Developing an open-source database for the Zambezi river basin

J.P. Matos and A.J. Schleiss, EPFL, Switzerland
J. Mertens, ETH Zurich, Switzerland
B. Wehrli, ETH Zurich and EAWAG, Switzerland

The Zambezi river basin, one of the largest in Africa, holds significant hydropower potential, which has yet to be exploited. To make the most of this opportunity and address the challenges it poses, local stakeholders must coordinate efforts, politically, institutionally, and technically. While recent progress on these fronts has been remarkable, legitimate concerns regarding data ownership, maintenance, and analysis still need to be addressed. This paper presents the African Dams Project's (ADAPT) water resources database, which relies on open source technology and is based on three key ideas: the total empowerment of stakeholders; the association of valuable analysis tools and data; and, an adaptable and simple architecture.

The Zambezi river basin in Southern Africa extends across eight countries and covers 1.4 × 10^6 km^2. It is home to approximately 30 million people, harbours a number of wildlife sites including areas of unique ecological and cultural value, some classified as UNESCO World Heritage and Ramsar Convention sites, and has an estimated potential hydropower capacity of 12 000 MW, of which only about 5000 MW is currently developed [The World Bank, 2010[1]]. An overview of the basin is presented in Fig. 1. It includes the basin's main features and the location of three discharge-gauging stations, the data from which are referred to later in this paper.

1. Current development

The Zambezi basin is one of the most valuable natural resources in Africa. Although it has significant unexploited hydropower potential, it is already one of the continent's most heavily dammed river systems, with major hydropower schemes such as Kariba (1830 MW), Cahora Bassa (2075 MW), and the Kafue hydropower system (900 MW at the Kafue Gorge dam). At present, a number of new projects are at various stages of development: the Mphanda Nkuwa run-of-river scheme (1500 MW) downstream of Cahora Bassa [COBA *et al.*, 2011[2]]; the Batoka Gorge dam (1600 MW) downstream of Victoria Falls; the Kafue Gorge lower dam (750 MW); and, the Cahora Bassa North Bank Extension (1250 MW) [Nippon Koei UK, 2012[3]]. As well as offering a source of hydropower, the Zambezi basin's river network acts as a source of sustenance and wealth to the region, through activities such as fishing, farming and tourism. Local communities are not only economically, but also culturally, linked to the main rivers and wetlands.

There are numerous opportunities to be seized within the Zambezi basin, and better management and cooperative development of the basin's water resources could significantly increase agricultural yields, hydropower outputs, and economic gains [The World Bank, 2010[1]; SWRSD Zambezi Basin Joint Venture, 2011[4]].

The future, however, also holds challenges. These include: climate change (with some estimates for the Zambezi predicting extreme variations) [Arnell, 1999[5]; IPCC, 2001[6]; Beilfuss, 2012[7]]; growing demands on water resources, which will increase the pressure on ecologically valuable areas; and, the definition and maintenance of ecological flows, which are essential to achieve sensible compromises between the economy and the environment. It is only through a proper evaluation of the basin's current state and future scenarios that an adequate balance between hydropower, water abstractions, and the preservation of natural ecosystems can be achieved.

At operational level, joint reservoir management, the scheduling of environmental flows, flood forecasting and flow control depend on sound hydrologic models, hydrological data and expertise. Even more so than planning efforts, such operational activities call for reliable and well established technical and institutional frameworks.

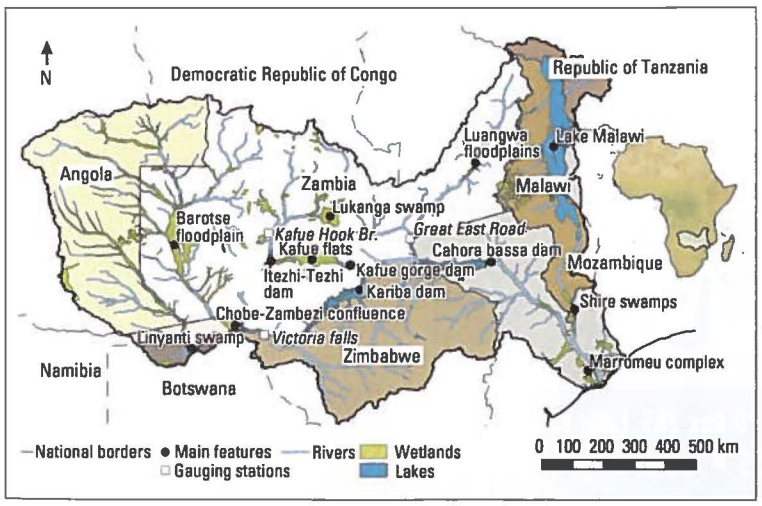


Fig. 1. Overview of the Zambezi river basin. The main features of the basin, along with three gauging stations, are highlighted.

1.1 The African Dams Project

Focusing on the Zambezi basin, the African Dams Project [ADAPT[8]] has been a fruitful interdisciplinary endeavour to enhance the scientific basis of integrated water resources management. From 2007 to 2013 it benefited from the collaboration of African and European partners in fields as diverse as ecology, economy, hydrology, and biogeochemistry [Mertens, 2013[9]; Mertens *et al.*, 2013[10]].

Operations were carried out on two levels:

- clearly defined research activities, which generated knowledge through data collection, concepts, and numerical models; and,

• a science-policy interface, which involved capacity building efforts (student exchanges and training activities), and interdisciplinary integration of research results, interacting with stakeholders and policy makers in the basin.

Follow-up activities to the main phase of ADAPT entail further contributions to the Zambezi's water resources community. One such activity is the establishment of an open-source online platform [EPFL[11]], which addresses two of the main obstacles faced by researchers and practitioners working in the region: the difficult access to water resources data and the need for customized tools to analyse it. With an aim to collect and divulge data, analysis tools, and knowledge about the basin, the platform contains a wealth of downloadable content, such as hydrological time series and a variety of documents ranging from reports to spreadsheets. It also includes an interface to calibrated Soil and Water Assessment Tool (SWAT) models [Arnold *et al.*, 2012[12]] of the Zambezi basin and sections dedicated to the presentation of third party's works. There is also potential to develop a functional discharge monitoring and forecasting system.

1.2 Motivation

The aim of the paper is to describe how the ADAPT water resources database can contribute to a more resourceful and efficient local water resources community, as well as ultimately contribute towards more informed and sustainable regional water resources policies.

Being designed to address the concerns of a variety of local stakeholders (concerns about data ownership, maintenance, and analysis) this open-source tool has the potential to be implemented beyond the Zambezi, particularly so where a lack of hydrological data and a difficult integration of multiple stakeholders' interactions might hinder effective water resources management.

2. Aggregating water resources data and information

2.1 The need for a water resources database and past developments

The optimal allocation and management of the Zambezi basin's water resources poses a number of challenges to decision makers, planners, managers, modellers, and researchers. Among these are the size and heterogeneity of the basin, the complexity of the local hydraulic and hydrological phenomena, and the difficulty in accessing existing data. In fact, with a plethora of institutions and agencies responsible for collecting measurements on the ground, the search for data is often a source of major concern to scientists and engineers, unfortunately not always with positive or satisfying results.

First, it is challenging to get an overview of all the existing data and the institutions responsible for it. Such information alone would already lead to more informed strategic decisions, management, and operations throughout the basin. Going further, a standardized and streamlined procedure for data retrieval from the relevant institutions is needed. Finally, a centralized water resources database, where data can be archived and remain easily accessible, would be optimal.

Aware of the large gains that stand to be made from data sharing and improved technical, institutional and political interactions [The World Bank, 2010[1], SWRSD Zambezi Basin Joint Venture, 2011[4]], water institutions acting across states such as the Zambezi Watercourse Commission (ZAMCOM) are promoting this. Despite considerable efforts, flagship projects such as the centralized Zambezi Water Information System (ZAMWIS) have yet to benefit a wide community of stakeholders.

In the past, ZAMWIS fell short of its objectives because of technical, financial, and political issues. In fact, the availability of data alone provides little motivation to overcome the burden of continuously updating and maintaining a centralized water resources database. For such a database to be worthwhile, the tools and know-how needed to analyse and put the data to good use must be present as well. Both have been greatly developed in recent times and, presently, ZAMCOM is undertaking a very promising effort to renew and operationalize ZAMWIS. Because of the large number of states and institutions acting as data providers in the Zambezi basin, it will be a major challenge to guarantee that the database is kept up to date. Ensuring a mandate to redistribute it will also be a complex process.

2.2 The ADAPT water resources database

As with most water resources endeavors carried out over the ZRB, ADAPT faced the difficulties raised by the lack of a centralized and regularly updated database. Although the project has taken advantage of its scale, as well as close cooperation with local institutions, which have been supportive, the data gathering process was still slow and complex.

The ADAPT database [EPFL[11]] started as an effort to organize and distribute data and information within the project. Later, it was extended as a vehicle to share the project's results and documentation.

At present, ADAPT's water resources database serves a broader purpose; it is a flexible open-source tool for the storage, organization and dissemination of time series data documentation. Along with this, it provides a platform for data analysis tools.

Three views of the time series module are presented in Fig. 2. Through the online interface, users can inspect processed daily satellite rainfall data from multiple sources and access historical series of river discharge, water level, rainfall, as well as other variables. These can be inspected in detail and, eventually, downloaded.

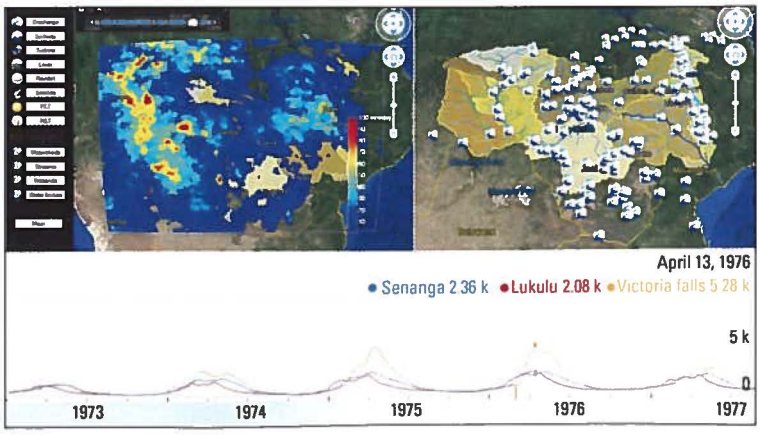


Fig. 2. View of the time series data selection tool. Left: daily satellite rainfall estimates; right: discharge gauge locations; below: time series interactive view.

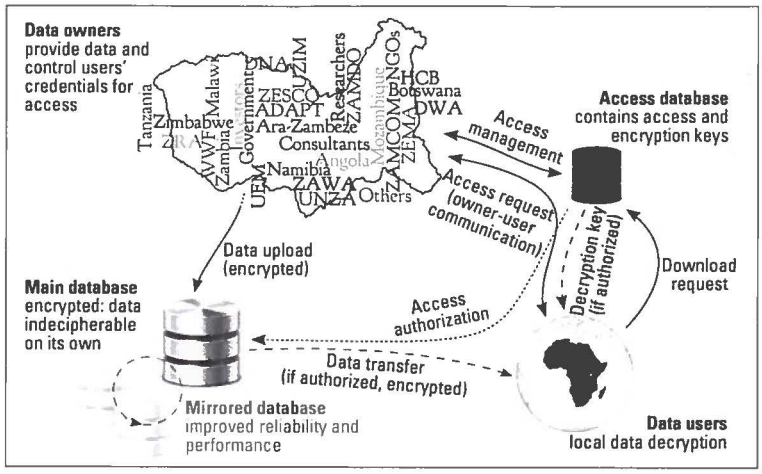


Fig. 3. Scheme of the ADAPT database's interactions.

During the development of the ADAPT water resources database, efforts such as ZAMWIS and solutions individually implemented by stakeholders were kept in mind. In response to these, a lightweight implementation and open-source solution was combined with a distributed architecture, in which each data provider retains complete control over its own data. It aims to be a solution that is:

- user friendly;
- cheap (based on open source software);
- easy to maintain;
- adaptable;
- safe (protecting data from loss and unintended usage); and,
- lightweight (requiring limited hardware resources).

The database is being developed in collaboration with local institutions to expand its capabilities in terms of the types of data it can host and the associated analysis tools.

Beyond technical concerns, it is thought that the main objections of data providers and other stakeholders to such a centralized database unfold mainly at the institutional and political levels. As in many other large international river basins, stakeholders share many objectives. Some of their common areas of interest are, however, not necessarily aligned, and sometimes even conflicting (for example, hydropower production and irrigation of farmlands upstream). It is therefore expected that stakeholders would adopt a cautious attitude regarding access to their data. In some cases, the revenue from data sales is important to the data provider. When this is the case, controlled access to the data understandably becomes paramount. Aiming to reduce political and institutional issues, the ADAPT database does not require a mandate or permission from the data owners for data redistribution. Instead, each data owner keeps complete control over its own data, managing access authorizations directly.

As data ownership is regarded as a major issue, the ADAPT water resources database adopts a distributed ownership policy. As such, all the contents of the database are encrypted and associated with a specific data owner. The only way to access specific contents is for users to obtain explicit authorization from the data owner. The proposed data request process can be synthesized in the following steps:

- The user is registered and signs a general terms of use policy.
- The user can access a list of the available data (and data owners).
- Choosing particular data, the user proceeds to request the relevant data owner directly for access (the form of the request is tailored to the needs of each data owner, and involves specific conditions and terms of use policies).
- Someone appointed by the data owner receives an e-mail with all the relevant information and uses a simple menu to choose if (and for how long) the user may have access to the requested data.
- The user receives e-mail feedback and, in the case of a positive reply, can proceed to download the data and decrypt it.

The concept underlying the database's interactions is further illustrated in Fig. 3.

Underpinning the database's design is the desire to achieve a solution tailored for the Zambezi river basin. On the technical side, the platform must be adaptable, sparing in its use of resources, and perform well. On the political/institutional side, it must strive to overcome blocks by not requiring any concessions from the data owners.

Focusing on data management, the gains from the adoption of such a database are quite self-evident. Benefits include:

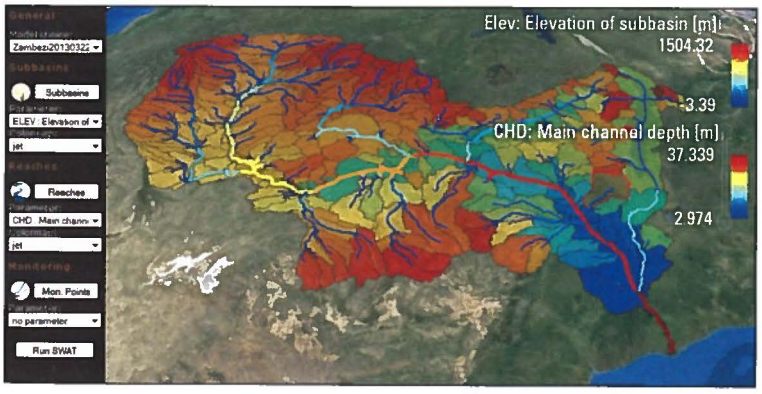
- easier assessment of the existing knowledge base;
- direct evaluation of real data gaps;
- effortless data sharing between stakeholders;
- more efficient data processing (fewer resources used as a result of easier and faster data access and standard data formats); and,
- added interest from research groups (the difficult access to water resources data within the basin is an obstacle in the way of sound scientific contributions).

A static hydrological database (one which is not consistently updated) has the capability to improve the quality and facilitate the validation of the hydrological models needed to support water resources planning. Consequently, it is of value to decision makers. If endorsed by key data providers and updated regularly, the database could provide additional benefits, particularly regarding operational tasks.

Activities such as the assessment of the effects of climate change, the study and planning of new hydropower schemes, the quantification of environmental flow requirements, and the evaluation of impacts associated with new water abstractions can benefit from such a database.

In addition, state-of-the-art operational flow forecasting models for the whole basin, integrating operations from an increasing number of major impound-

Fig. 4. The interactive online interface to the calibrated daily SWAT models of the Zambezi river basin.



ments, can only be achieved on the basis of a solid data-sharing infrastructure. Once implemented, they have the potential to make hydropower production more efficient and will be likely to contribute to better management of natural floods.

3. Customized modelling tools for water resources management

Data collection and dissemination is widely recognized as benefiting water resources management activities [Hannah *et al.*, 2011[13]]. According to the World Meteorological Organization [WMO, 2008[14]]:

"A data set is clearly of great value as it is inevitably collected through a huge commitment of time and money. The management of these data is therefore important work in itself and this work must be performed effectively in order to maximize the results of this investment. (...) A poor-quality archive, due to lack of forethought in its foundation or poor management, can lead to years of excess data collection or modelling work, and subsequent poor decision-making. (...) Moreover, poor-quality data and databases will result in suboptimal planning decisions and poorly designed engineering structures".

The maintenance and continued updating of a water resources database requires resources from data providers. However high the indirect value of the data may be, for some data providers day-to-day concerns are not necessarily bound to data retrieval or dissemination. In these cases, the data provider's support to the database is more easily justified if, in addition to indirect benefits, there are direct gains associated with its use.

The activities carried out within ADAPT led to the development of powerful analysis tools. These have been customized for the particular conditions found in the Zambezi basin and included the development and calibration of detailed daily hydrological models of the whole catchment using the Soil and Water Assessment Tool (SWAT) software [Arnold *et al.*, 2012[13]; Cohen Liechti, 2013[15]; Cohen Liechti *et al.*, 2014[16]; Matos, 2014[17]]. Models such as SWAT can provide a sound foundation for a broad range of water resources studies. Calibration and validation, presenting a difficult challenge given the size and lack of data in some parts of the basin, were only possible by resorting to the potential of the ADAPT database.

To make the SWAT models readily available to stakeholders, a customized online interface has been prepared (see Fig. 4). It allows users to change parameters, run simulations over a long period, compare results, and easily assess the impact of management of climatic changes within the basin.

While the SWAT models are useful for planning purposes, without data assimilation schemes their operational usefulness is limited. This step would be important to make the best possible use of new observations being made throughout the basin and, consequently, enhance results. Unfortunately, achieving such an integration is a time-consuming and technically challenging process.

To address the needs of local dam operators and present an immediate benefit to users, a new probabilistic flow forecasting technique is also being incorporated in ADAPT's database. It is highly adaptable and does not seek to replace, but rather ameliorate, the forecasting methods already in use. In Fig. 5, examples of forecasts upstream of the three major dams on the Zambezi river basin are shown: Kafue Hook Bridge, upstream of Itezhi-Tezhi and (a) 30-day lead time; Victoria Falls and (b) 30 day lead time; and, Great East Road bridge, upstream of Cahora Bassa and (c) and (d), variable lead times.

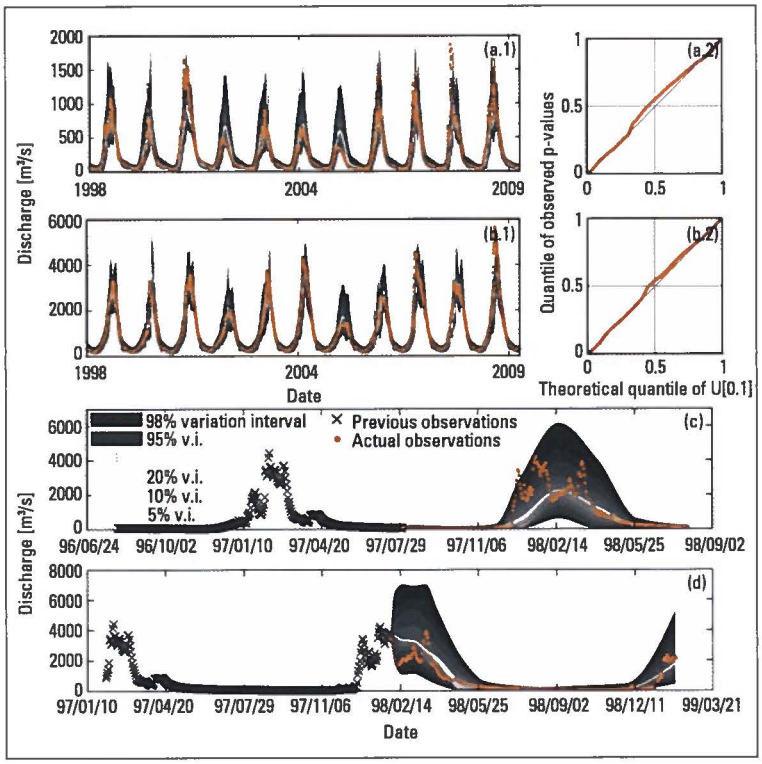


Fig. 5. Illustration of probabilistic flow forecasting models being developed. (a.1) Kafue Hook Bridge 30-days lead time forecast. (a.2) forecast probability distribution adjustment. (b.1) Victoria Falls 30-days lead time forecast. (b.2) forecast probability distribution adjustment. (c) 1 to 365-days lead time forecast from 1/8/1988 onwards at Great East Road Bridge. (d) 1 to 365-days lead time forecast from 1/2/1988 onwards at Great East Road Bridge.

4. Conclusion

One of the great river basins of the world, the Zambezi river basin, stands out because of its vast unexploited hydropower potential and the valuable wetland areas that can be found along its river network. In the future, challenges brought about by growing populations, developing economies, and climate change will put to the test water resources management experts and the coordination between eight riparian countries. By necessity, equitable water resource allocation decisions will be based on hydrological models and other analysis tools. As in many other river basins, water resources data will play a pivotal role in the calibration and validation of such models.

For years the materialization of an operational centralized water resources database for the basin has remained elusive. The Zambezi Water Information System (ZAMWIS), a decade-old top-down approach to assemble a basin-wide water resources database, is at present under a phase of renewed interest and active development, holding great promise.

The African Dams Project (ADAPT) online water resources database, the subject of this paper, provides a bottom-up complement to ZAMWIS. Relying on open-source technology and frugal with the use of resources, it is based on three key ideas: the total empowerment of stakeholders (as they keep control over the data at all times); the association of valuable analysis tools to the data (providing motivation for its maintenance and continued update); and, an adaptable and simple architecture (allowing customization to particular stakeholders' needs).

Currently under active development, the ADAPT database has caught the attention of local stakeholders, most notably dam operators, NGOs, and consultants. While it has been tailored for the Zambezi river basin, the adaptability of its structure and associated tools allow for its easy application to other river basins as well. ◊

References

1. **The World Bank**, "The Zambezi River basin. A multi-sector investment opportunities analysis," Washington DC, USA; 2010.
2. **COBA, Impacto, and ERM**, "Mphanda Nkuwa hydropower plant," Environmental and social impact assessment (draft), Hidroeléctrica Mphanda Nkuwa; 2011.
3. **Nippon Koei UK**, "Projecto Cahora Bassa central norte. Estudo do impacto sócio-ambiental," Hidroeléctrica de Cahora Bassa; 2012.
4. **SWRSD Zambezi Basin Joint Venture**, "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); 2011.
5. **Arnell, N.W.**, "Climate change and global water resources, Global Environmental Change 9," Supplement 1:S31-S49; 1999.
6. **Intergovernmental Panel on Climate Change (IPCC)**, "Africa Climate Change 2001: Impacts, Adaptation and Vulnerability," Cambridge University Press, UK; 2001.
7. **Beilfuss, R.**, "A risky climate for Southern African hydro-Assessing hydrological risks and consequences for Zambezi River Basin dams," International Rivers, USA; 2012.
8. **ADAPT**, www.cces.ethz.ch/projects/nature/adapt
9. **Mertens, J.**, "African Dams Project: an integrated water resources managament study," Final stakeholder report, ETH Zurich, Switzerland; 2013.
10. **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," *Hydropower & Dams*, Vol. 20, No. 2; 2013.
11. **EPFL**, http://zambezi.epfl.ch
12. **Arnold, J., Moriasi, D., Gassman, P., Abbaspour, K., White, M., Srinivasan, R., Santhi, C., Harmel, R., Van Griensven, A. and Van Liew, M.**, "SWAT: Model use, calibration, and validation," *Transactions of the ASABE*, Vol. 55, No. 4; 2012.
13. **Hannah, D.M., Demuth, S., van Lanen, H.A., Looser, U., Prudhomme, C., Rees, G., Stahl, K. and Tallaksen, L.M.**, "Large-scale river flow archives: importance, current status and future needs," *Hydrological Processes*, Vol. 25, No. 7; 2011.
14. **World Meteorological Organizaion**, "Guide to Hydrological Practices," Hydrology - from measurement to hydrological information, Vol. 1; 2008.
15. **Cohen Liechti, T.**, "Influence of dam operation on water resources management under different scenarios in the Zambezi River Basin considering environmental objectives and hydropower," PhD École Polytechnique Fédérale de Lausanne, Switzerland; 2013.
16. **Cohen Liechti, T., Matos, J.P., Boillat, J.L., Portela, M.M. and Schleiss, A.J.**, "Hydraulic–hydrologic model for water resources management of the Zambezi basin," *Journal of Applied Water Engineering and Research*, Vol. 2, No. 2; 2014.
17. **Matos, J.P.**, "Hydraulic-hydrologic model for the Zambezi River using satellite data and artificial intelligence techniques," PhD École Polytechnique Fédérale de Lausanne, Switzerland and the University of Lisbon, Portugal; 2014.

José Pedro Matos is a Postdoctoral Researcher at the Laboratory of Hydraulic Constructions of the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. His PhD focused on the hydrological modelling of the Zambezi river basin. He has worked as a consultant in hydrology, water supply, and sanitation, with a current focus in the fields of risk assessment, machine learning, remote sensing, and optimization of complex non-linear systems.

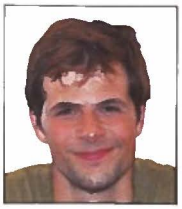
Prof Dr Anton J. Schleiss graduated in Civil Engineering from the Swiss Federal Institute of Technology (ETH) in Zurich, Switzerland, in 1978. After joining the Laboratory of Hydraulic, Hydrology and Glaciology at ETH as a research associate and senior assistant, he obtained a Doctorate of Technical Sciences on the topic of pressure tunnel design in 1986. He then worked for 11 years for Electrowatt Engineering Ltd in Zurich, and was involved in the design of many hydro projects around the world as an expert on hydraulic engineering and underground waterways. In 1997 he was nominated full professor and became Director of the Laboratory of Hydraulic Constructions (LCH) at the Swiss Federal Institute of Technology Lausanne (EPFL). He is also involved as an international expert on several dam and hydro projects throughout the world. From 2006 to 2012 he was Director of the Civil Engineering programme of EPFL and Chairman of the Swiss Committee on Dams. In 2006 he obtained the ASCE Karl Emil Hilgard Hydraulic Prize as well as the J.C. Stevens Award. He was listed in 2011 among the 20 international personalities that "have made the biggest difference to the sector of water power and dam construction over the past 10 years". In 2012, he was elected Vice-President of the International Commission on Large Dams (ICOLD).

École Polytechnique Fédérale de Lausanne (EPFL), Route Cantonale, CH-1015, Lausanne, Switzerland.

Jasmin Mertens is Research Project Coordinator at the Swiss Federal Institute of Technology in Zurich, Switzerland, and a Scientific Collaborator at the Swiss Federal Office of Transport. She holds a PhD in Environmental Sciences, and her interests include water resources management, water quality and, remediation. Dr. Mertens has been the coordinator of the African Dams Project, where she integrated interdisciplinary research efforts, and was involved in stakeholder interaction and outreach activities.

Bernhard Wehrli is Professor of Aquatic Chemistry at the Swiss Federal Institute of Technology (ETH) in Zurich. The research of his group is focused on the cycles of carbon and nutrients in large lakes and rivers. He was the principal investigator of the African Dams Project.

Swiss Federal Institute of Technology (ETH), Universitätstrasse 16, CH-8092 Zurich, Switzerland.



J. P. Matos

A. J. Schleiss

J. Mertens

B. Wehrli