Ecohydrological and Metacommunity Studies of Proliferative Kidney Disease Spread in Freshwater Salmonid Fish

THÈSE N° 8354 (2018)
PRÉSENTÉE LE 8 MAI 2018
À LA FACULTÉ DE L’ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE D’ÉCOHYDROLOGIE
PROGRAMME DOCTORAL EN GÉNIE CIVIL ET ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
POUR L’OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Luca CARRARO

acceptée sur proposition du jury:

Prof. A. Berne, président du jury
Prof. A. Rinaldo, Prof. E. Bertuzzo, directeurs de thèse
Prof. B. Okamura, rapporteuse
Prof. J. Jokela, rapporteur
Prof. T. I. Battin, rapporteur

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
Suisse
2018
All the rivers run into the sea;  
yet the sea is not full;  
unto the place from whence the rivers come,  
thither they return again.  
ECCLESIASTES, 1:7

Ad Alessandro, il tuo ricordo vive
Acknowledgements

These four years have been a fundamental path of growth, both from a professional and a personal viewpoint. This would not have been possible without the guidance of Professor Andrea Rinaldo, to whom my deep gratitude primarily goes. His unwavering enthusiasm and tireless motivation helped me overcome all the issues and doubts I had to face along the way. I am also grateful to Professor Enrico Bertuzzo for his kind, patient and thorough support throughout the progress of this Thesis, always aimed at moving me forward. Moreover, I wish to thank Professors Lorenzo Mari and Marino Gatto, whose help was substantial in the early development stages of the model, and whose prompt assistance was never lacking whenever needed. As this Thesis benefits from a joint effort of blending multidisciplinary expertise, my gratitude also goes to all the collaborators of the Sinergia project, in particular to Hanna Hartikainen, Ínes Fontes, Jukka Jokela, Nicole Strepparava and Thomas Wahli. I gratefully acknowledge financial support from the Swiss National Science Foundation through the Sinergia project CRSII3_147649: "Temperature driven emergence of Proliferative Kidney Disease in salmonid fish - role of ecology, evolution and immunology for aquatic diseases in riverine landscapes".

This Thesis owes much to all the current and former members of the ECHO laboratory, whose friendly and homelike atmosphere allowed me to work and live these years under the best possible conditions. Thanks again to Andrea Rinaldo, who permitted the creation of such a good environment; to Anna Rothenbühler, for untangling the jumble of administration; to Bernard Sperandio, for his eagerness in jumping into water to collect temperature data regardless of the weather conditions; to Damiano Pasetto, for keeping the tradition of the Italian polenta alive; to Flavio Finger, for enriching my French vocabulary about mountaineering and snow; to Javier Perez Saez, for teaching me the importance of paying attention to details, both in science and in people's clothes; to Jonathan Giezendanner, for the countless bike tours and the introduction to the Swiss-German language, culture and cuisine; to Paolo Benettin, for his example as a good scientist and a good scout; to Pierre Queloz, for introducing me to the fascinating mysteries of the lysimeter and the old water paradox; to Silvia Zaoli, for taking (or not understanding?) our continuous jokes. Thanks also to Ana Clara Santos, Andrea Giometto, Joseph Lemaitre and to the people gravitating around the ECHO lab (Francesco Comola, Giulia Castelli, Giulia Ruggeri, Lorenzo Gorla and many more) for the joyful moments spent together.

***
Acknowledgements

Desidero inoltre ringraziare tutti gli amici che, in diversi momenti, mi hanno accompagnato durante questa avventura losannese: Ale, Cecilia, Dona, Giorgia, Giorgio, Leonardo, Niccolò e molti altri. Auguro a tutti di fare fruttare le nostre esperienze svizzere per rendere l’Italia un posto migliore. Grazie a tutte le persone in quel di Piove che, ogni volta che torno, mi accolgono come se non me ne fossi mai andato: tra i tanti, Samu, Chiara (auguri!), Claudio, Poppe, Guido, Nicola e tutti i Santannesi che mi fanno ancora sentire parte di una comunità. Sono profondamente grato ai miei genitori per l’affetto e l’incoraggiamento che non hanno mai fatto mancare. Infine, grazie di cuore a Sofia per essere sempre stata al mio fianco accompagnando questo ulteriore grande passo, rispondendo alle mie inquietudini con amore e pazienza.

Lausanne, 30 Janvier 2018

Luca
PROLIFERATIVE kidney disease (PKD) is a high-mortality pathology affecting freshwater salmonid populations in Europe and North America. Its causative agent, the myxozoan Tetracapsuloides bryosalmonae, has a complex life-cycle exploiting freshwater bryozoans (mainly Fredericella sultana) and salmonids as hosts. PKD has recently increased in incidence and severity, causing remarkable declines in fish catches. In addition, environmental change is feared to cause PKD outbreaks in regions at higher latitude and altitude, as warmer water temperatures exacerbate disease development and fish mortality. In this perspective, this Thesis develops an integrated approach, involving field and modelling work, for the prediction of the incidence of PKD in river basins.

In particular, the Thesis develops a novel spatially-explicit model of PKD epidemiology in riverine host metacommunities. The model, summarizing the current knowledge on disease transmission modes and parasite’s life-cycle, accounts for both local population and disease dynamics of bryozoans and fish, as well as hydrodynamic dispersion of parasite spores and hosts along the river network. Model validation was attained through an integrated study of PKD in a prealpine Swiss river, where data on fish abundance, disease prevalence, concentration of primary hosts’ and parasite’s DNA in environmental samples (eDNA) and water temperatures were gathered at multiple locations within the catchment. In this context, a new method for predicting the spatial distribution of bryozoan density based on eDNA samples was developed. As water temperature is crucial to PKD severity, a deterministic, spatially-explicit water temperature model was formulated and tested on the case-study basin. The model, based on water and energy budgets at the reach scale, considers the effects of the spatial heterogeneity in environmental drivers, allowing the evaluation of gradients of daily mean water temperature along the river network.

Stability and sensitivity analyses performed on the local epidemiological model proved that the introduction of T. bryosalmonae in a disease-free community is very likely to trigger a PKD outbreak. Simulation experiments conducted on synthetic river network replicas showed that network connectivity engenders high PKD prevalence at downstream sites, while the speed of invasion fronts in disease-free environments is amplified by climate change. Moreover, patchily located bryozoan hotspots proved able to sustain the infection at the whole river scale. Validation on the case study confirmed the reliability of the epidemiological model, and identified the prediction of bryozoan distribution as a key direction for future research. In this regard, the developed eDNA transport model possibly opens avenues for the modeling of species distributions in freshwater ecosystems, also beyond the case of PKD. The spatially-explicit temperature model was shown to outperform traditional models based on local heat budgets. Such a tool is instrumental to several ecohydrological applications, from the identification of river reaches at highest PKD-related risk, to the assessment of habitat suitability of fish or other freshwater species. Overall, this Thesis bridges the fields of ecology, epidemiology, hydrometry and mathematical modelling in order to produce an integrated study of PKD, in the perspective of grasping the factors allowing disease persistence and devising mitigation strategies.
Summary

Key words: Disease ecology | Host-parasite interaction | Metacommunity model | Species distribution model | Hydrothermal regime
Résumé

La maladie rénale proliférative (MRP) est une pathologie mortelle affectant les populations de salmonidés d’eau douce en Europe et en Amérique du Nord. Son agent pathogène, le myxozoaire *Tetracapsuloides bryosalmonae*, a un cycle de vie complexe dans lequel les bryozoaires d’eau douce (principalement *Fredericella sultana*) jouent le rôle d’hôtes primaires, et les salmonidés celui d’hôtes secondaires. La fréquence et la sévérité de la MRP ont récemment augmenté, entraînant des diminutions notables des prises de poissons. En outre, il est à craindre que le changement climatique crée des foyers de MRP dans des régions à latitude et altitude plus élevées précédemment non affectées, à cause de l’augmentation de la température de l’eau qui favorise le développement de la maladie et augmente la mortalité des poissons. Dans ce contexte, ce travail développe une approche intégrée combinant des travaux de terrain et de modélisation pour prédire la propagation de la MRP dans les bassins fluviaux.

En particulier, cette thèse vise à construire un modèle mathématique novateur et spatialement explicite de l’épidémiologie de la MRP dans les métacommunautés riveraines. En capitalisant sur les connaissances existantes des modes de transmission de la maladie et du cycle de vie du parasite, le modèle prend en compte à la fois les dynamiques locales des bryozoaires et des poissons ainsi que la dispersion hydrodynamique des spores du parasite et des hôtes le long du réseau fluvial. Pour valider le modèle, une étude intégrée de la MRP a été réalisée dans une rivière préalpine suisse où des données sur l’abondance des poissons, la prévalence de la maladie, la concentration de l’ADN environnemental (ADNe) du parasite ainsi que la température de l’eau ont été collectées à différents endroits du bassin versant. Dans ce contexte, une nouvelle méthode pour prédire la distribution spatiale de la densité de bryozoaires basée sur des échantillons d’ADNe a été développée. Fort du constat que la température de l’eau joue un rôle majeur dans le degré de sévérité de la MRP, un modèle de température de l’eau déterministe et spatialement explicite a été développé et validé. Le modèle, fondé sur des bilans hydriques et énergétiques à l’échelle du tronçon fluvial, considère les effets de l’hétérogénéité spatiale des facteurs environnementaux, tout en permettant d’évaluer les gradients de température moyenne journalière le long du réseau fluvial.

Des analyses de stabilité et de sensibilité du modèle épidémiologique local ont montré que l’introduction de *T. bryosalmonae* dans une communauté saine mène quasi-systématiquement à une épidémie de MRP. Les simulations menées sur des réseaux fluviaux synthétiques (OCNs) ont montré que la connectivité hydrologique génère une prévalence élevée de MRP sur les sites en aval, tandis que la vitesse des fronts d’invasion dans les environnements sans maladie est accélérée par le changement climatique. En outre, il a été démontré qu’un nombre limité de colonies de bryozoaires éparses est capable d’entretenir l’infection dans l’ensemble de la rivière. La validation dans l’étude de cas a confirmé la fiabilité du modèle épidémiologique, tout en identifiant la prédiction de la distribution des bryozoaires comme l’une des directions clés pour la recherche future. À cet égard, le modèle de transport de ADNe développé dans cette thèse ouvre des voies pour la modélisation de la distribution des espèces dans les écosystèmes d’eau douce, aussi au-delà du cas de la MRP. Concernant le modèle de température de l’eau spatialement explicite, les résultats montrent que cette approche surpasse les...
modèles traditionnels fondés sur les budgets énergétiques locaux. Un tel outil est essentiel à plusieurs applications écohydrologiques, de l’identification des tronçons fluviaux au plus fort risque associé à la MRP jusqu’à l’évaluation de la qualité de l’habitat des poissons ou d’autres espèces d’eau douce.

**Mots clefs :** Écologie de la maladie | Interaction hôtes-parasites | Modèle de metacommenauté | Modèle de distribution d’espèces | Régime hydrothermique
Riassunto

La malattia renale proliferativa (PKD) è una patologia ad alta mortalità che colpisce le popolazioni di salmonidi d’acqua dolce in Europa e nel Nord America. Il suo agente causale, il mixozoo Tetracapsuloides bryosalmonae, ha un ciclo di vita complesso che sfrutta i briozoi d’acqua dolce (principalmente Fredericella sultana) come ospiti primari e i salmonidi come ospiti secondari. La PKD ha recentemente aumentato la sua incidenza e severità, causando una notevole diminuzione nel numero di pesci catturati nei fiumi temperati. Inoltre, si teme che i cambiamenti climatici possano causare epidemie di PKD in regioni a latitudine ed altitudine più elevate precedentemente non interessate dalla malattia, in quanto le alte temperature dell’acqua aggravano lo sviluppo della PKD e la mortalità dei pesci. In questa prospettiva, questa tesi sviluppa un approccio integrato che include attività di campo e di modellazione per la previsione dell’incidenza della PKD nei fiumi.

In particolare, questa tesi costruisce un modello spazialmente esplicito di epidemiologia di PKD nelle metacomunità fluviali. Riassumendo le conoscenze sui modi di trasmissione della malattia e il ciclo di vita del parassita, il modello rappresenta sia la dinamica delle popolazioni locali che la trasmissione della malattia nei briozoi e nei pesci, oltre alla dispersione idrodinamica delle spore del parassita e dei briozoi lungo la rete fluviale. Per validare il modello, è stato effettuato uno studio integrato di PKD in un fiume svizzero prealpino dove sono stati raccolti dati spazialmente distribuiti all’interno del bacino sull’abbondanza dei pesci, la prevalenza della malattia, la temperatura dell’acqua e la concentrazione di DNA ambientale (eDNA) di briozoi e parassita. In questo contesto, è stato sviluppato un nuovo metodo per la previsione della distribuzione spaziale della densità dei briozoi basata su campioni di eDNA. Poiché la temperatura dell’acqua è un fattore cruciale nel determinare la severità della PKD, è stato sviluppato un modello di temperatura dell’acqua deterministico e spazialmente esplicito. Il modello, basato su bilanci di massa ed energia a scala di tratto di fiume, considera gli effetti dell’eterogeneità spaziale delle variabili ambientali, permettendo di valutare gradienti di temperatura media giornaliera lungo la rete fluviale.

Analisi di stabilità e sensibilità condotte sul modello epidemiologico locale hanno mostrato che l’introduzione di T. bryosalmonae in una comunità libera da malattia porta in maniera quasi sistematica a un’epidemia di PKD. Esperimenti di simulazione condotti su repliche di reti fluviali sintetiche hanno dimostrato che la connettività di rete implica un’elevata prevalenza di PKD nei siti a valle, mentre la celerità dei fronti di invasione in ambienti liberi da malattia è accelerata dai cambiamenti climatici. Inoltre, è stato dimostrato che colonie di briozoi distribuite in maniera sparsa sono in grado di sostenere l’infezione nell’intera rete fluviale. La validazione effettuata nel caso di studio ha confermato l’affidabilità del modello epidemiologico, identificando nella predizione della distribuzione dei briozoi un’importante direzione per la ricerca futura. A tale scopo, il modello di trasporto di eDNA sviluppato in questa tesi illustra nuove prospettive per la modellazione della distribuzione delle specie in ecosistemi d’acqua dolce, anche al di là del caso della PKD. Infine, si è mostrato che il modello di temperatura dell’acqua spazialmente esplicito è capace di prestazioni superiori rispetto ai modelli tradizionali basati sul bilancio di energia locale. Tale strumento è funzionale a diverse applicazioni ecodidrologiche.
Riassunto

dall’individuazione dei tratti di fiume a maggior rischio legato alla PKD, alla valutazione dell’idoneità degli habitat dei pesci o di altre specie d’acqua dolce.

Parole chiave: Ecologia delle malattie | Interazione ospite-parassita | Modello di metacomunità | Modello di distribuzione di specie | Regime idrotermico
Zusammenfassung


Stabilitäts- und Sensitivitätsanalysen am lokalen Modell zeigten, dass die Einführung von *T. bryosalmonae* in eine krankheitsfreie Gemeinschaft höchstwahrscheinlich Auslöser eines PKD-Ausbruchs wird. Simulationsexperimente, die in synthetischen Flussnetz replikas durchgeführt wurden, zeigten, dass Netzwerkkonnektivität eine hohe PKD-Prävalenz an stromabwärts gelegenen Standorten erzeugt, während die Rate von Invasionsfronten in krankheitsfreien Umgebungen durch den Klimawandel beschleunigt wird. Darüber hinaus wurde gezeigt, dass fleckige Bryozoien-Hotspots Infektionen im gesamten Fluss aufrechterhalten können.

Zusammenfassung

**Stichwörter:** Krankheitsökologie | Wirt-Parasit-Interaktionen | Metacommunity-Modell | Artenverteilungsmodell | Hydrothermales Regime
# Contents

Acknowledgements i  
Summary iii  
Résumé v  
Riassunto vii  
Zusammenfassung ix  
List of figures xv  
List of tables xvii  
List of abbreviations xix  

1 Introduction 1  
1.1 PKD transmission cycle .................................................. 2  
1.2 On epidemiological models: complexity, spatiality, and a metacommunity perspective .................................................. 4  
1.3 Species distribution models and the use of environmental DNA .................................................. 6  
1.4 Stream temperature: a key ecohydrological driver .................................................. 8  
1.5 Aim and structure of the Thesis .................................................. 9  

2 A metacommunity epidemiological model for PKD 11  
2.1 Local PKD model .................................................. 11  
2.1.1 Intra-seasonal model .................................................. 12  
2.1.2 Inter-seasonal model .................................................. 17  
2.2 Spatially-explicit PKD model .................................................. 18  
2.2.1 Fish mobility rates .................................................. 20  
2.3 Spatially-explicit model with age-structure .................................................. 21  
2.4 List of symbols .................................................. 25  

3 Analysis of the local PKD model 27  
3.1 Methods .................................................. 28  
3.1.1 Model simulation .................................................. 28
## Contents

3.1.2 Analysis of the within-season model ........................................ 29  
3.1.3 Stability analysis of the full model ...................................... 31  
3.1.4 Sensitivity analyses ......................................................... 34  
3.2 Results ..................................................................................... 34  
  3.2.1 Model simulation ................................................................. 34  
  3.2.2 Analysis of the within-season model .................................... 36  
  3.2.3 Sensitivity analyses of the full model .................................... 37  
3.3 Discussion .................................................................................. 40  
3.4 List of symbols ........................................................................... 43  

4 PKD spread along river networks .................................................. 45  
  4.1 Methods .................................................................................. 46  
    4.1.1 Generation of synthetic river networks ................................ 46  
    4.1.2 Water temperature patterns .............................................. 49  
    4.1.3 Experiment 1: Connectivity effects on the heterogeneity of PKD impact ......................................................... 50  
    4.1.4 Experiment 2: Speed of disease propagation ....................... 51  
    4.1.5 Experiment 3: Influence of heterogeneity in bryozoan spatial distribution ......................................................... 51  
  4.2 Results ...................................................................................... 52  
  4.3 Discussion ................................................................................ 56  
  4.4 List of symbols ........................................................................ 59  

5 Case study: An integrated analysis of PKD spread in a Swiss prealpine river ......................................................... 61  
  5.1 Methods .................................................................................. 62  
    5.1.1 Study area and hydro-geomorphology ................................ 62  
    5.1.2 Field data collection ........................................................... 64  
    5.1.3 Water temperature monitoring and extrapolation ............... 66  
    5.1.4 Bryozoan habitat suitability .............................................. 67  
    5.1.5 Details on model simulations and calibration ..................... 70  
  5.2 Results ...................................................................................... 72  
  5.3 Discussion ................................................................................ 77  
  5.4 List of symbols ........................................................................ 79  

6 Using eDNA to track species distributions in river networks ......................................................... 81  
  6.1 Methods .................................................................................. 82  
  6.2 Results and discussion .............................................................. 84  
  6.3 eDNA modelling and the width function concept ...................... 85  
  6.4 List of symbols ........................................................................ 87  

7 A stream temperature model for ecohydrological applications ......................................................... 89  
  7.1 Introduction ............................................................................. 90  
  7.2 Methods .................................................................................. 91
Contents

7.2.1 Study area ................................................. 91
7.2.2 Formulation of the spatially-explicit model ................. 92
7.2.3 Spatial interpolation of air temperatures .................. 94
7.2.4 Spatial interpolation of soil temperatures .................. 96
7.2.5 Model parametrizations .................................. 97
7.2.6 Model calibration and comparison ......................... 100
7.3 Results and discussion ..................................... 100
7.4 List of symbols ............................................. 105

Conclusions and perspectives .................................... 109

A Use of survival models to estimate PKD transmission parameters 115

B Detection of F. sultana eDNA in water samples ............... 123

Bibliography .................................................. 127

Curriculum Vitae ............................................. 147
## List of Figures

1.1 A sketch of PKD transmission cycle. ................................. 2

1.1 Schematic representation of local PKD transmission dynamics. ............. 13

2.2 Schematic representation of the spatially-explicit PKD model. .............. 19

2.3 Schematic representation of the full epidemiological model. ................. 23

3.1 Functional forms for temperature-dependent parameters. .................. 29

3.2 Example of a model simulation. .................................... 35

3.3 Sensitivity analysis: values of the reproductive number $R$. ............... 36

3.4 Short-term risk of invasibility. ..................................... 37

3.5 Sensitivity analysis: Parasite invasion. ................................ 38

3.6 Sensitivity analysis: PKD-induced fish loss. ............................ 39

4.1 Example of an OCN and a distribution of mean fish residence times. ...... 48

4.2 Results from Experiment 1. ........................................ 53

4.3 Results from Experiment 2. ........................................ 54

4.4 Results from Experiment 3. ........................................ 55

5.1 Overview of the study area and data. ................................ 62

5.2 Brown trout population sampling data. ................................ 65

5.3 Modelled and observed mean daily water temperatures. .................... 68

5.4 Time series of the calibrated values of the linear regression parameters. .... 69

5.5 Maps of extrapolated mean water temperatures. .......................... 69

5.6 Correlation matrix plot. ............................................. 71

5.7 Results from the bryozoan habitat suitability study. ....................... 74

5.8 Results from the epidemiological metacommunity model. .................. 75

5.9 Maps of modelled fish density. ..................................... 76

5.10 Time evolution of modelled fish population in site #16. ................. 77

5.11 Posterior distributions of the calibrated parameters. ...................... 77

6.1 Overview of the dataset used in the study. ............................. 83

6.2 Results from application of the model on $T. bryosalmonae$ data. ....... 86

6.3 Test of Eq. (6.3) on optimal channel networks. .......................... 87

7.1 Map of the case study region. ...................................... 92
## List of Figures

7.2  Schematic representation of the water and energy budgets at the reach scale. . 93
7.3  Measured vs. modelled daily air temperatures. . . . . . . . . . . . . . . . . . 96
7.4  Measured vs. modelled daily soil temperatures. . . . . . . . . . . . . . . . . . 97
7.5  Daily global solar radiation data. . . . . . . . . . . . . . . . . . . . . . . . . . . . 99
7.6  Modelled vs. observed stream temperatures for all sampling sites. . . . . . . 102
7.7  Modelled vs. observed stream temperature time series in 2016. . . . . . . . . 103
7.8  Comparison among models in 2016. . . . . . . . . . . . . . . . . . . . . . . . . . 103
7.9  RMSE as a function of discharge. . . . . . . . . . . . . . . . . . . . . . . . . . . . 104
A.1  Results from survival models. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 119
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>List of model variables.</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>List of parameters for the local PKD model.</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>List of variables for the epidemiological model with age structure.</td>
<td>22</td>
</tr>
<tr>
<td>3.1</td>
<td>List of parameters involved in the local PKD model.</td>
<td>28</td>
</tr>
<tr>
<td>4.1</td>
<td>Reference set of parameters for the spatial PKD model.</td>
<td>47</td>
</tr>
<tr>
<td>4.2</td>
<td>Scale parameters for generation of OCNs and water temperature time series</td>
<td>50</td>
</tr>
<tr>
<td>5.1</td>
<td>Fish prevalence data.</td>
<td>65</td>
</tr>
<tr>
<td>5.2</td>
<td>List of covariates used for the model of bryozoan habitat suitability.</td>
<td>72</td>
</tr>
<tr>
<td>5.3</td>
<td>Reference set of parameters for the epidemiological model with age-structure.</td>
<td>73</td>
</tr>
<tr>
<td>7.1</td>
<td>Full list of tested models.</td>
<td>98</td>
</tr>
<tr>
<td>7.2</td>
<td>Best-fit parameters and performance indices for all models except the log series.</td>
<td>101</td>
</tr>
<tr>
<td>7.3</td>
<td>Best-fit parameters and performance indices for the log model series.</td>
<td>101</td>
</tr>
<tr>
<td>7.4</td>
<td>Root mean square errors for the calibrated C-series models.</td>
<td>104</td>
</tr>
<tr>
<td>A.1</td>
<td>Survival models. Best fit parameter values and performance scores.</td>
<td>119</td>
</tr>
<tr>
<td>B.1</td>
<td>Oligonucleotide primers and probe sequences for the qPCR assay.</td>
<td>124</td>
</tr>
<tr>
<td>B.2</td>
<td>Limits of quantification and detection for each qPCR assay.</td>
<td>125</td>
</tr>
</tbody>
</table>
List of abbreviations

AIC : Akaike information criterion
BIC : Bayesian information criterion
DFE : Disease-free equilibrium
DFT : Disease-free trajectory
DALR : Dry adiabatic lapse rate
dpe : days post exposure
eDNA : environmental deoxyribonucleic acid
FOEN : (Swiss) Federal office for the environment
GIS : Geographic information system
IPC : Internal positive control
LOD : Limit of detection
LOQ : Limit of quantification
LOOCV : Leave-one-out cross calibration
NS : Nash-Sutcliffe index
NSA : Neutral stability algorithm
OCN : Optimal channel network
PCR : Polymerase chain reaction
PKD : Proliferative kidney disease
RMSE : Root mean square error
qPCR : quantitative PCR
VIF : Variance inflation factor
YOF : Young-of-the-year fish
A major goal of disease ecology is to understand how the environment, hosts and pathogens interact to cause disease outbreaks [Hudson et al., 2002]. Consolidating the knowledge on the ecological and evolutionary interactions underlying disease, and predicting how and where disease outbreaks are likely to occur, remain important and often challenging endeavours. Such predictions are nevertheless crucial in guiding management decisions to prevent the spread and further emergence of major human and wildlife epidemics [Molli-son, 1995]. Indeed, epidemics of emerging diseases may have large economic and ecological impacts, threaten livelihoods, elicit biodiversity losses and affect key ecosystem services. Understanding patterns of disease emergence and their causes becomes even more topical as causative and correlative links with global climate and environmental change have been firmly established [Harvell et al., 2002; Altizer et al., 2013; Raffel et al., 2013]. Disease dynamics are typically driven by factors operating at multiple levels (such as biological, demographic and ecological ones), which interact among each other dynamically in time. Complex system modelling is thus essential to unravel causal relationships [Galea et al., 2010]. On the other hand, disease ecology cannot ignore long-established empirically-based evidence and knowledge. Therefore, mathematical modelling should play a chief, guiding role in comprehensive approaches encompassing fieldwork and laboratory work [Restif et al., 2012].

A paradigmatic example of complex interrelationships between pathogens, hosts and environmental drivers is constituted by proliferative kidney disease (PKD) of salmonid fish, caused by *Tetracapsuloides bryosalmonae* (phylum Cnidaria, class Malacosporea) [Hedrick et al., 1993]. PKD is highly problematic for hatcheries and fish farms, leading up to 100% infection prevalence and high mortalities (up to 90%) in affected fish [Clifton-Hadley et al., 1986; Feist and Longshaw, 2006]. Although the impacts of PKD in wild fish populations are still poorly known, PKD is considered as a key factor contributing to the decline of wild salmonid populations in Switzerland and Northern Europe [Wahl et al., 2002; Burkhardt-Holm et al., 2005; Wahli et al., 2007]. The parasite's life-cycle (see Fig. 1.1) alternates between freshwater bryozoans as primary hosts [Anderson et al., 1999; Longshaw et al., 1999; Okamura et al., 2001], and

1 This Chapter features contents from Carraro et al. [2016, 2017, 2018].
salmonid fish as secondary hosts, where infection can cause PKD [Feist and Bucke, 1993; Hedrick et al., 1993]. Transmission of the parasite between hosts occurs through spores that are released into the aquatic environment. Development and pathology of PKD are dependent on water temperature, as clinical signs and disease-related mortality generally increase with increasing water temperatures [Ferguson and Ball, 1979; Clifton-Hadley et al., 1986; Bettge et al., 2009a,b]. Water temperature is also key to the development of T. bryosalmonae in its freshwater bryozoan host. Increased temperatures have been shown to promote the production of spores infective to fish [Tops et al., 2006, 2009]. The connections between temperature, disease development and fish mortality suggest that, due to climate change, the prevalence and severity of PKD are likely to increase further [Okamura et al., 2011]. Hence, PKD must be considered as an emerging disease having a sizeable and growing impact on the health of salmonid populations and, consequently, on the economy of the fish farming industry and hatcheries. Particular concern stems from the spread of PKD to regions at high latitude or altitude that were previously unaffected. In this regard, a notable increase in PKD incidence in Northern Europe has recently been documented [Skovgaard and Buchmann, 2012; Dash and Vasemägi, 2014; Bruneaux et al., 2017; Debes et al., 2017]. The issue of pinpointing effective mitigation strategies for PKD becomes increasingly critical for environmental policy makers. A noteworthy example is given by the recent PKD outbreak occurred in the Yellowstone river (Montana, USA) [Robbins, 2016], where, due to an unprecedented fish kill, the local wildlife authorities imposed the closure of a 300 km-long river stretch and related tributaries to all recreational activities, to prevent farther spread of the parasite. In the perspective of understanding drivers and controls of the disease, and in the search for possible control measures, the development of a dynamical model of PKD transmission becomes crucial.

### 1.1 PKD transmission cycle

PKD has been detected among salmonid fish throughout Europe and North America. The first comprehensive reviews and surveys about PKD date back to the Eighties [Clifton-Hadley et al., 1984; Smith et al., 1984] although, in retrospective, PKD was probably first observed by Plehn
1.1. PKD transmission cycle

[1924], among trout in Germany [Hedrick et al., 1993]. Kent and Hedrick [1985] discovered that the causative agent of PKD is a myxozoan, initially termed PKX. A major breakthrough in the understanding of PKD transmission cycle occurred in 1999, when Canning et al. [1999] identified freshwater bryozoans as invertebrate hosts. This allowed a full description of the causative agent, and a new myxozoan class (the Malacosporea) was then introduced by Canning et al. [2000] to accommodate myxozoan parasites of freshwater bryozoans.

Freshwater bryozoans (Phylum Bryozoa, Class Phylactolaemata) are suspension-feeding animals usually found in benthic environments such as ponds, lakes and rivers. They grow in colonies attached to sheltered surfaces such as submerged roots of riparian trees, branches or boulders. Growth essentially occurs by incorporation of new modules (zooids) into the body of the colony [Wood and Okamura, 2005; Hartikainen and Okamura, 2015]. High water temperatures and eutrophication are known to play a major role in fostering bryozoan growth [Tops et al., 2009; Hartikainen and Okamura, 2012]. Bryozoans are viviparous hermaphrodites able to reproduce both sexually and asexually. Sexual reproduction is achieved via actively swimming larvae, shed in summer, which can attach to suitable locations and form new colonies. However, larval production is scant in *Fredericella sultana*, one of the most common bryozoan hosts of *T. bryosalmonae* [Wood, 1973]. On the other hand, colony fragmentation enables asexual propagation; furthermore, all phylactolaemates can produce statoblasts, namely asexual, seed-like propagules capable of surviving as dormant stages over winter and hatch new colonies at the beginning of a new season. Several bryozoan species survive through winter only as dormant stages, although the survival of colonies is not excluded [Wood and Okamura, 2005]. Bryozoan populations undergo considerable seasonal fluctuations and their density is usually patchy in both space and time, as typical of a metapopulation [Vernon et al., 1996; Okamura and Freeland, 2001; Hartikainen and Okamura, 2015].

*Tetracapsuloides bryosalmonae* has a complex life-cycle which exploits freshwater bryozoans and salmonids as hosts. Fish-to-fish and bryozoan-to-bryozoan transmissions are hindered [Tops et al., 2004]. Within bryozoans, the parasite can express either covert or overt infection stages. During covert infections, the parasite proliferates as single-cell stages [Canning et al., 2002; Morris and Adams, 2006a]. Covert infections have no effect on host morphology nor on fitness [Hartikainen et al., 2013], but they are mildly virulent, as pointed out by Tops et al. [2009] who found that growth of covertly infected bryozoans during statoblast production is reduced by some 27% with respect to uninfected colonies. The transition to overt infection implies increase in virulence, with the formation of multicellular sacs from which thousands of *T. bryosalmonae* spores (approximately 20 μm in diameter) are released into water. Spores are characterized by two amoeboideal cells and four polar capsules [Canning et al., 1999, 2000]. Parasite transmission from bryozoans to fish can thus take place only during the overt infection stage. On the other hand, covert infections are thought to be crucial in allowing long-term parasite persistence within bryozoan colonies, because observations of covert infection clearance are scarce and partial [Tops et al., 2009; Abd-Elfattah et al., 2014a]. Meiosis during developmental stages identifies bryozoans as primary hosts for *T. bryosalmonae* [Canning et al., 2000]. Overt infection hampers bryozoan growth, whereas covert stages pose low energetic cost to
bryozoans [Tops et al., 2009]. Peaks of overt infection have been observed in late spring and autumn [Tops, 2004]. In particular, overt stages develop when bryozoans are undergoing enhanced growth as a result of high temperatures or food levels [Tops et al., 2006; Hartikainen et al., 2009]. Overt infection also elicits temporary castration [Hartikainen and Okamura, 2012], with a severe reduction in the production of statoblasts. Covertly infected bryozoans can produce infected statoblasts thus allowing vertical transmission of *T. bryosalmonae* [Abd-Elfattah et al., 2014a].

Parasite spores released into water by overtly infected bryozoans infect fish through skin penetration and gills [Feist et al., 2001; Longshaw et al., 2002]. *T. bryosalmonae* subsequently migrates to the fishes’ kidney, where it undergoes multiplication and differentiation from extrasporegonic to sporogenic stages, entailing a massive granulomatous nephritis with vascular necrosis [Kent and Hedrick, 1985; Bettge et al., 2009a]. Spores developed in the lumen of kidney tubules contain one amoebid cell and two polar capsules [Morris and Adams, 2008], and are eventually excreted via urine [Hedrick et al., 2004]. Although PKD seems to develop in all salmonids to a varying degree, life-cycle completion may be highly species-specific. For example, in Europe, parasite spores infective to bryozoans develop in the brown trout (*Salmo trutta*) and the brook trout (*Salvelinus fontinalis*), but not in the rainbow trout (*Oncorhynchus mykiss*) [Morris and Adams, 2006b; Grabner and El-Matbouli, 2008; Abd-Elfattah et al., 2014b]. Fish infected with PKD often die owing to secondary infections [Feist and Bucke, 1993]; however, PKD alone has been shown to cause mortality [Bettge et al., 2009a,b]. When temperatures exceed 15 °C for a considerable period of time, the severity of renal pathology worsens and mortality may reach up to 80-90% [Bettge et al., 2009a; Schmidt-Posthaus et al., 2012; Bruneaux et al., 2017]. Fish that do not die during the acute phase of the infection may become long-term carriers of the parasite. Such carriers are reportedly able to infect *Fredericella sultana* up to 2 years after exposure [Abd-Elfattah et al., 2014b].

### 1.2 On epidemiological models: complexity, spatiality, and a meta-community perspective

Epidemiological models incorporating parasites with simple life-cycles have a long history [Anderson and May, 1991; Keeling and Rohani, 2007] and have been successfully applied to several human diseases and zoonoses, e.g. cholera [Capasso and Paveri-Fontana, 1979; Torres Codeço, 2001] and West Nile fever [Wonham et al., 2004; Bowman et al., 2005]. Conversely, models addressing parasites with complex life-cycles and multiple hosts such as *T. bryosalmonae* are less frequently developed, although many major human and wildlife diseases are caused by pathogens falling in this category. Successful application of epidemiological models for these diseases, including schistosomiasis [Gurarie and Seto, 2009] and malaria (see Koella [1991] for a review) fundamentally relies on the incorporation of the ecological dynamics of the different host populations. No epidemiological models for PKD embodying the complex life-cycle of its causative agent have been published prior to Carraro et al. [2016]. The only previous modeling attempt used a Bayesian probability network to assess the decline of brown
1.2. On epidemiological models: complexity, spatiality, and a metacommunity perspective


THE COMPLEXITY IN FORMULATING AN EPIDEMIOLOGICAL MODEL FOR DISEASES SUCH AS PKD ARISES NOT ONLY OWING TO THE INCLUSION OF THE ECOSYSTEMIC DYNAMICS OF BOTH HOSTS, BUT ALSO WITH REGARDS TO THE MANIFEST SPATIALLY HETEROGENEOUS CHARACTER OF ITS TRANSMISSION MODES. INDEED, T. BRYOSALMONAE SPORES ARE ADVECTED DOWNSTREAM BY STREAMFLOW, AND HENCE AN INFECTED BRYOZOAN COLONY GROWING AT A GIVEN SITE MIGHT BE RESPONSIBLE FOR THE INFECTION OF SALMONIDS FOR A CONSIDERABLE LENGTH OF FLUVIAL STRETCH ALONG THE FLOW DIRECTION. MOREOVER, MOVEMENTS OF INFECTED FISH INDUCED BY SPawning OR SEARCH FOR FOOD MIGHT CAUSE BRYOZOAN INFECTIONS IN PORTIONS OF THE RIVER NETWORK PREVIOUSLY UNAFFECTED BY THE DISEASE. BRYOZOAN COLONIES ARE TYPICALLY NOT MOBILE, ALTHOUGH PATHOGEN SPREAD CAN BE ACHIEVED BY COLONY FRAGMENTATION AND DRIFT OR STATOBLAST DISPERAL. IT IS NOTEWORTHY THAT ALL THESE SPATIAL DYNAMICS ARE STRONGLY AFFECTED BY THE HYDROLOGICAL AND THERMAL REGIME OF RIVERS, CHARACTERIZED BY TEMPORAL FLUCTUATIONS AND SPATIAL HETEROGENEITY. HENCE, UNDERSTANDING THE INTERACTIONS BETWEEN SPATIALLY DISTRIBUTED LOCAL COMMUNITIES AND ENVIRONMENTAL FACTORS IS CRUCIAL TO POSSIBLY PREDICT THE SPREAD OF A PKD EPIDEMIC AND TO ASSESS THE EFFECTS OF INTERVENTION STRATEGIES.


of the global carbon cycle [Battin et al., 2008], and on stream biofilm [Battin et al., 2016]. Moreover, fluvial landscapes govern species biodiversity and community dynamics [Carrara et al., 2012; Altermatt, 2013; Carrara et al., 2014].

In the case of PKD, river network connectivity regulates the spatial structure of salmonid and bryozoan communities which, in turn, plays a remarkable role in determining PKD dynamics at a regional scale. It is thus convenient to formulate the epidemiological model under the framework of metacommunity ecology, which addresses sets of local communities connected by dispersal of multiple interacting species [Leibold et al., 2004; Wilson, 1992]. In lotic systems, the prevailing mechanism describing metacommunity structure is species sorting (i.e. species abundances are mostly influenced by environmental suitability) [Heino et al., 2015]. However, the limited knowledge on the dispersal abilities of aquatic organisms limits the understanding of the role played by mass effect (i.e. source-sink dynamics fuelled by dispersal) on the structure of ecological communities. This is particularly relevant to the ecology of freshwater bryozoans [Bilton et al., 2001], which have a strong capacity to expand clonally via colony growth, fragmentation and statoblasts. Moreover, dispersal alters the probability of extinction of local bryozoan populations by increasing genetic diversity [Freeland et al., 2000]. The metacommunity approach also offers tools for understanding key local and dispersal effects governing parasite distributions, infection processes and host-parasite interactions occurring at multiple levels in spatially heterogeneous environments [Johnson et al., 2015; Suzán et al., 2015]. In the light of these considerations, a metacommunity framework (coupling local dynamics influenced by the heterogeneity of the environmental matrix and spatial dynamics governed by dispersal) represents a befitting tool for the study of PKD spread.

1.3 Species distribution models and the use of environmental DNA

Reliable predictions of the spread of infectious diseases and possible management strategies must be based on accurate assessments of the spatial distribution of the parasite host [Alexander et al., 2016]. Species distribution models [Guisan and Zimmermann, 2000; Guisan et al., 2017] constitute a fundamental branch of theoretical ecology, providing a variety of tools capable of interpreting spatially scattered and temporally distributed data on presence or absence of a target species. The choice of the appropriate model must rely on the type of available data, the degree of generality, and on the spatial and temporal scales on which habitat prediction is performed. In addition to several ecological applications, species distribution modeling also serves as a tool for the evaluation of the effect of human anthropization and climate change on organism distribution (e.g. Guisan and Theurillat [2000]). With regards to the case of PKD, assessing the spatial distribution of *Fredericella sultana* colonies, as well as the morphological and environmental variables affecting it, is pivotal to the prediction of disease prevalence and severity patterns at a catchment scale, as bryozoan density ultimately controls the force of infection of PKD.

To this end, a relevant contribution can be provided by the use of environmental DNA (eDNA).
eDNA is a widely used technique enabling the detection of rare or invasive species in freshwater environments [Thomsen and Willerslev, 2015]. Today, eDNA methods are used to measure upstream presence or absence of target species, with obvious asymmetries between them, as absence might be an artifact of sampling bias. Conversely, quantitative prediction of species location and biomass remains an unresolved issue. Therefore, a fundamental step to achieve this goal is the development of a spatially-explicit framework for interpreting eDNA data in rivers.

Applications using eDNA have rapidly gained popularity in ecology over the last decade, mostly because eDNA data potentially improves our quantitative understanding of trophic structure, species interactions, and dynamics of species distribution. eDNA-based applications are highly attractive as measuring biodiversity in a replicable and consistent manner remains challenging with traditional methods [Yoccoz, 2012]. In particular in aquatic environments, eDNA detection of rare, endangered and invasive species clearly outperforms traditional methods such as visual surveys or electrofishing [Ficetola et al., 2008; Jerde et al., 2011; Dejean et al., 2012; Mächler et al., 2014]. Traditional end-point polymerase chain reaction (PCR) techniques have been extensively employed to document presence or absence of target species, whereas quantitative PCR (qPCR) also allows to estimate species’ biomass, which is a crucial parameter in ecosystem studies [Takahara et al., 2012]. The use of qPCR in eDNA analyses has relevant applications in epidemiology and disease ecology, as the quantification of parasite loads [Huver et al., 2015] is key to estimating the force of infection and, ultimately, to capture dynamics of disease spread.

In flowing waters, due to downstream transport, eDNA may undergo decay, retention and resuspension, and the magnitude of these effects is highly affected by flow regime and substrate type [Deiner and Altermatt, 2014; Shogren et al., 2017]. Moreover, being constituted by a series of variously-sized organic sources (cells, tissues, feces, etc.), eDNA is polydisperse [Wilcox et al., 2015]. As a result, while it is straightforward to link a positive PCR test with the presence of the target species at some distance upstream, an effective understanding of the eDNA transport dynamics along a river stretch remains a crucial, yet poorly investigated issue [Jerde et al., 2016]. Further complications arise when eDNA is sampled at a cross-section of a dendritic basin, as in this case the measured eDNA concentration is the outcome of decay processes from the heterogeneously distributed species densities at any point of the drainage network along all downstream paths to the measuring site, which in turn may be characterized by heterogeneous hydro-morphological conditions. Hence, source areas of eDNA cannot be disentangled unless a hydrologically-based, spatially-explicit approach is used. Such a tool is still lacking to date, despite its obvious importance for improving our knowledge on freshwater species distribution and waterborne disease transmission.
Chapter 1. Introduction

1.4 Stream temperature: a key ecohydrological driver

As previously mentioned, water temperature critically controls the incidence and severity of PKD. Besides *T. bryosalmonae*, other pathogens of salmonids whose virulence is temperature-dependent are *Aeromanas salmonicida*, causing furunculosis [Nordmo and Ramstad, 1999], and sea lice *Lepeophtheirus salmonis* and *Caligus elongatus* [Boxaspen, 2006]. From a broader perspective, water temperature is a crucial variable in stream ecology governing the overall health of aquatic ecosystems. Stream temperature heavily impacts the metabolic balance of fluvial ecosystems [Yvon-Durocher et al., 2010; Demars et al., 2011] and determines the metabolic rates of organisms, their distribution along the river network and their interaction with other species [Allan and Castillo, 2007]. Organisms require specific temperature ranges to survive and proliferate, and the excess of critical temperature thresholds might have lethal effects [Lee and Rinne, 1980; Nielsen et al., 1994; Eaton et al., 1995]. Moreover, temperature exerts an effect on species geographical distribution [Parmesan and Yohe, 2003]. For instance, ectothermic animals are remarkably impacted by water temperature, which strongly influences their rates of biochemical reaction, development and growth [Angilletta et al., 2002; Jonsson and Jonsson, 2009].

In particular, the ecology of freshwater salmonids is highly impacted by stream temperature. Indeed, temperature represents one of the abiotic variables influencing habitat suitability, although it is argued that other factors such as water depth, current, substrate and cover might be more important than temperature in controlling the distribution and abundance of salmonids [Heggenes, 1990; Armstrong et al., 2003]. However, temperature plays a major role in determining mobility patterns of fish: salmon and trout perform behavioural thermoregulation by occupying cooler upstream reaches when temperatures rise above their upper tolerances [Theurer et al., 1985; Berman and Quinn, 1991]. Moreover, warm water temperatures are responsible for the shrinkage in seasonal migration patterns [Berman and Quinn, 1991], and for the fragmentation of watershed-wide fish populations by isolating suitable thermal habitats [Matthews and Zimmerman, 1990; Ebersole et al., 2001]. These zones, sheltering biotic communities from thermal disturbances, are referred to as thermal refugia [Sedell et al., 1990; Torgersen et al., 1999].

In light of the above considerations, great concern stems from perspective climate change, especially in view of possible local extinctions of iconic, endemic and/or commercially important fish species. Climate change affects freshwater ecosystems not only through increases in stream temperatures, but also via alterations of streamflow distributions. Implications for salmonids include northwards shift of thermal niches, modification of growth rates, increased mortality and alterations in timing for spawning and hatching (see Jonsson and Jonsson [2009] for a review focused on anadromous salmonids).

A good understanding of the river thermal processes and of the natural and anthropogenic factors controlling water temperature fluctuations is thus fundamental. As a matter of fact, a wide number of stream temperature models exists in the literature, spanning a variety of
1.5 Aim and structure of the Thesis

Lying at the interface between ecology, epidemiology and ecohydrology, this Thesis proposes and analyses a metacommunity framework for the study of PKD in riverine salmonid populations. The approach here proposed embeds spatially-explicit population dynamics of both hosts, epidemiological modelling of disease transmission processes, and the evaluation of ecohydrological drivers affecting PKD impact and propagation, such as river hydrothermal regimes and eDNA-based modelling of density patterns of freshwater bryozoans. This metacommunity framework, constituted by multiple and intertwined layers, enables to achieve a better understanding of PKD transmission modes in local communities and in river networks, and establishes a sound basis for designing effective mitigation policies. Nonetheless, as PKD is an epitome of complex systems in freshwater environments, the tools and concepts developed in this Thesis also contribute to address general questions in the field of ecohydrology, such as the issue of localizing and quantifying the biomass of a specific target species from downstream measurement of its eDNA in a river network, the assessment of fish habitat suitability, and the characterization of temperature gradients in rivers.

The Thesis is structured as follows. Chapter 2 introduces the metacommunity PKD model in three proposed formulations, depending on the extent of the spatial scale (local vs. spatial) and the inclusion of the age-structure of the fish population. These three variants constitute the specific focus of the three following Chapters. Chapter 3 is dedicated to a theoretical study of the local PKD model by means of simulations, stability and sensitivity analyses, in a bid to identify which disease dynamics (subsumed by corresponding model parameters) exert the highest control on disease prevalence, mortality and invasibility. A spatially-explicit version of the PKD model is analyzed in Chapter 4, where simulation experiments conducted on synthetic river networks allow the determination of the effects of network connectivity, climate change and distribution of bryozoans on spatial patterns of PKD prevalence. In Chapter 5, the complete formulation of the metacommunity model is tested on a case study river (Wigger, Switzerland), where spatially distributed data on prevalence in brown trout, presence of *F. sultana* and *T. bryosalmonae* in environmental samples and water temperatures

---

2Namely the capacity of the pathogen to invade and establish in disease-free communities.
Chapter 1. Introduction

have been collected in the period 2014-2016. Moving from the application introduced in this Chapter concerning the derivation of a bryozoan suitability map based on eDNA measurements, Chapter 6 presents perspectives on the modelling of eDNA data from water samples in rivers to formulate species distribution models. Chapter 7 is devoted to the formulation of a deterministic, spatially-explicit stream temperature model suitable for ecohydrological applications, where the determination of within-river temperature gradients cannot be disregarded. Such a model is tested on temperature data gathered at the case-study river. A discussion on conclusions and perspectives closes the Thesis.


2 A metacommunity epidemiological model for PKD

This Chapter deals with the development of a dynamical metacommunity model of PKD epidemiology in riverine host populations. This model, in its different formulations, constitutes the ground upon which the analyses proposed in the following Chapters are based. At a local scale, the model accounts for demographic and epidemiological dynamics of bryozoans and fish, explicitly incorporating the role of temperature, and coupling intra-seasonal and inter-seasonal dynamics. The former are described in a continuous-time domain, the latter in a discrete-time domain. This first formulation, representative of mean-field conditions, focuses solely on the temporal dynamics of the process, while neglecting spatial gradients. At the network scale, instead, the model couples the dynamics of each local community through hydrological transport of parasite spores and fish movement. Heterogeneity in habitat characteristics and hydrological conditions along a river network are also explicitly accounted for. The spatially-explicit epidemiological model is finally extended by including the age-structure of the fish population, as well as its spawning migration dynamics.

2.1 Local PKD model

The proposed model couples PKD transmission and population dynamics of fish and bryozoan populations. When reproduction processes are concentrated in time (as in the case of PKD, in which spawning and hatching for fish, and statoblast hatch for bryozoans occur mainly during the cold season), population dynamics are traditionally analyzed with discrete-time models (see the widely known Ricker model [Ricker, 1954], introduced to study stock and recruitment in fisheries). On the other hand, PKD transmission occurs continuously throughout the warm season, thus calling for a continuous modelling approach (e.g. a classical Susceptible-
Infected-Recovered model [Anderson and May, 1979]). Therefore, a discrete-continuous hybrid framework that couples intra-seasonal and inter-seasonal dynamics is here proposed. Seasons can be thought of as the periods of bryozoan proliferation (e.g. from April to November, depending on climate). Within-season dynamics are described by a set of coupled ordinary differential equations expressing how bryozoan biomass, fish density, and the abundances of infective spores and statoblasts change during the warm season. The transition between the end of a warm season and the beginning of the next one is modelled as a discrete-time update. Therefore, between-season dynamics is described by a set of difference equations. An interesting framework to formalize discrete-continuous hybrid models is the so-called time scale calculus [Agarwal et al., 2002]; however, the complexity of the model presented in the following paragraphs prevents an effective use of such formalism. In this section, the focus is limited to a local-scale model, able to mimic PKD dynamics in an isolated water body, where fish and bryozoans are well mixed, and there are no additional inputs or outputs. An outline of the model is shown in Fig. 2.1. The variables of the model are listed in Table 2.1; all state variables are referred to as concentrations.

### 2.1.1 Intra-seasonal model

The dynamics of the state variables during the warm season (from time $\tau_0$ to $\tau_1$) of year $y$ (see Fig. 2.1c) is described by the following system of first-order differential equations:

\[
\frac{dB_S}{d\tau} = g_S(B_S, B_C, B_O; T)B_S - \beta_B Z_F B_S + \psi B_C; \tag{2.1a}
\]

\[
\frac{dB_C}{d\tau} = g_C(B_S, B_C, B_O; T)B_C + \beta_B Z_F B_S - [d_{CO}(T) + \psi] B_C + d_{OC}(T)B_O; \tag{2.1b}
\]

\[
\frac{dB_O}{d\tau} = g_O(B_S, B_C, B_O; T)B_O + d_{CO}(T)B_C - d_{OC}(T)B_O; \tag{2.1c}
\]

\[
\frac{dS_S}{d\tau} = f_B(\tau, T)B_S + \phi f_B(\tau, T)B_C; \tag{2.1d}
\]
2.1. Local PKD model

\[
\frac{dS_I}{d\tau} = (1 - \phi) f_B(\tau, T) B_C; \quad (2.1e)
\]
\[
\frac{dF_S}{d\tau} = -\mu_F F_S - \beta_B Z_B F_S + \xi F_C; \quad (2.1f)
\]
\[
\frac{dF_E}{d\tau} = \beta_B Z_B F_S - (\mu_F + h(T)) F_E; \quad (2.1g)
\]
\[
\frac{dF_I}{d\tau} = (1 - \epsilon) h(T) F_E - (\mu_F + a(T) + \gamma) F_I; \quad (2.1h)
\]
\[
\frac{dF_C}{d\tau} = \epsilon h(T) F_E - (\mu_F + \zeta) F_C; \quad (2.1i)
\]
\[
\frac{dZ_B}{d\tau} = \pi_B B_O - \mu_Z Z_B; \quad (2.1j)
\]
\[
\frac{dZ_F}{d\tau} = \pi_F (F_I + \kappa F_C) - \mu_Z Z_F, \quad (2.1k)
\]
where the dependence of the parameters on temperature $T$ and time $\tau$ has been explicitly expressed.

The first terms in the right-hand sides of Eqs. (2.1a), (2.1b) and (2.1c) express the growth of the bryozoan biomass, which is assumed to be logistic-like. Thus one can set

$$g_X(B_S, B_C, B_O; T) = r_X(T)[1 - \rho(B_S + B_C + B_O)], \quad (2.2)$$

where $r_X(T)$ and $\rho$ are the baseline instantaneous growth rate of class $X = \{S, C, O\}$ and the inverse of the carrying capacity of the bryozoan population, respectively. It is assumed that susceptible and covertly infected bryozoans have the same baseline growth rate $r_S = r_C = r$, whereas $r_O \ll r$ since the growth of overtly infected bryozoans is strongly impaired [Hartikainen and Okamura, 2012]. Both $r$ and $r_O$ are assumed to be monotonically increasing functions of temperature, as experimental evidence suggests [Tops et al., 2009]. The term $\beta_B B_S Z_F$ in Eqs. (2.1a) and (2.1b) represents the flux of bryozoan biomass from the susceptible to the covertly infected class, with $\beta_B$ being the exposure rate of bryozoans to fish-released spores. The transition from covert to overt infection is assumed to occur at a rate $d_{CO}$ [see Eqs. (2.1b) and (2.1c)], which is defined as the inverse of the mean time necessary for the development of the overt stage of infection in a previously covertly infected bryozoan unit. The rate of transition from overt to covert infection is expressed by $d_{OC}$. Both parameters are temperature-dependent: in particular $d_{CO}$ increases with increasing temperature, while $d_{OC}$ decreases [Tops et al., 2006]. As previously stated, the development of an overt infection, which poses a high energetic cost to the bryozoan colony, also depends on host conditions and food availability. However, as a first approximation, temperature is considered as the only determinant of change of the two transition rates. It is hypothesized that covertly infected bryozoans can possibly clear the infection and become again susceptible at a rate $\psi$ [Eqs. (2.1a) and (2.1b)]. The production of statoblasts is only achieved by susceptible and covertly infected bryozoans, whereas overtly infected colonies do not produce statoblasts. In particular, susceptible bryozoans produce uninfected statoblasts $S_S$ [Eq. (2.1d)] at a rate $f_B$ that is assumed to depend on both temperature and time; specifically, the release of statoblasts is enhanced towards the end of the season. Uninfected bryozoans have been observed by Abd-Elfattah et al. [2014a] while producing infected statoblasts as well. However, these authors argued that their observation could be interpreted as a sign of a recent loss of infection. Therefore, for the sake of simplicity, this possibility is neglected. Instead, covertly infected bryozoans can produce both infected $S_I$ [Eq. (2.1e)] and uninfected statoblasts [Abd-Elfattah et al., 2014a]. The probability that infected bryozoans produce uninfected statoblasts is termed $\phi$. The total statoblast production rate of infected bryozoans is assumed to be equal to that of susceptible ones. $T. bryosalmonae$ spores [Eq. (2.1j)] are produced by overtly infected bryozoans at a constant rate $\pi_B$. A constant mortality rate for spores ($\mu_Z$, i.e. the inverse of the time span during which spores are viable) is also accounted for.

As for the fish population, no recruitment or immigration are considered to take place during the warm season. Therefore, the abundance of susceptible fish [Eq. (2.1f)] decays monotonically...
### Table 2.2 – List of parameters for the local PKD model. Note that Greek letters refer to constant parameters, while Latin letters indicate temperature-dependent parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bryozoans</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>Inverse of carrying capacity</td>
<td>$M^{-1}L^{-3}$</td>
</tr>
<tr>
<td>$\beta_B$</td>
<td>Exposure rate</td>
<td>$L^{3}T^{-1}$</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Fraction of $B_C$ generating uninfected statoblasts</td>
<td>-</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Rate of recovery from covert infection</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$\pi_B$</td>
<td>Rate of contamination operated by $B_O$</td>
<td>$M^{-1}T^{-1}$</td>
</tr>
<tr>
<td>$\sigma_B$</td>
<td>Probability of survival over winter for $B_S, B_C$</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma_O$</td>
<td>Probability of survival over winter for $B_O$</td>
<td>-</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Biomass generated by one statoblast</td>
<td>$M$</td>
</tr>
<tr>
<td>Temperature-dependent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>Baseline growth rate of $B_S, B_C$</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$r_O$</td>
<td>Baseline growth rate of $B_O$</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$d_{CO}$</td>
<td>Rate of covert-to-overt transition</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$d_{OC}$</td>
<td>Rate of overt-to-covert transition</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>Time and temperature-dependent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_B$</td>
<td>Statoblast production rate</td>
<td>$M^{-1}T^{-1}$</td>
</tr>
<tr>
<td><strong>Fish</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_F$</td>
<td>Natural mortality rate</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$\beta_F$</td>
<td>Exposure rate</td>
<td>$L^{3}T^{-1}$</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Fraction of acute infections</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Rate of recovery from acute infection</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Rate of total recovery</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$\pi_F$</td>
<td>Rate of contamination operated by $F_I$</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Relative rate of contamination operated by $F_C$</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma_F$</td>
<td>Probability of survival over winter</td>
<td>-</td>
</tr>
<tr>
<td>$\eta$</td>
<td>Baseline reproduction rate in Ricker model</td>
<td>-</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Parameter for density dependence in Ricker model</td>
<td>$L^{-3}$</td>
</tr>
<tr>
<td>Temperature-dependent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h$</td>
<td>Rate of development of the disease</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$a$</td>
<td>PKD-caused mortality rate</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td><strong>Tetracapsuloides bryosalmonae</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_Z$</td>
<td>Spore decay rate</td>
<td>$T^{-1}$</td>
</tr>
</tbody>
</table>

Cally throughout the season, due to natural mortality (at a rate $\mu_F$, defined as the inverse of the average lifespan of a fish) and infection from bryozoan-released spores $Z_B$. The term $\beta_F F_S Z_B$ in Eqs. (2.1f) and (2.1g) represents the flux of fish from the susceptible to the exposed compartment, with $\beta_F$ being the exposure rate of fish to bryozoan-released spores. Exposed fish contracted the disease but are not yet infective. The decrease in the number of these fish, besides natural mortality, is ruled by the temperature-dependent rate $h$ [Eq. (2.1g)]; namely, $h^{-1}$ is the average time for the development of the disease in fish. The parameter $h$ increases with increasing temperature [El-Matbouli and Hoffman, 1994]. It is assumed that a fraction $\epsilon$ of fish exiting from the exposed class does not show clinical symptoms and is not subject to PKD-related mortality, thus directly entering the carrier class. The remaining part $(1 - \epsilon)$
Chapter 2. A metacommunity epidemiological model for PKD

becomes infected. The abundance of infected fish [Eq. (2.1h)] decreases because of both natural and PKD-caused mortality. The latter is assumed to occur at a rate $\alpha$ that is positively correlated with temperature [Bettge et al., 2009a,b]. In addition, infected fish can enter the carrier compartment at a rate $\gamma$ [Eq. (2.1i)], which corresponds to the inverse of the average duration of the acute phase of the infection. Fish infected by the parasite that did not die in the first year can continue to release infective spores up to two years [Abd-Elfattah et al., 2014b]. However, field and experimental observations on the duration of this carrier stage are still scarce and further experiments are underway (see Strepparava et al. [2017] and Appendix A). Therefore, the possibility that carrier fish may recover is explored by introducing a recovery rate $\zeta$. Evidence of possible immunity are scant (e.g. immunity to second infection in rainbow trout has been reported by Foott and Hedrick [1987]), thus, as a safe assumption, recovered fish are assumed to become susceptible again [Eq. (2.1f)]. Spores [Eq. (2.1k)] are released by both infected and carrier fish at rates $\pi_F$ and $\kappa \pi_F$ respectively, with $\kappa$ being an appropriate coefficient ranging from 0 to 1. Spore mortality is accounted for through the parameter $\mu_Z$.

Since the dynamics of disease transmission between bryozoan and fish is driven by the product between the concentration of spores ($Z_B$ and $Z_F$) and the rates ($\beta_F$ and $\beta_B$) at which susceptible organisms are actually exposed to the infectious agents, two new state variables $Z_B^* = \beta_F Z_B$ and $Z_F^* = \beta_B Z_F$ can be introduced. These new quantities are termed equivalent spores: namely, $Z_F^*$ ($Z_B^*$) is the concentration of spores needed to infect a unit concentration of susceptible bryozoan biomass (a single susceptible fish in a unit water volume) per unit time. With this definition, the exposure rates $\beta_F$ and $\beta_B$ can be discarded and two synthetic contamination rates $\pi_B^* = \beta_F \pi_B$ [L$^3$M$^{-1}$T$^{-2}$] and $\pi_F^* = \beta_B \pi_F$ [L$^3$T$^{-2}$] are introduced. Note that both exposure and contamination rates as defined in model (2.1) are hardly measurable and would most likely need to be calibrated by contrasting model simulations against experimental or field data. The new parameter definitions thus reduce the number of parameters of the model and make the comparison between data and model predictions easier and more robust.

The new set of equations accounting for the rescaled state variables therefore reads:

\[ \frac{dS}{d\tau} = g_S(B_S, B_C, B_O; T) B_S - Z_F^* B_S; \]  
\[ \frac{dC}{d\tau} = g_C(B_S, B_C, B_O; T) B_C + Z_F^* B_S - [d_{CO}(T) + \psi] B_C + d_{OC}(T) B_O; \]  
\[ \frac{dS}{d\tau} = -\mu_F F_S - Z_F^* F_S + \zeta F_C; \]  
\[ \frac{dE}{d\tau} = Z_F^* F_S - [\mu_F + h(T)] F_E; \]  
\[ \frac{dZ_B^*}{d\tau} = \pi_B^* B_O - \mu_Z Z_B^*; \]  
\[ \frac{dZ_F^*}{d\tau} = \pi_F^*(F_I + \kappa F_C) - \mu_Z Z_F^*. \]  

These equations, coupled with Eqs. (2.1c), (2.1d), (2.1e), (2.1h) and (2.1i) constitute the intra-seasonal model hereafter applied.
2.1.2 Inter-seasonal model

The following difference equation system relates the state of the model variables at the end of a season \((y; \tau_1)\) with that at the beginning of the following season \((y+1; \tau_0)\).

\[
B_S(y+1; \tau_0) = \sigma_B B_S(y; \tau_1) + \nu S_S(y; \tau_1); \quad (2.4a)
\]
\[
B_C(y+1; \tau_0) = \sigma_B B_C(y; \tau_1) + \sigma_O B_O(y; \tau_1) + \nu S_I(y; \tau_1); \quad (2.4b)
\]
\[
B_O(y+1; \tau_0) = 0; \quad (2.4c)
\]
\[
S_S(y+1; \tau_0) = 0; \quad (2.4d)
\]
\[
S_I(y+1; \tau_0) = 0; \quad (2.4e)
\]
\[
F_S(y+1; \tau_0) = \sigma_F F_S(y; \tau_1) + f_F(F_S(y; \tau_1), F_E(y; \tau_1), F_I(y; \tau_1), F_C(y; \tau_1)); \quad (2.4f)
\]
\[
F_E(y+1; \tau_0) = 0; \quad (2.4g)
\]
\[
F_I(y+1; \tau_0) = 0; \quad (2.4h)
\]
\[
F_C(y+1; \tau_0) = \sigma_F \left[ p_E F_E(y; \tau_1) + p_I F_I(y; \tau_1) + F_C(y; \tau_1) \right]; \quad (2.4i)
\]
\[
Z_B(y+1; \tau_0) = 0; \quad (2.4j)
\]
\[
Z_F(y+1; \tau_0) = 0. \quad (2.4k)
\]

At the beginning of a new season \((y+1; \tau_0)\), no overtly infected bryozoans are present [Eq. (2.4c)]; the same condition applies for statoblasts [Eqs. (2.4d) and (2.4e)] and spores [Eq. (2.4j) and (2.4k)]. Statoblasts surviving for more than one year are here neglected for the sake of simplicity. The biomass of susceptible bryozoans [Eq. (2.4a)] is composed of a fraction \(\sigma_B\) of the susceptible bryozoan biomass at the end of the previous season \((y; \tau_1)\) that managed to survive over winter, and of newly hatched colonies from the uninfected statoblasts released during the previous season. The parameter \(\nu\) is defined as the mean amount of bryozoan biomass produced by a single statoblast.

The population of covertly infected bryozoan at the beginning of a new season [Eq. (2.4b)] is given by the sum of three contributions: the fraction of the covertly infected biomass at the end of the previous season that survives over winter (with probability \(\sigma_B\)), the fraction of overtly infected bryozoans at time \((y; \tau_1)\) with survival probability \(\sigma_O\), the newly established colonies generated by infected statoblasts, according to the parameter \(\nu\).

As for fish, it is assumed that a fraction \(\sigma_F\) of susceptible and carrier organisms at the end of the previous season survives over winter [Eqs. (2.4f) and (2.4i)]. Exposed and infected fish either die or enter the carrier class. This is modelled by computing additional coefficients \(p_E, p_I\) accounting for natural and PKD-induced deaths:

\[
p_I = \frac{\gamma}{\mu_F + \hat{\alpha} + \gamma}; \quad p_E = \frac{
hat{h}}{
\mu_F + \hat{h}} \left[ \epsilon + (1 - \epsilon) \frac{\gamma}{\mu_F + \hat{\alpha} + \gamma} \right]. \quad (2.5)
\]

Specifically, \(p_E (p_I)\) is the probability that an exposed (infected) fish survives and enters the carrier class in the first period of the winter season. In Eq. (2.5) \(\hat{\alpha}\) and \(\hat{h}\) are the rates of
Chapter 2. A metacommunity epidemiological model for PKD

PKD-caused mortality and disease development averaged over the first period of the winter season, respectively. For the sake of simplicity, it is assumed \( \hat{a} = a(\hat{T}) \) and \( \hat{h} = h(\hat{T}) \), with \( \hat{T} \) being a representative value of temperature over the considered period.

Fish reproduction occurs during winter; the amount of newborn uninfected fish (termed \( f_F \)) is assumed to depend on the total fish population at \( (y; \tau_1) \):

\[
f_F(\mathcal{F}) = \eta \mathcal{F} \exp(-\xi \mathcal{F}), \tag{2.6}
\]

where \( \mathcal{F} = F_S(y, \tau_1) + p_E F_E(y, \tau_1) + p_I F_I(y, \tau_1) + F_C(y, \tau_1) \). Eq. \((2.6)\) assumes a Ricker growth model [Ricker, 1954]. The parameter \( \eta \) is the baseline fertility rate, i.e. the average number of offspring produced by a single fish when the total fish population size is low, while \( \xi \) determines the strength of density dependence. For the aforementioned reasons, at the beginning of a season there are neither exposed nor infected fish [Eqs. \((2.4g)\) and \((2.4h)\)]. Table 2.2 summarizes all parameters of systems \((2.1)\) and \((2.4)\).

This local version of the PKD epidemiological model is used in the study presented in Chapter 3.

2.2 Spatially-explicit PKD model

To develop a spatially-explicit metacommunity model at the catchment scale, a river network is discretized into \( N_s \) stretches of suitable length (Fig. 2.2). Within each river stretch, the population and epidemiological dynamics are assumed to be well-mixed (i.e. no spatial description of the underlying variables and processes) and the environmental conditions (e.g. water temperature, velocity and depth) are approximated as uniform. Each river stretch constitutes a node of an oriented graph with edges following the flow direction. Mathematically, an oriented graph can be expressed by means of an adjacency matrix \( W \), whose entries \( w_{ij} \) are equal to unity if stretch \( i \) drains into stretch \( j \), and zero otherwise.

For each node of the network, a model akin to the one described in §2.1 is applied. In addition, a spatially-explicit modelling framework must account for the movement of fish along the river network and the hydrological transport of parasite spores. As disease transmission between the two hosts occurs only during the warm season, the spatial spreading of PKD is modelled only during such a season. Owing to their tiny size (about 20 \( \mu \)m in diameter [Canning et al., 1999, 2000]), parasite spores are assumed to be simply passively advected by the water flow, i.e. they travel downstream at the same velocity of water particles. Hence, the flux of spores, say released from fish, exiting from river stretch \( i \) can be computed as \( Q_i Z_{E,i}/V_i \), where \( Q_i \) is the discharge flowing out of the stretch and \( Z_{E,i}/V_i \) is the spore concentration, \( V_i \) being the water volume of the stretch.

2.2. Spatially-explicit PKD model

Fish movement is modelled as a diffusive process. This certainly represents a simplification of the complex set of behavioural, social and age-dependent determinants that control the mobility patterns of a fish population [Jonsson and Jonsson, 1993; Young, 1994; Forseth et al., 1999; Slavik et al., 2012; Frank et al., 2012]; however, it represents a useful working hypothesis to investigate large-scale processes induced by active movements of aquatic organisms in river networks (e.g. Fagan [2002]; Muneepeerakul et al. [2008]; Bertuzzo et al. [2008]; Mari et al. [2014b]). Moreover, diffusion represents the limiting behaviour of several more complex movement models, including correlated random walks and mixtures of correlated random walks [Patlak, 1953; Holmes, 1993; Bertuzzo et al., 2007]. The flux of fish exiting from stretch \(i\) is \(l_i F_i\), where \(F_i\) is the local fish abundance regardless of the epidemiological class and \(l_i [T^{-1}]\) is a mobility rate. The flux of fish from stretch \(i\) to stretch \(j\) reads \(d_{ij} l_i F_i\), where \(D\) is a diffusion matrix with entries \(d_{ij} > 0\) when either \(w_{ij}\) or \(w_{ji}\) are non-null, and \(d_{ij} = 0\) otherwise. The coefficients \(d_{ij}\) express the relative probability that a fish exiting from stretch \(i\) chooses stretch \(j\) as destination among all stretches connected to \(i\), either downstream or upstream. In addition, diagonal elements \(d_{ii}\) are positive if \(i\) represents either a headwater or the outlet, where, assuming that the river network is a closed system, fish that would exit the domain are assumed to remain in the same node \(i\). By continuity one has \(\sum_{j=1}^{N_i} d_{ij} = 1\). This formulation is rather general and requires additional assumptions (e.g. criteria for expressing fish preferences among river stretches) to fully determine the probability \(d_{ij}\). The diffusion matrix \(D\) henceforth used is derived using the following assumptions: fish have equal probability to move downstream or upstream; the probability to choose a given upstream stretch is proportional to its cross-sectional area; \(\sum_{j\neq i} d_{ij} = 0.5\) and \(d_{ii} = 0.5\) if stretch \(i\) is a headwater or the outlet. The assumption of equal probability for upstream/downstream movement seems fair when compared to field observation on brown trout seasonal movement patterns (e.g. James et al. [2007]).

According to the hypotheses introduced above, the set of ordinary differential equations describing the intra-annual disease dynamics in stretch \(i\) reads:

\[
\frac{dB_{S,i}}{dr} = r \left[ 1 - \frac{\rho_l}{V_i} (B_{S,i} + B_{C,i} + B_{O,i}) \right] B_{S,i} - \frac{Z_{F,i}^*}{V_i} B_{S,i}; \quad (2.7a)
\]

\[
\frac{dB_{C,i}}{dr} = r \left[ 1 - \frac{\rho_l}{V_i} (B_{S,i} + B_{C,i} + B_{O,i}) \right] B_{C,i} + \frac{Z_{F,i}^*}{V_i} B_{S,i} - d_{CO} B_{C,i} + d_{OC} B_{O,i}; \quad (2.7b)
\]
with the inter-annual set of difference equations (2.4), constitutes the spatially-explicit time-li

2.2.1 Fish mobility rates

The fish mobility rate \( l_i \) can be thought of as the inverse of the population-averaged residence time within stretch \( i \). In general, stretches may have different geometric and physical characteristics (e.g. length, water volume and depth, velocity, fish carrying capacity) and thus \( l_i \) is expected to change across the river network. This section illustrates how to compute \( l_i \) so that the stationary state of the underlying diffusion process is a specific spatial distribution of fish abundances \( F_i \). The underlying idea is that to apply the model, first a distribution of fish abundance at carrying capacity is assigned according to the characteristics of each stretch,
2.3. Spatially-explicit model with age-structure

then a set of mobility rates $l_i$ is derived so that fish movement leads, at steady state, to the target abundance distribution.

Once movements rules are determined (diffusion matrix $D$), the values of $l_i$ such that a distribution of fish abundances $F_i$ is an equilibrium state are obtained by solving the following linear system:

$$d_{ij}l_i F_i = d_{ji}l_j F_j \quad \forall i \leq N_s, j \leq N_s. \quad (2.8)$$

Note that only $N_s - 1$ of the above equations are not trivial identities: in fact, $d_{ij} \neq 0$ only if stretches $i$ and $j$ are directly connected, and every stretch has one downstream connection, with the exception of the outlet stretch. The system has $\infty^1$ solutions; indeed if a set of $l_i$ is a solution, also the same set multiplied by a scalar is a solution. It is thus possible to focus on a single solution by specifying the average mobility rate across the network:

$$l_{avg} = \frac{1}{N_s} \sum_{i=1}^{N_s} l_i.$$ 

An example of spatial distribution of population-average fish residence time is provided in Fig. 4.1d with regards to an idealized stream network built according to the optimal channel network principle (more details are provided in Chapter 4).

This spatial version of the PKD epidemiological model is used in the study presented in Chapter 4.

2.3 Spatially-explicit model with age-structure\(^3\)

To better match empirical features observed in the field, the previously introduced metacom- munity model is here expanded to account for the age-structure of the fish population. The time-hybrid discrete-continuous nature of the model directly enables splitting the total fish abundance into young-of-the-year (YOY) and adult (older than 1 year) individuals. Indeed, it is sufficient to consider that all newborn fish enter a YOY susceptible compartment ($Y_S$). During their first year of age, YOY individuals can then cycle across the exposed ($Y_E$), acutely infected ($Y_I$) and carrier ($Y_C$) compartments according to the dynamics already presented in §2.1. At the beginning of a new season, all survived susceptible YOY and adults enter the susceptible adult compartment $F_S$, while individuals belonging to all other classes that managed to survive through winter enter the adult carrier class $F_C$. The full epidemiological model is sketched in Fig. 2.3. A list of state variables is reported in Table 2.3.

Mobility rates and the diffusion matrix are as in §2.2 such that, given any spatial distribution

---

Symbol | Variable
--- | ---
$B_S$ | Biomass of susceptible bryozoans
$B_C$ | Biomass of covertly infected bryozoans
$B_O$ | Biomass of overtly infected bryozoans
$\delta_S$ | Non-infected statoblast abundance
$\delta_I$ | Infected statoblast abundance
$Y_S$ | Abundance of susceptible YOY
$Y_E$ | Abundance of exposed YOY
$Y_I$ | Abundance of acutely infected YOY
$Y_C$ | Abundance of carrier YOY
$F_S$ | Abundance of susceptible adult fish
$F_E$ | Abundance of exposed adult fish
$F_I$ | Abundance of acutely infected adult fish
$F_C$ | Abundance of carrier adult fish
$Z_B$ | Abundance of spores released by bryozoans
$Z_F$ | Abundance of spores released by fish

Table 2.3 – List of variables for the epidemiological model with age-structure. All state variables are dimensionless.

of fish at the beginning of the season, the system tends to reach the target equilibrium distribution towards the end of the season, provided that the average fish mobility rate $l_{avg}$ is large enough. During winter, the number of newborn fish generated by the female adults living in each stretch is estimated according to a Ricker model [Ricker, 1954]. As brown trout are subject to spawning migration to seek for suitable habitats and subsequent natal homing [Frank et al., 2012], newborn individuals generated by adults living in a specific stretch are assumed to hatch in suitable upstream stretches according to a gravity model [Erlander and Stewart, 1990]. For the sake of simplicity, the same set of epidemiological and mobility parameters is hereafter used for both YOY and adult fish, but such an hypothesis can be easily relaxed.

The set of ordinary differential equations describing the intra-season disease dynamics in stretch $i$ reads:

\[
\frac{d}{dt}B_{S,i} = r_i \left[ 1 - \frac{P_i}{V_i} (B_{S,i} + B_{C,i} + B_{O,i}) \right] B_{S,i} - \frac{Z_{F,i}^*}{V_i} B_{S,i}; \quad (2.9a)
\]

\[
\frac{d}{dt}B_{C,i} = r_i \left[ 1 - \frac{P_i}{V_i} (B_{S,i} + B_{C,i} + B_{O,i}) \right] B_{C,i} + \frac{Z_{F,i}^*}{V_i} B_{S,i} - d_{C,i} B_{C,i} + d_{O,i} B_{O,i}; \quad (2.9b)
\]

\[
\frac{d}{dt}B_{O,i} = r_{O,i} \left[ 1 - \frac{P_i}{V_i} (B_{S,i} + B_{C,i} + B_{O,i}) \right] B_{O,i} + d_{C,i} B_{C,i} - d_{O,i} B_{O,i}; \quad (2.9c)
\]

\[
\frac{d}{dt}S_{S,i} = f_B B_{S,i} + \phi f_B B_{C,i}; \quad (2.9d)
\]

\[
\frac{d}{dt}S_{I,i} = (1 - \phi) f_B B_{C,i}; \quad (2.9e)
\]

\[
\frac{d}{dt}Y_{S,i} = -\mu_F Y_{S,i} + \xi Y_{C,i} - \frac{Z_{B,i}^*}{V_i} Y_{S,i} + \sum_{k=1}^{N} d_k l_k Y_{S,k} - \sum_{k=1}^{N} d_k l_k Y_{S,i}; \quad (2.9f)
\]
2.3. Spatially-explicit model with age-structure

Figure 2.3 – Schematic representation of the full epidemiological model. a) Intra-seasonal local model. b) Inter-seasonal local model. Parameters are indicated in gray. c) Spatially-explicit model. B: bryozoan sub-model. Y,F: fish sub-model.

\[
\begin{align*}
\frac{dY_{E,i}}{dt} &= \frac{Z_{B,i}}{V_I} Y_{S,i} - (\mu_F + h_i) Y_{E,i} + \sum_{k=1}^{N_i} d_{kl} l_k Y_{E,k} - \sum_{k=1}^{N_i} d_{ik} l_i Y_{E,i}; \\
\frac{dY_{I,i}}{dt} &= (1-\epsilon) h_i Y_{E,i} - (\mu_F + a_i + \gamma) Y_{I,i} + \sum_{k=1}^{N_i} d_{kl} l_k Y_{I,k} - \sum_{k=1}^{N_i} d_{ik} l_i Y_{I,i}; \\
\frac{dY_{C,i}}{dt} &= \epsilon h_i Y_{E,i} + \gamma Y_{I,i} - (\mu_F + \zeta) Y_{C,i} + \sum_{k=1}^{N_i} d_{kl} l_k Y_{C,k} - \sum_{k=1}^{N_i} d_{ik} l_i Y_{C,i}; \\
\frac{dF_{S,i}}{dt} &= -\mu_F F_{S,i} + \zeta F_{C,i} - \frac{Z_{B,i}^*}{V_I} F_{S,i} + \sum_{k=1}^{N_i} d_{kl} l_k F_{S,k} - \sum_{k=1}^{N_i} d_{ik} l_i F_{S,i}; \\
\frac{dF_{E,i}}{dt} &= \frac{Z_{B,i}^*}{V_I} F_{S,i} - (\mu_F + h_i) F_{E,i} + \sum_{k=1}^{N_i} d_{kl} l_k F_{E,k} - \sum_{k=1}^{N_i} d_{ik} l_i F_{E,i}; \\
\frac{dF_{I,i}}{dt} &= (1-\epsilon) h_i F_{E,i} - (\mu_F + a_i + \gamma) F_{I,i} + \sum_{k=1}^{N_i} d_{kl} l_k F_{I,k} - \sum_{k=1}^{N_i} d_{ik} l_i F_{I,i}; \\
\frac{dF_{C,i}}{dt} &= \epsilon h_i F_{E,i} + \gamma F_{I,i} - (\mu_F + \zeta) F_{C,i} + \sum_{k=1}^{N_i} d_{kl} l_k F_{C,k} - \sum_{k=1}^{N_i} d_{ik} l_i F_{C,i};
\end{align*}
\]
Chapter 2. A metacommunity epidemiological model for PKD

\[ \frac{dZ_{B,i}^*}{d\tau} = \pi_{B} B_{O,i} - \mu_Z Z_{B,i}^* + \sum_{k=1}^{N_i} w_{k}\frac{Q_k}{V_{j}} Z_{B,k}^* - \sum_{k=1}^{N_i} w_{k}\frac{Q_l}{V_{l}} Z_{B,l}^*; \]  
(2.9n)

\[ \frac{dZ_{E,i}^*}{d\tau} = \pi_{E} (Y_{l,i} + F_{l,i}) + \kappa \pi_{E} (Y_{C,i} + F_{C,i}) - \mu_Z Z_{E,i}^* + \sum_{k=1}^{N_i} w_{k}\frac{Q_k}{V_{j}} Z_{E,j}^* - \sum_{k=1}^{N_i} w_{k}\frac{Q_l}{V_{l}} Z_{E,l}^*; \]  
(2.9o)

On the other hand, the following difference equation system relates the state of the model variables at the end of a season \((y)\) with that at the beginning of the following season \((y + 1)\).

\[
B_{S,i}(y + 1) = \sigma_B B_S(y) + \nu S_S(y); \quad (2.10a)
\]

\[
B_{C,i}(y + 1) = \sigma_B B_C(y) + \sigma_O B_O(y) + \nu S_I(y); \quad (2.10b)
\]

\[
B_{O,i}(y + 1) = 0; \quad (2.10c)
\]

\[
S_{S,i}(y + 1) = 0; \quad (2.10d)
\]

\[
S_{I,i}(y + 1) = 0; \quad (2.10e)
\]

\[
Y_{S,i}(y + 1) = \sum_{j=1}^{N_i} W_{j} \eta \bar{f}_j(y) \exp \left( -\frac{\xi_j}{\bar{V}_j} \bar{f}_j(y) \right); \quad (2.10f)
\]

\[
Y_{E,i}(y + 1) = 0; \quad (2.10g)
\]

\[
Y_{I,i}(y + 1) = 0; \quad (2.10h)
\]

\[
Y_{C,i}(y + 1) = 0; \quad (2.10i)
\]

\[
F_{S,i}(y + 1) = \sigma_F (Y_{S,i}(y) + F_{S,i}(y)); \quad (2.10j)
\]

\[
F_{E,i}(y + 1) = 0; \quad (2.10k)
\]

\[
F_{I,i}(y + 1) = 0; \quad (2.10l)
\]

\[
F_{C,i}(y + 1) = \sigma_F \{ p_E [Y_{E,i}(y) + F_{E,i}(y)] + p_I [Y_{I,i}(y) + F_{I,i}(y)] + Y_{C,i}(y) + F_{C,i}(y) \}; \quad (2.10m)
\]

\[
Z_{B,i}^*(y + 1) = 0; \quad (2.10n)
\]

\[
Z_{E,i}^*(y + 1) = 0. \quad (2.10o)
\]

where \(\bar{V}_j\) is the mean water volume in stretch \(j\), \(\bar{f}_j(y) = Y_{S,j}(y) + F_{S,j}(y) + p_E [Y_{E,j}(y) + F_{E,j}(y)] + p_I [Y_{I,j}(y) + F_{I,j}(y)] + Y_{C,j}(y) + F_{C,j}(y)\), while \(W_{ij}\) is the fraction of newborns generated by adults living in \(j\) that hatch in \(i\) \((i \in U_j, \text{ where } U_j \text{ contains all stretches upstream of } j \text{ and } j \text{ itself})\) and is calculated via a gravity model [Erlander and Stewart, 1990]:

\[
W_{ij} = \frac{W_i^A e^{-L_{ij}/\lambda_F}}{\sum_{i \in U_j} W_i^A e^{-L_{ij}/\lambda_F}} \quad (2.11)
\]

when \(i \in U_j\), and \(W_{ij} = 0\) otherwise. In Eq. (2.11), \(W_i^A\) is a dimensionless score for spawning suitability, \(L_{ij}\) the total length of the path connecting \(i\) to \(j\) (including their lengths), \(\lambda_F\) the shape factor of the exponential kernel representing the deterrence factor. An exponential function can also be used to express spawning suitability: \(W_i^A = e^{-A_i/\lambda_F}\), implying that eggs
2.4. List of symbols

The following Table lists all symbols introduced in this Chapter that were not included in Tables 2.1, 2.2 and 2.3.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>Upstream drainage area (of stretch $i$)</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$A_F$</td>
<td>Shape factor for spawning suitability $W^A$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$\hat{a}$</td>
<td>Characteristic value of the PKD-related mortality rate $a$ during early winter</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$d_{ij}$</td>
<td>Entry of the diffusion matrix $D$</td>
<td>-</td>
</tr>
<tr>
<td>$f_F$</td>
<td>Newborn fish</td>
<td>$-3$</td>
</tr>
<tr>
<td>$\mathcal{F}_i$</td>
<td>Fish population at the end of the year in stretch $i$ that will spawn</td>
<td>$L^{-3} \quad (\S 2.1)$</td>
</tr>
<tr>
<td>$g_X$</td>
<td>Growth rate of bryozoans belonging to class $X$</td>
<td>$L^{-3} \quad (\S 2.3)$</td>
</tr>
<tr>
<td>$\hat{h}$</td>
<td>Characteristic value of the disease development rate $h$ during early winter</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$l_{avg}$</td>
<td>Average fish mobility rate</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$l_i$</td>
<td>Fish mobility rate at stretch $i$</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$L_{ij}$</td>
<td>Downstream length from stretch $i$ to $j$</td>
<td>$L$</td>
</tr>
<tr>
<td>$N_s$</td>
<td>Total number of river stretches</td>
<td>-</td>
</tr>
<tr>
<td>$p_{E}$</td>
<td>Fraction of exposed fish at the end of the season that will enter the carrier class during winter</td>
<td>-</td>
</tr>
<tr>
<td>$p_I$</td>
<td>Fraction of acutely infected fish at the end of the season that will enter the carrier class during winter</td>
<td>-</td>
</tr>
<tr>
<td>$Q_i(\tau)$</td>
<td>Discharge at time $\tau$ in stretch $i$</td>
<td>$L^3T^{-1}$</td>
</tr>
<tr>
<td>$U_i$</td>
<td>Subset of stretches upstream of stretch $i$ (including $i$)</td>
<td>-</td>
</tr>
<tr>
<td>$V_i(\tau)$</td>
<td>Water volume at time $\tau$ in stretch $i$</td>
<td>$L^3$</td>
</tr>
<tr>
<td>$\overline{V}_i$</td>
<td>Mean water volume in stretch $i$</td>
<td>$L^3$</td>
</tr>
<tr>
<td>$w_{ij}$</td>
<td>Entry of the adjacency matrix $W$</td>
<td>-</td>
</tr>
<tr>
<td>$W_i^A$</td>
<td>Spawning suitability score</td>
<td>-</td>
</tr>
<tr>
<td>$W_{ij}$</td>
<td>Fraction of newborn fish hatched in stretch $i$ generated by fish dwelling in stretch $j$ at the end of the previous season</td>
<td>-</td>
</tr>
<tr>
<td>$y$</td>
<td>Counter of years (i.e. warm seasons)</td>
<td>-</td>
</tr>
</tbody>
</table>

This full version of the epidemiological model is used in the case study applications of Chapter 5. Further details on parameter specification are reported in §5.1.5.
## Chapter 2. A metacommunity epidemiological model for PKD

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_F$</td>
<td>Characteristic distance covered by <em>S. trutta</em> for spawning</td>
<td>L</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>Beginning of warm season</td>
<td>T</td>
</tr>
<tr>
<td>$\tau_1$</td>
<td>End of warm season</td>
<td>T</td>
</tr>
</tbody>
</table>
A theoretical study of the local epidemiological model is performed in this Chapter in order to investigate the key processes controlling parasite invasion and persistence. Such goal is achieved by means of model simulations, stability and sensitivity analyses. A simulation of the model where parameters are set to their reference values (obtained by literature review) produces seasonal patterns of PKD spread in both hosts that are in accordance with field observations. Two different procedures are employed to conduct stability analyses: the first is a simplified, linearized approach which neglects population dynamics and winter transitions; the second is based on a Poincaré map coupling inter-season updates with intra-season dynamics, the latter expressed via an integral operator derived from Floquet theory. These analyses reveal that, for realistic parameter ranges, a disease-free system is highly invasible, which implies that the introduction of the parasite in a susceptible community is very likely to trigger a disease outbreak. Sensitivity analysis shows that, when the disease is endemic, the impact of PKD outbreaks is mostly controlled by the rates of disease development in the fish population. The developed mathematical model, together with the insights gained through stability and sensitivity analyses, unravels the basic mechanisms involved in disease spread and its long-term persistence; and sets the basis for the forthcoming studies.¹

3.1 Methods

3.1.1 Model simulation

In order to understand possible patterns of disease dynamics, simulations of the model presented in §2.1 are run. The possible range of several model parameters was estimated based on literature values and experts’ knowledge. Reasonable values were assumed for the remaining ones. Reference parameter values and feasible ranges are reported in Table 3.1. In the absence of detailed information about recovery mechanisms for fish and bryozoans, slow recovery rates are assumed as reference values (average recovery time equal to half of the lifetime for fish and half of the yearly proliferation period for bryozoans) and a wide range of possible values is explored, including the absence of infection-clearing mechanisms ($\psi = \zeta = 0$). As for temperature-dependent parameters, linear ($r, r_O$) or parabolic ($dCO$, $dOC$).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
<th>Reference value</th>
<th>Unit</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>[0.01, 0.1]</td>
<td>0.064</td>
<td>d$^{-1}$</td>
<td>Tops et al. [2009]</td>
</tr>
<tr>
<td>$r_O$</td>
<td>[0.005, 0.05]</td>
<td>0.032</td>
<td>d$^{-1}$</td>
<td>Tops et al. [2009]</td>
</tr>
<tr>
<td>$\psi$</td>
<td></td>
<td>0.01</td>
<td>d$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td>200</td>
<td>gm$^{-3}$</td>
<td></td>
</tr>
<tr>
<td>$d_{OC}$</td>
<td>[0.02, 0.2]</td>
<td>0.0488</td>
<td>d$^{-1}$</td>
<td>Tops et al. [2007]</td>
</tr>
<tr>
<td>$d_{CO}$</td>
<td>[0.02, 0.2]</td>
<td>0.0848</td>
<td>d$^{-1}$</td>
<td>Tops et al. [2007]</td>
</tr>
<tr>
<td>$a$</td>
<td>[0.015, 0.1]</td>
<td>0.0456</td>
<td>d$^{-1}$</td>
<td>Bettge et al. [2009a,b]</td>
</tr>
<tr>
<td>$h$</td>
<td>[0.01, 0.075]</td>
<td>0.0334</td>
<td>d$^{-1}$</td>
<td>El-Matbouli and Hoffman [1994]</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>[0.01, 0.05]</td>
<td>0.02</td>
<td>d$^{-1}$</td>
<td>Feist et al. [2001]; Okamura et al. [2011]</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>[0.01, 0.05]</td>
<td>0.001</td>
<td>d$^{-1}$</td>
<td>Abd-Elfattah et al. [2014b]</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td></td>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{\pi_B}{\mu_B}$</td>
<td>$5 \cdot 10^{-3}$</td>
<td>m$^3$g$^{-1}$d$^{-2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\frac{\pi_F}{\mu_F}$</td>
<td>$5 \cdot 10^{-3}$</td>
<td>m$^3$d$^{-2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\kappa$</td>
<td></td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_F^{-1}$</td>
<td>~ 5 years</td>
<td>2000</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>$\mu_Z^{-1}$</td>
<td>≤ 24 h</td>
<td>0.75</td>
<td>d</td>
<td>De Kinkelin et al. [2002]</td>
</tr>
<tr>
<td>$\sigma_B$</td>
<td>[0, 0.3]</td>
<td>0.1</td>
<td>-</td>
<td>Okamura et al. [2011]</td>
</tr>
<tr>
<td>$\sigma_O$</td>
<td>[0, 0.1]</td>
<td>0.05</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\sigma_F$</td>
<td></td>
<td>0.9</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\eta$</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi$</td>
<td></td>
<td>0.01</td>
<td>m$^3$</td>
<td></td>
</tr>
<tr>
<td>$\nu$</td>
<td></td>
<td>0.04</td>
<td>g</td>
<td></td>
</tr>
<tr>
<td>$f_B$</td>
<td></td>
<td>0.1</td>
<td>g$^{-1}$d$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>[0.55, 0.8]</td>
<td>0.7</td>
<td>-</td>
<td>Abd-Elfattah et al. [2014a]</td>
</tr>
<tr>
<td>$\hat{T}$</td>
<td></td>
<td>15</td>
<td>oC</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1 – List of parameters involved in the local PKD model. Feasible ranges for some of the parameters are presented and the respective literature sources are provided. A reference set of parameters for sensitivity and stability analyses is defined. As regards temperature-dependent parameters, the reference values are obtained from the relationships of Fig. 3.1 by assuming $T = 15$ °C.
3.1. Methods

Figure 3.1 – Functional forms for temperature-dependent parameters. The maximum values of these parameters are set to match the upper limits proposed in Table 3.1. Furthermore, \( f_B = 0.005 T \text{[°C]} + 0.00025 \tau \text{[d]} \) is assumed, where \( \tau \) is the time elapsed since January 1.

\( d_{OC}, h, a \) relationships have been assumed; a superlinear dependence on \( T \) for the latter parameters was deduced from literature review (see references in Table 3.1). The functional forms used are illustrated in Fig. 3.1. Note that, for these parameters, the corresponding reference values are obtained from the relationships of Fig. 3.1 by assuming \( T = 15 \text{ °C} \). For the numerical simulations, a time series of stream water temperature measured in River Langete, Switzerland, in the period 2002-2013 is used (data provided by the Swiss Federal Office for the Environment - FOEN).

3.1.2 Analysis of the within-season model

Parasite invasion in a previously PKD-free system most likely occurs during the warm season. For this reason, the analysis is firstly focused on the continuous, within-season model and on the conditions under which the introduction of the parasite in a fully susceptible population of fish and bryozoans (say through the introduction of spores, infected fish or infected bryozoans) leads to an outbreak of PKD. To this end, the linear stability of the disease-free equilibrium (DFE) is analyzed in a simplified model that disregards population dynamics of both fish and bryozoans (i.e. both population sizes are treated as model parameters).

Consider a within-season system derived from Eqs. (2.1) and (2.4). In such a system, the equations expressing the time evolution of statoblasts are neglected; indeed, statoblasts only affect other state variables in the transition between two seasons. Moreover, in order to define a disease-free equilibrium, the equations related to the time evolution of susceptible bryozoan biomass and susceptible fish abundance are neglected, and the terms \( B_S \) and \( F_S \) appearing in other equations are treated as parameters. All parameters are assumed as constant in time.
Chapter 3. Analysis of the local PKD model

Hence, this system reads:

\[
\begin{align*}
\frac{dBC}{dr} &= r \left[ 1 - \rho (BS + BC + BO) \right] B_C - (d_{CO} + \psi) B_C + d_{OC} B_O + BS Z_F^*; \\
\frac{dBO}{dr} &= r_O \left[ 1 - \rho (BS + BC + BO) \right] B_O + d_{CO} B_C - (d_{OC} + \psi) B_O; \\
\frac{dFE}{dr} &= - (\mu_F + h) F_E + F_S Z_B^*; \\
\frac{dFI}{dr} &= (1 - \epsilon) h F_E - (\mu_F + a + \gamma) F_I; \\
\frac{dFC}{dr} &= \epsilon h F_E + \gamma F_I - (1 - \epsilon) \zeta F_C; \\
\frac{dZB}{dr} &= \pi^* B_O - \mu_Z Z_B^*; \\
\frac{dZF}{dr} &= \pi^* (F_I + \kappa F_C) - \mu_Z Z_F^*.
\end{align*}
\] (3.1a)

The disease-free equilibrium (DFE) of the system (3.1) is \( x_{DF} = \{ B_C; B_O; F_E; F_I; F_C; Z_B^*; Z_F^* \} = 0 \).

Upon linearization of (3.1) around \( x_{DF} \), its Jacobian can be calculated as:

\[
J_{DF} = \begin{bmatrix}
J_{DF,11} & d_{OC} & 0 & 0 & 0 & 0 & B_S \\
d_{CO} & J_{DF,22} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -h - \mu_F & 0 & 0 & F_S & 0 \\
0 & 0 & (1 - \epsilon) h & -a - \gamma - \mu_F & 0 & 0 & 0 \\
0 & 0 & \epsilon h & \gamma & -\zeta - \mu_F & 0 & 0 \\
0 & \pi_B^* & 0 & 0 & 0 & -\mu_Z & 0 \\
0 & 0 & \pi_F^* & \kappa \pi_F^* & 0 & 0 & -\mu_Z
\end{bmatrix}
\]

where \( J_{DF,11} = r (1 - \rho BS) - \psi - d_{CO} \) and \( J_{DF,22} = r_O (1 - \rho BS) - \psi - d_{OC} \). According to a classical result of dynamical system theory [Luenberger, 1979], the spectral abscissa \( \alpha_s (J_{DF}) \) of the Jacobian of (3.1) expresses the speed at which a perturbation of the equilibrium state will propagate; if \( \alpha_s (J_{DF}) < 0 \), the DFE is stable. The determinant of \( J_{DF} \) reads:

\[
\det(J_{DF}) = J_1 - J_2 T,
\]

where

\[
J_1 = B_S F_S d_{CO} h \pi_B^* \pi_F^* [\epsilon k a + \kappa \gamma + (1 - \epsilon) \zeta + (1 + \epsilon k - \epsilon) \mu_F];
\]

\[
J_2 = (h + \mu_F) (a + \gamma + \mu_F) (\zeta + \mu_F) \mu_Z^2;
\]

\[
T = \psi^2 + [d_{CO} + d_{OC} - (r + r_O) (1 - \rho BS)] \psi - (1 - \rho BS) \left[ r_O d_{CO} + r d_{OC} + r r_O (1 - \rho BS) \right].
\]

Note that \( J_1 \) and \( J_2 \) are non-negative by construction. Given that the order of \( J_{DF} \) is odd, a switch in the sign of \( \alpha_s (J_{DF}) \) from negative to positive corresponds to a switch in the sign of \( \det(J_{DF}) \) from negative to positive. Also, \( \det(J_{DF}) > 0 \) is a sufficient condition for the instability of the DFE, i.e. \( x_{DF} \). This condition is always true when \( T < 0 \); in this case, the parasite
can spread within the bryozoan population even in the absence of the fish population (i.e. when $F_S = 0$ and thus $J_1 = 0$). Indeed, it is straightforward to show that $\mathcal{T}$ is the determinant of the Jacobian of a system that considers only bryozoans [i.e. Eqs. (3.1a) with $Z_F^* = 0$ and (3.1b)], with $\mathcal{T} < 0$ representing the instability criterion for the DFE of such a system. If instead $\det(J_{DF}) > 0$ with $\mathcal{T} > 0$, $x_{DF}$ is still unstable, but in this case the parasite needs to cycle between the two hosts to invade the system.

A condition of incipient instability can be derived when $\det(J_{DF}) = 0$, which leads to an expression for the reproduction number $R$:

$$R = \frac{J_1}{J_2^*} = \frac{B_S F_S d_{CO} h \pi^*_F \pi^*_F [\epsilon \kappa a + \kappa \gamma + (1 - \epsilon) \zeta + (1 + \epsilon \kappa - \epsilon) \mu_F]}{\mathcal{T} (h + \mu_F)(a + \gamma + \mu_F)(\zeta + \mu_F) \mu^2_z}. \quad (3.2)$$

The DFE is unstable when $R$ is larger than unity. Note that the expression (3.2) is only valid if $\mathcal{T} > 0$; failing that, the DFE is always unstable and $R$ can not be defined.

The instability of the within-season DFE would indicate long-term persistence of the parasite only if model parameters were kept constant at the value used for the computation of the stability criterion. However, in system (2.1) population dynamics and water temperature variations constantly affect parameters values. Also, the alternation of warm and cold seasons, each endowed with different eco-epidemiological processes, *de facto* prevents a stable, nontrivial steady-state from being established in the system. Therefore, the instability of the within-season DFE can be seen as a useful indicator only of the short-term invasibility of the system. To investigate long-term parasite persistence and the establishment of endemic transmission conditions, a more complex analysis that accounts for both seasonal forcing and intra- and inter-seasonal population and epidemiological dynamics is required, as illustrated in the following section.

### 3.1.3 Stability analysis of the full model

As the seasonal cycle of water temperature critically controls PKD transmission and bryozoan population dynamics, it is crucial to study the stability of the disease-free trajectory (DFT), namely a succession of system states in which the disease is not present and endemic transmission is not possible. This aim is achieved by means of a Poincaré map coupling between-season update with the within-season dynamics, with the latter being described through an integral operator derived from Floquet theory [Bacaër and Ouifki, 2007; Wang and Zhao, 2008; Mari et al., 2014a], a mathematical framework that allows the analysis of systems of periodically forced differential equations. As this theory assumes periodic forcing, the time evolution of water temperature is approximated via a sinusoidal signal with yearly period.
Chapter 3. Analysis of the local PKD model

Disease-free trajectory

The DFT is defined as a discrete-continuous trajectory describing the yearly periodic evolution of biomass or abundances for the uninfected classes \(B_S, S_S, F_S\) in the absence of PKD. Therefore, it is convenient to first focus on a reduced system where only susceptible compartments are taken into account. Let \(x_S(y;\tau) = \{B_S(y;\tau); S_S(y;\tau); F_S(y;\tau)\}\) be the state of this system at season \(y\) and time \(\tau\). The time-evolution of this system can be represented as follows:

\[
x_S(y;\tau_0) \xrightarrow{v(\cdot)} x_S(y;\tau_1) \xrightarrow{w(\cdot)} x_S(y+1;\tau_0) \xrightarrow{v(\cdot)} x_S(y+1;\tau_1)
\]

where \(v(\cdot)\) is an operator describing the continuous-time dynamics of the system from \(\tau_0\) to \(\tau_1\), while \(w(\cdot)\) is an operator describing the discrete-time update of the system between the end of season \(y\) and the beginning of season \(y+1\). Therefore, one obtains:

\[
x_S(y+1;\tau_0) = w(v(x_S(y;\tau_0))) = (w \circ v)(x_S(y;\tau_0)),
\]

where \(w \circ v\) defines a so-called Poincaré map [Guckenheimer and Holmes, 1983]. A fixed point of the Poincaré map is defined as an equilibrium \(\bar{x}_S(y;\tau)\) such that \(\bar{x}_S(y;\tau_0) = \bar{x}_S(y+1;\tau_0) = (w \circ v)(\bar{x}_S(y;\tau_0))\).

The disease-free trajectory \(\bar{x}_S(y;\tau)\) for systems (2.1) and (2.4) is defined under the hypothesis that water temperature can be approximated by a sinusoidal function of time, with period equal to one year. We therefore obtain

\[
\bar{x}_S(y;\tau) = \left\{ \bar{B}_S(\tau), \int_{\tau_0}^{\tau} f_B(t') \bar{B}_S(t') dt', F_{S,0} \exp[-\mu_F(\tau - \tau_0)] \right\}, \tag{3.3}
\]

where

\[
\bar{B}_S(\tau) = \frac{B_{S,0}}{\rho B_{S,0} + (1 - \rho B_{S,0}) \exp[-R(\tau)]}; \quad R(\tau) = \int_{\tau_0}^{\tau} r dt
\]

and \(B_{S,0}\) \((F_{S,0})\) is the susceptible bryozoan (fish) population at \(\tau_0\) along the DFT. Note in fact that \(\bar{x}_S(y;\tau_0) = \{B_{S,0}, 0, F_{S,0}\}\). \(B_{S,0}\) is the solution of the following transcendental equation:

\[
1 - \frac{\sigma_B}{\rho B_{S,0} + (1 - \rho B_{S,0}) \exp[-R(\tau_1)]} - \int_{\tau_0}^{\tau_1} \rho f_B(1 - \rho B_{S,0}) \exp[-R(\tau)] \, dr = 0,
\]

The value of \(B_{S,0}\) can be obtained numerically for a given set of parameters. Instead, an analytical expression can be found for \(F_{S,0}\):

\[
F_{S,0} = \xi^{-1} \ln \left\{ \frac{\eta}{\exp[\mu_F(\tau_1 - \tau_0)] - \sigma_F} \right\} \exp[\mu_F(\tau_1 - \tau_0)]. \tag{3.4}
\]

Therefore, \(\bar{x}_S(y;\tau)\) is the trajectory such that its susceptible components are equal to (3.3), while its infected components \(x_I = \{B_C; B_O; S_I; F_E; F_C; Z_{O}^*; Z_{C}^*\}\) are null. Note that Eq. (3.4) requires \(\eta > \eta_{\text{min}} = \exp[\mu_F(\tau_1 - \tau_0)] - \sigma_F\) in order to avoid extinction of fish. For the reference
3.1. Methods

parameter set of Table 3.1, one gathers \( \eta_{\min} \approx 0.205 \).

**Stability of the disease-free trajectory**

In order to study the stability of the DFT \( \mathbf{x}_{\text{df}}(y;\tau) \) of model (2.1) and (2.4) with respect to small perturbations, one can consider a reduced system, where only infected compartments are considered. Hence, \( \mathbf{x}_I(y;\tau) \) denotes the state of such a system, whose time evolution is represented as follows:

\[
\begin{align*}
\mathbf{x}_I(y;\tau_0) & \overset{V}{\longrightarrow} \mathbf{x}_I(y;\tau_1) \overset{W}{\longrightarrow} \mathbf{x}_I(y+1;\tau_0) \overset{V}{\longrightarrow} \mathbf{x}_I(y+1;\tau_1).
\end{align*}
\]  

(3.5)

\( V \) and \( W \) are operators describing the evolution of the linearized system within and between seasons, respectively. Note that the fixed point for this system is the null vector.

The propagation matrix \( V \) is given by the solution of the following matrix ODE system:

\[
\frac{d\mathbf{Z}}{d\tau} = A(\tau)\mathbf{Z},
\]  

(3.6)

where \( A(\tau) \) is a matrix obtained from the Jacobian of the within-season system (2.1), evaluated along the trajectory \( \mathbf{x}_{\text{df}}(y;\tau) \), where all rows and columns related to the susceptible compartments have been removed:

\[
A(\tau) = \begin{bmatrix}
A_{11} & d_{OC} & 0 & 0 & 0 & 0 & B_S \\
\varphi f & A_{22} & 0 & 0 & 0 & 0 & 0 \\
(1-\varphi) f & 0 & 0 & 0 & 0 & 0 & 0 \\
-\mu F - h & 0 & 0 & 0 & 0 & 0 & F_S \\
0 & 0 & (1-\epsilon) h & A_{55} & 0 & 0 & 0 \\
0 & 0 & \epsilon h & \gamma & -\mu F - \zeta & 0 & 0 \\
0 & \pi_F^* & 0 & 0 & 0 & -\mu Z & 0 \\
0 & 0 & 0 & 0 & \pi_F^* & \kappa \pi_F^* & -\mu Z
\end{bmatrix},
\]

where \( A_{11} = r(1-\rho B) - (d_{CO} + \psi) \), \( A_{22} = r_0 (1 - \rho B) - d_{OC} \), \( A_{55} = -\mu F - \alpha - \gamma \). Eq. (3.6) must be integrated from \( \tau_0 \) to \( \tau_1 \), with initial condition \( \mathbf{Z}(\tau_0) = \mathbf{I} \) (identity matrix). One thus has \( V = \mathbf{Z}(\tau_1) \).

Matrix \( W \) can be obtained from the Jacobian matrix of the between-season system (2.4), by disregarding all rows and columns related to susceptible compartments. \( W \) is a sparse matrix of order 8 with the following non-null entries: \( W_{11} = \sigma_B \), \( W_{12} = \sigma_O \), \( W_{13} = \nu \), \( W_{64} = p_F \sigma_F \), \( W_{65} = p_I \sigma_F \), \( W_{66} = \sigma_F \).

The stability of the disease-free trajectory of systems (2.1) and (2.4) corresponds to the stability of the fixed point \( \mathbf{x}_I = 0 \) of the Poincaré map defined in (3.5), i.e. \( \mathbf{x}_I(y+1;\tau_1) = VW\mathbf{x}_I(y;\tau_1) \). Therefore, the necessary and sufficient condition for the exponential instability of the DFT
Chapter 3. Analysis of the local PKD model

reads

\[ \lambda_{max} = \max |\lambda(VW)| > 1, \quad (3.7) \]

where \( \lambda \) are the eigenvalues of matrix \( VW \). When condition (3.7) is met, the parasite can invade the system. Therefore, the analysis can be focused on \( \lambda_{max} \), i.e. the maximum modulus of the eigenvalues of \( VW \). For a given set of parameters, the corresponding value of \( \lambda_{max} \) can be computed numerically.

3.1.4 Sensitivity analyses

To understand how model parameters affect parasite invasibility and long-term PKD impact, sensitivity analyses are performed. Specifically, the goal is to investigate the effect of parameters on the value of \( \lambda_{max} \) and on the PKD-induced fish loss, here estimated as the percentage of fish at the end of the season with respect to the fish population size if the disease were absent. Computations are run by varying two focus parameters at a time while keeping the others at their reference value, as specified in Table 3.1. Temperature-dependent parameters are expressed via the functional forms of Fig. 3.1. The effect of these parameters is explored by varying their value at 15 °C, while keeping their minimum value at 0 °C (at 25 °C for \( d_{OC} \) constant. For parabolic functional forms, the null derivative at their minimum is also kept constant. The sinusoidal signal of water temperature is derived from the time series shown in the top panel of Fig. 3.2. Seasons are assumed to start on April 1 and last for 200 days. The effect of water temperature is also investigated by varying both the yearly mean value and the relative mid-amplitude of the sinusoidal signal, while using the relationships of Fig. 3.1.

3.2 Results

3.2.1 Model simulation

Fig. 3.2 shows an example of model simulation where it is assumed that at the beginning of the first season 1% of the bryozoans are covertly infected and 1% of the fish belong to the carrier class. The remaining fractions of the host populations are susceptible. This model setting mimics the invasion of the parasite in a fully susceptible system. The total bryozoan population tends to reach the carrying capacity \( \rho^{-1} \) towards the end of the season. Overt stages of infection are absent in early spring but peaks are observed in summer. The irregular shape of the curves for \( B_C \) and \( B_O \) mirrors the variability of water temperature. Regarding fish, peaks of infection are observed during the summer. After a few years, prevalence reaches about 80% at the end of the summer, where few susceptible fish are present and most of the survived individuals belong to the carrier class. This is in agreement with the fact that, although young-of-the-year fish surviving the infection are not likely to develop clinical PKD in the following year, the parasite can remain viable in the host for several seasons after initial exposure [Ferguson and Ball, 1979; Morris et al., 2000; Abd-Elfattah et al., 2014b]. In this
3.2. Results

model simulation, some 28% of the fish population which is alive at $r_0$ dies during the season because of PKD. About ten years after the initial invasion, the disease becomes endemic and the seasonal state variables’ trajectories remain virtually unchanged. Overall, the model reproduces patterns of disease spread in bryozoan and fish populations which are in good qualitative agreement with evidence from the literature and field observations (see Okamura et al. [2011]).
Chapter 3. Analysis of the local PKD model

3.2.2 Analysis of the within-season model

Fig. 3.3 identifies regions of stability for the disease-free equilibrium in the parameter space. Starting from the reference parameter set, \( R \) increases as parameters \( F_S, d_{CO}, h, \pi_B^*, \pi_F^* \) increase, and decreases as parameters \( \rho B_S, d_{OC}, a, \psi \) increase. Fig. 3.4 illustrates how \( R \) evolves during a season, as temperature affects parameter values, and the population sizes of fish and bryozoan follow the DFT. The effect of temperature dominates and maximizes the reproductive number, and the related risk of an outbreak, during the warmest period. With the reference parameter set, \( T > 0 \) and \( R > 1 \) throughout the season: the parasite requires the fish host to spread into the system and the DFE is always unstable. Fig. 3.4 also shows how reducing one of the transmission parameters (\( \pi_B^* \) or \( \pi_F^* \)) or the fish population size \( F_S \) by a factor 10 and 20 can curb the reproductive number below unity during the coolest periods and throughout the season, respectively.

Figure 3.3 – Values of the reproductive number \( R \) of system (3.1) as a function of the model parameters. Two parameters are varied at a time, the others are assumed as in Table 3.1. Regions where \( T < 0 \) (DFE unstable) are coloured in light red. Black dots refer to the reference parameter set. All rates are expressed as mean times (i.e., by their inverse). With regards to temperature-dependent parameters, their value at 15° C is displayed.
3.2. Results

Figure 3.4 – Short-term risk of invasibility. Reproductive number $R$ as a function of time. Water temperature is approximated as a sinusoidal signal (as for the computation of the DFT) and temperature-dependent parameters are taken as in Fig. 3.1. $F_S(\tau)$ and $B_S(\tau)$ follow the DFT. The red line refers to the reference parameter set reported in Table 3.1. The blue and green lines show the effect of reducing one of the transmission parameters ($\pi_B^*$ or $\pi_F^*$) or the fish population size $F_S$ by a factor 10 and 20, respectively.

3.2.3 Sensitivity analyses of the full model

Parasite invasion

An exploration of the values of $\lambda_{max}$ over wide ranges of the parameter space is reported in Fig. 3.5. Numerical results show high invasibility of the system ($\lambda_{max} > 1$) for wide ranges of realistic parameters. The DFT becomes stable if one of the contamination rates ($\pi_B^*$ and $\pi_F^*$, Fig. 3.5a) or if the characteristic sizes of bryozoan and fish population ($\rho^{-1}$ and $\xi^{-1}$, Fig. 3.5b) are small. Recovery mechanisms ($\psi$ and $\zeta$, Fig. 3.5f) can promote the stability of the DFT: in particular, the DFT is stable if infection-clearing in bryozoans is fast enough, whereas if only $\zeta$ is increased, the persistence of the infection in bryozoans hinders the stability of the DFT. The DFT is predicted to be stable also if the rate of transition from overt to covert infection $d_{OC}$ is considerably higher than the antithetic rate $d_{CO}$ (Fig. 3.5c), although this circumstance occurs for unlikely values of these parameters. Finally, stability of the DFT is observed for extremely low values of the fish reproduction and recovery rates ($\eta$ and $\gamma$, Fig. 3.5e).

When the DFT is unstable, the maximum modulus $\lambda_{max}$ of the eigenvalues of matrix $VW$ can be interpreted as the rate at which model trajectories depart from the disease-free trajectory. $\lambda_{max}$ thus represents a measure of how fast the parasite spreads into a susceptible population. When $\lambda_{max}$ is slightly greater than unity, the outbreak develops slowly and stochastic effects (e.g. due to the demographic stochasticity of fish and bryozoan infected populations), which are not accounted for in the current model formulation, may lead to the extinction of the disease. When the DFT is stable, the closer $\lambda_{max}$ is to unity, the slower perturbations to the DFT fade out and the system returns to a disease-free state. Therefore, $\lambda_{max}$ can also be seen as an approximate estimate of the actual risk of invasion. In this perspective, Fig. 3.5 provides information on the role of the model parameters in promoting the establishment of PKD in fish populations. $\pi_B^*$ and $\pi_F^*$ have analogous effects in enhancing the risk for an outbreak. $\rho^{-1}$, $\xi^{-1}$ and the fish baseline fertility rate $\eta$ are also positively correlated with $\lambda_{max}$. Concerning the effects of transmission and mortality rates, the highest values of $\lambda_{max}$
Chapter 3. Analysis of the local PKD model

Figure 3.5 – Sensitivity analysis: Parasite invasion. Values of $\lambda_{\text{max}}$ as a function of model parameters $\pi^*_B$ vs. $\pi^*_F$ (a); $\rho$ vs. $\xi$ (b); $d_{OC}$ vs. $d_{CO}$ (c); $b$ vs. $h$ (d); $\gamma$ vs. $e$ (e); $\psi$ vs. $\zeta$ (f); temperature (g). Pink dashed lines identify feasible parameter ranges. Black dots refer to the reference parameter set. For the sake of clarity, all rates are expressed as mean times (i.e. by their inverse). With regards to temperature-dependent parameters, their value at 15°C is displayed. Black solid lines in panel g identify levels of mean temperature during the warm season.

are found when the disease development rate $h$ is maximum and PKD-caused mortality $a$ is minimum (Fig. 3.5d). This condition maximizes the number of non-fatal infection in fish, thereby causing an increase in the release of spores $Z_F$. As for the thermal regime, the mean water temperature during the warm season stands as the main factor controlling $\lambda_{\text{max}}$ (Fig. 3.5g).

PKD-induced fish loss

Fig. 3.6 shows the results of the sensitivity analysis of PKD impact on fish population size. The residual population size is larger (i.e. PKD impact is reduced) for low values of the contamination rates and the bryozoan carrying capacity (Fig. 3.6a and b), whereas no sensitivity to high values of these parameters is observed. The bryozoan baseline growth rate (Fig. 3.6b) has no effect on PKD-induced fish loss. Fish population is preserved when $d_{OC}$ is high and $d_{CO}$ is low,
3.2. Results

Figure 3.6 – Sensitivity analysis: PKD-induced fish loss as a function of model parameters $\pi_B^*$ vs. $\pi_F^*$ (a); $\rho$ vs. $\xi$ (b); $d_{OC}$ vs. $d_{CO}$ (c); $a$ vs. $h$ (d); $\gamma$ vs. $\epsilon$ (e); $\psi$ vs. $\zeta$ (f); temperature (g). Simulations are run until convergence (100 seasons), with each season lasting for 200 days. Colors refer to the percentages of the population size at the end of the last season with respect to the same quantity calculated along the disease-free trajectory [computed as in Eqs. (3.3) and (3.4)]. For example, 50% means that at the end of the last season the population size is half of the population that would have survived if the disease were absent. State variables at the beginning of the first season are set as in the model simulation of Fig. 3.2. Pink dashed lines identify feasible parameter ranges. Black dots refer to the reference parameter set. All rates are represented as mean times. With regards to temperature-dependent parameters, their value at 15 °C is displayed. Black solid lines in panel g identify levels of mean temperature during the warm season.

while the opposite case does not result in a severe population decay (Fig. 3.6c). As expected, there is a positive correlation between PKD-induced fish loss and PKD-induced mortality rate $a$ (Fig. 3.6d). However, if $a$ is extremely high, the immediate death of the host limits the transmission of the disease and thus its impact, a common feature of epidemiological models. PKD impact is enhanced by the rate of disease development $h$. Fish population decay is also enhanced when both $\epsilon$ and $\gamma$ are low, as this condition maximizes the abundance of acutely infected fish (Fig. 3.6e); as previously pointed out, if $\gamma$ is extremely low, this effect is balanced by the augmented relative importance of $a$, which weakens PKD impact. As the recovery time of bryozoans, $\psi^{-1}$, decreases, PKD impact monotonically decreases (Fig. 3.6f). On the other hand, PKD impact reaches a peak for intermediate rates of fish recovery ($\zeta \approx 0.005$ d$^{-1}$),
presumably due to concomitant high level of contamination and increased availability of susceptible fish. As expected, higher water temperatures promote higher fish mortality (Fig. 3.6g). In particular, PKD impact is mostly related to the mean temperature during the warm season, whereas the amplitude of the temperature sinusoidal signal plays a lesser role. For instance, by assuming a relative mid-amplitude of 0.5, an increase of 5 °C (from 10 to 15 °C) in the seasonal mean water temperature would produce an overall reduction of 26% of the total fish population.

3.3 Discussion

According to the modelling framework here developed, the disease-free trajectory is found to be unstable over wide ranges of the ecological and epidemiological parameters. This property means that the introduction of the parasite in a fully susceptible community of salmonids and bryozoans would very likely lead to a PKD outbreak and to long-term parasite establishment. The ability of the parasite to cycle between covert, mildly virulent, and overt, transmissive phases in the bryozoan host allows the parasite to produce large numbers of transmissive stages without compromising the susceptible host population. Similarly, long-term infections in trout are also responsible for continuous infections of naive bryozoan hosts and promote parasite persistence. This result holds for the deterministic model (2.1) and (2.4), although stochastic effects, currently not accounted for in the model formulation, could actually prevent the invasion. Such findings obviously depend on the current knowledge on the transmission modes of PKD, and on how they have been translated into mathematical terms. Indeed, the analyses of the within-season and the full models show how recovery mechanisms for both bryozoans and fish reduce the risk of outbreak and can make the disease-free system stable. However, knowledge and investigations on these critical processes are still scant. Concerning bryozoans, Tops et al. [2009] found that infection cleared in 15 of 40 F. sultana colonies (37%) after overt infections had disappeared; Abd-Elfattah et al. [2014a] argued that observed production of infected statoblasts from uninfected F. sultana might have been due to a recent loss of infection of the maternal colony. As regards infection clearing in fish, experiments conducted on rainbow trout [Schmidt-Posthaus et al., 2012] revealed hints of a possible infection recovery in salmonids. Further studies on this topic are needed to better elucidate the mechanisms underlying the persistence of T. bryosalmonae in fish and bryozoan communities. Overall, these results underline the reasons for the emerging status of PKD throughout Europe – the extremely successful invasion mechanism of the parasite may have facilitated its historical spread in European streams, and the current changes in climate and other environmental conditions are causing it to proliferate.

These results highlight how both outbreak risk and long-term persistence of the parasite are critically enhanced by warmer water temperatures. The design of possible control strategies (e.g. the control of the population of one of the host or the interactions among them) should thus take into account that warmer periods pose a higher risk of a PKD outbreak. On the other hand, temperature effects offer the opportunity to design alternative intervention strategies.
3.3. Discussion

aimed at controlling stream water temperature through e.g. tree shading or the selective release of cold water from upstream reservoirs. The analysis of the within-season model also shows that, for certain parameter combinations, the parasite can initially spread even in the absence of the fish host. This possible behaviour is determined by the fact that the model assumes that infected bryozoan grow and that *T. bryosalmonae* can simultaneously proliferate inside them. This specific dynamics should be better scrutinized with additional experimental studies because the possibility of epidemics hosted solely by the bryozoan population has relevant implications for control strategies. It implies, for instance, that strategies focusing only on the fish population or on limiting the interactions between the two hosts might not be effective, under certain conditions, to prevent the invasion of the parasite.

Simulations and sensitivity analyses provided insights about the role and importance of the different parameters in parasite invasion and outbreak severity. In particular, the analysis of PKD-induced fish loss highlights the crucial role of the parameters that govern the transition between epidemiological classes in determining the impact of PKD outbreaks when the disease is endemic. While the carrying capacity of bryozoans and the contamination rates are the main factors controlling the stability of the disease-free trajectory, they appear less relevant in determining PKD impact in endemic settings. The underlying reason for this result is that these parameters solely control the rate at which susceptible fish are exposed to the parasite, whereas the infection develops over time scales ruled by the parameters governing the transition between epidemiological classes. A promising aspect is the fact that such parameters and their temperature dependence could be estimated in laboratory experiments by exposing hosts to the parasite and by monitoring the development of their infection status. Similar studies have already been carried out (see corresponding references in Table 3.1).

More are needed, however, to better elucidate the role of temperature on PKD dynamics. It should be noted, in fact, that most of the available literature on PKD is based on rainbow trout (*Oncorhynchus mykiss*), which is known to be a dead-end host for *Tetracapsuloides bryosalmonae*, as this species does not allow the completion of the parasite cycle by further infecting bryozoans [Kumar et al., 2013]. Conversely, only few studies have focused on brown trout [Grabner and El-Matbouli, 2009; Kumar et al., 2013; Abd-Elfattah et al., 2014b].

While it could be relevant to experimentally assess how the release of spores (by both infected fish and bryozoans) varies with temperature, the estimation of the actual value of the release rate may be less important. Indeed, the analysis of the model reveals that transmission dynamics are controlled by the product between contamination and exposure rates. The latter can hardly be estimated under field conditions as they depend on the probability of a successful encounter between viable spores and hosts. Therefore, efforts to precisely estimate contamination rates would be frustrated by the large uncertainties associated to exposure rates. The proposed rescaled model (2.3) features two parameters (one per host) that represent the product between contamination and exposure rates. Such parameters are key to disease dynamics, as discussed above. However, reasonable values can hardly be *a priori* estimated and they would need to be calibrated for each case study by contrasting model simulations with experimental or field data.
Chapter 3. Analysis of the local PKD model

Knowledge and literature on bryozoans are rather scant compared to the vast and traditional literature on population dynamics and habitat distribution of salmonids. To improve the current understanding of PKD as an emerging disease, such knowledge gap must be filled. In particular, knowledge of bryozoan habitat suitability needs to be improved in order to map the risk of PKD invasion. This task could be achieved by using species distribution models [Elith and Leathwick, 2009; Guisan et al., 2017] to relate the presence/absence of bryozoans to environmental variables (e.g. hydrological conditions, characteristics of the river-bed and banks, water quality, temperature). The few existing studies (e.g. Økland and Økland [2001]) focus on lakes, while large-scale studies of habitat suitability in streams and rivers are not available yet. Furthermore, many aspects of the relation between infection status and bryozoan proliferation are still to be elucidated: for example, population genetic diversity of bryozoans may be linked to their ability to resist infections; infective stages may propagate in partially infected bryozoan zooids even in the absence of further exposure to T. bryosalmonae.

The results here presented need be accompanied by an assessment of the limits of the model. It is acknowledged that young-of-the-year fish are highly likely to contract PKD when exposed to the parasite for the first time [Schager et al., 2007]. However, for the sake of simplicity, the age structure of the fish population was not accounted for in the model formulation analyzed in this Chapter. As formulated in §2.3, fish population can indeed be split into age-specific classes (e.g. juveniles and adults), each with its own epidemiological compartments with possibly different mortality, recovery and infection rates. The epidemiological model also relies on the simplifying hypothesis that all statoblasts hatch at the beginning of the following season. Hence, another aspect needing further investigation is the possibility of the existence of a statoblast bank (as suggested by Freeland et al. [2001]). A statoblast bank could form on the bed of the water body if statoblasts were able to survive and maintain their infection status for more than one season. Dormant infected statoblasts could thus trigger new PKD outbreaks even if the disease were no longer present in fish and bryozoans.

Although the described transmission processes imply proximity between spores and hosts, several mechanisms allow long-distance spreading of PKD along river networks and lakes, among which bryozoan fragmentation, buoyancy of surfaces with attached colonies, fish migration, and hydrodynamic dispersal of T. bryosalmonae spores and bryozoan statoblasts may play a remarkable role [Okamura et al., 2011]. To understand the role played by spatial processes in PKD transmission, the local model presented herein was extended by considering the hydrological connectivity among different river reaches (see §2.2). Modelling efforts at the river network scales (corroborated by field analyses) should try to understand whether PKD transmission is a spatially diffuse process or rather concentrated in transmission hot-spots. Such hot-spots could be represented by favorable habitats for both salmonids and bryozoans. The proximity of the hosts in these habitats could promote high transmission rates that may in turn be able to sustain the infection at the network scale through dispersal of fish and hydrodynamic transport of parasite spores. The possible identification of these hot-spots could offer alternative strategies to control the disease in the wild, like the confinement or the removal of one of the hosts in the isolated hot-spots. Such key issues constitute the focus of
3.4 List of symbols

The following table lists all symbols used in this Chapter. This table does not include state variables and parameters of the epidemiological model (presented in Tables 2.1 and 3.1, respectively).

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbf{A}$</td>
<td>Jacobian of a reduced within-season system including only non-susceptible classes</td>
<td>various</td>
</tr>
<tr>
<td>$B_S(\tau)$</td>
<td>Time evolution of the susceptible bryozoan population along the DFT</td>
<td>$\text{ML}^{-3}$</td>
</tr>
<tr>
<td>$B_{S,0}$</td>
<td>Susceptible bryozoan population at $\tau_0$ along the DFT</td>
<td>$\text{ML}^{-3}$</td>
</tr>
<tr>
<td>$F_{S,0}$</td>
<td>Susceptible fish population at $\tau_0$ along the DFT</td>
<td>$\text{L}^{-3}$</td>
</tr>
<tr>
<td>$J_{DF}$</td>
<td>Jacobian of the reduced within-season system (3.1)</td>
<td>various</td>
</tr>
<tr>
<td>$J_1$</td>
<td>Group of epidemiological parameters appearing in the definitions of $R$ and $T$</td>
<td>$T^{-7}$</td>
</tr>
<tr>
<td>$J_2$</td>
<td>Group of epidemiological parameters appearing in the definitions of $R$ and $T$</td>
<td>$T^{-5}$</td>
</tr>
<tr>
<td>$R$</td>
<td>Reproduction number for the linearized within-season system (3.1)</td>
<td>-</td>
</tr>
<tr>
<td>$R(\tau)$</td>
<td>Integral of bryozoan growth rate $r$ from $\tau_0$ to $\tau_1$</td>
<td>-</td>
</tr>
<tr>
<td>$T$</td>
<td>Threshold value allowing persistence of PKD (when $T &lt; 0$) in a local bryozoan community in the absence of fish population</td>
<td>$T^{-2}$</td>
</tr>
<tr>
<td>$V$</td>
<td>Operator describing within-season dynamics for a reduced system including only non-susceptible classes (from $\tau_0$ to $\tau_1$)</td>
<td>various</td>
</tr>
<tr>
<td>$v$</td>
<td>Operator describing within-season dynamics for a reduced system including only susceptible classes</td>
<td>various</td>
</tr>
<tr>
<td>$W$</td>
<td>Operator describing between-season dynamics for a reduced system including only non-susceptible classes</td>
<td>various</td>
</tr>
<tr>
<td>$w$</td>
<td>Operator describing between-season dynamics for a reduced system including only susceptible classes</td>
<td>various</td>
</tr>
<tr>
<td>$x_{DF}$</td>
<td>DFE state for the reduced within-season system (3.1)</td>
<td>various</td>
</tr>
<tr>
<td>$x_{DF}$</td>
<td>DFT for the full system (2.1) and (2.4)</td>
<td>various</td>
</tr>
<tr>
<td>$x_S$</td>
<td>State of a reduced system including only non-susceptible classes</td>
<td>various</td>
</tr>
<tr>
<td>$x_{S}$</td>
<td>State of a reduced system including only susceptible classes</td>
<td>various</td>
</tr>
<tr>
<td>$x_{S}$</td>
<td>Fixed point of the Poincaré map $w \circ v$</td>
<td>various</td>
</tr>
<tr>
<td>$Z(\tau)$</td>
<td>Operator describing within-season dynamics for a reduced system including only non-susceptible classes (from $\tau_0$ to $\tau$)</td>
<td>various</td>
</tr>
<tr>
<td>$\alpha_s$</td>
<td>Spectral abscissa of a matrix</td>
<td>various</td>
</tr>
<tr>
<td>$\eta_{\text{min}}$</td>
<td>Minimum value of the fish reproduction rate $\eta$ below which fish population is destined to extinction</td>
<td>-</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Eigenvalues of a matrix</td>
<td>various</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$</td>
<td>Maximum eigenvalue of matrix $VW$</td>
<td>-</td>
</tr>
</tbody>
</table>
4 PKD spread along river networks

The spatially-explicit metacomunity framework for PKD described in §2.2 is here used to study the spatial effects of the disease spread in idealized stream networks. In particular, network connectivity effects are investigated by running simulation experiments on synthetic river network replicas derived from Optimal Channel Networks (OCNs), spanning trees known to reproduce all the mutually connected topological and metric features of real rivers. Results show how network connectivity can produce heterogeneous patterns of PKD prevalence even when the underlying spatial distributions of fish and bryozoans are homogeneous. Prevalence is generally higher at the downstream sites: if fish mobility is neglected, the spatial distribution of prevalence follows that of the upstream drainage area; otherwise, prevalence patterns are correlated with the proximity to the outlet. Downstream invasion speed of PKD is generally high, due to the fast dynamics of hydrological spore transport. For the tested values, effects of water temperature on prevalence heterogeneity are minor. However, climate change may increase invasion speed in both downstream and upstream directions. PKD can establish in bryozoan-free river reaches, on the condition that the infection be sustained by upstream or downstream hot-spots. These results further the understanding of the drivers of disease incidence distribution in riverine ecosystems.¹

Chapter 4. PKD spread along river networks

4.1 Methods

The spatial PKD model presented in §2.2 is here applied to Optimal Channel Networks (OCNs), i.e. spanning dendritic structures obtained by minimizing total energy dissipation along drainage directions [Rodriguez-Iturbe et al., 1992a,b; Rinaldo et al., 1992, 2014]. OCNs reproduce all the related scaling features observed in real river networks and provide a robust characterization of the topological and metric features of channel networks [Rodriguez-Iturbe and Rinaldo, 2001]. The use of OCNs allows to run computational experiments on different river network replicas with the same statistics of morphological features, thus deriving results averaged over several realizations, i.e. particular shapes of a single (real) network grown within the same total contributing area. Such a feature has already favored the application of OCNs in the study of ecological and epidemiological dynamics in dendritic networks [Carrara et al., 2012, 2014; Gatto et al., 2013; Mari et al., 2014a,b; Bertuzzo et al., 2015, 2016].

This section presents the algorithm allowing the generation of Optimal Channel Networks, as well as the assumptions made concerning the geomorphological and hydrothermal features of the river stretches. The set-up of the numerical experiments is then described. If not specified otherwise, parameter values are taken from Table 4.1.

4.1.1 Generation of synthetic river networks

The first step for the generation of synthetic river networks draining a given area is the reliable replication of the topological and metric structure of the network. To this end, it is convenient to exploit the concept of OCNs, which are spanning trees minimizing a functional describing total energy dissipated along drainage directions by landscape-forming discharges [Rinaldo et al., 2014]. In a given drainage domain, the OCN constructs allow to generate replicas characterized by different configurations (owing to random search) with identical statistical features indistinguishable from those of real river networks. We consider landscapes formed by square lattices of side $\delta_p$ composed of $N_p = \delta_p^2$ pixels, whose side is $L_p$. Each pixel is connected to one of its eight nearest neighbours, forming a spanning tree where the only root a single outlet (multiple outlets might have also been employed, but arrangements of this type seemed unnecessary at this stage). The framework allowing the generation of OCNs is here briefly described for completeness.

Let $A_i$ be the total number of pixels upstream of each pixel $i$ of the landscape, i.e. the contributing area (in pixel units) to pixel $i$. A set of contributing areas $S = [A_1, \ldots, A_{N_p}]^T$ determines the tree configuration. Landscape-forming discharges are considered to be proportional to the contributing area ($Q_i \propto A_i$), a widely accepted assumption [Rodriguez-Iturbe and Rinaldo, 2001]. Along the $i$-th link of the network, energy dissipation is $H_i \propto Q_i \Delta z_i$, with $\Delta z_i$ being the drop in elevation, which in turn is assumed to depend on contributing area:

$$\Delta z_i \propto A_i^{-0.5},$$  (4.1)
### 4.1. Methods

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline growth rate of $B_S$, $B_C$</td>
<td>$r$</td>
<td>0.004 · $T$</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Baseline growth rate of $B_O$</td>
<td>$r_O$</td>
<td>0.002 · $T$</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Inverse of carrying capacity</td>
<td>$\rho$</td>
<td>0.05</td>
<td>m$^3$g$^{-1}$</td>
</tr>
<tr>
<td>Rate of covert-to-overt transition</td>
<td>$d_{OC}$</td>
<td>0.00032 · $T^2$</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Rate of covert-to-overt transition</td>
<td>$d_{CO}$</td>
<td>0.00032 · $T^2$ - 0.016 · $T$ + 0.2</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Recovery rate</td>
<td>$\psi$</td>
<td>0.01</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Rate of contamination operated by $B_O$</td>
<td>$\pi_B^*$</td>
<td>$\beta_B \pi_B$</td>
<td>m$^3$g$^{-1}$d$^{-2}$</td>
</tr>
<tr>
<td>Overwintering probability ($B_S$, $B_C$)</td>
<td>$\sigma_B$</td>
<td>0.1</td>
<td>-</td>
</tr>
<tr>
<td>Overwintering probability ($B_O$)</td>
<td>$\sigma_O$</td>
<td>0.05</td>
<td>-</td>
</tr>
<tr>
<td>Biomass generated by one statoblast</td>
<td>$\nu$</td>
<td>0.035</td>
<td>g</td>
</tr>
<tr>
<td>Statoblast production rate</td>
<td>$f_B$</td>
<td>0.005 · $T$ + 0.00025 · $\tau$</td>
<td>g$^{-1}$d$^{-1}$</td>
</tr>
<tr>
<td>Fraction of $B_C$ generating $S_S$</td>
<td>$\phi$</td>
<td>0.7</td>
<td>-</td>
</tr>
<tr>
<td>PKD-caused mortality rate</td>
<td>$a$</td>
<td>0.00016 · $T^2$</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Rate of disease development</td>
<td>$h$</td>
<td>0.00012 · $T^2$</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Rate of recovery from acute infection</td>
<td>$\gamma$</td>
<td>0.02</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Rate of complete recovery</td>
<td>$\zeta$</td>
<td>0.001</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Fraction of acute infections</td>
<td>$\epsilon$</td>
<td>0.1</td>
<td>-</td>
</tr>
<tr>
<td>Rate of contamination operated by $F_I$</td>
<td>$\pi_F^*$</td>
<td>$\beta_F \pi_F$</td>
<td>m$^3$d$^{-2}$</td>
</tr>
<tr>
<td>Relative cont. rate operated by $F_C$</td>
<td>$\kappa$</td>
<td>0.2</td>
<td>-</td>
</tr>
<tr>
<td>Natural mortality rate</td>
<td>$\mu_F$</td>
<td>5 · 10$^{-4}$</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Spore mortality rate</td>
<td>$\mu_Z$</td>
<td>1.33</td>
<td>d$^{-1}$</td>
</tr>
<tr>
<td>Overwintering probability</td>
<td>$\sigma_F$</td>
<td>0.9</td>
<td>-</td>
</tr>
<tr>
<td>Baseline reproduction rate</td>
<td>$\eta$</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Strength of density dependence</td>
<td>$\xi$</td>
<td>2</td>
<td>m$^3$</td>
</tr>
<tr>
<td>Relative overwintering probability ($F_E$)</td>
<td>$p_E$</td>
<td>0.57</td>
<td>-</td>
</tr>
<tr>
<td>Relative overwintering probability ($F_I$)</td>
<td>$p_I$</td>
<td>0.55</td>
<td>-</td>
</tr>
<tr>
<td>Duration of the warm season</td>
<td>$N_d$</td>
<td>250</td>
<td>d</td>
</tr>
</tbody>
</table>

Table 4.1 – Reference set of parameters for the spatial PKD model. Water temperature $T$ and time $\tau$ are in Celsius degrees and ordinal days, respectively.

as suggested by Rodriguez-Iturbe and Rinaldo [2001]. OCNs are thus network configurations such that the functional

$$H = \sum_{i=1}^{N_v} H_i \propto \sum_{i=1}^{N_v} A_i^{0.5}$$

is a local, dynamically accessible minimum. After picking an arbitrary pixel at a border of the lattice as the network outlet, the search for the set of contributing areas $S'$ that minimizes $H$ is operated by means of a simulated annealing strategy to mimic the ability of the OCN algorithm to find feasible (i.e. dynamically accessible) local optima. Interestingly, the search is confined to a comparison of spanning trees (that is, looping structures are excluded a priori) for a number of theoretical and practical reasons (e.g. Rinaldo et al. [2014]). The random search procedure warrants that each generated OCN is a different spanning tree.

Each derived OCN thus defines a tree structure that spans the whole landscape. In the real world, however, a drainage path becomes a stream only when certain hydrological conditions are met. The simplest method assumes that pixels form a channel when their contributing
Figure 4.1 – a) Example of an Optimal Channel Network (OCN). The elevation map has been obtained by extrapolating Eq. (4.1) to non-channelized pixels; while this hypothesis is not generally valid in real landscapes, it has no implication for this work. b) 3D landscape generated by the OCN depicted in panel a. c) Another OCN in the same domain with a different localization of the outlet pixel. d) Distribution of mean fish residence times for the OCN presented in panel c. Mean residence times are derived using Eq. (2.8), assuming $l_{avg} = 0.02 \text{ d}^{-1}$ and spatially uniform fish density. Residence times generally increase with increasing stretch length and distance from outlet. In the presented approach, these relationships stem from purely geomorphic factors: the former relates to the fact that the bigger the stretch, the longer it takes for a fish to exit from it; the latter to the imposition of reflecting boundary conditions for fish movement at headwater locations. Interestingly, such pattern is also supported by the findings of Zimmer et al. [2010], where it is argued that lower tendency to movement for fish located in upstream sites may be due also to smaller variations in water temperature, which make the habitat more suitable for trout.

area (a proxy of landscape-forming discharge) exceeds a certain threshold $A_T$ [O’Callaghan and Mark, 1984; Tarboton et al., 1991]. The network is then discretized into stretches, each defined as a sequence $T$ of channelized pixels starting from one pixel having either zero (river sources) or more than one channelized upstream pixels (confluences), and containing the downstream sequence of channelized pixels until another confluence or the outlet are reached. The obtained network of stretches is an oriented graph suitable for the application of the PKD metacommunity model.

The second step in the generation of synthetic river networks consists in the definition of the geomorphological properties of each river stretch. Cross-sections are approximated with...
4.1. Methods

rectangular shapes having width $B$, depth $D$ and average water velocity $U$. To account for how these geometric variables change along the network, classic results of river geomorphology are exploited [Leopold and Maddock, 1953], according to which $B \propto Q^{0.5}, D \propto Q^{0.4}$ and $U \propto Q^{0.1}$. Note that, for consistency’s sake, the sum of the three exponents must be equal to unity. By invoking the proportionality between landscape-forming discharge and contributing area at-a-site, one has

$$B_k = B_O \hat{A}_k^{0.5}; \quad D_k = D_O \hat{A}_k^{0.4}; \quad U_k = U_O \hat{A}_k^{0.1}.$$  (4.2)

where the subscript $k$ identifies a generic stretch $k$, and $B_O$, $D_O$ and $U_O$ are the maximum values (i.e. at the outlet) for width, depth and velocity, respectively; $\hat{A}_k = A_{ik} / \delta_p^2$ is the normalized contributing area to stretch $k$; subscript $i_k$ identifies the last downstream pixel of stretch $k$ (the last element of $T_k$). The discharge at the outlet of the catchment then reads $Q_O = B_O D_O U_O$; the water volume in a generic stretch $k$ is $V_k = B_k D_k L_{s,k}$, where $L_{s,k} = \sum_{i \in T_k} L_i$ and $L_i = L_p$ if the flow direction along pixel $i$ is parallel to a pixel side or $L_i = \sqrt{2} L_p$ if parallel to the pixel diagonal.

Eq. (4.1) allows constructing the landscape whose drainage directions match the network stretches. The elevation drop along a network pixel can indeed be expressed as $\Delta z_i = s_o \hat{A}_i^{-0.5} L_i$, where $s_o$ is the slope at the outlet pixel. This relationship can be iteratively applied starting from the outlet pixel (where an arbitrary elevation $z_o$ is imposed) towards all upstream paths, in order to reconstruct the elevation of each channelized pixel. The elevation of a stretch can be defined as the average elevation among all constituting pixels.

Some examples of OCNs obtained by changing only the location of the single outlet pixel chosen are shown in Figs. 4.1a, 4.1b, and 4.1c. Scale parameters (Table 4.2) that define the metric of the river network were chosen to be representative of a prealpine catchment of around 1000 km$^2$.

4.1.2 Water temperature patterns

The seasonal cycle of stream water temperature typically follows that of air temperature, albeit being damped and possibly delayed. However, notable deviations can be observed in streams with large impoundments or lakes upstream, or when the thermal regime is dominated by ice or snow melting [Mohseni et al., 1998; Caissie, 2006; Toffolon and Piccolroaz, 2015]. For this exercise, it is assumed that water temperature at the outlet reach $T_O(\tau)$ follows a sinusoidal function with period equal to one year (Table 4.2). To derive time series of water temperature for all network stretches, it is further assumed that water temperature mirrors the environmental lapse rate of air temperature. Water temperature in a generic stretch $k$ is then

$$T_k(\tau) = T_O(\tau) + \Gamma_w \Delta Z_k$$  (4.3)
Table 4.2 – Scale parameters for generation of Optimal Channel Networks and water temperature time series. Note that, with the chosen parameters, the value of the specific discharge \( \frac{Q_O}{A_{tot}} = 0.025 \text{ m}^3\text{s}^{-1}\text{km}^{-2} \) resembles that of several Swiss prealpine catchments (see Schädler and Weingartner [1992]). The sinusoidal signal for water temperature was fitted on data measured in River Langeten, Switzerland, in the period 2002-2013 (data provided by the Swiss Federal Office of the Environment - FOEN).

where \( \Gamma_w \) is the lapse rate and \( \Delta Z_k \) the difference in elevation with respect to the outlet. Lapse rates for air temperature can range from about -9.8 \( ^\circ\text{C}k\text{ m}^{-1} \) for dry air (dry adiabatic lapse rate - DALR) to about -4.0 \( ^\circ\text{C}k\text{ m}^{-1} \) for hot saturated air (saturated adiabatic lapse rate). It is then assumed \( \Gamma_w \) equal to -6.5 \( ^\circ\text{C}k\text{ m}^{-1} \), a typical value that is used as global mean environmental lapse rate for air temperature [Barry and Chorley, 2009].

4.1.3 Experiment 1: Connectivity effects on the heterogeneity of PKD impact

This first experiment aims at exploring the influence of connectivity in determining the spatial variability of PKD prevalence and fish loss within a river network. To isolate the effect of connectivity and morphology on prevalence heterogeneity, while excluding other possible confounding factors, a simplified set-up is used, in which fish and bryozoan abundances per unit volume are uniform in space. All stretches are initially infected.

Three possible connectivity processes that can control the spatial distribution of PKD are considered: hydrological transport of spores; fish mobility; and spatial heterogeneity of water temperature engendered by elevation gradients, which are determined by network configurations. Simulations are performed on several OCN replicas by alternatively perturbing these factors. In this regard, it is noteworthy that all OCN replicas have nearly the same elevation, i.e. every configuration that locally minimizes total energy dissipation also minimizes its mean elevation [Rinaldo et al., 1993].

Regarding fish movement, several values for \( l_{avg} \) have been tested, including \( l_{avg} = 0 \) which prevents fish mobility. Elevation, by controlling water temperature, indirectly affects several
model parameters. The effect of landscape elevation is investigated by comparing model results obtained in one landscape and in its flat counterpart, where water temperature would be equal at all sites. The artificially flat landscape is reproduced by attributing to all sites the temperature $T_f(\tau)$, which is chosen so that the total thermal energy of the river network equals that of the original landscape:

$$T_f(\tau) = \frac{\sum_{k=1}^{N_s} V_k T_k(\tau)}{\sum_{k=1}^{N_s} V_k}$$

where $T_k$ are defined as in Eq. (4.3). $T_f$ represents the weighted average of $T_k$, with weights equal to the water volumes of the various stretches.

Finally, the influence of climate change is investigated: the previous simulations are repeated with a 4 °C mean temperature increase in all stretches, as forecast by Scenario RCP8.5 proposed by the Intergovernmental Panel on Climate Change [IPCC, 2013; Collins et al., 2013].

4.1.4 Experiment 2: Speed of disease propagation

The goal of this experiment is to investigate the main factors controlling PKD propagation speed in both downstream and upstream directions. The model set-up is similar to the one used in Experiment 1: all sites have the same fish density and the same bryozoan carrying capacity per unit volume. By contrast, only one headwater stretch (the one at maximum distance from the outlet) is initially infected. Simulations are run until a steady annual cycle is reached. It is assumed that a stretch has been invaded when its prevalence attains 10% of the prevalence at equilibrium; propagation speed is then assessed by dividing the maximum headwater-outlet distance by the time elapsed from headwater invasion to outlet invasion. The same experiment is then reversed: the outlet stretch is the only initially infected site and the disease propagation up to the farthest stretch is evaluated. In analogy with the previous Experiment, the effect of climate change is also assessed.

4.1.5 Experiment 3: Influence of heterogeneity in bryozoan spatial distribution

This model set-up is designed to study the effects of uneven bryozoan spatial distributions on PKD spreading. The river network is thus partitioned in two distinct regions (downstream and upstream) based on a threshold distance to the outlet. In a first simulation, aimed at evaluating downstream disease propagation, only the upstream region is assumed to be suitable for bryozoans. Within this region, bryozoan density is constant and all sites are initially infected. Absence of bryozoans in the downstream region is ensured by setting $\rho = 10^5 \text{m}^3\text{g}^{-1}$ for the corresponding stretches, which implies that the bryozoan carrying capacity $\rho^{-1}$ is close enough to zero. A second simulation is performed by inverting the above-described pattern of bryozoan presence, which allows the analysis of upstream propagation. In both experiments, fish density is uniform over the whole network.


Chapter 4. PKD spread along river networks

4.2 Results

Experiment 1.  Fig. 4.2 shows how connectivity and fish mobility affect the distribution of disease prevalence in a flat landscape. Two main types of pattern arise: when fish mobility is negligible, the sites with higher prevalence are those whose contributing area is higher (Figs. 4.2b, 4.2c); when instead mobility rates are high, prevalence decreases as the distance to the outlet increases (Figs. 4.2a, 4.2d). Contributing area is indeed a proxy of the abundance of parasite spores that enter a given stretch, which is the main driver of infection in the absence of fish movement. In this case, small headwater stretches in the proximity of the outlet are not invaded by the parasite (Fig. 4.2b). Conversely, high mobility rates enhance the mixing process among local communities, with the result that the variability in prevalence among neighbouring stretches is low and the average network prevalence increases. When mobility rates are tuned to more realistic values (i.e., by assuming the distribution of mean residence times of Fig. 4.1d), the distribution of prevalence shows an intermediate behaviour (blue lines in Figs. 4.2c, 4.2d). Patterns of fish loss exhibit an analogous trend.

The inclusion of an elevation gradient has, at least in this setting, a minor impact on the distribution of prevalence (Fig. 4.2e). As a consequence of diminished temperature, a decrease of prevalence in upstream stretches is observed, which is not compensated by a corresponding increment in lower altitude stretches (which is already close to 100% with this parameter set). This causes a small reduction in the average network prevalence (about -1.5% regardless of fish mobility). As for the distribution of fish loss (Fig. 4.2f), the effects of non-uniform temperature are slightly accentuated. At the outlet, fish loss is generally higher (+4.3% when \(l_{avg} = 0\)) but the average fish loss does not change substantially with respect to the flat landscape case.

The application of climate change leads to enhanced fish loss, especially in the downstream region (Fig. 4.2f); with the reference parameter set and \(l_{avg} = 0\), the average network fish loss under climate change is equal to 87.4% (70.3% in the default case). The corresponding moderate decrease in prevalence is due to the augmented mortality rate \((a \gg \gamma\) in Eq. (2.7h)), which hampers most of the fish from surviving the acute phase of the infection and entering the carrier state.

Experiment 2.  A sensitivity analysis of disease propagation speed with respect to the contamination rate \(\pi^*\) and the fish mobility rate \(l_{avg}\) has been conducted. Downstream propagation (Fig. 4.3b) generally occurs after one to three years, much faster than upstream propagation (Fig. 4.3e), as a consequence of the bias in the hydrological transport of spores and of its fast dynamics. As expected, both fish mobility and contamination rates are positively correlated with propagation speed in both directions, although the role of \(l_{avg}\) in downstream propagation is minor. When both parameters are set to low values, PKD might not establish in the network (grey regions in Figs. 4.3b, 4.3e). Note that, while the absence of PKD at the outlet stretch implies that the whole network is disease-free, this is not necessarily true with regards to the headwaters (compare Figs. 4.3a, 4.3d with Fig. 4.3g). Effects of variations of the contamination
4.2. Results

Figure 4.2 – Experiment 1. Left and central columns: effect of the magnitude of fish mobility rates on PKD prevalence. Simulations are run for 50 years, the prevalence at the end of the 50th year is shown. A flat landscape is assumed. a) Prevalence map for a given OCN with $l_{\text{avg}} = 0.2 \text{ d}^{-1}$. b) Prevalence map in absence of fish mobility. c) Prevalence as a function of contributing area. For 10 different OCNs, prevalence at each stretch is evaluated. Solid lines represent mean trends; shaded areas identify 25th-75th percentiles of the distribution. d) Prevalence as a function of relative distance to the outlet. Note that, for the same value of $l_{\text{avg}}$, mean prevalences for groups of stretches with smallest contributing areas and highest distance from outlet are generally different, as these two subsets do not match exactly. Right column: effect of elevation gradient and climate change (CC) on prevalence (e) and fish loss (f) when $l_{\text{avg}} = 0 \text{ d}^{-1}$. Symbols are as in panel c. Epidemiological parameters are set to their reference value (Table 4.1).

Rate $\pi_B^*$ are similar to those observed for $\pi_F^*$ (not shown).

Downstream propagation consists of three main phases (Fig. 4.3c): during the first period (less than one year, for the reference parameter set and $l_{\text{avg}} = 0.02 \text{ d}^{-1}$), upstream prevalence is higher and the propagation front moves downstream (phase I in Fig. 4.3c); then, the outbreak in the downstream region implies a local increase in prevalence, while upstream prevalence tends to remain constant (phase II); finally, after downstream prevalence reaches almost 100%, upstream prevalence increases owing to fish movement (phase III). A steady distribution of prevalence is attained after around 10 years. If both $\pi_F^*$ and $l_{\text{avg}}$ were strongly increased, PKD would immediately invade the outlet stretch, hence skipping the first transmission phase (pale pink area in Fig. 4.3b). On the other hand, upstream invasion (Fig. 4.3f) is characterized by a
Chapter 4. PKD spread along river networks

**Figure 4.3** – Experiment 2. Left column: downstream propagation. Central column: upstream propagation. a), d) Prevalence at equilibrium in the stretch being farthest from the initially infected point (namely, a: prevalence at outlet; d: prevalence at farthest headwater) as a function of $\pi_F^*$ and $l_{avg}$. b), e) Propagation speed as a function of $\pi_F^*$ and $l_{avg}$. Target stretches are assumed to be disease-free if their prevalence at equilibrium is lower than 0.01 (grey areas). In the pale pink region, invasion of the target site (defined as the time when local prevalence reaches 10% of the local equilibrium prevalence at the end of the warm season) occurs earlier than that of the initially infected site. Sharp color transitions (as in panel b, for $\pi_F^* \approx 0.3 \text{ m}^3\text{d}^{-2}$) indicate that invasion of the target stretch occurs in different years. c), f) Evolution of prevalence along the path between the outlet and the farthest stretch (highlighted in dark colors in Fig. 4.4c), with $\pi_F^* = 0.2 \text{ m}^3\text{d}^{-2}$ and $l_{avg} = 0.02 \text{ d}^{-1}$. Dots represent stretches along the path. Dashed lines refer to the case where climate change is applied. Black arrows in panel c indicate phases of downstream invasion. g) Average network prevalence at equilibrium as a function of $\pi_F^*$ and $l_{avg}$; this panel refers to both directions of propagation. All simulations were conducted on the OCN showed in Fig. 4.1c, whose maximum path length is equal to 54.11 km. All non-specified parameters are as in Table 4.1.
4.2. Results

Figure 4.4 – Experiment 3. a) Downstream propagation: profile of prevalence at a steady state. Dots represent stretches along the path from the farthest site to the outlet. The dashed black line identifies the border between bryozoan-suitable and bryozoan-free areas. DO: Distance from outlet. b) Upstream propagation: profile of prevalence at a steady state. Symbols as in panel a. c) Partition of the network in upstream (violet - \( \text{DO} \geq 35 \text{ km} \)) and downstream (green - \( \text{DO} < 35 \text{ km} \)) regions. The path used in panels a and b is highlighted in darker colors. Bottom panels: prevalence maps. d) Downstream propagation, \( l_{\text{avg}} = 0.2 \text{ d}^{-1} \). e) Downstream propagation, \( l_{\text{avg}} = 0 \text{ d}^{-1} \). f) Upstream propagation, \( l_{\text{avg}} = 0.2 \text{ d}^{-1} \). Simulations were conducted on the OCN showed in Fig. 4.1c. The effect of elevation on temperature was included. All non-specified parameters are as in Table 4.1.

propagation front which reaches a steady state after around 30 years.

Climate change tends to accentuate invasion speed in both flow directions (see dashed lines in Figs. 4.3c, 4.3f): under a uniform 4 °C water temperature increase, steady distributions of prevalence are reached after about 5 years for downstream propagation and 20 years for upstream propagation. As pointed out in the previous Experiment, temperature rise results in a slight decrease of steady state prevalences with respect to the default case.

**Experiment 3.** Results for this experiment are reported in Fig. 4.4. When bryozoans are only present upstream (Fig. 4.4a), the distribution of prevalence at the corresponding stretches is governed by the mechanisms already described in Experiment 1 and generally increases downstream. Within the previously explored range of fish mobility rates, a maximum is
Chapter 4. PKD spread along river networks

reached in correspondence to the frontier between the two regions, followed by a decay towards the outlet. If fish mobility is absent (Fig. 4.4e), variations in prevalence are enhanced (with almost null prevalence at the headwaters of the upstream region, as in Experiment 1); this irregular trend (see green line in Fig. 4.4a) reflects the variability of the morphological features of the single stretches (in particular, the characteristics of the incoming tributaries). The inclusion of fish mobility (Fig. 4.4d) tends to flatten such variability. If $l_{avg}$ is set to extremely high values (violet line in Fig. 4.4a), prevalence becomes constant within the upstream region.

Upstream invasion from a downstream region suitable for bryozoans expectedly occurs only if fish mobility is invoked (Fig. 4.4b). In the stretches closer to the frontier with the bryozoan-free area, prevalence tends to decrease with respect to the case of uniform bryozoan density (e.g. compare Fig. 4.4f with Fig. 4.2a), owing to the lack of spore input from upstream.

4.3 Discussion

Fish movement and hydrological transport within a river network can produce a heterogeneous distribution of PKD prevalence and fish loss even in the absence of spatial gradients of fish and bryozoan densities, or of transmission rates. The typical lifetime of PKD spores (around 1 day) allows them to travel along with the river flow, and possibly infect fish, tens of kilometers downstream of the point where they are released. Stretches further downstream thus collect PKD spores from the whole (or a large portion of the) upstream area and are therefore more likely to exhibit higher PKD prevalence and fish loss. Conversely, headwaters and low-order streams are subject only to the spore load that is locally released and thus tend to be relatively less affected by PKD. Therefore, hydrological transport of spores tends to produce spatial patterns of PKD prevalence correlated to the upstream drainage area. Indeed, extensive surveys conducted in Swiss rivers [Wahli et al., 2002; Hari et al., 2006; Wahli et al., 2007; Zimmerli et al., 2007] have shown that PKD is mostly found in high order streams; Hari et al. [2006] found an altitude threshold of 800 m a.s.l. above which PKD was absent. Wahli et al. [2002] argued that within-stream prevalence gradients might be explained by different temperature regimes. Here, the obtained results show that pure network effects are sufficient to explain such a trend, although the role of temperature gradients can not be discarded. The dominant patterns of PKD prevalence can be partially affected by fish mobility. Indeed, a headwater connected directly to a high-order stream is subject to immigration of likely-infected fish that foster a local increase in prevalence. Overall, fish mobility promotes the mixing between low- and high-order streams resulting in a net increase of overall prevalence at the network scale.

The elevation of river stretches is intimately related to the aggregation structure of the underlying river network. Indeed, the slope-area relationships [Eq. (4.1)] dictate that a network configuration is uniquely associated to a fluvial topography, i.e. the relative elevation distribution of river sites. Thus network structure controls, as a byproduct, the distribution of elevations, a proxy of mean water temperature and in turn of PKD prevalence [Tops et al.,
Indeed, water temperature generally decreases with elevation and thus roughly with the distance from the outlet. Temperature gradients thus tend to produce distributions of PKD prevalence akin to those discussed above as driven by hydrological transport and fish mobility. However, results reported in Fig. 4.2 show that, at least for this example, the effect of temperature gradients is almost negligible compared to that produced by spatial coupling mechanisms. The river network analyzed herein spans an elevation relief of about 1000 m, which translates in about 6.5 °C difference in mean water temperature. Larger networks endowed with more pronounced elevation relief can possibly lead to more important effects of temperature gradients.

The invasion speed of PKD in the downstream direction is mostly controlled by hydrological transport of spores, whereas fish mobility has only a marginal effect. For transmission parameters leading to high PKD prevalence (above 90%, a value sometimes observed in affected river systems [Wahli et al., 2002, 2007]), the disease can invade from tens to hundreds of kilometers of river within a single proliferation season, provided that all sites are equally suitable for both fish and bryozoans. Under the same conditions, upstream invasion of PKD from a region of the network close to the outlet can occur only via fish swimming against the flow direction. The corresponding speed is much slower than the downstream one. With realistic values of fish mobility (e.g. \( l_{avg} = 0.02 \text{ d}^{-1} \)) and the reference value \( \pi^*_F = 0.2 \text{ m}^3 \text{d}^{-2} \), PKD can travel upstream about 4 km per year.

It is important to note that climate change would entail severe repercussions on the impact of PKD. With the setting described in Experiment 1, a 4 °C increment in water temperature would increase the average network fish loss by about 25%, with the most devastating effects occurring for high-order streams. For the reference parameter set, climate change provokes a minor decay in prevalence, owing to the augmented mortality of acutely infected fish; conversely, when \( \pi^*_F \) is set to lower values, a positive shift in temperature implies an increase in network prevalence (by setting \( \pi^*_F = 0.1 \text{ m}^3 \text{d}^{-2} \) and \( l_{avg} = 0 \text{ d}^{-1} \), average prevalence is 59.6% but rises up to 63.0% under climate change - result not shown). A similar feature was already observed with regards to the local PKD model (see §3.2.3). Notably, as shown in Experiment 2, higher water temperatures are responsible for a sensible increase in the speed of invasion fronts both downstream and upstream.

Modelling fish mobility as a diffusion process implicitly assumes that fish engage in trips with a thin-tailed distribution of distance [Chandrasekhar, 1943]. As discussed above, this is a useful working hypothesis. However, heavy-tailed distribution of travel distances, as suggested by some empirical studies (e.g. Skalski and Gilliam [2000]), could lead to an anomalous diffusion and to an enhanced speed of disease propagation [Baeumer et al., 2007]. Moreover, other factors not included in this analysis can boost the propagation of PKD in both directions. Disease transmission in bryozoans could also occur as a result of fragmentation of infected colonies [Fontes et al., 2017b], but this was here assumed to be a mainly local process. In addition, passive hydrodynamic dispersal of statoblasts was neglected due to their relatively large size (0.5-1 mm), compared to that of parasite spores; however, transport of statoblasts in
Chapter 4. PKD spread along river networks

ballast water or by waterfowl has been recognized as a cause of spread of bryozoan species worldwide, and hence of PKD infection within and between water bodies [Kipp et al., 2010; Okamura et al., 2011; Abd-Elfattah et al., 2014a]. Ecological models accounting for the role of ship ballast tanks in promoting species invasions in fluvial landscapes have already been developed (e.g. for zebra mussels [Mari et al., 2011]).

The results of the analysis presented in Experiment 3 highlight how the detection of PKD infection in a river site does not necessarily imply that PKD is locally transmitted, a process that requires the simultaneous presence of both fish and bryozoans. Indeed, the actual source of infection can be either downstream or upstream. In the former case, local PKD presence is due to infected fish swimming upstream. In the latter, local PKD presence is due to fish movement and, chiefly, to hydrological transport of infected spores, which can sustain disease prevalence for a long river segment in the downstream direction. The fact that hotspots of active PKD transmission can bolster PKD prevalence in a much wider range extending in both the downstream and, to a lesser extent, the upstream directions poses serious challenges to their identification and possible control. Indeed, extensive, network-scale monitoring campaigns of both fish and bryozoans, along with their infection status, represent a prohibitive effort in most cases. With this regard, the use of habitat models of fish distribution (e.g. González-Ferreras et al. [2016]) is beneficial. Moreover, innovative techniques in environmental monitoring like environmental DNA (eDNA) offer new opportunities. The issue of inferring the spatial distribution of bryozoans from eDNA data is tackled in detail in Chapters 5 and 6.

In this analysis, spatial coupling mechanisms, i.e. hydrological transport of spores and fish mobility, have been considered to take place only during the bryozoan proliferation season during which PKD can actually be transmitted (typically from April to November). While this hypothesis is sound for parasite spores, whose survival time-scale is much shorter than the duration of the winter season, it might be questionable as regards fish. Indeed, fish movement in the cold season, even if it cannot spread the disease, can enhance the redistribution of PKD carriers. It is known, for instance, that salmonids may move by several kilometers upstream in late summer or autumn for spawning, and an analogous distance is covered downstream in early spring. Mobility might depend on various factors, such as fish age, gender, size, water temperature, food availability, population density, growth rate, moon phase [Jonsson and Jonsson, 1993; Young, 1994; Slavík et al., 2012]. However, brown trout are usually subject to natal homing [Frank et al., 2012] and so they tend to return to the same place. Moreover, PKD is not vertically transmitted to newborn fish. Hence, one can argue that neglecting winter fish mobility does not frustrate this framework’s ability to capture the effect of this crucial process on the patterns of PKD prevalence.

The insights gained through this set of synthetic experiments on the spatial effects of PKD spread are fully exploited in the application to the case study presented in the next chapter.
### 4.4 List of symbols

The following table lists all symbols used in this Chapter. This table does not include state variables and parameters of the epidemiological model (presented in Tables 2.1 and 4.1, respectively). Generally, subscript $k$ identifies stretches, while $i$ represents pixels; subscript $O$ denotes the outlet stretch, while $o$ the outlet pixel.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>Upstream drainage area at pixel $i$ (in number of pixels)</td>
<td>-</td>
</tr>
<tr>
<td>$A_T$</td>
<td>Threshold area for stream definition</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$A_{tot}$</td>
<td>Total catchment area</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$\hat{A}_k$ [ $\hat{A}_i$ ]</td>
<td>Normalized upstream drainage area at stretch $k$ [pixel $i$]</td>
<td>-</td>
</tr>
<tr>
<td>$B_k$</td>
<td>River width of stretch $k$</td>
<td>$L$</td>
</tr>
<tr>
<td>$B_O$</td>
<td>Outlet stretch width</td>
<td>$L$</td>
</tr>
<tr>
<td>$D_k$</td>
<td>Water depth of stretch $k$</td>
<td>$L$</td>
</tr>
<tr>
<td>$D_O$</td>
<td>Outlet stretch depth</td>
<td>$L$</td>
</tr>
<tr>
<td>$DO$</td>
<td>Distance from outlet</td>
<td>$L$</td>
</tr>
<tr>
<td>$H_i$</td>
<td>Energy dissipation along pixel $i$</td>
<td>$ML^2T^{-2}$</td>
</tr>
<tr>
<td>$H$</td>
<td>Total energy dissipation along the river network</td>
<td>$ML^2T^{-2}$</td>
</tr>
<tr>
<td>$i_k$</td>
<td>Last downstream pixel of stretch $k$</td>
<td>-</td>
</tr>
<tr>
<td>$L_i$</td>
<td>Length of pixel $i$</td>
<td>$L$</td>
</tr>
<tr>
<td>$L_p$</td>
<td>Side of a pixel</td>
<td>$L$</td>
</tr>
<tr>
<td>$L_{s,k}$</td>
<td>Length of stretch $k$</td>
<td>$L$</td>
</tr>
<tr>
<td>$l_{avg}$</td>
<td>Mean fish mobility rate</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$N_p$</td>
<td>Number of pixels in lattices used for generation of OCNs</td>
<td>-</td>
</tr>
<tr>
<td>$N_s$</td>
<td>Total number of river stretches</td>
<td>-</td>
</tr>
<tr>
<td>$Q_i$</td>
<td>Water discharge at a pixel $i$</td>
<td>$L^3T^{-1}$</td>
</tr>
<tr>
<td>$Q_O$</td>
<td>Average discharge at the outlet stretch</td>
<td>$L^3T^{-1}$</td>
</tr>
<tr>
<td>$S$</td>
<td>Generic sequence of $A_i$ for all pixels of a lattice</td>
<td>-</td>
</tr>
<tr>
<td>$S'$</td>
<td>Sequence of $A_i$ for all pixels of a lattice minimizing $H_s$</td>
<td>-</td>
</tr>
<tr>
<td>$s_o$</td>
<td>Channel slope at the outlet pixel</td>
<td>-</td>
</tr>
<tr>
<td>$U_O$</td>
<td>Average water velocity at the outlet stretch</td>
<td>$LT^{-1}$</td>
</tr>
<tr>
<td>$T_f(\tau)$</td>
<td>Weighted average of $T_k(\tau)$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_k(\tau)$</td>
<td>Temperature at stretch $k$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_O(\tau)$</td>
<td>Temperature at the outlet stretch</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_k$</td>
<td>Sequence of pixels constituting stretch $k$</td>
<td>-</td>
</tr>
<tr>
<td>$V_k$</td>
<td>Water volume in stretch $k$</td>
<td>$L^3$</td>
</tr>
<tr>
<td>$z_o$</td>
<td>Elevation at the outlet pixel</td>
<td>$L$</td>
</tr>
<tr>
<td>$\Gamma_w$</td>
<td>Environmental lapse rate</td>
<td>$\Theta L^{-1}$</td>
</tr>
</tbody>
</table>
### Chapter 4. PKD spread along river networks

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta Z_k$</td>
<td>Drop in elevation along stretch $k$</td>
<td>$L$</td>
</tr>
<tr>
<td>$\Delta z_i$</td>
<td>Drop in elevation along pixel $i$</td>
<td>$L$</td>
</tr>
<tr>
<td>$\delta_p$</td>
<td>Side of the lattice</td>
<td>$-$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Time</td>
<td>$T$</td>
</tr>
</tbody>
</table>
Predicting how temperature, climate change and emerging infectious diseases interact to drive local extinction risk for natural populations requires complex integrated approaches involving field data, laboratory studies and metacommunity modelling. Here, the theoretical tools developed in the previous Chapters are applied to an integrated study of PKD in a prealpine Swiss river (the Wigger). During a three-year period, data on fish abundance, disease prevalence, concentration of primary hosts’ DNA in environmental samples (eDNA), hydrological variables and water temperatures, gathered at various locations within the catchment, were integrated into the developed metacommunity model, which includes ecological and epidemiological dynamics of fish and bryozoans, connectivity effects and hydrothermal drivers. Infection dynamics were captured well by the epidemiological model, especially with regards to the spatial prevalence patterns. PKD prevalence in the sampled sites for both young-of-the-year (YOY) and adult brown trout attained 100% at the end of summer, while seasonal population decay was higher in YOY than in adults. Moreover, a new method for the prediction of local species’ density based on decay distance of eDNA signal and accounting for variation in environmental drivers (such as morphology and geology) is here introduced. This study shows how spatial and environmental characteristics of river networks can be used to study epidemiology and disease dynamics of waterborne diseases. The provided framework, applicable to other aquatic pathogens, may function as a baseline for environmental management decisions aimed at preserving declining, possibly iconic salmonid species.¹

Chapter 5. Case study: An integrated analysis of PKD spread in a Swiss prealpine river

5.1 Methods

5.1.1 Study area and hydro-geomorphology

The river Wigger, located in the Swiss plateau, is a tributary of the river Aare and has a main channel length of 48.11 km. It drains a watershed of 382.4 km² which has an elevation range comprised between 396 and 1409 m a.s.l., at the Mount Napf (Fig. 5.1a, b). It is highly anthropized, with scarce restoration areas, regular channel widths and slopes engendered by a system of check dams along the main courses which does not hinder fish movement. The Wigger is subject to endemic PKD of several years (see Burkhardt-Holm et al. [2002]; Wahli et al. [2002]; a dataset of PKD occurrence across Switzerland is freely available at the website of the Swiss Federal Geoportal [Federal Office for the Environment FOEN, 2017]).

![Figure 5.1 – Overview of the study area and data. a) Digital terrain model of the Wigger and extracted river network with location of the sampling sites. b) Position of the Wigger catchment in Switzerland. c) Mean measured *F. sultana* eDNA concentration. Ungauged stretches are depicted in blue. d) Geological characterization of the catchment.](image)
5.1. Methods

The river network was extracted from a 25-m resolution digital terrain model using the Taudem routine in a GIS software [Tarboton, 1997]. Flow directions were determined by following steepest descent paths (D8 algorithm) [Tarboton, 1997]. Pixels belonging to the channeled portion of the landscape were those whose drained area was greater than or equal to 0.5 km² whereas details on slope- of curvature-dependent area thresholds prove immaterial at the scale of interest [Rodriguez-Iturbe and Rinaldo, 2001]. The extracted fluvial network was compared with the vectorial map provided by FOEN [Federal Office of Topography Swisstopo, 2016], and the overall match proved satisfactory. In order to subdivide the catchment into subunits where reasonably constant local morphological conditions apply, i.e. nodes of the metacommunity scheme, stream reaches longer than 5 km were split into equally long stretches. A subcatchment was defined by the direct contributing area drained by a single stream reach. Overall, the river network was divided into 166 subcatchments hierarchically arranged according to the network connectivity.

The geological characterization of the catchment (Fig. 5.1d) was obtained by a vectorized geological map of Switzerland provided by the Swiss Federal Office of Topography (Swisstopo) [Federal Office of Topography Swisstopo, 2005]. The map was processed on a GIS software and the initial 16 geological classes were grouped into five main classes, as displayed in Fig. 5.1d. Note that the ‘superficial water’ class only refers to ponds or small lakes, while all pixels covering the river network proper were conventionally classified as ‘alluvial rocks’. Daily mean discharges measured by the Swiss Federal Office for the Environment (FOEN) in Zofingen (corresponding to site #3 of Fig. 5.1a) were used to compute time series of discharge for all stretches. Owing to the relatively small extension of the catchment, it is reasonable to assume that the daily discharge at a station is proportional to the total contributing area therein [Rodriguez-Iturbe and Rinaldo, 2001]. This hypothesis was tested for the two discharge series available, namely measured by FOEN in Zofingen (drained area \( A = 366 \text{ km}^2 \)) and Nebikon (site #7 - \( A = 105 \text{ km}^2 \)). Comparing observed and modelled daily discharges in Zofingen for the period 2010-2015 (where the latter is obtained by multiplying the daily Nebikon discharge by \( 366/105 \approx 3.5 \)), a satisfactory fitting was achieved (Nash-Sutcliffe coefficient \( NS = 0.89 \)).

Average river width for each stretch was estimated from Swisstopo aerial images and checked against at-a-station scaling relationships linking landscape-forming discharges to total contributing area \( A \), which yield for consistency the scaling of widths and depths of the cross section to \( A \) [Leopold and Maddock, 1953; Leopold et al., 1964]. Shapes of cross-sections are approximated as rectangular and endowed with large width-to-depth ratios. Given the assumptions and the scales of interest, it is customary to hypothesize the maintenance of uniform flow conditions for individual river stretches. As such, a Manning's roughness coefficient \( n = 0.033 \text{ m}^{-1/3} \text{s} \) was assumed for the whole river network; according to the U.S. Federal Highway Authority [Arcement Jr and Schneider, 1989], this value corresponds to a fine gravel bedrock, with low vegetation and negligible cross-sectional variation, surface irregularities obstructions and sinuosity. Such features were qualitatively observed during in-situ campaigns. Through a systematic use of Manning's equation of uniform flow resistance, time series of water depths for all stretches were computed and used in the metacommunity model.
Chapter 5. Case study: An integrated analysis of PKD spread in a Swiss prealpine river

5.1.2 Field data collection

eDNA collection. In the period May 2014 - May 2015, stream water samples were collected in 15 different locations along the river network (Fig. 5.1a). For each site, 21 500-mL samples were taken at approximately bi-weekly (or monthly during December, January and February) intervals (except site #5, which was abandoned after 12 samples). For subsequent data analysis, the target DNA quantity estimated in each water sample was averaged over the 21 temporally distributed samples; the resulting mean concentration \( C_m \) (Fig. 5.1c) were then used as input for the determination of bryozoan habitat suitability. The 500-mL samples were filtered onto 5-cm diameter GF/F filters and filter papers frozen at -80 °C until extraction of eDNA. Environmental DNA was extracted from filtered water samples using methods described in Appendix B. The concentration of \( F. sultana \) qPCR target in each 500-mL sample was measured using an assay incorporating an internal positive control for PCR inhibition assessment. None of the negative control samples exceeded the limit of quantification for the \( F. sultana \) target and no inhibition exceeding three cycles was observed in any sample. Further details on eDNA collection and qPCR assay are reported in Appendix B.

Fish sampling and prevalence assessment. All young-of-the-year (YOY) trout sampled in this study were originated from natural reproduction, as there was no fish stocking in the Wigger in the period 2014-2016 (Aargau and Lucerne cantonal authorities, pers. comm.). Trout abundance estimation, collection of fish and kidney sampling were performed according to the Swiss regulations and the field setup was accepted by the relevant authorities (Veterinärdienst Kanton Luzern, Tierversuchskommission des Kantons Bern, Kantonaler Veterinärdienst des Kanton Aargau) under the number LU05/14+. Fish were caught by electrofishing in early and late summer on sites #4, #8, #16 (Fig. 5.1a), over a distance of 100 m. During each sampling trip, 25 YOY brown trout were collected outside the stretch used for density assessment. When available, 5 adult brown trout adult (>1 year old) were sampled during the late summer field campaign.

Electrofishing was performed by FORNAT AG (Forschung für Naturschutz und Naturnutzung, Zurich) with an EFKO FEG8000 machine (Power: 8 kW, number of machine 130708). Two runs of electrofishing per date and stretch were performed. For density assessment purpose, caught fish were kept in tanks supplied with oxygen and fish were set back into the river after the second fishing run. Prior to release fish length was measured. As for PKD prevalence assessment, the trout were euthanized using 3-aminobenzoic acid ethyl ester (MS 222®, Argent Chemical Laboratories) and length was recorded. The kidney was removed and preserved in RNAlater (Sigma Aldrich, Switzerland) for qPCR based assessment of \( T. bryosalmonae \)

\(^2\)Collection of water samples and qPCR analyses were carried out by Hanna Hartikainen, Inês Fontes and collaborators.

\(^3\)The author of the Thesis contributed to the fish sampling campaign and abundance assessment, jointly with Thomas Wahl, Nicole Strepparava, Inês Fontes and collaborators.

\(^4\)Fish euthanasia, kidney extraction and subsequent qPCR analyses were performed by Nicole Strepparava and collaborators.
5.1. Methods

infection presence. Fish kidneys were weighted and homogenized in a 1.5-mL tube with a 2-mm diameter steel bead (QIAGEN, Switzerland) with a tissue lyser (QIAGEN, Switzerland) at a shaking frequency of 30 shakes per second for 3 min. DNA was extracted using the Blood & Tissue DNA extraction kit (QIAGEN, Switzerland). Prevalence was assessed by qPCR using the \textit{T. bryosalmonae} specific TaqMan method according to Bettge et al. [2009b]. All reactions were carried out in duplicates. As positive control, a linearized plasmid containing the amplified fragment [Bettge et al., 2009a] was used. Non-target control (water) within the qPCR never showed any amplification while internal control was always amplified showing no qPCR inhibition.

Table 5.1 reports prevalence data; Fig. 5.2 displays results from the brown trout density assessment.

<table>
<thead>
<tr>
<th>Site</th>
<th>Age-class</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aug 14</td>
<td>Sep 18</td>
<td>Jun 10</td>
</tr>
<tr>
<td>#4</td>
<td>YOY</td>
<td>34/34</td>
<td>20/20</td>
<td>4/5</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>-</td>
<td>6/6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>-</td>
<td>4/4</td>
<td>-</td>
</tr>
<tr>
<td>#16</td>
<td>YOY</td>
<td>22/25</td>
<td>26/26</td>
<td>5/6</td>
</tr>
<tr>
<td></td>
<td>Adults</td>
<td>-</td>
<td>5/5</td>
<td>-</td>
</tr>
</tbody>
</table>

\textbf{Table 5.1} – Prevalence data. The first number counts the amount of PKD-positive individuals, the second refers to the total number of individuals tested.

\* This sampling was conducted in a different location (see Fig. 5.8e).

Figure 5.2 – Brown trout population sampling data. Blue indicates early summer sampling, red stands for late summer sampling, maroon colour results from overlay of blue and red bars. \textit{N}: total fish abundance; \textit{Y}: YOY abundance. Dashed lines highlight the threshold lengths used to distinguish between YOY and adults: 12 cm in early summer and 14 cm in late summer. Population sampling in site #4 (not shown) was only performed in late 2016 summer.
5.1.3 Water temperature monitoring and extrapolation

Water temperature were measured in 11 sites via HOBO Tidbit® v2 data loggers (see Fig. 5.1a). The sampling started on June 28, 2014, except for site #17, which was installed on September 27, 2015. Data are available until October 4, 2016. Two loggers, recording data at 15-minutes interval, have been deployed per each site. Water temperature is known to affect bryozoan growth rate [Tops et al., 2009] and may also impact spore shedding.

Daily mean temperatures were extrapolated to all ungauged subcatchments by means of a linear regression against five morphological covariates: local elevation \( (\text{LocElv}) \), local slope \( (\text{LocSlp}) \), upstream contributing area \( (\text{UpCAr}) \), upstream elevation \( (\text{UpElv}) \), upstream slope \( (\text{UpSlp}) \). All covariates were normalized in the range \([-1; 1]\), where bounds correspond to the lowest/highest values of the covariates among all subcatchments. The daily mean water temperature \( T_i(\tau) \) at day \( \tau \) is then expressed as:

\[
T_i(\tau) = T_b(\tau) + \alpha_{LE}(\tau)\text{LocElv}(i) + \alpha_{LS}(\tau)\text{LocSlp}(i) + \alpha_{UC}(\tau)\text{UpCAr}(i) + \alpha_{UE}(\tau)\text{UpElv}(i) + \alpha_{US}(\tau)\text{UpSlp}(i)
\]  

(5.1)

where the baseline temperature \( T_b(\tau) \) (i.e. the water temperature at day \( \tau \) of a reach whose covariates are at a null level) and the vector of coefficients \( \alpha(\tau) = [\alpha_{LE}(\tau); \alpha_{LS}(\tau); \alpha_{UC}(\tau); \alpha_{UE}(\tau); \alpha_{US}(\tau)] \) were calibrated separately for each day. Modelled temperatures were extended beyond the observation period via sinusoidal functions \( T_s(i) \) fitted for each subcatchment based on linear regression outputs:

\[
T_s(i)(\tau) = T_{m,i} - T_{a,i}\cos\left(2\pi\frac{\tau - \tau_{\text{min},i}}{365}\right)
\]

where \( T_{m,i} \) is the annual mean temperature in stretch \( i \); \( T_{a,i} \) the mid-amplitude of the sinusoidal signal; \( \tau_{\text{min},i} \) the day of minimum temperature (expressed in ordinal days). These three coefficients were estimated via a least squares technique. A comparison between observed and modelled water temperatures is presented in Fig. 5.3, while Fig. 5.4 shows the evolution of the calibrated parameters over time.

As expected, the baseline temperature follows a sinusoidal trend with yearly period (Fig. 5.4a). During warm periods, local and upstream elevations have a positive effect on water temperature (Fig. 5.4b, e), while local and upstream slope are negatively correlated with high temperatures (Fig. 5.4c, f). During winter, such trends are reversed. The effect of contributing area is generally positive (Fig. 5.4d) but its magnitude is modest.

Overall, the linear regression provides a straightforward and reliable tool for the spatial extrapolation of water temperature data at a catchment scale. The model performs well in reproducing the observed data (root mean square errors for all sites are generally below 0.5 °C) and the temperature patterns obtained for intermediate and downstream reaches appear reasonable. A shortcoming has been observed in the prediction of excessively low summer temperatures in short, upstream reaches (see Fig. 5.5), due to the fact that no sampling sites
5.1. Methods

are located close to the headwaters. However, this drawback has little impact in the outcomes of the PKD epidemiological model, since the reaches affected are those whose fish density is lower. On the other hand, one must acknowledge that there might be some repercussions in the determination of bryozoan suitability maps, as the mean water temperature during the warm season (LocMwt) is used as a covariate for the prediction of local bryozoan density. The prediction of spatio-temporal water temperature patterns could be improved by means of a deterministic model (see Caissie [2006]). A spatially-explicit model of stream temperatures is developed in Chapter 7 and tested against the same dataset here introduced.

5.1.4 Bryozoan habitat suitability

The whole river network is divided into a suitable number of stretches \( N_s \), some of which include the sampling sites. Let \( A_i^L \) be the direct drainage area of stretch \( i \), namely the sub-catchment pertaining to stretch \( i \) only, and \( A_i \) the upstream contributing area to stretch \( i \), that is the catchment surface upstream of \( i \). \( A_i \) obviously includes \( A_i^L \). Let \( c_i \) be the average eDNA concentration that would be attained in stretch \( i \) in absence of water flow. It is assumed that \( c_i \) is proportional to the density of bryozoan biomass pertaining to site \( i \), and that \( c_i \) does not change throughout the season. Moreover, it is hypothesized that the eDNA produced from bryozoans in stretch \( i \) does not decay until it reaches the exit cross-section of stretch \( i \). Instead, let \( C_i \) be the eDNA concentration measured in stretch \( i \), i.e. when waterflow is considered.

If one assumes that there are no bryozoans upstream of stretch \( i \), which is at hydraulic steady state, it follows that the flow of incoming eDNA (from the sole source, namely stretch \( i \)) equals the flow of outgoing eDNA through the exit cross-section of stretch \( i \). Since the contribution to discharge from subcatchment \( i \) is proportional to \( A_i^L \), while the outgoing flow is proportional to \( A_i \) (as previously assumed), one can state that \( C_i A_i = c_i A_i^L \). Notice that, because \( A_i^L \) can be much smaller than \( A_i \) (especially for downstream stretches), the measured concentration at site \( i \) can be small even though the bryozoan biomass in the subcatchment may be quite large. The deterioration of the eDNA signal from site \( i \) measured at a given downstream distance is modeled by a first-order kinetics with characteristic decay length \( \lambda_B \) [Jerde et al., 2016]; conversely, no longitudinal dispersion of the eDNA signal is assumed. The eDNA concentration measured at a given location \( j \) reads therefore:

\[
C_j = \frac{1}{A_j} \sum_{i=1}^{N_i} p_{ij} A_i^L \exp \left( -\frac{L_{ij}}{\lambda_B} \right) c_i
\]

(5.2)

where \( N_i \) is the number of stretches into which the river network is subdivided; \( p_{ij} \) is the generic entry of the connectivity matrix \( P \) (i.e. \( p_{ij} = 1 \) if there exists a downstream path connecting \( i \) to \( j \), zero otherwise), and \( L_{ij} \) the downstream distance between \( i \) and \( j \). Note that \( p_{jj} = 1 \) and \( L_{jj} = 0 \). The sum over all reaches upstream of \( j \) (weighted by their relative contribution to the total contributing area, a proxy of river discharge) accounts for the effect

---

5 The model presented in this Section is generalized and further discussed in Chapter 6.
Chapter 5. Case study: An integrated analysis of PKD spread in a Swiss prealpine river

![ Temperature vs Time Graphs ]

Figure 5.3 – Modelled and observed mean daily water temperatures. For each station, the root mean square error (in °C) achieved by the linear regression model is displayed. Sinusoidal fitted temperatures are also presented.

The local concentration $c_i$ is modelled as $c_i = c_0 \exp(\beta \cdot X(i))$, where $X(i)$ is a vector of covariates, while the scalar $c_0$ and the vector $\beta$, together with $\lambda_B$, are parameters that need calibration contrasting field data. The choice of an exponential link function is justified by the fact that $c_i$ must be non-negative. The chosen explanatory variables, listed in Table 5.2, refer to geomorphological, hydrothermal and geological features of the catchment. Covariates were
5.1. Methods

Figure 5.4 – Time series of the calibrated values of the linear regression parameters.

Figure 5.5 – Maps of mean water temperatures in winter (mean of winter 2015 and 2016) and summer (mean of summer 2014 and 2015) obtained by linear regression.
Chapter 5. Case study: An integrated analysis of PKD spread in a Swiss prealpine river

normalized in the range [-1; 1], where boundaries correspond to the lowest/highest values of the covariates among all subcatchments.

Starting from the initial set of eleven covariates of Table 5.2, possible multicollinearity effects were checked by computing the variance inflation factors (VIFs) and discarding the predictor with the highest VIF [Hair et al., 2013]. This procedure was repeated until all remaining covariates had VIFs all lower than 10. The discarded covariates were, in the order: GeoMls; UpSlp; GeoPea; LocElv. VIF values for the set of residual explanatory variables (LocMwt; LocSlp; UpElv; UpCar; GeoAll; GeoMrn; GeoWat) are: 5.8; 4.9; 2.4; 5.4; 7.8; 2.3; 1.9. None of the correlation coefficients $R$ between any of these seven variables were above the rule-of-thumb threshold $|R| = 0.8$ indicating strong correlation (Fig. 5.6). The performance of all models (total number: 127) generated by any subset of the residual explanatory variables was tested. Leave-One-Out-Cross-Validation (LOOCV) [Hastie et al., 2001] was used to evaluate the performance of each model. For a given subset of covariates, a site $k$ at a time was removed from the dataset; a calibration was then performed by maximizing the Nash-Sutcliffe index $NS_k$:

$$NS_k = 1 - \frac{\sum_{j=1, j \neq k}^{15} (C^m_j - C_j)^2}{\sum_{j=1, j \neq k}^{15} (C^m_j - C^m_k)^2},$$

where $C^m_j$ is the mean eDNA concentration measured at site $j$ and $C^m_k = \frac{1}{14} \sum_{j \neq k} C^m_j$. Once calibrated, the overall performance of the model was calculated as

$$s = 1 - \frac{\sum_{k=1}^{15} (C^m_k - C_k)^2}{\sum_{k=1}^{15} \left( C^m_k - \frac{1}{15} \sum_{j=1}^{15} C^m_j \right)^2}. $$

All models achieving $s > 0$ (i.e. 16 models) were deemed as behavioural (or acceptable, *sensu* Beven and Freer [2001]). These models were subsequently re-calibrated against the whole set of sites by maximizing the Nash-Sutcliffe index. All calibrations were performed via a particle swarm optimization algorithm [Kennedy and Eberhart, 1995].

5.1.5 Details on model simulations and calibration

Bryozoan biomass is expressed in dimensionless units, with local carrying capacities $\rho_i^{-1}$ assumed proportional to $c_i$ and equal to unity for the stretch characterized by the highest $c_i$ value (where $c_i$ values are taken as in Fig. 5.7c). Note that the sole effect of this re-scaling operation on the state variables is the modification of the unit of the parameter $\pi_B^*$ (from [m$^3$ g$^{-1}$ d$^{-2}$] to [m$^3$ d$^{-2}$], implying that the rate of production of equivalent spores is now referred to one dimensionless unit of bryozoan biomass rather than to 1 g). However, as already pointed out with regards to the introduction of the equivalent spore concept (see §2.1.1), parameters $\pi_B^*$ and $\pi_F^*$ cannot be measured and must be obtained via calibration. Therefore, such a modification proves immaterial for the purposes of this work, as its focus is not on
5.1. Methods

As for both YOY and adults, fish density at equilibrium is proportional to the mean stretch depth [Heggenes et al., 1999], which corresponds to an appropriate parametrization of the effect of density dependence $\xi$ (as specified in Table 5.3). As regards fish spawning migration patterns [Eq. (2.11)], the normalization parameters $A_F = 100 \text{ km}^2$ and $\lambda_F = 2000 \text{ m}$ were chosen such that, with the reference value $l_{avg} = 0.02 \text{ d}^{-1}$, the distribution of YOY at the end of the season is sufficiently close to the equilibrium distribution. Note that the value of $\lambda_F$ is congruent with observed values of typical migration distances covered by brown trout [Rustadbakken et al., 2004; Gosset et al., 2006].

In order to allow the system to lose memory of the initial conditions which are hardly determinable, the epidemiological model was run for 20 years (1996-2016) with real discharge data from the FOEN Zofingen station. Water temperatures before July 2014 were derived from sinusoidal interpolations of the regressed temperature time series at a local level (as shown in Fig. 5.3). At the beginning of the simulations, it is assumed that half of the bryozoan biomass at each site is covertly infected, while all fish are uninfected. Initial sizes of local YOY and adult fish populations were set to their dynamical disease-free trajectory value (as in Chapter 4). After around 10 years, spatial patterns of prevalence and population size reach an annual cycle with seasonal variations only due to hydrothermal variability.

The model is calibrated against both prevalence and seasonal population decline data. Fish

\[ \begin{array}{cccccccc}
\text{LocMwt} & \text{LocSlp} & \text{UpElv} & \text{UpCar} & \text{GeoAll} & \text{GeoMrn} & \text{GeoWat} \\
-1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \\
-0.70 & -0.62 & 0.46 & 0.47 & 0.52 & 0.54 & \\
0.25 & -0.51 & -0.77 & -0.52 & -0.25 & \\
-0.70 & -0.62 & 0.46 & 0.47 & 0.52 & 0.54 & \\
0.25 & -0.51 & -0.77 & -0.52 & -0.25 & \\
\end{array} \]

\[ \begin{array}{cccccccc}
\text{LocMwt} & \text{LocSlp} & \text{UpElv} & \text{UpCar} & \text{GeoAll} & \text{GeoMrn} & \text{GeoWat} \\
-1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \\
-0.70 & -0.62 & 0.46 & 0.47 & 0.52 & 0.54 & \\
0.25 & -0.51 & -0.77 & -0.52 & -0.25 & \\
-0.70 & -0.62 & 0.46 & 0.47 & 0.52 & 0.54 & \\
0.25 & -0.51 & -0.77 & -0.52 & -0.25 & \\
\end{array} \]

\[ \begin{array}{cccccccc}
\text{LocMwt} & \text{LocSlp} & \text{UpElv} & \text{UpCar} & \text{GeoAll} & \text{GeoMrn} & \text{GeoWat} \\
-1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \\
-0.70 & -0.62 & 0.46 & 0.47 & 0.52 & 0.54 & \\
0.25 & -0.51 & -0.77 & -0.52 & -0.25 & \\
-0.70 & -0.62 & 0.46 & 0.47 & 0.52 & 0.54 & \\
0.25 & -0.51 & -0.77 & -0.52 & -0.25 & \\
\end{array} \]

Figure 5.6 – Plot of the correlation matrix of the seven chosen predictors for the model of bryozoan suitability.

the spatial distribution of the absolute bryozoan biomass but rather on that of the relative biomass and on its implications for PKD spread and salmonid health status.
population decline was estimated only when two population samplings per year were conducted at the same site. A total of 33 data points is available (27 prevalence measurements and 6 fish decline); in the objective function, all data points have equal weight. Calibration was performed via a Metropolis-Hastings algorithm [Hastings, 1970]. As both prevalence and decline are constrained between 0 and 1, it is hypothesized that errors are distributed according to a normal distribution truncated between 0 and 1, with standard deviation equal to 0.1. Based on a previous sensitivity analysis (see §3.2.3), seven epidemiological parameters were chosen for calibration: $\pi^*_B$, $\pi^*_F$, $l_{avg}$, $a_i$, $h_i$, $\gamma$, $\epsilon$. As both $a_i$ and $h_i$ depend on temperature, the parabolic relationships of Fig. 3.1 were used. The calibrated parameters were then the values of $a_i$ and $h_i$ at 15 °C. Other parameters were set to their reference values (Table 5.3).

5.2 Results

5.2.1 Bryozoan habitat suitability

Local bryozoan density was positively correlated with the presence of moraines upstream ($GeoMrn$) (present with positive sign in 100% of the behavioural models - Fig. 5.7a). Environmental covariates such as mean water temperature ($LocMwt$) have a less clear effect on bryozoan density, as their correlation might be positive or negative depending on the particular model. Upstream mean elevation ($UpElv$) and local slope ($LocSlp$) are in most cases negatively associated with bryozoan density. In 43.8% of the behavioural models, the calibrated values of the decay length $\lambda_B$ lie between 10 and 12.5 km (Fig. 5.7b). Overall, the capability of the behavioural models to reproduce the mean $F. sultana$ eDNA concentrations measured in the 15 sampling sites ($C(m)$) seems satisfactory (Fig. 5.7d, e). The predicted spatial distribution of $F. sultana$ eDNA local concentrations (averaged over all behavioural models) is displayed in Fig. 5.7c. Confirming the sensitivity of the eDNA based detection method, multiple manual field searches for bryozoan colonies on all of the 15 eDNA sampling sites.
### Table 5.3 – Reference set of parameters for the spatially-explicit epidemiological model with age structure. Subscript $i$ identifies site-dependent parameters. Water temperature $T_i$ and time $\tau$ are in Celsius degrees and ordinal days, respectively. A list of references is reported in Table 3.1.
were able to locate populations only in the three sites (#9, #13, #15) corresponding to the highest eDNA concentrations and the most frequent detections (see Fig. 5.7d). Similarly, sites known to be bryozoan-negative through exhaustive field searches (e.g. #5) were consistently negative via eDNA.

### 5.2.2 Epidemiological model

Field surveys assessed PKD prevalence in fish at three sites (#4, #8, #16 - Fig. 5.1a) both in early (for YOY) and late summer (for both YOY and adults). In the downstream and intermediate sites (#4 and #8, respectively), the general pattern is that all fish of all age-classes in all sampling dates are infected (Fig. 5.8a), while in the upstream site #16 YOY prevalences in early summer were always lower than 100%. Notably, in the late 2015 summer sampling, no YOY were found at the usual location of the sampling site #16 (i.e., downstream of the junction of the three tributaries Änziwigger, Buechwigger and Seewag), hence the sampling was shifted upstream of the confluence with Seewag (see circle in Fig. 5.8e). None of the YOY sampled on that occasion tested positive for PKD.

The epidemiological metacommunity model is capable of reproducing the observed patterns of PKD prevalence (Fig. 5.8a) and, in particular, the late summer 100% prevalences observed for both age-classes at sites #4 and #8. The model forecasts prevalences close to 100% in large parts of the network in late summer for both adults (Fig. 5.8f) and, to a lesser extent, YOY.
5.2. Results

Figure 5.8 – Results from the epidemiological metacommunity model. Top row: results of calibration against a) prevalence data; b) seasonal fish decline, measured as the fractional decline of the estimated population size in late summer compared to early summer. In panel a, the left point of each year group corresponds to YOY early summer sampling. The observed 0-prevalence value was actually measured in a stretch upstream of site #16 (see panel e and main text). Bottom-left corner: time evolution of modelled prevalence in site #16 for c) YOY and d) adult fish. The intra-seasonal model is run for 200 days starting on April 1st. Note that YOY prevalence sample in October 2015 is missing. Bottom-right corner: maps of best fit modelled PKD prevalence evaluated at the end of summer 2016 for e) YOY; f) adults. Panel e features a zoom on site #16.

(Fig. 5.8e). Prevalence in adults tends to be high during the whole season (Fig. 5.8d), while the initially null prevalence level in YOY (Fig. 5.8c) is due to the absence of vertical transmission of PKD in fish. As expected, modelled PKD prevalence is lower in those parts of the network where there are no upstream stretches where predicted bryozoan abundance is high (Fig. 5.7c). This result agrees with the observed null prevalence upstream of the confluence with the Seewag: indeed, this tributary, unlike the Änziwigger and the Buechwigger, is characterized by high *F. sultana* density, according to the model of bryozoan suitability (see Fig. 5.7c).

Modelling spatial distributions of fish densities in the absence of PKD (Fig. 5.9; bottom row) expectedly follow the imposed patterns at equilibrium (see §2.2.1), according to which fish density is proportional to mean stretch depth that, in turn, scales with the contributing area. This is particularly evident for adults (Fig. 5.9g, h), where seasonal variations of the spatial density distributions appear negligible. Conversely, at the beginning of the warm season YOY are mainly concentrated in small, peripheral reaches, whose spawning suitability is higher, but
they later move towards the main river course (Fig. 5.9e, f), as confirmed by the bump in the YOY population in site #16 (see Fig. 5.10a). When PKD is present, the equilibrium distribution is perturbed by the enhanced disease severity at the sites characterized by large contributing area (see Chapter 4 and Fig. 4.2 in particular), thus resulting in a more spatially homogeneous distribution of adult and YOY fish density (Fig. 5.9a, c, d), whereas the initial YOY density remains higher in small headwaters (Fig. 5.9b).

The decrease of brown trout population size was estimated only for sites where two sampling campaigns in the same year were conducted. The decline in fish abundance is captured by the model (Fig. 5.8b), despite a tendency towards underestimation. This could be attributed to undocumented fishing activity or other possible stress factors not included in the model at the current stage. Overall, PKD engenders a remarkable decline in the fish population abundance with respect to the disease-free case (92.5% for YOY and 95.7% for adults at the end of the warm season for the best fit simulation shown in Fig. 5.9). The bottom row of Fig. 5.9
Figure 5.10 – Time evolution of modelled fish population in site #16.

Figure 5.11 – Posterior distributions of the calibrated parameters. The total length of the Markov Chain is 1250 steps. Dashed lines represent best-fit values.

represents a possible scenario of PKD complete eradication according to the present model and knowledge, although other density dependent effects regulating the fish population not included here might play a role in reducing fish abundance. Best fit and posterior distributions of the calibrated parameters are reported in Fig. 5.11.

5.3 Discussion

Disease emergence may occur variously, through either range expansion of existing pathogens or appearance of new, more virulent agents in existing endemic ranges [Keesing et al., 2010; Penczykowski et al., 2014]. Environmental change can also trigger the emergence of previously relatively avirulent, endemic parasites by altering the expression of virulence via e.g.
temperature-linked effects on host immune function [Engering et al., 2013]. Local abiotic and biotic conditions favoring parasite proliferation might vary in time owing to environmental change. In the case of parasites with complex life-cycles, such conditions must remain conducive to the persistence of multiple susceptible host classes. For example, the correct species of snail and vertebrate hosts are required to co-exist at appropriate points during the life-cycle of *Schistosoma mansoni* parasites to sustain the transmission of schistosomiasis [Gryseels et al., 2006; Perez-Saez et al., 2016]. Although *T. bryosalmonae* exhibits a similarly complex life-cycle with no fish-to-fish transmission, a notable difference is that that long-term parasite persistence in the bryozoan populations is possible, even in the absence of the fish host (see Chapter 4 and Fontes et al. [2017b]). Parasite propagation along the budding growth of the bryozoan and incorporation into asexually produced resting stages create an effective parasite reservoir with frequent spill-over effects to e.g. stocked and highly susceptible fish. For this type of pathogens, any environmental change favoring establishment of the key reservoir host increases disease risk in all the other hosts, including those that may be economically relevant. It also greatly complicates eradication of the disease through control measures.

The developed model suggests that increased bryozoan density is highly indicative of PKD severity and occurrence, and thus the key predictive factor for PKD was the habitat suitability for the bryozoan *F. sultana*. A strong correlation between the presence of upstream moraines and local abundance of *F. sultana* was recovered. This pattern, which was instrumental in generating the bryozoan density map for the Wigger, might imply that moraines create advantageous conditions for the proliferation of *F. sultana* by constituting the suitable substrate and/or by affecting geogenic solute concentrations, i.e. the chemical properties of the stream environment. However, moraines are present in a limited area of the watershed with few sampling sites with notably high *F. sultana* eDNA concentration; therefore a spurious correlation cannot be excluded. Understanding the causality and mechanisms explaining why moraines are priming bryozoan presence requires further survey studies in other catchments and laboratory experiments. Knowledge on habitat requirements of bryozoans might be crucial for disease control purposes, as the strong link between PKD severity and bryozoan density suggests that PKD management might rely on control of bryozoan populations. Other possible strategies might count on the production of resistant fish strains, also by eliminating fish stocking and thus allowing for selection of resistant fish strains in a natural way.

With regards to the epidemiological model, it is noteworthy that, even without specifying different epidemiological parameters for YOY and adults, the model predicts a population size decline of YOY that is almost three times larger than that of adults. Therefore, the higher susceptibility of YOY to PKD revealed by the data can be explained by the not yet acquired immunity rather than by an intrinsic severity of PKD for young fish. The observation that the decline in adults predicted by the model is lower than in observed values can be explained by an additional fishing mortality term not accounted for by the model (say, widespread anglers’ impacts or other stress factors, possibly of anthropic origin). Confidence intervals of fish decline are rather narrow because the main factor influencing adults’ reduction is the natural mortality rate, which was kept constant in the calibration procedure. Other possible stress
factors related to temperature increase were already implicitly taken into account, because the PKD-related mortality is expressed as a function of temperature. The seasonal decrease in the abundance of YOY observed at site #8 in 2016 was considerably low (see Fig. 5.8b). It is remarkable that the model is actually capable of partially reproducing this trend by predicting a lower YOY decrease compared to that of 2015. This is likely due to the shorter time lapse between the two seasonal sampling campaigns and the cooler summer temperatures during that season (see §5.1.3). Indeed, water temperature proves a key driver of PKD impacts on fish population abundance even though its influence on prevalence patterns is minor (as observed in Chapter 4).

In conclusion, this work highlights the profound influence of an emerging aquatic disease on the abundance and seasonal demography of threatened and economically important fish stocks. The integrated approaches resulted in a comprehensive spatial predictive framework of disease and identified key factors in driving disease patterns in the wild.

### 5.4 List of symbols

The following table lists all symbols used in this Chapter. This table does not include state variables and parameters of the epidemiological model (presented in Tables 2.3 and 5.3, respectively).

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>Upstream drainage area (of stretch $i$)</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$A'_i$</td>
<td>Watershed area directly drained by stretch $i$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$A_F$</td>
<td>Shape factor for spawning suitability $W^A$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$C_i$</td>
<td>$E. sultana$ eDNA concentration measurable in stretch $i$</td>
<td>$NL^{-3}$</td>
</tr>
<tr>
<td>$c_i$</td>
<td>$E. sultana$ eDNA concentration in an unconnected stretch $i$</td>
<td>$NL^{-3}$</td>
</tr>
<tr>
<td>$c_0$</td>
<td>Scale factor for exponential link function used for $c_i$</td>
<td>$NL^{-3}$</td>
</tr>
<tr>
<td>$C_i^{m}$</td>
<td>Mean measured $E. sultana$ eDNA concentration at stretch $i$</td>
<td>$NL^{-3}$</td>
</tr>
<tr>
<td>$\overline{C_i^{m}}$</td>
<td>Mean of $C_i^{m}$ at all sampling sites except $i$</td>
<td>$NL^{-3}$</td>
</tr>
<tr>
<td>$D_i$</td>
<td>Mean water depth in stretch $i$</td>
<td>$L$</td>
</tr>
<tr>
<td>$L_{ij}$</td>
<td>Downstream length from stretch $i$ to $j$</td>
<td>$L$</td>
</tr>
<tr>
<td>$N$</td>
<td>Measured adult fish abundance</td>
<td>-</td>
</tr>
<tr>
<td>$N_t$</td>
<td>Total number of river stretches</td>
<td>-</td>
</tr>
<tr>
<td>$NS_k$</td>
<td>Nash-Sutcliffe index used for calibration in LOOCV</td>
<td>-</td>
</tr>
<tr>
<td>$p_{ij}$</td>
<td>Entry of the connectivity matrix $P$</td>
<td>-</td>
</tr>
<tr>
<td>$Q_i(\tau)$</td>
<td>Discharge at time $\tau$ in stretch $i$</td>
<td>$L^3T^{-1}$</td>
</tr>
<tr>
<td>$s$</td>
<td>Performance index used to rank models of bryozoan habitat suitability</td>
<td>-</td>
</tr>
<tr>
<td>$T_i(\tau)$</td>
<td>Temperature in stretch $i$ at day $\tau$ obtained by Eq. (5.1)</td>
<td>$\Theta$</td>
</tr>
</tbody>
</table>
### Chapter 5. Case study: An integrated analysis of PKD spread in a Swiss prealpine river

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{a,i}$</td>
<td>Mid-amplitude of the sinusoidal signal $T_{s,i}$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{b}(\tau)$</td>
<td>Baseline temperature at day $\tau$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{m,i}$</td>
<td>Annual mean temperature in stretch $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{s,i}(\tau)$</td>
<td>Sinusoidal fitted temperature in stretch $i$ at day $\tau$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$\overline{V}_i$</td>
<td>Mean water volume in stretch $i$</td>
<td>L$^3$</td>
</tr>
<tr>
<td>$X_i$</td>
<td>Vector of normalized covariates evaluated at stretch $i$</td>
<td>-</td>
</tr>
<tr>
<td>$Y$</td>
<td>Measured YOY abundance</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha(\tau)$</td>
<td>Vector of parameters for the temperature model at day $\tau$</td>
<td>-</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Vector of parameters for the bryozoan habitat suitability model</td>
<td>-</td>
</tr>
<tr>
<td>$\lambda_B$</td>
<td>Characteristic decay length for $F.~sultana$ eDNA</td>
<td>L</td>
</tr>
<tr>
<td>$\lambda_F$</td>
<td>Characteristic distance covered by $S.~trutta$ for spawning</td>
<td>L</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Time in days</td>
<td>T</td>
</tr>
<tr>
<td>$\tau_{min,i}$</td>
<td>Ordinal day of annual minimum temperature in stretch $i$</td>
<td>T</td>
</tr>
</tbody>
</table>
Environmental DNA (eDNA) is a relatively new and widely used technique for detecting rare or invasive species in freshwater environments. The exact localization of the target species and assessment of its density based on eDNA surveys remain a demanding task, but would be highly valuable for management and monitoring. This Chapter proposes a general framework for using downstream eDNA measurements to reconstruct the upstream location and density of a target species across a river basin. This framework, based on mass conservation principles and hydrological transport concepts, can be coupled with other population and ecological models in order to detect the spatial location of the target species and its local biomass, as well as their correlations with environmental variables. The framework is tested by using quantitative eDNA measurements of Tetracapsuloides bryosalmonae at various locations within the Wigger watershed (Switzerland). T. bryosalmonae is excreted both by an immobile source (overtly infected bryozoans) and a mobile source (fish infected with PKD), and the two types of spores are genetically indistinguishable yet differentiated in terms of function. This aspect poses further challenges from a modelling perspective, as the two sources must be singled out. The case study represents a paradigm of the potentiality of this framework, as predicting the spatial pattern of bryozoan density is crucial for the assessment of the spatial distribution of PKD prevalence in fish and, ultimately, for the implementation of disease mitigation strategies. The conclusion of this Chapter is devoted to generalizations in the case of a target species with uniform spatial density, and to connections with the geomorphological theory of the hydrologic response.
Chapter 6. Using eDNA to track species distributions in river networks

6.1 Methods

The eDNA transport model here proposed stems from mass balance relationships. It is intended for use in river networks discretized into ‘units’ or ‘nodes’ i.e., stretches of suitable length within which hydrological conditions, as well as the target species density and its eDNA production can be considered homogeneous. As the discretization of the fluvial domain can be made arbitrarily small, one assumes that the choice of characteristic stretch could be suitably made small with respect to the correlation scale of any relevant heterogeneous property.

Within such units, the basis for the spatially-explicit source area model is the measured eDNA concentration, as follows (see §5.1.4):

\[
C_j = \frac{1}{A_j} \sum_{i \in \gamma(j)} A_i^L \exp \left( - \frac{L_{ij}}{\lambda} \right) c_i, \tag{6.1}
\]

where: \( C_j \) is the measured eDNA concentration at node \( j \); \( A_j \) is the total upstream catchment area at node \( j \), a proxy of landscape-forming discharges and of a range thereof [Rodriguez-Iturbe and Rinaldo, 2001]; \( \gamma(j) \) indexes the full set of nodes upstream of \( j \) connected by the river network\(^1\); \( A_i^L \) the directly contributing area to node \( i \) (i.e., if a watershed of total area \( A \) is divided into \( N_s \) stretches, \( \sum_{i=1}^{N_s} A_i^L = A \) [Muneepeerakul et al., 2008]); \( L_{ij} \) the length of the path connecting \( i \) to \( j \); \( \lambda \) a parameter termed decay length subsuming the damage of genetic material due to the hydrologic transport; \( c_i \) the eDNA concentration due to local production at \( i \). It is reasonable to assume that \( c_i \) is proportional to the target species density at node \( i \) [Takahara et al., 2012]. Eq. (6.1) is based on the following hypotheses: i) eDNA undergoes first-order exponential decay along the downstream path from the source possibly at \( i \) to the measuring node \( j \) [Jerde et al., 2016]; ii) water discharge is assumed proportional to contributing area for a broad range of streamflows, even if actual hydrologic conditions may be far removed from landscape-forming events [Rodriguez-Iturbe and Rinaldo, 2001]. The latter hypothesis allows the estimation of time-averaged spatial distributions of \( c_i \) at the \( i \)-th node starting from time-averaged measured eDNA concentrations \( C_j \). Note that this may be relaxed by replacing \( A_j \) and \( A_i^L \) with the actual water discharge at node \( j \) and time \( t \) (\( Q_j(t) \)) and the directly contributing discharge \( Q_i^L(t) \) (i.e. the runoff pertaining to the portion of catchment directly draining into stretch \( i \)), respectively. In the latter case, Eq. (6.1) would enable the evaluation of the instantaneous spatial distribution of \( c_i \), and even the time evolution of target species densities across the catchment. However, a significant added complexity stems from the need to use a spatially distributed hydrological model to estimate water fluxes, such as e.g. the one proposed in §7.2.2 in the context of stream temperature modelling. While all the above assumptions are refinable, the above framework should be taken as a solid building block towards a general theory of spatially-explicit eDNA source tracking.

The parameter \( \lambda \), incorporating eDNA degradation process through a characteristic length scale of an exponential decay, may assume values ranging from the order of magnitude of a few

\(^1\)Thus it is implicitly assumed that spores are transported as passive scalars, i.e. they are advected downstream by hydrologic velocities without active drifts, see e.g. Wilcox et al. [2015]
6.1. Methods

a) Measured eDNA concentration of *T. bryosalmonae* averaged over time. The precise location of the sampling sites is that of the pink dots of Fig. 5.1a. Ungauged stretches are depicted in blue. b) *F. sultana* local eDNA productions obtained as average of all behavioural models (see §5.1.4 and Fig. 5.7 for details). This distribution is assumed to be proportional to that of bryozoan density at carrying capacity. c) Map of mean seasonal PKD prevalence in bryozoans, evaluated via the calibrated epidemiological model of Chapter 5.

In Chapter 5, temporally averaged *Fredericella sultana* 18S rDNA concentrations measured at different sites were used to derive a map of its source density for the case study area. In that case, the assumption \( c_i = c_0 \exp(\beta'X_i) \) enabled to assess the effects of local, uncorrelated environmental variables (defined by the vector \( X_i \)) through calibration of the parameters \( c_0 \) and \( \beta \). Spatial patterns of PKD prevalence and fish decline were studied by means of a spatially-explicit epidemiological model using the reconstructed spatial map of bryozoans. Building on such results\(^2\), the local *T. bryosalmonae* eDNA production can be expressed as

\[
c_i = k_R p_{B,i} B_i + k_F p_{F,i} F_i,
\]

\(^2\)\(B_i, F_i, p_{B,i}, p_{F,i}\) are assumed as in the best-fit simulation of the epidemiological model presented in Chapter 5 (see Figs. 5.8 and 5.11). In particular, \( F_i \) and \( p_{F,i} \) refer to the overall fish population, without operating the distinction between YOY and adult individuals.
where: $B_i$ ($F_i$) is the modelled local average density of bryozoans (fish), displayed (rescaled in terms of bryozoan eDNA production) in Fig. 6.1b; $p_{B,i}$ ($p_{F,i}$) is the estimated average PKD prevalence\(^3\) in bryozoan (fish), shown in Fig. 6.1c; and $k_B$ and $k_F$ are scale parameters. Thus a proper modelling framework supports field data collections to track source areas of the target species and its average biomass.

The model (6.2) was calibrated against time-averaged measured $T. bryosalmonae$ eDNA concentrations collected\(^4\) at 15 locations (see Fig. 6.1a). For all locations except one, 21 water samples were taken at irregular intervals in the period May 2014 - May 2015. Site #5 was abandoned after 12 samples. The assessment of $T. bryosalmonae$ eDNA concentration in water samples was performed according to the protocol described in Fontes et al. [2017a]. Model calibration was performed by maximization of the Nash-Sutcliffe performance index

$$NS = 1 - \frac{\sum_{k=1}^{15} (C^m_k - C_k)^2}{\sum_{k=1}^{15} (C^m_k - \frac{1}{15} \sum_{j=1}^{15} C^m_j)^2} \tau,$$

where $C^m_k$ is the time-averaged measured eDNA concentration at the sampling site $k$, and $C_k$ the modelled eDNA concentration at the same site [obtained from Eq. (6.2)]. To this end, a particle swarm optimization algorithm [Kennedy and Eberhart, 1995] was used.

**6.2 Results and discussion**

The optimal value found for $k_F$ was null (within accuracy), implying that the amount of $T. bryosalmonae$ spores excreted by fish is negligible compared with that excreted by overtly infected bryozoans. While this result may appeal to previous evidence [Tops and Okamura, 2003], this may not be the case for disease spread of other waterborne diseases where mobile sources of pathogens are dominant. The same result was found after calibration of the alternative model\(^5\) where $\tau$ instead of $\lambda$ was kept as a free parameter. In the latter case, however, the fit was sensibly better (Nash-Sutcliffe index $NS = 0.89$ against $NS = 0.79$ obtained for the former model – see Figs. 6.2b, c). This suggests that the introduction of decay time $\tau$ improves prediction capabilities by adequately characterizing the extent of damage to the genetic material as a consequence of the variability in path-dependent flow conditions. The obtained maps for measurable eDNA concentration $C$ (Fig. 6.2a) and for source density $c$ (Fig. 6.2d) for

---

\(^3\)Unlike in the previous Chapters, prevalence is here intended as the fraction of bryozoans/fish belonging to spore-producing classes, namely overtly infected bryozoans $B_O$, or acutely infected ($Y_i,F_i$) and carrier ($Y_C,F_C$) fish.

\(^4\)Collection of water samples (performed as reported in §5.1.2) and assessment of $T. bryosalmonae$ eDNA concentration were carried out by Hanna Hartikainen, Inês Fontes and collaborators.

\(^5\)In this model, the distribution of velocities $v_k$ was obtained by rescaling the time-averaged discharge measured at Zofingen (corresponding to sampling site #3) to all stretches by assuming direct proportionality between discharge and area; the discharge $Q_k$ at stretch $k$ was then converted to $v_k$ by means of any suitable uniform-flow formula representative of average conditions within any single reach. In particular, the conversion to mean flow velocity was operated by assuming river’s cross-sections as rectangular with widths much larger than depths, and with a Manning coefficient $n = 0.033$ m$^{-1/3}$s. More details are provided in §5.1.1 and §7.2.2.
6.3 eDNA modelling and the width function concept

*Trichoplusia bryosalmonae* resemble, to some extent, those obtained for *F. sultana* (Figs. 5.7c and 6.1b, respectively), with the difference that sources of *T. bryosalmonae* are the network nodes richest in infected bryozoans. Hence the pattern displayed in Fig. 6.2d may be seen as a rescaled version of the maps in Figs. 6.1b and 6.1c. Therein, bryozoan prevalence generally increases downstream as an intrinsic hierarchical effect of network connectivity (as shown for fish prevalence in Chapter 4). Higher magnitudes of source density of *T. bryosalmonae* with respect to those of *F. sultana* reflect the kind of mass balance implicit in the models (6.1) and (6.2).

The choice of a parametrization for $c_i$ depends on the nature of the link between the target species density and its biological and environmental filters that one aims at investigating, and it can benefit from acquired knowledge or integrate other population or species distribution models, as showed here for the case of *T. bryosalmonae*. In order to successfully design an eDNA sampling campaign in river networks, the number of sampling sites should be much greater than the number of unknown parameters, to avoid issues with overfitting [Hawkins, 2004]. In order to obtain eDNA measurements not affected by the contribution of low-density or species-free areas, it is advisable to choose as sampling sites the outlets of tributaries where one expects to find high target species densities. Decay time can be held as an unknown parameter, although the inclusion of some physical elements might help define an appropriate range for $\tau$ and improve the confidence on the estimated values of parameters chosen to express $c_i$. For instance, eDNA half-life in water (equal to $\ln(2) \cdot \tau$) is impacted by water temperature, microbial loads and the pH, and could be assessed experimentally for the environmental conditions of interest [Lance et al., 2017].

In summary, tracking source areas and the local biomass density of target species’ via downstream eDNA measurements is possible provided that suitable quantitative techniques are used to track the biomass of target species. The use of qPCR in eDNA analyses, coupled to the type of hydro-geomorphological rescaling and parameter/model identification techniques employed here, is thus argued to open a whole new direction in ecohydrological studies.

6.3 eDNA modelling and the width function concept

Estimating source areas of eDNA from samples at the closure section of a catchment relates to one of the fundamental questions in hydrology, namely whether it is possible to recover the salient geomorphic features of a watershed by solely measuring its hydrologic response [Rinaldo et al., 1995]. Geomorphic controls on the hydrologic response of a catchment are embedded into travel time distributions $f(t)$ (i.e. the residence time distribution sampled at the outlet of the control volume assumed as an absorbing barrier) following an instantaneous injection uniformly distributed in space [Rodriguez-Iturbe and Rinaldo, 2001]. $f(t)$ may be reasonably assumed to be stationary because unchanneled pathways within the catchment are immaterial in the problem at hand. Travel times can be expressed in terms of the width function $W(x)$ (i.e. the fraction of nodes at distance $x$ from the fixed outlet) as $f(t) = u W(u t)/\delta$ [Rodríguez-Iturbe and Valdés, 1979; Rinaldo et al., 1995], where $u$ is a constant drift/celerity.
Chapter 6. Using eDNA to track species distributions in river networks

Figure 6.2 – Results from application of the model on *T. bryosalmonae* data. a) Map of modelled eDNA concentrations \( C_i \) obtained via the model with \( \tau \) as parameter. b) Modelled \( C_i \) vs. measured \( C_m \) eDNA concentrations for models with either decay length \( \lambda \) or decay time \( \tau \) as free parameters. Best fit values for these parameters are displayed. c) Zoom from panel b. d) Modelled *T. bryosalmonae* local eDNA productions \( c \).

along the network of links, and \( \delta \) a geometrical scale, i.e. the pixel size in the case of networks extracted from digital elevation models (see e.g. Tarboton et al. [1991])\(^6\). By considering the baseline assumptions of a target species whose density is uniformly distributed across the basin \( c_i = c_0 \), and a spatially constant downstream drift of weightless genetic material (i.e. the ratio \( \lambda / \tau \) is path-independent), Eq. (6.1) can be reformulated as:

\[
\frac{C_i}{c_0} = \frac{1}{\delta} \int_0^{L_{\text{max}}} W(x) e^{-x/\lambda} \, dx, \tag{6.3}
\]

where \( L_{\text{max}} \) is the maximum length from source to outlet.

Width functions reflect both universal (i.e. unrelated to exposed lithology, vegetation cover, climate and geology) and peculiar (pertaining to the particular constraints under which of capture and migration of divides occur) features, the former being linked to the high-frequency modes of the power spectrum of \( W(x) \), and the latter to its low-frequency modes controlled by the shape of the catchment at hand [Rinaldo et al., 1995; Rodriguez-Iturbe and Rinaldo, 2001]. In this perspective, Eq. (6.3) exhibits the footprint of river geomorphology on the transport of eDNA. Eq. (6.3) was tested on 50 optimal channel network (OCN) replicas\(^7\) generated from lattices of side \( L = 128 \) pixels (as in Fig. 6.3a), where the (single) outlet is placed at randomly

---

\(^6\)Generalizations in the case where water particles are subject to a Brownian motion in addition to the constant drift are straightforward, see Rinaldo et al. [1995], but beyond the scope of this work.

\(^7\)As recalled in Chapter 4, OCNs are spanning trees that reproduce all the mutually connected topological and metric features of real river networks [Rinaldo et al., 2014]. Details on the construction of the OCNs can be found in §4.1.1.
chosen sites at the lattice borders in each realization. The control that the basin shape exerts on eDNA transport is proportional to the ratio of the decay length $\lambda$ to the catchment size $L$ (see Fig. 6.3c). When $\lambda \ll L$, the eDNA concentration at the outlet tends to zero as all upstream contributions are depleted along most of their paths to the outlet. Conversely, when $\lambda \gg L$ no decay occurs. When $\lambda \approx L$, network connectivity has a sensible impact on $C_{\text{outlet}}/c_0$, and such a variability is negatively correlated with the average path length to the outlet (Fig. 6.3d). While model Eq. (6.1) is to be preferred due to its high degree of flexibility, Eq. (6.3) may serve as a baseline model, applicable for widely distributed target species. Any generalizations of Eq. (6.3) for patterns of species density that can be expressed as a function of the distance to the outlet $x$ are straightforward. This aspect seems particularly relevant, as the distance to the outlet acts as a proxy of relevant environmental or morphological variables, such as water temperature, elevation or channel slope/velocity.

### 6.4 List of symbols

The following table lists all symbols used in this Chapter.
### Chapter 6. Using eDNA to track species distributions in river networks

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>Upstream drainage area (of stretch $i$)</td>
</tr>
<tr>
<td>$A_i^L$</td>
<td>Watershed area directly drained by stretch $i$</td>
</tr>
<tr>
<td>$B_i$</td>
<td>Bryozoan abundance at carrying capacity (in dimensionless units) as displayed in Fig. 5.7</td>
</tr>
<tr>
<td>$C_i$</td>
<td>eDNA concentration measurable in stretch $i$</td>
</tr>
<tr>
<td>$C_{\text{outlet}}$</td>
<td>eDNA concentration measurable at the outlet of an OCN</td>
</tr>
<tr>
<td>$c_i$</td>
<td>eDNA concentration in an unconnected stretch $i$</td>
</tr>
<tr>
<td>$c_0$</td>
<td>eDNA production of a species uniformly distributed in space</td>
</tr>
<tr>
<td>$C_{m_i}^m$</td>
<td>Mean measured <em>T. bryosalmonae</em> eDNA concentration at stretch $i$</td>
</tr>
<tr>
<td>$F_i$</td>
<td>Mean fish abundance at stretch $i$ according to the best-fit epidemiological model of Chapter 5</td>
</tr>
<tr>
<td>$f(t)$</td>
<td>Distribution of residence times sampled at a river outlet</td>
</tr>
<tr>
<td>$k_B$</td>
<td>Scale parameter expressing contribution of infected bryozoans to <em>T. bryosalmonae</em> eDNA production</td>
</tr>
<tr>
<td>$k_F$</td>
<td>Scale parameter expressing contribution of infected fish to <em>T. bryosalmonae</em> eDNA production</td>
</tr>
<tr>
<td>$L_{avg}$</td>
<td>Average path length to the outlet of an OCN</td>
</tr>
<tr>
<td>$L_{max}$</td>
<td>Maximum path length to the outlet of an OCN</td>
</tr>
<tr>
<td>$L_{ij}$</td>
<td>Downstream length from stretch $i$ to $j$</td>
</tr>
<tr>
<td>$l_i$</td>
<td>Length of stretch $i$</td>
</tr>
<tr>
<td>$N_s$</td>
<td>Total number of river stretches</td>
</tr>
<tr>
<td>$P_{i\rightarrow j}$</td>
<td>Set of stretches connecting stretch $i$ to $j$</td>
</tr>
<tr>
<td>$p_{B,i}$</td>
<td>Mean prevalence in bryozoans at stretch $i$ according to the best-fit epidemiological model of Chapter 5</td>
</tr>
<tr>
<td>$p_{F,i}$</td>
<td>Mean prevalence in fish at stretch $i$ according to the best-fit epidemiological model of Chapter 5</td>
</tr>
<tr>
<td>$Q_i(t)$</td>
<td>Discharge in stretch $i$ (at time $t$)</td>
</tr>
<tr>
<td>$Q_i^L(t)$</td>
<td>Directly contributing discharge at stretch $i$ (at time $t$)</td>
</tr>
<tr>
<td>$u$</td>
<td>Constant velocity along all stretches of a basin</td>
</tr>
<tr>
<td>$v_i$</td>
<td>Water velocity in stretch $i$ (at steady state)</td>
</tr>
<tr>
<td>$v_{ij}$</td>
<td>Mean velocity across along the trajectory $L_{ij}$</td>
</tr>
<tr>
<td>$W(x)$</td>
<td>Width function of a basin</td>
</tr>
<tr>
<td>$\gamma(i)$</td>
<td>Set of stretches upstream of $i$</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Size of a pixel in an OCN landscape</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Characteristic decay length of eDNA material</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Decay time of eDNA material</td>
</tr>
</tbody>
</table>
The findings illustrated in the previous Chapters have singled out the importance of stream water temperature on PKD dynamics. Stream temperature is also a key variable controlling chemical, biological and ecological processes in freshwater environments. Stream temperature models have a long history; however, most of them only focus on given river’s cross-sections, whereas temperature gradients along stretches and tributaries of a river network are crucial to assess ecohydrological features such as e.g. aquatic species suitability, growth and feeding rates, or disease transmission. This Chapter proposes a deterministic, spatially-explicit stream temperature model for a whole river network, based on water and energy budgets at a reach scale and requiring only easily available spatially distributed datasets such as morphology and air temperature. Heat exchange processes at the air-water interface are modelled via the widely known equilibrium temperature concept (according to which the net heat flux exchanged between air and water is proportional to the difference between water temperature and the equilibrium temperature), but new parametrizations for this variable are here proposed. In the context of PKD studies that constitute the leit-motiv of this Thesis, a case study was conducted on the prealpine Wigger river (Switzerland), where water temperatures have been measured in the period 2014-2017 at eleven spatially distributed locations. Results show the advantages of accounting for water and energy budgets at the reach scale for the entire river network, compared to simpler, lumped formulations. The spatially-explicit model allows to trust regionalized values for the regression coefficient of the relationship between equilibrium temperature and air temperature, and enables the derivation of reliable estimates of mean daily stream temperatures for the whole catchment based on a limited number of conveniently located (namely, spanning the largest possible elevation range) measuring stations. Moreover, the use of a sinusoidal proxy for absorbed solar radiation in the computation of equilibrium temperature resulted in enhanced model performances compared with more classical parametrizations.
Chapter 7. A stream temperature model for ecohydrological applications

7.1 Introduction

Stream temperature is a crucial variable for the study of aquatic ecosystems, as it controls several chemical, biological and ecological processes, such as nutrient consumption rate, photosynthesis, respiration, oxygen solubility and sediment suspension [Webb et al., 2008; Ducharne, 2008]. Aquatic organisms have a limited temperature range in which they can grow and reproduce, and alterations to this range can have a significant effect. For instance, temperature variations related to climate change have devastating repercussions on salmonid populations [Hari et al., 2006; McCullough et al., 2009]. As extensively discussed in the introduction of this Thesis, water temperature is also a major driver of PKD incidence and related mortality in salmonid fish.

Stream temperature is controlled by a number of factors, subdivided by Caissie [2006] into four main groups: topography (such as stream aspect, latitude, altitude, slope, riparian vegetation), discharge (water volume, turbulence, inflows and outflows), atmospheric conditions (solar radiation, air temperature, wind speed, humidity, rainfall, snow, evaporation) and influence of the streambed (conduction, hyporheic exchange, groundwater input). Across a long history of stream temperature models [Smith, 1972], efforts have been made at isolating those few key factors that, depending on the spatial and temporal scales of interest, have the most impact on temperature dynamics. Models can be essentially divided into two main categories: statistical models and deterministic models. Statistical models mainly consist in linear or logistic regression between water temperature and few relevant predictors, among which air temperature plays a major role [Benyahya et al., 2007]. Despite their simplicity, these models typically show good performances, especially at larger time scales (such as weekly or monthly), therefore they are often employed in practical applications. When temperature has to be modelled at finer (hourly or daily) scales, stochastic or deterministic models might be preferred. Stochastic models (e.g. Caissie et al. [2001]) belong to the group of statistical models, but unlike regression models they take into account the auto-correlation of the water temperature time series. On the other hand, deterministic models consider all the relevant hydrological and heat exchange processes characterizing the energy budget.

In ungauged catchments, regression models are often employed to determine the impact of human activities and climate change on the environment [Gallice et al., 2015]. However, such models typically focus only on yearly aggregated temperature metrics, and therefore they are unsuited to predict the annual cycle of stream temperature, which is a key factor controlling the distribution of fish and other riverine species [Heggenes et al., 1999]. The existing modelling approaches span different degrees of spatial detail and physical representativeness [Comola et al., 2015; Piccolroaz et al., 2016], although most models predict water temperatures only at given river's cross-sections (typically, the outlet). Conversely, the prediction of spatial gradients of water temperature at a network scale is critical for models aimed at ecohydrological applications, such as assessing spatial niche variability for aquatic species and potential effects in propagation of diseases (of which PKD constitutes a paradigmatic example). In order to be conveniently applicable, such models should be based on easily accessible datasets.
such as river morphology and air temperature, as extended and detailed datasets on upland shading and riparian zone conditions may be hardly accessible.

In deterministic models, the most relevant components of the heat budget is the total heat flux exchanged at the water surface, which subsumes the processes of solar radiation, net long-wave radiation, evaporation and conduction. In order to simplify its expression and reduce the number of input variables required, many deterministic models successfully exploit the so-called equilibrium temperature concept [Edinger et al., 1968; Caissie et al., 2005], according to which the total heat flux is proportional to the difference between water temperature and the equilibrium temperature, an unknown variable which is typically related to air temperature and possibly other predictors. The drawback of this approach is that the regression coefficients typically reflect local conditions and hence they allow reliable predictions through time, but they would result in poor performances when used to predict stream temperatures at some distance from the site where they were calibrated.

The aim of this Chapter is to propose a deterministic stream temperature model for a whole river network, based on the application of water and energy budgets at a reach scale and on the equilibrium temperature concept to express local heat transfer. In particular, the spatial variability of water temperature dynamics is constrained into the processes of heat transport and mixing, while a spatially constant parametrization of the equation linking air and equilibrium temperature is maintained. Such a model is tested on a case study prealpine Swiss river (the Wigger), where water temperatures have been collected at 11 different locations in the period 2014-2017. Several types of parametrizations with increasing levels of complexity are tested, in order to assess to which degree increasing model complexity is worthwhile.

7.2 Methods

7.2.1 Study area

The river Wigger (see Fig. 7.1), located in the Swiss plateau, is a tributary of the river Aare and has a length of 48.11 km. It drains a watershed of 382.4 km² which has an elevation range comprised between 396 and 1409 m a.s.l., at the Mount Napf. Water temperature has been measured since July 2014 in 10 sites via HOBO TidbiT® v2 data loggers. An additional temperature gauge (site #11 in Fig. 7.1a) was added in September 2015. Two loggers, recording data at 15-minutes intervals, were deployed per each site. Air temperature data from 8 neighbouring stations were provided by MeteoSwiss. Daily mean discharges are measured by the Swiss Federal Office for the Environment (FOEN) in Zofingen (corresponding to site #1 in Fig. 7.1a).

1The river network extraction was operated as described in §5.1.1, to which the reader is referred for further information concerning the study area. The dataset of stream temperatures used in this Chapter is the same as the one presented in §5.1.3.
Chapter 7. A stream temperature model for ecohydrological applications

Figure 7.1 – a) Digital Elevation Model of the case study region. The black solid line identifies the extension of the Wigger catchment; light blue lines represent the extracted river network. The location of water temperature loggers and MeteoSwiss air temperature stations is also displayed. b) Position of the Wigger catchment in Switzerland with location of soil temperature stations used in this study.

7.2.2 Formulation of the spatially-explicit model

Let us consider a river network discretized into $N_i$ reaches of suitable length (Fig. 7.2). Reaches are assumed to be well mixed, i.e. the water temperature $T_i(\tau)$ at time $\tau$ is constant along reach $i$. The connectivity between reaches is expressed via an adjacency matrix $W$ with entries $w_{ji} = 1$ if reach $j$ drains directly into reach $i$ and null otherwise. River’ s cross-sections are approximated as rectangular shapes, with widths $B$ constant along the reach and much greater than river depths $D$. Therefore, the water volume at a reach $i$ reads $V_i(\tau) = L_i B_i D_i(\tau)$, where $L_i$ is the reach length. The water budget at reach $i$ reads:

$$\frac{dV_i}{d\tau} = \sum_{j=1}^{N_i} w_{ji} Q_j(\tau) + Q^L_i(\tau) - Q_i(\tau) \quad (7.1)$$

where $Q_i(\tau)$ is the water discharge flowing out of reach $i$, and $Q^L_i(\tau)$ the discharge contribution from the subcatchment directly draining into reach $i$.

On the other hand, the energy budget reads:

$$\frac{dH_i}{d\tau} = \rho c_p \frac{d(V_i T_i)}{d\tau} = \sum_{j=1}^{N_i} w_{ji} \phi_j(\tau) + \phi^L_i(\tau) - \phi_i(\tau) + \phi^na_i(\tau) \quad (7.2)$$

where: $H_i(\tau)$ is the total thermal energy of reach $i$; $\rho$ is the density of water; $c_p$ the specific heat of water at constant pressure; $\phi_i(\tau) = \rho c_p T_i(\tau) Q_i(\tau)$ the heat flux flowing through the exit cross-section of reach $i$; $\phi^L_i(\tau) = \rho c_p T^L_i(\tau) Q^L_i(\tau)$ the heat flux related to the water flux.
7.2. Methods

Figure 7.2 – Schematic representation of the water and energy budgets at the reach scale. Blue indicates water fluxes, red energy fluxes.

$Q^L_i (\tau)$, with $T^L_i (\tau)$ being the temperature of $Q^L_i (\tau)$; $\phi_{n}^{na}(\tau)$ the overall non-advective heat flux at reach $i$. $T^L_i (\tau)$ can be seen as the weighted average temperature of $Q^L_i (\tau)$, as this flux is typically made up of different contributions (e.g. groundwater, surface runoff, snowmelt). By combining Eqs. (7.1) and (7.2), one gathers

$$\frac{dT_i}{d\tau} = \frac{1}{V_i} \left[ \sum_{j=1}^{N_i} w_{ji}(T_j - T_i)Q_j + (T^L_i - T_i)Q^L_i + \frac{\phi_{na}^{n}}{\rho c_p} \right].$$

(7.3)

The dependence on time $\tau$ of the concerned variables is omitted for simplicity here and in the following.

It is reasonable to assume that propagation of flow perturbations is much faster than the daily time scale used in the model, at least for catchments up to $10^4$ km$^2$. Therefore flow regime can be approximated as a succession of steady states where at each time discharge is proportional to contributing area $A$ [Rodriguez-Iturbe and Rinaldo, 2001]: $Q_j = \beta_{1,ji}Q_i$, where $\beta_{1,ji} = A_j / A_i$. This allows writing

$$Q^L_i = \frac{dV_i}{d\tau} + \left( 1 - \sum_{j=1}^{N_i} w_{ji}\beta_{1,ji} \right) Q_i = \frac{dV_i}{d\tau} + \beta_{2,i}Q_i.$$  

(7.4)

Non-advective heat fluxes can be expressed as $\phi_{i}^{na} = \phi_{i}^{aw} + \phi_{i}^{sb}$, where $\phi_{i}^{aw}$ represents the heat flux at the air-water interface, while $\phi_{i}^{sb}$ is the heat flux at the streambed. $\phi_{i}^{aw}$ includes sensible (or conductive) heat, latent (or evaporative) heat, net solar (short-wave) radiation, atmospheric and emitted (long-wave) radiations. This term can conveniently be expressed as a function of the equilibrium temperature $T_i^{eq}$: $\phi_{i}^{aw} = B_i L_i (T_i^{eq} - T_i) k_i^j$, where $k_i^j$ [MT$^{-3}$Θ$^{-1}$] is a heat exchange coefficient. Hence, one has

$$\phi_{i}^{aw} \frac{\rho c_p V_i}{\rho c_p} = \frac{k_i}{D_i} (T_i^{eq} - T_i),$$

(7.5)

where $k_i = k_i^j / (\rho c_p)$ [LT$^{-1}$] (hereafter termed heat exchange velocity) represents the velocity of the heat exchange process between air and water. River depth can be linked to discharge
through Manning’s equation, which, owing to the above-mentioned assumptions of rectangular cross-section and large width, reads \( D_i = n_i^{0.6} B_i^{-0.6} s_i^{-0.3} Q_i^{0.6} \), where \( n_i \) is a Manning’s roughness coefficient, and \( s_i \) the average slope of the reach.

The streambed heat flux comprises the conductive heat flux with the groundwater and the friction dissipation. Both terms are in most cases negligible; however, the friction heat flux can be readily expressed via the already introduced variables without further parametrizations:

\[
\frac{\phi_{sb}^i}{\rho c_p V_i} = \frac{\rho g \Delta z_i \beta_{3,i} Q_i}{\rho c_p V_i} = \frac{g s_i L_i \beta_{3,i} Q_i}{c_p V_i},
\]

where \( g \) is the gravity acceleration; \( \Delta z_i = s_i L_i \) is the difference in elevation between the highest and lowest points of the reach; and \( \beta_{3,i} Q_i = (1 - 0.5 \beta_{2,i}) Q_i \) is the average discharge flowing through reach \( i \) at time \( \tau \). By inserting Eqs. (7.4), (7.5) and (7.6) into Eq. (7.3), one finally gathers

\[
\frac{dT_i}{d\tau} = \frac{Q_i}{V_i} \left[ \sum_{j=1}^{N_1} w_{ji} \beta_{1,j}(T_j - T_i) + \left( \beta_{2,i} + \frac{1}{Q_i} \frac{dV_j}{d\tau} \right)(T_L^i - T_i) + \beta_{3,i} \frac{g s_i L_i}{c_p} \right] + \frac{k_i}{D_i}(T_{eq}^i - T_i). \tag{7.7}
\]

Model (7.7) requires several types of input data. First, a morphological characterization of the catchment must be available. The extraction of the river network from a digital elevation model \[\text{Tarboton et al., 1991}\] readily allows the computation of coefficients \( \beta_{1,j,i}, \beta_{2,i}, \beta_{3,i} \) (which solely depend on the drainage area), the adjacency matrix \( W \), reach lengths and slopes. Reach widths can instead be estimated via in situ measurements, aerial images or, failing that, width-discharge relationships \[\text{Leopold and Maddock, 1953}\]. Second, a time series of discharge at a convenient cross-section of the river enables the derivation of water fluxes at all reaches thanks to the assumption of proportionality between \( Q \) and \( A \). Furthermore, time series of water depths can be derived via Manning’s equations\(^2\), allowing also to estimate the term \( dV_i/d\tau \). Third, other spatially heterogeneous environmental variables are required for the parametrization of \( T_{eq}^i \) and \( T_L^i \). In the following, equilibrium temperatures and temperatures of lateral water fluxes will be expressed as a function of air temperature and soil temperature, respectively.

### 7.2.3 Spatial interpolation of air temperatures

Key to any successful model of water temperatures is the full exploitation of the data available from field measurements. One issue is obviously the best use of spatial data for air temperature. Spatial interpolation of point-based air temperature data can be performed via a variety of methods, including kriging, inverse-distance weighting, Gaussian filters and 2-dimensional splines \[\text{Myers, 1994; Stahl et al., 2006}\]. However, these methods are typically conceived for flat and homogeneous terrains \[\text{Dodson and Marks, 1997}\], while in mountainous regions the

\(^2\)Similarly to Chapters 5 and 6, for this application a Manning’s roughness coefficient \( n = 0.033 \text{ m}^{-1/3}\text{s} \) was assumed for the whole river network.
relationship between elevation and air temperature must be accounted for prior to performing a spatial interpolation. Lapse rates (i.e. the rates at which air temperature changes following an increase in altitude) range between -9.8 °C km\(^{-1}\) (dry adiabatic lapse rate) and -4.0 °C km\(^{-1}\) (dry saturated lapse rate), with -6.5 °C km\(^{-1}\) being a typical value for the global mean environmental lapse rate [Barry and Chorley, 2009]. Especially in winter, the typical pattern of decreasing air temperature with increasing altitude can be altered by temperature inversions or cold air ponding [Csanady, 1974]. This phenomenon might hinder the achievement of great accuracies in air temperature prediction in small but topographically complex areas. As shown by Jarvis and Stuart [2001], empirical techniques of interpolation accounting for known autocorrelation in the temperature data can prove more efficient than the reliance on the selection of guiding variables or sophisticated spatial interpolation algorithms.

Stemming from these considerations, the approach here proposed does not specify an a priori value for the lapse rate, but rather evaluates air temperatures at the prediction point \(T^A_p\) as weighted averages of temperatures \(T^A_i\) recorded at \(N_a\) measuring stations:

\[
T^A_p = \frac{\sum_{i=1}^{N_a} W_{ip} T^A_i}{\sum_{i=1}^{N_a} W_{ip}},
\]

(7.8)

where

\[
W_{ip} = E_{ip}^{-\alpha_E} \Delta Z_{ip}^{-\alpha_Z}.
\]

(7.9)

Weights \(W_{ip}\) are function of both the Euclidean distance \(E_{ip}\) between the measuring site \(i\) and the prediction point \(p\) and their absolute difference in altitude \(\Delta Z_{ip} = |Z_i - Z_p|\). \(\alpha_E\) and \(\alpha_Z\) are positive parameters that were calibrated via Leave-One-Out-Cross-Validation [Hastie et al., 2001] on the daily air temperature time series measured in the period 2014-2017 at the eight MeteoSwiss stations presented in Fig. 7.1a. The prediction error was aggregated over all stations and expressed as root mean square error (RMSE). The best-fit parameters \(\alpha_E = 3\), \(\alpha_Z = 5\) attained RMSE = 1.05 °C. For comparison, the same calibration procedure was applied to spatially interpolated air temperatures obtained with the Neutral Stability Algorithm (NSA)\(^3\) proposed by Dodson and Marks [1997] (a method with specified lapse rate), obtaining RMSE = 2.42 °C. In particular, the NSA method showed a poor performance with regards to the higher-elevation stations A7 and A8 (see details in Fig. 7.3).

---

\(^3\)The Neutral Stability Algorithm converts air temperature data to sea-level potential temperatures (namely, the temperature attained by a fluid adiabatically brought to a pressure of 105 Pa), spatially interpolates sea-level potential temperatures, and then uses the inverse of the potential temperature function to retrieve air temperatures at the sought elevation. This algorithm requires two parameters: the lapse rate and the air temperature at sea level. RMSE values obtained for the NSA algorithm displayed in Fig. 7.3 were obtained by setting these two parameters to -0.0065 °C km\(^{-1}\) and 20 °C, respectively. Other values (within meaningful ranges) were tested but did not yield any improvement in the performance.
Chapter 7. A stream temperature model for ecohydrological applications

Figure 7.3 – Measured vs. modelled [via Eq.(7.9)] daily air temperatures for the eight sampling stations. For all stations, data span the period Jun 1, 2014 - May 3, 2017 (1069 days). For each station, the RMSE score is also reported. Numbers between round brackets indicate RMSE scores attained by the NSA model. The displayed quantiles are obtained through a binning procedure: each data series is sorted in ascending order and then divided in 20 equal parts; quantiles are then plotted at the midpoint of each bin.

7.2.4 Spatial interpolation of soil temperatures

The hydrological characteristics of a river, such as the presence of natural or artificial lakes or the role exerted by snowmelt, have a major impact in the regulation of its thermal regime [Piccolroaz et al., 2016]. In Eq. (7.7), such a type of information can be embedded by conveniently specifying the spatial pattern of lateral inflow temperatures $T_L^i$. As the Wigger is not regulated or substantially impacted by snowmelt, in this application the water temperature of the lateral flow $T_L^i$ is assumed to be equal to the soil temperature at 1-m depth $T_S^i$.

Soil temperature data were obtained for the set of MeteoSwiss stations shown in Fig. 7.1b. At the same stations, air temperatures are also recorded. Most time series of soil temperatures spanned periods comprised between years 1970 and 2000. Daily values of soil temperatures at 1-m depth were obtained by linear regression on moving averaged air temperatures. Calibration was performed by aggregating data over all stations and minimizing the overall root mean square error. The linear regression coefficients, as well as the lag for the moving average, were kept constant for all stations. The calibrated regression reads $T_S^i = 0.76 T_{A,MA}^i + 3.65 \degree C$, where $T_{A,MA}^i$ is the moving average of $T_A^i$ with a lag equal to 53 days. The corresponding RMSE is 1.24 $\degree$C. Fig. 7.4 displays the goodness of fit achieved by the regression model for each of the eight soil temperature stations.
7.2. Methods

Figure 7.4 – Measured vs. modelled daily soil temperatures at 1-m depth for the eight sampling stations. For each station, the elevation, the number (N) of data points and the RMSE score are also reported. The displayed quantiles are obtained through a binning procedure: each data series is sorted in ascending order and then divided in 20 equal parts; quantiles are then plotted at the midpoint of each bin.

7.2.5 Model parametrizations

The set of models here proposed (Table 7.1) is derived from Eq. (7.7) by properly specifying: i) further simplifying assumptions for the hydrothermal processes taking place in the reach; ii) a shape for the relationship between equilibrium temperature and air temperature; iii) a parametrization for the heat exchange velocity.

Model type. The full model of Eq. (7.7) is termed spatially-explicit (S). If the effect of water flow in the energy budget is negligible ($Q_i = 0$ and $dV_i/dt = 0$), only the last additive term of the r.h.s. of Eq. (7.7) is preserved, obtaining the local model (L):

$$\frac{dT_i}{dr} = k_i D_i (T_{eq}^i - T_i).$$  (7.10)

Note that, in the local model, the effect of hydrological conditions is still included in the water depth. Moreover, if one further assumes $D_i = 1$ m, one obtains the flat (F) model:

$$\frac{dT_i}{dr} = k_i (T_{eq}^i - T_i).$$  (7.11)

Finally, by further assuming $dT_i/dr = 0$, one gathers the regression (R) model $T_i = T_{eq}^i$. This approximation implies that the time scale for the adaptation of water temperature to the external forcing is much shorter than the time window of the observation, i.e. one day [Toffolon
Chapter 7. A stream temperature model for ecohydrological applications

<table>
<thead>
<tr>
<th>Model</th>
<th>( N_p )</th>
<th>Parameters</th>
<th>Model</th>
<th>( N_p )</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-lin-C</td>
<td>3</td>
<td>( a, b, k_0 )</td>
<td>F-lin-C</td>
<td>3</td>
<td>( a, b, k_0 )</td>
</tr>
<tr>
<td>S-lin-T</td>
<td>4</td>
<td>( a, b, k_0, k_1 )</td>
<td>F-lin-T</td>
<td>4</td>
<td>( a, b, k_0, k_1 )</td>
</tr>
<tr>
<td>S-day-C</td>
<td>4</td>
<td>( a, b, c, k_0 )</td>
<td>F-day-C</td>
<td>4</td>
<td>( a, b, c, k_0 )</td>
</tr>
<tr>
<td>S-day-T</td>
<td>5</td>
<td>( a, b, c, k_0, k_1 )</td>
<td>F-day-T</td>
<td>5</td>
<td>( a, b, c, k_0, k_1 )</td>
</tr>
<tr>
<td>S-sin-C</td>
<td>5</td>
<td>( a, b, c, \tau_{\text{max}}, k_0 )</td>
<td>F-sin-C</td>
<td>5</td>
<td>( a, b, c, \tau_{\text{max}}, k_0 )</td>
</tr>
<tr>
<td>S-sin-T</td>
<td>6</td>
<td>( a, b, c, \tau_{\text{max}}, k_0, k_1 )</td>
<td>F-sin-T</td>
<td>6</td>
<td>( a, b, c, \tau_{\text{max}}, k_0, k_1 )</td>
</tr>
<tr>
<td>S-log-C</td>
<td>5</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}}, k_0, k_1 )</td>
<td>F-log-C</td>
<td>5</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}}, k_0, k_1 )</td>
</tr>
<tr>
<td>S-log-T</td>
<td>6</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}}, k_0, k_1 )</td>
<td>F-log-T</td>
<td>6</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}}, k_0, k_1 )</td>
</tr>
<tr>
<td>L-lin-C</td>
<td>3</td>
<td>( a, b, k_0 )</td>
<td>R-lin</td>
<td>2</td>
<td>( a, b )</td>
</tr>
<tr>
<td>L-lin-T</td>
<td>4</td>
<td>( a, b, k_0, k_1 )</td>
<td>R-day</td>
<td>3</td>
<td>( a, b, c )</td>
</tr>
<tr>
<td>L-day-C</td>
<td>4</td>
<td>( a, b, c, k_0 )</td>
<td>R-sin</td>
<td>4</td>
<td>( a, b, c, \tau_{\text{max}} )</td>
</tr>
<tr>
<td>L-day-T</td>
<td>5</td>
<td>( a, b, c, k_0, k_1 )</td>
<td>R-log</td>
<td>4</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}} )</td>
</tr>
<tr>
<td>L-sin-C</td>
<td>5</td>
<td>( a, b, c, \tau_{\text{max}}, k_0 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-sin-T</td>
<td>6</td>
<td>( a, b, c, \tau_{\text{max}}, k_0, k_1 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-log-C</td>
<td>5</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}}, k_0, k_1 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-log-T</td>
<td>6</td>
<td>( T_{\text{low}}, \tau_{\text{up}}, \gamma, T_{\text{infl}}, k_0, k_1 )</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1 – Full list of tested models. \( N_p \): Number of parameters.

and Piccolroaz, 2015].

**Equilibrium temperature.** The relationship between \( T_i^{\text{eq}} \) and \( T_i^A \) can be linear (lin), as proposed by Caissie et al. [2005]:

\[
T_i^{\text{eq}}(\tau) = a T_i^A(\tau) + b
\]

(7.12)

where \( a \) and \( b \) are parameters requiring calibration. Following Bustillo et al. [2014], a second formulation (day) including the day-length factor \( DL \) is proposed:

\[
T_i^{\text{eq}}(\tau) = a T_i^A(\tau) + b + c DL,
\]

(7.13)

where \( c \) is an additional parameter, and

\[
DL(\tau) = \cos \left[ \frac{2\pi}{365} (j - 172) \right],
\]

(7.14)

with \( j \) being the ordinal day since January 1. As \( DL \) varies between -1 (when \( j \) corresponds to Dec 21) and +1 (for Jun 21, when \( j = 172 \)), this factor represents a proxy of incoming solar radiation, and thereby it is expected to take into account the non-uniqueness of the relationship \( T_i^{\text{eq}} = f(T_i^A) \), which is also influenced by solar radiation. A third formulation (sin), derived from Eqs. (7.13) and (7.14), is based on the fact that, while the amount of solar radiation reaching Earth peaks at the summer solstice in the Northern Hemisphere, the global solar radiation absorbed at the Earth surface tends to keep increasing into July (see data from station A1 in Fig. 7.5).
Figure 7.5 – Daily global solar radiation data from the Buchs-Aarau MeteoSwiss station (corresponding to station A1 in Fig. 7.1). The dataset spans the period 1985-2016. Percentiles were computed for each day of the year but leap days, which were discarded. Although the sinusoidal trend, fitted on the median, peaks on June 24, the distribution of median values tends to be slightly right-skewed, with higher monthly averaged values (see yellow lines) in July than in June and in August than in May. On the other hand, maximum values of solar radiation tend to a sinusoidal peaking at the summer solstice.

By adding a parameter $\tau_{\text{max}}$ identifying the ordinal day of yearly peak of absorbed solar radiation, the sin formulation of the relationship $T^e_i = f(T^A_i)$ reads:

$$T^e_i(\tau) = a T^A_i(\tau) + b + c \cos\left(\frac{2\pi}{365}(\tau - \tau_{\text{max}})\right).$$

(7.15)

A fourth type of link between equilibrium temperature and air temperature assumes a logistic (log) relationship (after Mohseni et al. [1998]):

$$T^e_i(\tau) = T_{\text{low}} + \frac{T_{\text{up}} - T_{\text{low}}}{1 + \exp[\gamma(T^\text{infl}_i - T^A_i(\tau))]},$$

(7.16)

where $T_{\text{low}}$, $T_{\text{up}}$ (lower and upper bounds for $T^e_i$, respectively), $T^\text{infl}_i$ and $\gamma$ (equilibrium temperature and slope at the inflection point, respectively), are parameters to be calibrated.

Heat exchange velocity. The simplest parametrization (C) consists in setting a spatially constant value for $k_i = k_0$. Bustillo et al. [2014] proposed an alternative formulation (T) where $k_i$ is expressed as a function of $T^A_i$, thereby introducing non-linearities in the heat exchange process between air and water even if the relationship $T^e_i = f(T^A_i)$ is linear. Unlike the above-mentioned authors, who proposed a linear relationship between $k_i$ and $T^A_i$, an exponential link is here proposed:

$$k_i(\tau) = k_0 \exp[k_1 T^A_i(\tau)].$$

(7.17)

Such a formulation bounds $k_i$ to be non-negative (when $k_0$ is bounded in $\mathbb{R}^+$) regardless of the value of $T^A_i(\tau)$, while the parameter $k_1$ is unbounded.
7.2.6 Model calibration and comparison

Mean daily measured water temperatures are available from Jun 28, 2014 to May 3, 2017 at sites #1 to #10, and from Sep 28, 2015 to Oct 4, 2016 at site #11. The lack of data after October 2016 for this site is due to the loss of both loggers after a large sediment transport event. In order to test the predictive performance of the models in both space and time, the subset of data points used for validation is made up of the full-length time series at sites #2, #6, #9, in addition to the last fifth of the time series at all other sites (i.e., from Oct 06, 2016 to May 3, 2017 at all remaining sites excluding #11, where data are missing in this period). Hence, the calibration dataset comprises 58% of the whole dataset. In order to lose memory of the initial condition, models are run starting from Jun 1, 2014 with an initial temperature of 10 °C at all stretches.

Model performance is expressed via root mean square error (RMSE_cal) on the aggregated calibration dataset. Calibration is performed via the Particle Swarm Optimization algorithm [Kennedy and Eberhart, 1995]. Best-fit models are then compared by means of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The root mean square error for the validation subset (RMSE_val) is also computed. Up to an additive constant that only depends on the number of observations \( N_{obs} \) and not on the model, AIC and BIC read, respectively:

\[
\text{AIC} = 2N_p + N_{obs}\log(RMSE_{cal}^2); \\
\text{BIC} = N_p\log(N_{obs}) + N_{obs}\log(RMSE_{cal}^2). 
\] (7.18) (7.19)

BIC contains a larger penalty term for the number of parameters \( N_p \) with respect to AIC. Because the absolute values of AIC and BIC are not significant, model comparison is actually performed through \( \Delta \text{AIC} \) (and \( \Delta \text{BIC} \)), i.e. the differences of AIC (BIC) scores with respect to the minimum AIC (BIC) achieved by the best-fit model.

7.3 Results and discussion

Tables 7.2 and 7.3 summarize calibration results for all models. According to both AIC and BIC criteria, the model that attains the best performance is S-sin-C, one of the only two models (the other being S-sin-T) achieving a root mean square error for the calibration period below 1 °C. Results for this model are detailed in Figs. 7.6 and 7.7. It is remarkable that model S-sin-C also achieves the best performance with regards to the validation period (lowest RMSE_val), which excludes overfitting issues. In general, the inclusion of the day-length factor (day model series) results in an enhanced fitting with respect to a linear (lin) parametrization for the equilibrium temperature. When the additional parameter \( \tau_{max} \) is included (sin model series), model performances are further improved. Conversely, the use of a logistic (log) parametrization

---

4The temperature sensor was replaced on Sep 6, 2017. All other stations are currently recording data without interruptions. In the perspective of a future publication, the analyses presented in this Chapter will be repeated on the extended datasets.
7.3. Results and discussion

Table 7.2 – Best-fit parameters and performance indices for all models except the log series.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>$\tau_{\text{max}}$</th>
<th>RMSE$_{\text{cal}}$</th>
<th>RMSE$_{\text{val}}$</th>
<th>$\Delta$AIC</th>
<th>$\Delta$BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-lin-C</td>
<td>0.64</td>
<td>3.94</td>
<td>0.663</td>
<td>-</td>
<td>-</td>
<td>1.102</td>
<td>1.174</td>
<td>1322</td>
<td>1309</td>
<td></td>
</tr>
<tr>
<td>S-lin-T</td>
<td>0.64</td>
<td>3.94</td>
<td>0.670</td>
<td>-0.001</td>
<td>-</td>
<td>1.102</td>
<td>1.174</td>
<td>1324</td>
<td>1318</td>
<td></td>
</tr>
<tr>
<td>S-day-C</td>
<td>0.52</td>
<td>5.30</td>
<td>0.597</td>
<td>-1.75</td>
<td>-</td>
<td>1.005</td>
<td>1.121</td>
<td>182</td>
<td>176</td>
<td></td>
</tr>
<tr>
<td>S-day-T</td>
<td>0.51</td>
<td>5.41</td>
<td>0.643</td>
<td>-0.007</td>
<td>1.80</td>
<td>1.005</td>
<td>1.121</td>
<td>181</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>S-sin-C</td>
<td>0.44</td>
<td>6.04</td>
<td>0.934</td>
<td>2.26</td>
<td>Jul14</td>
<td>0.990</td>
<td>1.106</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>S-sin-T</td>
<td>0.48</td>
<td>5.66</td>
<td>0.754</td>
<td>-0.007</td>
<td>1.97</td>
<td>0.999</td>
<td>1.110</td>
<td>152</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>L-lin-C</td>
<td>0.64</td>
<td>4.09</td>
<td>0.046</td>
<td>-</td>
<td>-</td>
<td>1.333</td>
<td>1.358</td>
<td>3687</td>
<td>3674</td>
<td></td>
</tr>
<tr>
<td>L-lin-T</td>
<td>0.65</td>
<td>4.12</td>
<td>0.060</td>
<td>-0.013</td>
<td>-</td>
<td>1.330</td>
<td>1.360</td>
<td>3658</td>
<td>3652</td>
<td></td>
</tr>
<tr>
<td>L-day-C</td>
<td>0.60</td>
<td>4.51</td>
<td>0.041</td>
<td>-0.59</td>
<td>-</td>
<td>1.313</td>
<td>1.358</td>
<td>3505</td>
<td>3499</td>
<td></td>
</tr>
<tr>
<td>L-day-T</td>
<td>0.60</td>
<td>4.62</td>
<td>0.067</td>
<td>-0.038</td>
<td>0.63</td>
<td>1.314</td>
<td>1.372</td>
<td>3515</td>
<td>3516</td>
<td></td>
</tr>
<tr>
<td>L-sin-C</td>
<td>0.36</td>
<td>6.84</td>
<td>1.476</td>
<td>-2.90</td>
<td>Aug 1</td>
<td>1.072</td>
<td>1.179</td>
<td>992</td>
<td>992</td>
<td></td>
</tr>
<tr>
<td>L-sin-T</td>
<td>0.36</td>
<td>6.88</td>
<td>3.750</td>
<td>-0.011</td>
<td>2.92</td>
<td>1.073</td>
<td>1.181</td>
<td>996</td>
<td>1003</td>
<td></td>
</tr>
<tr>
<td>F-lin-C</td>
<td>0.66</td>
<td>3.95</td>
<td>0.259</td>
<td>-</td>
<td>-</td>
<td>1.284</td>
<td>1.321</td>
<td>3215</td>
<td>3203</td>
<td></td>
</tr>
<tr>
<td>F-lin-T</td>
<td>0.66</td>
<td>3.95</td>
<td>0.259</td>
<td>0</td>
<td>-</td>
<td>1.284</td>
<td>1.323</td>
<td>3219</td>
<td>3213</td>
<td></td>
</tr>
<tr>
<td>F-day-C</td>
<td>0.61</td>
<td>4.44</td>
<td>0.216</td>
<td>-0.56</td>
<td>-</td>
<td>1.271</td>
<td>1.327</td>
<td>3100</td>
<td>3094</td>
<td></td>
</tr>
<tr>
<td>F-day-T</td>
<td>0.63</td>
<td>4.24</td>
<td>0.216</td>
<td>-0.001</td>
<td>-</td>
<td>1.271</td>
<td>1.320</td>
<td>3100</td>
<td>3101</td>
<td></td>
</tr>
<tr>
<td>F-sin-C</td>
<td>0.37</td>
<td>6.71</td>
<td>2.500</td>
<td>-2.78</td>
<td>Aug 1</td>
<td>1.067</td>
<td>1.166</td>
<td>922</td>
<td>923</td>
<td></td>
</tr>
<tr>
<td>F-sin-T</td>
<td>0.37</td>
<td>6.75</td>
<td>4.046</td>
<td>-0.027</td>
<td>2.79</td>
<td>1.067</td>
<td>1.170</td>
<td>928</td>
<td>935</td>
<td></td>
</tr>
<tr>
<td>R-lin</td>
<td>0.62</td>
<td>4.43</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.403</td>
<td>1.476</td>
<td>4316</td>
<td>4296</td>
<td></td>
</tr>
<tr>
<td>R-day</td>
<td>0.58</td>
<td>4.81</td>
<td>-</td>
<td>0.53</td>
<td>-</td>
<td>1.382</td>
<td>1.463</td>
<td>4135</td>
<td>4122</td>
<td></td>
</tr>
<tr>
<td>R-sin</td>
<td>0.38</td>
<td>6.70</td>
<td>-</td>
<td>2.78</td>
<td>Aug 4</td>
<td>1.001</td>
<td>1.128</td>
<td>132</td>
<td>126</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.3 – Best-fit parameters and performance indices for the log model series.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model</th>
<th>$T_{\text{low}}$</th>
<th>$T_{\text{up}}$</th>
<th>$T_{\text{infl}}$</th>
<th>$\gamma$</th>
<th>$k_0$</th>
<th>$k_1$</th>
<th>RMSE$_{\text{cal}}$</th>
<th>RMSE$_{\text{val}}$</th>
<th>$\Delta$AIC</th>
<th>$\Delta$BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-log-C</td>
<td>-1.50</td>
<td>22.00</td>
<td>9.70</td>
<td>0.124</td>
<td>0.734</td>
<td>-</td>
<td>1.086</td>
<td>1.160</td>
<td>1030</td>
<td>1064</td>
<td></td>
</tr>
<tr>
<td>S-log-T</td>
<td>-1.25</td>
<td>24.90</td>
<td>12.20</td>
<td>0.109</td>
<td>0.865</td>
<td>-0.020</td>
<td>1.092</td>
<td>1.173</td>
<td>1099</td>
<td>1139</td>
<td></td>
</tr>
<tr>
<td>L-log-C</td>
<td>-3.00</td>
<td>24.80</td>
<td>10.60</td>
<td>0.102</td>
<td>0.058</td>
<td>-</td>
<td>1.300</td>
<td>1.338</td>
<td>3269</td>
<td>3303</td>
<td></td>
</tr>
<tr>
<td>L-log-T</td>
<td>-0.73</td>
<td>22.50</td>
<td>11.60</td>
<td>0.139</td>
<td>0.086</td>
<td>-0.028</td>
<td>1.298</td>
<td>1.391</td>
<td>3244</td>
<td>3285</td>
<td></td>
</tr>
<tr>
<td>F-log-C</td>
<td>-2.12</td>
<td>24.20</td>
<td>10.80</td>
<td>0.109</td>
<td>0.310</td>
<td>-</td>
<td>1.260</td>
<td>1.326</td>
<td>2873</td>
<td>2907</td>
<td></td>
</tr>
<tr>
<td>F-log-T</td>
<td>-2.94</td>
<td>22.00</td>
<td>8.51</td>
<td>0.119</td>
<td>0.243</td>
<td>0.022</td>
<td>1.265</td>
<td>1.316</td>
<td>2927</td>
<td>2968</td>
<td></td>
</tr>
<tr>
<td>R-log</td>
<td>-3.00</td>
<td>27.50</td>
<td>12.70</td>
<td>0.089</td>
<td>-</td>
<td>-</td>
<td>1.373</td>
<td>1.459</td>
<td>3940</td>
<td>3967</td>
<td></td>
</tr>
</tbody>
</table>

for the link between $T^{eq}$ and $T^A$ does not yield better results, despite the increased number of parameters. For all models of the sin series, the calibrated values of $\tau_{\text{max}}$ range between July and August, thereby confirming the hypothesis that the equilibrium temperature is more influenced by absorbed solar radiation than by incoming solar radiation.

As for the model type, all spatially-explicit (S) models perform substantially better than local (L) or flat (F) models. Best-fit values for the baseline heat exchange velocity $k_0$ for the spatially-explicit models are higher than those for L and F models (excluding models of the sin series), and therefore these models are more capable of following day-to-day variations in the
measured signal (see Fig. 7.8). Conversely, local models tend to produce temperature fluctuations that are too damped at downstream sections (where river depths are generally higher) and too spiky at upstream sections. As a result, F models, which neglect the effect of river depth on the heat exchange process, perform slightly better than L models. Spatially-explicit models also include river depth at the denominator of the term accounting for air-water heat exchanges, but the impact of this variable is here mitigated by the inclusion of water fluxes from upstream sections and hillslopes. Regression (R) models perform quite poorly compared to the others, with the exception of model R-$\sin$, which ranks as the second best model (Tables 7.2 and 7.3) although with a $\Delta$AIC of 132 (a $\Delta$AIC of 2-4 is typically assumed to select a model over another one [Burnham and Anderson, 2004]). As shown in Fig. 7.9, all model types of the $\sin$ series perform best when discharge is close to its median value. Model R-$\sin$ overcomes the others for high flows, while its performance resembles that of the less efficient L and F models when discharge is low. In this latter case, the spatially-explicit model proves more reliable in capturing water temperature dynamics. Presumably, this is due to the fact that the chosen simplifications for the hydrological model (i.e. proportionality between discharge and contributing area, rectangular river cross-sections with large widths and uniform
7.3. Results and discussion

![Figure 7.7 – Relationship between modelled (S-sin-C) and observed temperature time series in 2016 for all sampling sites. Residuals are calculated as the difference between modelled and observed values.]

![Figure 7.8 – Comparison among results from models S-lin-C, L-lin-C, F-lin-C in 2016 for sites #1 and #11.]

Flow conditions as postulated by Manning’s equation) do not hold during discharge peaks. A more sophisticated hydrological model would likely attain better performances during large flow events, with the drawback that additional spatially distributed datasets would then be required.

The parametrization of the heat exchange velocity as a function of air temperature (T model series) did not produce any improvement in the fitting for any of the models tested, as compared to the assumption of a constant value (C). Indeed, the best-fit values found for the
Chapter 7. A stream temperature model for ecohydrological applications

Table 7.4 – Root mean square errors (in °C) evaluated at the 11 sampling sites for the calibrated C-series models. These RMSE values are computed for the whole available data time series at each station (i.e. without distinguishing between calibration and validation periods).

<table>
<thead>
<tr>
<th>Model</th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
<th>#9</th>
<th>#10</th>
<th>#11</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-lin-C</td>
<td>1.01</td>
<td>1.00</td>
<td>0.91</td>
<td>0.91</td>
<td>1.27</td>
<td>0.86</td>
<td>1.16</td>
<td>1.23</td>
<td>1.54</td>
<td>1.18</td>
<td>1.39</td>
</tr>
<tr>
<td>S-day-C</td>
<td>0.93</td>
<td>0.93</td>
<td>0.85</td>
<td>0.83</td>
<td>1.13</td>
<td>0.77</td>
<td>1.11</td>
<td>1.23</td>
<td>1.53</td>
<td>1.03</td>
<td>1.05</td>
</tr>
<tr>
<td>S-sin-C</td>
<td>0.91</td>
<td>0.90</td>
<td>0.83</td>
<td>0.77</td>
<td>1.17</td>
<td>0.75</td>
<td>1.06</td>
<td>1.25</td>
<td>1.51</td>
<td>1.03</td>
<td>0.97</td>
</tr>
<tr>
<td>S-log-C</td>
<td>0.99</td>
<td>0.97</td>
<td>0.90</td>
<td>0.85</td>
<td>1.25</td>
<td>0.84</td>
<td>1.12</td>
<td>1.24</td>
<td>1.53</td>
<td>1.17</td>
<td>1.39</td>
</tr>
<tr>
<td>L-lin-C</td>
<td>1.33</td>
<td>1.29</td>
<td>1.16</td>
<td>1.05</td>
<td>1.57</td>
<td>1.07</td>
<td>1.27</td>
<td>1.52</td>
<td>1.65</td>
<td>1.34</td>
<td>1.61</td>
</tr>
<tr>
<td>L-day-C</td>
<td>1.30</td>
<td>1.26</td>
<td>1.17</td>
<td>1.05</td>
<td>1.53</td>
<td>1.08</td>
<td>1.27</td>
<td>1.57</td>
<td>1.67</td>
<td>1.28</td>
<td>1.47</td>
</tr>
<tr>
<td>L-sin-C</td>
<td>0.99</td>
<td>0.94</td>
<td>0.89</td>
<td>0.76</td>
<td>1.36</td>
<td>0.84</td>
<td>1.05</td>
<td>1.36</td>
<td>1.54</td>
<td>1.15</td>
<td>1.26</td>
</tr>
<tr>
<td>L-log-C</td>
<td>1.26</td>
<td>1.22</td>
<td>1.14</td>
<td>1.03</td>
<td>1.55</td>
<td>1.06</td>
<td>1.26</td>
<td>1.53</td>
<td>1.64</td>
<td>1.28</td>
<td>1.61</td>
</tr>
<tr>
<td>F-lin-C</td>
<td>1.20</td>
<td>1.16</td>
<td>1.15</td>
<td>1.06</td>
<td>1.51</td>
<td>1.06</td>
<td>1.25</td>
<td>1.51</td>
<td>1.64</td>
<td>1.28</td>
<td>1.37</td>
</tr>
<tr>
<td>F-day-C</td>
<td>1.20</td>
<td>1.17</td>
<td>1.16</td>
<td>1.05</td>
<td>1.47</td>
<td>1.05</td>
<td>1.24</td>
<td>1.55</td>
<td>1.67</td>
<td>1.26</td>
<td>1.18</td>
</tr>
<tr>
<td>F-sin-C</td>
<td>0.98</td>
<td>0.93</td>
<td>0.88</td>
<td>0.76</td>
<td>1.35</td>
<td>0.83</td>
<td>1.04</td>
<td>1.35</td>
<td>1.52</td>
<td>1.13</td>
<td>1.25</td>
</tr>
<tr>
<td>F-log-C</td>
<td>1.19</td>
<td>1.14</td>
<td>1.12</td>
<td>1.02</td>
<td>1.50</td>
<td>1.05</td>
<td>1.24</td>
<td>1.51</td>
<td>1.64</td>
<td>1.28</td>
<td>1.33</td>
</tr>
<tr>
<td>R-lin</td>
<td>1.20</td>
<td>1.18</td>
<td>1.08</td>
<td>1.23</td>
<td>1.74</td>
<td>1.29</td>
<td>1.41</td>
<td>1.60</td>
<td>1.77</td>
<td>1.50</td>
<td>1.80</td>
</tr>
<tr>
<td>R-day</td>
<td>1.19</td>
<td>1.18</td>
<td>1.08</td>
<td>1.21</td>
<td>1.72</td>
<td>1.29</td>
<td>1.38</td>
<td>1.61</td>
<td>1.77</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>R-sin</td>
<td>0.92</td>
<td>0.87</td>
<td>0.74</td>
<td>0.70</td>
<td>1.30</td>
<td>0.78</td>
<td>1.00</td>
<td>1.30</td>
<td>1.50</td>
<td>1.07</td>
<td>1.26</td>
</tr>
<tr>
<td>R-log</td>
<td>1.17</td>
<td>1.15</td>
<td>1.07</td>
<td>1.19</td>
<td>1.69</td>
<td>1.27</td>
<td>1.41</td>
<td>1.60</td>
<td>1.77</td>
<td>1.47</td>
<td>1.72</td>
</tr>
</tbody>
</table>

Table 7.4 reports RMSE values for the full-length time series at the various site. Site #9, which is the second site after #11 in terms of elevation, always shows the poorest performance regardless of model parametrization; in particular, as displayed in Fig. 7.6 with regards to model S-sin-C, the model tends to overestimate negative peaks and, to a lesser extent, underestimate positive peaks. However, it is worth recalling that the temperatures measured in this site were only used in validation. On the other hand, the RMSE values obtained for site #6, which was also used in validation, are generally the lowest among all sites. Moreover, RMSE values for the third site (#2) used in validation are also rather low (below 1 °C for all models of the sin series). This suggests that the prediction capabilities of this suite of models are particularly remarkable with regards to the flatter portion of the Wigger catchment, where indeed (see Fig. 7.1a) most of the sampling sites where located.
In summary, this Chapter develops a spatially-explicit, deterministic water temperature model able to evaluate within-river temperature gradients and therefore suitable for ecohydrological applications. Such a model enables the derivation of reliable estimates of stream temperatures for the whole catchment based on a limited number of conveniently located (namely, spanning the largest possible elevation range) measuring stations. Results showed how considering water and energy budgets at a reach scale considerably improves model performances with respect to local models, allowing to trust regionalized values for the regression coefficient of the relationship between equilibrium temperature and air temperature. The introduction of a sinusoidal proxy for the effect of absorbed solar radiation in the evaluation of the equilibrium temperature constitutes a major enhancement of model performances. At small time scales as in the case here explored, the correct evaluation of air temperature fields across the region is crucial to the production of reliable stream temperature estimates. The sole use of a lapse rate would not take into account the phenomenon of temperature inversion, which may heavily characterize the climate in the portions of the watershed at higher elevation. Therefore, distributed air temperature measuring stations are needed; furthermore, the spatial interpolation approach should be as much close to the data as possible, rather than theoretically robust.

Among the foremost limitations of the approach, however suitable for the case study at hand, it must be noted that the model needs to be tested on other catchments, possibly with higher elevation gradients and/or characterized by snowmelt or presence of lakes. Given the relevance of water temperatures to mortality and incidence patterns of PKD, the model presented in this Chapter represents a fundamental tool to perform ecohydrological studies of PKD spread in rivers.

### 7.4 List of symbols

The following table lists all symbols used in this Chapter. For the sake of clarity and completeness, symbols already defined in previous Chapters are here repeated.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>Slope of the linear relationship $T_{eq}^i = f(T_A^i)$</td>
<td>-</td>
</tr>
<tr>
<td>$A_i$</td>
<td>Upstream area at stretch $i$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$A_i^L$</td>
<td>Directly contributing area at stretch $i$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>$b$</td>
<td>Intercept of the linear relationship $T_{eq}^i = f(T_A^i)$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$B_i$</td>
<td>River width of stretch $i$</td>
<td>$L$</td>
</tr>
<tr>
<td>$c$</td>
<td>Mid-amplitude of the sinusoidal term in the relationship $T_{eq}^i = f(T_A^i)$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$c_p$</td>
<td>Specific heat of water at constant pressure</td>
<td>$L^2T^{-2}\Theta^{-1}$</td>
</tr>
<tr>
<td>$D_i(\tau)$</td>
<td>Water depth of stretch $i$</td>
<td>$L$</td>
</tr>
<tr>
<td>$DL(\tau)$</td>
<td>Day-length sinusoidal term in the relationship $T_{eq}^i = f(T_A^i)$</td>
<td>-</td>
</tr>
<tr>
<td>$E_{ip}$</td>
<td>Eucledian distance between stations $i$ and $p$</td>
<td>$L$</td>
</tr>
<tr>
<td>$H_i(\tau)$</td>
<td>Thermal energy of stretch $i$</td>
<td>$ML^2T^{-2}$</td>
</tr>
</tbody>
</table>
# Chapter 7. A stream temperature model for ecohydrological applications

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_i(\tau)$</td>
<td>Heat exchange velocity at stretch $i$</td>
<td>LT$^{-1}$</td>
</tr>
<tr>
<td>$k_i'(\tau)$</td>
<td>Heat exchange coefficient at stretch $i$</td>
<td>MT$^3\Theta^{-1}$</td>
</tr>
<tr>
<td>$k_0$</td>
<td>Baseline value for $k_i$</td>
<td>LT$^{-1}$</td>
</tr>
<tr>
<td>$k_1$</td>
<td>Strength of dependency on $T_{A}^i$ for $k_i$</td>
<td>$\Theta^{-1}$</td>
</tr>
<tr>
<td>$L_i$</td>
<td>Length of stretch $i$</td>
<td>L</td>
</tr>
<tr>
<td>$n$</td>
<td>Manning’s roughness coefficient at stretch $i$</td>
<td>TL$^{1/3}$</td>
</tr>
<tr>
<td>$N_A$</td>
<td>Number of data points used for air temperature interpolation</td>
<td>-</td>
</tr>
<tr>
<td>$N_{obs}$</td>
<td>Total number of observed values of mean daily stream temperature</td>
<td>-</td>
</tr>
<tr>
<td>$N_p$</td>
<td>Number of parameters</td>
<td>-</td>
</tr>
<tr>
<td>$N_s$</td>
<td>Number of river stretches</td>
<td>-</td>
</tr>
<tr>
<td>$Q_i(\tau)$</td>
<td>Water discharge at stretch $i$</td>
<td>L$^3T^{-1}$</td>
</tr>
<tr>
<td>$Q_{L}^i(\tau)$</td>
<td>Directly contributing discharge at stretch $i$</td>
<td>L$^3T^{-1}$</td>
</tr>
<tr>
<td>RMSE$\text{cal}$</td>
<td>RMSE score in calibration</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>RMSE$\text{val}$</td>
<td>RMSE score in validation</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$s_i$</td>
<td>Mean slope of stretch $i$</td>
<td>-</td>
</tr>
<tr>
<td>$T_i(\tau)$</td>
<td>Stream temperature at stretch $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{A}^i(\tau)$</td>
<td>Air temperature at the subcatchment $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{A,M}^i(\tau)$</td>
<td>Moving average of $T_{A}^i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{eq}^i(\tau)$</td>
<td>Equilibrium temperature at the subcatchment $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{inf}$</td>
<td>Temperature at the inflection point for the logistic relationship $T_{i}^{eq} = f(T_{A}^i)$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{L}^i(\tau)$</td>
<td>Temperature of the lateral inflow at the subcatchment $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{low}$</td>
<td>Lower bound of the logistic relationship $T_{i}^{eq} = f(T_{A}^i)$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{S}^i(\tau)$</td>
<td>Soil temperature (at 1-m depth) at the subcatchment $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$T_{up}$</td>
<td>Upper bound of the logistic relationship $T_{i}^{eq} = f(T_{A}^i)$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$V_i(\tau)$</td>
<td>Water volume of stretch $i$</td>
<td>L$^3$</td>
</tr>
<tr>
<td>$w_{ij}$</td>
<td>Entry of the adjacency matrix $W$</td>
<td>-</td>
</tr>
<tr>
<td>$W_{ip}$</td>
<td>Contribution of air temperature measured at station $i$ to the inferred air temperature at point $p$</td>
<td>L$^{-2\alpha_E-\alpha_Z}$</td>
</tr>
<tr>
<td>$Z_{p}$</td>
<td>Elevation of point $p$</td>
<td>L</td>
</tr>
<tr>
<td>$\alpha_E$</td>
<td>Exponent for the dependency on $E_{ip}$ in $W_{ip}$</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_Z$</td>
<td>Exponent for the dependency on $\Delta Z_{ip}$ in $W_{ip}$</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{1,i,j}$</td>
<td>Geometric coefficient expressing the ratio $Q_j/Q_i$</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{2,i}$</td>
<td>Geometric coefficient expressing the ratio $Q_{L}^j/Q_i$</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_{3,i}$</td>
<td>Geometric coefficient equal to $1 - 0.5\beta_{2,i}$</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Slope at the inflection point for the logistic relationship $T_{i}^{eq} = f(T_{A}^i)$</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta z_i$</td>
<td>Elevation gain along stretch $i$</td>
<td>L</td>
</tr>
<tr>
<td>$\Delta Z_{ip}$</td>
<td>Absolute difference in elevation between points $i$ and $p$</td>
<td>L</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Dimension</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>$\phi_i(\tau)$</td>
<td>Heat flux associated to $Q_i(\tau)$</td>
<td>$ML^2T^{-3}$</td>
</tr>
<tr>
<td>$\phi_i^d(\tau)$</td>
<td>Heat flux associated to $Q_i^d(\tau)$</td>
<td>$ML^2T^{-3}$</td>
</tr>
<tr>
<td>$\phi_{i,aw}(\tau)$</td>
<td>Heat flux exchanged at the air-water interface of stretch $i$</td>
<td>$ML^2T^{-3}$</td>
</tr>
<tr>
<td>$\phi_{i,na}(\tau)$</td>
<td>Non-advective heat flux at stretch $i$</td>
<td>$ML^2T^{-3}$</td>
</tr>
<tr>
<td>$\phi_{i,bb}(\tau)$</td>
<td>Heat flux exchanged at the streambed of stretch $i$</td>
<td>$ML^2T^{-3}$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Water density</td>
<td>$ML^{-3}$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Time</td>
<td>$T$</td>
</tr>
<tr>
<td>$\tau_{max}$</td>
<td>Day of positive peak in the sinusoidal term of $T_i^{eq} = f(T_i^A)$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

Water density $\rho$: $ML^{-3}$

Time $\tau$: $T$
Conclusions and perspectives

Proliferative kidney disease constitutes a serious threat for stocked and wild salmonid populations in Europe and North America. The ecohydrological and metacommunity studies presented in this Thesis, all joined by a clear conceptual thread, show that PKD, owing to its strong influence in the regulation of the fish population size, can potentially create catchment-wide stock reductions. As high PKD transmission and mortality rates are associated to warmer water temperatures, climate change is feared to exacerbate the risk of local extinction for fish populations, even iconic ones, in the forthcoming years. Some characteristics of the life-cycle of its myxozoan causative agent *Tetracapsuloides bryosalmonae* make this disease an enduring and resisting menace for riverine wildlife: in particular, the parasite can persist in its primary host, the bryozoan *Fredericella sultana*, across unfavorable conditions and winters under the form of statoblasts, asexually produced propagules. Moreover, the apparent dearth of infection clearing mechanisms in salmonids (and, to a lesser extent, in bryozoans) poses further challenges to disease eradication.

In the light of the above considerations, an appraisal of catchment-scale risk of PKD infection allowing predictions under environmental change scenarios cannot forgo a comprehensive modelling approach, which includes disease transmission dynamics, the ecology of bryozoan and salmonid populations, the role of the hydrological and thermal drivers and the effect of river network connectivity. In this context, not only does modelling benefit from field and laboratory studies inasmuch as it can merge all information under a common framework, but it also serves as a guiding tool helping the design of further studies by highlighting the weaknesses of the current knowledge. This Thesis aimed at bridging the diverse but entwined fields of ecology, epidemiology, hydrology and mathematical modelling in order to produce an integrated study of PKD, in the perspective of grasping the main factors allowing disease persistence and spread, and possibly devising mitigation strategies.

A first local model of PKD was developed in a bid to dig out the basic mechanisms involved in disease spread and its long-term persistence. Possible ranges for many of the epidemiological parameters were pointed out through a literature review, while the effects of other parameters in determining parasite invasibility and seasonal population decay were assessed via sensitivity analyses. Owing to a stability analysis based on Floquet theory, PKD was shown to be a disease with high potential of invasion in disease-free communities; this is especially due
Conclusions and perspectives

to the fact that recovered fish are able to shed infective spores for long periods after the first infection. This epidemiological phase, here termed the carrier state, can potentially last until fish death, as the existence of an infection-clearing dynamics is still uncertain. Some recent analyses (see Appendix A) suggested that this state can last for at least one year, a time scale much longer than that of all other epidemiological dynamics. Although to a lesser extent, the absence of complete recovery dynamics in bryozoans was also shown to enhance the invasibility of *T. bryosalmonae*. On the other hand, when the disease is endemic, the size of the seasonal decay in the fish population size is mostly controlled by disease transmission rates, which are known to be temperature-dependent. Consequently, even small increases in water temperature can lead to large population reductions, as observed in Swiss rivers [Wahli et al., 2002; Burkhardt-Holm et al., 2005].

By including the processes of hydrodynamic dispersion of parasite spores and fish mobility, the local model was made spatially explicit. Simulation experiments run on synthetic river networks enabled the assessment of network effects on disease propagation celerity and on spatial patterns of PKD prevalence and population decay, as well as the effects of spatial heterogeneity of bryozoan populations. Due to hydrological transport of *T. bryosalmonae*, both prevalence and population size decay naturally tend to increase in the downstream direction, while high fish mobility promotes the mixing of populations between headwater and main-course reaches. As observed with regard to the local model, climate change yields severe population reductions, while its effect on the distribution of prevalence is minor. This is mainly due to the fact that spore shedding rates were not assumed to be temperature-dependent, as recently supported by the findings of Strepparava et al. [2017]. Another observed effect of the increase in water temperature is the enhancement of disease propagation celerity in both downstream and upstream directions. Downstream propagation is generally fast, owing to the passive transport of parasite spores, and consists of three main phases: first, the propagation front moves downstream; second, the disease invades the main course of the river, where prevalence increases, possibly attaining 100%; third a second propagation front moves in the upstream direction increasing the prevalence of the headwater sites until a steady state is reached, typically after some years from the first introduction of the parasite. Moreover, it was showed that infected bryozoan colonies located in restrained river areas are able to sustain the infection in the downstream reaches and, to a lesser extent, in the upstream direction. This last aspect prompts great concern from the point of view of control strategies, as it questions the effectiveness of bryozoan eradication campaigns.

The spatial model of PKD epidemiology, completed with the inclusion of the age structure of the fish population and the spawning migration movements, was then validated by means of the application to a case study catchment (the Wigger), where an extensive three-year long field campaign allowed to gather spatially distributed data on fish abundance and prevalence, eDNA concentrations of *F. sultana* in water samples, and water temperatures. The model showed good performances in reproducing observed spatial and temporal patterns of prevalence. In addition, it was demonstrated that the previously highlighted higher dangerousness constituted by PKD for young-of-the-year fish (Y0Y) with respect to adult individuals is not
Conclusions and perspectives

necessarily due to an intrinsic higher mortality rate, but could be simply explained by a lack of acquired immunity. As a matter of fact, prevalence in YOY tends to be lower than in adults (since vertical transmission of PKD in fish is inhibited) but seasonal population decay in YOY is around three times higher. Observed prevalences were generally high, and in most cases equal to 100%, save for one sampling in a stretch for which the estimated bryozoan density in its upstream reaches was very low, and where sampled prevalence was null. This observation constitutes an important validation of both the epidemiological model and the model used to estimate bryozoan suitability based on eDNA samples.

Overall, the analyses here presented underline the severe danger constituted by PKD to stocked and wild salmonid fish populations. However, the key roles of temperature and the invertebrate host distribution suggest possible avenues for management solutions. As a reservoir host of *T. bryosalmonae*, bryozoans retain the parasite in the ecosystem even in the case of severe virulence in the fish populations. Therefore, frequent and repeated fish stocking is unlikely to result in improvement in fish population health. However, the obtained results hint that approaches promoting local breeding success of fish (such as management of flow conditions and water temperatures to prevent severe mortality events) and natural selection for PKD-resistant fish strains may potentially relieve the long-term impacts of PKD, without any directed intervention on the bryozoan populations. Research along this direction is currently underway (T. Wahli, pers. comm.). Although it seems that the eradication of the bryozoans on the whole-river scale is hardly possible or advisable, the spatial PKD model identifies key areas where habitat management, river restoration and directed management of the bryozoan populations may relieve PKD status in fish. However, the impact of such interventions on the spatial patterns of PKD occurrence may be difficult to predict or measure. Indeed, local conditions may be highly relevant for controlling the dispersal and long-term infectivity of spores released from bryozoans, and, for instance, high frequency of local infection events may increase severity of disease locally but reduce the number of infective spores available at larger, regional scales.

The knowledge obtained from the proposed framework allows to enlighten fundamental research directions for future improvement of the understanding of PKD emergence and spread. Chiefly, the analyses presented in Chapter 5 demonstrate that knowledge on the density pattern of bryozoan colonies in river basins is pivotal to predicting the spatial distribution of disease prevalence in fish. The strong correlation found between bryozoan abundance and the presence of moraines upstream for the case-study river should be taken with some degree of reservation, as it definitely needs further validation in other catchments with different hydro-morphological and geological characteristics. To this end, the eDNA transport model developed in this context represents a promising tool to discriminate key environmental variables governing the presence of bryozoans. The degree of complexity of such a model can readily be augmented by including additional covariates (in particular water quality pa-

---

1 Note that, while the species distribution model presented in Chapter 6 allows the identification of the river stretches associated to highest density of bryozoans, the actual localization of bryozoan colonies within a single stretch would require extensive manual search along the river banks.
rameters such as dissolved oxygen content and pollution indicators or their proxies, e.g. land use and location of urban settlements) or more physically-based decay assumptions (e.g. the adoption of decay time instead of decay length as unknown parameter – see Chapter 6). From a theoretical or modelling perspective, other interesting paths to explore would consist in performing a full stability analysis of the spatial epidemiological model in a bid to unravel the major factors driving parasite invasion in river networks, or in simulating the effect of bryozoan eradication from key hot-spots in the diminution of disease prevalence in fish. Furthermore, it is worth mentioning some caveats of the current formulation of the local PKD model here proposed, whose inclusion would increase the accuracy of the modelling framework. First, while the model only considers disease transmission dynamics during the warm season, it has been shown [Gay et al., 2001] that fish can actually develop PKD when they are exposed to *T. bryosalmonae* in rivers during winter. Second, the encounter rates between the bryozoan or fish host and the parasite spores were here expressed via the mass-action assumption, i.e. as the product between the local (at stretch scale) predator and prey densities, as if the system was well-mixed. This hypothesis might be questionable, as many bryozoan populations occur on submerged tree roots and are thus located where fish typically take refuge, thereby enhancing transmission with respect to the well-mixed case. To overcome this issue, one could formulate predator-prey dynamics by means of spatial moment equations [Barraquand and Murrell, 2013]. Third, the current formulation of the model only allows the estimation of macroscopic indicators of PKD impact, such as prevalence or population decay. In order to assess dose-dependent effects on transmission and virulence in fish, a state variable identifying the mean parasite load in the fish kidney should be added to the model. Such an approach is borrowed from models of schistosomiasis, where the mean worm burden in the host population is accounted for [Macdonald, 1965; Gurarie and Seto, 2009].

Research on several aspects of the life cycle of *T. bryosalmonae* and its interactions with the two necessary hosts is rapidly advancing via a number of field studies and laboratory experiments. The insights gained through these investigations would be highly valuable to the improvement of the prediction capabilities of the epidemiological model developed in this Thesis, and, from a reversed perspective, results from the model permit to identify relevant directions of investigation. In particular, further laboratory studies should assess the extent and the drivers of potential clearance of infection in both bryozoans and fish, as this aspect has been shown to be dominant in promoting parasite persistence. Furthermore, absolute quantification of the number of released spores from both hosts and encounter rates would help define a proper value (or, more reasonably, a feasible range) for parameters $\pi_B^*$ and $\pi_F^*$, which constitute the largest unknown in the epidemiological model. The conjecture that PKD transmission from fish to bryozoans is minor compared to vertical transmission found substantiation both from the modelling exercise of Chapter 6 and from previous pilot laboratory studies [Tops and Okamura, 2003], but certainly requires additional inquiry. If confirmed, it would imply that measured (or model-inferred) eDNA concentrations of *T. bryosalmonae* could be directly linked to the force of infection in fish, with evident benefits from a modelling viewpoint. Additionally, several features of the interaction between the myxozoan parasite
Conclusions and perspectives

and the bryozoan host still need to be elucidated, as pointed out by Hartikainen and Okamura [2015] and Fontes et al. [2017c]: in particular, the role of fission and fragmentation of bryozoan colonies in transmitting covert infection (but see Fontes et al. [2017b]), or alternatively, in isolating uninfected portion of the zooids; the extent to which local adaptation and strain differentiation could produce less susceptible (or possibly immune) bryozoan genotypes; the role played by infected statoblasts in prompting PKD outbreaks in disease-free environments, possibly caused by dispersal via waterfowl or through long-term survival in regions where PKD was previously eradicated. A deeper understanding of these issues would be of paramount importance for the enhancement of the current epidemiological model and the reliability of its predictions.

Beyond the case of PKD, the approach developed in this Thesis can possibly be extended to the study of other waterborne epizootics, such as whirling disease and infection by Ceratomyxa shasta, both caused by myxozoan parasites with complex life cycles. Some of the tools presented in this work are readily available for broader ecohydrological applications: the species distribution model presented in Chapter 6 introduces a hydrological viewpoint to the field of eDNA sampling in freshwater environments, allowing the detection and quantification of sources of genetic material in a river network, and the assessment of the environmental variables associated with the presence or absence of the target species. The spatially-explicit temperature model introduced in Chapter 7 represents a powerful, ecohydrologically-oriented tool in the vast field of stream temperature models, particularly aimed at applications such as the identification of river areas subject to the highest risk related to PKD, the assessment of habitat suitability of fish or of other freshwater species.

Although the identification of the causative agent of PKD is relatively recent [Canning et al., 1999, 2000], research in this field, understandably prompted by the relevant environmental and economic implications of this disease, has successfully striven to unravel the main features of its transmission dynamics. Nonetheless, as mentioned above, several issues are still far from being fully resolved. The metacommunity framework proposed in this Thesis forms part of this research process, as it subsumes the current parasitological and epidemiological knowledge enabling the formulation of catchment-wide predictions of PKD risk and promoting new directions of investigation, while at the same time allowing a description of realistic features of the fluvial environment from data. In this regard, more effort should be put in place to formulate effective disease control strategies, which are currently scarce, and whose design and deployment would be greatly supported by the tools developed in this Thesis.
A Use of survival models to estimate PKD transmission parameters

Survival analysis consists in a widely applied variety of methods for the estimation of the time until a given event occurs. Areas of application include, among others, medicine, biostatistics, economics, reliability analysis in engineering, or event-history analysis in sociology. In this Appendix, survival analysis tools are exploited to gain insights on temperature-dependent local PKD transmission dynamics for the brown trout. Three different modelling approaches are tested against data obtained in a laboratory experiment, where brown trout kept at 12 and 15 °C were experimentally infected with *T. bryosalmonae*. In particular, the goals of this study are to determine the rate of disease development and to assess whether brown trout are able to clear their infection status. Results show that the estimated values of the rate of disease development are compatible with those chosen as reference in Chapter 3; moreover, infection clearing apparently takes place at a slow (< 0.01 d\(^{-1}\)) rate, which is negatively correlated with temperature.

Methods

The experiment. The analyses here presented are based on a laboratory experiment conducted by Strepparava et al. [2017]. For the sake of completeness, the experiment is here briefly described. Specific-pathogen free YOY brown trout (*n* = 1340) were held in six tanks with water at a constant temperature (three tanks at 12 °C and three at 15 °C, with one tank per temperature serving as uninfected control). Approval for animal experiments was obtained from the Cantonal Veterinary Office (Bern, Switzerland) (Authorization BE60/14). A homogeneous infection solution was formed by disrupting bryozoans overtly infected with *T. bryosalmonae* and suspended in tap water. The infection solution was mixed to the four exposure tanks before re-activating water flow. Fish sampling for parasite detection in the kidney was performed daily for the first 7 days post exposure (dpe), then at larger intervals until 320 dpe, where all remaining fish were sampled. In all other samples, either 5 or 10 fish were removed from each tank. Sampled fish were euthanized before kidney extraction. During
Appendix A. Use of survival models to estimate PKD transmission parameters

the experiment, cumulative mortality ranged from 1% in the control group, to 2.3% and 3.5% for the 12-°C and 15-°C groups, respectively, with no significant differences among any of these groups. No fish showed external signs of PKD, although prevalence in sampled fish was always greater than zero starting from 7 dpe for tanks at \( T = 15 \) °C (15 dpe for \( T = 12 \) °C).

The models. Before introducing the models, a brief digression on fundamental survival analysis concepts and notation is proposed, and affinities with the fish infection experiment are highlighted. Survival analysis (see Bewick et al. [2004] for a review) focuses on the time at which the event of interest (death, failure, fish infection in the present case) occurs in each of the \( n \) individuals followed during a certain period. Events can be fully observed when the exact time \( t \) of event occurrence is known, right-censored when the individual did not experience the event during the period when this subject was observed\(^1\), or left-censored if at time \( t \) it is only possible to ascertain that the event occurred in an undetermined past instant. Note that, in the case of the experiment, all observations are either left- (when a sampled fish is found infected) or right-censored (when a sampled fish is found uninfected).

Generally, survival analysis terminology considers all events as deaths. The survival function \( S(t) \) represents the probability that an individual is still alive at time \( t \): \( S(t) = \Pr(\tau > t) \), where \( \tau \) is a random variable representing the time of death. Its complement to unity is defined as the lifetime distribution function: \( F(t) = \Pr(\tau < t) \), whose corresponding probability density function is indicated with \( f: f(t) = F'(t) \), expressing the probability density of experiencing death around time \( t \). Finally, the hazard function is defined as

\[
H(t) = \lim_{dt \to 0} \frac{\Pr(t < \tau < t + dt)}{S(t)} = \frac{1 - f(t)}{S(t)}.
\]

Survival models can be partitioned into two main categories: semiparametric models (such as the Cox proportional hazards model [Cox, 1972] or the accelerated time failure model [Wei, 1992]), in which the effect of explanatory variable is assessed without specifying a functional form for the hazard function \( H(t) \); and parametric models, where instead a given expression for \( H(t) \) is assumed. In the case at stake, the goal is to find a functional form for \( H(t) \) to which the disease development rate \( h \) introduced in Chapter 2 (see Table 2.2) can be related; hence, parametric models are chosen. A typical probability distribution used in parametric models is the Weibull distribution

\[
f(t) = \lambda pt^{p-1} \exp(-\lambda t^p)
\]

where \( \lambda \) and \( p \) are parameters. The corresponding survival and hazard function read, respectively:

\[
S(t) = \exp(-\lambda t^p); \quad H(t) = \lambda pt^{p-1}.
\]

\(^1\)In medicine, this can also happen if a patient withdraws from a clinical trial or becomes unreachable before the trial is completed; in this case the patient is said to be lost to follow-up.
With \( p = 1 \), the Weibull distribution collapses into the exponential distribution. The resulting model is then analogous to the epidemiological model presented in §2.1, if one assumes that all fish have successfully been exposed to the parasite, but none of them developed acute PKD infection.\(^2\) Indeed, in this case \( S(t) \) decays exponentially in time, mirroring the exponential decay of the abundance of exposed\(^3\) fish \( F_E(t) \) due to infection development; furthermore, under the assumption of exponential distribution, one has \( H = \lambda \) and, through comparison with the epidemiological model, \( H \) equals the disease development rate \( h \). Unlike the epidemiological model (where a parabolic form was used, see Fig.2.1), the dependence of the hazard function on temperature is here expressed via an exponential link:

\[
H(t \mid T) = \lambda(T) = \exp(\beta_0 + \beta_1 T),
\]
as customary in survival models. \( \beta_0 \) and \( \beta_1 \) are parameters that need to be calibrated.

By introducing the dummy variable \( \delta_i \), which equals 0 when the observation for the individual \( i \) is right-censored, while \( \delta_i = 2 \) in the case of left-censoring\(^4\), the likelihood for this first model (hereafter termed EXP) can be written as:

\[
\mathcal{L}(\beta_0, \beta_1) = \prod_{i=1}^{n} \left[ 1 - \exp(\lambda(T_i) \cdot t_i) \right]^{0.5\delta_i} \left[ \exp(\lambda(T_i) \cdot t_i) \right]^{1-0.5\delta_i}. \tag{A.1}
\]
The first term under square brackets represents the probability that the model correctly predicts a left-censored observation, while the second term refers to right-censoring.

An enhancement of model EXP is obtained by following the suggestion of Corbière and Joly [2007]. These authors introduced the concept of mixture cure models, in which the assumption that all individuals will eventually experience the event is relaxed. A new dummy variable \( U_i \) is introduced: one has \( U_i = 1 \) if individual \( i \) is susceptible\(^5\), \( U_i = 0 \) otherwise. The fraction of susceptible individuals, termed \( \pi \), can be expressed as a function of explanatory variables (temperature in the case at stake), by means of a logit\(^6\) function:

\[
\pi(T) = \frac{\exp(\gamma_0 + \gamma_1 T)}{1 + \exp(\gamma_0 + \gamma_1 T)},
\]
\( \gamma_0, \gamma_1 \) being two additional parameters. The survival function for the mixture cure model is

---

\(^2\)This assumption is supported by the fact that in the experiment fish mortality was negligible, and it corresponds to setting \( \epsilon = 1 \) in the epidemiological model.

\(^3\)In this case it is hypothesized that infected fish are only those belonging to \( F_I \) and \( F_C \) classes. Conversely, in the previous Chapters, prevalence was calculated by also including the exposed compartment \( F_E \).

\(^4\)The usual notation uses \( \delta_i = 1 \) to label fully observed events.

\(^5\)The term 'susceptible' is here used in the sense of 'bound to experience the event of interest', as intended by Corbière and Joly [2007]. In the analogy with the epidemiological model, susceptible individuals are identified with the exposed class \( F_E \), while individuals characterized by \( U = 0 \) correspond to the susceptible compartment \( F_S \). See also note 3.

\(^6\)Another typically used link function is the probit, namely the cumulative distribution function of the standard normal distribution whose argument is the linear predictor \( \gamma_0 + \gamma_1 T \) in the present case.
Appendix A. Use of survival models to estimate PKD transmission parameters

given by

\[ S(t \mid T) = \pi(T)S(t \mid U = 1, T) + 1 - \pi(T) \]

where \( S(t) \) is the unconditional survival function for the entire population, while \( S(t \mid U = 1) \) is the survival function for susceptible individuals. The likelihood of the resulting model (termed MIX) reads:

\[
L(\beta_0, \beta_1, \gamma_0, \gamma_1) = \prod_{i=1}^{n} \left\{ \pi(T_i) \left[ 1 - \exp(-\lambda(T_i) \cdot t_i) \right] \right\}^{0.5\delta_i} \cdot \left[ 1 - \pi(T_i) + \pi(T_i) \exp(-\lambda(T_i) \cdot t_i) \right]^{1-0.5\delta_i}.
\]  

(A.2)

Note that, for \( \pi = 1 \), Eq. (A.2) reduces to (A.1).

Model MIX can be further extended by introducing the possibility that fish clear the infection. Let \( f_R(t) \) be the probability density of recovering around time \( t \), given previous infection. In accordance with the previous formulations, \( f_R \) is assumed to follow an exponential distribution with mean \( \frac{1}{\lambda_R} \), possibly depending on temperature:

\[
f_R(t \mid T) = \lambda_R(T) \exp(-\lambda_R(T) \cdot t),
\]

while \( S_R(t \mid T) = 1 - F_R(t \mid T) = \exp(-\lambda_R(T) \cdot t) \) is the probability of being still infected at \( t \). Note that \( \lambda_R \) mirrors the infection clearing rate \( \zeta \) introduced in the epidemiological model. Again, an exponential link can be used to express the relationship \( \lambda_R(T) \):

\[
\lambda_R(T) = \exp(\epsilon_0 + \epsilon_1 T).
\]

Finally, the likelihood for this third model (termed REC) reads:

\[
L(\beta_0, \beta_1, \gamma_0, \gamma_1, \epsilon_0, \epsilon_1) = \prod_{i=1}^{n} \left\{ \pi(T_i)F(t_i \mid T_i)S_R(t_i \mid T_i) \right\}^{0.5\delta_i} \cdot \left[ 1 - \pi(T_i) + \pi(T_i) \left[ S(t_i \mid T_i) + F(t_i \mid T_i)F_R(t_i \mid T_i) \right] \right]^{1-0.5\delta_i}.
\]  

(A.3)

In words, the probability of a left-censored observation is the product of the probabilities of being susceptible (\( \pi \)), having got infected before time \( t \) (\( F \)) and not having cleared the infection at time \( t \) (\( S_R \)). On the other hand, the probability of a right-censored observation is made up by the probability of not being susceptible (\( 1 - \pi \)) and the probability of having cleared the infection at \( t \). The latter is in turn equal to the product of the probabilities of being susceptible (\( \pi \)), have become infected before \( t \) (\( F \)) and having recovered given previous infection (\( F_R \)). This latter model formulation is a novelty of this study.
Table A.1 – Survival models. Best fit parameter values and performance scores.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\beta_0$</th>
<th>$\beta_1$</th>
<th>$\gamma_0$</th>
<th>$\gamma_1$</th>
<th>$\delta_0$</th>
<th>$\delta_1$</th>
<th>$\log L$</th>
<th>$\Delta$AIC</th>
<th>$\Delta$BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXP</td>
<td>-7.73</td>
<td>0.15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-809.4</td>
<td>426.2</td>
<td>416.1</td>
</tr>
<tr>
<td>MIX</td>
<td>-4.51</td>
<td>0.12</td>
<td>-2.56</td>
<td>0.18</td>
<td>-</td>
<td>-</td>
<td>-616.9</td>
<td>45.2</td>
<td>40.2</td>
</tr>
<tr>
<td>REC</td>
<td>-5.91</td>
<td>0.15</td>
<td>$\pi = 1$</td>
<td>-9.09</td>
<td>-0.12</td>
<td>-592.3</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Figure A.1 – Results from survival models. a) Comparison between functional forms for the disease development rate $h$ and the infection clearing rate $\zeta$ for the three survival models here tested and the reference values of Table 3.1. Dots indicate values at 12 and 15 °C that were actually estimated with the survival models. b), c): Output of the epidemiological model run with the values of $h$ and $\zeta$ estimated at 12 and 15 °C, respectively. Dots indicate sampled prevalences averaged over the two tanks.

Results and discussion

Table A.1 displays best fit parameter values and maximum likelihood values obtained for the three models. Optimization was performed via a simulated annealing algorithm [Kirkpatrick et al., 1983]. Models’ performances are compared via both the Akaike Information Criterion (AIC)\(^7\) and the Bayesian Information Criterion (BIC)\(^8\). Table A.1 also presents differences in AIC and BIC scores with respect to the best performing models, namely model REC according to both criteria. Calibration of model REC failed to find unambiguous optimal values for parameters $\gamma_0$ and $\gamma_1$ expressing the relationship between the fraction of susceptible individuals and temperature; in fact, all pairs $[\gamma_0; \gamma_1]$ such that $\pi = 1$ enabled likelihood maximization, provided that the other 4 parameters ($\beta_0$, $\beta_1$, $\delta_0$, $\delta_1$) are set to their optimal values of Table A.1. Instead, model MIX predicts that only 40% of the fish kept at 12 °C and 53% of those at 15 °C were successfully exposed to the parasite. Fig. A.1a shows the values of parameters $h$ and $\zeta$ estimated for the three survival models at both temperature levels. Figs. A.1b, c show the output of the epidemiological model run with the estimated parameters at both 12 and 15 °C. The poor performance of model EXP suggests that a simple exponential formulation for the rate of disease development does not reproduce adequately the observed data. The hypothesis

\(^7\)AIC = $2N_p - 2\log(L)$, where $N_p$: number of parameters
\(^8\)BIC = $N_p \log(n) - 2 \log(L)$, where $n$: number of observations (here equal to the number of individuals).
Appendix A. Use of survival models to estimate PKD transmission parameters

of insufficient exposition to the parasite for a fraction of the fish examined (model MIX) yields better results; moreover, in this case the estimated values of $h$ are comparable with the reference values of the epidemiological model (see Fig. A.1a). By also including an infection clearing mechanism (model REC), the goodness of fit is further improved: estimated $h$ values are even closer to the reference curve with respect to the MIX model, while $\zeta$ values are fairly greater than the hypothesized value $\zeta = 0.001 \, d^{-1}$, although, as expected, they are one order of magnitude lower than $h$ values (see Chapter 3). The infection clearing rate $\zeta$ appears negatively correlated with temperature, as the parameter $\delta_1$ expressing the magnitude of this dependence is significantly different from zero (details not shown). This evidence suggest a possible improvement for the local PKD epidemiological model, where $\zeta$ was assumed constant as a proof-of-concept value.

Strepparava et al. [2017] claimed that there was no statistically significant decay in prevalence over the period between 80 and 200 dpe, while the prevalence decrease was significant after 200 dpe, and concluded that brown trout at stable temperatures do not clear the infection. However, the results derived from survival analysis actually lead to the opposite conclusion. It appears evident how the choice of the type of statistical analysis or model used to interpret the data has a great influence on the conclusions that can be drawn. Hence, the possibility that brown trout clear the infection can not be excluded, and the presumable order of magnitude of the clearing period is one year (estimated values for model REC are 252 d at 12°C and 361 d at 15°C). It is likely that if the experiment lasted longer, or if more samples were taken between 200 and 320 dpe, more evidence would have supported the conclusion that brown trout are able to clear the infection.

List of symbols

The following table lists all symbols used in this Appendix. For the sake of clarity and completeness, symbols already defined in previous Chapters are here repeated.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F$</td>
<td>Lifetime distribution function</td>
<td>-</td>
</tr>
<tr>
<td>$f$</td>
<td>Lifetime probability density function</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$F_R$</td>
<td>Cumulative distribution function of recovery time (conditional to previous infection)</td>
<td>-</td>
</tr>
<tr>
<td>$f_R$</td>
<td>Probability density function of recovery time (conditional to previous infection)</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$H$</td>
<td>Hazard function</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$h$</td>
<td>Rate of disease development (= $H$)</td>
<td>$T^{-1}$</td>
</tr>
<tr>
<td>$\mathcal{L}$</td>
<td>Likelihood of a model</td>
<td>-</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of individuals considered in the study</td>
<td>-</td>
</tr>
<tr>
<td>$N_p$</td>
<td>Number of parameters</td>
<td>-</td>
</tr>
<tr>
<td>$p$</td>
<td>Shape parameter for the Weibull distribution</td>
<td>-</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Dimension</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>$S$</td>
<td>Survival function</td>
<td>-</td>
</tr>
<tr>
<td>$t$</td>
<td>Realization of the random variable $\tau$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T_i$</td>
<td>Water temperature corresponding to individual $i$</td>
<td>$\Theta$</td>
</tr>
<tr>
<td>$U$</td>
<td>Dummy variable identifying susceptible individuals</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>Baseline parameter for the hazard function $H$</td>
<td>$\log(T)$</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>Parameter expressing dependency on temperature in $H$</td>
<td>$\log(T)$</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>Baseline parameter for the probability of being susceptible $\pi$</td>
<td>-</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>Parameter expressing dependency on temperature in $\pi$</td>
<td>$\Theta^{-1}$</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Dummy variable distinguishing left- and right-censored events</td>
<td>-</td>
</tr>
<tr>
<td>$\epsilon_0$</td>
<td>Baseline parameter for the relationship $\lambda_R$</td>
<td>$\log(T)$</td>
</tr>
<tr>
<td>$\epsilon_1$</td>
<td>Parameter expressing dependency on temperature in $\lambda_R$</td>
<td>$\log(T)\Theta^{-1}$</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Scale parameter for the Weibull distribution $f(t)$</td>
<td>$T^{-p}$</td>
</tr>
<tr>
<td>$\lambda_R$</td>
<td>Mean value of the probability density function $f_R$</td>
<td>$d^{-1}$</td>
</tr>
<tr>
<td>$\pi_i$</td>
<td>Probability that individual $i$ is susceptible</td>
<td>-</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Random variable expressing event time</td>
<td>$T$</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Rate of infection clearing ($= \lambda_R$)</td>
<td>$T^{-1}$</td>
</tr>
</tbody>
</table>
B Detection of *F. sultana* eDNA in water samples

**eDNA detection.** Pre-sterilized plastic bottles (with 10% bleach followed by UV-B treatment) were used to collect water from the river by submerging the bottle with a gloved hand. The samples were transported to the laboratory on ice and filtered within the same day on to 5-cm diameter, 0.45-μm pore size individually packaged sterile membrane filters (Merck Millipore). A vacuum pump with a borosilicate glass filtration setup was used and sterilized in 10% bleach between each sample. Negative controls were created by filtering MilliQ water through a sterile filter at the start and end of the filtration session, as well as once during the filtration (after sample 7). Filter papers were placed in 2-mL bead beating tubes (obtained from the kit described below) and frozen at 80 °C until extraction. Filter papers were cut with sterilized scissors to break them up and eDNA was extracted from all filter papers, including controls, using a PowerSoil® DNA kit (MO BIO Laboratories) in a dedicated clean laboratory (free of PCR products). The kit includes a bead beating step and a separate inhibitor removal step. The eDNA was eluted in 60 μL of Solution C6 and subsequently preserved at -20 °C. eDNA samples were only removed from the -20 °C freezer for screens and remained at room temperature for a maximum of 2 h.

**qPCR assay.** Specific primers for bryozoan detection (Fs_16S_F1q and Fs_16S_R1q) were designed based on inspection of 16S mitochondrial sequences of all major clades of phylactolaemates. The primers were designed to amplify fredericellid phylactolaemates, with 100% identity in primer and probe sequence with the most common host in Europe, *F. sultana*. The size of the *F. sultana* PCR fragment was 71 bp. A custom internal positive control (IPC) template was spiked into all reactions and amplified in multiplex with the *F. sultana* probe assay. The inclusion of the IPC allowed the detection of possible PCR inhibition. The IPC primers and probe are reported in Table B.1.

The 71-bp fragment of *F. sultana* was amplified from genomic DNA samples derived field collected colonies. PCR reaction composition and cycling conditions were the same as those

---

1This Appendix, extracted from Carraro et al. [2017], was developed and written by Ines Fontes and Hanna Hartikainen and is here included for the sake of completeness.
Appendix B. Detection of *F. sultana* eDNA in water samples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Fredericella sultana</td>
<td>Fs_16S_F1q</td>
<td>CATTGAGCTTC-GGAATGTT</td>
<td>20</td>
<td>45.0</td>
<td>54.4</td>
<td>49.4</td>
<td>60</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>Fs_16S_R1q</td>
<td>ATGAAACCTCG-TCCCTGTG</td>
<td>20</td>
<td>50.0</td>
<td>56.3</td>
<td>49.4</td>
<td>60</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>Fs_Probe_16S_1</td>
<td>Cy5-GGGGTCAG-GTTGCTAAGC-CATGA-BHQ-2</td>
<td>23</td>
<td>56.5</td>
<td>62.9</td>
<td>-</td>
<td>60</td>
<td>200</td>
</tr>
<tr>
<td>Internal Positive Control</td>
<td>MIMr</td>
<td>ATGTGACTGGAC-TCCGTATCG-Gy3-CGACGGCC-AGTGAATTGTA-ATACGA-BHQ-2</td>
<td>22</td>
<td>50.0</td>
<td>57.8</td>
<td>52.8</td>
<td>60</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>IPC_probe</td>
<td>Cy3-CGACGGCC-AGTGAATTGTA-ATACGA-BHQ-2</td>
<td>25</td>
<td>48.0</td>
<td>59.9</td>
<td>-</td>
<td>60</td>
<td>250</td>
</tr>
</tbody>
</table>

Table B.1 – Oligonucleotide primers and probe sequences for the qPCR assay. Calculated melting ($T_m$) and annealing ($T_a$) temperatures are at 0.5 μM for primers and at 0.25 μM for probes. Reporter and quencher dyes for each probe are presented in the sequences: Cy®3 = cyanine 3; Cy®5 = cyanine 5; BHQ-1 = Black Hole Quencher®-1; and BHQ-2 = Black Hole Quencher®-2. The optimal primer and probe concentrations (Optimal conc) in multiplex reactions are reported in nM.

used to test primer pairs. The amplification was performed in 50-μL reactions and 5 reactions were pooled prior to gel purification using the QIAquick Gel Extraction Kit (QIAGEN). The concentration of the pooled purified products and the IPC template stock solution (100 mM) were measured using a Qubit® 2.0 Fluorometer double-stranded DNA (dsDNA) high-sensitivity Assay (in ng/μL; Invitrogen) and adjusted to 1 nM. A 1:10 serial dilution of the standardized 1-nM solution of each of the two fragments was performed and pooled. Seven standards of the resulting serial dilution were included in each run. A master standard curve for each target obtained from a single multiplex experiment (*F. sultana* and IPC) was applied to all subsequent experiments using the first standard point as a calibrator (i.e. 8.33 · 10⁻¹² mol/L per PCR reaction).

PCR amplification and target quantification were performed using a LightCycler® 480 II (Roche Diagnostics) with color compensation. The detection format, analysis mode and color compensation data were the same across all experiments. The optimal forward primer concentrations are reported in Table B.1. The amplification was performed in a final volume of 10 μL containing 7.5 μL of master mix with probes and 2.5 μL of template DNA. Each PCR reaction contained 1x LightCycler® 480 Probes Master (Roche Diagnostics) and 1x optimized primer-probe mix. Two master mixes were prepared per experiment: one without IPC to generate the standard curve and test the non-template control (NTC); and a second one with IPC for the quantification of the sample and IPC calibrator reactions, which contained 1 μL of 1.00 · 10⁻¹³ mol/L of IPC. The reaction volumes of both master mixes were made up to 10 μL with PCR-grade water (Sigma-Aldrich). PCR amplification was performed in triplicates with an initial DNA polymerase activation/denaturation step of 95 °C for 10 minutes (4.8 °C/s ramp rate) followed by 45 cycles of denaturation at 95 °C for 10 s (4.8 °C/s ramp rate) and annealing/extension at 60 °C for 1 minute (2.5 °C/s ramp rate), followed by a cooling step of
Table B.2 – Limits of quantification (LOQ) and detection (LOD) for each qPCR assay, given as Cq values, calculated molar concentration, molecular weight (MW) and copy number. Results are given as mean, standard deviation is included between parentheses. IPC: internal positive control. The LOQ and LOD for the IPC assay are the same.

<table>
<thead>
<tr>
<th>Target</th>
<th>Cq</th>
<th>Calculated conc. [mol/L]</th>
<th>Q [mol]</th>
<th>MW [ng]</th>
<th>Copy number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fredericella sultana</td>
<td>LOQ</td>
<td>34.977 (0.191)</td>
<td>8.68·10^{-19}</td>
<td>8.68·10^{-24}</td>
<td>4.07·10^{-10}</td>
</tr>
<tr>
<td></td>
<td>LOD</td>
<td>37.097 (0.967)</td>
<td>4.02·10^{-20}</td>
<td>4.02·10^{-25}</td>
<td>1.88·10^{-11}</td>
</tr>
<tr>
<td></td>
<td>LOQ/LOD</td>
<td>34.930 (0.645)</td>
<td>9.12·10^{-20}</td>
<td>9.12·10^{-25}</td>
<td>5.84·10^{-11}</td>
</tr>
</tbody>
</table>

40 °C for 10 s (2.0 °C/s ramp rate). PCR amplification took a total of 1.5 h. Fluorescence was acquired once per cycle at the end of the annealing/extension phase. A QIAgility pipetting robot (QIAGEN) was used to prepare and dispense the master mix and template DNA into 384-well white qPCR microplates (TreffLab) in a maximum of 2 h to avoid evaporation. All samples used in this study were kept at -20 °C and defrosted in a fridge at 4 °C for 30 minutes before mixing and pipetting into qPCR plates. At the end of each experiment, an absolute quantification of samples was performed using the second derivative maximum method using a high confidence algorithm on the LightCycler® software.

The Limit of Detection (LOD, see Table B.2) of the assay was 37.097 for *F. sultana*, defined as the highest Cq mean observed for a truly positive sample with all triplicates fluorescing. Limit of Quantification (LOQ) was set at Cq value where target concentration no longer exhibited linear relationship with Cq readings. The LOQ of the assay was 34.977 for *F. sultana*. Reactions were considered to be positive when Cq value was not greater than 34.98 for *F. sultana*, corresponding to 0.78 template DNA copies. NTCs either did not generate fluorescence, this was higher than the LOQ or it was not present in all triplicates. Experimental samples were considered positive if there was fluorescence in all three triplicates and the Cq mean was not greater than the LOQ.


Bibliography


Bibliography


Bibliography


Bibliography


Pavlovsky, E. N. et al. (1966). Natural nidality of transmissible diseases with special reference to the landscape epidemiology of zooanthroponoses. *Natural Nidality of Transmissible Diseases with special reference to the Landscape Epidemiology of Zooanthroponoses*. 

140
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


Curriculum vitae

Luca Carraro
Avenue de Chailly 14, 1012 Lausanne, Switzerland
(+41) 787 976 750  (+41) 21 69 33777
luca.carraro@epfl.ch
Skype luca.carraro.6

Gender Male | Date of birth 20 March 1989 | Nationality Italian

EDUCATION AND TRAINING

Apr 2014 – present
Ph.D. candidate - Civil and Environmental Engineering
Laboratory of Ecohydrology – École Polytechnique Fédérale de Lausanne, Switzerland
Thesis title
Ecohydrological and Metacommunity Studies of Proliferative Kidney Disease Spread in Freshwater Salmonid Fish
Fields Ecology, Epidemiology, Hydrology, Mathematical modelling
Synopsis
This thesis develops a comprehensive approach, involving field and modelling work, for the prediction of the incidence of proliferative kidney disease (PKD) in river basins. PKD is a temperature-dependent, high-mortality pathology critically affecting salmonids in temperate rivers. The modelling framework comprises disease transmission dynamics, the role of the hydrological and thermal drivers, habitat suitability and migration patterns of fish and the assessment of the spatial distribution of the parasite host based on environmental DNA sampling.
Advisors
Prof. Andrea Rinaldo, Prof. Enrico Bertuzzo

Jan 2012 – Apr 2014
Master of Science in Civil Engineering (specialization in Hydraulics)
Università degli Studi di Padova, Padua, Italy
Final grade 108/110
Thesis title
Inferring Age Mixing Processes and Travel Time Distributions from a Lysimeter Experiment
Field: Hydrology. Advisors: Prof. Gianluca Botter, Prof. Andrea Rinaldo
Main subjects
Hydraulic Constructions, Hydrodynamics, Hydrology, Maritime Structures, River Hydraulics

Oct 2008 – Nov 2011
Bachelor of Science in Civil Engineering
Università degli Studi di Padova, Padua, Italy
Final grade 110/110
Main subjects
Calculus, Physics, Numerical Analysis, Hydraulics, Soil Mechanics, Mechanics of Materials and Structures, Structural Analysis and Design

TEACHING

Apr 2014 – present
Teaching assistant
École Polytechnique Fédérale de Lausanne, Switzerland
Material preparation and exercise sessions for the course “Water Resources Engineering” (M.Sc. in Environmental Science and Engineering).

Feb 2017 – Jul 2017
Teaching assistant
École Polytechnique Fédérale de Lausanne, Switzerland
Exercise sessions for the course “Mathématiques 2A, 2B” (B.Sc., remedial classes Mise à niveau – subjects: linear algebra and geometry).
PERSONAL SKILLS

Mother tongue Italian

Other languages

<table>
<thead>
<tr>
<th>UNDERTANDING</th>
<th>SPEAKING</th>
<th>WRITING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening</td>
<td>Reading</td>
<td>Spoken interaction</td>
</tr>
<tr>
<td>English</td>
<td>C1</td>
<td>C1</td>
</tr>
<tr>
<td>French</td>
<td>B2</td>
<td>C1</td>
</tr>
<tr>
<td>German</td>
<td>A2</td>
<td>A2</td>
</tr>
</tbody>
</table>


Communication / Organisational skills

Well-organised, able to meet deadlines and prioritize tasks, detail-oriented, flexible, inclined to teamwork, able to cope with stress, disposed to learn.

Computer skills

– Text editor & office suite: Microsoft Office (expert), LATEX (expert)
– Programming: MATLAB (expert), Mathematica (beginner), R (beginner)
– Graphics & Design: Adobe Illustrator (expert), AutoCAD (intermediate)
– Geographic information system: ArcGIS (intermediate)

ADDITIONAL INFORMATION

Publications


Conference posters

