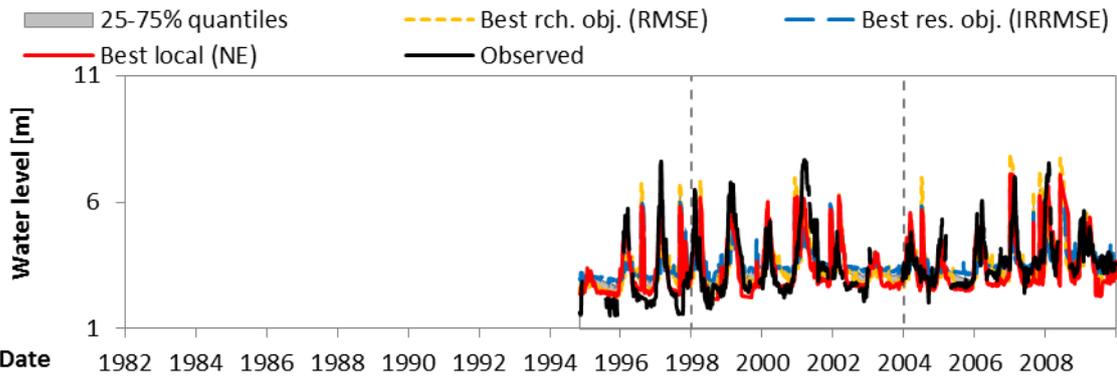
Figure AI.145. Synthesis water level series of the cascading-regularized calibration of the detailed model at Marromeu.
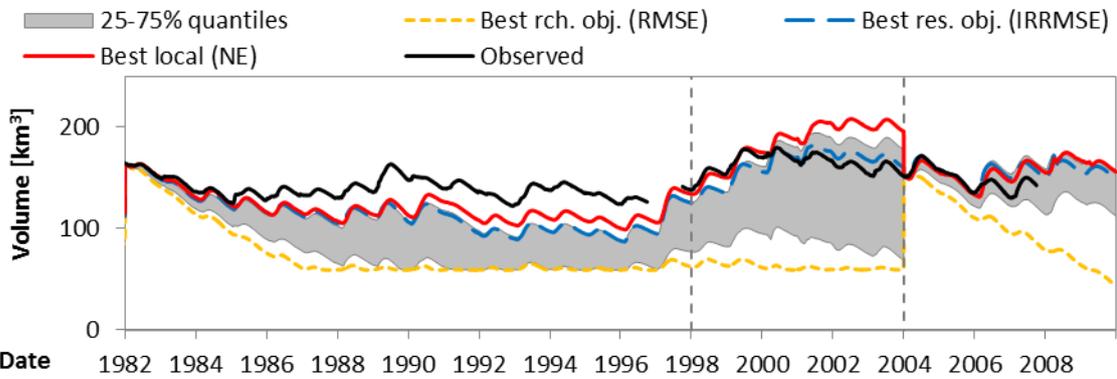
Figure AI.146. Synthesis volume series of the cascading-regularized calibration of the detailed model at the Kariba Dam.
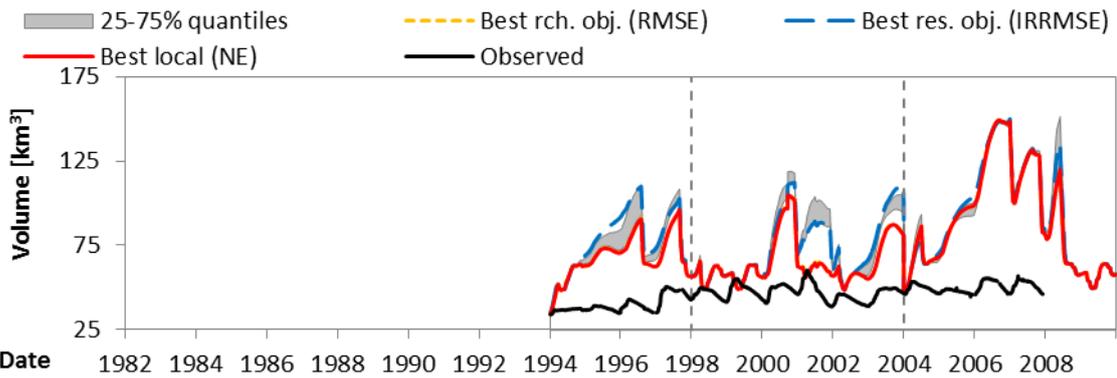
Figure AI.147. Synthesis volume series of the cascading-regularized calibration of the detailed model at the Cahora Bassa Dam.
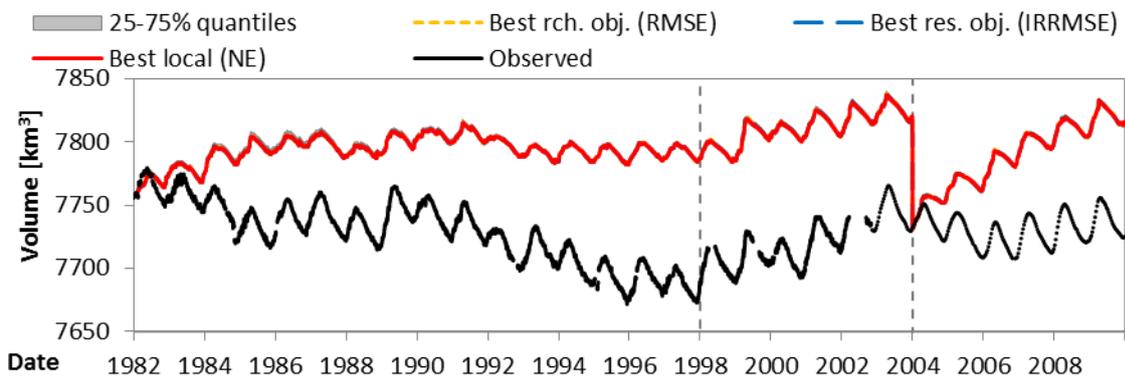
Figure AI.148. Synthesis volume series of the cascading-regularized calibration of the detailed model at the Malawi Lake.

APPENDIX I.Q. ADDITIONAL ERROR MEASURES FOR COMPARING CALIBRATION STRATEGIES
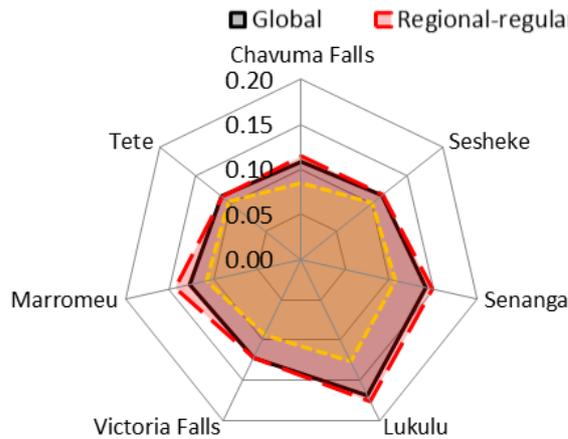
Figure AI.149. Best calibration IRRMSE of the global, regional-regularized, and cascading-regularized models at points along the main reach of the Zambezi River.
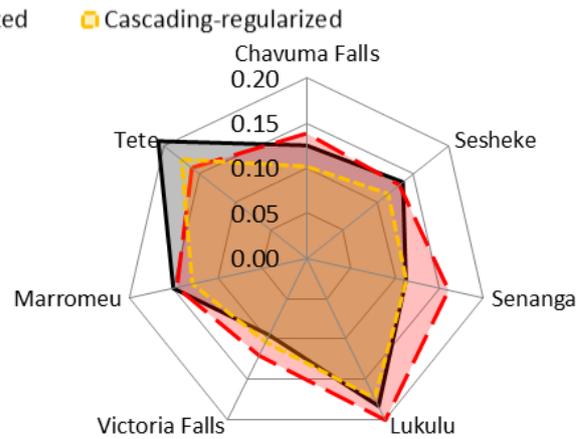
Figure AI.150. Best validation IRRMSE of the global, regional-regularized, and cascading-regularized models at points along the main reach of the Zambezi River.
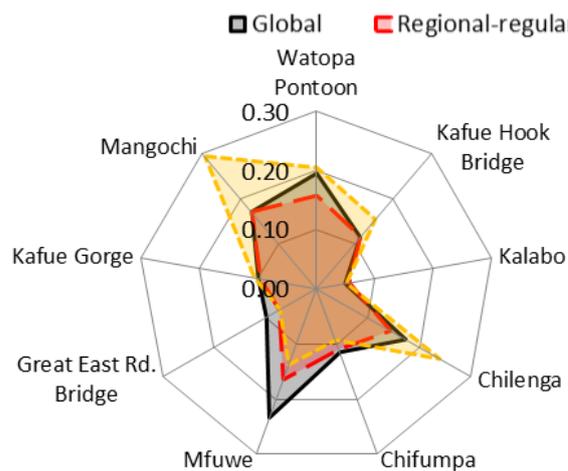
Figure AI.151. Best calibration IRRMSE of the global, regional-regularized, and cascading-regularized models at points along the main tributaries.
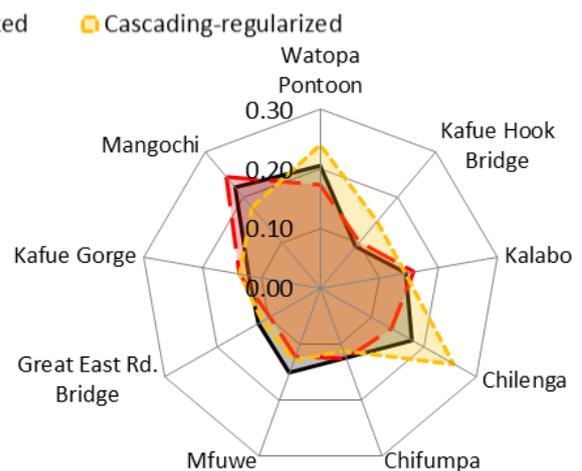
Figure AI.152. Best validation IRRMSE of the global, regional-regularized, and cascading-regularized models at points along the main tributaries.
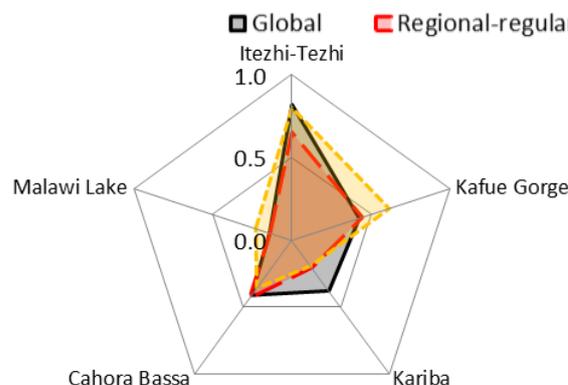
Figure AI.153. Best calibration IRRMSE of the global, regional-regularized, and cascading-regularized models at the monitored reservoirs.
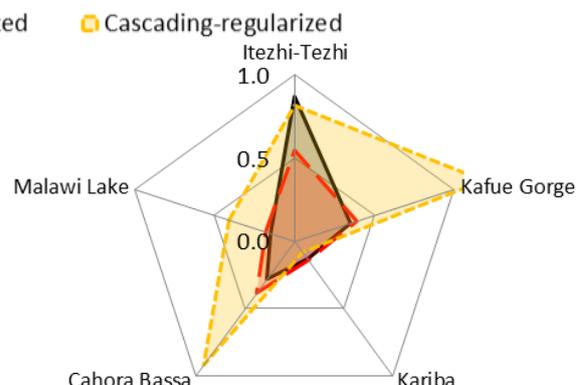
Figure AI.154. Best validation IRRMSE of the global, regional-regularized, and cascading-regularized models at the monitored reservoirs.

Below, the plots referring to the NE coefficient are presented. For convenience, the values are actually displayed as 1-NE, or NE*. Consequently, values closer to zero indicate the best performances. A logarithmic scale is used.
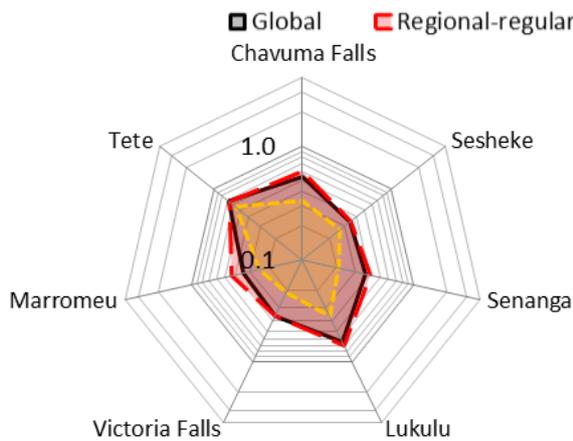

Figure AI.155. Best calibration NE* of the global, regional-regularized, and cascading-regularized models at points along the main reach of the Zambezi River.
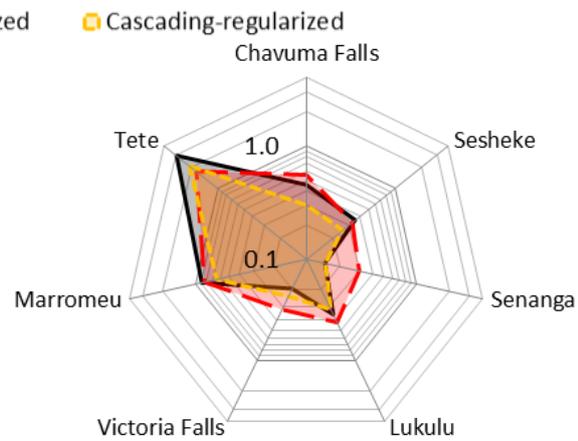

Figure AI.156. Best validation NE* of the global, regional-regularized, and cascading-regularized models at points along the main reach of the Zambezi River.
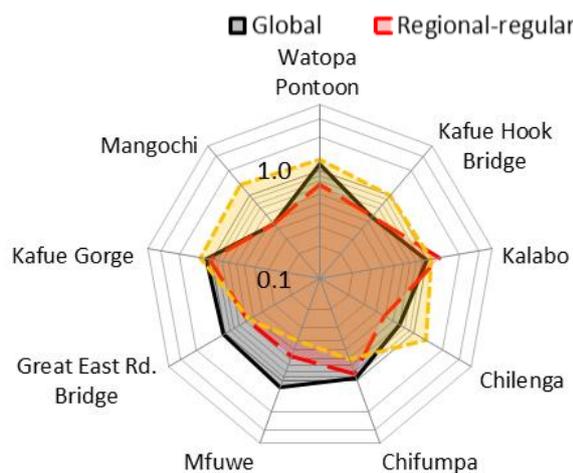

Figure AI.157. Best calibration NE* of the global, regional-regularized, and cascading-regularized models at points along the main tributaries.
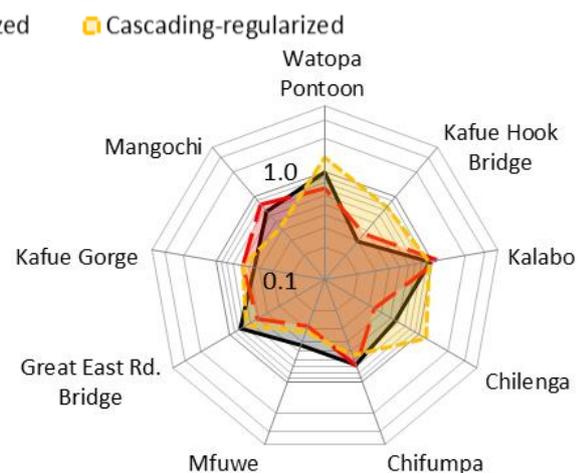

Figure AI.158. Best validation NE* of the global, regional-regularized, and cascading-regularized models at points along the main tributaries.
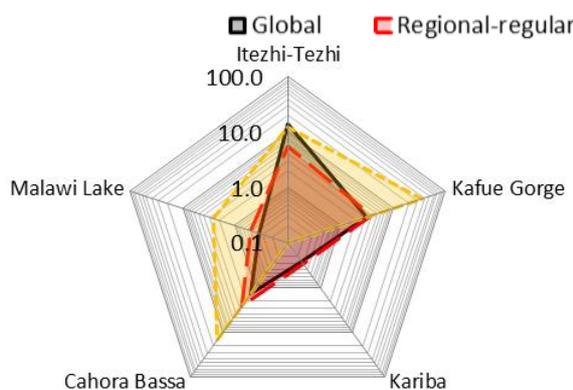

Figure AI.159. Best calibration NE* of the global, regional-regularized, and cascading-regularized models at the monitored reservoirs.
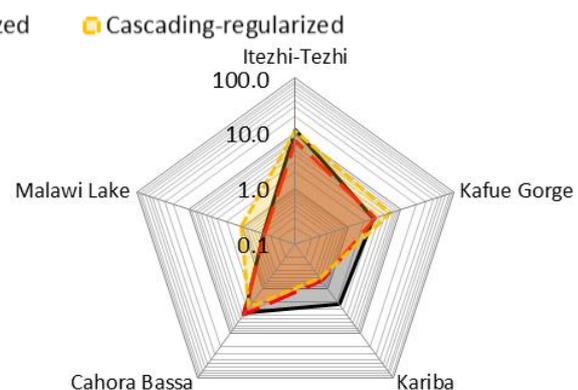

Figure AI.160. Best validation NE* of the global, regional-regularized, and cascading-regularized models at the monitored reservoirs.

APPENDIX II

APPENDIX II.A. THE ADAPT ONLINE DATABASE

Motivation and objectives

The present chapter very briefly presents the ADAPT online database, a web resource developed in the framework of the thesis and the Swiss Competence Center Environment and Sustainability (CCES) funded African Dams Project (ADAPT)[1]. The reader is welcome to visit its website and register at:

http://zambezi.epfl.ch/

The database was initially developed as a tool for storing and displaying the great amounts of hydrological data gathered by fellow ADAPT researchers. It primarily included observed discharges, but was also fitted to store other hydrologically relevant time series.

It soon became evident that the database could be useful, not only within ADAPT, but also to third parties, which motivated further developments. These included the storage and display of daily satellite data from several sources, the programming of a water resources section promoting the sharing of documents of various sorts and the inclusion of a models section. In this latter section the developed SWAT models, described in Chapters 7, 8 and 9, along with the model developed by Cohen Liechti (2013) can be modified and run[2].

Because most of the data remains property of the stakeholders who provided it and should not be made available to the general public, the database's access relies on a credential system designed to give the data owners complete control over which users can access it.

Engaging the data owners is essential to the long-term success of the database. Once protocols for regular data updates are in place, exciting opportunities will arise. One regarded as particularly promising is the inclusion of short to long term discharge forecast functionalities on the website, which can be done by implementing the powerful methodology described in Section 6.6 and alternative methodologies developed within ADAPT – notably the machine learning models developed for the Kafue River by Florian Köck.

The database has been endorsed by ADAPT's coordinators and was met with interest at an ADAPT workshop organized within the Zambian Water Forum and Exhibition (ZAWAFE), held in November 2013 in Lusaka, Zambia. Current plans are to further develop and showcase it during the next two years within the framework of a grant from CCES on the subject of knowledge transfer and conservation.

It is hoped that this resource can contribute to improve data sharing practices over the basin and thus potentiate new and better works on water resources assessment and planning. Also, it is regarded as a means to augment the impact of ADAPT's outcomes and through its models, provide and "in kind" return to contributing stakeholders.

Overview

The website has four main areas focusing on the introduction to the website itself and the ZRB, the data, modeling, and account management and administration. These sections are accessible through a simple front page, displayed in Figure AII.1.

---

[1] More information about ADAPT can be found at http://www.cces.ethz.ch/projects/nature/adapt.
[2] At the time of writing this functionality has not been made available to the public mostly due to computational concerns on the server which still need to be addressed.

Figure AII.1. Front-page of the ADAPT online database.

The sections under "introduction" contain information about the website itself, a walkthrough with several videos that illustrates how the website works (Figure AII.2), an interactive map identifying the main river-related features of the ZRB, links, and a frequently asked questions (FAQ) area.
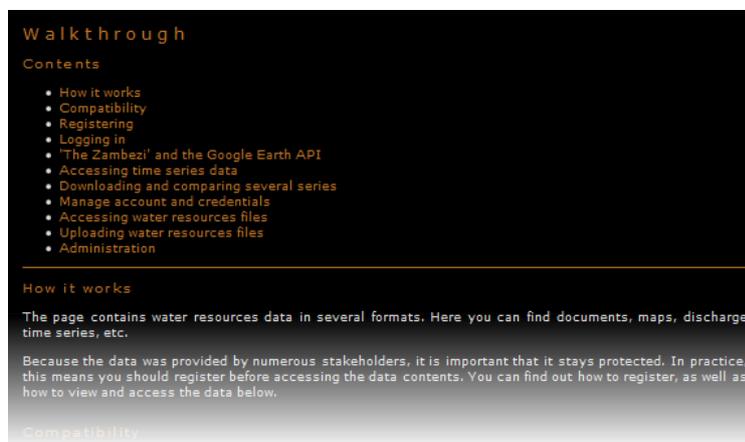

Figure AII.2. Detail of the walkthrough section of the website.

The database hosts several types of time series georeferenced data. Beyond this, it also serves as a repository for documents, spreadsheets, maps, etc. These latter types of information are accessible in the "water resources database" section. The information provided can be sorted by using queries related to their title's contents, authors or keywords. An example of query results is shown in Figure AII.3.

By clicking an entry's title the user gets access to detailed information about the data and the possibility to download it. Under the "upload data" section users can easily upload new entries to the database.

Figure AII.3. Detail of the water resources database" section. The query fields and first results are illustrated.

The handling of time series information is more complex. Presented with a 3-D World Globe, users can opt to load several layers, ranging from maps to the locations of chosen data types. The distribution of river discharge series across the basin is depicted in Figure AII.4.
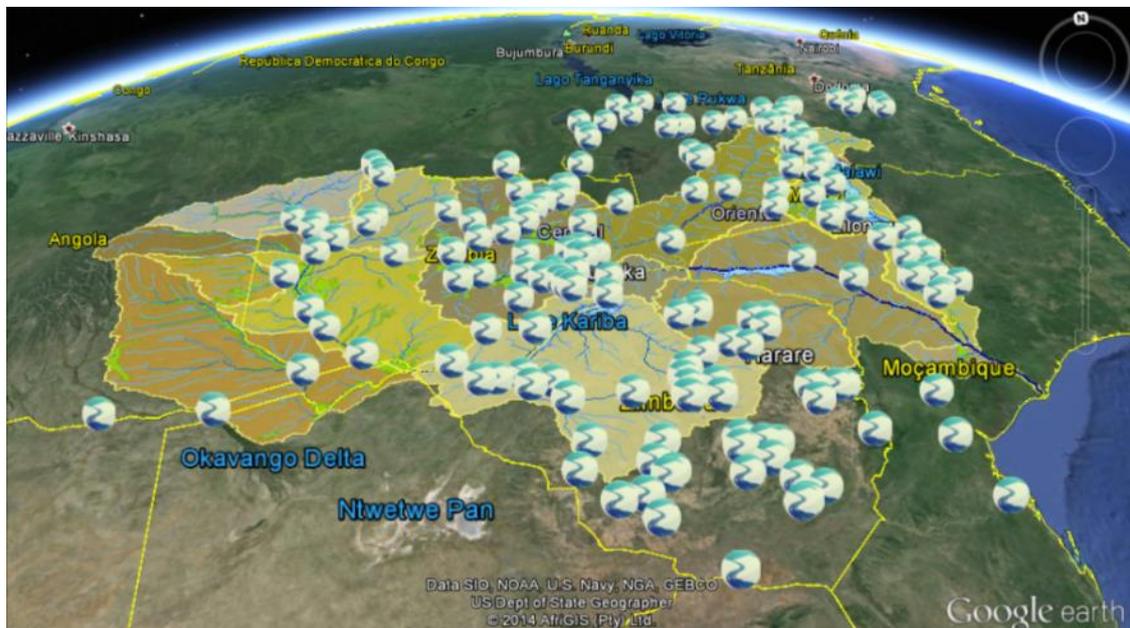
Figure AII.4. Locations of the stored river discharge time series over the ZRB's area (background map from Google Earth).

Easily extendable, presently the database is prepared to handle the following types of time series data:

- river discharge;
- spillway discharge;
- turbine discharge;
- water levels;
- ground rainfall;
- satellite rainfall;
- potential evapotranspiration; and
- real evapotranspiration.

Each of the points contains metadata about the corresponding series. The metadata is displayed in an information balloon once a given series icon is clicked. One such balloon is illustrated in Figure AII.5.
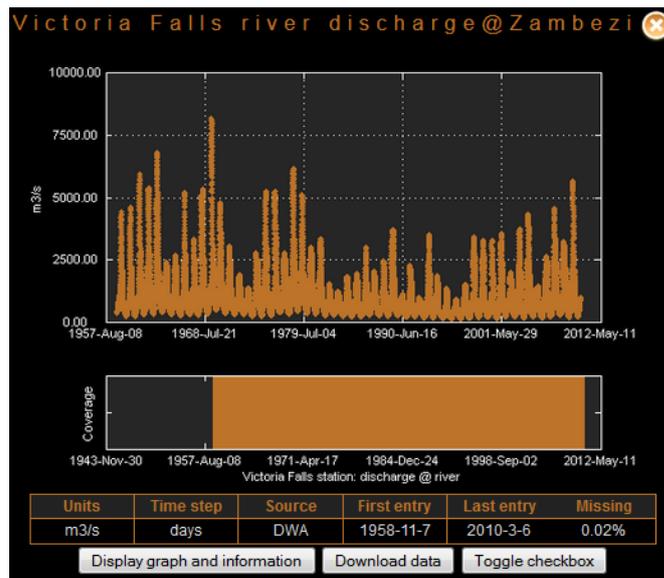
Figure AII.5. Example of an information balloon which is displayed by clicking a given data icon on the map.

Either through the information balloon or an organized tree-view list of the series currently displayed on the globe, a number of series can be selected in parallel. The user can then choose to view a chart containing the selected series or proceed to download them. The chart, shown in Figure AII.6, is fully interactive and allows for the detailed comparison of the selected series. Both chart and downloadable data are updated in real-time along with the database.
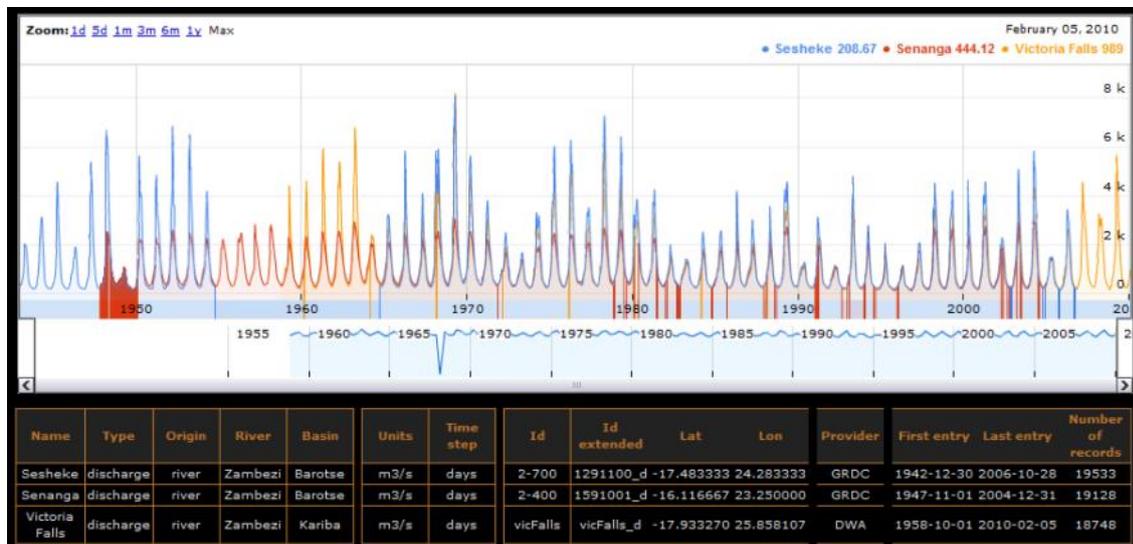
**Figure AII.6. Interactive chart allowing for the detailed inspection and comparison of data series.**

Rainfall satellite estimates are also a time series but they differ from all of the others stored in the database as they are distributed in space. They can be loaded according to data product, year and month, and followingly browsed on a daily time step. An example of satellite rainfall display is given by Figure AII.7.
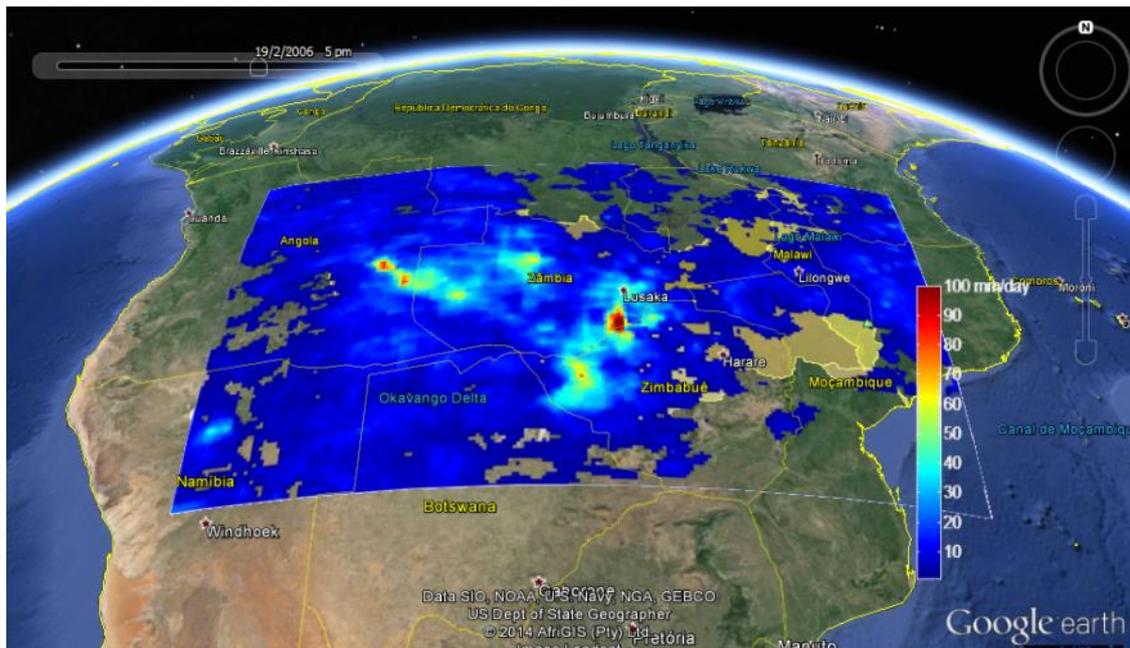
Figure AII.7. Example of daily rainfall distribution over the basin (19th February 2006). Data based on the NOAA/FEWS RFE2.0 product. Background map from Google Earth.

The third area of the website is mostly an interface for SWAT. It is prepared to host several models. Once one of these models is chosen, subbasins, reaches, and monitoring points can be plotted on the map. The main parameters associated with the subbasins and reaches can be easily analyzed through a color scale. In Figure AII.8, one of the regional-regularized models, described in Chapter 9, is depicted. The colors of subbasins and reaches reflect average tributary and main channel width, respectively.
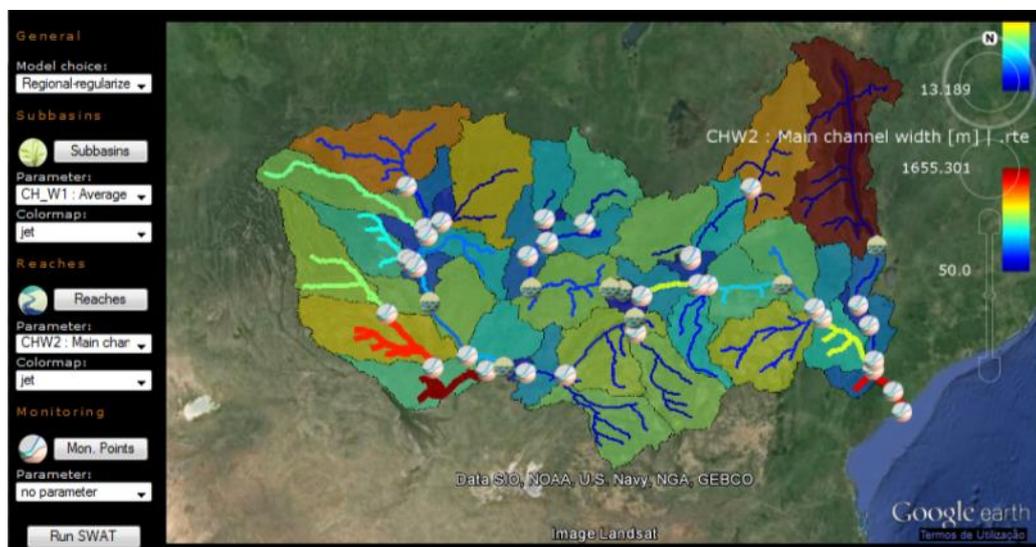

Figure AII.8. Illustration of a calibrated SWAT model of the ZRB. Feature colors translate average tributary and main channel widths.

As mentioned before, the interface is interactive. By clicking any subbasin or reach the SWAT files that characterize it can be accessed. Changes can be made to any file and submitted to the server (Figure AII.9).
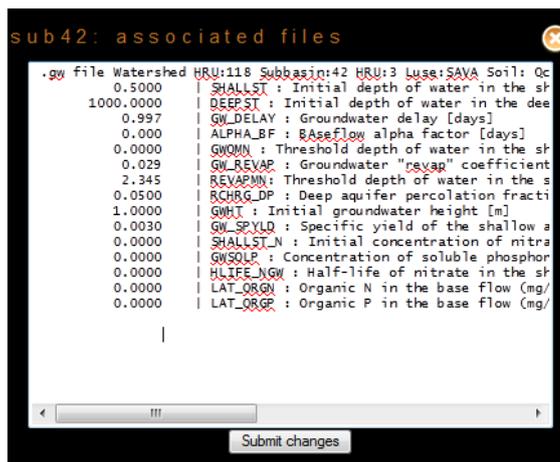
Figure AII.9. Balloon which allows viewing and changing the SWAT's simulation files associated with a given map feature.

Once and if any changes are made, SWAT can also be rerun on server (Figure AII.10).

The model's results are accessible by clicking any of the monitoring points. If an observed series corresponding to the selected monitoring point is available, it will be simultaneously displayed. The interactive chart presented in Figure AII.6 is used in order to this end.

In the future, it is a similar interface that will be used in order to publish discharge forecasts at different locations of the basin.

Technical aspects

The database was conceived with the goal of eventually being transmitted to African stakeholders. Along this line, it is mostly based in open source or free technology.

As aforementioned, a major concern during programming was to ensure that the property of the data is maintained. To that end, a distributed credential system was developed. Once users register in the website and gain access to protected areas, they can request access to proprietary data. The infrastructure is then prepared to pass these requests to "data administrators". The "data administrators", on behalf of the data owners, have then the capability to grant or withdraw access rights to their data in particular.

Figure AII.10. Dialog following a successful SWAT run.

The website runs on an open-source Apache Tomcat[3] server, which implements Java Servlet and JavaServer Pages (JSP) technologies. Maps and charts rely on freely available Application Programming Interfaces (API) from Google[4]. The also free MySQL[5] was the chosen database solution. The intervening programming languages are HTML, Java, JavaScript, JavaServer Pages Standard Tag Library (JSTL), SQL and Matlab.

The only costly software in use is Matlab, mostly associated with running and displaying the SWAT models. While there would be only limited technical difficulties in developing the needed

---

[3] http://tomcat.apache.org/.
[4] https://developers.google.com/earth/ and https://developers.google.com/chart/.
[5] http://www.mysql.com/.

functionalities as server capabilities using solely the Java programming language or, alternatively, a free external alternative such as the R software, these are time consuming efforts. Under the time constraints of the Ph.D. research, the easier choice of using Matlab during an initial stage was made.

Hosting is provided by EPFL under a (also not free) Windows Server operating system. Plans exist to migrate the structure to a location within the ZRB and few stakeholders, contacted by ADAPT coordinators, have shown interest in this opportunity. The acquisition of hard and software to equip these servers is accommodated in the already referenced CCES grant. Notwithstanding, the aim is to adapt the code infrastructure in order for it to run on a Linux server, thus using exclusively free software.

In line with CCES' grant proposal, the ADAPT online database will also be integrated with broader data platforms such as the Online Support Platform for Environmental Research (OSPER)[6] or Envidat.ch.

While containing a plethora of historical data, the database lacks up-to-date inputs, which reduces its capabilities. In the future, every effort will be made in order to engage local data providers, establish data sharing protocols, and put in place software for real-time data transfer.

---

[6] http://www.swiss-experiment.ch/.

CURRICULUM VITAE

Work experience

Dates	**Since March 2014.**
Occupation or position held	Post-doc.
Main activities and responsibilities	- Research (risk, hydrology, machine learning, and optimization).
Name and address of employer	EPFL: Laboratoire de Constructions Hydrauliques (LCH), EPFL-ENAC-IIC-LCH, GC A3 504, Station 18, CH-1015 Lausanne, Switzerland.

Dates	**October 2009 to 2014**
Occupation or position held	PhD student.
Main activities and responsibilities	- Research (hydrology, machine learning, and remote sensing). - Consulting (hydrology, hydraulics). - Teaching assistant (hydrology, hydraulics and statistics).
Name and address of employer	EPFL: Laboratoire de Constructions Hydrauliques (LCH), EPFL-ENAC-IIC-LCH, GC A3 504, Station 18, CH-1015 Lausanne, Switzerland. IST: CEHIDRO, Departamento de Engenharia Civil e Arquitectura, Instituto Superior Técnico (IST) Av. Rovisco Pais, 1049-001, Lisboa, Portugal.

Dates	**May, 2008 to September 2009.**
Occupation or position held	Junior Engineer.
Main activities and responsibilities	- Development of working tools for the analysis of telemetry data for water supply systems. - Participation in a leak detection program. - Cooperation in the preparation of commercial proposals for new concessions for municipal water and wastewater services.
Name and address of employer	AGS, Administração e Gestão de Sistemas de Salubridade, S.A. SintraCascais Escritórios, Rua da Tapada da Quinta de Cima – Linhó, 2714-555 Sintra, Portugal.
Type of business or sector	Water and wastewater services company.
Dates	**October, 2005 to May, 2008**
Occupation or position held	Junior Engineer (part-time).
Main activities and responsibilities	Collaboration in the following projects: - Design of small-scale sewerage systems. - Preliminary design of a long sea outfall. - Calculation of necessary storage volumes for the peak flow reduction using HEC-HMS (*Hydrologic Engineering Center - Hydrologic Modeling System*). - Drainage system's performance assessment using EPA (*U. S. Environment Protection Agency*) SWWM (*Storm Water Management Model*).
Name and address of employer	Hidra, Hidráulica e Ambiente, Lda Av. Defensores de Chaves nº.31, 1º Esq. 1000-110, Lisboa
Type of business or sector	Consulting of urban hydraulics and environmental engineering.
Dates	**March to August, 2007**
Main activities and responsibilities	- Cooperation in a hydrologic study with the objective of estimating minimum flows in the catchment of a regional solid waste treatment system (Meia Serra, Madeira island, Portugal).
Name and address of employer	Assistant Professor Maria Manuela Portela Silva. Secção de Hidráulica e Recursos Hídricos do Instituto Superior Técnico, Rua Rovisco Pais, 1049-001 Lisboa.
Dates	**January to April, 2008.**
Main activities and responsibilities	- Cooperation in a hydraulic/hydrologic study for estimating peak flows in the Tagus river, Portugal.
Name and address of employer	Assistant Professor Maria Manuela Portela Silva. Secção de Hidráulica e Recursos Hídricos do Instituto Superior Técnico, Rua Rovisco Pais, 1049-001 Lisboa.

Education and training	
Dates	**Since October 2009.**
Principal subjects/occupational skills covered	Ongoing Ph.D. research. Doctoral Program in Environment; IST-EPFL joint doctoral program. Thesis title: *Hydraulic-hydrologic model for the Zambezi River using satellite data and artificial intelligence techniques*.
Name and type of organisation providing education and training	École Polytechnique Fédérale de Lausanne (EPFL) / University of Lisbon – Instituto Superior Técnico (IST).
Level in national or international classification	ISCED 8
Dates	**December, 2008.**
Principal subjects/occupational skills covered	Civil Engineering MSc. Thesis title: *Using artificial neural networks to model the degradation of water supply systems*.
Name and type of organisation providing education and training	Technical University of Lisbon - Instituto Superior Técnico (IST).
Level in national or international classification	ISCED 7 (17/20).
Dates	**September, 2007.**
Principal subjects/occupational skills covered	Degree in Civil Engineering.
Name and type of organisation providing education and training	Technical University of Lisbon - Instituto Superior Técnico (IST).
Level in national or international classification	ISCED 3 (15/20).
Personal skills and competences	
Computer skills and competences	Familiarity with the following modeling software: - Soil & Water Assessment Tool (coding experience), HEC-HMS, HEC-RAS, EPANET (coding experience), EPA SWWM, and FLOW-3D. Programming experience in: - Java, Fortran, JavaScript, VB.Net, SQL, HTML, VBA (Microsoft Excel, Microsoft Word, AutoCAD), Matlab, R and KML (Google Earth).
Awards	Alberto Abecassis Manzanares Award (4th edition, December 2007), granted to the 3 best students in hydraulics and water resources courses amongst civil engineering students at IST.

Scientific publications

Master's thesis	Using artificial neural networks to model the degradation of water supply systems, MSc. Thesis, Instituto Superior Técnico (IST), Technical University of Lisbon, 2008.
Papers in international scientific periodicals with referees	Cohen Liechti, T., Matos, J. P., Boillat, J.-L., and Schleiss, A. J.: Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi River Basin, Hydrol. Earth Syst. Sci., 16, 489-500, doi:10.5194/hess-16-489-2012, 2012.

Marchi, A., Salomons, E., Ostfeld, A., Kapelan, Z., Simpson, A., Zecchin, A., Maier, H., Wu, Z., Elsayed, S., Song, Y., Walski, T., Stokes, C., Wu, W., Dandy, G., Alvisi, S., Creaco, E., Franchini, M., Saldarriaga, J., Páez, D., Hernández, D., Bohórquez, J., Bent, R., Coffrin, C., Judi, D., McPherson, T., van Hentenryck, P., Matos, J., Monteiro, A., Matias, N., Yoo, D., Lee, H., Kim, J., Iglesias-Rey, P., Martínez-Solano, F., Mora-Meliá, D., Ribelles-Aguilar, J., Guidolin, M., Fu, G., Reed, P., Wang, Q., Liu, H., McClymont, K., Johns, M., Keedwell, E., Kandiah, V., Jasper, M., Drake, K., Shafiee, E., Barandouzi, M., Berglund, A., Brill, D., Mahinthakumar, G., Ranjithan, R., Zechman, E., Morley, M., Tricarico, C., de Marinis, G., Tolson, B., Khedr, A., and Asadzadeh, M.: The Battle of the Water Networks II (BWN-II), Journal of Water Resources Planning and Management, doi:10.1061/(ASCE)WR.1943-5452.0000378, 2013.

Mertens, J., Wehrli, B., Tilmant, A., Schleiss, A., Cohen Liechti, T., and Matos, J. P.: Adapted reservoir management in the Zambezi river basin to meet environmental needs, International Journal on Hydropower & Dams, 20(2), 80-84, 2013.

Matos, J. P., Cohen Liechti, T., Juízo, D., Portela, M. M., and Schleiss, A. J.: Can satellite based pattern-oriented memory improve the interpolation of sparse historical rainfall records?, Journal of Hydrology, 492, 7, 102-116, doi:10.1016/j.jhydrol.2013.04.014, 2013.

Sousa, V., Matos, J. P., Almeida, N., and Saldanha Matos, J.: Risk assessment of sewer condition using artificial intelligence tools: application to the SANEST sewer system, Water Science & Technology, In Press, doi:10.2166/wst.2013.758.

Matos, J. P., Cohen Liechti, T., Portela, M. M., and Schleiss, A. J.: Pattern-oriented memory interpolation of sparse historical rainfall records, Journal of Hydrology, 510, 493-503, doi:10.1016/j.jhydrol.2014.01.003.

Cohen Liechti, T., Matos, J. P., Ferràs Segura, D., Boillat, J.-L., and Schleiss, A. J.: Hydrological modelling of the Zambezi River Basin taking into account floodplain behaviour by a modified reservoir approach, International Journal of River Basin Management, 12, 1, 29-41, doi: 10.1080/15715124.2014.880707

Papers in national periodicals with referees	Matos, J. P., Monteiro, A. J.: Utilização de Redes Neuronais Artificiais para a Modelação da Degradação de Sistemas de Distribuição de Água, Águas e Resíduos, III-9, 38-49, 2009.
	Portela, M. M., Fragoso dos Santos, J. F., Naghettini, M., Matos, J. P., Silva, A. T.: Superfícies de limiares de precipitação para identificação de secas em Portugal continental: uma aplicação complementar do Índice de Precipitação Padronizada, SPI, Recursos Hídricos, 33, 2, 2012.
Papers in conference proceedings	Matos, J. P., Monteiro, A. J.: Utilização de Redes Neuronais Artificiais para a Modelação da Degradação de Sistemas de Distribuição de Água, 13º Encontro Nacional de Saneamento Básico, Covilhã, Portugal, Outubro 2008.
	Matos, J. P., Matias, N., Matos, J. S.: Potencialidades e Limitações do Método Holandês para a Estimativa de Volumes de Retenção de Águas Pluviais em Meio Urbano, 13º Encontro Nacional de Saneamento Básico, Covilhã, Portugal, Outubro 2008.
	Matos J. P., Cohen T., Boillat J.-L., Portela M. M., Schleiss A. J.: Analysis of flow regime changes due to operation of artificial reservoirs on the Zambezi River, proceedings of the 6th International Symposium on Environmental Hydraulics, Athens, Greece, June 2010.
	Cohen Liechti, T., Matos, J. P., Boillat. J.-L., Schleiss, A. J., Portela, M.M.: Semi-distributed hydraulic-hydrological modeling of the Zambezi River Basin, International Conference on the Status and Future of the World's Large Rivers, Vienna, April 2011.
	Matos, J. P., Matias, N., Monteiro, A. J.: Redesigning water distribution networks through a structured evolutionary approach, 14th Water Distribution Systems Analysis Conference, Adelaide, Australia, September 2012.
	Matos, J. P., T., Portela, M.M., Juízo, D. and Schleiss A. J.: Base de dados e modelo hidrológico da bacia do Zambeze, 11º Simpósio de Hidráulica e Recursos Hídricos dos Países de Língua Oficial Portuguesa, Maputo, Mozambique, May 2013.
	Matos, J. P., T., Portela M. M. and Schleiss A. J.: Cálculo e previsão melhorada de caudais afluentes a grandes albufeiras utilizando filtros de Kalman, 11º Simpósio de Hidráulica e Recursos Hídricos dos Países de Língua Oficial Portuguesa, Maputo, Mozambique, May 2013.
	Sousa, V., Matos, J. P., Almeida, N. and Saldanha Matos, J.: Risk assessment of sewer condition using artificial intelligence tools: Application to the SANEST sewer system, Proceedings of the 7th International Conference on Sewer Processes and Networks, Sheffield, United Kingdom, August 2013.
	Matos, J. P., Cohen Liechti, T., Portela, M.M. and Schleiss A. J.: Coupling satellite rainfall estimates and machine learning techniques for flow forecast: application to a large catchment in Southern Africa, Proceedings of the 2013 IAHR Congress, Chengdu, China, September 2013.