

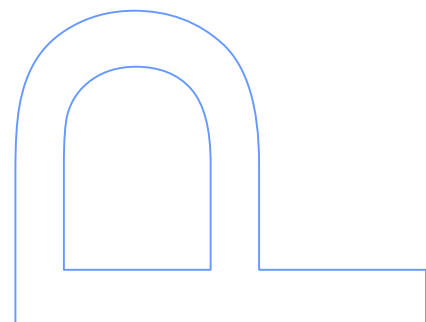
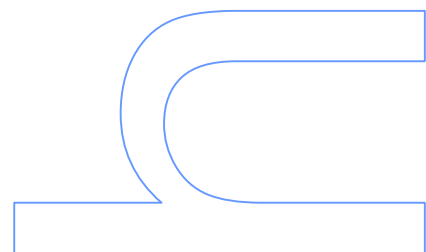
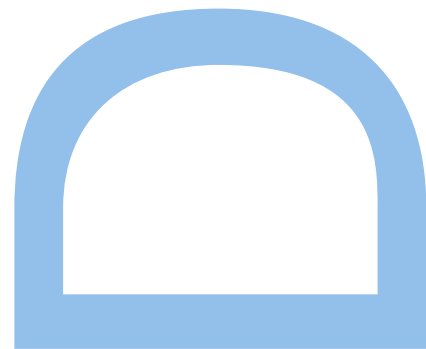


# Modelling biodiversity patterns and processes to support conservation in stream networks

Mário Rui Mota Ferreira

Doutoramento em Biodiversidade, Genética e Evolução  
Departamento de Biologia  
2021

**Orientador**  
Pedro Beja, CIBIO-InBio





## Nota Prévia

Na elaboração desta dissertação, e nos termos do número 2 do Artigo 4º do Regulamento Geral dos Terceiros Ciclos de Estudos da Universidade do Porto e do Artigo 31º do D.L. 74/2006, de 24 de Março, com a nova redação introduzida pelo D.L. 230/2009, de 14 de Setembro, foi efetuado o aproveitamento total de um conjunto coerente de trabalhos de investigação já publicados ou submetidos para publicação em revistas internacionais indexadas e com arbitragem científica, os quais integram alguns dos capítulos da presente tese. Tendo em conta que os referidos trabalhos foram realizados com a colaboração de outros autores, o candidato esclarece que, em todos eles, participou ativamente na sua conceção, na obtenção, análise e discussão de resultados, bem como na elaboração da sua forma publicada.

Lista de artigos:

**Capítulo 3 – Ferreira, M.**, Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>

**Capítulo 4 – Filipe, A. F.**, Quaglietta, L., **Ferreira, M.**, Magalhães, M. F., & Beja, P. (2017). Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks. *Biological Invasions*, 19(10), 2899–2912. <https://doi.org/10.1007/s10530-017-1492-3>

**Capítulo 5 – Mota-Ferreira, M.**, & Beja, P. (2020). Combining geostatistical and biotic interaction modeling to predict amphibian refuges under crayfish invasion across dendritic stream networks. *Diversity and Distributions*, 26(6), 699–714. <https://doi.org/10.1111/ddi.13047>

**Capítulo 6 – Mota-Ferreira, M.**, Filipe, A. F., Magalhães, M. F., Carona, S., & Beja, P. (2021). Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change. *Diversity and Distributions*, 27(2), 313–326. <https://doi.org/10.1111/ddi.13189>

Este trabalho foi apoiado pela Fundação para a Ciência e Tecnologia (FCT) através de fundos nacionais do Ministério da Educação e Ciência e pelo Fundo Social Europeu através do Programa Operacional Potencial Humano (POPH) – QREN – Tipologia 4.1 – Formação Avançada (bolsa de doutoramento SFRH/BD/95202/2013). Foi ainda financiado pelo projeto de investigação SABOR LTER (FCT: LTER/BIA-BEC/0004/2009) e pela EDP – Energias de Portugal através da Cátedra em Biodiversidade.



Este trabalho foi desenvolvido em:



This thesis should be cited as:

Mota-Ferreira (2021). *Modelling biodiversity patterns and processes to support conservation in stream networks*. Ph.D. Thesis, Faculdade de Ciências da Universidade do Porto, Portugal

Cover, and chapters separators illustrations by Inês Pontes (all rights reserved).

## Acknowledgements

Frist and foremost an acknowledgement to my supervisor Pedro Beja. He created all the conditions to the fulfilment this work, he funded the Sabor LTER site, obtained the funds to support the immense field work and gave me every condition to finish this thesis. You're also a remarkable scientist, a great supervisor, and I never cease to be amazed by your clinical eye, insightful remarks; and writing skills. In the words of Isaac Newton "If I have seen further it is by standing on the shoulders of Giants."

An appreciation for the co-authors of the papers, specially to professor Maria Filomena Magalhães, the two field "Ladies" that was my privilege to fish with: Filipa Filipe (aka princesa ervilha) and Sara Carona (senhora dos bordalos), and my mate of many adventures Lorenzo Quaglietta (aka Lontra, Lontrenzo, etc...). Hope to continue to work with you in the future.

This Ph.D. had an insane amount of field work that begun even before I won the scholarship. I'm very glad that I did very little work alone. You made my work, not only possible, but also easier and cheerful! A great thank you to Pedro Henrique, Rita Severino, Sara Ivone, Sergio Henrique, Duarte Prata, João Lopes, João Guilherme, José Barradas, Rui Cereja, Ninda Batista, Joana Pinto, Erika Almeida, Giorgio Pace, Miguel Pinto, William Agudelo, Guilherme Buzzo, Filipe Sá, and Nuno Pereira. A special regard to Ivo Rosa, that have accompanied more than most. I am humbled by your sensibility for biodiversity and really admire your enthusiasm, but for the love of god, do finish your course!!

For the great part of the PhD the field station was located in the small town of Cerejais, thanks for all community there that welcome our teams there. Also thank you to all colleagues that share the field station with me, Francisco Amorim (We are falling behind on the sumol meetings), Andreia Penado, Marta Sampaio, Vanessa Mata, Helena Geraldés and Pedro among others.

To all the colleagues of the CIBIO Lisboa, past and present since the beginning in the facilities of IICT to actual location in ISA, Joana Santana, Luis Reino, Hugo Rebelo, Miguel Porto (your kung fu is still the mightiest!) , Sasha Vasconcelos, Patricia Rodrigues, Silvia Pina, Luís Borda de Água, Margarida Ferreira, António Ferreira, Fernando Ascensão, Marcelo d'Amico, Rafael

Barrientos (não acredito que está feito!), Francisco Moreira, Miguel Monteiro, Marta Neto, Inês Catry, Joana Bernardino, Joana Marcelino, Filipe Dias (We will catch them all!) , Marta Neto, Ricardo Rocha, Fabio Marcolino, and many others!

To all my friends around the world. To all my mates in scout movement.

To my family, you've been my emotional support. My parents, brother and sister and their companions, and my nephews.

To my wife Inês and my son Artur. It's been an incredible journey!

# Table of contents

Nota Prévia.....	i
Acknowledgements.....	iii
Table of contents .....	v
List of tables .....	ix
List of figures.....	xiii
Resumo .....	1
Palavras chave:.....	4
Abstract.....	5
Keywords:.....	8
Chapter 1 – General Introduction.....	9
1.1    Biodiversity and its decline .....	11
1.2    The freshwater biodiversity crisis .....	12
1.3    Stream Networks .....	14
1.4    Biodiversity modelling in aquatic systems.....	15
1.4.1    Species distribution models .....	15
1.4.2    The challenges of distribution modelling in river networks .....	19
1.5    Modelling dynamic community patterns.....	22
1.6    Objectives.....	24
1.7    Thesis outline .....	25
1.8    References .....	28
Chapter 2 – Study system .....	39
2.1    Study Area.....	41
2.2    Studied organisms.....	43
2.2.1    Non-native invasive crayfish .....	43
2.2.2    Stream fish .....	46
2.2.3    Amphibians .....	51
2.3    References .....	56
Chapter 3 – Modelling stream fish distributions using interval-censored detection times .....	65
Abstract.....	67
3.1    Introduction .....	68
3.2    Methods.....	70
3.2.1    Fish and environmental data .....	70

3.2.2	Neighborhood effects .....	71
3.2.3	Species distribution models .....	72
3.2.4	Simulations for the detectability model .....	73
3.2.5	Model building and evaluation .....	73
3.2.6	Species distribution mapping.....	75
3.2.7	Model Fit .....	75
3.3	Results.....	76
3.4	Discussion.....	79
	Acknowledgements.....	84
3.5	References .....	84
3.6	Supplementary Materials.....	89
	Appendix 3.S1 Supplementary Methods .....	89
	Appendix 3.S2 Interval-censored time to detection model.....	94
	Appendix 3.S3 Code used to fit the time to detection model using WinBUGS .....	98
	Appendix 3.S4 Response curves to environmental variables .....	102
	Appendix S5 Maps of prediction uncertainty .....	103
	Appendix S6 Supplementary references.....	104
	Chapter 4 – Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks .....	107
	Abstract.....	109
4.1	Introduction .....	110
4.2	Methods.....	113
4.2.1	Study area and species.....	113
4.2.2	Crayfish survey .....	113
4.2.3	Environmental and spatial data .....	114
4.2.4	Geostatistical modelling.....	116
4.2.5	Species distribution mapping.....	117
4.3	Results.....	117
4.4	Discussion.....	121
4.4.1	Environmental effects.....	122
4.4.2	Spatial effects.....	124
4.4.3	Distribution modelling of invasive species across stream networks .....	125
	Acknowledgments.....	126
4.5	References .....	127
4.6	Supplementary material .....	131



Chapter 5 – Combining geostatistical and biotic interactions modelling to predict amphibian refuges under crayfish invasion across dendritic stream networks .....	143
Abstract.....	145
5.1 Introduction .....	146
5.2 Methods.....	148
5.2.1 Study area .....	148
5.2.2 Study design.....	149
5.2.3 Field sampling .....	150
5.2.4 Environmental and spatial variables.....	150
5.2.5 Distribution modelling .....	152
5.2.6 Distribution mapping under current and future scenarios.....	153
5.3 Results.....	155
5.4 Discussion.....	162
5.4.1 Study limitations .....	163
5.4.2 Effects of biotic interaction and environmental effects on amphibian distributions.....	164
5.4.3 The role of spatial dependencies across the stream network.....	165
5.4.4 Predicting amphibian refuges under crayfish invasion.....	166
5.4.5 Conservation and management implications .....	167
5.5 Conclusions .....	168
Acknowledgements.....	168
5.6 References .....	168
Data Accessibility Statement: .....	174
5.7 Supplementary Material .....	175
Supplementary Methods: Workflow detailing the methodological procedures.....	175
Supplementary Tables .....	185
Supplementary Figures .....	187
Chapter 6 - Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change.....	193
Abstract.....	195
6.1 Introduction .....	196
6.2 Methods.....	199
6.2.1 Study area .....	199
6.2.2 Fish sampling.....	199
6.2.3 Environmental and spatial variables.....	201
6.2.4 Community trajectory analysis and modelling .....	201

6.2.5	Geostatistical modelling and mapping .....	203
6.3	Results.....	204
6.3.1	Overall assemblage patterns .....	204
6.3.2	Changes in species richness and fish catches .....	206
6.3.4	Community trajectories .....	206
6.3.5	Spatial variation in community dynamics.....	209
6.4	Discussion.....	211
6.4.1	Temporal drivers of local community change .....	211
6.4.2	Patterns and environmental drivers of spatial variation in community dynamics.....	212
6.4.3	Spatial dependencies in community temporal dynamics .....	214
6.4.4	Mapping community temporal dynamics to guide management .....	215
6.5	Conclusions .....	216
	Acknowledgements.....	216
6.6	References .....	217
6.7	Supplementary Material .....	222
	Supplementary Tables .....	222
	Supplementary Figures .....	228
	Supplementary references.....	229
	Chapter 7 – General discussion.....	231
7.1	In a nutshell.....	233
7.2	Challenges of spatial distribution modelling in dendritic networks .....	234
7.3	Modelling expanding distributions of invasive species and their consequences .....	236
7.4	Modelling temporally dynamic ecological processes in stream networks .....	238
7.5	Stream network biodiversity modelling as an essential tool for conservation .....	239
7.6	Conclusions and future research .....	241
7.7	References .....	244

## List of tables

Table 3.1 - Performance of the interval-censored time to detection model in retrieving parameter from simulated data. ....	77
Table 3.2 - Mean parameter estimates used in the distribution mapping of six freshwater fish species.....	78
Table 3.S.1 - List of fish species sampled in the Sabor catchment in the summer of 2012. .	90
Table 3.S.2 - Description and summary statistics of environmental variables used to model variation in detection rate and occupancy probability. ....	93
Table 4.1 - Parameter estimates and summary statistics for the environmental component of crayfish distribution models in the Sabor watershed.....	118
Table 4.2 - Comparison of mixture models relating the probability of occurrence of two crayfish species to environmental and spatial components. ....	119
Table 4.3 - Percentage of variance accounted for by models relating the probability of two crayfish species to environmental and spatial components. ....	120
Table 4.S1 - Summary of the variable selection procedure for models relating the probability of occurrence of <i>Procambarus clarkii</i> and environmental variables. ....	131
Table 4.S2 - Summary of the variable selection procedure for models relating the probability of occurrence of <i>Pacifastacus leniusculus</i> and environmental variables.....	132
Table 4.S3 - Summary of the selection of the mixture covariance model for <i>Procambarus clarkii</i> . ....	133
Table 4.S4 - Summary of the selection of the mixture covariance model for <i>Pacifastacus leniusculus</i> . ....	138

Table 5.1 - Parameter estimates and summary statistics for models relating the probability of occurrence of amphibian species in the Sabor watershed.....	156
Table 5.2 - Covariance structure selected for the geostatistical models of each species, indicating the percentage of variation accounted by the fixed and each spatial component of the final models.....	161
Table 5.S1 - Amphibian species detected in surveys carried out in the Sabor watershed in January-December 2015.....	185
Table 5.S2 - Percentage decline in the length of streams available for each amphibian species, in relation to expansion scenarios for the crayfish..	186
Table 6.1 - Fish species recorded through electrofishing in the River Sabor watershed between 2012 and 2019.....	205
Table 6.2 - Summary results of the best AICc models explaining variation in mean velocity of community change, directionality and pairwise distances between community trajectories. ....	209
Table 6.3 - Partitioning of the proportion of explained variation by geostatistical linear mixed models relating community directionality and mean velocity of change to environmental variables while accounting for spatial random effects.....	209
Table 6.S1 - Summary statistics of fish communities sampled at 30 sites in the Sabor watershed, between 2012-2019.....	222
Table 6.S2 - Description and summary statistics of environmental variables used to characterize the 30 stream reaches sampled for fish in the Sabor watershed, between 2012 and 2019.....	223
Table 6.S3 - Summary results of univariate analysis screening for the presence of linear or quadratic relations between variables describing fish community dynamics and environmental variables. ....	224

Table 6.S4 - Ranking of alternative candidate models relating the directionality of community change to environmental variables..... 225

Table 6.S5 - Ranking of alternative candidate models relating the mean velocity of community change to environmental variables..... 227



## List of figures

Figure 1.1 - Visualisation of the distribution of water on Earth. ....	13
Figure 1.2 - Similarities between survival analysis and time to first detection. ....	18
Figure 1.3 - Different distances used to compute spatial autocorrelation in stream networks. .....	21
Figure 1.4 - Hypothetical trajectories reflecting the temporal dynamics of a community in a multivariate species-space. ....	23
Figure 2.1 – Study Area .....	42
Figure 2.2 Invasive crayfish species present in the Sabor river catchment.....	45
Figure 2.3 – Native fish fauna of the Sabor river basin. ....	47
Figure 2.4 - Exotic fish fauna of the Sabor river basin. ....	49
Figure 2.5 - Amphibians associated with streams of the Sabor river region.....	55
Figure 3.1 - Variation in median times to first detection of each species as a function of stream depth and width.....	79
Figure 3.2 - Predicted occupancy probabilities of six fish species across the river Sabor catchment .....	80

Figure 3.S1 - Map of the study area, showing the location of sites visited in the summer of 2012..... 91

Figure 3.S2 - Relationships estimated from hierarchical occupancy-detection models between environmental variables and both the probability of water and fish species presence. .... 102

Figure 3.S3 - Maps of model prediction uncertainty for predicted distributions of six stream fish species in the Sabor catchment. .... 103

Figure 4.1- Map of the Sabor watershed and the location of the sites where each crayfish species was detected..... 115

Figure 4.2 - Torgegrams describing spatial dependencies in the distribution of *Procambarus clarkii* and *Pacifastacus leniusculus* along the stream network of the Sabor watershed. ... 118

Figure 4.3 - Response curves estimated from a logistic mixed model relating the probability of occurrence of *Pacifastacus leniusculus* to environmental variables..... 121

Figure 4.4 - Maps of potential distribution of *Procambarus clarkii* and *Pacifastacus leniusculus* in the river Sabor watershed. .... 123

Figure 5.1 - Maps showing the observed presences/absences of four amphibian species in the river Sabor watershed in 2015..... 158

Figure 5.2 - Maps of potential distributions of four amphibian species under the worst case scenario of future crayfish expansion in the river Sabor watershed. .... 159

Figure 5.S1 - Map of the study area showing The Sabor watershed and the sites sampled for amphibians in January-December 2015..... 187

Figure 5.S2 - Receiver operating characteristic curve and the area under the curve values for the geostatistical model for *Procambarus clarkii* and *Pacifastacus leniusculus*. .... 188

Figure 5.S3 - Torgegrams showing spatial dependencies in the residuals of models relating the probability of amphibian occurrence to variables reflecting environmental factors and biotic interactions ..... 189



Figure 5.S4 - Predicted distribution of <i>Procambarus clarkii</i> across the Sabor watershed at present and under different invasion scenarios. ....	190
Figure 5.S5 - Predicted distribution of <i>Pacifastacus leniusculus</i> across the Sabor watershed at present and under different invasion scenarios. ....	191
Figure 6.1 - Map of sites sampled for fish in 2012-2019 in the Sabor watershed.....	200
Figure 6.2 - Temporal variation in annual fish catches in the Sabor watershed. ....	204
Figure 6.3 - Temporal fish community trajectories at 30 sites sampled in the Sabor watershed in the period 2012-2019, represented in the first two axis of a Principal Coordinate Analysis. ....	207
Figure 6.4 - Trend lines describing the relations inferred from models relating the directionality and the mean velocity of community change.....	208
Figure 6.5 - Maps showing variation across the Sabor watershed of the directionality and mean velocity of fish community change.....	210
Figure 6.S1 - Mantel correlograms showing variations in the correlations between pairwise distances in community trajectories and spatial distances, considering either Euclidean or hydrologic distances.....	228
Figure 6.S2 - Mantel correlograms showing variations in the correlations between pairwise environmental and spatial distances between sites, considering either Euclidean or hydrologic distances.	
Figure 7.1 - Structure for the hierarchical model developed for chapter 3.....	235



## Resumo

A biodiversidade não está uniformemente distribuída pelo nosso planeta. Os ecossistemas de água doce possuem uma quantidade desproporcionada de biodiversidade quando comparados com outros biomas, embora cubram apenas uma pequena porção da superfície do nosso planeta. Sendo a água essencial para atividades humanas, o crescimento populacional e o desenvolvimento económico exercem uma enorme pressão sobre os ecossistemas de água doce. Para além das pressões antropogénicas diretas, tais como a extração excessiva, a construção de barragens, a modificação do habitat e a poluição, as espécies de água doce apresentam geralmente distribuições restritas a uma bacia hidrográfica ou a uma região, e enfrentam a ameaça de centenas de espécies invasoras que foram introduzidas nos ecossistemas de água doce. Devido a todos estes fatores, a biodiversidade dos ecossistemas de água doce é das mais ameaçadas do nosso planeta.

As redes hidrográficas merecem atenção especial porque são habitats particularmente ameaçados e ricos em biodiversidade. As redes hidrográficas são corpos lineares de água com uma organização dendrítica, ou em forma de árvore, que flui das cabeceiras até à foz do rio. A distribuição de organismos em redes hidrográficas não é aleatória, resultando de vários processos que funcionam a diferentes escalas, como o clima, a hidrologia, ou interações bióticas. A diversidade e abundância de peixes e outros organismos estão geralmente associadas a um aumento da ordem dos rios, mas há também processos de dispersão que devem ser tidos em conta. A distribuição de algumas espécies, como as espécies invasoras, é frequentemente mais um reflexo de processos espaciais, tais como múltiplas introduções e expansão posterior, do que de filtros ambientais que limitam a distribuição. Para espécies adaptadas à alternância entre fases de secas e inundadas, as cabeceiras dos rios podem funcionar como refúgios para interações bióticas adversas. As comunidades ribeirinhas, tal como os peixes, podem persistir no tempo num estado de equilíbrio dinâmico que varia entre estados alternativos, sem qualquer direção discernível de mudança. Desvios a este equilíbrio podem indicar perturbações na comunidade de origem natural, como secas ou inundações, ou de origem antropogénica.

Os modelos de distribuição de espécies (em inglês Species Distribution Models - SDM) são um conjunto de ferramentas utilizadas para obter previsões espacialmente explícitas de adequação ambiental, relacionando as ocorrências das espécies com dados ambientais

relevantes. Devido à sua natureza e à natureza das redes hidrográficas, o desenvolvimento de SDM para organismos que estão associados a habitats ribeirinhos é um desafio. Os organismos aquáticos raramente estão disponíveis para observação direta, e mesmo com a ajuda de técnicas padrão, como a pesca elétrica, podemos assumir que não conseguimos detetar todas as espécies presentes num determinado local. Esta questão, conhecida como detetabilidade imperfeita, é uma fonte comum de viés nos SDM e tende a ser ignorada pelos investigadores de água doce. Também, incorporar a autocorrelação espacial (em inglês spatial autocorrelation - SAC) melhora o desempenho dos SDM, mas as estruturas dendríticas das redes hidrográficas, juntamente com fortes gradientes ambientais, criam dependências espaciais com estruturas complexas que não são completamente descritas por distâncias euclidianas entre observações. As interações bióticas, tais como competição e/ou predação, podem ser uma potencial fonte de discrepância entre a distribuição real e a prevista das espécies, particularmente se as interações forem entre espécies invasoras e espécies nativas. A monitorização a longo prazo das comunidades é essencial para compreender o impacto das pressões antropogénicas nos ecossistemas ribeirinhos, mas as observações são geralmente recolhidas em locais espacialmente discretos. Uma visão espacial contínua da dinâmica temporal das comunidades é essencial para estudar as pressões antropogénicas e para planear ações de conservação e gestão.

O principal objetivo desta tese é desenvolver e testar novas ferramentas de modelação estatística para ajudar ecologistas e conservacionistas a obter uma representação mais realista da distribuição das espécies e da dinâmica temporal das comunidades à escala da paisagem fluvial. Concentrei-me principalmente nas soluções para as questões relacionadas com o problema da detetabilidade imperfeita, incorporação de SAC em SDM aplicados a redes hidrográficas, na utilização de interações bióticas em SDM, e na extrapolação de dinâmicas temporais de comunidades para uma previsão espacial contínua à escala da rede hidrográfica. Para abordar estas questões, recolhi dados sobre a distribuição de peixes, lagostins e anfíbios num sistema de estudo específico, uma bacia hidrográfica Mediterrânica no Nordeste de Portugal, o rio Sabor.

Para descrever a distribuição das espécies de peixe com dados recolhidos num grande trabalho de pesca elétrica no rio Sabor em 2012, desenvolvemos um modelo hierárquico Bayesiano em que modelámos a probabilidade de presença de água em segmentos de rio, e a probabilidade de ocupação de espécies condicionada à presença de água, em relação a variáveis ambientais e espaciais de grande escala. Também modelámos o tempo à primeira deteção das espécies, utilizando um modelo de sobrevivência com distribuição exponencial. Como é difícil determinar o tempo exato de deteção de uma espécie quando se procede à

amostragem de peixes com a técnica de pesca elétrica, modificámos o modelo de sobrevivência para incluir observações censuradas em intervalos. Para incorporar as dependências espaciais, incluímos um termo espacial de autocovariância na estimativa da probabilidade de presença de água e da probabilidade de ocupação das espécies. As ocupações das espécies foram consistentemente afetadas pela ordem do rio, elevação, e precipitação anual, enquanto o tempo de deteção de espécies foi influenciada principalmente pela profundidade e, em menor medida, pela largura dos cursos de água.

O pressuposto de equilíbrio entre os organismos e o seu ambiente é um postulado padrão de trabalho com SDM que raramente é cumprido, particularmente para espécies que estão em processo de expansão como as espécies invasoras. Além disso, para espécies que invadem os sistemas fluviais, a estrutura dendrítica da rede hidrográfica limita os padrões da expansão. Para descrever a distribuição de duas espécies de lagostins invasores no rio Sabor, utilizámos uma classe de modelos geoestatísticos desenvolvidos para lidar com SAC em redes hidrográficas (em inglês, Spatial Stream Network Models - SSNM). Estes modelos obtiveram melhor performance do que modelos puramente ambientais, evidenciando que a distribuição destes lagostins invasores é mais um produto de processos espaciais do que de filtragem ambiental.

As interações biológicas são fatores importantes que moldam a distribuição das espécies. Quando as espécies nativas são deslocadas de parte da sua área de distribuição, podem persistir em refúgios ecológicos. Estes refúgios podem ser manchas de habitat inadequadas para espécies invasoras, ou áreas que as espécies invasoras ainda não colonizaram devido à distância, barreiras físicas ou atraso na expansão. Identificar a localização e os fatores que influenciam a distribuição destes refúgios é importante para conservação destas espécies. Para modelar a distribuição dos refúgios ecológicos de anfíbios na bacia hidrográfica do rio Sabor, incluímos como variáveis preditoras a probabilidade de presença dos dois lagostins invasores entre outros preditores ambientais e espaciais. Verificámos que os refúgios de anfíbios estão localizados principalmente nas cabeceiras do rio Sabor, sendo provável que estes refúgios contraiam no futuro em função de cenários plausíveis de expansão dos lagostins invasores.

A gestão dos ecossistemas fluviais é geralmente planeada à escala da bacia hidrográfica e, como tal, é importante desenvolver ferramentas que permitam a extrapolação das dinâmicas das comunidades que são observadas em segmentos discretos de rios para uma previsão espacial contínua de toda a rede hidrográfica. Para analisar dados de abundância de peixes recolhidos em 30 locais na bacia do rio Sabor entre 2012 e 2019, utilizámos uma nova abordagem que descreve e compara a dinâmica das comunidades usando as propriedades

geométricas das suas trajetórias num espaço escolhido de dissemelhança. Calculámos a velocidade média e a direccionalidade global de mudança das comunidades de peixes e utilizámos modelos SSNM para relacionar estas métricas com os fatores ambientais e extrapolar a dinâmica da comunidade para toda a bacia hidrográfica. Não encontramos qualquer evidência de direccionalidade na mudança das comunidades de peixes do Sabor, apoiando a hipótese de que estas existem num estado de equilíbrio frouxo (em inglês, Loose Equilibrium). No entanto, a velocidade de mudança foi maior nas linhas de água que correm para a albufeira do aproveitamento hidroelétrico localizado perto da foz do rio Sabor. Estes cursos de água estão provavelmente sob uma pressão crescente a partir da albufeira, devido a alterações do regime hídrico e expansão de espécies alóctones.

De um modo geral, esta tese faz avançar a nossa compreensão dos fatores que regulam a distribuição das espécies e comunidades em redes hidrográficas, o que constitui um conhecimento essencial para a conservação destes ecossistemas. O novo conjunto de ferramentas apresentado pode ajudar ecologistas e conservacionistas a obter uma representação mais realista da distribuição da biodiversidade e da sua dinâmica temporal à escala da paisagem fluvial.

### Palavras chave:

anfíbios, análise de trajetória de comunidades, autocorrelação espacial, deteção imperfeita, dinâmica de comunidades, ecologia de rios, geoestatística, invasões biológicas, lagostins invasores, modelo de deteção-ocupação, modelos bayesianos hierárquicos, modelos de distribuição de espécies, peixes, redes ecológicas dendríticas, tempo à primeira deteção

## Abstract

Biodiversity is not evenly distributed across our planet. Freshwater ecosystems hold a disproportionate amount of biodiversity when compared with other biomes, though it only covers a small portion of our planet surface. Because water is essential for human activities, population growth and economic development put an enormous pressure on freshwater ecosystems. Besides the direct anthropogenic pressures, such as over exploitation, damming, habitat modification and pollution, the freshwater species usually present restricted distributions, to a watershed or a region, and face the threat of hundreds of invasive species that have been introduced to freshwater ecosystems. Due to all these factors, freshwater biodiversity is among the most threatened of our planet.

Stream networks deserve special attention because they are particularly threatened and rich in biodiversity. Streams networks are linear bodies of water with a dendritic, or tree shape, organization, flowing from the headwaters to a single outlet. The distribution of organisms in stream networks are not random, resulting from several processes that work at different scales, like climate, hydrology and biotic interactions. The diversity and abundance of fish and many other organisms are usually associated with an increase in stream order, but there are also dispersal processes that should be taken into account. The distribution of some species, like invasive species, is often more a reflection of spatial processes, such as multiple introductions and posterior expansion, than environmental filters that limit the distribution. Headwaters can function as refuges from adverse biotic interactions for species that support water intermittency. Stream communities, like fish, usually persist in time in a state of dynamic equilibrium, varying between alternate states with no discernible direction of change. Deviations from this equilibrium may reflect disturbances to the community from natural states, like droughts or floods, or from anthropogenic sources. For proper conservation and management of stream networks, it is essential to understand the drivers of the spatial patterns and dynamics of stream biodiversity.

Species distribution models (SDM's) are the set of tools used to derive spatially-explicit predictions of environmental suitability, by relating species occurrences to relevant environmental data. Due to their nature and the nature of the stream network habitats, the development of SDM's for organisms that are associated with streams is challenging. Aquatic organisms are rarely available for direct observation, and even with the help of standard

techniques, like electrofishing, it is fair to assume that we will fail to detect some of the species present at any given location. This issue, known as imperfect detectability, is a common source of bias in SDM's, and tends to be ignored by freshwater researchers. Accounting for spatial autocorrelation (SAC) improves SDM performance, but the dendritic structure of stream networks, together with strong environmental gradients, create spatial dependences with complex structures that are not completely described by Euclidean distances. Biotic interactions, such as competition or predation, are also a potential source of mismatch between the actual and the predicted distribution of species, particularly if the interactions are between invasive species and native species. Long term monitoring of communities is essential to understand the impact of anthropogenic pressures in stream ecosystems, but usually rely on data collected on any given number of discrete locations. A spatial continuous view of the temporal dynamics would be essential to study such pressures, and of value to plan conservation and management actions.

The main aim of this thesis is to develop new tools and frameworks to help ecologists and conservationists to obtain a more realistic depiction of the distribution of species, and the temporal dynamics of communities at the riverscape scale. I mainly focused on solutions to the issues related to dealing with imperfect detectability, accounting for SAC in stream networks, accounting for biotic interactions, and extrapolating the community temporal dynamics to a continuous spatial prediction. To address these issues, I have collected data on the distribution of fish, crayfish, and amphibians on a specific study system, the Sabor River, a Mediterranean watershed in the Northeast of Portugal.

To describe the distribution of fish species with data collected in a comprehensive electrofishing survey, while accounting for imperfect detectability, we extended the time-to-detection occupancy-detection model to include interval-censored observations, because it is difficult to ascertain the exact time-to-detection of a species when sampling fish with electrofishing techniques. Using a Bayesian hierarchical framework, we modelled the probability of water presence in stream segments, and the probability of species occupancy conditional on water presence, in relation to environmental and spatial variables. We also modelled time-to-first detection conditional on occupancy in relation to local factors, using a modified interval-censored exponential survival models. To account for SAC, we included a spatial autocovariate term in the estimation of the probability of water presence and the probability of species occupancy. Species occupancies were consistently affected by stream order, elevation and annual precipitation, while species detection rate was primarily influenced by depth and, to a lesser extent, stream width.



The assumption of equilibrium between organisms and their environment is a standard working postulate in SDM's that is seldom met, particularly for species that are expanding their range like invasive species. Furthermore, for species invading river systems, the dendritic structure of the stream network will constrain the patterns of the expansion. In this thesis, I addressed these issues by describing the distribution of two invasive crayfish in the Sabor river stream network, using a class of geostatistical models developed to deal with SAC in stream networks, known as spatial stream network models (SSNM). Accounting for SAC greatly improved model performance, evidencing that the distribution of these invasive crayfish was more of a product of spatial process than environmental filtering.

Biotic interactions are important drivers of species distributions. When native species are displaced from part of their distributional range, they may persist in ecological refuges. These refuges may be patches of habitat that are unsuitable for invasive species or areas where invasive species have not reached due to distance, physical barriers or time lags in the expansion. Identifying the distribution and the environmental drivers of these refuges is of conservation concern. We modelled the distribution of amphibian ecological refuges in the Sabor river catchment, by including as predictor variables the probability of presence of the two invasive crayfish, among other environmental and spatial predictors. We found that the refuges of amphibians are located mainly in the headwaters, and that, under plausible expansion scenarios of the crayfish species, these refuges are likely to contract in the future.

Management of stream networks is usually planned at the river basin scale, and as such, it is important to develop frameworks that allow the extrapolation of the community dynamics observed at discrete segments of rivers to a continuous spatial view of the entire river basin. We collected stream fish data on 30 locations on the Sabor river basin, between 2012 and 2019, and used a novel framework to describe and compare the trajectories of the fish communities using their geometric properties in a given dissimilarity space. We computed the mean velocity and the overall directionality of change of the fish community, and used the SSNM framework to relate these metrics to environmental drivers and extrapolate the community dynamics to the entire watershed. We found no evidence of directionality in the change of the Sabor fish communities, supporting the hypothesis that these communities exist in a loose equilibrium state. However, the rate of change was higher in streams draining into the hydroelectric reservoir located near the mouth of the Sabor River. These streams are likely under increased stress from the reservoir, due to alterations of the flow regime and/or expansion of alien species from the reservoir.

Overall this thesis advances our understanding of the drivers that govern the distribution of species in stream networks, providing key information for the conservation of these

ecosystems. The new set of tools presented here can aid ecologists and conservationists to obtain a more realistic depiction of species distribution and their temporal dynamics at the riverscape scale.

### Keywords:

amphibians, biological invasions, community dynamics, community trajectory analysis, dendritic ecological networks, geostatistics, hierarchical Bayesian models, imperfect detection, invasive crayfish, occupancy-detection modelling, spatial autocorrelation, species distribution models, stream ecology, stream fish, time-to-first-detection

## Chapter 1 – General Introduction





# Chapter 1 – General Introduction

## 1.1 Biodiversity and its decline

Biodiversity is a term that entered the lexicon of biologists by the second half of the 1980's decade (Tangley, 1985; Wilson, 1988), and achieved widespread use in academia (Harper & Hawksworth, 1994) and in our societies by the turn of the century. The term is the contraction of "Biological Diversity" and is attributed by Wilson (1988) to Walter G. Rosen. There are several definitions (Wilcox, 1984; Wilson, 1988) but they can be summed by Gaston & Spicer (2004) definition: "variation of life at all levels of biological organization".

Biodiversity as the variation of life at all levels of biological organization, refers not only to the total number of life forms across an area, but also to the range of differences between those forms (DeLong, 1996). It can be seen as the combined evolutionary capital amassed by all species since the beginning of life at least 3500 million years ago (Schopf, 2006). Every successful solution to each evolutionary challenge was imprinted in the DNA of the survivors and was passed onto the next generations. We depend directly on biodiversity for food, but also for regulating the processes that recycle the air we breathe and the water we drink. Several human activities are directly dependent on biodiversity, like agriculture, forestry, fishing, and hunting, while the world's economy is based on fuels that resulted from the fossilization of living organisms (Guo et al., 2010; Isbell et al., 2017; McNeely, 1994; Scherr & McNeely, 2008). Several of our technological advances are drawn from biological systems (biomimicry), being active principles for medicine, vehicle design, textiles, and others (Benyus, 1997). The squander of such resources, besides limiting new technological advances, hinder our survival as species. Biodiversity is correlated with the stability of ecosystems and loss of biodiversity is related to a decrease in the efficiency of ecosystems processes (Cardinale et al., 2012).

The decline in biodiversity, as well as habitat destruction, was acknowledged even before the term was coined (Wilson, 1985). Nowadays, we witness a rate of species extinction that could be up to 100 times steeper than the baseline rate of species extinctions, indicating that we could have entered the sixth mass extinction event (Ceballos et al., 2015). Human activities are the main cause of biodiversity loss (Dirzo et al., 2014; Ceballos et al., 2015), mainly due to habitat alteration and destruction. Other major causes are direct exploitation of organisms (harvesting, hunting, fishing), pollution, the

introduction of alien species, emerging diseases, and climate change. These factors have a negative impact on their own. but are particularly serious when acting simultaneously. Therefore, understanding and finding solutions to the decline of biodiversity is a critical challenge facing humanity.

## 1.2 The freshwater biodiversity crisis

Biodiversity is not evenly distributed across the planet. Freshwater ecosystems hold only 0.01% of the world's water and cover only about 0.8 % of the planet surface (Gleick, 2014), but provide habitats for about 6% of total species and a third of the vertebrate species (Dudgeon et al., 2006). There are over 126 000 animal species belonging to 17 phyla described for freshwater ecosystems (Balian et al., 2008), the majority of which are insects (60.4%), followed by vertebrates (14,5%) and crustaceans (10%). The number of strictly freshwater fish species currently known is around 13 000, corresponding to about 45% of the total fish species recorded (Lévêque et al., 2008). Amphibians are highly dependent on water to complete their life cycle and, and as such, the majority (73%) are dependent on freshwater habitats (Balian et al., 2008). The number of endangered species in freshwater is disproportionately high when comparing with land and marine environments (Strayer & Dudgeon, 2010). This is the result of external threats and the fact that many species of freshwater are endemic to certain watershed or regions, often presenting small distribution ranges (Balian et al., 2008). The insular nature of freshwater habitats results in a high level of fragmentation and endemism, and low ability to migrate, making these species very sensitive to human impacts and climate changes (Strayer & Dudgeon, 2010).

Water is essential for human activities, and following population growth and economic development, water usage has increased by a factor of six in the last 100 years (UNESCO World Water Assessment Programme, 2020), putting enormous pressure on freshwater ecosystems. Freshwaters hold just 0.01% of the world's water (Gleick, 2014), but are the bulk of the water that is available for human activities, with mankind already using uses more than half of the accessible runoff water (Vitousek, 1997). The overexploitation of a habitat that is concentrated in less than 1% of the world surface (Figure 1.1, Gleick, 2014) makes freshwater ecosystems one of the most threatened ecosystems of the world (Dudgeon et al., 2006; Sala et al., 2000; Strayer & Dudgeon, 2010). Nowadays, more than 58000 large dams (International Commission on Large Dams [ICOLD], 2020) and over a million smaller impoundments fragment river networks (Jackson et al., 2001). There are several impacts of dams and reservoirs in river basins:

interruption the river continuum, blocking fish migration and altering sediment and nutrient transport; Siltation and habitat homogenization; Downstream flow and water quality alteration (Schmutz & Moog, 2018). Rivers have also been confined by levees, dredged and straightened for navigation and flood control (Strayer & Dudgeon, 2010). Alterations to riparian zones alter the inputs of water and materials and may cause water eutrophication, which together with pollution can have big impacts on biotic communities of freshwaters. Several hundreds of species have been introduced in freshwater ecosystems (Strayer, 2010) and freshwater fisheries are seriously overexploited (Dudgeon et al., 2006; Strayer & Dudgeon, 2010). It is predicted that climate change, with drought episodes becoming more frequent and more intense, will exacerbate the human pressure on freshwater habitats (UNESCO World Water Assessment Programme, 2020). Therefore, understanding and conserving the biodiversity of freshwater has become a top priority in recent years.

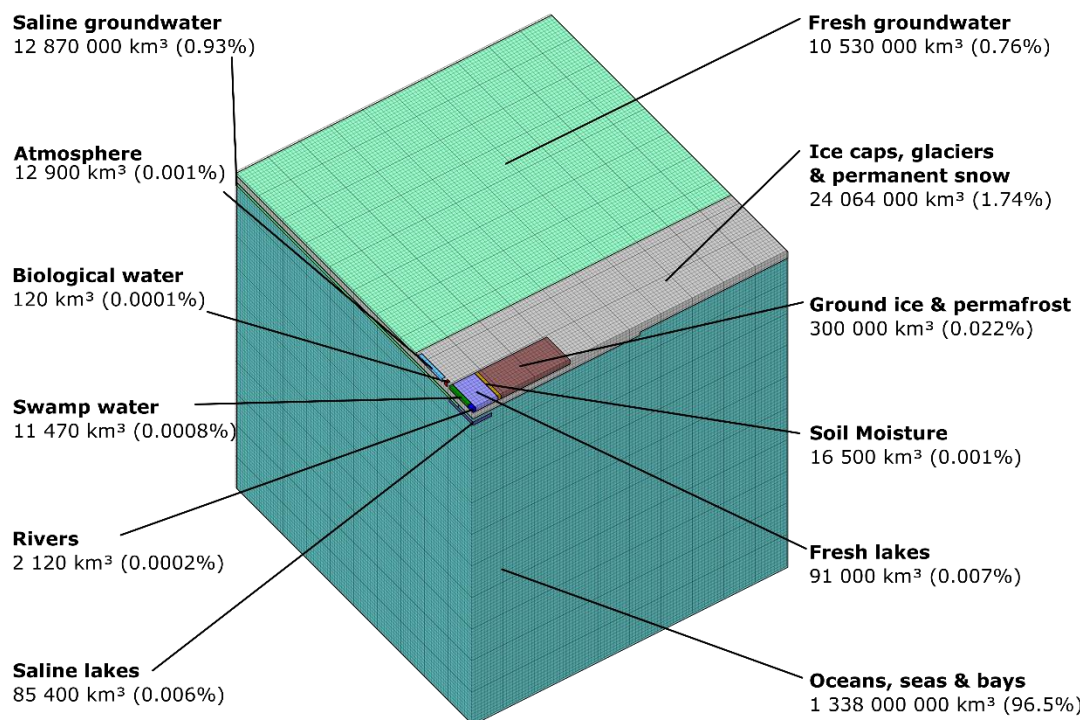


Figure 1.1 - Visualisation of the distribution (by volume) of water on Earth. Each cube (as the one representing biological water) corresponds to approximately 1400 cubic km of water. The figure comprises 1 million small blocks. Adapted from Cmglee (2020) with data taken from Shklomanov (1993).

### 1.3 Stream Networks

In the context of biodiversity research in freshwater ecosystems, running waters (streams, rivers) deserve special attention, as they are both particularly threatened and particularly rich in aquatic organisms (Dudgeon et al., 2006; Sala et al., 2000; Strayer & Dudgeon, 2010). When precipitation exceeds evaporation, or evapotranspiration, the residual water either infiltrates into the ground or runs superficially to the ocean, lakes or inland basins (Dai & Trenberth, 2002). The running superficial water gathers downhill, forming linear bodies of flowing water known as streams. Small streams join other streams becoming larger streams, that eventually become known as rivers flowing to an ocean or an endorheic basin. The structure of a stream network can be described as dendritic, tree-shaped, and is characterized by flow directionality, from headwaters to a single downstream outlet (Peterson et al., 2013). Although rivers worldwide only hold 0.0002% of the total water and 0.006% of the freshwater of the world (about 2 120 km<sup>3</sup>, Figure 1.1, Shiklomanov, 1993) at any given time, the volume of water that passes through rivers annually is far greater. Each year, rivers worldwide discharge about 37 288 km<sup>3</sup> of freshwater to the oceans, which corresponds to about 35% of the inland precipitation (Dai & Trenberth, 2002). The interannual variation of precipitation causes temporal and spatial variation in the flow, with streams in arid and semiarid zones showing no superficial flow during at least part of the year and in at least some sectors of the watershed. (Cid et al., 2017). For instance, in Mediterranean river basins the small headwater streams of first and second order can be completely dry in the hot summer period (Kerezszy et al., 2017; Magalhães et al., 2007).

From headwaters to mouth, stream networks form a continuous biophysical system. The biological communities within stream networks respond to the change of continuous physical variables, as stream width and depth, as well as inputs and transportation of matter and energy (Vannote et al., 1980). Headwaters present a predictable shift between wet and dry phases that harbour a rich and diverse community (Meyer et al., 2007). Species that dwell in headwater have developed morphological, physiological, and behavioural adaptations to cope with the intermittency of flow (Sánchez-Montoya et al., 2017). Headwaters can function as refuges from harsher environmental conditions or adverse biotic relations like competitors, predators that could be invasive species (Meyer et al., 2007).

While headwater communities rely on allochthonous organic matter inputs from the terrestrial environment, such as leaves and other detritus that are washed to the stream,



and communities of medium rivers rely on autochthonous production from algae and aquatic plants, downstream communities in large rivers are adapted to capitalize on the organic matter that leaked and is transported through the stream network with little or no production of their own (Vannote et al., 1980). The large rivers usually present more species of fish than headwaters and can enhance species richness in nearby tributaries, regardless of their size, through mass effects and dispersal processes (Erös, 2017; Matthews, 1998). Large rivers are also more likely to be affected by human activities, such as damming or the introduction of exotic species (Johnson et al., 1995).

Stream networks are more than ribbons of linear bodies of water. In river catchments, usually, they present a large network of heterogeneous habitats that are connected by complex dendritic structure creating strong environmental gradients over a range of spatial scales (Domisch et al., 2015; Erös & Schmera, 2010; Fausch et al., 2002; Peterson et al., 2013; Vannote et al., 1980; Ward & Stanford, 1995). This complexity makes the study of the distribution of organisms in streams challenging, thus adding to relevance of this thesis.

## 1.4 Biodiversity modelling in aquatic systems

### 1.4.1 Species distribution models

#### 1.4.1.1 The basic approach

An understanding of the drivers governing variation of biodiversity distribution in space and time is key to freshwater ecosystems conservation (Geist, 2011). Freshwater ecosystems, such as stream networks, are complex systems for which our observations and knowledge could be insufficient to establish casual relationships (Wang & Grant, 2019b). Therefore, models should be regarded as a scientific process that is used to structure our observations, knowledge and assumptions, using an abstraction of reality with a formal description of the essential elements and their relations, to enhance our understanding of a system (Wang & Grant, 2019a, 2019b).

Species distribution models (SDM's) is the set of statistical tools that use occurrence or abundance records of any given species from field samples, museums, literature, and/or on-line databases, to derive spatial explicitly predictions of environmental suitability for species in a given range and time period (Guisan et al., 2013; Guisan & Thuiller, 2005). Usually, this is achieved by establishing species-environmental relations using relevant environmental data from the locations where the species was recorded and, desirably,

from locations where the species is presumably absent. Once the data is compiled and a model fitted, assuming that the model is valid, it is possible to generate predictions and then apply them to inform conservation planning and management (Guisan & Thuiller, 2005; Rodríguez et al., 2007).

There have been several applications of SDM's for conservation biology: improving the maps in atlas (Bustamante & Seoane, 2004); detecting potential suitable habitat and populations in locations without data (Jarnevich & Reynolds, 2011; Seoane et al., 2003); predicting future range shifts due to climate change (Araújo et al., 2006); identifying species that are underrepresented in the current network of protected areas using gap analysis (Araújo et al., 2007; Hermoso et al., 2016); designing reserves (Dobson et al., 2013); predicting areas of conflict between human activities and wildlife, like agriculture (Ficetola et al., 2014) and transportation (Santos et al., 2018); identifying barriers to species dispersion (Meixler et al., 2009); selecting areas for species reintroduction (Olsson & Rogers, 2009); and habitat restoration (Adame et al., 2015).

The use of SDM's has also been valuable for understanding the impact of biological invasions (Franklin, 2013; Guisan et al., 2013; Rodríguez et al., 2007; Srivastava et al., 2019), as the dispersion of alien species by human activities confuses biogeographical patterns (Capinha et al., 2015). SDM's have been used to identify areas under the risk of biological invasions (Capinha et al., 2011; Hortal et al., 2010; Jiménez-Valverde et al., 2011); predict the distribution of invasive species in newly invaded areas (Capinha & Anastácio, 2011); assess the factors associated with colonization success (Capinha et al., 2012); measure the impact on native species (Ficetola et al., 2011); and select possible areas for conservation actions (Préau et al., 2020).

Although there has been an increase in the need and use of SDM's to support conservation planning and decision-making, actions are most often taken without the support of these tools (Guisan et al., 2013). Their difficult adoption by practitioners, besides the relative complexity of the statistical skills needed, maybe is related to a misplaced notion that the theoretical assumptions behind SDM's are unrealistic (Araújo & Peterson, 2012). SDM's have limitations and assumptions that have been extensively discussed, such as niche conservatism of species (i.e., the lack of consideration of the adaptation potential of species), the static nature of the models concerning the biotic and abiotic interactions, and the tendency of SDMs to over-simplify species-environment relationships (Araújo & Guisan, 2006; Araújo & Peterson, 2012; Guisan et al., 2013; Guisan & Thuiller, 2005). When developing SDM's, one should address the following critical issues: choose the adequate statistical framework and adopt the best standards;

use an appropriate spatial and temporal scale that match the modelling purpose and fit the available data; choose sound and ecological meaningful predictors; and collect enough data to cover the environmental gradients (Domisch et al., 2015).

#### 1.4.1.2 The problem of imperfect detection

The failure to detect a species when a species is present is a very common source of bias in datasets used to produce SDM's (Dorazio, 2012; Guillera-Arroita et al., 2014; M. Kéry & Schmidt, 2008; Marc Kéry, 2011; Lahoz-Monfort et al., 2014; Lobo et al., 2010). When conducting field surveys, non-detecting a species when a species is present in a habitat patch, sampling unit, etc., could be a result of several factors: the species of interest may be notorious cryptic and hard to find due to a small size, cryptic habits, or camouflage (Mazerolle et al., 2007); the efficiency of the sampling protocol (Einoder et al., 2018); and the habitat characteristics (Ferreira & Beja, 2013). This issue is known as imperfect detectability and results from the fact that the probability of detecting a species, when it is present in a patch of habitat or sampling unit, is usually less than one (MacKenzie et al., 2002).

To account for imperfect detectability MacKenzie et al. (2002) developed a new class of models that jointly models the ecological process of the species presence (occurrence) and observational process (detectability). These models are known as "site occupancy" or "occupancy-detection" models, and normally rely on replicated surveys conducted at a subset of sampling units (MacKenzie et al., 2006). These replicated surveys can be made by visiting sites more than once, or in a single visit divided by time, observer or method, as long as each survey/observation remain independent of other observations (independence assumption), and that the state of the sampling unit (site) doesn't change between observations (closure assumption, i.e., a site remains occupied or vacant across all surveys; Guillera-Arroita et al., 2011; MacKenzie, 2006).

If undertaking repeated surveys is not an option, an alternative is to model the observations as a continuous process using the time needed to detect the first individual of a species (time-to-first detection, Garrard et al., 2008). Time-to-first detection is a decreasing function of detectability and is known to be affected by factors such as species abundance, species traits, and sampling efficiency (Bornand et al., 2014; Garrard et al., 2013; McCarthy et al., 2013). Time-to-first detection is modelled as a parametric survival model (Kleinbaum & Klein, 2012). The rationale being that if a species is present at a site, finding an individual is a matter of time, the more time spent

the higher the probability to find the species. When sampling is halted, usually due to logistic constraints, it's possible to estimate the probability of a failed detection (Figure 1.2).

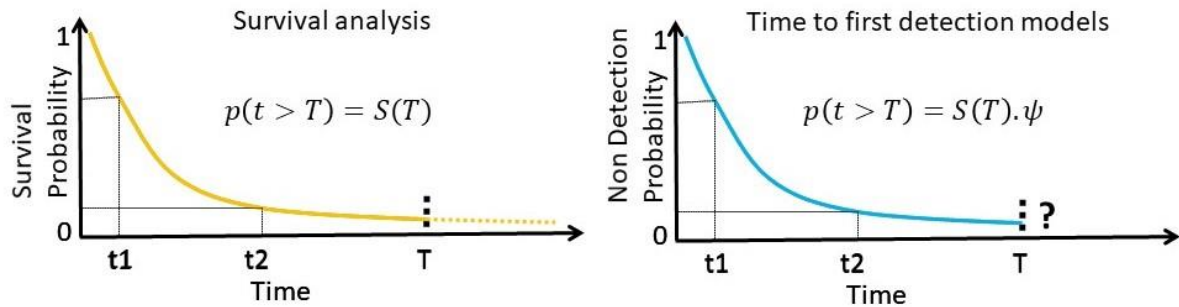


Figure 1.2 - Similarities between survival analysis and time to first detection. In survival analysis the event of interest is referred as failure. The probability of the failure occurring after time  $T$ , i.e. the failure occurring at time  $t > T$ , is given by the survival probability for time  $T$ . In time to first detection models, the probability of species being present given that it was not detect until time  $T$ , i.e. it would be detect at time  $t > T$ , is given by survival function for time  $T$  multiplied by the occupancy probability ( $\psi$ ).

### 1.4.1.3 Accounting for biotic interactions

Another potential source of mismatch between the actual and predicted distribution of a species is related to the effects of biotic interactions, which can contract the used niche in relation to the potential niche defined only by abiotic drivers. Biotic interactions can shape species distribution from local to continental scale (Araújo & Luoto, 2007; Wisz et al., 2013) in different ways: competition among species affect range limits as for instance in the case of mutual exclusion over large extents of the European ranges of hedgehog species (Hewitt, 1999); Predators, especially apex consumers, are key species in ecosystems and can have a strong influence on the abundance and distribution of their prey, and can also have cascading effects across the ecosystems (Estes et al., 2011); Invasive species can also be vectors of pathogens that once introduced in naïve populations can have a decimating effect, as for example, is the case of North American crayfish carrying subclinical infections of the oomycete *Aphanomyces astaci*, the agent of the crayfish plague, that have devastated the populations of European crayfish since the XIX century (Gherardi, 2007). If an ecological link between a pair species is known or suspected, the most straightforward approach to account for biotic interactions in SDM's is to use the distribution of one species and other abiotic predictors to predict the distribution of the second species, but many other approaches have been proposed (Wisz et al., 2013).

## 1.4.2 The challenges of distribution modelling in river networks

The distribution of freshwater species in streams results from top-down and bottom-up processes. Climate may be the main top-down driver of species distribution at continental or basin-scale with cascading effects to catchment scale and reach scale, but we may observe a mismatch of the predicted species distribution at more local scales due to habitat heterogeneity derived from local topography and hydrology, or human impact in the surrounding landscape or directly in the stream network (e.g.: impoundments). The choice of scale and resolution, as well as the environmental predictors, is not trivial and maybe a potential source bias (Domisch et al., 2015).

Riverine ecosystems and freshwater species have several properties that challenge the application of SDM's, thereby making it more difficult than in other systems to understand the patterns of species distribution (Olden et al., 2010). This is because stream ecosystems differ from their terrestrial and marine counterparts in fundamental aspects. Streams present a dendritic linear network of habitats connected laterally, longitudinally and vertically, with strong environmental gradients creating highly heterogeneous systems over a range of spatial scales (Domisch et al., 2015; Erös & Schmera, 2010; Fausch et al., 2002; Peterson et al., 2013; Vannote et al., 1980; Ward & Stanford, 1995). Moreover, while most fish and other aquatic species are strictly dependent on streams for habitat and dispersion, other organisms, although preferably using streams, can disperse overland like crayfish (Cruz & Rebelo, 2007), amphibians (Grant et al., 2010) and aquatic mammals (Quaglietta et al., 2014), further confusing patterns of species distribution in stream networks.

The multiple spatial influences affecting aquatic organisms needs to be incorporated in distribution models to correctly understand and predict their distributions, though the way to do it correctly differs greatly from that used in SDMs applied to terrestrial species. Ecological data usually shows spatial autocorrelation (SAC), due for instance to the presence of environmental gradients or habitat patches that affect species distributions (Legendre, 1993). SAC thus arises because nearby pairs of points tend to be more similar in physical conditions and present a similar species composition and abundance than pairs further apart (Legendre, 1993), which is a potential source of bias because it violates SDM's assumption of independence among observations (Diniz-Filho et al., 2003; Legendre, 1993; Record et al., 2013; Václavík & Meentemeyer, 2012). SDM's that do not account for SAC are more prone to type 1 errors (Diniz-Filho et al., 2003; Record

et al., 2013), i.e., failing to predict the presence of a species where it is present, while SDM's that do account for SAC usually have enhanced predictive performance (Václavík et al., 2012).

To incorporate SAC in generalised linear models (GLM's), it is necessary to relax the independence assumption by modelling the covariance matrix in the residuals as a function of nearby observations (Ver Hoef et al., 2018). Two common approaches that account for SAC are geostatistical and autoregressive models. In geostatistical models, SAC is modelled as a function of the distance between observations, while in autoregressive models SAC is modelled as conditional to the observations in a network of neighbours (Ver Hoef et al., 2018). Autoregressive models, like the autologistic models, include an autocovariate coding the influence of neighbouring locations and a corresponding parameter to estimate the strength of neighbourhood effects (Augustin et al., 1996; Bardos et al., 2015; Besag, 1974; Gumpertz et al., 1997; Hoeting et al., 2000). To construct the autocovariate, it is necessary to select the neighbourhood of each location, the set of locations that have influence on the value of the variable at the focal location, and weight the influence of each neighbour usually using distance. Weighted means are known to produce invalid weighting schemes, namely by not assuring symmetric weights among neighbours, and as such it is preferable to use weighted sums (Bardos et al., 2015).

Geostatistical models usually rely on Euclidean distances to account for SAC, but Euclidean distances alone are of little application when modelling species distributions that are constrained to streams like fish or aquatic macroinvertebrates (Peterson & Ver Hoef, 2010), or even species that move preferably along streams but may disperse across the land as amphibians (Grant et al., 2010) and semi-aquatic mammals (Quaglietta et al., 2014). To apply geostatistical models to streams networks, it is thus more appropriate to use hydrological (along the channel) distances (Isaak et al., 2014; Peterson et al., 2007). The hydrological distances can be symmetric, among all locations within the stream network, or asymmetric, if only the locations connect by flow are taken into consideration (Figure 1.3, Peterson et al., 2007). Spatial stream network models (SSNM) were proposed by Ver Hoef et al. (2006) as a framework to deal with SAC in stream networks that take into account the dendritic structure and the direction of flow. Similar to previous geostatistical models, SSNM can fit in the random component an autocovariance function based on Euclidean distances, which is useful to account for spatial patterns that arise from the linkage with the terrestrial landscape, including due to factors like climate, land cover or overland dispersal by semi-aquatic organisms (Isaak

et al., 2014). These models, however, can also include two autocovariance functions based on hydrological distances: a tail-up and a tail down model based on moving-average over a dendritic structure (Isaak et al., 2014; Peterson & Ver Hoef, 2010; Ver Hoef & Peterson, 2010). Tail-down models have the autocovariance function pointed downstream, and allow for spatial dependence among all locations in the stream network. Tail-up models have the autocovariance function pointed upstream, and only allow for spatial dependence among locations that are connected by flow. Spatial weights can be used to split the tail up autocovariance function at stream confluences using a relevant metric, like flow discharge or catchment area, to give a more accurate representation of tributary influence downstream (Isaak et al., 2014; Peterson & Ver Hoef, 2010). Tail-up models were designed to predict the distribution of organisms that move passively downstream while tail down are useful to predict organisms that actively colonise up-streams (Peterson & Ver Hoef, 2010). Autocovariance functions may be combined to create a flexible covariance structure that simultaneously account for several types of spatial relationships across several scales, and to assess the importance of each component as a function of the variance explained (Isaak et al., 2014; Peterson & Ver Hoef, 2010).

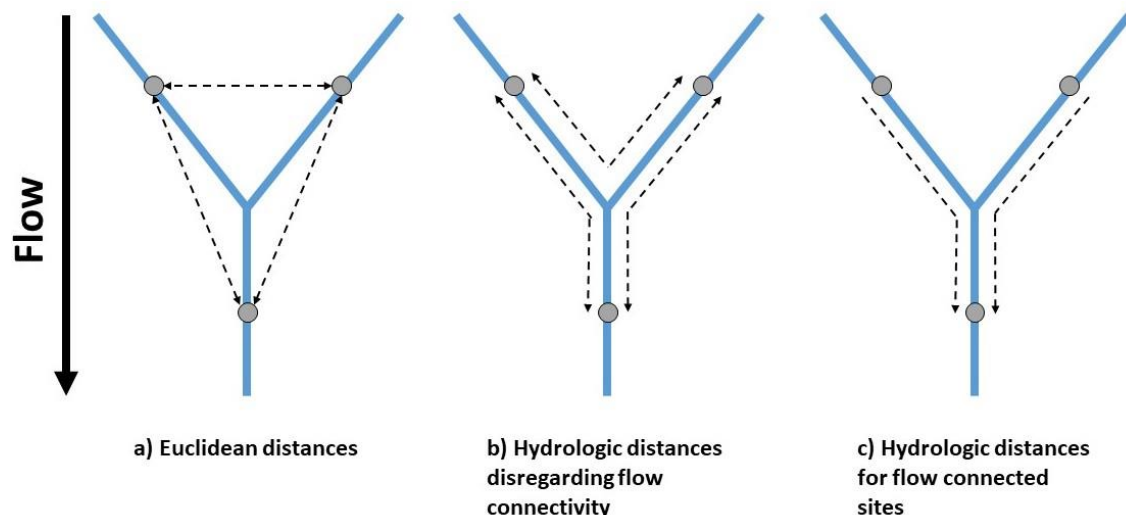


Figure 1.3 - Different distances used to compute spatial autocorrelation in stream networks. Distances a) and b) are symmetric and c) is asymmetric. Adapted from Peterson et al. (2007)

## 1.5 Modelling dynamic community patterns

Freshwater communities<sup>1</sup> are changing worldwide due to increasing anthropogenic pressures (Albert et al., 2020; Reid et al., 2019). Several studies have been devoted to understanding how and why freshwater communities change over time, but are usually based on data collected over an extended time period at a small number of discrete locations (e.g. Baranov et al., 2020; Bêche et al., 2009; Erős et al., 2020; Jourdan et al., 2018; Matthews & Marsh-Matthews, 2016). While long term data is essential for understanding community dynamics (Matthews & Marsh-Matthews, 2017), it is crucial to have a spatial continuous view of such dynamics in order to prioritize conservation actions (Cid et al., 2020; Fausch et al., 2002).

Changes in biological communities may result from local extinctions and/or colonizations that affect their composition and richness (Heino, 2013), but more usually they result from variations in species abundances (Grossman et al., 1990; Magalhães et al., 2007; Matthews & Marsh-Matthews, 2017). These temporal variations may be gradual or more abrupt (saltatory), and may alternate between alternative stable states, revolving around a loose equilibrium, or presenting a directional departure from the initial state as a result of a disturbance or ecological succession (Figure 1.4, Collins, 2000; DeAngelis et al., 1985; Matthews et al., 2013; Matthews & Marsh-Matthews, 2017). The usual metrics that quantify these changes in species composition, as the Kendall's coefficient or the coefficient of variation (Grossman et al., 1990), do not reflect the temporal patterns. These patterns have been inferred by visual examination of plotted trajectories in a chosen space of community resemblance (Magalhães et al., 2007; Matthews & Marsh-Matthews, 2017).

Based on the geometric properties of such trajectories, De Cáceres et al. (2019) developed an analytical framework to analyse and describe the temporal dynamics of

---

<sup>1</sup> In the literature, the terms “assemblage” and “community” have been used interchangeably, although they don't necessarily have the same meaning (Fauth et al., 1996). In this thesis, I use the term “community” (sensu Matthews & Marsh-Matthews, 2017) to describe the fish (or other faunal groups) found in a particular place and time. In this context, there's a predictability about the species that are found together, and there are biotic and abiotic factors that shape the communities through time and space (Matthews & Marsh-Matthews, 2017).



ecological communities (Community Trajectory Analysis). To define the trajectory of a community, each observation serves as a point in a multidimensional space and the trajectory is formalized as the set of directed segments connecting the observations in order. The multidimensional space is defined by the dissimilarity coefficient (e.g. Bray-Curtis) chosen to evaluate the resemblance between observations (De Cáceres et al., 2013, 2019; Sturbois et al., 2021). The geometric properties of the trajectories provide information about the community dynamics, for instance, the length of the segments and the trajectory can be used to quantify how much the community changes and if changes are abrupt or gradual, and the angles between segments and overall directionality quantify if the change is directional or not (De Cáceres et al., 2019; Sturbois et al., 2021). The synchrony of the variation of the communities in pairs of locations can be assessed by quantifying the convergence/divergence of the trajectories (De Cáceres et al., 2019). All these metrics could be used to model the spatial patterns of community dynamics in streams using the SSNM framework (Peterson & Ver Hoef, 2010).

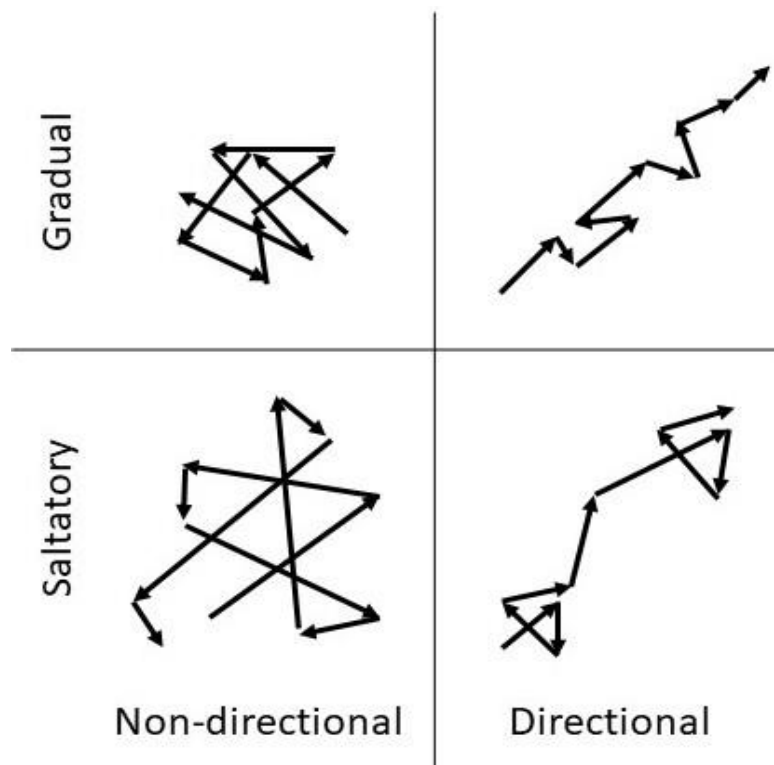


Figure 1.4 - Hypothetical trajectories reflecting the temporal dynamics of a community in a multivariate species-space. The framework depicts gradual versus saltatory change crossed with non-directional and directional change. Figure adapted from Matthews et al., 2013.

## 1.6 Objectives

Modelling the distribution of aquatic and semi-aquatic organisms in streams networks is challenging to say the least, but it is also critical to address the current freshwater biodiversity crisis. Therefore, the general objective of this thesis is to develop new ideas, frameworks and tools that can aid ecologists and conservationists to obtain a more realistic depiction of species distribution and their temporal dynamics at the riverscape scale, as well as a better understanding of biotic and abiotic factors that drive such distributions. To address this general goal, I have focused on a study system involving a Mediterranean watershed, the Sabor River, where I have undertaken several empirical studies related to distribution, biotic interactions and community dynamics of several animal groups, including crayfish, fish and amphibians (details in Chapter 2). Based on this system, the thesis pursued the following specific objectives:

1. To develop new approaches to account for imperfect detectability in distribution modelling across stream networks. Aquatic species are notoriously difficult to observe, and even standard protocols like electrofishing have a varying efficiency (Penczak & Głowacki, 2008; Reynolds, 1996; Zalewsky & Cowx, 1990), thus making imperfect detection an important issue that needs to be duly considered in species distribution modelling. While the toolbox of detection-occupancy models is rich and varied, most of the techniques available are difficult to use with sampling schemes often adopted for stream fish. Therefore, I have aimed at developing new tools to occupancy-detection modelling tailored for stream fish sampling, taking the Bayesian framework and a time to first detection approach, borrowing techniques from parametric survival analysis.
2. To develop and test new approaches to deal with spatially autocorrelated (SAC) data in species distribution models across stream networks. SAC is pervasive to ecological data and river ecosystems are no exception. To account for SAC in river ecosystems is necessary to take into account the dendritic structure of the streams networks (Peterson & Ver Hoef, 2010), but the methods available have rarely been used in combination with species distribution models in stream networks, particularly in the case of occupancy-detection models. Therefore, in the presented thesis I tested and developed new empirical frameworks to integrate SAC in distribution models, including:
  - a. Autoregressive models using hydrological distances to define the neighbourhood, in occupancy-detection models based on the time to first detection approach. For this component, I have used stream fish as the model organisms.

- b. Geostatistical models that explicitly account for the dendritic structure of stream networks (Peterson & Ver Hoef, 2010; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010), to improve spatial modelling of aquatic species distributions and community dynamics. For this component, I have used crayfish, fish and amphibians as model organisms.
3. To develop and test new approaches to model the effects of invasive species on present and future distributions of native species. Biotic interactions can be important drivers of species distribution (Araújo & Luoto, 2007; Wisz et al., 2013), and so they need to be duly considered. This is particularly important in the case of invasive species affecting native species, where spatial modelling can be used to identify in what areas should be targeted for conservation. For this component, I have used crayfish and amphibians as model organisms.
4. To develop and test new approaches to model spatial patterns in the community dynamics of aquatic organisms. The dynamics of communities is usually inferred from a relatively small number of discrete locations, but it is poorly known how to model the spatial pattern of temporal variation in the communities of aquatic organisms. To address this, in this thesis I aimed to combine for the first time, new geostatistical techniques developed for modelling processes in stream networks, with new methods designed to quantify community dynamics based on trajectory analysis (De Cáceres et al., 2019), For this component, I have used fish as model organisms.

## 1.7 Thesis outline

This thesis is structured in seven chapters, which include a general introduction (**chapter 1**), a brief description of the study area and of the studied species (**chapter 2**), a general discussion (**chapter 7**), and four chapters that correspond to four scientific papers that are already published in international peer-reviewed papers. **Chapter 3, 4, and 5** describe the modelling of the distribution of fish, invasive crayfish, and amphibians, respectively, while dealing with particular challenges of species distribution modelling in stream networks. **Chapter 6** describes a framework to develop a continuous spatial view of the dynamic of stream fish communities. The content of each chapter is summarised below.

**Chapter 1** is an introduction to the main themes this thesis addresses. It starts by describing the biodiversity crisis with a particular emphasis on freshwater ecosystems.

Then, it makes a brief description of stream networks and summarises the challenges of species distribution modelling in stream networks. It ends with the thesis objectives and outlines.

**Chapter 2** makes a brief description of the system at study. It starts with a brief description of the study area, the Sabor river catchment, and then proceed to make a small description of the studied species giving notes on conservation status and biogeography.

**Chapter 3** addresses **objectives 1** and **2a**, combining a novel occupancy-detection model with an autoregressive spatial structure, to describe and understand the drivers of distribution of the most common fish species in the Sabor river network, while accounting for imperfect detectability and spatial autocorrelation. The data was collected in a large survey in the summer of 2012, with a total of 184 sites visited, of which 95 were dry. Using a Bayesian framework, I developed a hierarchical occupancy-detection model that estimates the probability of the presence of water in stream channels, and conditional on the presence of water the probability of species presence related to large scale environmental predictors and spatial factors. The observation model estimates the detection probability using time-to-first-detection conditional on species presence and environmental predictors, using a modified interval-censored exponential survival model. The posterior probability of species occurrence derived from the model was used to produce distribution maps. This study was published in *Ecology and Evolution*: Ferreira, M., Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>

**Chapter 4** addresses **objective 2b**, providing one of the first ever applications regarding species distributions of geostatistical approaches to the modelling of ecological processes in stream networks. The approach is applied to data on the occurrence of two crayfish species across the River Sabor in the summer 2012. Because the dendritic structure of stream networks may constrain the dispersion of the invasive crayfish, we used geostatistical mixed models to relate the occurrence of crayfish to large scale environmental predictors, while specifying three spatial autocorrelation components in

the random errors. These components describe spatial dependencies between sites as a function of Euclidean distances, hydrological distances among flow connected sites, and hydrological distances irrespective of flow connection. This study was published in *Biological Invasions*: Filipe, A. F., Quaglietta, L., Ferreira, M., Magalhães, M. F., & Beja, P. (2017). Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks. *Biological Invasions*, 19(10), 2899–2912. <https://doi.org/10.1007/s10530-017-1492-3>

**Chapter 5** addresses **objectives 2a** and **3**, investigating the spatial distribution and environmental drivers of amphibians refuges under invasion by two species of crayfish. Data on amphibian occurrence was collected at 168 200-m stream stretches during 2015. Geostatistical models were used to relate the amphibian species occurrence to environmental factors and biotic interactions with crayfish, while controlling for Euclidean and hydrological spatial dependencies. The biotic interactions were specified using the crayfish occurrence probability, computed in the previous chapter, as a predictor in the amphibian distribution models. We then used a novel approach to predict future amphibian distributions under different scenarios of crayfish expansion across the stream network, thereby identifying areas that will likely remain free from crayfish (refuges), and thus should be primary targets for amphibian conservation efforts. This study was published in *Diversity and Distributions*: Mota-Ferreira, M., & Beja, P. (2020). Combining geostatistical and biotic interaction modeling to predict amphibian refuges under crayfish invasion across dendritic stream networks. *Diversity and Distributions*, 26(6), 699–714. <https://doi.org/10.1111/ddi.13047>

**Chapter 6** addresses **objective 4**, describing a framework to create a continuous spatial depiction of species community dynamics across a river network. Based on stream fish data collected at 30 sites between 2012 and 2019, we first used community trajectory analysis to quantify the directionality and velocity of community change in the River Sabor watershed. We then used the geostatistical approach already applied in previous chapters to model descriptors of community dynamics in relation to large scale environmental predictors while accounting for Euclidean and hydrological spatial dependencies. We then projected the models to the entire watershed, thereby providing a basis to visualise spatial patterns in community change. This study was published in *Diversity and Distributions*: Mota-Ferreira, M., Filipe, A. F., Magalhães, M.F., Carona, S.,

& Beja, P. (2021). Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change. *Diversity and Distributions*, 27(2), 313–326. <https://doi.org/10.1111/ddi.13189>

**Chapter 7** provides the main conclusions obtained from these studies, and contextualise them concerning existing research in the field. It also discusses the main shortcomings and caveats of the studies, as well the main challenges that were met in the making of this thesis. In the end, it raises specific questions that still need to be addressed, as well as promising research pathways for the future.

## 1.8 References

- Adame, M. F., Hermoso, V., Perhans, K., Lovelock, C. E., & Herrera-Silveira, J. A. (2015). Selecting cost-effective areas for restoration of ecosystem services. *Conservation Biology*, 29(2), 493–502. <https://doi.org/10.1111/cobi.12391>
- Albert, J. S., Destouni, G., Duke-Sylvester, S. M., Magurran, A. E., Oberdorff, T., Reis, R. E., Winemiller, K. O., & Ripple, W. J. (2020). Scientists' warning to humanity on the freshwater biodiversity crisis. *Ambio*. <https://doi.org/10.1007/s13280-020-01318-8>
- Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33(10), 1677–1688. <https://doi.org/10.1111/j.1365-2699.2006.01584.x>
- Araújo, M. B., Lobo, J. M., Moreno, J. C., & Ujo, M. B. A. R. A. (2007). The effectiveness of Iberian protected areas in conserving terrestrial biodiversity. *Conservation Biology: The Journal of the Society for Conservation Biology*, 21(6), 1423–1432. <https://doi.org/10.1111/j.1523-1739.2007.00827.x>
- Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16(6), 743–753. <https://doi.org/10.1111/j.1466-8238.2007.00359.x>
- Araújo, M. B., & Peterson, A. T. (2012). Uses and misuses of bioclimatic envelope modeling. *Ecology*, 93(7), 1527–1539. <https://doi.org/10.1890/11-1930.1>
- Araújo, M. B., Thuiller, W., & Pearson, R. G. (2006). Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, 33(10), 1712–1728. <https://doi.org/10.1111/j.1365-2699.2006.01482.x>
- Augustin, N. H., Muggleston, M. A., & Buckland, S. T. (1996). An Autologistic Model for the Spatial Distribution of Wildlife. *The Journal of Applied Ecology*, 33(2), 339. <https://doi.org/10.2307/2404755>
- Balian, E. V., Segers, H., Lévêque, C., & Martens, K. (2008). The Freshwater Animal Diversity Assessment: an overview of the results. *Hydrobiologia*, 595(1), 627–637. <https://doi.org/10.1007/s10750-007-9246-3>
- Baranov, V., Jourdan, J., Pilotto, F., Wagner, R., & Haase, P. (2020). Complex and nonlinear

- climate-driven changes in freshwater insect communities over 42 years. *Conservation Biology*, 0(0), 1–11. <https://doi.org/10.1111/cobi.13477>
- Bardos, D. C., Guillera-Aroita, G., & Wintle, B. A. (2015). Valid auto-models for spatially autocorrelated occupancy and abundance data. *Methods in Ecology and Evolution*, 6(10), 1137–1149. <https://doi.org/10.1111/2041-210X.12402>
- Bêche, L. A., Connors, P. G., Resh, V. H., & Merenlender, A. M. (2009). Resilience of fishes and invertebrates to prolonged drought in two California streams. *Ecography*, 32(5), 778–788. <https://doi.org/10.1111/j.1600-0587.2009.05612.x>
- Benyus, J. M. (1997). *Biomimicry: Innovation Inspired by Nature*. Harper Perennial.
- Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society - Series B (Methodological)*, 36(2), 192–236. <http://www.jstor.org/stable/2984812>
- Bornand, C. N., Kéry, M., Bueche, L., & Fischer, M. (2014). Hide-and-peek in vegetation: time-to-detection is an efficient design for estimating detectability and occurrence. *Methods in Ecology and Evolution*, 5(5), 433–442. <https://doi.org/10.1111/2041-210X.12171>
- Bustamante, J., & Seoane, J. (2004). Predicting the distribution of four species of raptors (Aves: Accipitridae) in southern Spain: Statistical models work better than existing maps. *Journal of Biogeography*, 31(2), 295–306. <https://doi.org/10.1046/j.0305-0270.2003.01006.x>
- Campbell Grant, E. H., Nichols, J. D., Lowe, W. H., & Fagan, W. F. (2010). Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. *Proceedings of the National Academy of Sciences of the United States of America*, 107(15), 6936–6940. <https://doi.org/10.1073/pnas.1000266107>
- Capinha, C., & Anastácio, P. (2011). Assessing the environmental requirements of invaders using ensembles of distribution models. *Diversity and Distributions*, 17(1), 13–24. <https://doi.org/10.1111/j.1472-4642.2010.00727.x>
- Capinha, C., Brotons, L., & Anastácio, P. (2012). Geographical variability in propagule pressure and climatic suitability explain the European distribution of two highly invasive crayfish. *Journal of Biogeography*, n/a-n/a. <https://doi.org/10.1111/jbi.12025>
- Capinha, C., Essl, F., Seebens, H., Moser, D., & Pereira, H. M. (2015). The dispersal of alien species redefines biogeography in the Anthropocene. *Science*, 348(6240), 1248–1251. <https://doi.org/10.1126/science.aaa8913>
- Capinha, C., Leung, B., & Anastácio, P. (2011). Predicting worldwide invasiveness for four major problematic decapods: an evaluation of using different calibration sets. *Ecography*, 34(3), 448–459. <https://doi.org/10.1111/j.1600-0587.2010.06369.x>
- Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., MacE, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, 486(7401), 59–67. <https://doi.org/10.1038/nature11148>
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5), e1400253. <https://doi.org/10.1126/sciadv.1400253>
- Cid, N., Bonada, N., Carlson, S. M., Grantham, T. E., Gasith, A., & Resh, V. H. (2017). High variability is a defining component of mediterranean-climate rivers and their biota. *Water*

- (Switzerland), 9(1), 1–24. <https://doi.org/10.3390/w9010052>
- Cid, N., Bonada, N., Heino, J., Cañedo-Argüelles, M., Crabot, J., Sarremejane, R., Soininen, J., Stubbington, R., & Datry, T. (2020). A Metacommunity Approach to Improve Biological Assessments in Highly Dynamic Freshwater Ecosystems. *BioScience*, 70(5), 427–438. <https://doi.org/10.1093/biosci/biaa033>
- Cmglee. (2020). Earth water distribution. Wikimedia Commons, the free media repository. [https://commons.wikimedia.org/wiki/File:Earth\\_water\\_distribution.svg](https://commons.wikimedia.org/wiki/File:Earth_water_distribution.svg)
- Collins, S. L. (2000). Disturbance frequency and community stability in native tallgrass prairie. *American Naturalist*, 155(3), 311–325. <https://doi.org/10.1086/303326>
- Cruz, M. J., & Rebelo, R. (2007). Colonization of freshwater habitats by an introduced crayfish, *Procambarus clarkii*, in Southwest Iberian Peninsula. *Hydrobiologia*, 575(1), 191–201. <https://doi.org/10.1007/s10750-006-0376-9>
- Dai, A., & Trenberth, K. E. (2002). Estimates of Freshwater Discharge from Continents: Latitudinal and Seasonal Variations. *Journal of Hydrometeorology*, 3(6), 660–687. [https://doi.org/10.1175/1525-7541\(2002\)003<0660:EOFDfC>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0660:EOFDfC>2.0.CO;2)
- De Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wiser, S. K., Fortin, M.-J., Condit, R., & Hubbell, S. (2019). Trajectory analysis in community ecology. *Ecological Monographs*, 89(2), e01350. <https://doi.org/10.1002/ecm.1350>
- De Cáceres, M., Legendre, P., & He, F. (2013). Dissimilarity measurements and the size structure of ecological communities. *Methods in Ecology and Evolution*, 4(12), 1167–1177. <https://doi.org/10.1111/2041-210X.12116>
- DeAngelis, D. L., Waterhouse, J. C., Post, W. M., & O'Neill, R. V. (1985). Ecological modelling and disturbance evaluation. *Ecological Modelling*, 29(1–4), 399–419. [https://doi.org/10.1016/0304-3800\(85\)90063-8](https://doi.org/10.1016/0304-3800(85)90063-8)
- Delong, D. C. (1996). Defining Biodiversity. *Wildlife Society Bulletin*, 24(4), 738–749. <https://www.jstor.org/stable/3783168>
- Diniz-Filho, J. A. F., Bini, L. M., & Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, 12(1), 53–64. <https://doi.org/10.1046/j.1466-822X.2003.00322.x>
- Dobson, A. P., Nowak, K., & Rodríguez, J. P. (2013). Conservation Biology, Discipline of. *Encyclopedia of Biodiversity: Second Edition*, 2, 238–248. <https://doi.org/10.1016/B978-0-12-384719-5.00272-0>
- Domisch, S., Jähnig, S. C., Simaika, J. P., Kuemmerlen, M., & Stoll, S. (2015). Application of species distribution models in stream ecosystems: The challenges of spatial and temporal scale, environmental predictors and species occurrence data. *Fundamental and Applied Limnology*, 186(1–2), 45–61. <https://doi.org/10.1127/fal/2015/0627>
- Dorazio, R. M. (2012). Predicting the Geographic Distribution of a Species from Presence-Only Data Subject to Detection Errors. *Biometrics*, 68(4), 1303–1312. <https://doi.org/10.1111/j.1541-0420.2012.01779.x>
- Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., Naiman, R. J., Prieur-Richard, A.-H., Soto, D., Stiassny, M. L. J., & Sullivan, C. A. (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews*, 81(02), 163. <https://doi.org/10.1017/S1464793105006950>



- Einoder, L. D., Southwell, D. M., Lahoz-Monfort, J. J., Gillespie, G. R., Fisher, A., & Wintle, B. A. (2018). Occupancy and detectability modelling of vertebrates in northern Australia using multiple sampling methods. *PLoS ONE*, 13(9). <https://doi.org/10.1371/journal.pone.0203304>
- Erös, T. (2017). Scaling fish metacommunities in stream networks: Synthesis and future research avenues. *Community Ecology*, 18(1), 72–86. <https://doi.org/10.1556/168.2017.18.1.9>
- Erös, T., Comte, L., Filipe, A. F., Ruhi, A., Tedesco, P. A., Brose, U., Fortin, M. J., Giam, X., Irving, K., Jacquet, C., Larsen, S., Sharma, S., & Olden, J. D. (2020). Effects of nonnative species on the stability of riverine fish communities. *Ecography*, 1–11. <https://doi.org/10.1111/ecog.04985>
- Erös, T., & Schmera, D. (2010). Spatio-temporal scaling of biodiversity and the species-time relationship in a stream fish assemblage. *Freshwater Biology*, 55(11), 2391–2400. <https://doi.org/10.1111/j.1365-2427.2010.02438.x>
- Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S. R., Essington, T. E., Holt, R. D., Jackson, J. B. C., Marquis, R. J., Oksanen, L., Oksanen, T., Paine, R. T., Pikitch, E. K., Ripple, W. J., Sandin, S. A., Scheffer, M., Schoener, T. W., ... Soulé, M. E. (2011). Trophic Downgrading of Planet Earth. *Science*, 333(July), 301–307. <https://doi.org/10.1126/science.1205106>
- Fausch, K. D., Torgersen, C. E., Baxter, C. V., & Li, H. W. (2002). Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Fishes. *BioScience*, 52(6), 483. [https://doi.org/10.1641/0006-3568\(2002\)052\[0483:lrbtj\]2.0.co;2](https://doi.org/10.1641/0006-3568(2002)052[0483:lrbtj]2.0.co;2)
- Fauth, J. E., Bernardo, J., Camara, M., Resetarits, W. J., Van Buskirk, J., & McCollum, S. A. (1996). Simplifying the jargon of community ecology: A conceptual approach. *American Naturalist*, 147(2), 282–286. <https://doi.org/10.1086/285850>
- Ferreira, M., & Beja, P. (2013). Mediterranean amphibians and the loss of temporary ponds: Are there alternative breeding habitats? *Biological Conservation*, 165(null), 179–186. <https://doi.org/10.1016/j.biocon.2013.05.029>
- Ficetola, G. F., Bonardi, A., Mairota, P., Leronna, V., & Padoa-Schioppa, E. (2014). Predicting wild boar damages to croplands in a mosaic of agricultural and natural areas. *Current Zoology*, 60(2), 170–179. <https://doi.org/10.1093/czoolo/60.2.170>
- Ficetola, G. F., Siesa, M. E., Manenti, R., Bottoni, L., De Bernardi, F., & Padoa-Schioppa, E. (2011). Early assessment of the impact of alien species: Differential consequences of an invasive crayfish on adult and larval amphibians. *Diversity and Distributions*, 17(6), 1141–1151. <https://doi.org/10.1111/j.1472-4642.2011.00797.x>
- Franklin, J. (2013). Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions*, 19(10), 1217–1223. <https://doi.org/10.1111/ddi.12125>
- Garrard, G. E., Bekessy, S. A., McCarthy, M. A., & Wintle, B. A. (2008). When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology*, 33(8), 986–998. <https://doi.org/10.1111/j.1442-9993.2008.01869.x>
- Garrard, G. E., McCarthy, M. a., Williams, N. S. G., Bekessy, S. a., & Wintle, B. a. (2013). A general model of detectability using species traits. *Methods in Ecology and Evolution*, 4(1), 45–52. <https://doi.org/10.1111/j.2041-210x.2012.00257.x>
- Gaston, K. J., & Spicer, J. I. (2004). *Biodiversity: An Introduction* (2nd ed.). Wiley-Blackwell.
- Geist, J. (2011). Integrative freshwater ecology and biodiversity conservation. *Ecological Indicators*, 11(6), 1507–1516. <https://doi.org/10.1016/j.ecolind.2011.04.002>

- Gherardi, F. (2007). Understanding the impact of invasive crayfish. In F. Gherardi (Ed.), *Biological invaders in inland waters: Profiles, distribution, and threats* (pp. 507–542). Springer Netherlands. [https://doi.org/10.1007/978-1-4020-6029-8\\_28](https://doi.org/10.1007/978-1-4020-6029-8_28)
- Gleick, Peter H. (2014). *The World's Water Volume 8: Vol. The Bienni* (Peter H. Gleick (ed.); 1st ed.). Island Press/Center for Resource Economics. <https://doi.org/10.5822/978-1-61091-483-3>
- Grossman, G. D., Dowd, J. F., & Crawford, M. (1990). Assemblage stability in stream fishes: A review. *Environmental Management*, 14(5), 661–671. <https://doi.org/10.1007/BF02394716>
- Guillera-Arroita, G., Lahoz-Monfort, J. J., MacKenzie, D. I., Wintle, B. A., & McCarthy, M. A. (2014). Ignoring Imperfect Detection in Biological Surveys Is Dangerous: A Response to 'Fitting and Interpreting Occupancy Models'. *PLoS ONE*, 9(7), e99571. <https://doi.org/10.1371/journal.pone.0099571>
- Guillera-Arroita, G., Morgan, B. J. T., Ridout, M. S., & Linkie, M. (2011). Species Occupancy Modeling for Detection Data Collected Along a Transect. *Journal of Agricultural, Biological, and Environmental Statistics*, 16(3), 301–317. <https://doi.org/10.1007/s13253-010-0053-3>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16(12), 1424–1435. <https://doi.org/10.1111/ele.12189>
- Gumpertz, M. L., Graham, J. M., & Ristaino, J. B. (1997). Autologistic Model of Spatial Pattern of Phytophthora Epidemic in Bell Pepper: Effects of Soil Variables on Disease Presence. *Journal of Agricultural, Biological, and Environmental Statistics*, 2(2), 131. <https://doi.org/10.2307/1400400>
- Guo, Z., Zhang, L., & Li, Y. (2010). Increased Dependence of Humans on Ecosystem Services and Biodiversity. *PLoS ONE*, 5(10), e13113. <https://doi.org/10.1371/journal.pone.0013113>
- Harper, J. L., & Hawksworth, D. L. (1994). Biodiversity: measurement and estimation. Preface. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 345(1311), 5–12. <https://doi.org/10.1098/rstb.1994.0081>
- Heino, J. (2013). The importance of metacommunity ecology for environmental assessment research in the freshwater realm. *Biological Reviews*, 88(1), 166–178. <https://doi.org/10.1111/j.1469-185X.2012.00244.x>
- Hermoso, V., Filipe, A. F., Segurado, P., & Beja, P. (2016). Catchment zoning to unlock freshwater conservation opportunities in the Iberian Peninsula. *Diversity and Distributions*, 22(9), 960–969. <https://doi.org/10.1111/ddi.12454>
- Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68(1–2), 87–112. <https://doi.org/10.1111/j.1095-8312.1999.tb01160.x>
- Hoeting, J. A., Leecaster, M., & Bowden, D. (2000). An Improved Model for Spatially Correlated Binary Responses. *Journal of Agricultural, Biological, and Environmental Statistics*, 5(1), 102. <https://doi.org/10.2307/1400634>
- Hortal, J., Borges, P. A. V., Jiménez-Valverde, A., de Azevedo, E. B., & Silva, L. (2010). Assessing the areas under risk of invasion within islands through potential distribution modelling: The case

- of *Pittosporum undulatum* in São Miguel, Azores. *Journal for Nature Conservation*, 18(4), 247–257. <https://doi.org/10.1016/j.jnc.2009.11.002>
- International Commission on Large Dams [ICOLD]. (2020). ICOLD world register of dams.
- Isaak, D. J., Peterson, E. E., Ver Hoef, J. M., Wenger, S. J., Falke, J. A., Torgersen, C. E., Sowder, C., Steel, E. A., Fortin, M.-J., Jordan, C. E., Ruesch, A. S., Som, N., & Monestiez, P. (2014). Applications of spatial statistical network models to stream data. *Wiley Interdisciplinary Reviews: Water*, 1(3), 277–294. <https://doi.org/10.1002/wat2.1023>
- Isbell, F., Gonzalez, A., Loreau, M., Cowles, J., Díaz, S., Hector, A., Mace, G. M., Wardle, D. A., O'Connor, M. I., Duffy, J. E., Turnbull, L. A., Thompson, P. L., & Larigauderie, A. (2017). Linking the influence and dependence of people on biodiversity across scales. *Nature*, 546(7656), 65–72. <https://doi.org/10.1038/nature22899>
- Jackson, R. B., Carpenter, S. R., Dahm, C. N., McKnight, D. M., Naiman, R. J., Postel, S. L., & Running, S. W. (2001). Water in a changing world. *Ecological Applications*, 11(4), 1027–1045. [https://doi.org/10.1890/1051-0761\(2001\)011\[1027:WIACW\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2001)011[1027:WIACW]2.0.CO;2)
- Jarnevich, C. S., & Reynolds, L. V. (2011). Challenges of predicting the potential distribution of a slow-spreading invader: a habitat suitability map for an invasive riparian tree. *Biological Invasions*, 13(1), 153–163. <https://doi.org/10.1007/s10530-010-9798-4>
- Jiménez-Valverde, A., Peterson, A. T., Soberón, J., Overton, J. M., Aragón, P., & Lobo, J. M. (2011). Use of niche models in invasive species risk assessments. *Biological Invasions*, 13(12), 2785–2797. <https://doi.org/10.1007/s10530-011-9963-4>
- Johnson, B. L., Richardson, W. B., & Naimo, T. J. (1995). Past, present, and future concepts in large river ecology - How rivers function and how human activities influence river processes. *BioScience*, 45(3), 134–141. <https://doi.org/10.2307/1312552>
- Jourdan, J., O'Hara, R. B., Bottarin, R., Huttunen, K. L., Kuemmerlen, M., Monteith, D., Muotka, T., Ozoliņš, D., Paavola, R., Pilotto, F., Springe, G., Skuja, A., Sundermann, A., Tonkin, J. D., & Haase, P. (2018). Effects of changing climate on European stream invertebrate communities: A long-term data analysis. *Science of the Total Environment*, 621, 588–599. <https://doi.org/10.1016/j.scitotenv.2017.11.242>
- Kerezszy, A., Gido, K., Magalhães, M. F., & Skelton, P. H. (2017). The Biota of Intermittent Rivers and Ephemeral Streams: Fishes. In *Intermittent Rivers and Ephemeral Streams* (pp. 273–298). Elsevier. <https://doi.org/10.1016/B978-0-12-803835-2.00010-3>
- Kéry, M., & Schmidt, B. (2008). Imperfect detection and its consequences for monitoring for conservation. *Community Ecology*, 9(2), 207–216. <https://doi.org/10.1556/ComEc.9.2008.2.10>
- Kéry, Marc. (2011). Towards the modelling of true species distributions. *Journal of Biogeography*, 38(4), 617–618. <https://doi.org/10.1111/j.1365-2699.2011.02487.x>
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis*. Springer New York. <https://doi.org/10.1007/978-1-4419-6646-9>
- Lahoz-Monfort, J. J., Guillera-Aroita, G., & Wintle, B. a. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, 23(4), 504–515. <https://doi.org/10.1111/geb.12138>
- Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6), 1659–1673. <https://doi.org/10.2307/1939924>

- Lévêque, C., Oberdorff, T., Paugy, D., Stiassny, M. L. J., & Tedesco, P. A. (2008). Global diversity of fish (Pisces) in freshwater. *Hydrobiologia*, 595(1), 545–567. <https://doi.org/10.1007/s10750-007-9034-0>
- Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33(1), 103–114. <https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- MacKenzie, D. I. (2006). Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. *Journal of Wildlife Management*, 70(2), 367–374. [https://doi.org/10.2193/0022-541X\(2006\)70\[367:MTPORU\]2.0.CO;2](https://doi.org/10.2193/0022-541X(2006)70[367:MTPORU]2.0.CO;2)
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8), 2248–2255. [https://doi.org/10.1890/0012-9658\(2002\)083\[2248:ESORWD\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2)
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier Academic Press. <http://www.google.pt/books?hl=pt-PT&lr=&id=pf-w-JAUOd0C&pgis=1>
- Magalhães, M. F., Beja, P., Schlosser, I. J., & Collares-Pereira, M. J. (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, 52(8), 1494–1510. <https://doi.org/10.1111/j.1365-2427.2007.01781.x>
- Matthews, W. J. (1998). *Patterns in Freshwater Fish Ecology*. Springer.
- Matthews, W. J., & Marsh-Matthews, E. (2016). Dynamics of an upland stream fish community over 40 years: Trajectories and support for the loose equilibrium concept. *Ecology*, 97(3), 706–719. <https://doi.org/10.1890/14-2179.1>
- Matthews, W. J., & Marsh-Matthews, E. (2017). Stream fish community dynamics: A critical synthesis. In *Stream Fish Community Dynamics: A Critical Synthesis*. JHU Press.
- Matthews, W. J., Marsh-Matthews, E., Cashner, R. C., & Gelwick, F. (2013). Disturbance and trajectory of change in a stream fish community over four decades. *Oecologia*, 173(3), 955–969. <https://doi.org/10.1007/s00442-013-2646-3>
- Mazerolle, M. J. M. J., Bailey, L. L., Kendall, W. L., Andrew Royle, J., Converse, S. J. S. J., & Nichols, J. D. (2007). Making great leaps forward: accounting for detectability in herpetological field studies. *Journal of Herpetology*, 41(4), 672–689. <https://doi.org/10.1670/07-061.1>
- McCarthy, M. a., Moore, J. L., Morris, W. K., Parris, K. M., Garrard, G. E., Vesk, P. a., Rumpff, L., Giljohann, K. M., Camac, J. S., Bau, S. S., Friend, T., Harrison, B., & Yue, B. (2013). The influence of abundance on detectability. *Oikos*, 122(5), 717–726. <https://doi.org/10.1111/j.1600-0706.2012.20781.x>
- McNeely, Jeffrey A. (1994). Lessons from the past: forests and biodiversity. *Biodiversity and Conservation*, 3(1), 3–20. <https://doi.org/10.1007/BF00115329>
- Meixler, M. S., Bain, M. B., & Todd Walter, M. (2009). Predicting barrier passage and habitat suitability for migratory fish species. *Ecological Modelling*, 220(20), 2782–2791. <https://doi.org/10.1016/j.ecolmodel.2009.07.014>
- Meyer, J. L., Strayer, D. L., Wallace, J. B. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks. *JAWRA Journal*

- of the American Water Resources Association, 43(1), 86–103. <https://doi.org/10.1111/j.1752-1688.2007.00008.x>
- Olden, J. D., Kennard, M. J., Leprieur, F., Tedesco, P. A., Winemiller, K. O., & García-Berthou, E. (2010). Conservation biogeography of freshwater fishes: Recent progress and future challenges. *Diversity and Distributions*, 16(3), 496–513. <https://doi.org/10.1111/j.1472-4642.2010.00655.x>
- Olsson, O., & Rogers, D. J. (2009). Predicting the distribution of a suitable habitat for the white stork in Southern Sweden: identifying priority areas for reintroduction and habitat restoration. *Animal Conservation*, 12(1), 62–70. <https://doi.org/10.1111/j.1469-1795.2008.00225.x>
- Penczak, T., & Głowacki, Ł. B. (2008). Evaluation of electrofishing efficiency in a stream under natural and regulated conditions. *Aquatic Living Resources*, 21(3), 329–337. <https://doi.org/10.1051/alr:2008044>
- Peterson, E. E., Theobald, D. M., & Ver Hoef, J. M. (2007). Geostatistical modelling on stream networks: Developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology*, 52(2), 267–279. <https://doi.org/10.1111/j.1365-2427.2006.01686.x>
- Peterson, E. E., & Ver Hoef, J. M. (2010). A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91(3), 644–651. <https://doi.org/10.1890/08-1668.1>
- Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>
- Préau, C., Nadeau, I., Sellier, Y., Isselin-Nondedeu, F., Bertrand, R., Collas, M., Capinha, C., & Grandjean, F. (2020). Niche modelling to guide conservation actions in France for the endangered crayfish *Austropotamobius pallipes* in relation to the invasive *Pacifastacus leniusculus*. *Freshwater Biology*, 65(2), 304–315. <https://doi.org/10.1111/fwb.13422>
- Quaglietta, L., Fonseca, V. C., Mira, A., & Boitani, L. (2014). Sociospatial organization of a solitary carnivore, the Eurasian otter (*Lutra lutra*). *Journal of Mammalogy*, 95(1), 140–150. <https://doi.org/10.1644/13-MAMM-A-073.1>
- Record, S., Fitzpatrick, M. C., Finley, A. O., Veloz, S., & Ellison, A. M. (2013). Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change. *Global Ecology and Biogeography*, 22(6), 760–771. <https://doi.org/10.1111/geb.12017>
- Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94(3), 849–873. <https://doi.org/10.1111/brv.12480>
- Reynolds, J. B. (1996). *Electrofishing. Fisheries Techniques* (B. R. Murphy & D. W. Willis (eds.); 2nd ed.). American Fisheries Society.
- Rodríguez, J. P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions*, 13(3), 243–251. <https://doi.org/10.1111/j.1472-4642.2007.00356.x>

- Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L. F., Jackson, R. B., Kinzig, A., Leemans, R., Lodge, D. M., Mooney, H. A., Oesterheld, M., Poff, N. L. R., Sykes, M. T., Walker, B. H., Walker, M., & Wall, D. H. (2000). Global biodiversity scenarios for the year 2100. In *Science* (Vol. 287, Issue 5459, pp. 1770–1774). <https://doi.org/10.1126/science.287.5459.1770>
- Sánchez-Montoya, M. M., Moleón, M., Sánchez-Zapata, J. A., & Escoriza, D. (2017). The Biota of Intermittent and Ephemeral Rivers: Amphibians, Reptiles, Birds, and Mammals. In *Intermittent Rivers and Ephemeral Streams* (pp. 299–322). Elsevier. <https://doi.org/10.1016/B978-0-12-803835-2.00011-5>
- Santos, R. A. L., Mota-Ferreira, M., Aguiar, L. M. S., & Ascensão, F. (2018). Predicting wildlife road-crossing probability from roadkill data using occupancy-detection models. *Science of The Total Environment*, 642, 629–637. <https://doi.org/10.1016/j.scitotenv.2018.06.107>
- Scherr, S. J., & McNeely, J. a. (2008). Biodiversity conservation and agricultural sustainability: towards a new paradigm of “ecoagriculture” landscapes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1491), 477–494. <https://doi.org/10.1098/rstb.2007.2165>
- Schmutz, S., & Moog, O. (2018). Dams: Ecological Impacts and Management. In S. Schmutz & J. Sendzimir (Eds.), *Riverine Ecosystem Management* (pp. 111–127). Springer International Publishing. [https://doi.org/10.1007/978-3-319-73250-3\\_6](https://doi.org/10.1007/978-3-319-73250-3_6)
- Schopf, J. W. (2006). Fossil evidence of Archaean life. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1470), 869–885. <https://doi.org/10.1098/rstb.2006.1834>
- Seoane, J., Viñuela, J., Díaz-Delgado, R., & Bustamante, J. (2003). The effects of land use and climate on red kite distribution in the Iberian peninsula. *Biological Conservation*, 111(3), 401–414. [https://doi.org/10.1016/S0006-3207\(02\)00309-9](https://doi.org/10.1016/S0006-3207(02)00309-9)
- Shiklomanov, L. A. (1993). World Freshwater Resources. In P.H. Gleick (Ed.), *Water in Crisis: A Guide to World’s Freshwater Resources* (pp. 13–24). Oxford University Press.
- Srivastava, V., Lafond, V., & Griess, V. C. (2019). Species distribution models (SDM): Applications, benefits and challenges in invasive species management. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 14(020). <https://doi.org/10.1079/PAVSNNR201914020>
- Strayer, D. L. (2010). Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. *Freshwater Biology*, 55(SUPPL. 1), 152–174. <https://doi.org/10.1111/j.1365-2427.2009.02380.x>
- Strayer, D. L., & Dudgeon, D. (2010). Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society*, 29(1), 344–358. <https://doi.org/10.1899/08-171.1>
- Sturbois, A., De Cáceres, M., Sánchez-Pinillos, M., Schaal, G., Gauthier, O., Le Mao, P., Ponsero, A., & Desroy, N. (2021). Extending community trajectory analysis: New metrics and representation. *Ecological Modelling*, 440(December 2020), 109400. <https://doi.org/10.1016/j.ecolmodel.2020.109400>
- Tangley, L. (1985). A New Plan to Conserve the Earth’s Biota: Spurred by congressional concern, federal officials have drafted a US government strategy to preserve biological diversity in the Third World. *BioScience*, 35(6), 334–341. <https://doi.org/10.1093/bioscience/35.6.334>
- UNESCO World Water Assessment Programme. (2020). The United Nations world water

- development report 2020: Water and climate change, executive summary. <https://en.unesco.org/themes/water-security/wwap/wwdr/2020>
- Václavík, T., Kupfer, J. A., & Meentemeyer, R. K. (2012). Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography*, 39(1), 42–55. <https://doi.org/10.1111/j.1365-2699.2011.02589.x>
- Václavík, T., & Meentemeyer, R. K. (2012). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73–83. <https://doi.org/10.1111/j.1472-4642.2011.00854.x>
- Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(1), 130–137. <https://doi.org/10.1139/f80-017>
- Ver Hoef, J. M., & Peterson, E. E. (2010). A Moving Average Approach for Spatial Statistical Models of Stream Networks. *Journal of the American Statistical Association*, 105(489), 6–18. <https://doi.org/10.1198/jasa.2009.ap08248>
- Ver Hoef, J. M., Peterson, E. E., Hooten, M. B., Hanks, E. M., & Fortin, M. J. (2018). Spatial autoregressive models for statistical inference from ecological data. *Ecological Monographs*, 88(1), 36–59. <https://doi.org/10.1002/ecm.1283>
- Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13(4), 449–464. <https://doi.org/10.1007/s10651-006-0022-8>
- Vitousek, P. M. (1997). Human Domination of Earth's Ecosystems. *Science*, 277(5325), 494–499. <https://doi.org/10.1126/science.277.5325.494>
- Wang, H.-H., & Grant, W. E. (2019a). Systems, models, and modeling. In H.-H. Wang & W. E. Grant (Eds.), *Ecological Modeling: An Introduction to the Art and Science of Modeling Ecological Systems* (Vol. 31, pp. 3–20). <https://doi.org/10.1016/B978-0-444-64163-2.00001-3>
- Wang, H.-H., & Grant, W. E. (2019b). What good is modeling? In H.-H. Wang & W. E. Grant (Eds.), *Ecological Modeling: An Introduction to the Art and Science of Modeling Ecological Systems* (Vol. 31, Issue 4, pp. 215–224). <https://doi.org/10.1016/B978-0-444-64163-2.00011-6>
- Ward, J. V., & Stanford, J. A. (1995). Ecosystems and Its Disruption By Flow Regulation. *Regulated Rivers: Research & Management*, II(July 1994), 105–119.
- Wilcox, B. A. (1984). In situ conservation of genetic resources: determinants of minimum area requirements. In J.A. McNeely & K. R. Miller (Eds.), *National parks, conservation and development: the role of protected areas in sustaining society* (pp. 18–30). Smithsonian Institution Press.
- Wilson, Edward O. (1985). The biological diversity crisis: A challenge to science. *Issues in Science and Technology*, 2(1), 20–29.
- Wilson, Edward O. (1988). *Biodiversity* (E O Wilson & Harvard University (eds.)). National Academies Press. <https://doi.org/10.17226/989>
- Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M. C., Normand, S., Öckinger, E., Schmidt, N. M., ... Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88(1), 15–30.

<https://doi.org/10.1111/j.1469-185X.2012.00235.x>

Zalewsky, M., & Cowx, I. G. (1990). Factors affecting the efficiency of electric fishing. In *Fishing with electricity: applications in freshwater fisheries management* (pp. 89–111). Fishing News Books.



## Chapter 2 – Study system



## Chapter 2 – Study system

### 2.1 Study Area

The fieldwork to collect the data used in this thesis was conducted in the Sabor river catchment, a tributary of Douro river in NE Portugal (N41°090–42°000, W7°150–6°150, Figure 2.), mostly within the Baixo Sabor Long Term Ecological Research Site (LTER\_EU\_PT\_002) and Natura 2000 site of Sabor-Maçãs (PTCON0021). It is a scarcely populated region (8.5 – 28.7 inhabitants / km<sup>2</sup> ; <https://www.pordata.pt/Municipios>) following a process of land abandonment since the 1970s (Azevedo et al., 2011; Hoelzer, 2003). The study area is in the transition between the meso- and supra-Mediterranean bioclimatic zones (Monteiro-Henriques, 2010), with cold winters and hot dry summers. Precipitation is mainly concentrated in October-March and virtually absent in the hot summer months of June-August. Because of this, most of the headwater streams dry out, or become a series of isolated pools, in the summer, though the main river and largest tributaries are permanent. The stream network covers a wide range of environmental conditions in terms of elevation (1000 – 1500 meters above sea level), total annual precipitation (443 – 1163 mm), and mean annual temperature (6.9 – 16.6°C). Two hydroelectric dams (Feiticeiro: 181 ha; Baixo Sabor: 2 820 ha) located near the mouth of the Sabor river started to be built in 2009 with the main reservoir filled during the winter of 2014/2015 (Jackson, 2011; Santos et al., 2017). Besides the two hydroelectric dams and a few small impoundments, the streams are free-flowing.

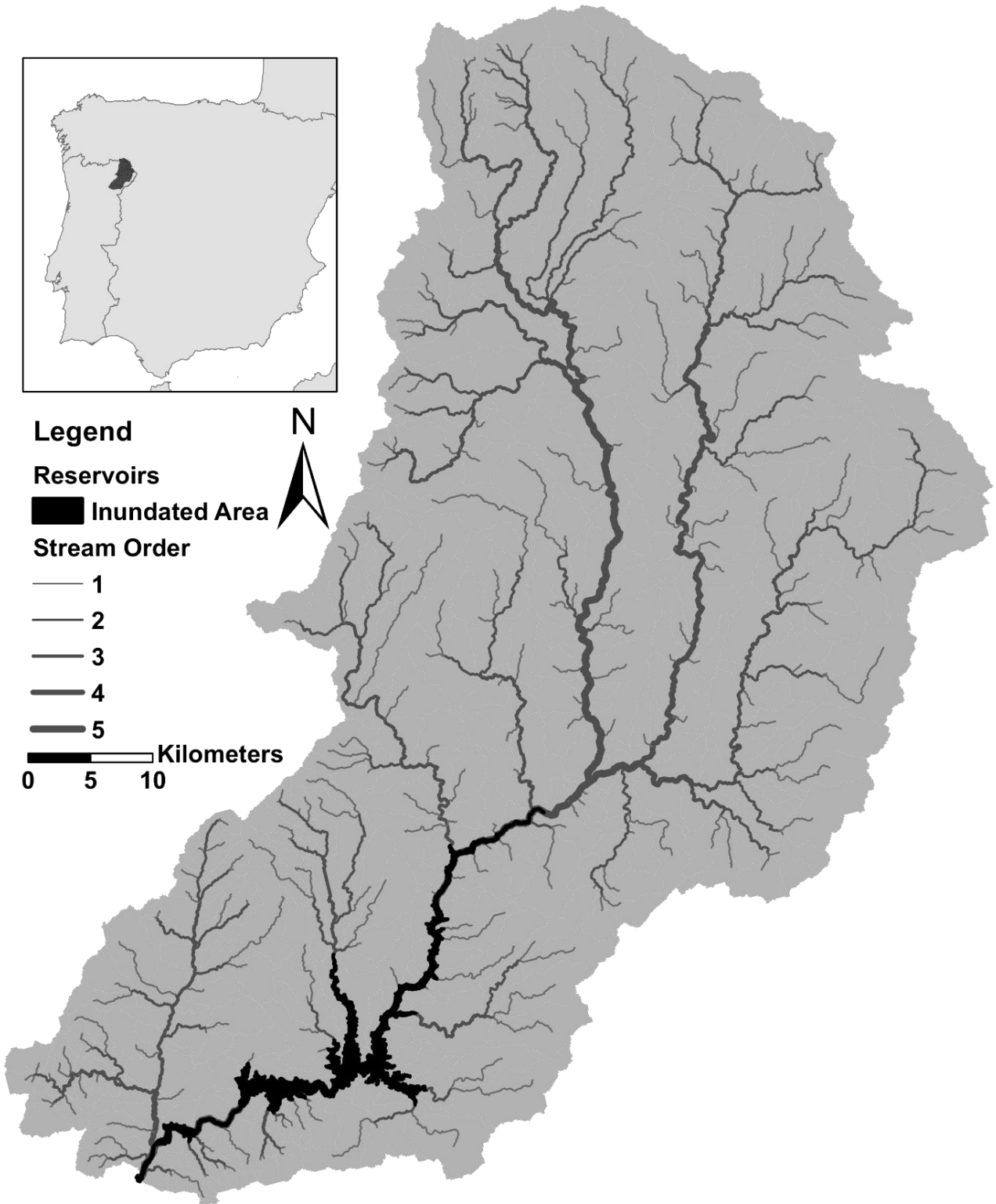


Figure 2.1 – Study Area

## 2.2 Studied organisms

### 2.2.1 Non-native invasive crayfish

Crayfish are among the largest and longest-lived freshwater invertebrates (Gherardi et al., 2011). Most species are omnivorous and ecologically key species, consuming plants and detritus, and being opportunistic predators of invertebrates, amphibians eggs and larvae, and even crayfish (Gherardi, 2007; Nyström et al., 1996). Crayfish are an important socioeconomic resource; in 2018, the worldwide production of crayfish aquaculture was about 1.7 million tonnes, corresponding to 1.49% of the worldwide production of aquaculture species, but corresponding to 5.49% of the value (Junning et al., 2020). Due to stocking and introductions, much of the present distribution of crayfish is due to human intervention (Gherardi, 2007), with at least eleven non-native invasive crayfish species (NICS) in Europe (Kouba et al., 2014). These include some populations that were once regarded as native, but that are in fact historical introductions (Gherardi, 2007).

According to Gherardi (2007), NCIS can impact biodiversity at different levels. NCIS can impact biodiversity at individual and population level by preying on native species (Cruz, Pascoal, et al., 2006; Cruz & Rebelo, 2005; Ilhéu et al., 2007; Sousa et al., 2019), modifying or inhibiting native species behaviour (Gamradt et al., 1997; Nunes, 2011; Nyström & Åbjörnsson, 2000), and being a vector for diseases for native crayfish (Alderman et al., 1990; Diéguez-Uribeondo & Söderhäll, 1993) and for other organisms (Brannelly et al., 2015). They also have become important prey for native species (Correia, 2001). NCIS are known to compete and even hybridize with native crayfish (review by Gherardi, 2007). The impact of NCIS in aquatic ecosystems can be devastating, causing the local extirpation of native species such as amphibians (Cruz et al., 2008). Due to diet plasticity, NCIS can consume in several trophic levels, becoming the central species on the nutrient cycle of the invaded habitats (Nyström et al., 1996). NCIS also alter the physicochemical properties of the water in invaded habitats. By walking, tail flipping, and burrowing, NCIS promote the suspension of sediments and nutrients and are thus linked to blooms of cyanobacteria (Welch & Pintor, 2014; Yamamoto, 2010). Options to control and mitigate NCIS exist, but are expensive, labour intensive, and so far have met limited success (Gherardi et al., 2011; Stebbing et al.,

2014). Once a population of NCIS is established in the wild, eradication is nearly impossible. The focus should be on raising public awareness about the impact of NCIS and halting the spread of NCIS to non-invaded habitats that serve as refuges for native species (Gherardi, 2012).

In the Iberian Peninsula, there are no native crayfish, though invasive species are widespread in most freshwater systems (Clavero et al., 2016; Kouba et al., 2014). The earliest crayfish invader was probably the white-clawed crayfish, *Austropotamobius italicus*, which was once regarded as a naturally occurring species and an important socio-economic resource (Gutiérrez-Yurrita et al., 1999), but now it has been recognised as a species introduced from Italy, in the XVI century (Clavero et al., 2016). Following a crash in *A. italicus* populations, the Spanish authorities tried to introduce several species of exotic crayfish in the decade of 1970 (Gutiérrez-Yurrita et al., 1999), but only two species became widespread: the red swamp crayfish, *Procambarus clarkii* (Girard, 1852), and the signal crayfish, *Pacifastacus leniusculus* (Dana 1852), (Figure 2.2, Kouba et al., 2014). *P. clarkii* was introduced in the south of Spain (Gutiérrez-Yurrita et al., 1999) and has invaded almost every basin in the Iberian Peninsula (Kouba et al., 2014). *P. leniusculus* was introduced in the Madrid region and as spread to the colder waters of the historical distribution of *A. italicus* (Gutiérrez-Yurrita et al., 1999; Kouba et al., 2014).

The red swamp crayfish, *P. clarkii* is a typically dark red crayfish and is generally regarded as the most ubiquitous NICS in the world. It is native to the north-eastern Mexico and the southcentral USA (Crandall, 2010; Gherardi, 2006). It is capable of reaching over 50g within 3 to 5 months of age, with adults measuring from 5.5 to 12cm (Global Invasive Species Database, 2011). Reproductively, *P. clarkii*, exhibits traits of an r-selected species, with early maturity at a small body size, rapid growth rates, large egg production (400 pleopodal eggs for an average-sized female), and a short life span. Due to its commercial value, this species has been introduced across the North American continent and later to South America, East Africa, Asia and Europe (Crandall, 2010; Gherardi, 2006). Although *P. clarkii* is usually associated with larger, warmer lowland streams, there are increasing reports of occurrence at higher altitudes and latitudes (Chucholl, 2013).

The signal crayfish, *P. leniusculus*, is a larger crayfish native to the north-western USA and south-western Canada (Johnsen & Taugbøl, 2010), and is the most widespread

NICS in Europe (Kouba et al., 2014). The common name of this specie comes from the white-turquoise patch on the upper-side of the chelae, near the junction of the fix with the movable finger (Lewis, 2002). *P. leniusculus* males can reach up to 16 cm in length and females up to 12 cm, and the weight is typically 60 and 110g, at 50 and 70mm carapace length, respectively (Global Invasive Species Database, 2005). Females produce 200 to 400 eggs that are carried under the tail until hatching. Juvenile crayfishes stay with the mother during three stages (two moults), before they became independent. Sexual maturity is reached at 6-9 cm, corresponding to an age of 2-3 years.

In the watershed of the Sabor river where the study was conducted, these two NICS were first reported in the 1990s. while *P. clarkii* was probably introduced by local people (Bernardo et al., 2011). *P. leniusculus* was introduced by the Spanish authorities in 1994, in two small streams near the international section of Maçãs river (Bernardo et al., 2011).



Figure 2.2 Invasive crayfish species present in the Sabor river catchment. On left the red swamp crayfish, *Procambarus clarkii* (Girard, 1852), and on right the signal crayfish, *Pacifastacus leniusculus* (Dana1852). *P. clarkii* photo by Mertz (2017) and *P. leniusculus* by MMF.

## 2.2.2 Stream fish

Fish are among the most charismatic organisms inhabiting freshwater ecosystems. In stream networks, the fish diversity and abundances tend to be correlated with stream order (Harrel et al., 1967; Paller, 1994; Platts, 1979). The main stems of rivers, usually perennial, tend to have richer and more abundant communities, while headwaters usually present a subset of the species found downstream (Cook et al., 2004) or a community dominated by a few headwater specialists (Meyer et al., 2007). In the absence of major sources of disturbances, stream networks can present very stable fish communities over time (Gorman & Karr, 1978). Although the numbers of the observed species can vary through time, the composition tends to remain stable, with few extinctions and additions disregarding species introductions, also few changes in the more abundant species (Magalhães et al., 2007; Matthews et al., 2013; Matthews & Marsh-Matthews, 2016, 2017). Matthews & Marsh-Matthews (2017) proposes that most fish communities are in a state of 'Loose Equilibria' (S. L. Collins, 2000; DeAngelis et al., 1985), where communities vary across time, but given enough time communities return towards a central condition instead of remaining in an altered state.

In the Iberian Peninsula, as well as the rest of the Mediterranean region, the native ichthyofauna is dominated by the *Cyprinidae* family (Tierno de Figueroa et al., 2013). They occur mainly in rivers that were formed in the transition from the Oligocene to the Pliocene, together with the rise of the main mountains. The absence of canals between rivers during the quaternary left a clear phylogeographic structure in the native Iberian ichthyofauna, which presents a high level of endemism (Gómez & Lunt, 2007). Because of this, closely related endemic species usually have different distributions associated with the main rivers, with a few zones of sympatry that resulted from occasional headwater captures among river basins or human-mediated introduction (Gómez & Lunt, 2007). Besides the native fish fauna, the Iberian Peninsula, there is a large number of introduced freshwater fish in Iberian waters. For instance, in Portugal there are at least twenty non-native fish species recorded (Anastácio et al., 2019). Non-native fish may have multiple impacts on native species and ecosystems, including hybridization, disease transmission, disruption of biotic interactions, predation and competition (Ribeiro & Leunda, 2012). Most of the introductions have been intentional, often to enhance fisheries activities (Anastácio et al., 2019).

In the Sabor River studied in this thesis, the native ichthyofauna belongs to the families *Cyprinidae*, represented by four *genera* and five species, all endemic to the Iberian Peninsula, and *Salmonidae*, with a single species, *Salmo trutta* (Figure 2.3):





Figure 2.3 – Native fish fauna of the Sabor river basin. All photos by MMF.

- *Luciobarbus bocagei* (Steindachner, 1864) – It is the largest of the species observed in the Sabor basin, reaching a length of up to over 60 cm, but it is usually under 30 cm (MMF personal observation). It can have a lifespan of over a decade, with males becoming sexually mature between the third and fourth years, and females between the sixth and eight years (Doadrio et al., 2011; Salvador Milla, 2017). The species has two pairs of barbells located near the mouth that probably helps in the benthonic diet; *L. bocagei* mainly consumes detritus and aquatic insect larvae (Salvador Milla, 2017). Although it is listed as Least Concern in Portugal and by IUCN (Cabral et al., 2005; IUCN, 2020), several of the populations of *L. bocagei* have disappeared from the Douro basin in Spain (Doadrio et al., 2011).

- *Pseudochondrostma duriense* (Coelho, 1985) – It is a medium-sized species, with lengths up to 50 cm (Doadrio et al., 2011), but usually does not surpass the 20 cm (MMF personal observation). It feeds mainly in detritus and vegetal material and occasionally macroinvertebrates using a conspicuous straight corneous lower lip (Sánchez-Hernández et al., 2011). This species makes small upriver migrations to spawn (Doadrio et al., 2011). *P. duriense* is listed as Least Concern in Portugal (Cabral et al., 2005) and as Vulnerable by the IUCN (IUCN, 2020).
- *Squalius alburnoides* (Steindachner, 1866) – This is a complex that resulted from unidirectional hybridization involving females of *Squalius pyrenaicus* with males from an already extinct species from the lineage of *Anaocypris hispanica* (Collares-Pereira et al., 2013; Collares-Pereira & Coelho, 2010). The individuals of this complex are small, the length is usually less than 10 cm (MMF personal observation), and present sexual and asexual reproduction modes that are neither strictly clonal nor hemiclinal (Collares-Pereira et al., 2013). The persistence in natural populations is dependent on the mating with sympatric species of *Squalius*, which in the Sabor basin is *S. carolitertii*, leading to the replacement of the ancestral genome of *S. pyrenaicus* by the genome of the extant species (Collares-Pereira et al., 2013). *S. alburnoides* is listed as Vulnerable in Portugal and by the IUCN (Cabral et al., 2005; IUCN, 2020).
- *Squalius carolitertii* (Doadrio, 1988) – It is a larger and more robust species than *S. alburnoides*, but seldom surpasses the 15 cm of total length (MMF personal observation). The diet consists mainly of small arthropods and small fishes caught on the water column (Doadrio et al., 2011). *S. carolitertii* is listed as Least Concern in Portugal and by the IUCN (Cabral et al., 2005; IUCN, 2020).
- *Achondrostoma* sp. This is a species assigned to *Achondrostoma arcasii* in a paper of this thesis, but molecular evidence suggests that it's a taxon yet to be named (Robalo et al., 2007). This species is morphological similar to *A. arcasii*, presenting the same reddish spots in the fins and not exceeding the 13 cm in length (MMF personal observation). It's an omnivorous species, feeding mainly on detritus and plant material but with aquatic macroinvertebrates as an important complement (Sánchez-Hernández et al., 2011). Because this species is yet to be formally described, its conservation status has not been evaluated yet, but due to its restrict distribution it should be a classification of high threat, like endangered (EN) or critical endangered (CR).

- *Salmo trutta* (Linnaeus, 1758) – is a salmonid species associated with cold headwaters, with a native distribution across Europe, but that has been introduced worldwide in temperate areas (Filipe et al., 2013). In the Iberian Peninsula, there are two coexisting populations with distinct life strategies: migrant populations that complete the life cycle in the ocean, and, more common, resident populations, as the population found in the Sabor basin (Alonso et al., 2017). It is a very appreciated species for sport fishing and as such has been widely targeted for re-stocking and introductions (Piccolo, 2011). *S. trutta* mainly feeds on fish and terrestrial invertebrates that fall or are washed into the water (Alonso et al., 2017). IUCN list this species as Least Concern (IUCN, 2020); in Portugal, while the migrant populations are classified as Critically Endangered (CR), the resident populations are listed as Least Concern (Cabral et al., 2005).

Besides the native species, in the Sabor river basin there are five exotic species registered, including two species of cyprinids, and one species each of *Cobitidae*, *Centrarchidae*, and *Poeciliidae* (Figure 2.4):

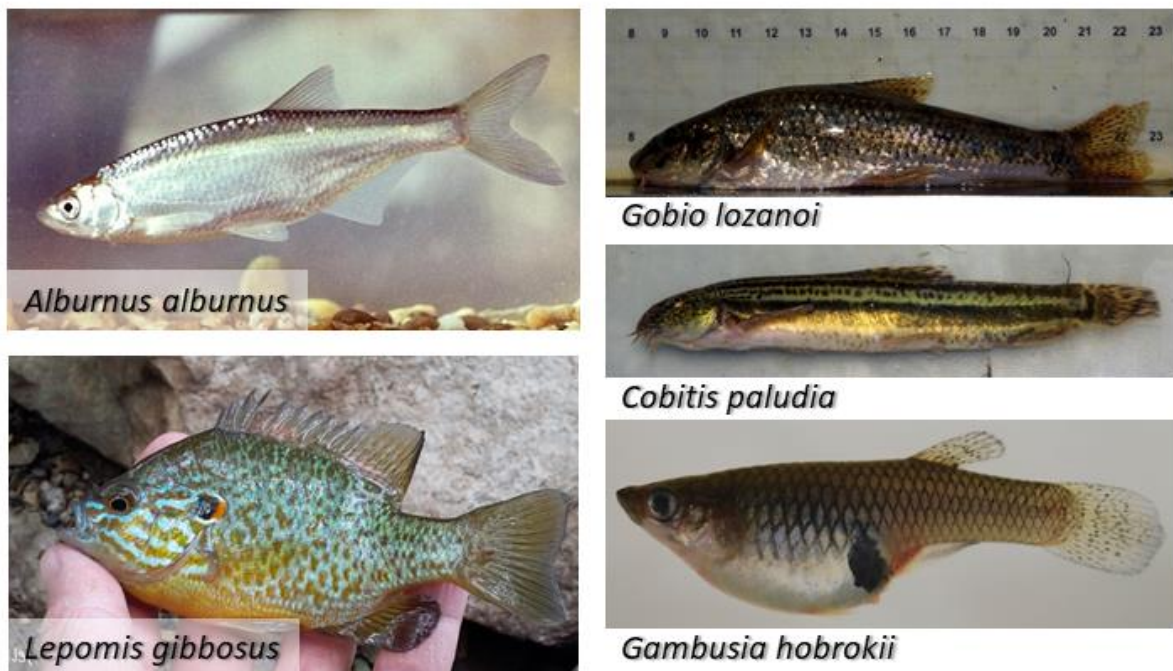


Figure 2.4 - Exotic fish fauna of the Sabor river basin. Photo of *A. alburnus* by Harka (2011). Photo of *G. lozanoi* and *C. paludica* by Perez (2010b, 2010a). Photo of *L. gibbosus* by Wissink (2018). Photo of *G. holbrooki* by Aguilar (2006).

- *Alburnus alburnus* (Linnaeus, 1758) – A medium cyprinid, usually under 17 cm total length (MMF personal observation). It is native to central Europe, but was introduced in reservoirs in Spain as forage fish in the decade of 1990's, and since then has spread to the main basins of the Iberian Peninsula (Vinyoles et al., 2007). The plasticity of the diet and habitat requirements are key in this species success as an invader (David Almeida et al., 2017; Latorre et al., 2020; Matono et al., 2018). *A. alburnus* can have an impact on other cyprinids of similar size at several levels: It is a strong competitor for habitat and food, and a predator of eggs clutches (Latorre & Almeida, 2019). It is also capable of hybridization with endemic *Squalius* species (Almodóvar et al., 2012).
- *Gobio lozanoi* (Doadrio & Madeira, 2004) – A small cyprinid, usually not larger than 12 cm (MMF personal observation). It is native to the south of France and east of the Iberian Peninsula, and has spread to the main basins of the Iberian Peninsula during the XXth century (Doadrio et al., 2011). It is a benthonic feeder and consumes mainly small arthropods (Oscoz et al., 2006). The spread of this species has been favoured by impoundments and river regulation (Muñoz-Mas et al., 2016).
- *Cobitis paludica* (de Buen, 1930) – A small species native to the south of the Iberian Peninsula, which is regarded as introduced in the Douro basin (Doadrio et al., 2011), although this status has been contested by some authors (Sánchez-Hernández et al., 2018). Accidental release due to the traditional use of this species as bait should account for the majority of the introductions. It is a benthonic feeder that consumes mainly chironomid larvae and ostracods (Sánchez-Carmona et al., 2008). *C. paludica* is a native of the south of Portugal, as such it doesn't have an exotic status (Anastácio et al., 2019) and is listed by the IUCN as Vulnerable (IUCN, 2020).
- *Lepomis gibbosus* (Linnaeus, 1758) – A small centrarchid species, usually less than 10 cm (MMF personal observation). It is native to east-central North America, and is presently distributed across most of western and central Europe as a result of introductions going back more than 100 years (Bhagat et al., 2006). *L. gibbosus* was introduced by Spanish authorities, over 100 years, as forage fish to previously introduced piscivore species (Elvira & Almodóvar, 2001). In Portugal, the expansion of this species is related to introduction in reservoirs by anglers and subsequent expansion (Godinho et al., 1998). *L. gibbosus* is an opportunistic and plastic forager (D. Almeida et al., 2009), and can be more efficient forager than native species

(Anastácio et al., 2019). The males of this species present parental care behaviour (Naspleda et al., 2017).

- *Gambusia holbrooki* (Girard, 1859) – A small poeciliid species native to the eastern coast of North America that has been introduced worldwide as a mosquito control agent (Pyke, 2008). The species was introduced in Spain in 1921, and has spread to almost every basin in the Iberian Peninsula (Doadrio et al., 2011; Moreno-Valcárcel & Ruiz-Navarro, 2017). Although its efficiency as a control agent has been disputed, it is widely regarded as an aggressive invader with a negative impact on invertebrates, amphibians and other fish (Pyke, 2008).

### 2.2.3 Amphibians

Amphibians populations worldwide are declining fast (Alford & Richards, 1999; McCallum, 2007; Wake, 1991), and are presently considered the most threatened faunal group (IUCN, 2020). Although no single global cause has been identified, six leading causes have been advanced, which may act independently or in multiple interactions (J. P. Collins & Storer, 2003): alien species introduction, over-exploitation, land-use change, global alterations, contaminants, and emerging diseases. In recent years, the panzootic chytridiomycosis, caused by the fungus *Batrachochytrium dendrobatidis*, which emerged in Asia and has spread with the help of global trade of amphibians, has been linked to the decline of amphibian populations worldwide, particularly in stream associated populations on remote and pristine habitats (Fisher & Garner, 2020; O'Hanlon et al., 2018). In Europe, however, amphibian declines are commonly associated with habitat destruction and/or alteration, and invasive species (Cruz et al., 2008; Ferreira & Beja, 2013; Nyström et al., 2007; Stuart et al., 2004).

The most common life cycle of amphibians involves three distinct stages: aquatic eggs and larvae, and usually more water independent adult stage (Pough et al., 1998). Females usually lay jelly-covered eggs in aquatic habitats that hatch into larvae that subsequently metamorphose into adults. Anuran larvae (tadpoles) are hatched limbless, developing hind limbs first and forelimbs just before the end of the metamorphosis. The digestive track also suffers important changes, both morphological and physiological, as the individual transitions from a suspension-feeding tadpole to a carnivorous adult (Hourdry et al., 1996). Urodele larvae develop both pairs of limbs in the early stages of development and are always carnivore (Pough et al., 1998). The changes that occur

during metamorphosis are controlled by hormones, but are also influenced by external factors such as temperature and conspecific larvae density (Pough et al., 1998).

While the highest diversity of amphibians is usually associated with ponds and pools, particularly in the Mediterranean (Beja & Alcazar, 2003; Ferreira & Beja, 2013; Gómez-Rodríguez et al., 2010; Snodgrass et al., 2000), small streams can sustain high amphibian diversity, particularly in regions where ponds are scarce (De Vries & Marco, 2017). Larger river and streams are regarded as unsuitable habitat for most amphibians due to hydrology and predation risk (De Vries & Marco, 2017), but amphibians are particularly adapted to small intermittent streams (Sánchez-Montoya et al., 2017). The predictable shift between dry and wet phases of headwaters and small streams that are unsuitable for predators, such as fish and crayfish, but can provide food resources such as algae and macroinvertebrates, makes these habitats particularly valuable for amphibians (De Vries & Marco, 2017; Meyer et al., 2007; Sánchez-Montoya et al., 2017).

In the Iberian Peninsula, the fauna of amphibians includes a large number of endemic or quasi-endemic species. The current distribution of the Iberian amphibians results from a complex process of contractions during the quaternary ice ages, and posterior expansion during the interglacial periods (Gómez & Lunt, 2007). Most of the current genera of Iberian amphibians were already present in what is now the Iberian Peninsula during the Miocene (23.8 – 5.3 MY, Pinho & de Almeida, 2008).

The batrachofauna in the system studied in this thesis includes at least thirteen species (Loureiro et al., 2008), of which only eight species present a strong association with stream habitats: three urodele species from the family *Salamandridae*, and five anuran species from three families, two species each from *Alytidae* and *Ranidae*, and one species from *Bufo* (Figure 2.5).

- *Salamandra salamandra* (Linnaeus, 1758) – It is a medium-sized salamandrid species, which usually does not reach more than 200mm in length, and is distributed across Europe (Sillero et al., 2014; Velo-Antón & Buckley, 2015). This species presents high geographical variation, with at least ten subspecies recognized for the Iberian Peninsula (Burgon et al., 2021). *S. s. galaica*, the subspecies present in the Sabor river basin, presents oviparism as the most common reproduction strategy (Velo-Antón & Buckley, 2015). In mesocosm experiments, larvae of *S. salamandra*

were predated by *P. clarkii*, although it did not significantly reduce larvae survival (Cruz & Rebelo, 2005).

- *Triturus marmoratus* (Latreille, 1800) – With a usual length between 110 and 160 mm, this species is distributed from Portugal to France (Montori, 2014; Sillero et al., 2014). The aquatic phase of the adults usually begins in November or December, and may extend until May at lower altitudes, while at higher altitudes it often lasts from May to November (Caetano & Castanet, 1993). The oviposition strategy of females of this species, which involves laying single eggs in submerged vegetation by bending the leaves (Montori, 2014), may render this species more susceptible to predation by crayfish. Possibly in association with this, the introduction of *P. clarkii* can drive local populations of *T. marmoratus* to extinction (Cruz et al., 2008).
- *Lissonotriton boscai* (Lataste, 1879) – A small salamandrid that usually does not reach more than 100 mm in length, and is endemic to the western part of the Iberian Peninsula (Díaz-Paniagua, 2014; Sillero et al., 2014). Adults of this species may stay in aquatic habitats for a large part of the year if conditions are favourable (Caetano & Leclair Jr., 1999). The oviposition strategy of this species is similar to *T. marmoratus* (Díaz-Paniagua, 2014). In mesocosm experiments, predation by *P. clarkii* significantly reduces larvae survival (Cruz & Rebelo, 2005).
- *Alytes cisternasii* (Boscá, 1879) – A small anuran, usually smaller than 5 cm, which is endemic to the southwest of the Iberian Peninsula (Márquez, 2017; Reino et al., 2017; Sillero et al., 2014). Males of this genus wrap the fertilized egg chord around their hind legs, and carries it on land during the embryonic development, only releasing the eggs mass in a water body when the tadpoles are fully formed (Márquez, 1992). This species is more associated with xeric habitats and temporary water masses than its sister species *A. obstetricans*, and as such its reproduction activity tends to be explosive after the first autumn rains (Márquez, 1992, 2017; Reino et al., 2017). In mesocosm experiments, predation by *P. clarkii* significantly reduced larvae survival (Cruz & Rebelo, 2005).
- *Alytes obstetricans* (Laurenti, 1758) – Very similar to its sister species, *A. cisternasii* has a wider distribution from Portugal to Germany (Bosch, 2014; Sillero et al., 2014). In the Iberian Peninsula, the species from this genus present a parapatric distribution with a small sympatric zone, with *A. obstetricans* being distributed in the northern half and associated with regions with high precipitation (Reino et al., 2017). The peak of

the reproduction activity is at the end of spring, and larvae may stay in the water for several years (Bosch, 2014; Márquez, 1992).

- *Bufo spinosus* (Daudin, 1803) – The largest of the Iberian anurans, this species is distributed from the Northwest of Africa, across the Iberian Peninsula, to the south and west of France (Ortiz-Santaliestra, 2014; Recuero et al., 2012). It reproduces preferably in streams and permanent water bodies, including irrigation reservoirs (Ferreira & Beja, 2013). In mesocosm experiments, *P. clarkii* did not predate on eggs and embryos, but rapidly consumes newly hatched tadpoles (Cruz & Rebelo, 2005).
- *Pelophylax perezi* (López Seoane, 1885) – A medium-size ranid, usually larger than 5 cm, widely distributed in the Iberian Peninsula and south and west coast of France (Egea-Serrano, 2014; Sillero et al., 2014). *P. perezi* is mostly associated with aquatic habitats, although it is capable of land dispersion. It is the most common amphibian in the Iberian Peninsula, occupying almost every existent water body, including man-made (Egea-Serrano, 2014; Ferreira & Beja, 2013). In mesocosm experiments, *P. clarkii* predate eggs and tadpoles of *P. perezi*, and significantly reduces their survival (Cruz & Rebelo, 2005). However, in natural populations, a positive correlation has been found between the presence of *P. clarkii* and *P. perezi* (Cruz, Rebelo, et al., 2006).
- *Rana iberica* (Boulenger, 1879) – A small ranid, usually not larger than 5 cm, endemic to the Northwest part of the Iberian Peninsula, where the Atlantic climatic influence is stronger (Salvador, 2015; Sillero et al., 2014). Like *P. perezi*, it is a mostly aquatic species, but usually only occupies small, fast, and cold mountain streams (Salvador, 2015).

In the IUCN Red List assessment, *Rana iberica* is classified as vulnerable and is the only species with a threat conservation status, while all others are classified as least concern. However, with the exceptions of *B. spinosus* and *L. boscai*, all species show a decreasing trend (IUCN, 2020). This assessment listed as major threats to these species the loss and modification of habitats, invasive species and emergent diseases (IUCN, 2020). In Portugal, none of these species is considered threatened (Cabral et al., 2005).



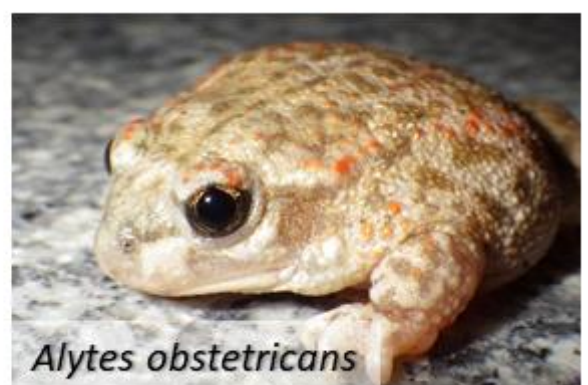


Figure 2.5 - Amphibians associated with streams of the Sabor river region. All photos by MMF.

## 2.3 References

- Aguilar, R. (2006). *Gambusia holbrokii*. Wikimedia Commons, the free media repository. [https://commons.wikimedia.org/wiki/File:Gambusia\\_holbrooki\\_\(S0833\)\\_\(12597857535\).jpg](https://commons.wikimedia.org/wiki/File:Gambusia_holbrooki_(S0833)_(12597857535).jpg)
- Alderman, D. J., Holdich, D., & Reeve, I. (1990). Signal crayfish as vectors in crayfish plague in Britain. *Aquaculture*, 86(1), 3–6. [https://doi.org/10.1016/0044-8486\(90\)90216-A](https://doi.org/10.1016/0044-8486(90)90216-A)
- Alford, R. A., & Richards, S. J. (1999). Global amphibian declines: a problem in applied ecology. *Annual Review of Ecology and Systematics*, 30, 133–165. <http://www.jstor.org/stable/221682>
- Almeida, D., Almodóvar, A., Nicola, G. G., & Elvira, B. (2009). Feeding tactics and body condition of two introduced populations of pumpkinseed *Lepomis gibbosus*: taking advantages of human disturbances? *Ecology of Freshwater Fish*, 18(1), 15–23. <https://doi.org/10.1111/j.1600-0633.2008.00317.x>
- Almeida, David, Fletcher, D. H., Rangel, C., García-Berthou, E., & da Silva, E. (2017). Dietary traits of invasive bleak *Alburnus alburnus* (Actinopterygii, Cyprinidae) between contrasting habitats in Iberian fresh waters. *Hydrobiologia*, 795(1), 23–33. <https://doi.org/10.1007/s10750-016-3052-8>
- Almodóvar, A., Nicola, G. G., Leal, S., Torralva, M., & Elvira, B. (2012). Natural hybridization with invasive bleak *Alburnus alburnus* threatens the survival of Iberian endemic calandino *Squalius alburnoides* complex and Southern Iberian chub *Squalius pyrenaicus*. *Biological Invasions*, 14(11), 2237–2242. <https://doi.org/10.1007/s10530-012-0241-x>
- Alonso, C., Gortázar, J., & García de Jalón, D. (2017). Trucha común – *Salmo trutta* Linnaeus, 1758. In *Enciclopedia Virtual de Vertebrados Españoles* (3rd ed.). CSIC - Museo Nacional de Ciencias Naturales (MNCN). <https://doi.org/10.20350/digitalCSIC/8847>
- Anastácio, P. M., Ribeiro, F., Capinha, C., Banha, F., Gama, M., Filipe, A. F., Rebelo, R., & Sousa, R. (2019). Non-native freshwater fauna in Portugal: A review. *Science of The Total Environment*, 650, 1923–1934. <https://doi.org/10.1016/j.scitotenv.2018.09.251>
- Azevedo, J. C., Moreira, C., Castro, J. P., & Loureiro, C. (2011). Agriculture Abandonment, Land-use Change and Fire Hazard in Mountain Landscapes in Northeastern Portugal. In C. Li, R. Laforzezza, & J. Chen (Eds.), *Landscape Ecology in Forest Management and Conservation: Challenges and Solutions for Global Change* (pp. 329–351). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-12754-0\\_14](https://doi.org/10.1007/978-3-642-12754-0_14)
- Beja, P., & Alcazar, R. (2003). Conservation of Mediterranean temporary ponds under agricultural intensification: an evaluation using amphibians. *Biological Conservation*, 114(3), 317–326. [https://doi.org/10.1016/S0006-3207\(03\)00051-X](https://doi.org/10.1016/S0006-3207(03)00051-X)
- Bernardo, J. M., Costa, a. M., Bruxelas, S., & Teixeira, a. (2011). Dispersal and coexistence of two non-native crayfish species ( *Pacifastacus leniusculus* and *Procambarus clarkii* ) in NE Portugal over a 10-year period. *Knowledge and Management of Aquatic Ecosystems*, 401, 28. <https://doi.org/10.1051/kmae/2011047>
- Bhagat, Y., Fox, M. G., & Ferreira, M. T. (2006). Morphological differentiation in introduced pumpkinseed *Lepomis gibbosus* (L.) occupying different habitat zones in Portuguese reservoirs. *Journal of Fish Biology*, 69(sc), 79–94. <https://doi.org/10.1111/j.1095-8649.2006.01267.x>
- Bosch, J. (2014). Sapo partero común – *Alytes obstetricans* (Laurenti, 1768). In *la Enciclopedia Virtual de Vertebrados Españoles* (10th ed.). CSIC - Museo Nacional de Ciencias Naturales

(MNCN). <http://www.vertebradosibericos.org/>

- Brannelly, L., McMahon, T., Hinton, M., Lenger, D., & Richards-Zawacki, C. (2015). Batrachochytrium dendrobatidis in natural and farmed Louisiana crayfish populations: prevalence and implications. *Diseases of Aquatic Organisms*, 112(3), 229–235. <https://doi.org/10.3354/dao02817>
- Burgon, J. D., Vences, M., Steinfartz, S., Bogaerts, S., Bonato, L., Donaire-Barroso, D., Martínez-Solano, I., Velo-Antón, G., Vieites, D. R., Mable, B. K., & Elmer, K. R. (2021). Phylogenomic inference of species and subspecies diversity in the Palearctic salamander genus Salamandra. *Molecular Phylogenetics and Evolution*, 157, 107063. <https://doi.org/10.1016/j.ympev.2020.107063>
- Cabral, M. J., Almeida, J., Almeida, P. R., Dellinger, T., Ferrand de Almeida, N., Oliveira, M. E., Palmeirim, J. M., Queirós, A. I., Rogado, L., & Santos-Reis, M. (2005). *Livro Vermelho dos Vertebrados de Portugal*. Instituto da Conservação da Natureza.
- Caetano, M. H., & Castanet, J. (1993). Variability and microevolutionary patterns in Triturus marmoratus from Portugal: age, size, longevity and individual growth. *Amphibia-Reptilia*, 14(2), 117–129. <https://doi.org/10.1163/156853893X00291>
- Caetano, M. H., & Leclair Jr., R. (1999). Comparative Phenology and Demography of Triturus boscai from Portugal. *Journal of Herpetology*, 33(2), 192. <https://doi.org/10.2307/1565714>
- Churchill, C. (2013). Feeding ecology and ecological impact of an alien “warm-water” omnivore in cold lakes. *Limnologia*, 43(4), 219–229. <https://doi.org/10.1016/j.limno.2012.10.001>
- Clavero, M., Nores, C., Kubersky-Piredda, S., & Centeno-Cuadros, A. (2016). Interdisciplinarity to reconstruct historical introductions: solving the status of cryptogenic crayfish. *Biological Reviews*, 91(4), 1036–1049. <https://doi.org/10.1111/brv.12205>
- Collares-Pereira, M. J., & Coelho, M. M. (2010). Reconfirming the hybrid origin and generic status of the Iberian cyprinid complex Squalius alburnoides. *Journal of Fish Biology*, 76(3), 707–715. <https://doi.org/10.1111/j.1095-8649.2009.02460.x>
- Collares-Pereira, M. J., Matos, I., Morgado-Santos, M., & Coelho, M. M. (2013). Natural pathways towards polyploidy in animals: The Squalius alburnoides fish complex as a model system to study genome size and genome reorganization in polyploids. *Cytogenetic and Genome Research*, 140(2–4), 97–116. <https://doi.org/10.1159/000351729>
- Collins, J. P., & Storer, A. (2003). Global amphibian declines: sorting the hypotheses. *Diversity & Distributions*, 9(2), 89–98. <https://doi.org/10.1046/j.1472-4642.2003.00012.x>
- Collins, S. L. (2000). Disturbance frequency and community stability in native tallgrass prairie. *American Naturalist*, 155(3), 311–325. <https://doi.org/10.1086/303326>
- Cook, R. R., Angermeier, P. L., Finn, D. S., Poff, N. L. R., & Krueger, K. L. (2004). Geographic variation in patterns of nestedness among local stream fish assemblages in Virginia. *Oecologia*, 140(4), 639–649. <https://doi.org/10.1007/s00442-004-1618-z>
- Correia, A. M. (2001). Seasonal and interspecific evaluation of predation by mammals and birds on the introduced red swamp crayfish Procambarus clarkii (Crustacea, Cambaridae) in a freshwater marsh (Portugal). *Journal of Zoology*, 255(4), 533–541. <https://doi.org/10.1017/s0952836901001625>
- Crandall, K. A. (2010). Procambarus clarkii, Red Swamp Crayfish. In *The IUCN Red List of*

*Threatened Species 2010* (Vol. 8235).

- Cruz, M. J., Pascoal, S., Tejedo, M., & Rebelo, R. (2006). Predation by an exotic crayfish, *Procambarus clarkii*, on natterjack toad, *Bufo calamita*, embryos: its role on the exclusion of this amphibian from its breeding ponds. *Copeia*, 2(2), 274–280. [https://doi.org/10.1643/0045-8511\(2006\)6\[274:PBAECP\]2.0.CO;2](https://doi.org/10.1643/0045-8511(2006)6[274:PBAECP]2.0.CO;2)
- Cruz, M. J., & Rebelo, R. (2005). Vulnerability of Southwest Iberian amphibians to an introduced crayfish, *Procambarus clarkii*. *Amphibia-Reptilia*, 26(3), 293–303. <https://doi.org/10.1163/156853805774408577>
- Cruz, M. J., Rebelo, R., & Crespo, E. G. (2006). Effects of an introduced crayfish, *Procambarus clarkii*, on the distribution of south-western Iberian amphibians in their breeding habitats. *Ecography*, 29(3), 329–338. <https://doi.org/10.1111/j.2006.0906-7590.04333.x>
- Cruz, M. J., Segurado, P., Sousa, M., & Rebelo, R. (2008). Collapse of the amphibian community of the Paul do Boquilobo Natural Reserve (central Portugal) after the arrival of the exotic American crayfish *Procambarus clarkii*. *Herpetological Journal*, 18(4), 197–204. <http://cat.inist.fr/?aModele=afficheN&cpsidt=21982001>
- De Vries, W., & Marco, A. (2017). The importance of fluvial habitats for amphibian conservation in the Mediterranean climate. *Basic and Applied Herpetology*, 31, 5–16. <https://doi.org/10.11160/bah.56>
- DeAngelis, D. L., Waterhouse, J. C., Post, W. M., & O'Neill, R. V. (1985). Ecological modelling and disturbance evaluation. *Ecological Modelling*, 29(1–4), 399–419. [https://doi.org/10.1016/0304-3800\(85\)90063-8](https://doi.org/10.1016/0304-3800(85)90063-8)
- Díaz-Paniagua, C. (2014). Tritón ibérico – *Lissotriton boscai*. In *Enciclopedia Virtual de los Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org/>
- Diéguez-Uribeondo, J., & Söderhäll, K. (1993). *Procambarus clarkii* Girard as a vector for the crayfish plague fungus, *Aphanomyces astaci* Schikora. *Aquaculture Research*, 24(6), 761–765. <https://doi.org/10.1111/j.1365-2109.1993.tb00655.x>
- Doadrio, I., Perea, S., Garzón-Heydt, P., & González, J. L. (2011). *Ictiofauna continental española. Bases para su seguimiento*. (DG Medio Natural y Política Forestal (ed.)). Ministerio de Medio Ambiente y Medio Rural y Marino, Centro de Publicaciones.
- Egea-Serrano, A. (2014). Rana común – *Pelophylax perezi*. In *Enciclopedia Virtual de los Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org>
- Elvira, B., & Almodóvar, A. (2001). Freshwater fish introductions in Spain: Facts and figures at the beginning of the 21st century. *Journal of Fish Biology*, 59(SUPPL. A), 323–331. <https://doi.org/10.1006/jfbi.2001.1753>
- Ferreira, M., & Beja, P. (2013). Mediterranean amphibians and the loss of temporary ponds: Are there alternative breeding habitats? *Biological Conservation*, 165(null), 179–186. <https://doi.org/10.1016/j.biocon.2013.05.029>
- Filipe, A. F., Markovic, D., Pletterbauer, F., Tisseuil, C., De Wever, A., Schmutz, S., Bonada, N., & Freyhof, J. (2013). Forecasting fish distribution along stream networks: brown trout (*Salmo trutta*) in Europe. *Diversity and Distributions*, n/a-n/a. <https://doi.org/10.1111/ddi.12086>
- Fisher, M. C., & Garner, T. W. J. (2020). Chytrid fungi and global amphibian declines. *Nature*

- Reviews Microbiology*, 18(6), 332–343. <https://doi.org/10.1038/s41579-020-0335-x>
- Gamradt, S. C., Kats, L. B., & Anzalone, C. B. (1997). Aggression by non-native crayfish deters breeding in California Newts. *Conservation Biology*, 11(3), 793–796. <https://doi.org/10.1046/j.1523-1739.1997.96230.x>
- Gherardi, Francesca. (2006). Crayfish invading Europe: The case study of *Procambarus clarkii*. In *Marine and Freshwater Behaviour and Physiology* (Vol. 39, Issue 3, pp. 175–191). Taylor & Francis Group. <https://doi.org/10.1080/10236240600869702>
- Gherardi, Francesca. (2007). Understanding the impact of invasive crayfish. In Francesca Gherardi (Ed.), *Biological invaders in inland waters: Profiles, distribution, and threats* (pp. 507–542). Springer Netherlands. [https://doi.org/10.1007/978-1-4020-6029-8\\_28](https://doi.org/10.1007/978-1-4020-6029-8_28)
- Gherardi, Francesca. (2012). Public Perception of Invasive Alien Species in Mediterranean Europe. In I. D. Rotherham & R. A. Lambert (Eds.), *Invasive and Introduced Plants and Animals Human Perceptions, Attitudes and Approaches to Management*. Routledge.
- Gherardi, Francesca, Aquiloni, L., Diéguez-Urbeondo, J., & Tricarico, E. (2011). Managing invasive crayfish: Is there a hope? In *Aquatic Sciences* (Vol. 73, Issue 2, pp. 185–200). SP Birkhäuser Verlag Basel. <https://doi.org/10.1007/s00027-011-0181-z>
- Global Invasive Species Database. (2005). *Species profile Pacifastacus leniusculus*. <http://www.iucngisd.org/gisd/species.php?sc=725>
- Global Invasive Species Database. (2011). *Species profile Procambarus carkii* (Issue December). <http://www.iucngisd.org/gisd/species.php?sc=608>
- Godinho, F. N., Ferreira, M. T., & Castro, M. I. P. (1998). Fish assemblage composition in relation to environmental gradients in Portuguese reservoirs. *Aquatic Living Resources*, 11(5), 325–334. [https://doi.org/10.1016/S0990-7440\(98\)80003-X](https://doi.org/10.1016/S0990-7440(98)80003-X)
- Gómez-Rodríguez, C., Díaz-Paniagua, C., Bustamante, J., Serrano, L., & Portheault, A. (2010). Relative importance of dynamic and static environmental variables as predictors of amphibian diversity patterns. *Acta Oecologica*, 36(6), 650–658. <https://doi.org/10.1016/j.actao.2010.10.002>
- Gómez, A., & Lunt, D. H. (2007). Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula. In S. Weiss & N. Ferrand (Eds.), *Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity* (pp. 155–188). Springer Netherlands. [https://doi.org/10.1007/1-4020-4904-8\\_5](https://doi.org/10.1007/1-4020-4904-8_5)
- Gorman, O. T., & Karr, J. R. (1978). Habitat Structure and Stream Fish Communities. *Ecology*, 59(3), 507–515. <https://doi.org/10.2307/1936581>
- Gutiérrez-Yurrita, P., Martínez, J., Ilhéu, M., Bravo-Utrera, M., Bernardo, J., & Montes, C. (1999). The status of crayfish populations in Spain and Portugal. In F. Gherardi & D. Holdich (Eds.), *Crayfish in Europe as Alien Species* (1st ed., Vol. 11, pp. 161–192). Balkema, Rotterdam/Brookfield.
- Harka, A. (2011). *Alburnus alburnus*. Wikimedia Commons, the free media repository. [https://commons.wikimedia.org/wiki/File:Alburnus\\_alburnus\\_Hungary.jpg](https://commons.wikimedia.org/wiki/File:Alburnus_alburnus_Hungary.jpg)
- Harrel, R. C., Davis, B. J., & Dorris, T. C. (1967). Stream Order and Species Diversity of Fishes in an Intermittent Oklahoma Stream. *American Midland Naturalist*, 78(2), 428. <https://doi.org/10.2307/2485240>

- Hoelzer, A. (2003). *Vegetation ecological studies at the lower course of Sabor River (Tras-os-Montes, NE-Portugal)* [University of Bremen]. <http://www.grin.com/en/e-book/28444/vegetation-ecological-studies-at-the-lower-course-of-sabor-river-tras-os-montes>
- Hourdry, J., L'Hermite, A., & Ferrand, R. (1996). Changes in the Digestive Tract and Feeding Behavior of Anuran Amphibians during Metamorphosis. *Physiological Zoology*, 69(2), 219–251. <https://doi.org/10.1086/physzool.69.2.30164181>
- Ilhéu, M., Bernardo, J. M., & Fernandes, S. (2007). Predation of invasive crayfish on aquatic vertebrates: the effect of *Procambarus clarkii* on fish assemblages in Mediterranean temporary streams. In Francesca Gherardi (Ed.), *Biological invaders in inland waters: Profiles, distribution, and threats* (pp. 543–558). Springer Netherlands. [https://doi.org/10.1007/978-1-4020-6029-8\\_29](https://doi.org/10.1007/978-1-4020-6029-8_29)
- IUCN. (2020). The IUCN Red List of Threatened Species. Version 2020-3. <https://www.iucnredlist.org>
- Jackson, A. L. R. R. (2011). Renewable energy vs. biodiversity: Policy conflicts and the future of nature conservation. *Global Environmental Change*, 21(4), 1195–1208. <https://doi.org/10.1016/J.GLOENVCHA.2011.07.001>
- Johnsen, S. I., & Taugbøl, T. (2010). NOBANIS - Invasive Alien Species Fact Sheet *Pacifastacus leniusculus*. *Online Database of the European Network on Invasive Alien Species – NOBANIS* [Www.Nobanis.Org](http://www.Nobanis.Org), 2624, 1–9.
- Junning, C., Zhou, X., Xue, Y., Lucentea, D., & Laganaa, C. (2020). Top 10 species groups in global aquaculture 2017. In *Food and Agriculture Organization of the United Nations: Vol. June*. <http://www.fao.org/3/ca5224en/ca5224en.pdf>
- Kouba, A., Petrussek, A., & Kozák, P. (2014). Continental-wide distribution of crayfish species in Europe: update and maps. *Knowledge and Management of Aquatic Ecosystems*, 413, 05. <https://doi.org/10.1051/kmae/2014007>
- Latorre, D., & Almeida, D. (2019). Alburno – *Alburnus alburnus* (Linnaeus, 1758). In *Enciclopedia Virtual de los Vertebrados Españoles*. CSIC - Museo Nacional de Ciencias Naturales (MNCN). <https://doi.org/10.20350/digitalCSIC/8940>
- Latorre, D., Masó, G., Hinckley, A., Verdiell-Cubedo, D., Castillo-García, G., González-Rojas, A. G., Black-Barbour, E. N., Vila-Gispert, A., García-Berthou, E., Miranda, R., Oliva-Paterna, F. J., Ruiz-Navarro, A., da Silva, E., Fernández-Delgado, C., Cucherousset, J., Serrano, J. M., & Almeida, D. (2020). Interpopulation Variability in Dietary Traits of Invasive Bleak *Alburnus alburnus* (Actinopterygii, Cyprinidae) Across the Iberian Peninsula. *Water*, 12(8), 2200. <https://doi.org/10.3390/w12082200>
- Lewis, S. D. (2002). *Pacifastacus*. In *Biology of freshwater crayfish* (pp. 511–540). Blackwell Science Ltd.
- Loureiro, A., Ferrand, N., Carretero, M. a., & Paulo, O. (2008). Atlas dos anfíbios e répteis de Portugal. In Armando Loureiro, N. Ferrand, M. Carretero, & O. Paulo (Eds.), *Instituto da Conservação da Natureza e da Biodiversidade, Lisboa*. ICNB. <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Atlas+dos+Anf?bios+e+R?pteis+de+Portugal#0>
- Magalhães, M. F., Beja, P., Schlosser, I. J., & Collares-Pereira, M. J. (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, 52(8), 1494–1510. <https://doi.org/10.1111/j.1365-2427.2007.01781.x>

- Márquez, R. (1992). Terrestrial paternal care and short breeding seasons: reproductive phenology of the midwife toads *Alytes obstetricans* and *A. cisternasii*. *Ecography*, 15(3), 279–288. <https://doi.org/10.1111/j.1600-0587.1992.tb00036.x>
- Márquez, R. (2017). Sapo partero ibérico – *Alytes cisternasii* Boscá, 1879. In *Enciclopedia Virtual de los Vertebrados Españoles* (p. 18). Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org/>
- Matono, P., da Silva, J., & Ilhéu, M. (2018). How does an invasive cyprinid benefit from the hydrological disturbance of Mediterranean temporary streams? *Diversity*, 10(2), 1–22. <https://doi.org/10.3390/D10020047>
- Matthews, W. J., & Marsh-Matthews, E. (2016). Dynamics of an upland stream fish community over 40 years: Trajectories and support for the loose equilibrium concept. *Ecology*, 97(3), 706–719. <https://doi.org/10.1890/14-2179.1>
- Matthews, W. J., & Marsh-Matthews, E. (2017). Stream fish community dynamics: A critical synthesis. In *Stream Fish Community Dynamics: A Critical Synthesis*. JHU Press.
- Matthews, W. J., Marsh-Matthews, E., Cashner, R. C., & Gelwick, F. (2013). Disturbance and trajectory of change in a stream fish community over four decades. *Oecologia*, 173(3), 955–969. <https://doi.org/10.1007/s00442-013-2646-3>
- McCallum, M. L. (2007). Amphibian Decline or Extinction? Current Declines Dwarf Background Extinction Rate. *Journal of Herpetology*, 41(3), 483–491. [https://doi.org/10.1670/0022-1511\(2007\)41\[483:ADOECD\]2.0.CO;2](https://doi.org/10.1670/0022-1511(2007)41[483:ADOECD]2.0.CO;2)
- Meyer, J. L., Strayer, D. L., Wallace, J. B. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks. *JAWRA Journal of the American Water Resources Association*, 43(1), 86–103. <https://doi.org/10.1111/j.1752-1688.2007.00008.x>
- Monteiro-Henriques, T. (2010). Landscape and phytosociology of the Paiva river's hydrographical basin and contiguous basins of the Douro River's left margin, from the Paiva to the Tedo River. Instituto Superior de Agronomia – TULisbon.
- Montori, A. (2014). Tritón jaspeado – *Triturus marmoratus*. In *Enciclopedia Virtual de los Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org/>
- Moreno-Valcárcel, R., & Ruiz-Navarro, A. (2017). Gambusia – *Gambusia holbrooki*. In *Enciclopedia Virtual de Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org/>
- Muñoz-Mas, R., Fukuda, S., Vezza, P., & Martínez-Capel, F. (2016). Comparing four methods for decision-tree induction: A case study on the invasive Iberian gudgeon (*Gobio lozanoi*; Doadrio and Madeira, 2004). *Ecological Informatics*, 34, 22–34. <https://doi.org/10.1016/j.ecoinf.2016.04.011>
- Naspleda, J., Zamora, L., & Vila-Gispert, A. (2017). Pez sol – *Lepomis gibbosus*. In *Enciclopedia virtual de los Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org>
- Nunes, A. L. (2011). Predator-prey interactions between an invasive crayfish predator and native anurans: ecological and evolutionary outcomes. Faculdade de Ciências da Universidade de Lisboa.

- Nyström, P., & Åbjörnsson, K. (2000). Effects of fish chemical cues on the interactions between tadpoles and crayfish. *Oikos*, *88*(1), 181–190. <https://doi.org/10.1034/j.1600-0706.2000.880120.x>
- Nyström, P., Brönmark, C., & Granéli, W. (1996). Patterns in benthic food webs: A role for omnivorous crayfish? *Freshwater Biology*, *36*(3), 631–646. <https://doi.org/10.1046/j.1365-2427.1996.d01-528.x>
- Nyström, P., Hansson, J., Månsson, J., Sundstedt, M., Reslow, C., & Broström, A. (2007). A documented amphibian decline over 40 years: Possible causes and implications for species recovery. *Biological Conservation*, *138*(3–4), 399–411. <https://doi.org/10.1016/j.biocon.2007.05.007>
- O'Hanlon, S. J., Rieux, A., Farrer, R. A., Rosa, G. M., Waldman, B., Bataille, A., Kosch, T. A., Murray, K. A., Brankovics, B., Fumagalli, M., Martin, M. D., Wales, N., Alvarado-Rybak, M., Bates, K. A., Berger, L., Böll, S., Brookes, L., Clare, F., Courtois, E. A., ... Fisher, M. C. (2018). Recent Asian origin of chytrid fungi causing global amphibian declines. *Science*, *360*(6389), 621–627. <https://doi.org/10.1126/science.aar1965>
- Ortiz-Santaliestra, M. (2014). Sapo común – Bufo spinosus. In *Enciclopedia Virtual de los Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid.
- Oscos, J., Leunda, P. M., Miranda, R., & Escala, M. C. (2006). Summer feeding relationships of the co-occurring Phoxinus phoxinus and Gobio lozanoi (Cyprinidae) in an Iberian river. *Folia Zoologica*, *55*(4), 418–432.
- Paller, M. H. (1994). Relationships between Fish Assemblage Structure and Stream Order in South Carolina Coastal Plain Streams. *Transactions of the American Fisheries Society*, *123*(2), 150–161. [https://doi.org/10.1577/1548-8659\(1994\)123<0150:rbfasa>2.3.co;2](https://doi.org/10.1577/1548-8659(1994)123<0150:rbfasa>2.3.co;2)
- Perez, D. (2010a). *Cobitis paludica*. Wikimedia Commons, the free media repository. [https://commons.wikimedia.org/wiki/File:Cobitis\\_paludica\\_03\\_by-dpc.jpg](https://commons.wikimedia.org/wiki/File:Cobitis_paludica_03_by-dpc.jpg)
- Perez, D. (2010b). *Gobio lozanoi*. Wikimedia Commons, the free media repository. [https://commons.wikimedia.org/wiki/File:Gobio\\_lozanoi\\_03\\_by-dpc.jpg](https://commons.wikimedia.org/wiki/File:Gobio_lozanoi_03_by-dpc.jpg)
- Piccolo, J. J. (2011). Challenges in the conservation, rehabilitation and recovery of native stream salmonid populations: beyond the 2010 Luarca symposium. *Ecology of Freshwater Fish*, *20*(3), 346–351. <https://doi.org/10.1111/j.1600-0633.2010.00463.x>
- Pinho, C., & de Almeida, N. F. (2008). Biogeografia e Evolução Recente dos Anfíbios e Répteis de Portugal. In Armando Loureiro, N. F. de Almeida, M. A. Carretero, & O. S. Paulo (Eds.), *Atlas dos Anfíbios e Répteis de Portugal* (pp. 56–69). Instituto de Conservação da Natureza e das Florestas.
- Platts, W. S. (1979). Relationships among Stream Order, Fish Populations, and Aquatic Geomorphology in an Idaho River Drainage. *Fisheries*, *4*(2), 5–9. [https://doi.org/10.1577/1548-8446\(1979\)004<0005:rasofp>2.0.co;2](https://doi.org/10.1577/1548-8446(1979)004<0005:rasofp>2.0.co;2)
- Pough, F. H., Andrews, R. M., Cadle, J. E., Crump, M. L., Savitzky, A. H., & Wells, K. D. (1998). *Herpetology*. Prentice-Hall, Inc.
- Pyke, G. H. (2008). Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia* species. In *Annual Review of Ecology, Evolution, and Systematics* (Vol. 39, pp. 171–191). <https://doi.org/10.1146/annurev.ecolsys.39.110707.173451>
- Recuero, E., Canestrelli, D., Vörös, J., Szabó, K., Poyarkov, N. A., Arntzen, J. W., Crnobrnja-



- Isailovic, J., Kidov, A. A., Cogălniceanu, D., Caputo, F. P., Nascetti, G., & Martínez-Solano, I. (2012). Multilocus species tree analyses resolve the radiation of the widespread *Bufo bufo* species group (Anura, Bufonidae). *Molecular Phylogenetics and Evolution*, 62(1), 71–86. <https://doi.org/10.1016/j.ympev.2011.09.008>
- Reino, L., Ferreira, M., Martínez-Solano, Í., Segurado, P., Xu, C., & Márcia Barbosa, A. (2017). Favourable areas for co-occurrence of parapatric species: niche conservatism and niche divergence in Iberian tree frogs and midwife toads. *Journal of Biogeography*, 44(1), 88–98. <https://doi.org/10.1111/jbi.12850>
- Ribeiro, F., & Leunda, P. M. (2012). Non-native fish impacts on Mediterranean freshwater ecosystems: current knowledge and research needs. *Fisheries Management and Ecology*, 19(2), 142–156. <https://doi.org/10.1111/j.1365-2400.2011.00842.x>
- Robalo, J. I., Almada, V. C., Levy, A., & Doadrio, I. (2007). Re-examination and phylogeny of the genus *Chondrostoma* based on mitochondrial and nuclear data and the definition of 5 new genera. *Molecular Phylogenetics and Evolution*, 42(2), 362–372. <https://doi.org/10.1016/j.ympev.2006.07.003>
- Salvador, A. (2015). *Rana patilarga* – *Rana iberica* (Boulenger, 1879). In *Enciclopedia virtual de los vertebrados españoles*. (9th ed., p. 12). CSIC - Museo Nacional de Ciencias Naturales (MNCN). <https://doi.org/10.1016/j.jtcvs.2014.02.071>
- Salvador Milla, A. (2017). *Barbo común* – *Luciobarbus bocagei* (Steindachner, 1864). In *Enciclopedia Virtual de Vertebrados Españoles* (2nd ed.). CSIC - Museo Nacional de Ciencias Naturales (MNCN). <https://doi.org/10.20350/digitalCSIC/8713>
- Sánchez-Carmona, R., Encina, L., Rodríguez-Ruíz, A., & Rodríguez-Sánchez, V. (2008). Age, growth and diet of the Iberian loach, *Cobitis paludica* in two different environments. *Folia Zoologica*, 57(4), 420–434.
- Sánchez-Hernández, J., Vieira-Lanero, R., Barca, S., Silva, S., Lago, L., Gómez, P., Cobo, M. D. C., & Cobo, F. (2018). An update on the distribution of *Cobitis paludica* (De Buen, 1930) in the NW Iberian Peninsula. *Limnetica*, 37(2), 181–185. <https://doi.org/10.23818/limn.37.15>
- Sánchez-Hernández, J., Vieira-Lanero, R., Servia, M. J., & Cobo, F. (2011). Feeding habits of four sympatric fish species in the Iberian Peninsula: keys to understanding coexistence using prey traits. *Hydrobiologia*, 667(1), 119–132. <https://doi.org/10.1007/s10750-011-0643-2>
- Sánchez-Montoya, M. M., Moleón, M., Sánchez-Zapata, J. A., & Escoriza, D. (2017). The Biota of Intermittent and Ephemeral Rivers: Amphibians, Reptiles, Birds, and Mammals. In *Intermittent Rivers and Ephemeral Streams* (pp. 299–322). Elsevier. <https://doi.org/10.1016/B978-0-12-803835-2.00011-5>
- Santos, R. M. B., Sanches Fernandes, L. F., Cortes, R. M. V., Varandas, S. G. P., Jesus, J. J. B., & Pacheco, F. A. L. (2017). Integrative assessment of river damming impacts on aquatic fauna in a Portuguese reservoir. *Science of the Total Environment*, 601–602, 1108–1118. <https://doi.org/10.1016/j.scitotenv.2017.05.255>
- Sillero, N., Campos, J., Bonardi, A., Corti, C., Creemers, R., Crochet, P. A., Isailović, J. C., Denoël, M., Ficetola, G. F., Gonçalves, J., Kuzmin, S., Lymberakis, P., De Pous, P., Rodríguez, A., Sindaco, R., Speybroeck, J., Toxopeus, B., Vieites, D. R., & Vences, M. (2014). Updated distribution and biogeography of amphibians and reptiles of Europe. *Amphibia Reptilia*, 35(1), 1–31. <https://doi.org/10.1163/15685381-00002935>
- Snodgrass, J. W., Komoroski, M. J., Bryan, A. L., & Burger, J. (2000). Relationships among

- Isolated Wetland Size, Hydroperiod, and Amphibian Species Richness: Implications for Wetland Regulations. *Conservation Biology*, 14(2), 414–419. <https://doi.org/10.1046/j.1523-1739.2000.99161.x>
- Sousa, R., Nogueira, J. G., Ferreira, A., Carvalho, F., Lopes-Lima, M., Varandas, S., & Teixeira, A. (2019). A tale of shells and claws: The signal crayfish as a threat to the pearl mussel *Margaritifera margaritifera* in Europe. *Science of the Total Environment*, 665, 329–337. <https://doi.org/10.1016/j.scitotenv.2019.02.094>
- Stebbing, P., Longshaw, M., & Scott, A. (2014). Review of methods for the management of non-indigenous crayfish, with particular reference to Great Britain. *Ethology Ecology and Evolution*, 26(2–3), 204–231. <https://doi.org/10.1080/03949370.2014.908326>
- Stuart, S. N., Chanson, J. S., Cox, N. A., Young, B. E., Rodrigues, A. S. L., Fischman, D. L., & Waller, R. W. (2004). Status and trends of amphibian declines and extinctions worldwide. *Science*, 306(5702), 1783–1786. <https://doi.org/10.1126/science.1103538>
- Tierno de Figueroa, J. M., López-Rodríguez, M. J., Fenoglio, S., Sánchez-Castillo, P., & Fochetti, R. (2013). Freshwater biodiversity in the rivers of the Mediterranean Basin. *Hydrobiologia*, 719(1), 137–186. <https://doi.org/10.1007/s10750-012-1281-z>
- Velo-Antón, G., & Buckley, D. (2015). Salamandra común – Salamandra salamandra. In *Enciclopedia Virtual de los Vertebrados Españoles*. Museo Nacional de Ciencias Naturales, Madrid. <http://www.vertebradosibericos.org/>
- Vinyoles, D., Robalo, J., Sostoa, A., Almodóvar, A., Elvira, B., Nicola, G., Fernández-Delgado, C., Santos, C., Doadrio, I., Sardá-Palamera, F., & Almada, V. (2007). Spread of the alien bleak *Alburnus alburnus* (Linnaeus, 1758) (Actinopterygii cyprinidae) in the Iberian Peninsula: The role of reservoirs. *Graellsia*, 63(1), 101–110. <https://doi.org/10.3989/graellsia.2007.v63.i1.84>
- Wake, D. B. (1991). Declining amphibian populations. *Science*, 253(5022), 860. <https://doi.org/10.1126/science.253.5022.860>
- Welch, C., & Pintor, L. (2014). Bioturbation by the Invasive Rusty Crayfish (*Orconectes rusticus*) Affects Turbidity and Nutrients: Implications for Harmful Algal Blooms. The Ohio State University. <https://kb.osu.edu/handle/1811/60394>
- Wissink, J. (2018). *Lepomis gibbosus*. Wikimedia Commons, the free media repository. [https://commons.wikimedia.org/wiki/File:Zonnebaars\\_\(Lepomis\\_gibbosus\).png](https://commons.wikimedia.org/wiki/File:Zonnebaars_(Lepomis_gibbosus).png)
- Yamamoto, Y. (2010). Contribution of bioturbation by the red swamp crayfish *Procambarus clarkii* to the recruitment of bloom-forming cyanobacteria from sediment. *Journal of Limnology*, 69(1), 102. <https://doi.org/10.4081/jlimnol.2010.102>

## Chapter 3 – Modelling stream fish distributions using interval-censored detection times





## Chapter 3 – Modelling stream fish distributions using interval-censored detection times

**Mário Ferreira**, Ana Filipa Filipe, David C. Bardos, Maria Filomena Magalhães and Pedro Beja

This chapter was published in the journal *Ecology and Evolution* and should be cite as:

**Ferreira, M.**, Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>

Keywords:

Distribution modelling; hierarchical Bayesian models; imperfect detection; occupancy-detection modelling; stream fish; survival analysis; time to first detection.

### Abstract

Controlling for imperfect detection is important for developing species distribution models (SDMs). Occupancy-detection models based on the time needed to detect a species can be used to address this problem, but this is hindered when times to detection are not known precisely. Here we extend the time to detection model to deal with detections recorded in time intervals, and illustrate the method using a case study on stream fish distribution modelling.

We collected electrofishing samples of six fish species across a Mediterranean watershed in Northeast Portugal. Based on a Bayesian hierarchical framework, we modelled the probability of water presence in stream channels, and the probability of species occupancy conditional on water presence, in relation to environmental and spatial variables. We also modelled time to first detection conditional on occupancy in relation to local factors, using modified interval-censored exponential survival models. Posterior distributions of occupancy probabilities derived from the models were used to produce species distribution maps.

Simulations indicated that the modified time to detection model provided unbiased parameter estimates despite interval-censoring. There was a tendency for spatial variation in detection

rates to be primarily influenced by depth and, to a lesser extent, stream width. Species occupancies were consistently affected by stream order, elevation and annual precipitation. Bayesian P-values and AUCs indicated that all models had adequate fit and high discrimination ability, respectively. Mapping of predicted occupancy probabilities showed widespread distribution by most species, but uncertainty was generally higher in tributaries and upper reaches.

The interval-censored time to detection model provides a practical solution to model occupancy-detection when detections are recorded in time intervals. This modelling framework is useful for developing SDMs while controlling for variation in detection rates, as it uses simple data that can be readily collected by field ecologists.

### 3.1 Introduction

Species distribution models (SDMs) are widely used for research on biodiversity patterns and processes, and for informing conservation action and wildlife management (Guisan & Thuiller, 2005). Despite their value, SDMs may often be biased due to the use of datasets including false absences (Dorazio, 2012; Kéry, 2011; Lahoz-Monfort et al., 2014; Lobo et al., 2010) because failure to detect a species where it is present is a common source of error in biological surveys (Guillera-Arroita et al. 2014, and references therein). This problem may be solved using occupancy-detection modelling, whereby presence-absence and detectability given presence are jointly modelled in relation to covariates (MacKenzie et al., 2006), though only recently this approach has been considered in SDM development (G. Chen et al., 2013; Kéry et al., 2013; Lahoz-Monfort et al., 2014).

Occupancy-detection modelling is generally based on data from replicate discrete surveys conducted at, at least, a subset of sampling units (sites) (MacKenzie et al., 2006). Replicated surveys may be made by visiting sites more than once, but they may also be conducted at the same site on a single visit but partitioned by time, observer or method, or they can be conducted at different locations within a site (Guillera-Arroita, 2011; MacKenzie et al., 2006). In the removal design (MacKenzie et al. 2006), surveying is halted at a site once the species is detected and it was proposed that detection probabilities could be modelled as functions of covariates that vary across sites and also those ("such as local environmental conditions, time of day, or survey or experience") that vary across surveys. This removal design is therefore a very general approach to modelling first-detections where survey effort is treated as a series of discrete surveys. As such a crucial issue is exactly how the detection probabilities are modelled parametrically;

MacKenzie et al. (2006) suggested a logistic model using a combination of covariates that varied only between sites and those that vary between both sites and surveys.

A potentially more natural approach for developing SDMs while controlling for imperfect detection is to model the observation process as continuous process (e.g. a Poisson point process in time) and use the time needed to first detect a species, rather than a detection/non-detection history (Garrard et al., 2008; Guillera-Aroita et al., 2011). Time to first detection is an inverse function of detectability, and is known to be affected by factors such as species abundance, species traits, and sampling efficiency (Bornand et al., 2014; Garrard et al., 2013; McCarthy et al., 2013). The method is based on survival analysis (Kleinbaum & Klein, 2012), using distributions of times to first detection to parameterize a survival curve  $S(t)$  (i.e., the probability of a species remaining undetected before a given time  $t$ ), and to separate the probability of occupancy from the probability of detection given occupancy. The method has been mostly used in visual surveys of vascular plants (e.g. Garrard et al. 2008, 2013; McCarthy et al. 2013; Bornand et al. 2014), but it is likely useful for a wide range of taxonomic groups and sampling methods.

One problem of time to detection approaches is that the exact time when a species was first detected may be difficult to estimate precisely in some circumstances due, for instance, to sampling or recording constraints. In the case of bird point counts it is common practice to divide the count in time intervals, and recording species detections in each interval rather than at specific points in time (e.g. Alldredge et al. 2007). Likewise, in surveys of aquatic organisms using for instance nets, electrofishing or traps, species detections can often be recorded only within time intervals, and so the exact time to first detection is not known precisely (e.g. Beja & Alcazar 2003). In conventional survival analysis this problem has been described as interval-censoring, commonly resulting when periodic assessments (e.g. clinical or laboratory examinations) are used to assess if an event of interest has occurred (D.-G. Chen et al., 2012; Kleinbaum & Klein, 2012; Radke, 2003). In these circumstances, the event is known only to have occurred before a given assessment (right-censoring) and after the previous assessment (left-censoring), corresponding to the upper and lower bounds of a time interval. Common approximations for dealing with interval-censored data assume exact times (e.g., events occurring at the lower-bound, mid-point or upper bound of the interval); these approaches are arbitrary and can result in biased estimates of the survival curve and the effects of covariates (D.-G. Chen et al., 2012; Radke, 2003). We therefore avoid such approximations in applying interval-censored survival analysis to occupancy-detection modelling.

In this study, we developed a time-to-first detection framework in the context of SDMs, using a modified formulation of interval-censored survival analysis to deal with detections recorded in

time intervals (Kleinbaum & Klein, 2012). This provides a natural and coherent parameterization of detection probabilities for the removal design (MacKenzie et al. 2006) as a function of site covariates and arbitrary time intervals, by way of a detection rate that may be constant (exponential survival model) or vary with elapsed time (e.g. the 2-parameter Weibull survival model), and can be modelled as a function of site-covariates using a log-linear model.

The approach is illustrated using stream fish distributions, for which detection may greatly vary across sampling sites, and times to detection are difficult to estimate precisely due to sampling constraints (Penczak & Głowacki, 2008; Reynolds, 1996; Zalewsky & Cowx, 1990). In detail, we examined if the interval-censored time to detection approach allows building reliable models when imperfect detection is a potential drawback. We then used these models to extrapolate distributions of fish throughout the catchment streams. Finally, we discuss potential applications of the interval-censored time to detection model to different datasets that may often be collected by field ecologists.

## 3.2 Methods

### 3.2.1 Fish and environmental data

Descriptions of the study area, and of methodological details for species surveys and, the collection of environmental data are provided in Appendix 3.S1 Supplementary Methods in Supporting Information. We studied time to detection data for freshwater fish species sampled using electrofishing (Reynolds, 1996), in 50-m reaches (hereafter sites) distributed across the river Sabor catchment (NE Portugal). Sampling was conducted in the summer of 2012 at 89 sites, while no conditions for fish occurrence due to lack of surface water were recorded at another 95 sites. The study focused on the six most prevalent species (>20 sampling sites), including four natives (*Luciobarbus bocagei*, *Pseudochodrostoma duriense*, *Squalius alburnoides*, and *Squalius carolitertti*) and two exotics (*Gobio lozanoi* and *Lepomis gibbosus*). At each site we carried out an electrofishing session lasting for 15 to 25 minutes, with longer surveys used in wider streams to cover adequately the entire sector. The first detection of each species was recorded in 5-minute intervals due to practical constraints associated with electrofishing sampling.

Detection probabilities were modelled in relation to stream width and depth, because these variables strongly affect detectability by inducing variations in electrofishing efficiency (Reynolds, 1996) and in fish abundances (MacKenzie et al., 2006; McCarthy et al., 2013). Occupancy probabilities were modelled in relation to annual precipitation, elevation, and



Strahler's stream order, because these variables are known to strongly influence the distribution of stream fish in Mediterranean regions (M. T. Ferreira et al., 2007; Filipe et al., 2004; Magalhães et al., 2002), and they could be readily used to project the distribution models for the entire watershed.

### 3.2.2 Neighborhood effects

Modelling included neighbourhood effects to account for potential biases resulting from spatial autocorrelation of the data, i.e., lack of independence between the values of variables sampled at nearby locations (Legendre, 1993). We employed autologistic models (Augustin et al., 1996; Bardos et al., 2015; Besag, 1974; Gumpertz et al., 1997; Hoeting et al., 2000) for species occurrence and surface water presence;  $W_i = 1$  denotes water presence at site  $i$ , while  $Z_i = 1$  indicates true species presence. These models include an autocovariate that models the distance-weighted influence on response variables of surrounding response values, and a corresponding parameter allowing estimation of the strength of neighbourhood effects. The autocovariate was constructed as a weighted sum over neighborhood responses, not as a weighted mean, following the work of Bardos et al. (2015). We used an inverse-distance weighting, based on hydrological distance (stream length) in km, with a long-distance cut-off of 30km (above which the weighting is zero) and a short-distance cut-off of 5km, below which the weighting remains at 1/5, encoding the idea that the influence of particularly close sites does not increase without limit. The autocovariates at site  $i$  are therefore:

$$Wsp_i = \sum_{\substack{k \neq i \\ k \leq N_s \\ d_{ik} \leq 30}} \min\left(\frac{1}{d_{ik}}, \frac{1}{5}\right) W_k$$

$$Zsp_i = \sum_{\substack{k \neq i \\ k \leq N_s \\ d_{ik} \leq 30}} \min\left(\frac{1}{d_{ik}}, \frac{1}{5}\right) Z_k$$

Equation 3.1

Where  $N_s = 184$  is the total number of sampling sites (including 'dry' sites) and  $d_{i,k}$  is the hydrological distance in km between sites  $i$  and  $k$ . Different long-distance cut-offs were tested but the 30-km limit was retained because each site had at least two other sampling sites in its 30-km neighbourhood, and because it efficiently removed autocorrelation in model residuals as judged through Moran's I correlograms (Legendre & Legendre, 2012).

### 3.2.3 Species distribution models

We use WinBUGS to estimate autologistic models for water availability and true species presence-absence;  $\chi_i = \Pr(W_i = 1 \mid W_{-i})$  denotes the conditional probability of water presence at site  $i$ , given water presence-absence at all other sites (denoted  $W_{-i}$ ) and similarly  $\psi_i = \Pr(Z_i = 1 \mid Z_{-i})$  is the conditional probability of true occurrence at site  $i$ .  $Z_i$  depends on  $W_i$  and each depends on a common set  $X_j, j=1,2,\dots,n$ , of covariates, via autologistic models

$$\begin{aligned} \text{logit}(\chi_i) &= \alpha_0 + \alpha_1 X_{1i} + \dots + \alpha_n X_{ni} + \alpha_{\text{auto}} W_{sp_i} \\ \text{logit}(\psi_i) &= \beta_0 + \beta_1 X_{1i} + \dots + \beta_n X_{ni} + B(W_i - 1) + \beta_{\text{auto}} Z_{sp_i} \end{aligned}$$

Equation 3.2

Where  $\alpha_{\text{auto}}, \alpha_0, \alpha_1, \dots$  and  $\beta_{\text{auto}}, \beta_0, \beta_1, \dots$  are regression coefficients and  $B$  is a large positive constant (e.g.  $10^9$ ) that ensures the probability of presence  $\psi_i$  is effectively zero when water is absent ( $W_i = 0$ ).

We related true occupation to observed species presence and detection times via a model based on interval-censored exponential survival models (D.-G. Chen et al., 2012; Kleinbaum & Klein, 2012). Under interval-censoring (see Appendix 3.S2 Interval-censored time to detection model in Supporting Information), the likelihood of detecting a species at each sampling site, in the time interval  $(t_{1,i}, t_{2,i}]$ , during a survey of duration  $T_i$ , is given in terms of parametric detection-time distributions  $S(t) = S(t, \theta)$ :

$$\begin{aligned} l(\delta = 1 \mid t_{1,i}, t_{2,i}, \theta_i, \psi_i) &= \psi_i \left( S(t_{1,i}, \theta_i) - S(t_{2,i}, \theta_i) \right) \\ l(\delta = 0 \mid T_i, \theta_i, \psi_i) &= \psi_i S(T_i, \theta_i) + (1 - \psi_i) \end{aligned}$$

Equation 3.3

For  $i \in \{1, 2, \dots, N_S\}$ , where  $\delta_i$  is an indicator variable specifying whether the species was detected ( $\delta_i = 1$ ) or not ( $\delta_i = 0$ ) at site  $i$ ,  $\theta_i$  is a vector of detection-time distribution parameters at site  $i$ ,  $t_{1,i}$  and  $t_{2,i}$  are the lower and upper bounds of the time interval in which the species was detected at site  $i$ ,  $T_i$  is the total survey time at site  $i$ .

For the analysis here we use the exponential detection-time distribution  $S(t) = e^{-\lambda t}$ , where the detection rate  $\lambda$  is the sole parameter, so that the likelihood is then

$$l(\delta = 1 | t_{1,i}, t_{2,i}, \lambda_i, \psi_i) = \psi_i (e^{-\lambda_i t_{1,i}} - e^{-\lambda_i t_{2,i}})$$

$$l(\delta = 0 | T_i, \lambda_i, \psi_i) = \psi_i e^{-\lambda_i T_i} + (1 - \psi_i)$$

Equation 3.4

and we use a log-linear model for the detection rate  $\lambda_i$  at site  $i$

$$\log(\lambda_i) = \gamma_0 + \gamma_1 Y_{1i} + \dots + \gamma_m Y_{mi}$$

Equation 3.5

where  $Y_j$ ,  $j=1,2,\dots,m$ , comprise linear and quadratic terms for environmental covariates and  $\lambda_1, \lambda_2, \dots$  are regression coefficients.

### 3.2.4 Simulations for the detectability model

We conducted simulations to evaluate the performance of the interval-censored exponential model for detection data resulting from a study design comparable to ours, using an approach similar to Garrard (2009). For a set of  $K = 150$  sampling sites, we used a Bernoulli trial with a probability  $\psi$  to generate the “known” occupancy status at each site. Detection times given occupancy were generated using a random generator of exponential distribution times, with detection rate  $\lambda$ . We set a maximum time for sampling at each site of  $T_{\max} = 15$  minutes, with non-detections occurring when sites were vacant or when time to detection exceeded  $T_{\max}$ . Simulations were performed considering nine combinations of parameters, with occupancy set to  $\psi = 0.25, 0.5$  and  $0.75$ , and the detection rate set to  $\lambda = 0.20, 0.10$  and  $0.07$ . These detection rates correspond to mean detection times of 5, 10 and 15 minutes respectively. For each combination of parameters, we ran 1000 times.

### 3.2.5 Model building and evaluation

To avoid model instability and allow comparisons between parameters, all environmental covariates were standardized to zero mean and unit standard deviation. The detection component was fit to the full model, including second order polynomials of both depth and width, thereby allowing for non-linear changes in detection in relation to covariates. The occupancy and water presence components were also fit to the full model, including the three large-scale environmental variables and the neighbourhood effects. We fitted full models instead of seeking more parsimonious models because there is at present considerable uncertainty on the most

reliable methods to undertake selection in Bayesian models (e.g. Kéry 2010), the number of variables was low relative to sample sizes, and modelling was based on a small set of variables described in the literature to affect stream fish detection and occupancy. The effects of variables were judged from the 95% credible intervals, assuming that evidence for an effect is ambiguous when the credible interval of a parameter estimate includes zero (Kéry, 2010).

Overall model fit was assessed using posterior predictive checks based on standard Bayesian P-values (Gelman et al., 1996), measuring the discrepancy between observed and predicted detections at sampling sites. Extreme P-values (e.g., >0.95 or <0.05) are indicative of poor fit, whereas values near 0.5 indicate well-fitting models. Model discrimination ability was evaluated using an elaboration of the area under the receiver operating characteristic curve (AUC) in which posterior AUC distributions are calculated (Zipkin et al., 2012). Predicted probabilities of species presence cannot be directly compared to observed presences/absences, because false absences may occur (Garrard et al., 2013). In our study, AUC was based on comparisons between predicted detection probabilities and actual detections/non-detections at sites that were sampled (i.e., sites that were not dry), thereby providing an evaluation of the time to detection model fit. Probability of detecting a species at each site  $i$ , conditional on the sampling duration,  $T_i$ , was based on the second part of Equation 3.4, as follows:

$$\Pr(t_i < T_i | \psi_i, \lambda_i) = \psi_i (1 - e^{-\lambda_i T_i})$$

Equation 3.6

This unconditional probability of detection integrates both the probability of the species being present at the site, and the conditional probability of detection given presence. We performed a 5-fold cross-validation, in order to obtain a true predictive performance measure (Broms et al., 2016): (i) we randomly divided the data in 5 sets; (ii) withholding one set, we fitted the model to the remaining sets; (iii) computed AUC for the withheld set; and (iv) we repeated the process for every sub-set. We used all draws of the estimates of equation 5 to estimate posterior distributions and credible intervals of AUC values (ranging 0-1, where values > 0.5 indicate progressively better discrimination ability) using the R package ROCR (Sing et al., 2005).

The posterior probabilities of species detection were also used against actual detections/non-detections to estimate spatial autocorrelation in model residuals. For each model, we constructed a Moran's I correlogram using the mean values of the residuals posterior distributions, and evaluated the significance of Moran's I coefficients with Monte Carlo permutation tests using the R package APE (Paradis et al., 2004). To build the correlogram, pairwise distances were divided in classes such that a similar number of pairs was assigned to

each class, thereby assuring comparable power in tests of significance across all distance classes (Legendre & Legendre, 2012).

### 3.2.6 Species distribution mapping

We developed occupancy probability maps comprising (i) posterior autologistic occupancy probabilities  $\bar{\psi}_i$  for sampled sites  $i \leq N_s$  and (ii) extrapolated probabilities  $\bar{\psi}_i$  for a further 1861 unsampled sites (with labels  $i > N_s$ ) across the stream network of the Sabor catchment, for which neighbourhood effects are extrapolated by treating sampled sites as though they are neighbours of each unsampled site, i.e. by applying equation 1 to sites  $i > N_s$ . In the case where equation 1 reduces to a logistic model (i.e.  $\alpha_{\text{auto}} = \beta_{\text{auto}} = 0$ ), then for each extrapolation site  $i > N_s$ ,  $\bar{\psi}_i$  reduces to a posterior logistic occupancy probability. We used this extrapolation approach for neighbourhood effects because including the unsampled sites as missing data in the autologistic model was computationally impractical in WinBUGS.

For computational convenience, the stream network was segmented according to the following criteria: (i) each first order stream was one segment; (ii) one segment in higher order streams was the reach between two successive tributaries; and (iii) long reaches were divided so that all segments were  $< 1000\text{m}$ . Each segment was then assigned with the environmental characteristics of the corresponding centroid. At each segment we thus assumed that environmental conditions and neighbourhood effects were constant, and there was no variation in the probabilities of water presence and species occupancy at 50-m stream reaches. We used the mean estimated probabilities of species occupancy, and the standard deviation of the posterior distribution to produce the maps of predicted species distribution, and the uncertainty of model predictions. All spatial analysis and data manipulation were performed in ArcMap 10.0 (ESRI, 2011).

### 3.2.7 Model Fit

Models were fit in WinBUGS (Lunn et al., 2000), by calling WinBUGS through the package R2WinBUGS (Sturtz et al., 2005) in R (R Core Team, 2019), and handling the results back in R. Following a sensitivity analysis (Cressie et al., 2009), prior distributions of parameters were specified as normal distributions with zero mean and 10 units of standard deviation. We ran five chains of 100,000 iterations after a burn in of 50,000, and thinned the chains by 20 resulting in 12,500 simulations for each parameter. Convergence was assessed with the R-hat statistic, which examines the variance ratio of the MCMC algorithm within and between chains across

iterations. WinBUGS code is provided in Appendix 3.S3 Code used to fit the time to detection model using WinBUGS in Supporting Information.

### 3.3 Results

The simulation results (Table 3.1) showed that at sample sizes similar to ours the interval-censored model performed well. The simulated parameters were always well within the estimated credible intervals, and they were generally very close to the median parameter estimates. However, the occupation probability tended to be overestimated for lower levels of occupancy especially for lower detection rates.

The occupation-detection models for the six species showed adequate convergence of parameter estimates as judged from the R-hat statistics. Bayesian P-values were far from zero and one, ranging from 0.43 (*L. gibbosus*) to 0.64 (*S. alburnoides*), and thus model fit was considered adequate. Median AUCs estimated through cross-validation ranged between 0.67 and 0.93 indicating that the discrimination ability between detection and non-detection sites was particularly high (AUC > 0.80) for all species but *L. gibbosus* (Table 3.2). Moran's I correlograms indicated that there was no significant autocorrelation in the residuals of species occupancy-detection models. There was evidence for depth influencing the detection probabilities of *L. bocagei*, *P. duriense*, and *S. carolitertii*, as the credible intervals of parameter estimates for the linear (except *S. carolitertii*) and quadratic terms did not overlap zero (Table 3.2). These results suggested a U-shaped relationship with the median time to first detection, with shorter detection times when the stream was neither too shallow nor too deep (Figure 3.1). In the case of width, the credible intervals did not overlap zero in the model developed for *S. alburnoides*, suggesting also a U-shaped relationship (Figure 3.1). The probability of the stream channel having surface water during the sampling visit was positively related to stream order, elevation and precipitation, but the latter two effects were ambiguous because the credible intervals overlapped zero (Table 3.2, Appendix 3.S4 Response curves to environmental variables in Supporting Information). The probability of occupancy in sites with surface water was positively related to stream order for *L. bocagei* and *S. carolitertii*; elevation had a positive effect on *S. alburnoides*, and a negative effect on *G. lozanoi* and *L. gibbosus*; and precipitation had a negative effect on *L. bocagei*, *S. alburnoides*, and *S. carolitertii* (Table 3.2, Appendix 3.S4 Response curves to environmental variables in Supporting Information). Evidence for positive neighbourhood effects was found for *S. carolitertii* (Table 3.2).

Maps of predicted distribution indicated that *L. bocagei*, *P. duriense*, and *S. carolitertti* were widespread, occupying most of the main river and its two largest tributaries (Figure 3.2). *S. alburnoides* was more restricted, occurring primarily in the upper reach of the Sabor and the two main tributaries. From the two exotic species, *G. lozanoi* occurred primarily in the downstream reaches of the main river and its largest tributary, whereas *L. gibbosus* was more widespread, though it was also absent from upstream reaches and smallest tributaries (Figure 3.2). Uncertainty in model predictions was low to moderate, and it was highest for *P. duriense*, *L. gibbosus*, and *S. alburnoides* (Appendix S5 Maps of prediction uncertainty in Supporting Information). In most cases, uncertainty in species occupancy probability tended to be higher in the tributaries and in upper river reaches, where it was affected by uncertainties in whether the watercourses were dry or not.

Table 3.1 - Performance of the interval-censored time to detection model in retrieving parameter from simulated data. The simulated data was generated using 9 combinations of parameters, including three levels each of occupancy probability ( $\Psi$ ) and detection rate ( $\lambda$ ). For each simulated condition we present the median and credible intervals (in brackets) of parameter estimates based on the medians from 1000 simulations.

Simulated Parameters		Estimated parameters	
$\Psi$	$\lambda$	$\Psi$	$\lambda$
0.25	0.20	0.26 (0.15 - 0.36)	0.19 (0.09 - 0.33)
	0.10	0.28 (0.17 - 0.60)	0.09 (0.02 - 0.20)
	0.07	0.35 (0.16 - 0.63)	0.04 (0.01 - 0.14)
0.5	0.20	0.49 (0.39 - 0.60)	0.20 (0.15 - 0.27)
	0.10	0.48 (0.34 - 0.77)	0.10 (0.05 - 0.19)
	0.07	0.50 (0.32 - 0.73)	0.07 (0.03 - 0.15)
0.75	0.20	0.74 (0.65 - 0.83)	0.20 (0.16 - 0.26)
	0.10	0.72 (0.60 - 0.86)	0.11 (0.07 - 0.15)
	0.07	0.69 (0.52 - 0.86)	0.07 (0.05 - 0.13)

Table 3.2 - Mean parameter estimates and the corresponding 95% credible intervals (in brackets) of the best-supported models used in the distribution mapping of six freshwater fish species. Values are shown for each level of the hierarchical model: water availability - probability of a site having water; occupancy – probability of species occupying a site; detection – detection rate of the species in sites where it is present. AUC is the area under the curve of the receiver operating characteristic. Highlighted in grey and in bold are parameters (except the intercept) with credible intervals excluding zero.

Parameters	<i>L. bocagei</i>	<i>P. duriense</i>	<i>S. alburnoides</i>	<i>S. carolitertii</i>	<i>G. lozanoi</i>	<i>L. gibbosus</i>
<b>Water availability</b>						
Intercept	-0.44 (-1.69 ; 0.8)	-0.44 (-1.68 ; 0.81)	-0.43 (-1.67 ; 0.82)	-0.44 (-1.69 ; 0.82)	-0.43 (-1.68 ; 0.84)	-0.42 (-1.68 ; 0.82)
Elevation	0.37 (-0.40 ; 1.14)	0.37 (-0.4 ; 1.16)	0.37 (-0.4 ; 1.16)	0.37 (-0.41 ; 1.14)	0.37 (-0.4 ; 1.14)	0.37 (-0.4 ; 1.14)
Stream Order	<b>2.55 (1.86 ; 3.35)</b>	<b>2.55 (1.86 ; 3.35)</b>	<b>2.55 (1.86 ; 3.35)</b>	<b>2.55 (1.86 ; 3.35)</b>	<b>2.55 (1.86 ; 3.35)</b>	<b>2.55 (1.86 ; 3.35)</b>
Precipitation	0.37 (-0.13 ; 0.90)	0.38 (-0.12 ; 0.89)	0.38 (-0.12 ; 0.89)	0.38 (-0.13 ; 0.89)	0.38 (-0.12 ; 0.91)	0.37 (-0.13 ; 0.89)
Neighbourhood	-0.11 (-1.36 ; 1.12)	-0.10 (-1.36 ; 1.13)	-0.11 (-1.39 ; 1.09)	-0.11 (-1.38 ; 1.12)	-0.11 (-1.37 ; 1.12)	-0.12 (-1.35 ; 1.11)
<b>Occupancy</b>						
Intercept	-3.34 (-6.33 ; -0.72)	-0.87 (-3.95 ; 4.91)	1.40 (-1.39 ; 5.72)	-2.77 (-4.46 ; -1.21)	-4.93 (-8.19 ; -2.25)	-1.71 (-3.99 ; 1.67)
Elevation	-0.39 (-2.59 ; 1.02)	0.49 (-4.76 ; 2.50)	<b>3.06 (-0.05 ; 5.83)</b>	1.55 (0.78 ; 2.44)	<b>-5.25 (-8.39 ; -2.56)</b>	<b>-0.96 (-5.33 ; 1.61)</b>
Stream Order	<b>3.79 (1.83 ; 6.93)</b>	2.18 (-0.88 ; 6.07)	-1.08 (-3.53 ; 0.93)	<b>1.34 (0.48 ; 2.35)</b>	1.72 (0.18 ; 3.70)	-0.96 (-3.06 ; 1.15)
Precipitation	<b>-1.89 (-4.02 ; -0.36)</b>	-1.84 (-7.09 ; 0.38)	<b>-4.02 (-7.46 ; -0.58)</b>	<b>-1.03 (-2.24 ; -0.13)</b>	-1.23 (-3.44 ; 0.87)	-1.11 (-3.5 ; 0.56)
Neighbourhood	-0.34 (-4.11 ; 3.65)	1.97 (-2.37 ; 7.43)	0.25 (-4.91 ; 4.97)	<b>3.06 (0.54 ; 5.30)</b>	-1.12 (-4.75 ; 2.42)	3.66 (-2.49 ; 7.46)
<b>Detection</b>						
Intercept	-1.15 (-2.03 ; -0.19)	-2.47 (-3.05 ; -1.14)	-3.06 (-3.81 ; -2.13)	-1.00 (-1.51 ; -0.54)	-1.54 (-2.05 ; -1.08)	-3.17 (-3.86 ; -2.18)
Width	-1.03 (-4.00 ; 1.82)	0.56 (-2.57 ; 2.27)	<b>4.01 (1.28 ; 6.37)</b>	-0.53 (-2.29 ; 1.11)	-0.13 (-2.05 ; 1.64)	0.55 (-2.06 ; 2.44)
Width <sup>2</sup>	2.06 (-0.79 ; 5.85)	-0.01 (-1.56 ; 2.60)	<b>-5.40 (-8.43 ; -2.11)</b>	0.36 (-1.03 ; 1.96)	0.31 (-1.22 ; 2.16)	0.45 (-1.27 ; 3.05)
Depth	<b>2.84 (0.18 ; 5.25)</b>	<b>1.80 (0.46 ; 3.25)</b>	0.46 (-1.63 ; 2.16)	1.42 (-0.19 ; 3.12)	0.41 (-1.81 ; 2.64)	0.15 (-2.21 ; 2.34)
Depth <sup>2</sup>	<b>-3.01 (-5.33 ; -0.42)</b>	<b>-1.59 (-2.82 ; -0.45)</b>	-0.30 (-1.88 ; 2.18)	<b>-1.49 (-3.02 ; -0.01)</b>	-0.57 (-2.93 ; 1.92)	0.32 (-1.94 ; 3.4)
AUC	0.92 (0.80 - 1.00)	0.83 (0.53 - 0.96)	0.83 (0.63 - 0.94)	0.83 (0.61 - 1.00)	0.93 (0.68 - 1.00)	0.67 (0.14 - 0.94)
Bayesian P-value	0.54	0.58	0.64	0.55	0.5	0.43



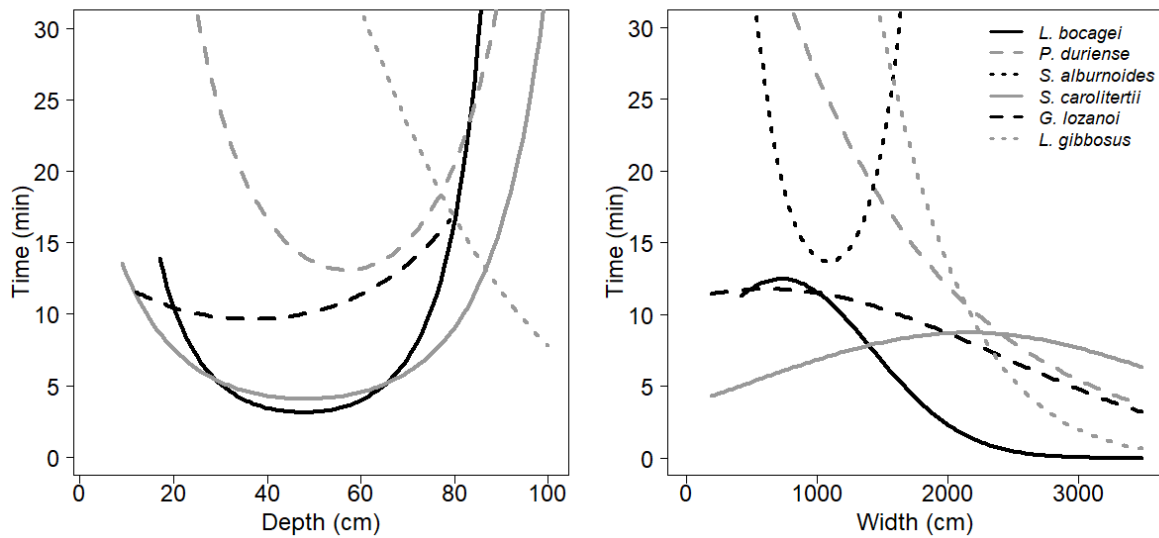


Figure 3.1 - Variation in median times to first detection of each species with 0.9 success probability if species is present, as a function of stream depth and width. Curves were derived from the detection models in Table 2, by varying the values of one variable conditioning on the mean values of other covariates in the model.

### 3.4 Discussion

Our work expanded the time to detection model (Garrard et al., 2008, 2013) to deal with detections collected in time intervals (interval-censoring), and illustrated its value for modelling species distribution using stream fish as a case study. The environmental correlates of occupancy identified for each species were in line with previous research on Mediterranean stream fish (e.g. Magalhães et al. 2002; Filipe et al. 2004; Ferreira et al. 2007), suggesting that models successfully incorporated key factors influencing species distributions. For most species we found significant spatial variation in detectability, supporting the importance to control for imperfect detection in distribution modelling studies (Guillera-Arroita et al., 2014; Lahoz-Monfort et al., 2014). Overall, our approach should provide a useful addition to the toolbox of field ecologists modelling species distributions while controlling for imperfect detection (G. Chen et al., 2013; Lahoz-Monfort et al., 2014).

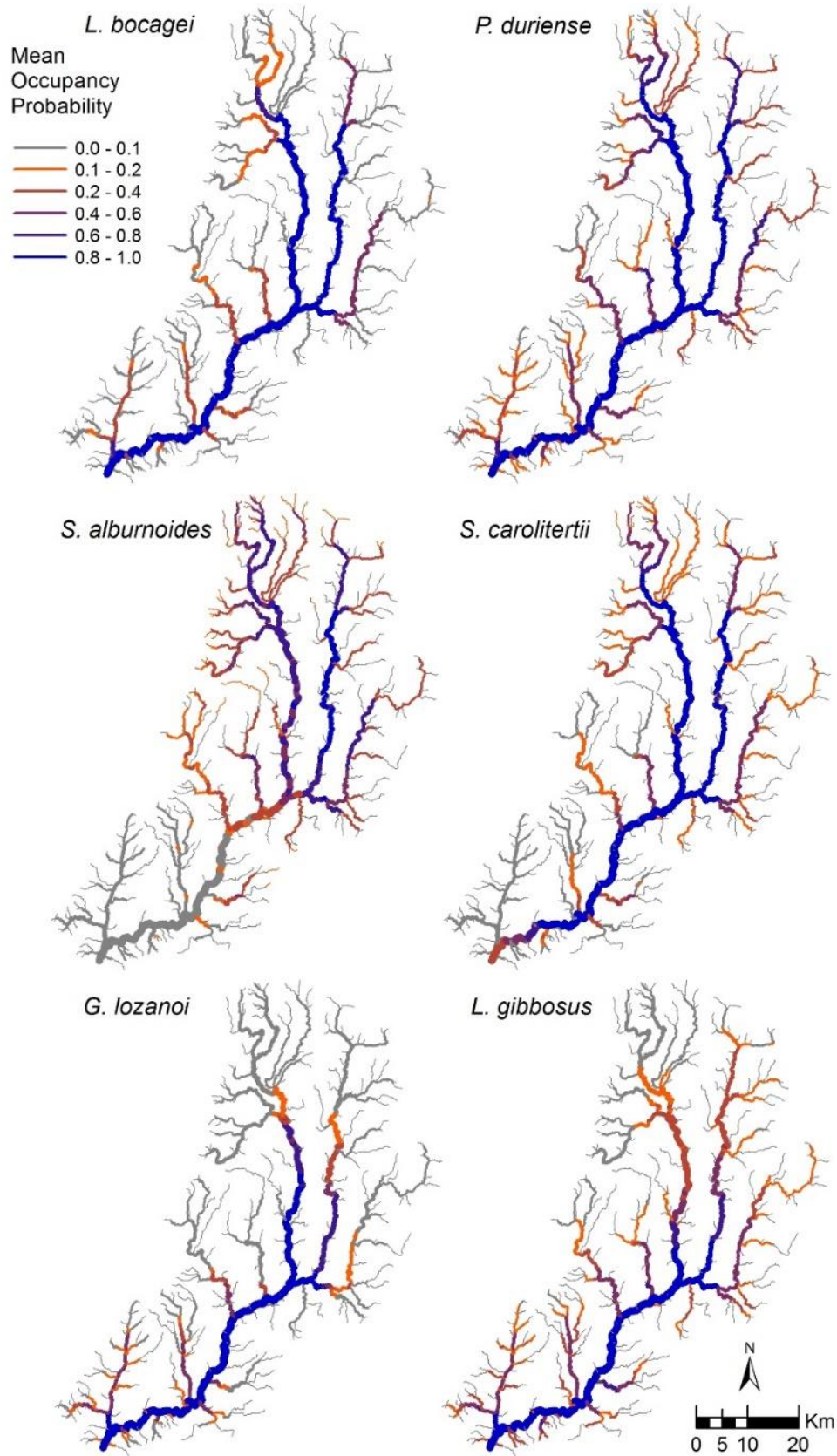


Figure 3.2 - Predicted occupancy probabilities of six fish species across the river Sabor catchment, combining the probabilities of surface water being present in the watercourse, and the conditional probabilities of occupancy given water presence. Line width is proportional to stream order.

Our study was based on the exponential model, which has been used in time to detection studies (Garrard et al., 2008, 2013), and it was considered a convenient choice due to its simplicity and its wide applicability (Kleinbaum & Klein, 2012). The exponential is the simplest of the parametric survival models (Kleinbaum & Klein, 2012), where times to detection are described by only one parameter and detections are assumed to occur at a constant rate (Garrard et al., 2008). Due to its memoryless property (K. E. Murphy et al., 2002), the non-detections on previous intervals do not provide information for a starting sampling interval, and thus it cannot model increases or decreases in detectability during each survey. This limitation may be overcome using other parametric survival models, but exploring these possibilities were beyond the scope of our study. In contrast to previous time to detection studies (Garrard et al., 2008, 2013), our study was based on detections recorded in 5-minute time intervals rather than continuously. This was unavoidable, because during electrofishing it is nearly impossible to keep a continuous track of each species captured, due to logistic constraints and difficulties in species identification.

Therefore, we have used a modification of the time to detection approach based on interval-censored survival analysis (D.-G. Chen et al., 2012; Kleinbaum & Klein, 2012), because common approximations assuming for instance events occurring at the lower-bound, mid-point or upper bound of the interval may result in biased estimates of the survival curve and the effects of covariates (Radke, 2003). Simulations showed that our approach provides unbiased estimates of detection rates and occupancy probabilities, suggesting that the method perform well in retrieving simulated values under conditions similar to our sampling design. It should be noted, however, that the occupancy probability tended to be overestimated for rare species (values of 0.25 in prevalence).

In four out of six species, we found that variation in detectability across sites was influenced by stream depth, stream width, or both, and that responses to these variables varied across species. These effects may reflect variation in electrofishing efficiency, which is generally expected to be lower when water is too deep or too shallow, and when rivers are very wide (e.g., Zalewsky & Cowx 1990; Penczak & Głowacki 2008). Electrofishing efficiency is known to be affected by factors such fish size, shape and behaviour (e.g., benthic versus pelagic) (e.g., Zalewsky & Cowx 1990; Penczak & Głowacki 2008), which may explain to at least some extent the differences observed across species. It is also possible that effects of width and depth were mediated by their strong influence on Mediterranean stream fish abundances (e.g., Ferreira et al. 2007), which in turn may have major effects on species detection probabilities (MacKenzie et

al., 2006; McCarthy et al., 2013). Different species reach the highest abundances in stream sectors of different width and depth (e.g., Ferreira et al. 2007), which may also contribute to explain changes in detection across species. Whatever the reasons, the results obtained provide empirical support to the view that accounting for imperfect detection is important when undertaking species distribution modelling (Lahoz-Monfort et al., 2014). This may be particularly relevant when focusing on aquatic species such as fish and amphibians, because organisms living underwater are notoriously difficult to sample and may be highly affected by imperfect detection (Głowacki, 2011), thereby calling for the use of modelling techniques controlling for variation in detectability (Comte & Grenouillet, 2013; M. Ferreira & Beja, 2013).

Modelling results revealed relationships between occupancy probabilities and environmental variables that are in line with the results from other studies carried out in Mediterranean streams, highlighting in particular the strong effect of stream order on occupancy (M. T. Ferreira et al., 2007; Filipe et al., 2004; Magalhães et al., 2002). For instance, we found that occupancy by *L. bocagei* and *P. duriensis* strongly increased with stream order, which is in line with observations elsewhere showing that barbel and straight-mouse nase tend to be more prevalent in higher order streams. Overall, results suggest that time to detection modelling was successful in identifying key factors affecting fish distribution, while controlling for variation in detectability. It is noteworthy, however, that this component of the hierarchical model accounted only for the probability of occupancy when there is water in the watercourse, because part of the streams were dry and thus unavailable for occupation by the target species. This was dealt with by modelling the probability of water presence in relation to environmental variables as an additional component of the hierarchical model, using binary draws from this probability to simulate surface water availability, then predicting the probability, given water availability, of fish occupancy of any 50-m reach of the stream network. Results indicated that the probability of water presence was mainly related to stream order, with headwater streams of order one and two tending to be dry and thus without conditions for fish, while streams and rivers of order three and above had a high probability of having water. This pattern is common in Mediterranean streams and elsewhere, where headwaters dry and as the stream channel increases in size downstream, surface water remains in pools or in surface flowing (Lake, 2003; Robson et al., 2013). We thus suggest that both the presence of water and the detection of species given water presence should be routinely considered when modelling the distribution of aquatic organisms along stream networks and in other waterbodies (e.g., pond breeding amphibians; Ferreira & Beja 2013),

providing a more realistic account of two potentially distinct processes affecting occupancy.

Evaluation of model discrimination ability for occupancy-detection models is difficult, because true absences are unknown, and so predicted probabilities of species occupancy cannot be directly compared with observed presences/absences (Garrard et al., 2013). To circumvent this problem, Garrard et al. (2013) evaluated occupancy-detection models by comparing the observed and predicted proportion of sites where each species was detected. Here we expanded this approach, using a variant of the AUC method described by Zipkin et al. (2012) to compare predicted detection probabilities with observed detections/non-detections, which avoided any assumptions about the characteristics of non-detections. AUC is a standard method for evaluating species distribution models (e.g., Kharouba et al. 2013) that provides a more complete characterization of model discrimination ability than the simple comparison of the observed and predicted proportion of species detections. In contrast to Zipkin et al. (2012) we used AUC to estimate the discrimination ability between detections and non-detections, and not between presences and absences.

The approach described here may find wide applicability where time to detection approaches are sought to control for imperfect detection in occupancy studies (e.g. Garrard et al. 2008, 2013), but where a species detection can only be determined to lie in an interval obtained from a sequence of sampling intervals. This may be generally the case in electrofishing studies such as ours, but the problem may also occur over a wide range of circumstances. For instance, sampling of aquatic organisms in shallow waters often involve dip-netting during fixed time intervals (Beja & Alcazar, 2003). Also, during bird counts it is common to register detections in time intervals (Alldredge et al., 2007), because it is impractical to register the exact moment when each individual was seen or heard. Finally, in studies involving periodic checking of traps (e.g., drift nets, mist nets, live traps for small mammals) it is possible to know that a capture event occurred after the trap was set but before it was checked, but the exact moment of capture it is often unknown. In all these cases, time to detection modelling may benefit from a wealth of methods developed to deal with interval-censored data, which have been particularly well explored in the medical and veterinary sciences (e.g., Radke 2003; Chen et al. 2012). These methods allow extending the relatively simple case described in our study, by accommodating for instance variation in the duration of time intervals across sampling units, or by replacing the exponential by a more flexible model (e.g., Weibull) that can account for changes in detectability within each sampling occasion (e.g., Kleinbaum &

Klein 2012; Chen et al. 2012). Overall, the interval-censored time to detection model framework revealed as a promising approach for developing SDMs, that could accommodate variation in detection rates, and we expect this approach to be tested in other case studies where time of first detection is not known precisely.

## Acknowledgements

This project was funded by the Portuguese Science and Technology Foundation (FCT) through project LTER/BIA-BEC/0004/2009 and through a PhD grant to MF (SFRH/BD/95202/2013), and by the Electrical Company EDP. We thank the collaboration in field work by Pedro Silva, Rita Severino, Sara Ivone and Sérgio Henriques, and the critical reading of the manuscript by Hugo Rebelo and Tiago Marques. We also thank the contribution of M.A. Amaral Turkman for help in formulating the interval-censored model.

## 3.5 References

- Aldredge, M. W., Pollock, K. H., Simons, T. R., Collazo, J. a., & Shriener, S. A. (2007). Time-of-Detection Method for Estimating Abundance from Point-Count Surveys. *The Auk*, *124*(2), 653. [https://doi.org/10.1642/0004-8038\(2007\)124\[653:TMFEAF\]2.0.CO;2](https://doi.org/10.1642/0004-8038(2007)124[653:TMFEAF]2.0.CO;2)
- Augustin, N. H., Muggleston, M. A., & Buckland, S. T. (1996). An Autologistic Model for the Spatial Distribution of Wildlife. *The Journal of Applied Ecology*, *33*(2), 339. <https://doi.org/10.2307/2404755>
- Bardos, D. C., Guillera-Aroita, G., & Wintle, B. A. (2015). Valid auto-models for spatially autocorrelated occupancy and abundance data. *Methods in Ecology and Evolution*, *6*(10), 1137–1149. <https://doi.org/10.1111/2041-210X.12402>
- Beja, P., & Alcazar, R. (2003). Conservation of Mediterranean temporary ponds under agricultural intensification: an evaluation using amphibians. *Biological Conservation*, *114*(3), 317–326. [https://doi.org/10.1016/S0006-3207\(03\)00051-X](https://doi.org/10.1016/S0006-3207(03)00051-X)
- Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society - Series B (Methodological)*, *36*(2), 192–236. <http://www.jstor.org/stable/2984812>
- Bornand, C. N., Kéry, M., Bueche, L., & Fischer, M. (2014). Hide-and-seek in vegetation: time-to-detection is an efficient design for estimating detectability and occurrence. *Methods in Ecology and Evolution*, *5*(5), 433–442. <https://doi.org/10.1111/2041-210X.12171>
- Broms, K. M., Hooten, M. B., & Fitzpatrick, R. M. (2016). Model selection and assessment for multi-species occupancy models. *Ecology*. <https://doi.org/10.1890/15-1471.1>
- Chen, D.-G., Sun, J., & Peace, K. (2012). *Interval-Censored Time-to-Event Data: Methods and*

- Applications* (D.-G. Chen, J. Sun, & K. Peace (eds.); Vol. 52). Chapman and Hall/CRC. <https://doi.org/10.1201/b12290>
- Chen, G., Kéry, M., Plattner, M., Ma, K., & Gardner, B. (2013). Imperfect detection is the rule rather than the exception in plant distribution studies. *Journal of Ecology*, *101*(1), 183–191. <https://doi.org/10.1111/1365-2745.12021>
- Comte, L., & Grenouillet, G. (2013). Species distribution modelling and imperfect detection: comparing occupancy versus consensus methods. *Diversity and Distributions*, *19*(8), 996–1007. <https://doi.org/10.1111/ddi.12078>
- Cressie, N. A. C., Calder, C. a, Clark, J. S., Ver Hoef, J. M., & Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications: A Publication of the Ecological Society of America*, *19*(3), 553–570. <http://www.ncbi.nlm.nih.gov/pubmed/19425416>
- Dorazio, R. M. (2012). Predicting the Geographic Distribution of a Species from Presence-Only Data Subject to Detection Errors. *Biometrics*, *68*(4), 1303–1312. <https://doi.org/10.1111/j.1541-0420.2012.01779.x>
- ESRI. (2011). *ArcMap* (10.2). Environmental Systems, Research Institute.
- Ferreira, M., & Beja, P. (2013). Mediterranean amphibians and the loss of temporary ponds: Are there alternative breeding habitats? *Biological Conservation*, *165*(null), 179–186. <https://doi.org/10.1016/j.biocon.2013.05.029>
- Ferreira, M. T., Sousa, L., Santos, J. M., Reino, L., Oliveira, J., Almeida, P. R., & Cortes, R. V. (2007). Regional and local environmental correlates of native Iberian fish fauna. *Ecology of Freshwater Fish*, *16*(4), 504–514. <https://doi.org/10.1111/j.1600-0633.2007.00241.x>
- Filipe, A. F., Marques, T. A., Tiago, P., Ribeiro, F., Da Costa, L. M., Cowx, I. G., & Collares-Pereira, M. J. (2004). Selection of Priority Areas for Fish Conservation in Guadiana River Basin, Iberian Peninsula. *Conservation Biology*, *18*(1), 189–200. <https://doi.org/10.1111/j.1523-1739.2004.00620.x>
- Garrard, G. E. (2009). *Dealing with imperfect detectability in biological surveys for native grassland management* (Issue November). School of Global Studies, Social Science and Planning RMIT University.
- Garrard, G. E., Bekessy, S. A., McCarthy, M. A., & Wintle, B. A. (2008). When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology*, *33*(8), 986–998. <https://doi.org/10.1111/j.1442-9993.2008.01869.x>
- Garrard, G. E., Mccarthy, M. a., Williams, N. S. G., Bekessy, S. a., & Wintle, B. a. (2013). A general model of detectability using species traits. *Methods in Ecology and Evolution*, *4*(1), 45–52. <https://doi.org/10.1111/j.2041-210x.2012.00257.x>
- Gelman, A., Meng, X., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *6*, 733–807. <http://www.cs.princeton.edu/courses/archive/fall09/cos597A/papers/GelmanMengStern1996.pdf>
- Głowacki, Ł. B. (2011). Accuracy of species richness estimators applied to fish in small and large temperate lowland rivers. *Biodiversity and Conservation*, *20*(6), 1365–1384. <https://doi.org/10.1007/s10531-011-0032-1>
- Guillera-Arroita, G. (2011). Impact of sampling with replacement in occupancy studies with spatial

- replication. *Methods in Ecology and Evolution*, 2(4), 401–406. <https://doi.org/10.1111/j.2041-210X.2011.00089.x>
- Guillera-Arroita, G., Lahoz-Monfort, J. J., MacKenzie, D. I., Wintle, B. A., & McCarthy, M. A. (2014). Ignoring Imperfect Detection in Biological Surveys Is Dangerous: A Response to ‘Fitting and Interpreting Occupancy Models’.’ *PLoS ONE*, 9(7), e99571. <https://doi.org/10.1371/journal.pone.0099571>
- Guillera-Arroita, G., Morgan, B. J. T., Ridout, M. S., & Linkie, M. (2011). Species Occupancy Modeling for Detection Data Collected Along a Transect. *Journal of Agricultural, Biological, and Environmental Statistics*, 16(3), 301–317. <https://doi.org/10.1007/s13253-010-0053-3>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Gumpertz, M. L., Graham, J. M., & Ristaino, J. B. (1997). Autologistic Model of Spatial Pattern of Phytophthora Epidemic in Bell Pepper: Effects of Soil Variables on Disease Presence. *Journal of Agricultural, Biological, and Environmental Statistics*, 2(2), 131. <https://doi.org/10.2307/1400400>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hoeting, J. A., Leecaster, M., & Bowden, D. (2000). An Improved Model for Spatially Correlated Binary Responses. *Journal of Agricultural, Biological, and Environmental Statistics*, 5(1), 102. <https://doi.org/10.2307/1400634>
- Kéry, M. (2010). Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses (1st ed.). Elsevier.
- Kéry, M. (2011). Towards the modelling of true species distributions. *Journal of Biogeography*, 38(4), 617–618. <https://doi.org/10.1111/j.1365-2699.2011.02487.x>
- Kéry, M., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2013). Analysing and mapping species range dynamics using occupancy models. *Journal of Biogeography*, 40(8), 1463–1474. <https://doi.org/10.1111/jbi.12087>
- Kharouba, H. M., McCune, J. L., Thuiller, W., & Huntley, B. (2013). Do ecological differences between taxonomic groups influence the relationship between species’ distributions and climate? A global meta-analysis using species distribution models. *Ecography*, 36(6), 657–664. <https://doi.org/10.1111/j.1600-0587.2012.07683.x>
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis*. Springer New York. <https://doi.org/10.1007/978-1-4419-6646-9>
- Lahoz-Monfort, J. J., Guillera-Arroita, G., & Wintle, B. a. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, 23(4), 504–515. <https://doi.org/10.1111/geb.12138>
- Lake, P. S. (2003). Ecological effects of perturbation by drought in owing waters. *Freshwater Biology*, 48, 1161–1172. <https://doi.org/10.1046/j.1365-2427.2003.01086.x>
- Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6), 1659–1673. <https://doi.org/10.2307/1939924>
- Legendre, P., & Legendre, L. F. J. (2012). *Numerical Ecology* (3rd ed.). Elsevier.



- Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, 33(1), 103–114. <https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- Lunn, D. J. J. D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10(4), 325–337. <http://www.springerlink.com/index/M832242U211K4437.pdf>
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier Academic Press.
- Magalhães, M. F., Beja, P., Canas, C., & Collares-Pereira, M. J. (2002). Functional heterogeneity of dry-season fish refugia across a Mediterranean catchment: the role of habitat and predation. *Freshwater Biology*, 47(10), 1919–1934. <https://doi.org/10.1046/j.1365-2427.2002.00941.x>
- Magalhães, M. F., Beja, P., Schlosser, I. J., & Collares-Pereira, M. J. (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, 52(8), 1494–1510. <https://doi.org/10.1111/j.1365-2427.2007.01781.x>
- McCarthy, M. a., Moore, J. L., Morris, W. K., Parris, K. M., Garrard, G. E., Vesk, P. a., Rumpff, L., Giljohann, K. M., Camac, J. S., Bau, S. S., Friend, T., Harrison, B., & Yue, B. (2013). The influence of abundance on detectability. *Oikos*, 122(5), 717–726. <https://doi.org/10.1111/j.1600-0706.2012.20781.x>
- Murphy, K. E., Carter, C. M., & Brown, S. O. (2002). The exponential distribution: the good, the bad and the ugly. A practical guide to its implementation. *Annual Reliability and Maintainability Symposium. 2002 Proceedings*, 550–555. <https://doi.org/10.1109/RAMS.2002.981701>
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20(2), 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Penczak, T., & Głowacki, Ł. B. (2008). Evaluation of electrofishing efficiency in a stream under natural and regulated conditions. *Aquatic Living Resources*, 21(3), 329–337. <https://doi.org/10.1051/alr:2008044>
- Podani, J., & Csontos, P. (2006). Quadrat size dependence, spatial autocorrelation and the classification of community data. *Community Ecology*, 7(1), 117–127. <https://doi.org/10.1556/ComEc>
- R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation Statistical Computing* (3.6.1). R Foundation for Statistical Computing. <https://doi.org/ISBN-3-900051-07-0>
- Radke, B. R. (2003). A demonstration of interval-censored survival analysis. *Preventive Veterinary Medicine*, 59(4), 241–256. [https://doi.org/10.1016/S0167-5877\(03\)00103-X](https://doi.org/10.1016/S0167-5877(03)00103-X)
- Reynolds, J. B. (1996). *Electrofishing. Fisheries Techniques* (B. R. Murphy & D. W. Willis (eds.); 2nd ed.). American Fisheries Society.
- Robson, B. J., Chester, E. T., Mitchell, B. D., & Matthews, T. G. (2013). Disturbance and the role of refuges in mediterranean climate streams. *Hydrobiologia*, 719(1), 77–91. <https://doi.org/10.1007/s10750-012-1371-y>
- Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROCr: visualizing classifier performance in R. *Bioinformatics (Oxford, England)*, 21(20), 3940–3941. <https://doi.org/10.1093/bioinformatics/bti623>

- Sturtz, S., Gelman, A., & Ligges, U. (2005). R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical Software*, 12(3), 1–16. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.161.3477&rep=rep1&type=pdf>
- Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, 165(6), 710–718. <https://doi.org/10.1093/aje/kwk052>
- Vogt, J. V., Soille, P., Jager, A. De, Rimavičiūtė, E., Mehl, W., Haastrup, P., Paracchini, M. M. L., Dusart, J., Bódis, K., Foisneau, S., & Bamps, C. (2007). Developing a pan-European Database of Drainage Networks and Catchment Boundaries from a 100 Metre DEM. *Proceedings 10th AGILE Int. Conference on Geographic Information Science, May*, 8–11. [http://ams.jrc.it/publications/pdfs/JVogt\\_etal\\_AGILE2007.pdf](http://ams.jrc.it/publications/pdfs/JVogt_etal_AGILE2007.pdf)
- Zalewsky, M., & Cowx, I. G. (1990). Factors affecting the efficiency of electric fishing. In *Fishing with electricity: applications in freshwater fisheries management* (pp. 89–111). Fishing News Books.
- Zipkin, E. F., Grant, E. H. C., & Fagan, W. F. (2012). Evaluating the predictive abilities of community occupancy models using AUC while accounting for imperfect detection. *Ecological Applications*, 22(7), 1962–1972. <https://doi.org/10.1890/11-1936.1>

## 3.6 Supplementary Materials

### Appendix 3.S1 Supplementary Methods

#### 3.S1.1 Study area

The study was conducted in NE Portugal, in the river Sabor catchment (Figure 3.S), which covers a wide range of environmental conditions in terms of elevation (100-1500 meters above sea level), total annual precipitation (443 - 1163 mm), and mean annual temperature (6.9 – 15.6°C). Climate is Mediterranean, with precipitation largely concentrated in October-March, while it is virtually absent in the hot summer months (June-August). Flow regime is highly seasonal, with most headwater streams drying out or being reduced to a series of disconnected pools in summer, though the main watercourse and some of the tributaries are permanent.

#### 3.S1.2 Fish sampling

Fish sampling was carried out at 50-m reaches of streams (hereafter referred to as sites) distributed across the watershed, during the summer of 2012 (June 13 to July 15, and August 28 to September 15). The reach length was selected based on previous studies showing that it is appropriate for describing spatial and temporal variations in Mediterranean stream fish assemblages (Magalhães et al., 2002, 2007). We visited 184 sites across the watershed, and sampled the 89 sites with flowing water or isolated pools. The remaining sites were totally dry, and so they could not sustain fish populations during the sampling period. The sites were selected in the field based on accessibility and representativeness, as long they were 4 to 10km apart from each other, and provided a thorough coverage of the watershed (Figure 3.S). Within each site, we sampled fish using a single anode electrofishing gear (350-750 V, 3–5A, DC), following standard procedures (Penczak & Głowacki, 2008; Zalewsky & Cowx, 1990). Electrofishing was always conducted by the same operator (M.F.), accompanied by a second operator to net fish displaced by electrofishing. Sampling was carried out during 15 to 25 minutes, with longer surveys used in wider streams to cover adequately the entire sector (Reynolds, 1996). During each sampling session, we electrofished progressively in the upstream direction, and transferred fish alive to containers distributed along the margins. Because most fish captured were small-sized or otherwise difficult to identify to species level without careful examination, identification and the estimate of times to first detection

were made at the end of the sampling session. To this end, fish captured in each interval of 5 minutes were transferred to a different set of marked containers, which allowed the recording of species detections in 5-minute intervals. After identification fish were returned alive to the stream.

Sampling yielded 12 fish species (Table 3.S.1), of which four natives (*Luciobarbus bocagei*, *Pseudochodrostoma duriense*, *Squalius alburnoides*, and *Squalius carolitertii*) and two exotics (*Gobio lozanoi* and *Lepomis gibbosus*) were used in occupation-detection modelling. The remaining six species occurred too rarely (1 to 13 sites), and so they were discarded to reduce potential problems associated with a low number of events per variable in occupancy and detection models (e.g., Vittinghoff & McCulloch, 2007).

Table 3.S.1 - List of fish species sampled in the Sabor catchment in the summer of 2012. The percentage of sites with detections is provided for each species (frequency of detection; n=89 sites). The percentage of sites with detections is provided for each species (frequency of detection; n=89 sites). Species used in distribution modelling (occurring in > 20 sites) are highlighted in bold type.

Species	Frequency of detection (%)
<b>Native</b>	
<b>Pseudochondostoma duriense</b>	56.2
<b>Squalius carolitertii</b>	53.9
<b>Luciobarbus bocagei</b>	52.8
<b>Squalius alburnoides</b>	33.7
Achondrostoma arcasii	13.5
Salmo trutta	11.2
Cobitis paludica	10.1
<b>Exotic</b>	
<b>Gobio lozanoi</b>	41.6
<b>Lepomis gibbosus</b>	31.5
Gambusia holbrookii	14.6
Alburnus	4.5
Carassius auratus	1.1

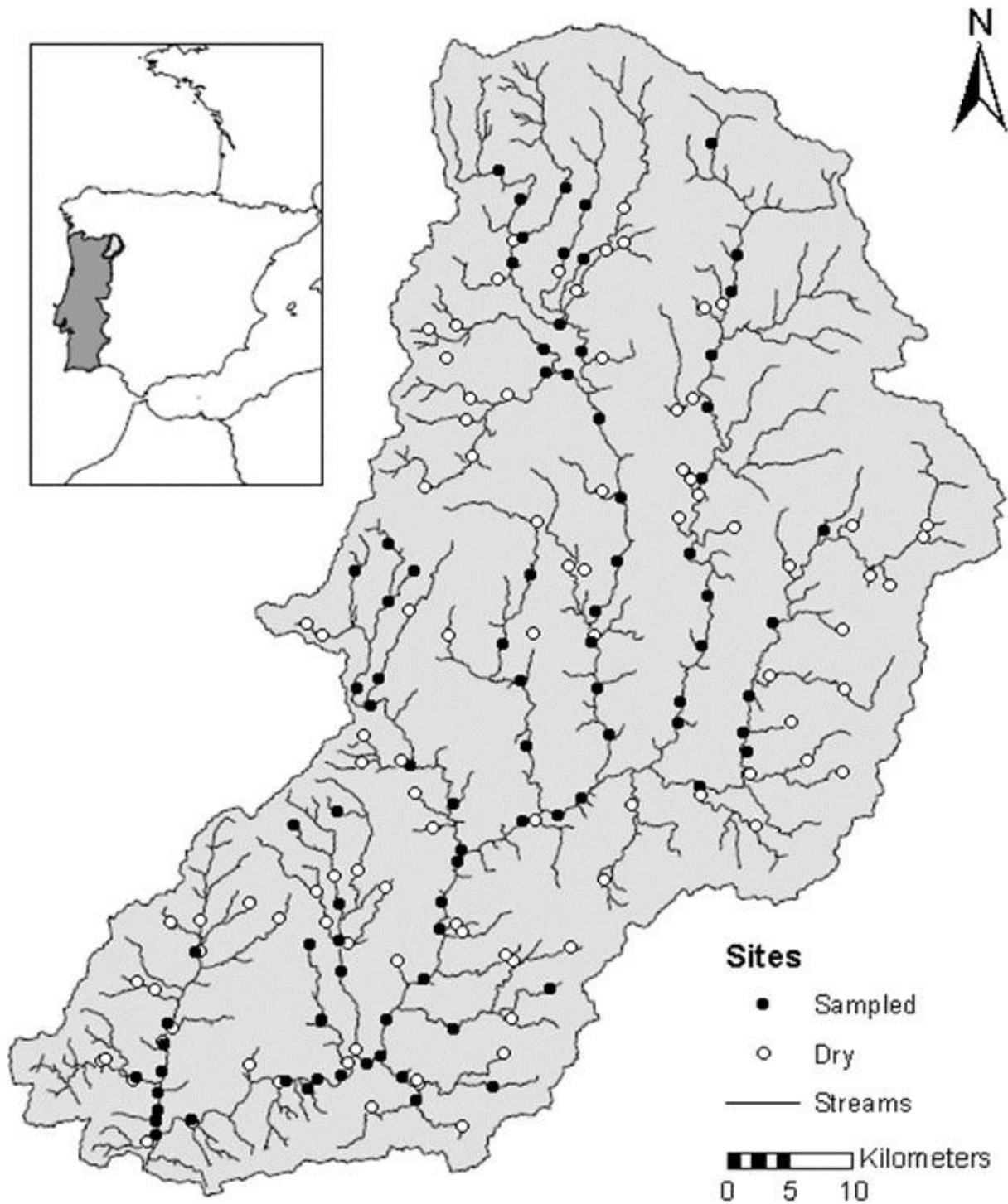


Figure 3.S1 - Map of the study area, showing the location of sites visited in the summer of 2012 (June 13 to July 15, and August 28 to September 15). Fish sampling was carried in 89 sites that had flowing water or isolated pools, while the remaining 95 sites were dry.

### 3.S1.3 Environmental covariates

Detection probabilities were modelled in relation to stream width and depth (Table 3.S.2), because electrofishing efficiency tends to be lower in larger and deeper watercourses (Reynolds, 1996). Also, these variables may strongly influence fish abundances, which in turn may positively affect detectability (MacKenzie et al., 2006; McCarthy et al., 2013). Mean reach width was estimated from measurements taken along four equally spaced transversal segments, whereas mean depth was estimated from three measurements taken in each of the four segments (Table 3.S.2).

Covariates for site occupancy included total annual precipitation, elevation, and Strahler's stream order, which have been widely reported to strongly influence the distribution of stream fish in Mediterranean regions (M. T. Ferreira et al., 2007; Filipe et al., 2004; Magalhães et al., 2002). The same covariates were used to model the probability of water presence at each site. Elevation at each sampling site was calculated from a 10-m resolution Digital Elevation Model (DEM) derived from 1: 25,000 topographic maps using ArcMap 10.0 (ESRI, 2011). Stream order was used to define stream size based on the hierarchy of tributaries, and it was extracted from the Catchment Characterization and Modelling database (CCM2), based on a 100-m resolution DEM (Vogt et al., 2007). Precipitation was extracted from WorldClim current climate predictors, which are based on the monthly mean interpolations from records collected over a 50-year period (1950-2000), with a 30 arc-seconds grid resolution (approximately 1km<sup>2</sup>; Hijmans et al., 2005). The resolution of this variable was converted to match the 10-m resolution of the DEM.

Table 3.S.2 - Description and summary statistics (mean  $\pm$  SD, range) of environmental variables used to model variation in detection rate and occupancy probability.

Variables	Description	Mean $\pm$ SD	Range
<b>Detection model</b>			
Stream width (m)	Mean of four width measurements taken at equally spaced transversal segments along the 50-m sampling reach.	9.7 $\pm$ 6.8	1.5-34.8
Stream depth (cm)	Mean of depth measurements taken at, 25%, 50%, and 75% of the length of each transversal segment.	34 $\pm$ 19	9-100
<b>Occupancy model</b>			
Elevation (m)	Altitude above sea level extracted from a 10-m resolution digital elevation model derived from 1: 25,000 topographic maps	469 $\pm$ 202	110-800
Precipitation (mm)	Average of total annual precipitation (1950-2000), extracted from WorldClim and downscaled to a 10-m resolution.	686 $\pm$ 82	568-894
Stream order	Strahler's stream order, extracted from the River and Catchment Database CCM2	2.6 $\pm$ 1.4	1-6

## Appendix 3.S2 Interval-censored time to detection model

The modelling procedure was based on the exponential time to detection model developed by (Garrard et al., 2008, 2013), using a modified formulation of interval-censored parametric survival models to deal with cases when detections are recorded in time intervals instead of continuously (D.-G. Chen et al., 2012; Kleinbaum & Klein, 2012). Under the general parametric survival model for the time  $T$  observed for a certain event of interest (e.g. the first detection of a species), the probability of observing an event after time  $t$  (i.e. time of the event  $T$  is greater than  $t$ ) equals the survival distribution for time  $t$ :

$$\Pr(T > t) = S(t)$$

Equation 3.S1

and the probability of observing an event before a time  $t$  equals the complementary probability of observing the event after time  $t$ :

$$\Pr(T \leq t) = 1 - S(t)$$

Equation 3.S2

Therefore, the probability of the event occurring in a given time interval defined by a lower bound ( $t_1$ ) and an upper bound ( $t_2$ ) is the probability of observing the event before the time  $t_2$  minus the probability of observing the event before time  $t_1$ :

$$\Pr(t_1 < T \leq t_2) = 1 - S(t_2) - (1 - S(t_1)) = S(t_1) - S(t_2)$$

Equation 3.S3

This very general formulation can be parameterized using one of several available distributions of survival times, including in the simplest case the exponential model, which is fully described by a single parameter – the detection rate ( $\lambda$ ):

$$S(t) = e^{-\lambda t}$$

Equation 3.S4



In the context of species detection, the previous equations can be combined to estimate the likelihood of species detection in a given time interval. Considering that a species occupies a site  $i$ , and that first detections follows an exponential distribution with detection rate  $\lambda$ , the likelihood of observing a first-detection event (denoted  $\delta_i = 1$ ) during a survey interval defined by  $(t_{1,i}, t_{2,i}]$ , is

$$l(\delta_i = 1, t_{1,i}, t_{2,i} | \lambda) = e^{-\lambda t_{1,i}} - e^{-\lambda t_{2,i}}$$

Equation 3.S5

and the likelihood of not detecting ( $\delta_i = 0$ ) the species during a survey of duration  $T_i$  is

$$l(\delta_i = 0 | \lambda, T_i) = e^{-\lambda T_i}$$

Equation 3.S6

These equations assume that the event will occur, even if it is not detected during the survey time. However, in contrast to survival analysis, it is uncertain whether a species is present or absent, and so it may remain unrecorded either because it is absent or because it is present but remained undetected. This possibility is considered in time to detection models by including in equations S5 and S6 the probability that the species actually occupy the site. That is, the probability of detection in a given time interval under unknown occupancy is given by

$$l(\delta = 1 | t_{1,i}, t_{2,i}, \lambda, \psi) = \psi(e^{-\lambda t_{1,i}} - e^{-\lambda t_{2,i}})$$

$$l(\delta_i = 0 | T_i, \lambda, \psi) = \psi(e^{-\lambda T_i}) + (1 - \psi)$$

Equation 3.S7

where  $\lambda$  is the rate at which detection events occur,  $\psi$  is the probability of the species occupying a site,  $T_i$  is the survey time at site  $i$ . This formulation implies that the likelihood of not recording a species during a survey ( $\delta = 0$ ) is now a function of both imperfect detection ( $\psi \cdot e^{-\lambda T_i}$ ) and true absence ( $1 - \psi$ ) (Garrard et al., 2008, 2013). The model assumes that the species is available for detection during the entire sampling period, which is a reasonable assumption, considering that all fish occurring in a stream reach are exposed to electrofishing sampling (Reynolds, 1996). However, a more general treatment of imperfect detection would have to describe both the probability of a species being available for sampling and the probability of detection given availability (e.g., Kéry & Schmidt, 2008), but this was beyond the scope of this study.

In the above we have specified the observation process as the exponential model from Equation 3.S4 onward; returning now to full generality we give the likelihood expressions for any given parametric detection-time distribution  $S(t) = S(t, \theta)$  with vector of parameters  $\theta$ :

$$l(\delta = 1 | t_{1,i}, t_{2,i}, \theta, \psi) = \psi \cdot (S(t_{1,i}, \theta) - S(t_{2,i}, \theta))$$

$$l(\delta_i = 0 | T_i, \theta, \psi) = \psi(S(T_i, \theta)) + (1 - \psi)$$

Equation 3.S8

Whereas the exponential model has the property that the detection probability for any time interval  $(t_1, t_2]$  depends only on the length  $t_1 - t_2$  of the interval (the memoryless property, Murphy et al., 2002), more general distributions such as the 2-parameter Weibull distribution allow detection probabilities for equal intervals to increase or decrease with later times. This allows, for example, for inclusion of changing detection probability over time due for instance to disturbances from survey efforts.

Due to this property, when using the exponential model, i.e. the simplest of parametric survival models (Kleinbaum & Klein, 2012), the interval-censored time to detection model with equal intervals is the mathematical equivalent of the occupation-detection models with removal design (MacKenzie et al., 2006) with equal detection probabilities for each of a series of discrete surveys. In this approach researchers record species detections at the end of intervals of length  $\Delta t$ , stopping after the 1st detection or once a predefined maximum number ( $K$ ) of intervals has elapsed. If time to first detection of species follows an exponential survival time model with detection rate  $\lambda$  then the detection probability in each interval is  $p = 1 - e^{-\lambda\Delta t}$ .

When the first detection at a site happens within the interval defined by times  $t_1$  and  $t_2$ , and calling that interval the  $k_i^{th}$  repeat visit we can developed the likelihood in Equation 3.S7 as follows:

$$l(\delta = 1 | t_{1,i}, t_{2,i}, \lambda, \psi) = \psi \cdot (e^{-\lambda t_1} - e^{-\lambda t_2}) = \psi \cdot [(e^{-\lambda})^{t_1} - (e^{-\lambda})^{t_2}]$$

$$= \psi \cdot [(1 - p)^{t_1/\Delta t} - (1 - p)^{t_2/\Delta t}]$$

Equation 3.S9

If we keep  $\Delta t$  constant then by definition of the  $k_i^{th}$  repeat visit, we know that  $t_2/\Delta t = k_i$  and  $t_1/\Delta t = k_i - 1$ , so we have:

$$\begin{aligned} l(\delta = 1 | k_i, p, \psi) &= \psi \cdot [(1 - p)^{k_i - 1} - (1 - p)^{k_i}] = \psi \cdot (1 - p)^{k_i - 1} \cdot (1 - 1 + p) \\ &= \psi \cdot (1 - p)^{k_i - 1} \cdot p \end{aligned}$$

Equation 3.S10

which is the expression for removal sampling detection data for a site where 1st detection happens at the  $k_i$ -th repeat visit, i.e. we get  $(k_i - 1)$  non-detections followed by the single detection (MacKenzie et al., 2006).

When there are no detections after a total survey time  $T_i$  (i.e.  $K$  repeat visits with  $K = T/\Delta t$ ):

$$l(\delta_i = 0 | T_i, \theta, \psi) = \psi(e^{-\lambda T_i}) + (1 - \psi) = \psi(1 - p)^K + (1 - \psi)$$

Equation 3.S11

The first-detection time models considered here generalize the above scheme to allow (i) uneven sampling times/intervals, (ii) variation of detection rates between sites and (iii) variation of detection rates as a function of time by employing non-exponential parametric survival models (e.g. Weibull model Kleinbaum & Klein, 2012).

## Appendix 3.S3 Code used to fit the time to detection model using WinBUGS

```
model {  
  
# priors  
  
a0 ~ dnorm(0, .1) I(-10, 10) # Intercept for Water availability  
  
b0 ~ dnorm(0, .1) I(-10, 10) # Intercept for occupation  
  
g0 ~ dnorm(0, .1) I(-10, 10) # Intercept for detection  
  
for (n in 1:Xocc) { # mean effects in occupation  
    a[n] ~ dnorm(0, .1) I(-10, 10) #Effects Water availability  
    b[n] ~ dnorm(0, .1) I(-10, 10) #Effects occupancy  
}  
  
for (m in 1:Xdet) {  
    g[m]~ dnorm(0, .1) I(-10, 10) #Effects Detection  
}  
  
a.sp ~ dnorm(0, .1) I(-10, 10) #Spatial Effect on Water availability  
  
b.sp ~ dnorm(0, .1) I(-10, 10) #Spatial Effect on occupation  
  
#Spatial Autologistic term computation  
  
for (i in 1:nsite) {  
    for(j in 1:nnb[i]) {
```

```
autoZ[i,j] <- Z[nblists[i,j]]
  }
Z.sp[i]<-inprod(autoW[i,1:nnb[i]],autoZ[i,1:nnb[i]])
}

#Model
for (i in 1:nsite) {

  #Water availability
  IW[i] <- a0 + inprod(a[ ], X1[i, ]) + a.sp * W.sp[i]
  pW[i] <- 1/(1+exp(-IW[i]))
  W[i] ~ dbern(pW[i])

  #Occupancy Model
  lpsi[i] <- b0 + inprod(b[ ], X1[i, ]) - (1 - W[i]) * pow(10, 9) + b.sp * Z.sp[i]
  psi[i] <- 1/(1+exp(-lpsi[i]))

  Z[i] ~ dbern(psi[i]) #True state occupation

  #Detection rate
  lambda[i] <- exp(g0 + inprod(g[], X2[i,]))

  #Survival function for Left bound
  S1[i] <- exp(-lambda[i] * y1[i])
```

```
#Survival function for Right bound

S2[i] <- exp(-lambda[i] * y2[i])

#when species is detected

pp[i] <- (S1[i] - S2[i]) * psi[i]

#when species is not detected

pn[i] <- S2[i] * psi[i] + (1 - psi[i])

#Select likelihood

p[i] <- (d[i]*pp[i] + (1-d[i])*pn[i]) #d[i] = 1 where detected, d[i]= 0 where not
detected

ones[i] ~ dbern(p[i]) #ones trick

# Probability of detecting an individual at site i (Evaluation purposes)

p1[i] <- psi[i] * (1 - exp(-lambda[i] * TT[i]))

Res[i] <- d[i] - p1[i] #Residuals

#Replicate observations

d_rep[i] ~ dbern(p1[i]) #Generate replicate observations

Res_rep[i] <- d_rep[i] - p1[i] #Replicate residuals

}

fit <- sum(Res[]) # Sum of residuals for actual data set

fit.new <- sum(Res_rep[]) # Sum of residuals for new data set
```

```
test <- step(fit.new - fit)           # Test whether new data set more extreme

bpvalue <- mean(test)                # Bayesian p-value

# Catchment area extrapolation

#Spatial Autologistic term computation

for (i in 1:nsite2) {
  for(j in 1:nnb2[i]) {
    autoZ2[i,j] <- Z[nblists2[i,j]]
  }
  Z2.sp[i]<-inprod(autoW2[i,1:nnb2[i]],autoZ2[i,1:nnb2[i]])
}

for (j in 1:nsite2) {
  #Water availability

  IW2[j] <- a0 + inprod(a[ ], X3[j, ]) + a.sp * W.sp2[j]

  pW2[j] <- 1/(1+exp(-IW2[j]))

  W2[j] ~ dbern(pW2[j])

  #Occupancy Model

  lpsi2[j] <- b0 + inprod(b[ ], X3[j, ]) - (1 - W2[j]) * pow(10, 9) + b.sp * Z.2sp[j]

  psi[j] <- 1/(1+exp(-lpsi2[j]))
}

}# end of model
```

Appendix 3.S4 Response curves to environmental variables

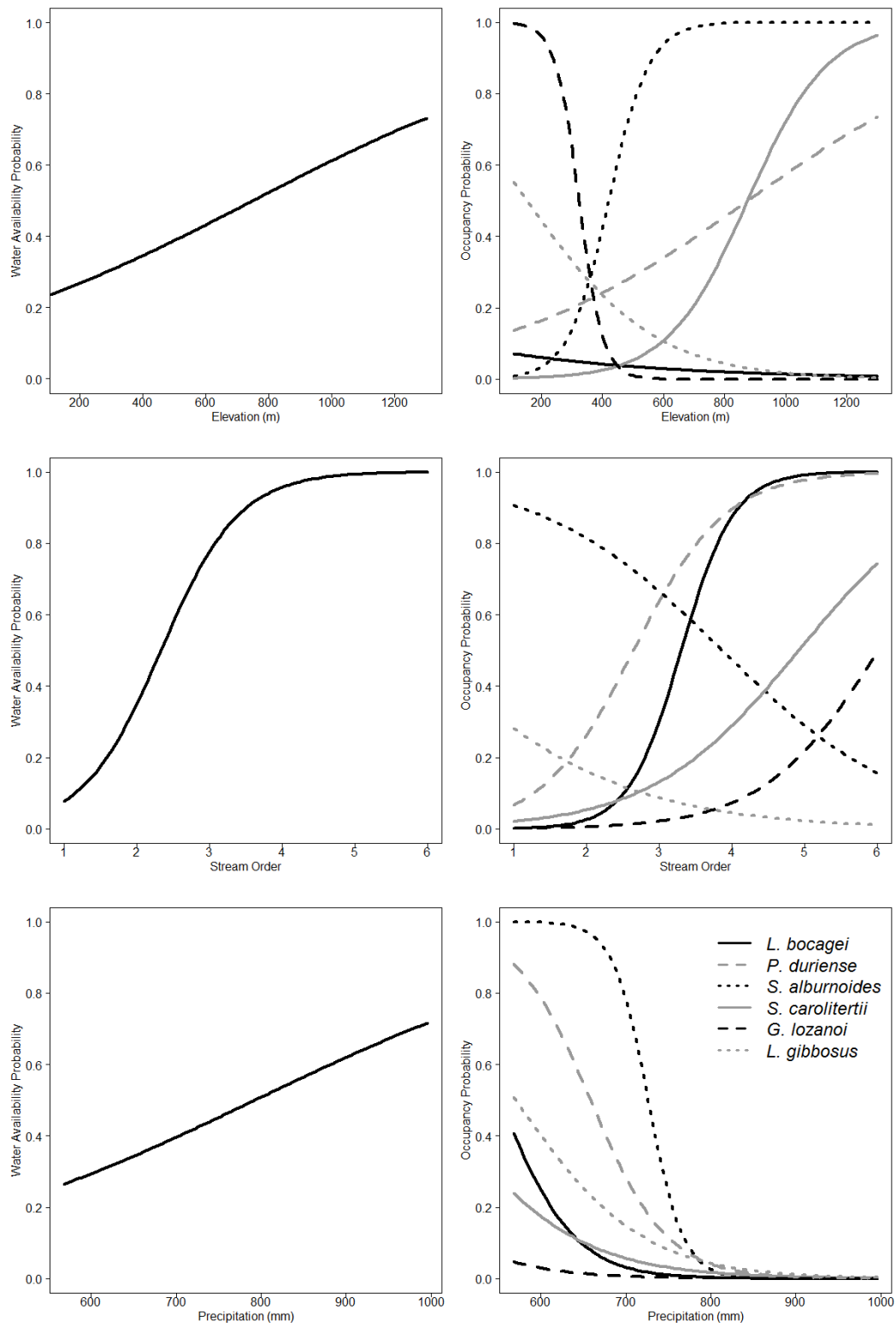


Figure 3.S2 - Relationships estimated from hierarchical occupancy-detection models between environmental variables and both the probability of the watercourse having water during the sampling visit, and the probability of occupancy for each fish species when surface water is present.



## Appendix S5 Maps of prediction uncertainty

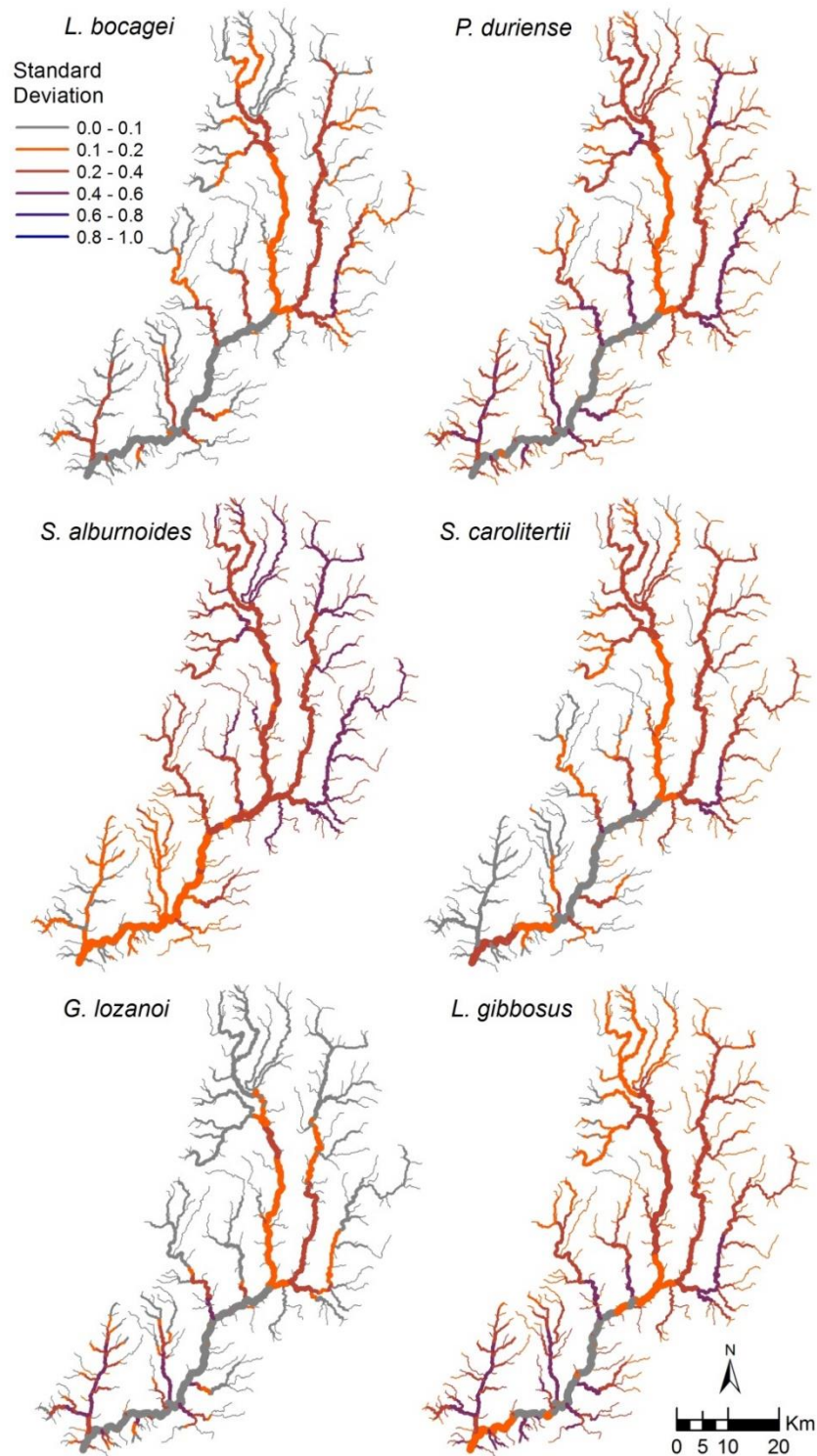


Figure 3.S3 - Maps of model prediction uncertainty for predicted distributions of six stream fish species in the Sabor catchment. Uncertainty was estimated from the standard deviation of the posterior distribution of occupancy probabilities derived from the best-supported model for each species

## Appendix S6 Supplementary references

- Chen, D.-G., Sun, J., & Peace, K. (2012). *Interval-Censored Time-to-Event Data: Methods and Applications* (D.-G. Chen, J. Sun, & K. Peace (eds.); Vol. 52). Chapman and Hall/CRC. <https://doi.org/10.1201/b12290>
- ESRI. (2011). *ArcMap* (10.2). Environmental Systems, Research Institute.
- Ferreira, M. T., Sousa, L., Santos, J. M., Reino, L., Oliveira, J., Almeida, P. R., & Cortes, R. V. (2007). Regional and local environmental correlates of native Iberian fish fauna. *Ecology of Freshwater Fish*, 16(4), 504–514. <https://doi.org/10.1111/j.1600-0633.2007.00241.x>
- Filipe, A. F., Marques, T. A., Tiago, P., Ribeiro, F., Da Costa, L. M., Cowx, I. G., & Collares-Pereira, M. J. (2004). Selection of Priority Areas for Fish Conservation in Guadiana River Basin, Iberian Peninsula. *Conservation Biology*, 18(1), 189–200. <https://doi.org/10.1111/j.1523-1739.2004.00620.x>
- Garrard, G. E., Bekessy, S. A., McCarthy, M. A., & Wintle, B. A. (2008). When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology*, 33(8), 986–998. <https://doi.org/10.1111/j.1442-9993.2008.01869.x>
- Garrard, G. E., McCarthy, M. a., Williams, N. S. G., Bekessy, S. a., & Wintle, B. a. (2013). A general model of detectability using species traits. *Methods in Ecology and Evolution*, 4(1), 45–52. <https://doi.org/10.1111/j.2041-210x.2012.00257.x>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis*. Springer New York. <https://doi.org/10.1007/978-1-4419-6646-9>
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier Academic Press.
- Magalhães, M. F., Beja, P., Canas, C., & Collares-Pereira, M. J. (2002). Functional heterogeneity of dry-season fish refugia across a Mediterranean catchment: the role of habitat and predation. *Freshwater Biology*, 47(10), 1919–1934. <https://doi.org/10.1046/j.1365-2427.2002.00941.x>
- Magalhães, M. F., Beja, P., Schlosser, I. J., & Collares-Pereira, M. J. (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, 52(8), 1494–1510. <https://doi.org/10.1111/j.1365-2427.2007.01781.x>
- McCarthy, M. a., Moore, J. L., Morris, W. K., Parris, K. M., Garrard, G. E., Vesk, P. a., Rumpff, L., Giljohann, K. M., Camac, J. S., Bau, S. S., Friend, T., Harrison, B., & Yue, B. (2013). The influence of abundance on detectability. *Oikos*, 122(5), 717–726. <https://doi.org/10.1111/j.1600-0706.2012.20781.x>
- Murphy, K. E., Carter, C. M., & Brown, S. O. (2002). The exponential distribution: the good, the bad and the ugly. A practical guide to its implementation. *Annual Reliability and Maintainability Symposium. 2002 Proceedings*, 550–555. <https://doi.org/10.1109/RAMS.2002.981701>
- Penczak, T., & Głowacki, Ł. B. (2008). Evaluation of electrofishing efficiency in a stream under natural and regulated conditions. *Aquatic Living Resources*, 21(3), 329–337. <https://doi.org/10.1051/alr:2008044>

- Podani, J., & Csontos, P. (2006). Quadrat size dependence, spatial autocorrelation and the classification of community data. *Community Ecology*, 7(1), 117–127. <https://doi.org/10.1556/ComEc>.
- Reynolds, J. B. (1996). *Electrofishing. Fisheries Techniques* (B. R. Murphy & D. W. Willis (eds.); 2nd ed.). American Fisheries Society.
- Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, 165(6), 710–718. <https://doi.org/10.1093/aje/kwk052>
- Vogt, J. V, Soille, P., Jager, A. De, Rimavičiūtė, E., Mehl, W., Haastrup, P., Paracchini, M. M. L., Dusart, J., Bódis, K., Foisneau, S., & Bamps, C. (2007). Developing a pan-European Database of Drainage Networks and Catchment Boundaries from a 100 Metre DEM. *Proceedings 10th AGILE Int. Conference on Geographic Information Science, May*, 8–11. [http://ams.jrc.it/publications/pdfs/JVogt\\_etal\\_AGILE2007.pdf](http://ams.jrc.it/publications/pdfs/JVogt_etal_AGILE2007.pdf)
- Zalewsky, M., & Cowx, I. G. (1990). Factors affecting the efficiency of electric fishing. In *Fishing with electricity: applications in freshwater fisheries management* (pp. 89–111). Fishing News Books.



## Chapter 4 – Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks





## Chapter 4 – Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks

Ana Filipa Filipe, Lorenzo Quaglietta, **Mário Ferreira**, Maria Filomena Magalhães, Pedro Beja

This chapter was published in the journal *Biological Invasions* and should be cite as:

Filipe, A. F., Quaglietta, L., Ferreira, M., Magalhães, M. F., & Beja, P. (2017). Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks. *Biological Invasions*, 19(10), 2899–2912. <https://doi.org/10.1007/s10530-017-1492-3>

Keywords:

Dendritic ecological networks; freshwater invasions; geostatistics; spatial autocorrelation; species distribution models; stream ecology

### Abstract

Species distribution models combining environmental and spatial components are increasingly used to understand and forecast species invasions. However, modelling distributions of invasive species inhabiting stream networks requires due consideration of their dendritic spatial structure, which may strongly constrain dispersal and colonization pathways. Here we evaluate the application of novel geostatistical tools to species distribution modelling in dendritic networks, using as case study two invasive crayfish (*Procambarus clarkii* and *Pacifastacus leniusculus*) in a Mediterranean watershed. Specifically, we used logistic mixed models to relate the probability of occurrence of each crayfish to environmental variables, while specifying three spatial autocorrelation components in random errors. These components described spatial dependencies between sites as a function of (i) straight-line distances (Euclidean model)

between sites, (ii) hydrologic (along the waterlines) distances between flow-connected sites (tail-up model), and (ii) hydrologic distances irrespective of flow connection (tail-down model). We found a positive effect of stream order on *P. clarkii*, indicating an association with the lower and mid reaches of larger streams, while *P. leniusculus* was affected by an interaction between stream order and elevation, indicating an association with larger streams at higher altitude. For both species, models including environmental and spatial components far outperformed the pure environmental models, with the tail-up and the Euclidean components being the most important for *P. clarkii* and *P. leniusculus*, respectively. Overall, our study highlighted the value of geostatistical tools to model the distribution of riverine and aquatic invasive species, and stress the need to specify spatial dependencies representing the dendritic network structure of stream ecosystems.

## 4.1 Introduction

The invasion of stream ecosystems by exotic species is a cause of concern worldwide due to their negative environmental and economic impacts (Strayer, 2010; Walsh et al., 2016). Therefore, understanding how invasive freshwater species spread into novel areas is essential to prevent further expansion, promote eradication, or adapt to their continued presence. This problem has often been addressed using distribution modelling tools, aiming to understand invasion drivers, predict species range expansions or contractions in relation to natural and anthropogenic factors, and to guide early detections (Capinha et al., 2013; Capinha & Anastácio, 2011; Larson & Olden, 2012; Siesa et al., 2011; Václavík & Meentemeyer, 2012). However, studies often ignore the specificities of aquatic organisms, particularly those living along dendritic stream networks, which may bias model results and ultimately mislead management prescriptions (Peterson et al., 2013).

While terrestrial species live and move in a two-dimensional space, the movements of stream species are constrained by the topology of the dendritic stream network, which strongly affects their distribution, persistence and diversity (Altermatt, 2013; Carrara et al., 2012). This applies to strictly aquatic species such as freshwater molluscs and fish, but also to species that move preferentially along waterlines but that can also travel overland such as amphibians and semi-aquatic mammals (e.g., Grant et al. 2010; Quaglietta et al. 2014). In invasive species, the stream spatial structure may constrain



the patterns of expansion from initial founder populations, with individuals dispersing up- and downstream and progressively colonizing favourable habitats across the stream network (Bernardo et al., 2011; Bronnenhuber et al., 2011; Hein et al., 2011). Stream reaches that are connected and close to founder populations may thus become occupied first, while farther or unconnected reaches may take more time to be colonized, even if there are favourable habitat conditions. Early during expansion, the distribution is unlikely to be in equilibrium with the environment, as a species may be absent from potentially suitable sites because of colonization time lag and dispersal limitations (Václavík & Meentemeyer, 2009, 2012). Over time, the species may progressively spread across the river network and eventually colonize all suitable habitats, thereby converging to an equilibrium with prevalent environmental conditions. At this stage, the stream network topology may still be important because, for instance, reaches sharing the same headwaters are likely to have similarities in terms of flow regime and water chemistry, thus providing similar habitat conditions (Carrara et al., 2012; Mcguire et al., 2014; Peterson et al., 2013). Failure to incorporate these spatial processes may introduce errors and biases in distribution modelling, such as over-estimating the importance of environmental factors (Diniz-Filho et al., 2003; Václavík & Meentemeyer, 2012), under-estimating potential distribution ranges (Václavík & Meentemeyer, 2012), or over-estimating actual distribution ranges (De Marco, Diniz-Filho, et al., 2008; Václavík et al., 2012; Václavík & Meentemeyer, 2009).

The use of geostatistical modelling to account for spatial dependencies in physical and ecological processes across stream networks was first introduced by Ver Hoef et al. (2006), providing a valuable tool to improve distribution modelling of aquatic invasive species. Geostatistical models are similar to conventional linear mixed models, with spatial autocorrelation specified in the random errors. In ecological applications, the deterministic mean of the dependent variable is modelled as a linear function of explanatory variables, and local deviations from the mean are modelled using the spatial autocorrelation between nearby sites (Peterson & Ver Hoef, 2010; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). This is specified using covariance functions, which represent the strength of the influence between sites as a function of the distance separating such sites. Distances can be straight-line (Euclidean) distances measured overland, or hydrologic distances measured along the flow lines (i.e. longitudinal connection). Hydrologic distances can represent flow-connected relations, assuming that a point downstream may be influenced by a point upstream, but not the reverse (tail-up models), or both flow-connected and flow-unconnected relations, assuming that

influences are not limited by flow direction (tail-down models). The stream network models of Ver Hoef & Peterson (2010) account for these multiple spatial relationships based on a mixture of covariances, each of which may be specified using a moving-average function (e.g. exponential, spherical, linear-with-sill). Covariance functions differ on how they specify the distance-decay of spatial influences, but typically they have three parameters: the nugget effect, representing the variation between sites when their separation distance approaches zero; the sill, representing the variance found among spatially independent sites; and the range, representing how fast the covariance decays with distance (Peterson et al., 2007). Application of these models to aquatic invasive species might provide information on the relative role of environmental versus spatial processes on current distributions, which can help clarifying whether the species is in equilibrium with the environment. Also, they might provide a basis to infer likely mechanisms of expansion (e.g. Siesa et al., 2011), by comparing for instance the relative importance of different spatial components. Finally, they might allow more robust inferences on species occurrences at unobserved sites, using observed values at nearby sites.

This study explores the use of geostatistical distribution modelling to clarify the conditions associated with the establishment of two invasive crayfish and to predict their potential expansion range in dendritic stream networks. We focused on the red swamp crayfish *Procambarus clarkii* and the signal crayfish *Pacifastacus leniusculus*, which are the two most widely distributed invasive crayfish worldwide (Gherardi et al., 2011; Hänfling et al., 2011). Previous studies have shown that these crayfish have different environmental requirements and that their invasion range may still not be in equilibrium with the environment (Capinha et al., 2013; Capinha & Anastácio, 2011), but they disregarded the potentially important role of spatial processes occurring across dendritic stream networks. Here we use a case study in a Mediterranean watershed of NE Portugal, aiming to: i) identify the main environmental factors explaining the current crayfish distributions, ii) quantify the relative importance of environmental and spatial processes influencing the distributions; iii) compare the relative importance of spatial processes described by the Euclidean, tail-up and tail-down models; and iv) predict the potential invasion ranges by combining environmental and spatial predictors. Results are used to discuss the value of geostatistical approaches in relation to conventional species distribution modelling to deal with aquatic invasive species.

## 4.2 Methods

### 4.2.1 Study area and species

The study was conducted in NE Portugal, in the river Sabor watershed (N 41°09'–42°00', W 7°15'–6° 15'; Figure 4.1), which covers a wide range of environmental conditions in terms of elevation (100-1500 meters above sea level), total annual precipitation (443 - 1163 mm), and mean annual temperature (6.9 – 15.6°C). Climate is Mediterranean, with precipitation largely concentrated in October-March, while it is virtually absent in the hot summer months (June-August). Flow regime is highly seasonal, with most headwater streams drying out or being reduced to a series of disconnected pools in summer, though the main watercourse and the largest tributaries are permanent. A large hydroelectric power plant involving two dams located near the mouth of the Sabor river were under construction during the study period (Jackson, 2011), but otherwise the river was largely free flowing except for a few small impoundments. See Ferreira et al. (2016) for details of the study area. The two crayfish species studied were intentionally introduced in Europe in the 1960s (*P. leniusculus*) and in the 1970s (*P. clarkii*), due to their economic value (Clavero, 2016), and currently have well-established populations (Capinha et al., 2013; Souty-Grosset et al., 2006). In the Sabor watershed, *P. clarkii* was first recorded in the 1990s (Bernardo et al., 2011), but possibly was present earlier because it has spread rapidly in the Iberian Peninsula since the first introduction in 1973 (Habsburgo-Lorena, 1978; Ramos & Pereira, 1981). Although *P. leniusculus* was also introduced in the Iberian Peninsula in the 1970s, it probably reached the Sabor only in 1994, following a deliberate introduction in the Spanish sector of the watershed (Bernardo et al., 2011). Recent studies suggest that the two species are still expanding in the Sabor watershed, which may be a consequence of dispersal from source populations along the stream network, in both the upstream and downstream directions (Anastácio et al., 2015; Bernardo et al., 2011). Dispersal overland may also occur in at least *P. clarkii* (Cruz & Rebelo, 2007; Ramalho & Anastácio, 2015), but its contribution to range expansion in the study area is unknown.

### 4.2.2 Crayfish survey

We visited 167 50-m stream reaches (hereafter sites) providing a homogeneous coverage of the Sabor catchment during the summer of 2012 (June 13 to July 15, and August 28 to September 15; Figure 4.). At sites where there was flowing water or isolated pools, we estimated the presence/absence of crayfish using a single anode electrofishing

gear, always operated by the same person (MF), and following standard procedures outlined in Ferreira et al. (2016). One additional researcher with a net was always present to pick up stunned crayfishes. After identification, all individuals were eliminated following national regulations for invasive species. The presence of crayfish in dry stream stretches was not assessed, though they may be able to persist therein by burrowing in soft sediments (Gherardi et al., 2011; Guan, 1994). Therefore, distribution modelling focused on crayfish that are active in the water during the dry summer season. Although this could potentially underestimate the true crayfish distribution, we believe this was not a serious issue, because incidental observations throughout the year suggest that crayfish are actually absent from headwater streams drying out in summer (Mário Ferreira and Lorenzo Quaglietta, Unpublished Data).

### 4.2.3 Environmental and spatial data

To model crayfish distribution, we used three potentially influential environmental variables (Capinha et al., 2013; Capinha & Anastácio, 2011; Moreira et al., 2015) that can be easily extracted from topographic online maps, and could thus be used to extrapolate the species distribution models to the entire watershed. Strahler's stream order was used as a proxy for habitat size and heterogeneity (Ferreira et al., 2016; Hughes et al., 2011), and it was extracted from CCM2 (Catchment Characterization and Modelling database), which is based on a 100-m resolution digital elevation model (DEM) (Vogt et al., 2007). Elevation and slope at each site were also estimated from information available in CCM2. We did not use climate variables, because they were often highly correlated with the topographic variables, and the range of conditions within the study area was well within the much wider climate niches of both species (Capinha et al., 2013; Capinha & Anastácio, 2011). All variables were screened for potential outliers and influential points.

Spatial data necessary for geostatistical modelling was obtained in a Geographic Information System (GIS) using the Sabor watershed network extracted from CCM2 and the layer of sampling locations. Estimates included the Euclidean and hydrologic distances (total and downstream hydrologic distances) between every pair of sampling sites (Peterson & Ver Hoef, 2010). We also estimated the watershed area draining to each site to be used in tail-up models (see below). Estimates were made using the Spatial Tools for the Analysis of River Systems (STARS) toolbox version 2.0.0 (Peterson & Ver Hoef, 2014) for ArcGIS 10.2 (ESRI, 2011).

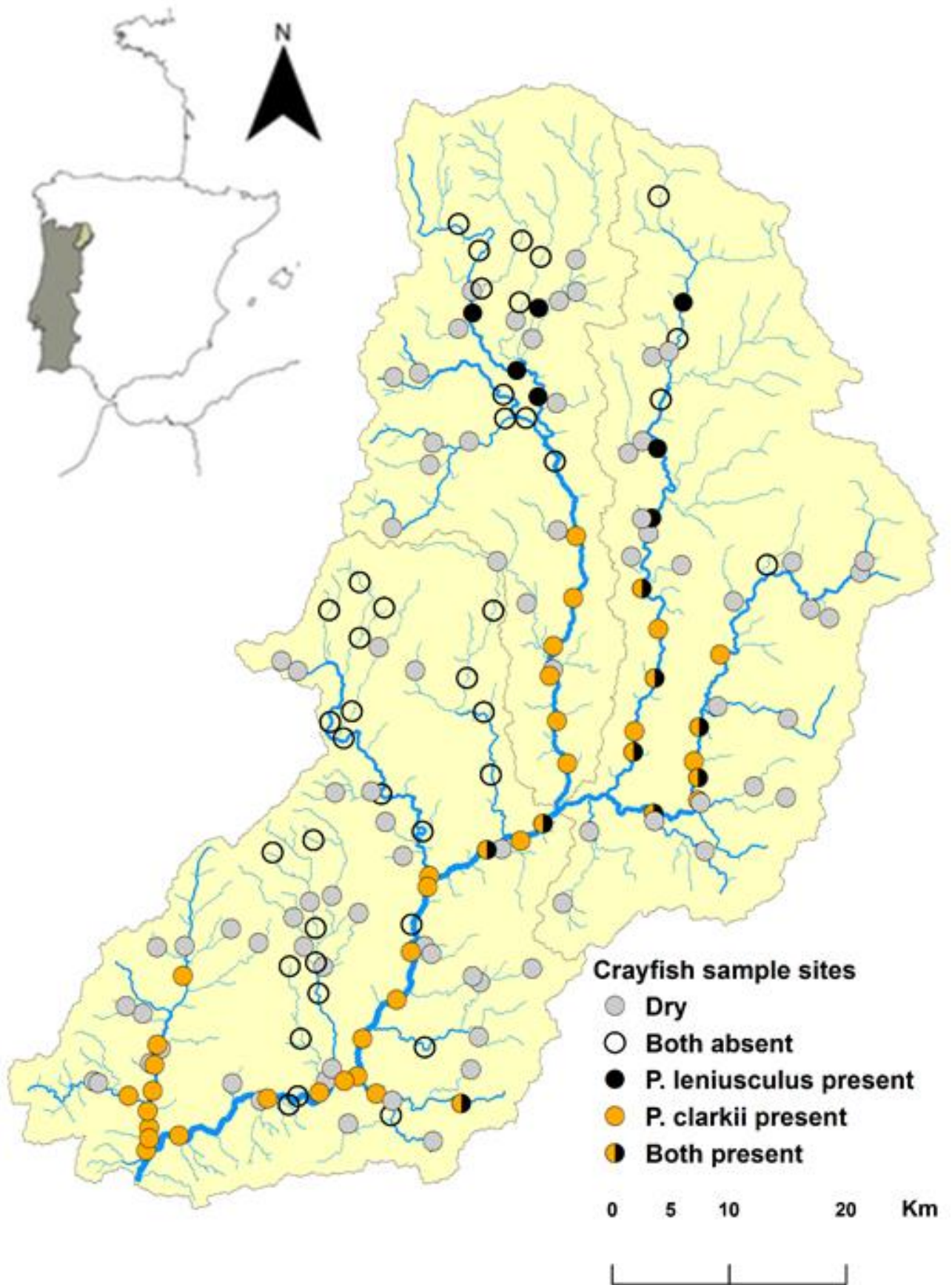


Figure 4.1- Map of the Sabor watershed (NE Portugal) and its location in the Iberian Peninsula, showing the spatial distribution of the sites visited, the sites sampled using electrofishing, and the sites where each crayfish species was detected.

#### 4.2.4 Geostatistical modelling

To visualise spatial dependencies along the stream network, we built empirical semivariograms depicting how semivariance in species presence/absence between pairs of sampling sites changed in relation to the hydrologic distances separating them (Torgegrams), considering either flow-connected or flow-unconnected sites (Peterson et al., 2013). The distribution (presence/absence) of each crayfish was then modelled in relation to environmental variables using a logistic function, with spatial autocorrelation specified through a full covariance mixture-model in the random component (Ver Hoef & Peterson, 2010). Therefore, the model incorporated spatial dependencies represented by tail-up (TU), tail-down (TD) and Euclidean (EUC) models. To deal with confluences in tail-up models, the spatial weights to allocate the moving-average function between upstream segments were based on watershed areas (e.g. Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010).

For each species, model building followed a two-step procedure (Peterson & Ver Hoef, 2010). First, we selected the environmental component while maintaining constant the spatial component, which included exponential tail-up, linear-with-sill tail-down, and Gaussian Euclidean spatial covariance functions. We tested all combinations of environmental variables and their interaction terms, and selected in each case the model with the lowest root-mean-squared-prediction error (RMSPE) estimated through leave-one-out cross-validation (Frieden et al., 2014; Peterson & Ver Hoef, 2010; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). Interactions were only specified when the main effects were also included in the model. We then built the spatial component while maintaining the best environmental model selected in the previous step. We tested all combinations of spatial components, using exponential, spherical, Gaussian and Cauchy functions as alternatives for the Euclidean model, and exponential, linear-with-sill, spherical and Mariah functions as alternatives for the tail-up and tail-down models (Ver Hoef & Peterson, 2010). We selected the best function for each spatial component based on the minimization of RMSPE criterion, and kept the three components in the model to allow estimates of the proportion of variation explained by each one. Besides this full “spatial/environmental model”, we also built a pure “environmental model” based on a simple logistic regression, and a pure “spatial model”, including in each case the variables and functions selected in model building. The discrimination ability of each model was estimated using the area under the receiver operating characteristic curve (AUC) (e.g., Václavík & Meentemeyer, 2009), and Cohen’s kappa using species prevalence as the threshold for predicted presences (Titus et al., 1984).

We performed all analyses in the R version 3.1.2 (R Core Team, 2014) and the SSN (Spatial Stream Networks) packages (Ver Hoef et al., 2014), and used ArcGIS 10.2 for mapping (ESRI, 2011). AUC was computed using the package pROC (Robin et al., 2011), and Cohen's kappa was computed using the package irr (Gamer et al., 2012).

#### 4.2.5 Species distribution mapping

The Sabor stream network was divided in 1716 segments, which were used as a basis to predict the distribution of each crayfish using the species distribution models (Ferreira et al., 2016). For each segment, we extracted the environmental variables, the area of the watershed discharging into the segment, and the in-stream and Euclidean distances to every other segment. Prediction of the probability of each species being present in each segment was then computed using universal kriging (Cressie, 1993). Maps of species potential distribution across the watershed were then produced, using the prevalence of each species as the threshold for separating segments with predicted presence or absence. Maps were built using either the "spatial/environmental model" or the "environmental model", to assess whether considering the spatial network structure improved the species distribution mapping.

### 4.3 Results

From the 167 visited sites, 87 had flowing water or isolated pools and were thus sampled for crayfish (Figure 4.1). From these, we detected *P. clarkii* at 41 (46.1%) and *P. leniusculus* at 16 (18.0%) sites. Visual analysis of Torgegrams indicated that there were spatial dependencies in species occurrences along the stream network, as suggested by the patterns of change in semivariance between sites in relation to the hydrologic distances separating them (Figure 4.2). For *P. clarkii*, there was a rapid and nearly linear increase in semivariance with distance between flow-connected sites, levelling off at about 30km, while semivariance between flow-unconnected sites was generally much smaller and it increased slowly with distance between sites. For *P. leniusculus*, the toregrams showed much less marked patterns than for *P. clarkii*, though semivariance was also smaller between flow-unconnected than flow-connected sites.

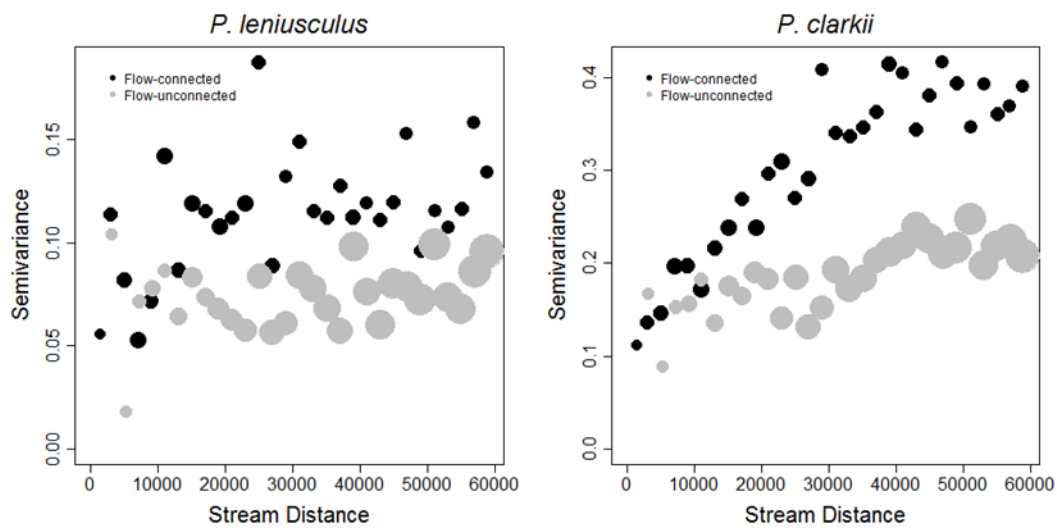


Figure 4.2 - Torgegrams describing spatial dependencies in the distribution (presence/absence) of *Procambarus clarkii* and *Pacifastacus leniusculus* along the stream network of the Sabor watershed (NE Portugal). The graphs show changes in semivariance between sampling sites in relation to the hydrologic distances separating them, considering either flow-connected (black circles) or flow-unconnected (grey circles) sites. The size of circles is proportional to the number of pairs of sites used to estimate the semivariance.

Table 4.1 - Parameter estimates and summary statistics for the environmental component of crayfish distribution models in the Sabor watershed (NE Portugal). For each species we present the parameter estimates for the pure environmental, the pure spatial, and the spatial/environmental models. For each variable included in the best model, we provide the regression coefficient (Coef.), the standard error of the estimate (SE), the corresponding t- and P-values. For each model, we also provide the area under the receiver operating characteristic curve (AUC), and the Cohen's kappa (k).

Models	Variables	Coef.	SE	t-values	P-values	AUC	k
<i>Procambarus clarkii</i>							
Environmental	Intersect	-6.71	0.88	-7.60	<0.001	0.86	0.45
	Stream order	1.35	0.20	6.78	<0.001		
Spatial	Intersect	-1.12	0.18	-6.23	<0.001	0.98	0.63
Spatial/Environmental	Intersect	-5.52	1.11	-4.97	<0.001	0.96	0.84
	Stream order	1.15	0.20	5.83	<0.001		
<i>Pacifastacus leniusculus</i>							
Environmental	Intersect	-0.97	2.61	-0.37	0.711	0.78	0.32
	Stream order	-0.76	0.52	-1.46	0.147		
	Elevation	-0.01	0.00	-2.32	0.022		
	Stream order × Elevation	0.00	0.01	3.44	<0.001		
Spatial	Intersect	-2.25	0.26	-8.51	<0.001	0.90	0.22
Spatial/Environmental	Intersect	-0.14	1.25	-0.11	0.912	0.82	0.38
	Stream order	-0.53	0.26	-2.05	0.042		
	Elevation	-0.01	0.01	-2.90	0.004		
	Stream order × Elevation	0.00	0.00	3.67	<0.001		



Table 4.2 - Comparison of mixture models relating the probability of occurrence of two crayfish species to environmental and spatial components. The models shown represent the best model fit for each mixture type based on the root mean square prediction error (RMSPE). Models are Euclidean (EUC), tail-up (TU) and tail-down (TD), and they were tested using exponential, spherical, Gaussian and Cauchy functions as alternatives for EUC, and exponential, linear-with-sill, spherical and Mariah functions as alternatives for TU and TD.

Mixture	Model 1	Model 2	Model 3	RMSPE
<i>Procambarus clarkii</i>				
Nonspatial				3.081
Spatial				
TU	Linear-sill			3.004
TD	Linear-sill			2.798
EUC	Cauchy			2.378
TU/TD	Linear-sill	Linear-sill		2.356
TU/EUC	Linear-sill	Spherical		2.235
TD/EUC	Mariah	Spherical		2.336
TU/TD/EUC	Linear-sill	Linear-sill	Spherical	2.237
<i>Pacifastacus leniusculus</i>				
Nonspatial				3.005
Spatial				
TU	Mariah			3.079
TD	Mariah			2.324
EUC	Exponential			1.623
TU/TD	Mariah	Linear-sill		1.826
TU/EUC	Mariah	Gaussian		1.595
TD/EUC	Exponential	Gaussian		1.412
TU/TD/EUC	Linear-sill	Spherical	Cauchy	1.424

In model building, the environmental component with the lowest RMSPE for *P. clarkii* included only the positive effect of stream order on the probability of species occurrence (Table 4.1, Table 4.S1). The environmental component for *P. leniusculus* included stream order, elevation, and their interaction term (Table 4.1, Table 4.S2), indicating that the probability of occurrence increased in higher order streams at higher elevation (Table 4.1, Figure 4.3). Regarding the spatial component, the best covariance structure for *P. clarkii* included a tail-up linear-with-sill function and a Euclidean spherical function (Table 4.2, Table 4.S3). The model with the three spatial components included also a tail-down linear-with-sill function, and it was very close to the best considering the RMSPE criterion (Table 4.2, Table 4.S3). The range of the tail-up (31.8km) and Euclidean (188.2km) components were similar in both models, and much smaller than the range estimated for

the tail-down component (627.4km). For *P. leniusculus*, the best spatial model included a tail-down exponential function, and a Euclidean Gaussian function (Table 4.2). The full autocovariance mixture model was also the second best, including a tail-up linear-with-sill-function, a tail-down spherical function, and a Euclidean Cauchy function (Table 4.2, Table 4.S4). The range estimate for the tail-up component was very small (1.4km),

Table 4.3 - Percentage of variance accounted for by models relating the probability of two crayfish species to environmental and spatial components. For each species we present the values for the model including only environmental variables (nonspatial), for the spatial/environmental model minimising the root mean square prediction error (Best model), and for the full mixture spatial/environmental model (Full covariance mixture). Separate percentages are provided for the Euclidean (EUC), tail-up (TU) and tail-down (TD) components of the spatial model. The nugget is the variation unexplained in models.

<sup>a</sup> Given the very small range estimated (1.4km), this component was acting like a nugget effect.

	Environmental	Spatial			Nugget
		TU	TD	EUC	
<i>Procambarus clarkii</i>					
Nonspatial	21.8	-	-	-	78.2
Best model	17.1	20.8	-	13	49.2
Full covariance mixture	17.1	20.8	≈ 0.0	12.8	49.3
<i>Pacifastacus leniusculus</i>					
Nonspatial	16.2	-	-	-	83.8
Best model	10.8	-	≈ 0.0	26	63.2
Full covariance mixture	11	63.0 <sup>a</sup>	0.2	24.5	1.3

while those of the tail-down (634.8-645.93km) and Euclidean (371.1km) components were very large.

For both crayfish species, the model combining environmental and spatial components outperformed both the pure environmental and the pure spatial model, in terms of variance explained, AUC and Cohen's kappa (Table 4.1 and Table 4.3). In the spatial/environmental model, the environmental component always explained less variance than the spatial component (Table 4.3). The full covariance mixture model for *P. clarkii* accounted for about half the variance in species presence/absence, of which about 41% corresponded to tail-up spatial dependencies, another 25% to Euclidean spatial dependencies, and 34% to environmental variables (Table 4.3). The predictive accuracy of the model was very satisfactory, as measured using either AUC or Cohen's kappa (Table 4.1). For *P. leniusculus*, the full mixture model accounted for almost all variance, most of which (64%) corresponding to the tail-up component, and the rest by

the Euclidean (25%) and environmental (11%) components. However, the very small range estimated for the tail-up component (1.4km) indicates that autocorrelation was essentially zero between all sampling sites, so this component is acting like a nugget effect. Predictive accuracy was reasonable when measured through AUC, but low when measured using Cohen's kappa (Table 4.1).

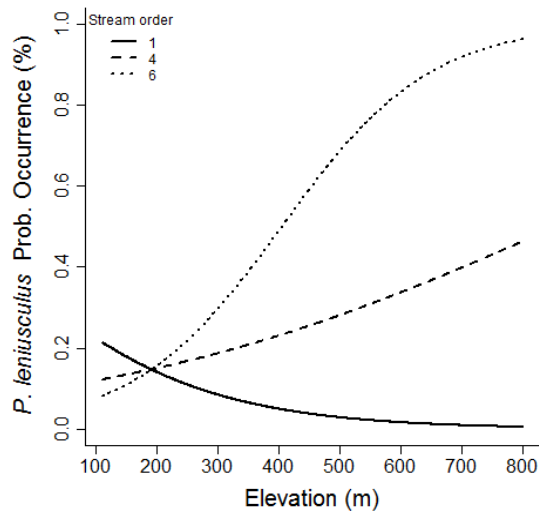


Figure 4.3 - Response curves estimated from a logistic mixed model relating the probability of occurrence of *Pacifastacus leniusculus* to environmental variables, describing the interaction effect between stream order and

The distribution models based solely on environmental variables produced poor results for the potential distribution of both species, while improving markedly when considering the spatial component (Figure 4.4). For *P. clarkii*, the environmental model correctly predicted a continuous distribution along the mid and lower reaches of the Sabor river, but it produced many false absences along the main tributaries. In contrast, the environmental/spatial model predicted a distribution much closer to that observed, highlighting a continuous occurrence in the main river and large tributaries. The environmental

model for *P. leniusculus* also predicted a distribution much restricted than that observed, while the environmental/spatial model produced a larger distribution that was closer to that observed. However, this model missed sections of river where the species was detected, while predicting a far more continuous distribution than that observed.

#### 4.4 Discussion

Our study found that the occurrence of *P. clarkii* and *P. leniusculus* in a Mediterranean stream network was strongly related to environmental variables such as stream order and elevation, but also that species distributions appeared strongly shaped by spatial processes. Actually, we found that spatial dependencies accounted for a larger proportion of variation in species occurrences than environmental variables, and that disregarding spatial effects seemed to strongly underestimate potential species distributions. These general patterns were observed for both species, though model performance was much lower for the species with a more restricted and patchy

distribution (*P. leniusculus*), than for the species with a larger a more continuous distribution (*P. clarkii*). Overall, our results point out the importance of considering both environmental and spatial effects when modelling the distribution of invasive aquatic organisms, and stress the need to specify spatial dependencies representing the dendritic network structure of stream ecosystems.

#### 4.4.1 Environmental effects

The only variable included in the environmental model for *P. clarkii* was stream order, indicating that probability of occurrence increased from the headwaters to the large streams and the main river. This is in line with previous studies suggesting that this species often occurs in lowland rivers, with lower slopes, larger width, slower flowing waters, abundant aquatic vegetation and finer sediments (e.g., Anastácio et al., 2015; Cruz & Rebelo, 2007; Moreira et al., 2015), which often are associated with high-order streams (Allan & Castillo, 2007). On the other hand, absence from low order streams (i.e., small tributaries) was probably a consequence of these drying out in summer (Ferreira et al., 2016), and so they were not occupied by active crayfish at the time of sampling. The pure environmental model correctly predicted the extent of species distribution in the main Sabor river, but it underestimated considerably the extent of occurrence in its main tributaries. This could be a consequence of this study using only a limited set of environmental variables, thereby failing to fully identify the habitat conditions required by the species. It is also possible, however, that the distribution in the tributaries was strongly influenced by spatial processes associated with expansion from the core distribution in the main river (Bernardo et al., 2011), thereby confounding the operation of environmental drivers (see below).

In the case of *P. leniusculus*, the environmental model included the interactive effects of stream order and elevation, suggesting that the species prefers high-order streams at higher elevation, while being absent from lowland high-order streams. This is in line with previous studies suggesting that the species is highly adapted to mountain rivers, where it prefers riffle habitats with abundant riparian vegetation (Anastácio et al., 2015; Rallo & García-Arberas, 2002). However, the species was absent from the small, lower order streams, possibly because the headwaters of the Sabor watershed generally dry out in summer (Ferreira et al., 2016). The pure environmental model provided a relatively crude picture of the species distribution, producing both false absences and false presences. This may be a consequence, at least partly, of the relatively small number of presences detected in our study (16), which can cause problems in logistic models (Vittinghoff &

McCulloch, 2007). In alternative, this may be a consequence of the species being still in rapid expansion from several points of introduction (Bernardo et al., 2011; Amilcar Teixeira, personal communication), and thus remaining out of equilibrium with the environmental conditions. Whatever the mechanisms, these results suggest that the distribution of both species in stream networks may be poorly predicted by purely environmental models.

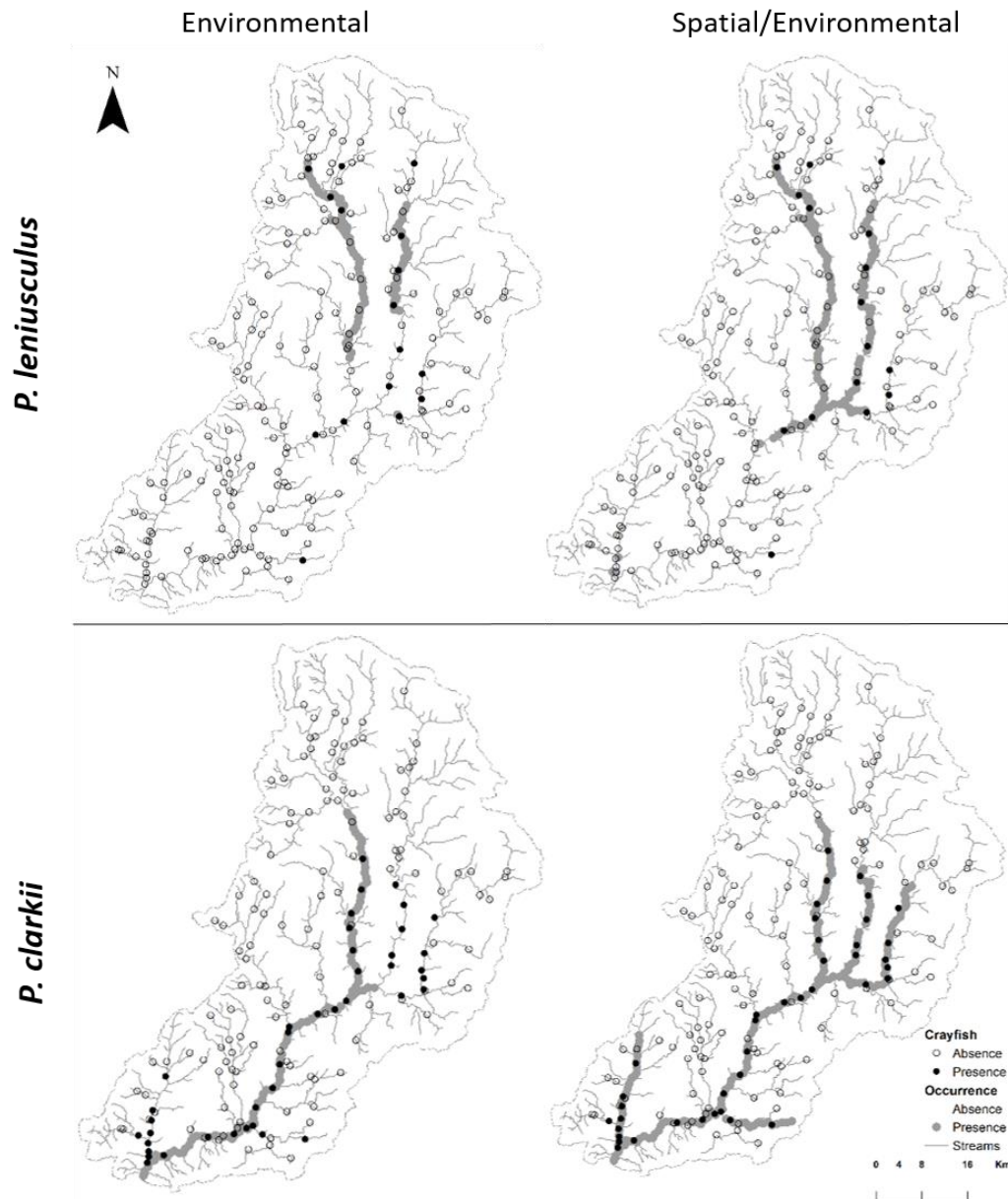


Figure 4.4 - Maps of potential distribution of *Procambarus clarkii* and *Pacifastacus leniusculus* in the river Sabor watershed (NE Portugal), predicted from either simple environmental models (left panels), or models including both environmental and spatial components (right panels). The threshold for predicted presences was set equal to the observed prevalence of each species.

#### 4.4.2 Spatial effects

The models of both species greatly improved when autocovariance functions accounting for spatial processes were considered. For *P. clarkii*, there was a particularly strong tail-up component, suggesting that although presences were spatially autocorrelated along streams, the prevalence of the species can vary widely between stream branches just above a confluence. This result, in combination with the spatial distribution pattern inferred from the environmental/spatial model, is compatible with the idea that the species expands in an invasion front that moves upstream from the lowland sections of the main river. However, expansion may advance primarily through some stream branches, possibly corresponding to higher order streams, as underlined by the environmental model, while colonization may be slower or do not occur in other branches, possibly the smaller tributaries. Previous studies in the Sabor watershed have indeed shown a progressive upstream expansion (Bernardo et al., 2011), and radiotracking revealed that individual crayfish have the ability to move considerably in the upstream direction (up to 250m in half day; Anastácio et al., 2015). In contrast to our results, previous studies found that *P. clarkii* occurrences were spatially autocorrelated only up to about 2-2.5 km (Cruz & Rebelo, 2007; Siesa et al., 2011), which is much smaller than the  $\approx 30$ km range estimated in our study. This may be due to differences in habitat characteristics, with previous studies largely focusing on naturally disconnected waterbodies such as ponds and small lakes, where dispersal may be more difficult than in a continuous stream network such as that of our study area.

The model for *P. clarkii* also included a significant Euclidean component, but the underlying mechanism was unclear. However, the large range estimated ( $\approx 190$ km) suggests that spatial autocorrelation was mainly due to broad scale distribution trends, rather than small scale processes. This may be a consequence of the topology of the watershed, as the species was consistently absent in zones occupied by headwater streams, while it was present in three waterlines that run parallel and at short distance to each other. These circumstances probably determined a positive relation between occupancy status and straight-line distances between sites, albeit with no ecological meaning. It could not be ruled out, however, that the Euclidean pattern was at least partly due to the effect of an unmeasured, spatially-structured environmental variable, though it is uncertain what this variable might be as all the area is within the climate niche of the species (Capinha et al., 2013). It is also unlikely that the Euclidean component reflected colonisations occurring by individuals moving through terrestrial habitats. Although this process has been described for *P. clarkii* inhabiting lowland habitats (Cruz & Rebelo,

2007; Ramalho & Anastácio, 2015), it is unlikely that it contributed much to species expansion in a dry and mountainous area such as ours, where dispersal overland would imply long movements across very dry habitats and steep slopes.

The spatial models developed for *P. leniusculus* differed considerably and were generally poorer than those of *P. clarkii*. Furthermore, there appeared to be some model instability, as shown by the full covariance mixture model, where the tail-up component acted like a nugget effect, with a range estimate (1.4km) that was smaller than the typical distance between nearest sites ( $\approx 5$ km). Nevertheless, it is noteworthy that there were spatial dependencies described by the Euclidean component, which explained over twice the variation in *P. leniusculus* distribution as the environmental component. As for *P. clarkii*, the large range estimated for the Euclidean component ( $\approx 370$ km) suggests that it described primarily broad scale trends in species distribution, rather than small scale processes. Including the spatial component improved the performance of the pure environmental model, but even so the predictive ability of the best environmental/spatial model was low, with a large number of false presences and absences. These patterns are probably a consequence of the low number of presences detected for this species, as noted for the environmental component, but they may also result from its patchy distribution, which was possibly caused by multiple introductions followed by progressive expansions that are still far from complete. This likely affected both environmental and spatial relationships, because the species was likely absent from sites with adequate environmental conditions, while the spatial dependencies were inconsistent, possibly due to the spatial scattering of the introductions. As this species is still rapidly spreading in the study area (Anastácio et al., 2015; Bernardo et al., 2011; Amilcar Teixeira, personal communication), we expect that the spatial distribution will keep changing for some more time, eventually reaching a stable pattern corresponding to the equilibrium with environmental conditions. In the meantime, we predict that sites that were identified in our study as false absences will soon be occupied, and that the species will progressively occupy most, if not all larger streams at higher elevation.

#### 4.4.3 Distribution modelling of invasive species across stream networks

Our study adds to recent research showing the importance of explicitly accounting for dendritic spatial structures when modelling the distribution of physical and ecological processes across stream networks, and the value of geostatistical tools to undertake

such modelling (Mcguire et al., 2014; Peterson et al., 2013). Also, our study provides novel insights on the application of this approach to invasive species, whose distribution modelling is affected by a specific set of challenging processes such as non-equilibrium with environmental conditions and highly dynamic distribution patterns (De Marco, Alexandre, et al., 2008; Václavík & Meentemeyer, 2012). First, our results reinforce the idea that both environmental and spatial processes need to be incorporated to predict the distribution of invasive species (De Marco, Alexandre, et al., 2008; Siesa et al., 2011; Václavík & Meentemeyer, 2012). Second, we confirmed that Euclidean distances are insufficient to incorporate spatial structure in distribution models for invasive stream organisms, as they may fail to account for their strongly constrained dispersal along waterlines and may produce biased and biologically meaningless results (Altermatt, 2013; Filipe et al., 2010; Peterson et al., 2013). Third, considering the tail-up autocovariance function may be generally recommended when testing for spatial dependencies, as it accounts for biological meaningful ecological processes such as connectivity along flow-connected waterlines (Carrara et al., 2012). Finally, geospatial models such as those applied here may improve inferences on the invasion processes of aquatic organisms from distribution data (e.g., Siesa et al., 2011; Václavík & Meentemeyer, 2012), by accounting for biologically more meaningful spatial dependencies along waterlines, though careful interpretation of results is required. Care should particularly be taken when dealing with species at the early stages of invasion, as it was the case of *P. leniusculus* in our study, where a small number of occurrences and a very fragmented distribution may produce models that are difficult to interpret. Overall, we suggest that geostatistical modelling across stream networks provides an important addition to the toolbox of researchers interested in biological invasions of aquatic organisms, which may contribute to address this global environmental problem by helping to understand driving mechanisms and to predict future distributions (Strayer, 2010).

## Acknowledgments

This work is part of the Baixo Sabor Long Term Ecological Research (LTER) project, funded by Portuguese Science and Technology Foundation (FCT) through LTER/BIA-BEC/0004/2009, and by EDP Energias de Portugal. AFF was supported by the FRESHING Project funded by FCT and COMPETE (PTDC/AAG-MAA/2261/2014 – POCI-01-0145-FEDER-356 016824), and MF was supported by FCT PhD grant



SFRH/BD/95202/2013. We thank Pedro Silva, Rita Severino, Sara Ivone and Sérgio Henriques for their collaboration in field work, and Duarte Prata for preliminary analysis of the data,

## 4.5 References

- Allan, J. D., & Castillo, M. M. (2007). Stream Ecology. In *Stream Ecology*. Springer Netherlands. <https://doi.org/10.1007/978-1-4020-5583-6>
- Altermatt, F. (2013). Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology*, *47*(3), 365–377. <https://doi.org/10.1007/s10452-013-9450-3>
- Anastácio, P. M., Banha, F., Capinha, C., Bernardo, J. M., Costa, A. M., Teixeira, A., & Bruxelas, S. (2015). Indicators of movement and space use for two co-occurring invasive crayfish species. *Ecological Indicators*, *53*, 171–181. <https://doi.org/10.1016/j.ecolind.2015.01.019>
- Bernardo, J. M., Costa, a. M., Bruxelas, S., & Teixeira, a. (2011). Dispersal and coexistence of two non-native crayfish species ( *Pacifastacus leniusculus* and *Procambarus clarkii* ) in NE Portugal over a 10-year period. *Knowledge and Management of Aquatic Ecosystems*, *401*, 28. <https://doi.org/10.1051/kmae/2011047>
- Bronnenhuber, J. E., Dufour, B. A., Higgs, D. M., & Heath, D. D. (2011). Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round goby, *Neogobius melanostomus*, in Great Lakes tributaries. *Molecular Ecology*, *20*(9), 1845–1859. <https://doi.org/10.1111/j.1365-294X.2011.05030.x>
- Capinha, C., & Anastácio, P. (2011). Assessing the environmental requirements of invaders using ensembles of distribution models. *Diversity and Distributions*, *17*(1), 13–24. <https://doi.org/10.1111/j.1472-4642.2010.00727.x>
- Capinha, C., Brotons, L., & Anastácio, P. (2013). Geographical variability in propagule pressure and climatic suitability explain the European distribution of two highly invasive crayfish. *Journal of Biogeography*, *40*(3), 548–558. <https://doi.org/10.1111/jbi.12025>
- Carrara, F., Altermatt, F., Rodriguez-Iturbe, I., & Rinaldo, A. (2012). Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences*, *109*(15), 5761–5766. <https://doi.org/10.1073/pnas.1119651109>
- Clavero, M. (2016). Species substitutions driven by anthropogenic positive feedbacks: Spanish crayfish species as a case study. *Biological Conservation*, *193*, 80–85. <https://doi.org/10.1016/j.biocon.2015.11.017>
- Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119115151>
- Cruz, M. J., & Rebelo, R. (2007). Colonization of freshwater habitats by an introduced crayfish, *Procambarus clarkii*, in Southwest Iberian Peninsula. *Hydrobiologia*, *575*(1), 191–201. <https://doi.org/10.1007/s10750-006-0376-9>
- De Marco, P., Alexandre, J., Diniz-Filho, F., & Bini, L. M. (2008). Spatial analysis improves species distribution modelling during range expansion. *Biol. Lett*, *4*, 577–580. <https://doi.org/10.1098/rsbl.2008.0210>
- De Marco, P., Diniz-Filho, J. A. F., & Bini, L. M. (2008). Spatial analysis improves species

- distribution modelling during range expansion. *Biology Letters*, 4(5), 577–580. <https://doi.org/10.1098/rsbl.2008.0210>
- Diniz-Filho, J. A. F., Bini, L. M., & Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, 12(1), 53–64. <https://doi.org/10.1046/j.1466-822X.2003.00322.x>
- ESRI. (2011). *ArcMap* (10.2). Environmental Systems, Research Institute.
- Ferreira, M., Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>
- Filipe, A. F., Magalhães, M. F., & Collares-Pereira, M. J. (2010). Native and introduced fish species richness in Mediterranean streams: the role of multiple landscape influences. *Diversity and Distributions*, 16(5), 773–785. <https://doi.org/10.1111/j.1472-4642.2010.00678.x>
- Frieden, J. C., Peterson, E. E., Angus Webb, J., & Negus, P. M. (2014). Improving the predictive power of spatial statistical models of stream macroinvertebrates using weighted autocovariance functions. *Environmental Modelling & Software*, 60, 320–330. <https://doi.org/10.1016/j.envsoft.2014.06.019>
- Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various Coefficients of Interrater Reliability and Agreement*. CRAN. <https://cran.r-project.org/web/packages/irr/>
- Gherardi, F., Aquiloni, L., Diéguez-Uribeondo, J., & Tricarico, E. (2011). Managing invasive crayfish: Is there a hope? In *Aquatic Sciences* (Vol. 73, Issue 2, pp. 185–200). SP Birkhäuser Verlag Basel. <https://doi.org/10.1007/s00027-011-0181-z>
- Grant, E. H. C., Nichols, J. D., Lowe, W. H., & Fagan, W. F. (2010). Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. *Proceedings of the National Academy of Sciences*, 107(15), 6936–6940. <https://doi.org/10.1073/pnas.1000266107>
- Guan, R.-Z. (1994). Burrowing behaviour of signal crayfish, *Pacifastacus lemjucuujs* (Dana), in the river Great Ouse, England. *Freshwater Forum*, 4, 144–168.
- Habsburgo-Lorena, A. S. (1978). Present situation of exotic species of crayfish introduced into Spanish continental waters. *Freshwater Crayfish*, 4, 175–184.
- Hänfling, B., Edwards, F., & Gherardi, F. (2011). Invasive alien Crustacea: dispersal, establishment, impact and control. *BioControl*, 56(4), 573–595. <https://doi.org/10.1007/s10526-011-9380-8>
- Hein, C. L., Öhlund, G., & Englund, G. (2011). Dispersal through stream networks: modelling climate-driven range expansions of fishes. *Diversity and Distributions*, 17(4), 641–651. <https://doi.org/10.1111/j.1472-4642.2011.00776.x>
- Hughes, R. M., Kaufmann, P. R., & Weber, M. H. (2011). National and regional comparisons between Strahler order and stream size. *Journal of the North American Benthological Society*, 30(1), 103–121. <https://doi.org/10.1899/09-174.1>
- Jackson, A. L. R. R. (2011). Renewable energy vs. biodiversity: Policy conflicts and the future of nature conservation. *Global Environmental Change*, 21(4), 1195–1208. <https://doi.org/10.1016/J.GLOENVCHA.2011.07.001>
- Larson, E. R., & Olden, J. D. (2012). Using avatar species to model the potential distribution of emerging invaders. *Global Ecology and Biogeography*, 21(11), 1114–1125.

<https://doi.org/10.1111/j.1466-8238.2012.00758.x>

Mcguire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., & Bailey, S. W. (2014). *Network analysis reveals multiscale controls on streamwater chemistry*. 111(19). <https://doi.org/10.1073/pnas.1404820111>

Moreira, F. D., Ascensão, F., Capinha, C., Rodrigues, D., Segurado, P., Santos-Reis, M., & Rebelo, R. (2015). Modelling the risk of invasion by the red-swamp crayfish (*Procambarus clarkii*): incorporating local variables to better inform management decisions. *Biological Invasions*, 17(1), 273–285. <https://doi.org/10.1007/s10530-014-0725-y>

Peterson, E. E., Theobald, D. M., & Ver Hoef, J. M. (2007). Geostatistical modelling on stream networks: Developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology*, 52(2), 267–279. <https://doi.org/10.1111/j.1365-2427.2006.01686.x>

Peterson, E. E., & Ver Hoef, J. M. (2010). A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91(3), 644–651. <https://doi.org/10.1890/08-1668.1>

Peterson, E. E., & Ver Hoef, J. M. (2014). Stars: An ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *Journal of Statistical Software*, 56(2), 1–17. <https://doi.org/10.18637/jss.v056.i02>

Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>

Quaglietta, L., Fonseca, V. C., Mira, A., & Boitani, L. (2014). Sociospatial organization of a solitary carnivore, the Eurasian otter (*Lutra lutra*). *Journal of Mammalogy*, 95(1), 140–150. <https://doi.org/10.1644/13-MAMM-A-073.1>

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. (3.1.2). R Foundation for Statistical Computing.

Rallo, A., & García-Arberas, L. (2002). Differences in abiotic water conditions between fluvial reaches and crayfish fauna in some northern rivers of the Iberian Peninsula. *Aquatic Living Resources*, 15(2), 119–128. [https://doi.org/10.1016/S0990-7440\(02\)01156-7](https://doi.org/10.1016/S0990-7440(02)01156-7)

Ramalho, R. O., & Anastácio, P. M. (2015). Factors inducing overland movement of invasive crayfish (*Procambarus clarkii*) in a ricefield habitat. *Hydrobiologia*, 746(1), 135–146. <https://doi.org/10.1007/s10750-014-2052-9>

Ramos, M. A., & Pereira, T. M. G. (1981). Um Novo Astacidae para a Fauna Portuguesa: *Procambarus clarkii* (Girard, 1852). In *Bol. Inst. Nac. Invest. Pescas: Vols. 6 Jul.-Out*.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12(1), 77. <https://doi.org/10.1186/1471-2105-12-77>

Siesa, M. E., Manenti, R., Padoa-Schioppa, E., De Bernardi, F., & Ficetola, G. F. (2011). Spatial autocorrelation and the analysis of invasion processes from distribution data: a study with the crayfish *Procambarus clarkii*. *Biological Invasions*, 13(9), 2147–2160. <https://doi.org/10.1007/s10530-011-0032-9>

Souty-Grosset, C., Holdich, D. M., Noël, P. Y., Reynolds, J. D., & Haffner, P. (Eds.). (2006). *Atlas*

*of crayfish in Europe*. Musé um national d'Histoire naturelle.

- Strayer, D. L. (2010). Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. *Freshwater Biology*, 55(SUPPL. 1), 152–174. <https://doi.org/10.1111/j.1365-2427.2009.02380.x>
- Titus, K., Mosher, J. A., & Williams, B. K. (1984). Chance-corrected Classification for Use in Discriminant Analysis: Ecological Applications. *American Midland Naturalist*, 111(1), 1. <https://doi.org/10.2307/2425535>
- Václavík, T., Kupfer, J. A., & Meentemeyer, R. K. (2012). Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography*, 39(1), 42–55. <https://doi.org/10.1111/j.1365-2699.2011.02589.x>
- Václavík, T., & Meentemeyer, R. K. (2009). Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling*, 220(23), 3248–3258. <https://doi.org/10.1016/j.ecolmodel.2009.08.013>
- Václavík, T., & Meentemeyer, R. K. (2012). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73–83. <https://doi.org/10.1111/j.1472-4642.2011.00854.x>
- Ver Hoef, J. M., & Peterson, E. E. (2010). A Moving Average Approach for Spatial Statistical Models of Stream Networks. *Journal of the American Statistical Association*, 105(489), 6–18. <https://doi.org/10.1198/jasa.2009.ap08248>
- Ver Hoef, J. M., Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *Journal of Statistical Software*, 56(3), 1–45. <https://doi.org/10.18637/jss.v056.i03>
- Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13(4), 449–464. <https://doi.org/10.1007/s10651-006-0022-8>
- Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, 165(6), 710–718. <https://doi.org/10.1093/aje/kwk052>
- Vogt, J., Soille, P., De Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., Bódis, K., Dusart, J., Paracchini, M. L., Haastrop, P., & Bamps, C. (2007). *A pan-European River and Catchment Database*. <https://doi.org/10.2788/35907>
- Walsh, J. R., Carpenter, S. R., & Vander Zanden, M. J. (2016). Invasive species triggers a massive loss of ecosystem services through a trophic cascade. *Proceedings of the National Academy of Sciences*, 113(15), 4081–4085. <https://doi.org/10.1073/pnas.1600366113>

## 4.6 Supplementary material

Table 4.S1 - Summary of the variable selection procedure for models relating the probability of occurrence of *Procambarus clarkii* and environmental variables. We tested all combinations of variables and their interaction terms, indicating for each candidate model the variables included and the root mean square prediction error (RMSPE). Models are ranked according to their RMSPE

<b>Model rank</b>	<b>Formula</b>	<b>RMSPE</b>
1	Stream order	2.2967
2	Stream order + Slope	2.4248
3	Stream order + Elevation x Slope	2.4847
4	Stream order x Slope	2.6526
5	Stream order x Slope + Elevation	2.6903
6	Stream order + Elevation + Slope	2.6989
7	Stream order x Elevation x Slope	2.9117
8	Stream order x Elevation	3.5527
9	Stream order + Elevation	3.8612
<b>10</b>	Elevation x Slope	4.3464
<b>11</b>	Stream order x Elevation + Slope	6.4626
<b>12</b>	Elevation	10.4346
<b>13</b>	Elevation + Slope	12.7393
<b>14</b>	Slope	13.7309
<b>15</b>	Intercept	18.2796

Table 4.S2 - Summary of the variable selection procedure for models relating the probability of occurrence of *Pacifastacus leniusculus* and environmental variables. We tested all combinations of variables and their interaction terms, indicating for each candidate model the variables included and the root mean square prediction error (RMSPE). Models are ranked according to their RMSPE.

<b>Model Rank</b>	<b>Variables</b>	<b>RMSPE</b>
1	Stream order x Elevation	1.6177
2	Stream order x Elevation + Slope	2.0767
3	Elevation x Slope	2.8515
4	Stream order + Elevation x Slope	3.9787
5	Stream order + Elevation	4.1513
6	Stream order + Elevation + Slope	4.1578
7	Stream order x Elevation x Slope	4.9630
8	Slope	5.2188
9	Stream order x Slope + Elevation	5.9410
10	Stream order x Slope	9.6485
11	Stream order	15.9638
12	Elevation + Slope	21.9376
13	Intercept	39.9514
14	Stream order + Slope	486.3754
15	Elevation	1.97x10 <sup>34</sup>

Table 4.S3 - Summary of the selection of the mixture covariance model for *Procambarus clarkii*. Alternative candidate models included combinations of the Euclidean (EUC), tail-up (TU), and tail-down (TD) spatial components, which were tested using exponential, spherical, Gaussian and Cauchy functions as alternatives for EUC, and exponential, linear-with-sill, spherical and Mariah functions as alternatives for TU and TD. For each candidate model, we indicate the spatial components and functions tested, and the corresponding root mean square prediction error (RMSPE). Models are ranked according to their RMSPE.

<b>Model rank</b>	<b>Variance Components</b>	<b>RMSPE</b>
1	Linear-sill (TU) + Spherical (EUC)	2.2351
2	Linear-sill (TU) + Linear-sill (TD) + Spherical (EUC)	2.2371
3	Linear-sill (TU) + Mariah (TD) + Spherical (EUC)	2.2377
4	Linear-sill (TU) + Spherical (TD) + Spherical (EUC)	2.2377
5	Linear-sill (TU) + Exponential (TD) + Spherical (EUC)	2.2378
6	Spherical (TU) + Linear-sill (TD) + Gaussian (EUC)	2.2398
7	Spherical (TU) + Spherical (EUC)	2.2417
8	Spherical (TU) + Mariah (TD) + Spherical (EUC)	2.2418
9	Spherical (TU) + Spherical (TD) + Spherical (EUC)	2.2419
10	Spherical (TU) + Linear-sill (TD) + Spherical (EUC)	2.2419
11	Linear-sill (TU) + Mariah (TD) + Exponential (EUC)	2.2432
12	Linear-sill (TU) + Mariah (TD) + Gaussian (EUC)	2.2441
13	Linear-sill (TU) + Spherical (TD) + Gaussian (EUC)	2.2443
14	Linear-sill (TU) + Linear-sill (TD) + Gaussian (EUC)	2.2445
15	Exponential (TU) + Linear-sill (TD) + Spherical (EUC)	2.2457
16	Exponential (TU) + Spherical (EUC)	2.2459
17	Exponential (TU) + Spherical (TD) + Spherical (EUC)	2.2460
18	Exponential (TU) + Mariah (TD) + Spherical (EUC)	2.2460
19	Linear-sill (TU) + Mariah (TD) + Cauchy (EUC)	2.2471
20	Linear-sill (TU) + Spherical (TD) + Cauchy (EUC)	2.2472
21	Linear-sill (TU) + Exponential (TD) + Cauchy (EUC)	2.2473

22	Linear-sill (TU) + Linear-sill (TD) + Cauchy (EUC)	2.2473
23	Spherical (TU) + Mariah (TD) + Exponential (EUC)	2.2497
24	Linear-sill (TU) + Linear-sill (TD) + Exponential (EUC)	2.2582
25	Linear-sill (TU) + Exponential (TD) + Exponential (EUC)	2.2584
26	Spherical (TU) + Spherical (TD) + Gaussian (EUC)	2.2612
27	Spherical (TU) + Mariah (TD) + Gaussian (EUC)	2.2613
28	Mariah (TU) + Spherical (EUC)	2.2616
29	Mariah (TU) + Spherical (TD) + Spherical (EUC)	2.2616
30	Mariah (TU) + Linear-sill (TD) + Spherical (EUC)	2.2618
31	Mariah (TU) + Exponential (TD) + Spherical (EUC)	2.2618
32	Spherical (TU) + Exponential (EUC)	2.2638
33	Linear-sill (TU) + Exponential (EUC)	2.2671
34	Spherical (TU) + Linear-sill (TD) + Exponential (EUC)	2.2793
35	Spherical (TU) + Spherical (TD) + Exponential (EUC)	2.2796
36	Exponential (TU) + Linear-sill (TD) + Gaussian (EUC)	2.2967
37	Exponential (TU) + Mariah (TD) + Gaussian (EUC)	2.2968
38	Exponential (TU) + Exponential (TD) + Gaussian (EUC)	2.3003
39	Linear-sill (TU) + Exponential (TD) + Gaussian (EUC)	2.3072
40	Exponential (TU) + Linear-sill (TD) + Exponential (EUC)	2.3114
41	Exponential (TU) + Mariah (TD) + Exponential (EUC)	2.3155
42	Spherical (TU) + Exponential (TD) + Spherical (EUC)	2.3185
43	Spherical (TU) + Exponential (TD) + Exponential (EUC)	2.3243
44	Exponential (TU) + Exponential (TD) + Spherical (EUC)	2.3312
45	Spherical (TU) + Exponential (TD) + Gaussian (EUC)	2.3332
46	Mariah (TD) + Spherical (EUC)	2.3360
47	Spherical (TU) + Exponential (TD) + Cauchy (EUC)	2.3361
48	Exponential (TU) + Exponential (EUC)	2.3398



49	Exponential (TU) + Spherical (TD) + Exponential (EUC)	2.3481
50	Exponential (TU) + Exponential (TD) + Exponential (EUC)	2.3490
51	Linear-sill (TU) + Linear-sill (TD)	2.3559
52	Mariah (TU) + Spherical (TD) + Gaussian (EUC)	2.3622
53	Spherical (TU) + Linear-sill (TD)	2.3657
54	Mariah (TU) + Mariah (TD) + Cauchy (EUC)	2.3724
55	Mariah (TU) + Exponential (TD) + Cauchy (EUC)	2.3725
56	Mariah (TU) + Linear-sill (TD) + Cauchy (EUC)	2.3726
57	Mariah (TU) + Spherical (TD) + Cauchy (EUC)	2.3726
58	Mariah (TU) + Cauchy (EUC)	2.3727
59	Linear-sill (TD) + Cauchy (EUC)	2.3775
60	Cauchy (EUC)	2.3776
61	Exponential (TD) + Cauchy (EUC)	2.3777
62	Exponential (TU) + Exponential (TD) + Cauchy (EUC)	2.3785
63	Exponential (TU) + Linear-sill (TD) + Cauchy (EUC)	2.3785
64	Exponential (TU) + Spherical (TD) + Cauchy (EUC)	2.3786
65	Exponential (TU) + Mariah (TD) + Cauchy (EUC)	2.3787
66	Exponential (TU) + Cauchy (EUC)	2.3787
67	Spherical (TD) + Cauchy (EUC)	2.3807
68	Mariah (TU) + Mariah (TD) + Exponential (EUC)	2.3833
69	Mariah (TU) + Exponential (TD) + Exponential (EUC)	2.3834
70	Mariah (TU) + Spherical (TD) + Exponential (EUC)	2.3834
71	Mariah (TU) + Exponential (EUC)	2.3835
72	Mariah (TU) + Linear-sill (TD) + Exponential (EUC)	2.3836
73	Exponential (TU) + Linear-sill (TD)	2.3869
74	Mariah (TD) + Cauchy (EUC)	2.3972
75	Spherical (TU) + Mariah (TD) + Cauchy (EUC)	2.3988

76	Spherical (TU) + Linear-sill (TD) + Cauchy (EUC)	2.3988
77	Spherical (TU) + Spherical (TD) + Cauchy (EUC)	2.3989
78	Spherical (TU) + Cauchy (EUC)	2.3991
79	Linear-sill (TU) + Cauchy (EUC)	2.3996
80	Exponential (EUC)	2.4059
81	Exponential (TD) + Exponential (EUC)	2.4062
82	Linear-sill (TD) + Exponential (EUC)	2.4062
83	Spherical (TD) + Exponential (EUC)	2.4063
84	Linear-sill (TD) + Spherical (EUC)	2.4162
85	Mariah (TD) + Exponential (EUC)	2.4165
86	Spherical (TD) + Gaussian (EUC)	2.4178
87	Linear-sill (TU) + Spherical (TD) + Exponential (EUC)	2.4205
88	Mariah (TU) + Linear-sill (TD)	2.4491
89	Linear-sill (TD) + Gaussian (EUC)	2.4550
90	Spherical (EUC)	2.5191
91	Spherical (TD) + Spherical (EUC)	2.5191
92	Exponential (TD) + Spherical (EUC)	2.5192
93	Mariah (TU) + Exponential (TD) + Gaussian (EUC)	2.5457
94	Mariah (TU) + Gaussian (EUC)	2.5460
95	Mariah (TU) + Mariah (TD) + Gaussian (EUC)	2.5460
96	Mariah (TU) + Linear-sill (TD) + Gaussian (EUC)	2.5460
97	Gaussian (EUC)	2.5594
98	Exponential (TD) + Gaussian (EUC)	2.5597
99	Exponential (TU) + Spherical (TD) + Gaussian (EUC)	2.5610
100	Exponential (TU) + Gaussian (EUC)	2.5614
101	Mariah (TD) + Gaussian (EUC)	2.5617
102	Mariah (TU) + Mariah (TD) + Spherical (EUC)	2.5750

<b>103</b>	Linear-sill (TU) + Gaussian (EUC)	2.5804
<b>104</b>	Spherical (TU) + Gaussian (EUC)	2.5805
<b>105</b>	Spherical (TU) + Spherical (TD)	2.5944
<b>106</b>	Exponential (TU) + Spherical (TD)	2.5962
<b>107</b>	Linear-sill (TU) + Spherical (TD)	2.6050
<b>108</b>	Mariah (TU) + Exponential (TD)	2.7734
<b>109</b>	Linear-sill (TD)	2.7982
<b>110</b>	Exponential (TU) + Exponential (TD)	2.8062
<b>111</b>	Spherical (TD)	2.8989
<b>112</b>	Linear-sill (TU) + Exponential (TD)	2.9410
<b>113</b>	Exponential (TD)	2.9629
<b>114</b>	Linear-sill (TU)	3.0039
<b>115</b>	Linear-sill (TU) + Mariah (TD)	3.0040
<b>116</b>	Mariah (TU) + Mariah (TD)	3.0174
<b>117</b>	Mariah (TU) + Spherical (TD)	3.0184
<b>118</b>	Mariah (TD)	3.0237
<b>119</b>	Spherical (TU)	3.0285
<b>120</b>	Spherical (TU) + Exponential (TD)	3.0285
<b>121</b>	Spherical (TU) + Mariah (TD)	3.0285
<b>122</b>	Nugget	3.0813
<b>123</b>	Exponential (TU)	3.0929
<b>124</b>	Exponential (TU) + Mariah (TD)	3.0930
<b>125</b>	Mariah (TU)	3.2350

---

Table 4.S4 - Summary of the selection of the mixture covariance model for *Pacifastacus leniusculus*. Alternative candidate models included combinations of the Euclidean (EUC), tail-up (TU), and tail-down (TD) spatial components, which were tested using exponential, spherical, Gaussian and Cauchy functions as alternatives for EUC, and exponential, linear-with-sill, spherical and Mariah functions as alternatives for TU and TD. For each candidate model, we indicate the spatial components and functions tested, and the corresponding root mean square prediction error (RMSPE). Models are ordered according to their RMSPE. <sup>a</sup> The model with the Cauchy EUC component alone did not converge.

Model rank	Variance Components	RMSPE
1	Exponential (TD) + Gaussian (EUC)	1.4116
2	Linear-sill (TU) + Spherical (TD) + Cauchy (EUC)	1.4238
3	Spherical (TD) + Cauchy (EUC)	1.4346
4	Linear-sill (TU) + Linear-sill (TD) + Cauchy (EUC)	1.5283
5	Mariah (TU) + Linear-sill (TD) + Cauchy (EUC)	1.5285
6	Exponential (TD) + Spherical (EUC)	1.5329
7	Mariah (TU) + Exponential (TD) + Gaussian (EUC)	1.5332
8	Linear-sill (TD) + Spherical (EUC)	1.5338
9	Linear-sill (TU) + Exponential (TD) + Gaussian (EUC)	1.5442
10	Linear-sill (TD) + Gaussian (EUC)	1.5499
11	Mariah (TU) + Linear-sill (TD) + Gaussian (EUC)	1.5622
12	Spherical (TD) + Gaussian (EUC)	1.5628
13	Exponential (TU) + Exponential (TD) + Cauchy (EUC)	1.5800
14	Exponential (TD) + Cauchy (EUC)	1.5804
15	Mariah (TU) + Spherical (TD) + Gaussian (EUC)	1.5938
16	Mariah (TU) + Gaussian (EUC)	1.5947
17	Linear-sill (TU) + Spherical (TD) + Gaussian (EUC)	1.5992
18	Mariah (TU) + Mariah (TD) + Gaussian (EUC)	1.6085
19	Spherical (TU) + Gaussian (EUC)	1.6132
20	Spherical (TU) + Exponential (TD) + Gaussian (EUC)	1.6136
21	Linear-sill (TU) + Exponential (TD) + Exponential (EUC)	1.6166

22	Exponential (TU) + Linear-sill (TD) + Gaussian (EUC)	1.6177
23	Exponential (TU) + Gaussian (EUC)	1.6184
24	Linear-sill (TD) + Cauchy (EUC)	1.6189
25	Exponential (TU) + Mariah (TD) + Gaussian (EUC)	1.6195
26	Exponential (TU) + Exponential (TD) + Exponential (EUC)	1.6222
27	Exponential (TU) + Exponential (TD) + Gaussian (EUC)	1.6223
28	Linear-sill (TD) + Exponential (EUC)	1.6224
29	Linear-sill (TU) + Linear-sill (TD) + Gaussian (EUC)	1.6225
30	Exponential (EUC)	1.6234
31	Spherical (TD) + Exponential (EUC)	1.6236
32	Linear-sill (TU) + Mariah (TD) + Gaussian (EUC)	1.6238
33	Linear-sill (TU) + Gaussian (EUC)	1.6274
34	Spherical (TU) + Spherical (TD) + Gaussian (EUC)	1.6347
35	Mariah (TU) + Cauchy (EUC)	1.6360
36	Mariah (TD) + Spherical (EUC)	1.6369
37	Linear-sill (TU) + Exponential (TD) + Cauchy (EUC)	1.6402
38	Mariah (TU) + Mariah (TD) + Cauchy (EUC)	1.6457
39	Exponential (TU) + Spherical (TD) + Cauchy (EUC)	1.6458
40	Spherical (TU) + Linear-sill (TD) + Gaussian (EUC)	1.6493
41	Exponential (TU) + Mariah (TD) + Cauchy (EUC)	1.6557
42	Linear-sill (TU) + Cauchy (EUC)	1.6578
43	Exponential (TU) + Linear-sill (TD) + Cauchy (EUC)	1.6585
44	Spherical (TU) + Spherical (TD) + Cauchy (EUC)	1.6627
45	Spherical (TU) + Mariah (TD) + Cauchy (EUC)	1.6630
46	Spherical (TU) + Exponential (TD) + Cauchy (EUC)	1.6658
47	Spherical (TU) + Linear-sill (TD) + Cauchy (EUC)	1.6775
48	Spherical (TU) + Mariah (TD) + Gaussian (EUC)	1.7084

49	Exponential (TU) + Cauchy (EUC)	1.7219
50	Spherical (TU) + Spherical (TD) + Spherical (EUC)	1.7299
51	Mariah (TD) + Exponential (EUC)	1.7301
52	Spherical (TD) + Spherical (EUC)	1.7371
53	Mariah (TU) + Spherical (EUC)	1.7565
54	Mariah (TU) + Exponential (TD) + Spherical (EUC)	1.7619
55	Mariah (TU) + Mariah (TD) + Spherical (EUC)	1.7670
56	Linear-sill (TU) + Mariah (TD) + Cauchy (EUC)	1.7684
57	Linear-sill (TU) + Linear-sill (TD) + Spherical (EUC)	1.7688
58	Mariah (TU) + Spherical (TD) + Spherical (EUC)	1.7713
59	Mariah (TU) + Linear-sill (TD) + Spherical (EUC)	1.7811
60	Exponential (TU) + Linear-sill (TD) + Spherical (EUC)	1.8049
61	Exponential (TU) + Mariah (TD) + Spherical (EUC)	1.8131
62	Exponential (TU) + Spherical (EUC)	1.8160
63	Exponential (TU) + Spherical (TD) + Spherical (EUC)	1.8176
64	Exponential (TD) + Exponential (EUC)	1.8212
65	Mariah (TU) + Linear-sill (TD)	1.8258
66	Exponential (TU) + Exponential (TD) + Spherical (EUC)	1.8336
67	Spherical (TU) + Exponential (TD) + Exponential (EUC)	1.8348
68	Linear-sill (TU) + Mariah (TD) + Spherical (EUC)	1.8361
69	Spherical (TU) + Mariah (TD) + Spherical (EUC)	1.8368
70	Linear-sill (TU) + Spherical (TD) + Spherical (EUC)	1.8526
71	Mariah (TU) + Exponential (EUC)	1.8604
72	Exponential (TU) + Linear-sill (TD) + Exponential (EUC)	1.8609
73	Mariah (TU) + Linear-sill (TD) + Exponential (EUC)	1.8748
74	Mariah (TU) + Exponential (TD) + Exponential (EUC)	1.8749
75	Spherical (TU) + Exponential (TD) + Spherical (EUC)	1.8758

76	Mariah (TU) + Mariah (TD) + Exponential (EUC)	1.8778
77	Spherical (TU) + Linear-sill (TD) + Spherical (EUC)	1.8820
78	Spherical (TU) + Spherical (EUC)	1.8861
79	Exponential (TU) + Spherical (TD) + Exponential (EUC)	1.9200
80	Spherical (TU) + Linear-sill (TD) + Exponential (EUC)	1.9377
81	Mariah (TU) + Spherical (TD) + Exponential (EUC)	1.9414
82	Linear-sill (TU) + Mariah (TD) + Exponential (EUC)	1.9491
83	Linear-sill (TU) + Exponential (EUC)	1.9608
84	Exponential (TU) + Mariah (TD) + Exponential (EUC)	1.9631
85	Linear-sill (TU) + Exponential (TD) + Spherical (EUC)	1.9922
86	Linear-sill (TU) + Spherical (TD) + Exponential (EUC)	1.9998
87	Spherical (TU) + Spherical (TD) + Exponential (EUC)	2.0028
88	Mariah (TU) + Spherical (TD) + Cauchy (EUC)	2.0503
89	Mariah (TU) + Spherical (TD)	2.0768
90	Mariah (TD) + Gaussian (EUC)	2.0972
91	Spherical (TU) + Mariah (TD) + Exponential (EUC)	2.1856
92	Linear-sill (TU) + Spherical (TD)	2.2904
93	Mariah (TD)	2.3243
94	Mariah (TD) + Cauchy (EUC)	2.3259
95	Gaussian (EUC)	2.8031
96	Linear-sill (TU) + Exponential (TD)	2.8065
97	Spherical (TU) + Exponential (EUC)	2.8244
98	Exponential (TD)	2.8246
99	Spherical (EUC)	2.8828
100	Spherical (TD)	2.9130
101	Spherical (TU) + Cauchy (EUC)	2.9604
102	Linear-sill (TD)	2.9637

<b>103</b>	Mariah (TU) + Exponential (TD) + Cauchy (EUC)	2.9740
<b>104</b>	Mariah (TU) + Mariah (TD)	2.9981
<b>105</b>	Nugget	3.0054
<b>106</b>	Exponential (TU) + Spherical (TD) + Gaussian (EUC)	3.0361
<b>107</b>	Exponential (TU) + Exponential (TD)	3.0401
<b>108</b>	Exponential (TU) + Exponential (EUC)	3.0404
<b>109</b>	Spherical (TU) + Linear-sill (TD)	3.0614
<b>110</b>	Linear-sill (TU) + Spherical (EUC)	3.0700
<b>111</b>	Linear-sill (TU) + Linear-sill (TD) + Exponential (EUC)	3.0722
<b>112</b>	Mariah (TU)	3.0785
<b>113</b>	Mariah (TU) + Exponential (TD)	3.0786
<b>114</b>	Exponential (TU) + Linear-sill (TD)	3.0887
<b>115</b>	Exponential (TU) + Mariah (TD)	3.0903
<b>116</b>	Exponential (TU)	3.0904
<b>117</b>	Exponential (TU) + Spherical (TD)	3.0906
<b>118</b>	Linear-sill (TU) + Linear-sill (TD)	3.1012
<b>119</b>	Spherical (TU)	3.1017
<b>120</b>	Spherical (TU) + Spherical (TD)	3.1018
<b>121</b>	Spherical (TU) + Mariah (TD)	3.1018
<b>122</b>	Linear-sill (TU) + Mariah (TD)	3.1018
<b>123</b>	Spherical (TU) + Exponential (TD)	3.1019
<b>124</b>	Linear-sill (TU)	3.1044

---



## Chapter 5 – Combining geostatistical and biotic interactions modelling to predict amphibian refuges under crayfish invasion across dendritic stream networks





# Chapter 5 – Combining geostatistical and biotic interactions modelling to predict amphibian refuges under crayfish invasion across dendritic stream networks

Mário Mota-Ferreira, Pedro Beja

This chapter was published in the journal *Diversity and Distributions* and should be cite as:

**Mota-Ferreira, M., & Beja, P. (2020).** Combining geostatistical and biotic interaction modeling to predict amphibian refuges under crayfish invasion across dendritic stream networks. *Diversity and Distributions*, 26(6), 699–714. <https://doi.org/10.1111/ddi.13047>

Keywords:

Alien invasive species, biological invasions, ecological refuges, frog, geostatistics, newt, species distribution models, stream ecology, toad

## Abstract

**Aim:** Biological invasions are pervasive in freshwater ecosystems, often causing native species to contract into areas that remain largely free from invasive species impacts. Predicting the location of such ecological refuges is challenging, because they are shaped by the habitat requirements of native and invasive species, their biotic interactions, and the spatial and temporal invasion patterns. Here we investigated the spatial distribution and environmental drivers of refuges from invasion in river systems, by considering biotic interactions in geostatistical models accounting for stream network topology. We focused on Mediterranean amphibians negatively impacted by the invasive crayfishes *Procambarus clarkii* and *Pacifastacus leniusculus*.

**Location:** River Sabor, NE Portugal

**Methods:** We surveyed amphibians at 168 200-m stream stretches in 2015. Geostatistical models were used to relate the probabilities of occurrence of each species to environmental and biotic variables, while controlling for linear (Euclidean) and hydrologic spatial dependencies. Biotic interactions were specified using crayfish probabilities of occurrence extracted from previously developed geostatistical models. Models were used to map the distribution of potential refuges for the most common amphibian species, under current conditions and future scenarios of crayfish expansion.

**Results:** Geostatistical models were produced for eight out of 10 species detected, of which five species were associated with lower stream orders and only one species with higher stream orders. Six species showed negative responses to one or both crayfish species, even after accounting for environmental effects and spatial dependencies. Most amphibian species were found to retain large expanses of potential habitat in stream headwaters, but current refuges will likely contract under plausible scenarios of crayfish expansion.

**Main conclusions:** Incorporating biotic interactions in geostatistical modelling provides a practical and relatively simple approach to predict present and future distributions of refuges from biological invasion in stream networks. Using this approach, our study shows that stream headwaters are key amphibian refuges under invasion by alien crayfish.

## 5.1 Introduction

Biological invasions are pervasive in freshwater ecosystems, where they are major drivers of native species declines (Strayer, 2010; Walsh et al., 2016). Addressing this threat is challenging, because once fully established the control of invasive species is often nearly impossible, which limits the management options to protect native species. In some circumstances, the impacts of biological invasions may be partly offset by the presence of ecological refuges, which are habitats where a species can retreat, persist in for up to a few decades, and eventually expand from under changing environmental conditions (Davis et al., 2013). Such refuges correspond to freshwater habitats unsuitable for invasive species, or areas where their spread is prevented by physical barriers such as waterfalls or culverts (Kerby et al., 2005; Rahel, 2013). Refuges may thus allow the persistence of at least some remnant populations of native species (e.g., Chapman et al., 1996; Grabowski et al., 2009; Habit et al., 2010; Radinger et al., 2019), making it a priority to understand where, why and how refuges can contribute to species conservation under biological invasion.

Species distribution models (SDM) incorporating biotic interactions provide a simple framework to quantify how one or more species influence the distribution of others (e.g., Wisz et al., 2013),

making them useful to predict the location and drivers of refuges from invasions. A straightforward approach is to take the distribution patterns of invasive species together with abiotic variables to model the occurrence of native species, and then use the ensuing models to predict the distribution of refuges under current and future invasion scenarios (Araújo & Luoto, 2007; Wisz et al., 2013). One problem is that SDMs for geographical range prediction assume equilibrium between species distribution and the environment, which is unwarranted when modelling range contractions by native species in face of biological invasion (De Marco et al., 2008; Elith et al., 2010; Václavík et al., 2012). A native species may occur in areas that will later become unsuitable due to the expansion of an invasive species, or it may eventually be able to coexist only temporarily with the invasive species due to time-lags in negative impacts (Crooks, 2005). In either case, SDMs built using snapshots of species distributions may overestimate the extent of refuges, eventually misdirecting conservation efforts towards areas where native species persistence is unlikely.

Incorporating predictors describing spatial autocorrelations in invasive species occurrences to account for unmeasured dispersal and colonization processes may help mitigating, albeit not solving, problems associated with non-equilibrium conditions in SDMs (De Marco et al., 2008; Filipe et al., 2017; Václavík et al., 2012). For alien species invading river ecosystems, SDMs can be improved using geostatistical models accounting for spatial dependencies in physical and ecological processes across stream networks (Filipe et al., 2017; Lois et al., 2015; Lois & Cowley, 2017). These models are similar to conventional mixed models, with species occurrence modelled in relation to environmental variables using a logistic function, and spatial autocorrelation considered in the random errors (Peterson et al., 2013; Peterson & Ver Hoef, 2010; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). The latter are specified as a mixture of covariance functions representing the strength of influence between sites as a function of their (i) straight-line (Euclidean) distances calculated overland; (ii) hydrologic distances (i.e., distances along the waterlines) representing flow-connected relations (tail-up models); and (iii) hydrologic distances irrespective of flow connection (tail-down models) (Ver Hoef & Peterson, 2010). This approach can easily incorporate biotic interactions by including predictors describing the occurrence or abundance of potentially interacting species in the fixed component (Lois et al., 2015; Lois & Cowley, 2017). Another possibility is to develop a geostatistical model for the invasive species itself, and then use the fitted response (i.e., the probability of occurrence) in the native species model. This should be useful for predicting the location of refuges, as it would consider not only the current distribution of the invasive species, but also suitable areas that will eventually be colonised during the expansion process.

This study investigates the location and environmental drivers of refuges in dendritic stream networks, combining biotic interactions and geostatistical modelling to predict their spatial distribution under current and future scenarios of invasive species expansion. We focused on interactions between amphibians and the exotic crayfish *Procambarus clarkii* and *Pacifastacus leniusculus* in the Iberian Peninsula, where there are no native crayfish (Clavero et al., 2016). These crayfishes are among the most widespread and damaging aquatic invaders (Lodge et al., 2012; Twardochleb et al., 2013), which have expanded widely in Iberia since the 1970s due to multiple introductions for commercial purposes and subsequent natural dispersal (Bernardo et al., 2011; Clavero, 2016; Gutiérrez-Yurrita et al., 1999). Invasive crayfish predate on amphibian eggs and larvae (Axelsson et al., 1997; Cruz, Pascoal, et al., 2006; Cruz & Rebelo, 2005; Gamradt & Kats, 1996), and seem to have strong negative impacts on native amphibian populations in Iberian waters (Cruz et al., 2008; Cruz, Pascoal, et al., 2006; Cruz, Rebelo, et al., 2006) and elsewhere (Ficetola et al., 2011). The main amphibian refuges are probably temporary ponds far from permanent waters (Beja & Alcazar, 2003; Cruz, Rebelo, et al., 2006; Ferreira & Beja, 2013), where crayfish cannot persist (Cruz & Rebelo, 2007). Refuges may also exist in small and intermittent Mediterranean streams, which often hold rich amphibian communities (de Vries & Marco, 2017) and where crayfishes are usually absent (e.g., Cruz & Rebelo, 2007; Filipe et al., 2017; Gil-Sánchez & Alba-Tercedor, 2002). To investigate amphibian refuges from crayfish invasion, we (i) made a detailed survey of amphibian occurrence in a Mediterranean watershed; (ii) developed geostatistical models relating the occurrence of each amphibian species to environmental variables, the probabilities of crayfish occurrence (Filipe et al., 2017) and spatial dependencies; and predicted the spatial distribution of refuges under (iii) current and (iv) future scenarios of crayfish expansion. Our study can help improve conservation strategies for amphibians negatively affected by crayfish invasions, and more generally it provides a framework for identifying refuges from biological invasions in dendritic stream networks.

## 5.2 Methods

### 5.2.1 Study area

The study was conducted in the river Sabor watershed (NE Portugal; N41°090–42°000, W7°150–6°150; Figure 5.S1 in the supplementary figures). Human population is low (8.5 to 28.7 inhabitants/km<sup>2</sup>; <https://www.pordata.pt/Municipios>) following a process of land abandonment since the 1970s (Azevedo et al., 2011). Land cover is dominated by extensive agriculture and pastureland, forest plantations, and natural vegetation (Caetano et al., 2018;

Hoelzer, 2003). A large proportion of the watershed is included in the Natura 2000 network (Costa et al., 2007). The watershed covers a wide range of elevations (100–1500 m above sea level), total annual precipitation (443–1163 mm), and mean annual temperature (6.9–15.6 C°). The climate is Mediterranean, with precipitation concentrated in October–March and virtually none in June–August. Most small streams dry out or become reduced to a series of disconnected pools during the dry months, though the main watercourse and the largest tributaries are permanent (Ferreira et al., 2016). Two large hydroelectric reservoirs were built and filled just before this study, but otherwise the river is largely free-flowing. Crayfish were first reported in the Sabor watershed in the 1990s, with *P. clarkii* probably introduced by local people, while *P. leniusculus* was introduced in 1994 by Spanish authorities (Bernardo et al., 2011). *P. clarkii* is far more widespread than *P. leniusculus*, but both species seem to still be spreading, possibly through natural dispersal along the stream network (Anastácio et al., 2015; Bernardo et al., 2011).

### 5.2.2 Study design

The study was designed to obtain a comprehensive snapshot of stream-dwelling amphibian distributions, considering both species such as Iberian frog (*Rana iberica*) and Iberian green frog (*Pelophylax perezii*) that occupy streams during their entire life cycle, and species such as fire salamander (*Salamandra salamandra*) and midwife toads (*Alytes obstetricans* and *A. cisternasii*) that have both terrestrial and aquatic phases, occupying streams mostly during the breeding and larval development periods. To adequately cover all species, surveys encompassed the main environmental gradients represented in the watershed, and they were carried out monthly during one year to account for differences in activity peaks and breeding phenology across species (e.g., Diaz-Paniagua, 1992; Ferreira & Beja, 2013). Yet, because it was logistically unfeasible to sample a sufficiently large number of sites each month, the sampling effort was distributed over the year, with different sites sampled in different months. Considering these constraints, we initially selected >200 potential sampling sites, based on previous studies (Ferreira et al., 2016; Filipe et al., 2017; Quaglietta et al., 2018) and new field surveys. Sites were constrained to cover all Strahler stream orders and to be at >1km from each other. All sites were in wadable stream reaches (water depth < 1.20) to facilitate amphibian surveys. From the overall set of potential sampling sites, we selected each month a subset of 30 sites, following a stratified random procedure to guarantee that comparable environmental conditions were covered each month, and thus avoiding space x time interactions. To do this, we divided the Sabor watershed in three sub-basins (Figure 5.S1), and randomly selected each month two sites of each Strahler stream order represented in

each sub-basin. A total of 168 sites were surveyed (Figure 5.S1), with each site visited at most twice, except sites in higher orders that were visited more often because they were relatively scarce in the watershed. The river network, sub-basins and stream orders were obtained from CCM 2.1 (Catchment Characterization and Modelling database; Vogt et al., 2007) (Vogt et al., 2007). A detailed workflow of the procedures used to analyse the data is provided in the

Supplementary Methods: Workflow detailing the methodological procedures in the supplementary materials.

### 5.2.3 Field sampling

Sampling was carried out monthly in 2015, except in May due to logistic constraints. At each sampling site and date, a 200-m stream reach was thoroughly surveyed for amphibians, including both adults and larvae. The survey was conducted by two observers walking slowly along the banks or wading in shallow water along the stream. Observers used dip nets to collect aquatic larvae and adults, and they systematically searched the stream banks for terrestrial adults, using torches where necessary to survey cavities and other shaded areas. All amphibians found were identified to species level *in situ* and released thereafter. In a few cases, small unidentified larvae were preserved and identified in the laboratory. During surveys, crayfish were also recorded and identified to species, and all individuals collected were eliminated following guidelines established by the Portuguese biodiversity conservation agency.

### 5.2.4 Environmental and spatial variables

Sampling sites were characterised using variables potentially affecting the distribution of stream-dwelling amphibians that could be extracted from digital maps (e.g., Cruz & Rebelo, 2007; de Vries & Marco, 2017), making it possible to predict species distributions across the entire watershed in relation to potential expansions of crayfish ranges. Each site was characterised using four environmental variables (elevation [Alt], total annual precipitation [Prec], Strahler's stream order [SO] and the probability of water presence during the dry season [Water]) and two variables describing the potential for biotic interactions between amphibians and either *P. clarkii* [Pclar] or *P. leniusculus* [Plen] (Cruz, Rebelo, et al., 2006; de Vries & Marco, 2017). We also included a multiplicative interaction between elevation and Strahler's stream order (SOxAlt), which was used to distinguish between small streams in the lowlands and small streams in mountain areas. Initially we also considered other climatic variables, but they were discarded because of strong correlations with precipitation and/or elevation.



Although features of the surrounding landscape are known to affect stream-dwelling amphibians (Ficetola et al., 2011; Riley et al., 2005), these were not considered because preliminary analysis showed very minor effects of land cover variables in our study area, possibly due to the dominance of natural vegetation and low-intensity land uses that are generally suitable for amphibians. Furthermore, adding land cover variables often caused model instability and convergence problems, possibly due to redundancies with other environmental variables already included in the models.

All variables were computed in a Geographic Information System (GIS) using ArcGis (ESRI, 2016). Elevation was taken from a DEM built from 1:25,000 topographic maps as in Ferreira et al. (2016). Total annual precipitation was extracted from WorldClim 2 with a 30' ( $\approx$  1km) resolution (Fick & Hijmans, 2017). Strahler's stream order was used as a proxy for habitat size and heterogeneity (Ferreira et al., 2016; Hughes et al., 2011), and it was extracted from CCM 2.1, which is based on a 100-m resolution digital elevation model (DEM) (Vogt et al., 2007). The probability of water presence in the dry period was used because many Mediterranean amphibian species are associated with temporary water bodies (Beja & Alcazar, 2003; de Vries & Marco, 2017; Ferreira & Beja, 2013), and it was extracted from a model developed in a previous study (Ferreira et al., 2016). Variables describing biotic interactions were specified considering the probability of occurrence of either *P. clarkii* or *P. leniusculus*, extracted from a previous geostatistical modelling of crayfish distribution in the Sabor watershed (Filipe et al., 2017). These models were built using electrofishing data collected on 167 sites in summer 2012, and they showed that crayfish distributions were mainly associated with stream order, elevation and spatial dependencies across the stream network (Filipe et al., 2017). The models had reasonable predictive accuracy, for both *P. clarkii* (AUC=0.963) and *P. leniusculus* (AUC=0.823). Probabilities derived from distribution models were used instead of the actual crayfish presences/absences recorded at sampling sites, in order to project the distribution of each species across the entire watershed, as well as to build scenarios of future crayfish expansion. All variables were standardized to have a mean of zero and a standard deviation of one to improve the interpretability of model coefficients (Schielzeth, 2010), and we screened for outliers and influential points that might bias coefficient estimates.

Spatial data necessary to account for spatial autocorrelation (see below) were obtained in a GIS using the Sabor watershed stream network extracted from CCM2.1 (Vogt et al., 2007), and the layer of survey sites. Estimates included the Euclidean and hydrologic distances (total and downstream hydrologic distances) between every pair of sampling sites (Peterson & Ver Hoef, 2010). To deal with confluences in tail-up models (see below), we also estimated watershed areas to weight the relative influence of the branching upstream segments (e.g.

Peterson & Ver Hoef, 2010). Spatial estimates were made using the Spatial Tools for the Analysis of River Systems (STARS) toolbox version 2.0.0 (Peterson & Ver Hoef, 2014) for ArcGIS 10.5 (ESRI, 2016).

### 5.2.5 Distribution modelling

Distribution models (SDM) were developed considering for each species only the occurrence data from the sampling months encompassing its aquatic phase, and thus the periods when the species is detectable during stream surveys (Table 5.S1 in the supplementary tables). For instance, we considered data from all sampling months for *R. iberica* and *P. perezii*, because they are largely aquatic and strongly attached to stream habitats all year round, while we only considered data from April-November and October-April for *A. obstetricans* and *S. Salamandra*, respectively, corresponding to their aquatic phases. This approach aimed at avoiding false negatives caused by species not using potentially suitable stream habitats during the terrestrial phase.

For each species, we developed three logistic models relating its presence/absence to either (i) only environmental predictors, (ii) only biotic predictors or (iii) environmental + biotic predictors, and a (iv) geostatistical distribution model (Peterson et al., 2013). The logistic models were used to evaluate how considering biotic interactions affected species perceived responses to environmental variables, and as preliminary steps in geostatistical model building. The geostatistical model for each species included a fixed component corresponding to a logistic function linking its probability of occurrence to environmental and biotic predictors, and random components accounting for spatial dependencies in the stream network (Peterson et al., 2013; Peterson & Ver Hoef, 2010; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). As random components we considered the Euclidean model, assuming that spatial dependencies among sites can occur overland, and the tail-up and tail-down covariance models, assuming that spatial dependencies can also occur along the hydrological network independently of flow and/or only between flow connected sites, respectively (Peterson & Ver Hoef, 2010).

To build the logistic models for each set of predictors and species, we considered all combinations of predictors of each set and retained for inference the best subset model minimising AIC (Murtaugh, 2009). Autocorrelation in model residuals was visualised using Torgegrams, depicting how semivariance in the residuals of the best logistic models between pairs of sampling sites changed in relation to their hydrologic distances (Peterson et al., 2013). In Torgegrams, increasing semivariance reflects declining spatial dependency between points. The fixed component of the geostatistical model for each species was then built considering

the predictors included in the best environmental + biotic logistic model. The structure of the random components was assessed using the residuals of the environmental + biotic model following Quaglietta et al. (2018), by testing all combinations of alternative functions available in the R package 'SSN' (Ver Hoef et al., 2014) for the Euclidean and the hydrologic autocovariance functions, and retaining the model structure minimising AIC. We then combined the variables selected for the fixed component with the best spatial structure to build the final model for each species. In all logistic models and in the fixed component of the geostatistical model, we considered significance level for individual predictors at  $P < 0.10$ , to reduce the likelihood of type II errors and thus the probability of missing true negative effects of crayfish invasion.

The crayfish models previously developed by Filipe et al. (2017) were validated with presence/absence data from the 2015 survey, using the area under the receiver operating curve (AUC) (Allouche et al., 2006). Discrimination ability of amphibian models was estimated using predictions obtained by the "leave-one-out" cross validation method, considering the overall prediction success, the AUC, Cohen's kappa and the true skill statistics (TSS) (Allouche et al., 2006; Václavík et al., 2012). Prediction success was estimated using prevalence as the threshold for predicted presences (Liu et al., 2005). All analyses were carried out in R (R Core Team, 2017), using MuMIn (Barton, 2016), 'SSN' (Ver Hoef et al., 2014) and 'modEVA' (Barbosa et al., 2016) packages.

### 5.2.6 Distribution mapping under current and future scenarios

To map the predicted distribution of each amphibian species under current conditions, we projected the distribution models on the stream network of the entire Sabor watershed. First, we divided the stream network into segments of a maximum length of 1000 meters using ArcGIS desktop (ESRI, 2016), and we extracted the value of environmental variables from the centroid of each segment. We then predicted the probability of each species occurring in the segment using universal kriging within the 'SSN' package (Ver Hoef et al., 2014). We considered segments occupied if the predicted occupancy probability was above the species prevalence threshold (Liu et al., 2005).

To simulate how crayfish expansion might affect amphibians, we changed the value of variables describing biotic interactions assuming a two, three and five-fold increase of the relative risk ( $R_r$ ) of each crayfish species occurring at each site, using as baseline the predictions from the geostatistical models of Filipe et al. (2017). The relative risk was defined as the odds ratio of the probabilities of crayfish occurrence under future and current conditions, where odds are the ratio of the probability of occurrence and the probability of absence. The

probability of occurrence at each site under each scenario of future crayfish expansion was then computed using the expression,

$$\frac{\frac{p'_i}{(1-p'_i)}}{\frac{p_i}{(1-p_i)}} = Rr \Leftrightarrow p'_i = \frac{Rr \cdot p_i}{1 + p_i(Rr - 1)}, \text{ for } Rr = \{2,3,5\}$$

Equation 5.1

where  $p_i$  and  $p'_i$  are the probabilities of crayfish occurrence at present and in the future at site  $i$ . We considered 16 different invasion scenarios, assuming changes in the distribution of each crayfish species at a time, and both crayfish species simultaneously. These scenarios of crayfish expansion were built considering empirical observations showing that both species are still expanding in the watershed (Bernardo et al., 2011), and assume that populations will expand from the areas currently occupied and will progressively colonise streams with habitat conditions most suitable for each species based on Filipe et al. (2017). Although this is a simplistic model, it can still provide approximate indications on potential amphibian refuges under crayfish expansion.

Future distributions of each amphibian species were predicted using either the non-spatial environmental + biotic logistic model or the spatial geostatistical model, which reflect different assumptions on range change processes (Record et al., 2013). The non-spatial model assumes that amphibian distributions will change along with changes in crayfish occurrence, irrespective of amphibian current distributions. However, spatial structure is still implicit in predictions, because probabilities of crayfish occurrence across the dendritic stream network were themselves predicted using geostatistical models (Filipe et al., 2017). In the geostatistical model, spatial random effects act to draw the projected distributions back toward the observed distribution used to calibrate the model (Record et al., 2013). Therefore, the current and future distributions will be similar, unless there are strong negative effects of biotic interactions. Prediction of future distributions were only made for amphibian species showing significant negative effects of crayfish occurrence. We did not consider climate change effects due to uncertainties regarding how climate will change in our relatively small area and how crayfish and amphibians will respond to such changes, though this should be the subject of further research due to potential interactions between climate change and biological invasion (Hulme, 2017).

### 5.3 Results

We detected a total of 10 amphibian species, the most widespread (>20% of sites) of which were *P. perezi* (69%), *S. salamandra* (28%), and *R. iberica* (26%) (Table 5.S1). The frogs *Discoglossus galganoi* and *Hyla molleri* were excluded from further analysis because they occurred at just one and two sites, respectively. *P. clarkii* and *P. leniusculus* were detected at 28% and 22% of sites surveyed for amphibians, with the models of Filipe et al. (2017) successfully predicting the presence/absence of each species (*P. clarkii*: AUC= 0.96; *P. leniusculus*; AUC = 0.92; Figure 5.S2).

Models including only environmental effects showed significant negative effects of stream order on the probability of occurrence of *A. cisternasii*, *R. iberica*, *Lissotriton boscai*, *S. salamandra* and *Triturus marmoratus*, and positive effects on *P. perezi* (Table 5.1). Altitude was negatively related to *A. cisternasii*, and positively so with *A. obstetricans* and *R. iberica*. Significant interactions between stream order and altitude were found for *A. cisternasii* and *R. iberica*, indicating in both cases that negative effects of stream order were weaker at higher elevations. Precipitation was negatively related to *A. cisternasii*, and positively so with *A. obstetricans*, *R. iberica*, and *S. salamandra*. The probability of a stream segment retaining water in summer showed a negative relation with *P. perezi* and a positive relation with *R. iberica*. Models including only biotic interactions showed significantly negative relations between *P. clarkii* and *A. obstetricans*, *R. iberica*, *L. boscai*, *S. salamandra* and *T. marmoratus* (Table 5.). *P. leniusculus* was negatively related to *A. cisternasii* and *S. salamandra*, and positively to *P. perezi*. *Bufo spinosus* was the only species showing no significant environmental or biotic effects.

When combining environmental and biotic variables (Table 5.1), the effects of biotic interactions were retained for all species except *P. perezi*, while the type and significance of environmental effects often changed considerably. For *L. boscai*, *S. salamandra* and *T. marmoratus*, only the negative effects of *P. clarkii* were retained in the best model. For *R. iberica*, the combined model highlighted a negative effect of *P. leniusculus* and stream order, and a positive effect of the probability of water presence. For *A. obstetricans*, there was a negative effect of *P. clarkii* and a positive effect of water presence, while for *A. cisternasii* there were negative effects of *P. leniusculus* and precipitation, and an interaction between stream order and elevation suggesting that the species was most likely to occur in higher stream orders at higher elevation, and the reverse at lower elevations.

Table 5.1 - Parameter estimates and summary statistics for models relating the probability of occurrence of amphibian species to variables describing environmental effects and biotic interactions in the Sabor watershed (NE Portugal). For each species we provide four different models incorporating (i) environmental variables (Env); (ii) biotic interactions (Bio); (iii) environmental variables + biotic interactions (Env/Bio); and (iv) environmental variables + biotic interactions + spatial dependencies across stream networks (Full). Variables included were: elevation (Alt), total annual precipitation (Prec), Strahler's stream order (SO), the water presence probability (Water), and the probability of the presence of either *Procambarus clarkii* (Pclar) or *Pacifastacus leniusculus* (Plen). We also included a multiplicative interaction between elevation and stream order (SOxAlt). For each model we provide the regression coefficient estimates (B), their standard errors (SE) and P-values, and three measures of model predictive ability: area under the receiver operating curve (AUC), Cohen's kappa and true skill statistics (TSS). Species observed prevalence was used as the threshold to compute Cohen's kappa, and the true skill statistics

Species/ models	Intercept		Alt		Prec		SO		Water		SOxAlt		Pclar		Plen		AUC	Kappa	TSS
	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P			
<i>Alytes cisternasii</i>																			
Env	-2.04 (0.27)	0.00	-1.23 (0.31)	0.00	-0.57 (0.32)	0.07	-1.00 (0.29)	0.00	-	-	-	-	-	-	-	-	0.72	0.23	0.36
Bio	-1.76 (0.23)	0.00	-	-	-	-	-	-	-	-	-	0.36 (0.23)	0.12	-0.66 (0.33)	0.05	-	0.45	0.24	0.39
Env + Bio	-1.80 (0.32)	0.00	-1.09 (0.31)	0.00	-0.90 (0.36)	0.01	1.06 (0.7)	0.13	-	-	1.25 (0.46)	0.01	-	-2.15 (0.78)	0.01	-	0.75	-0.02	0.05
Full	-1.31 (0.78)	0.09	-0.12 (0.44)	0.79	-0.64 (0.55)	0.24	0.88 (0.63)	0.17	-	-	1.08 (0.49)	0.03	-	-1.11 (0.64)	0.08	-	0.82	0.26	0.44
<i>Alytes obstetricans</i>																			
Env	-1.64 (0.26)	0.00	0.67 (0.39)	0.09	0.51 (0.24)	0.04	-0.07 (0.31)	0.82	-	-	0.79 (0.37)	0.03	-	-	-	-	0.66	0.15	0.26
Bio	-2.57 (0.68)	0.00	-	-	-	-	-	-	-	-	-	-	-2.20 (1.16)	0.06	-	-	0.53	0.14	0.31
Env + Bio	-3.83 (1.41)	0.01	-	-	-	-	-	-	0.60 (0.31)	0.06	-	-	-4.73 (2.55)	0.07	-	-	0.66	0.12	0.26
Full	-1.65 (0.31)	0.00	-	-	-	-	-	-	0.07 (0.28)	0.8	-	-	-0.36 (0.24)	0.14	-	-	0.8	0.27	0.48
<i>Pelophylax perezii</i>																			
Env	0.89 (0.18)	0.00	-	-	-	-	1.39 (0.53)	0.01	-0.89 (0.48)	0.06	-	-	-	-	-	-	0.57	0.11	0.13
Bio	0.85 (0.18)	0.00	-	-	-	-	-	-	-	-	-	-	-	0.49 (0.20)	0.02	-	0.52	0.12	0.14
Env + Bio	0.89 (0.18)	0.00	-	-	-	-	1.39 (0.53)	0.01	-0.89 (0.48)	0.06	-	-	-	-	-	-	0.57	0.11	0.13
Full	0.87 (0.28)	0.00	-	-	-	-	1.29 (0.62)	0.04	-0.71 (0.54)	0.19	-	-	-	-	-	-	0.68	0.17	0.21
<i>Rana iberica</i>																			
Env	-1.42 (0.27)	0.00	1.44 (0.51)	0.01	0.62 (0.25)	0.02	-1.83 (1.08)	0.09	1.61 (0.85)	0.06	0.79 (0.55)	0.15	-	-	-	-	0.83	0.46	0.52
Bio	-3.22 (0.95)	0.00	-	-	-	-	-	-	-	-	-	-	-4.54 (1.63)	0.01	-	-	0.66	0.34	0.5
Env + Bio	-4.81 (2.56)	0.06	-	-	-	-	-3.25 (1.02)	0.00	4.69 (0.84)	0.00	-	-	-6.27 (4.83)	0.2	-2.05 (0.93)	0.03	0.89	0.58	0.65
Full	-1.74 (0.69)	0.01	-	-	-	-	-1.61 (0.61)	0.01	2.62 (0.65)	0.00	-	-	-0.91 (0.48)	0.06	-1.87 (0.62)	0.00	0.91	0.61	0.68

Table 5.1 – (Cont.)

Species/ models	Intercept		Alt		Prec		SO		Water		SOxAlt		Pclar		Plen		AUC	Kappa	TSS
	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P	B (SE)	P			
<b>Lissotriton boscai</b>																			
Env	-2.39 (0.32)	0.00	-		-		-1.11 (0.34)	0	-		-		-		-		0.61	0.16	0.38
Bio	-5.34 (1.8)	0.00	-		-		-		-		-		-6.37 (2.99)	0.03	-		0.64	0.12	0.29
Env + Bio	-5.34 (1.8)	0.00	-		-		-		-		-		-6.37 (2.99)	0.03	-		0.64	0.12	0.29
Full	-5.34 (1.8)	0.00	-		-		-		-		-		-6.37 (2.99)	0.03	-		0.64	0.11	0.29
<b>Salamandra</b>																			
Env	-1.31 (0.25)	0.00	-		0.38 (0.21)	0.07	-1.28 (0.3)	0.00	-		-		-		-		0.78	0.39	0.46
Bio	-1.62 (0.39)	0.00	-		-		-		-		-		-1.19 (0.63)	0.06	-1.09 (0.44)	0.01	0.77	0.35	0.43
Env + Bio	-1.62 (0.39)	0.00	-		-		-		-		-		-1.19 (0.63)	0.06	-1.09 (0.44)	0.01	0.77	0.35	0.43
Full	-2.34 (1.78)	0.19	-		-		-		-		-		-2.12 (2.7)	0.43	-0.84 (0.55)	0.13	0.83	0.4	0.5
<b>Triturus marmoratus</b>																			
Env	-1.96 (0.27)	0.00	-		-		-0.92 (0.29)	0.00	-		-		-		-		0.59	0.14	0.28
Bio	-3.42 (1.10)	0.00	-		-		-		-		-		-3.65 (1.9)	0.05	-		0.59	0.2	0.43
Env + Bio	-3.42 (1.10)	0.00	-		-		-		-		-		-3.65 (1.9)	0.05	-		0.59	0.2	0.43
Full	-2.60 (0.85)	0.00	-		-		-		-		-		-2.22 (1.31)	0.09	-		0.66	0.04	0.08

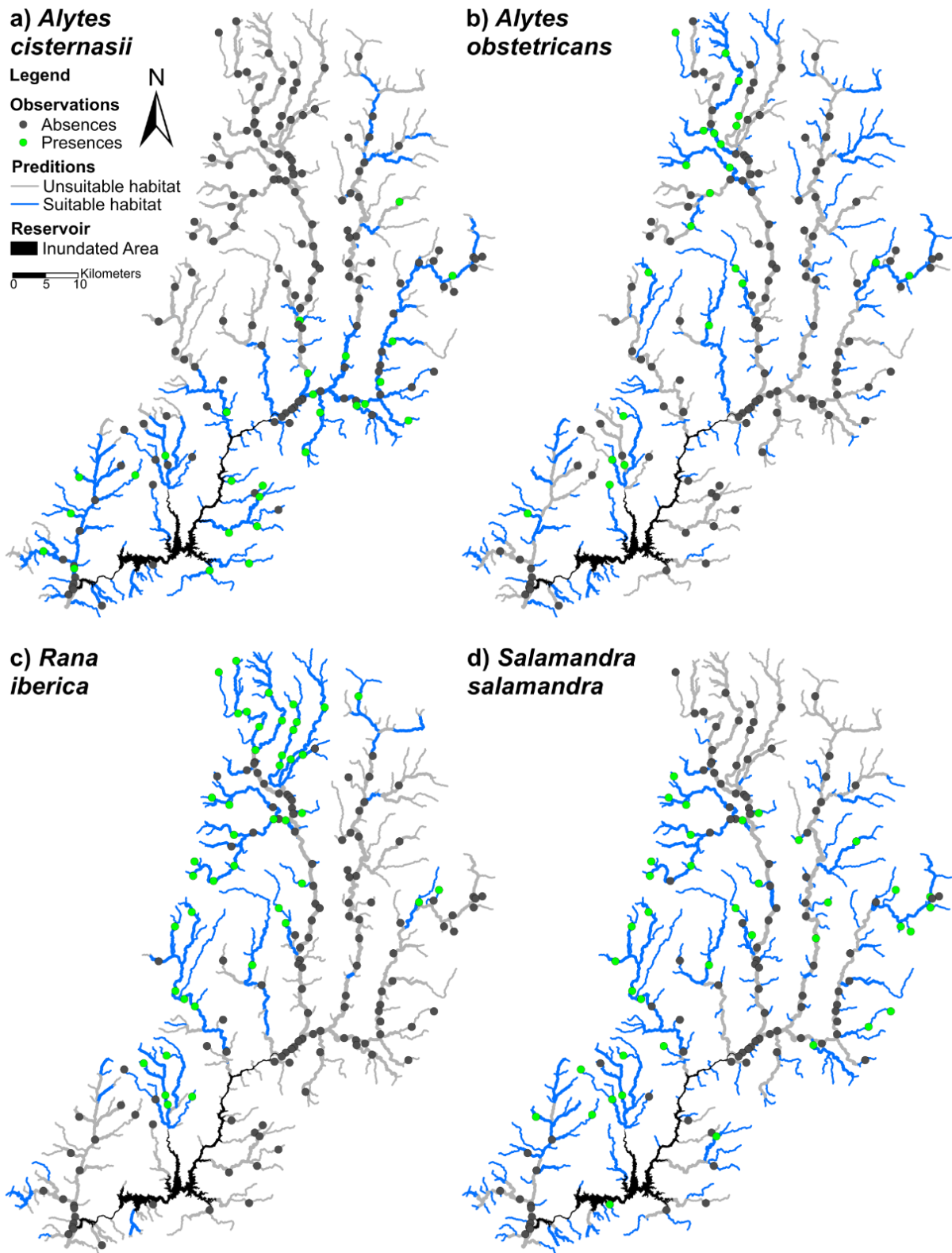
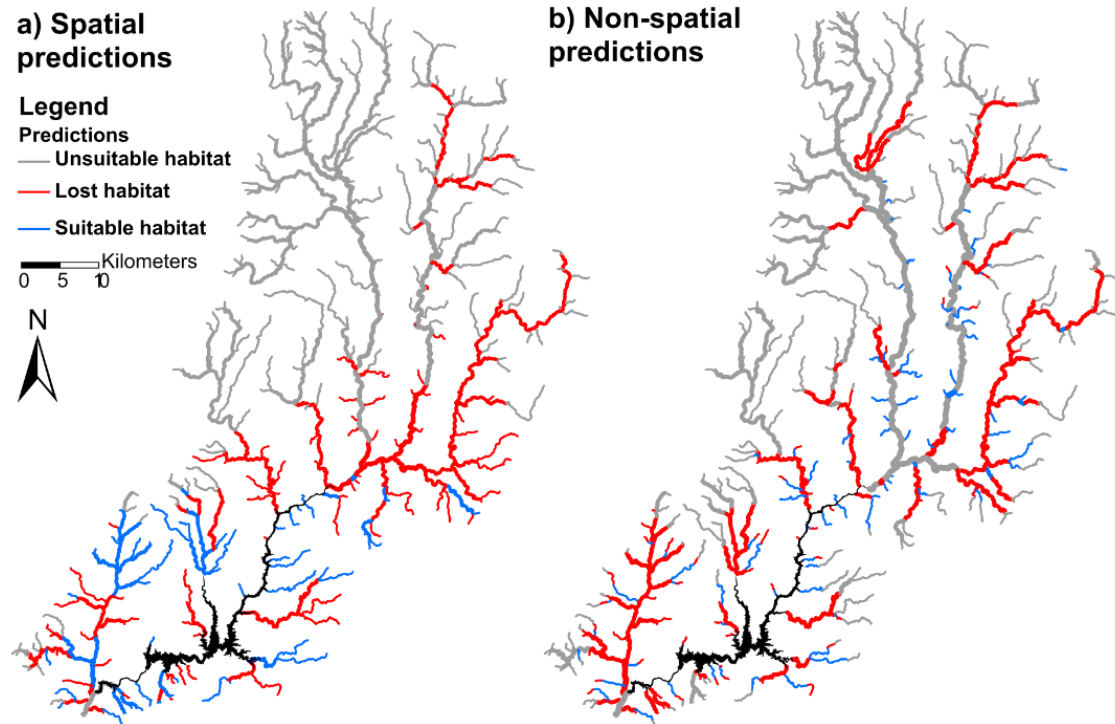


Figure 5.1 - Maps showing the observed presences/absences of four amphibian species in the river Sabor watershed (NE Portugal) in 2015, and their potential distributions predicted from geostatistical models combining environmental effects, biotic interactions and spatial dependencies across the dendritic stream network. The threshold for predicted presences was set equal to the observed prevalence of each species



### *Alytes cisternasii*



### *Alytes obstetricans*

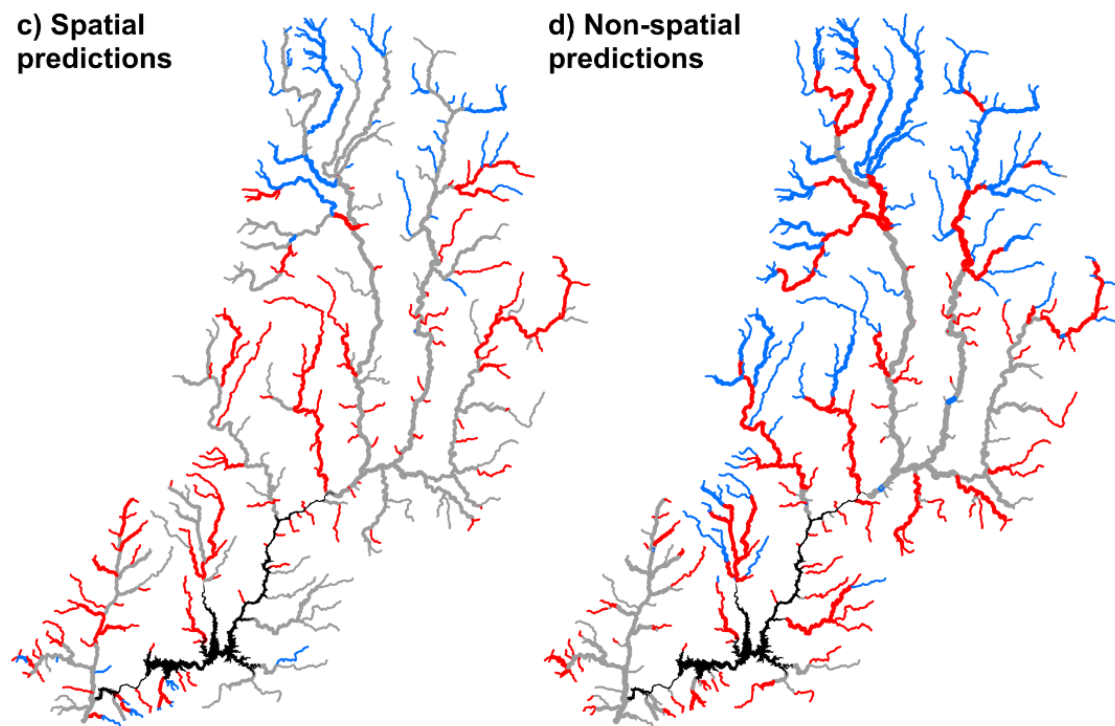


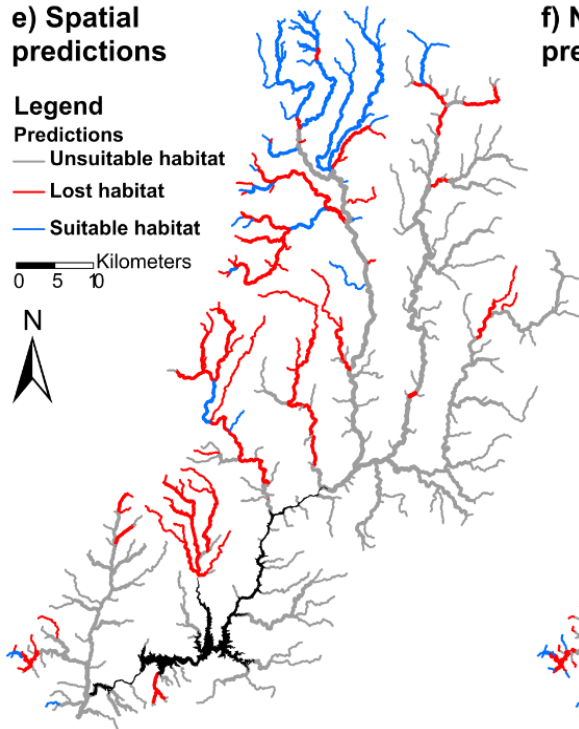
Figure 5.2 - Maps of potential distributions of four amphibian species under the worst case scenario of future crayfish expansion in the river Sabor watershed (NE Portugal). The scenario was built considering a five-fold increase in the relative risk of crayfish occurrence (both *P. clarkii* and *P. leniusculus*) at each stream segment in relation to the baseline scenario corresponding to the predicted distribution of each species in 2012 (Filipe et al., 2017). In each map we indicate the waterlines where potential habitat will remain available (suitable habitat), and those where potential habitat will be lost (lost habitat) in relation to the baseline scenario. Maps were produced using predictions from either spatial or non-spatial (i.e. including only environmental effects and biotic interactions) models.

***Rana iberica***

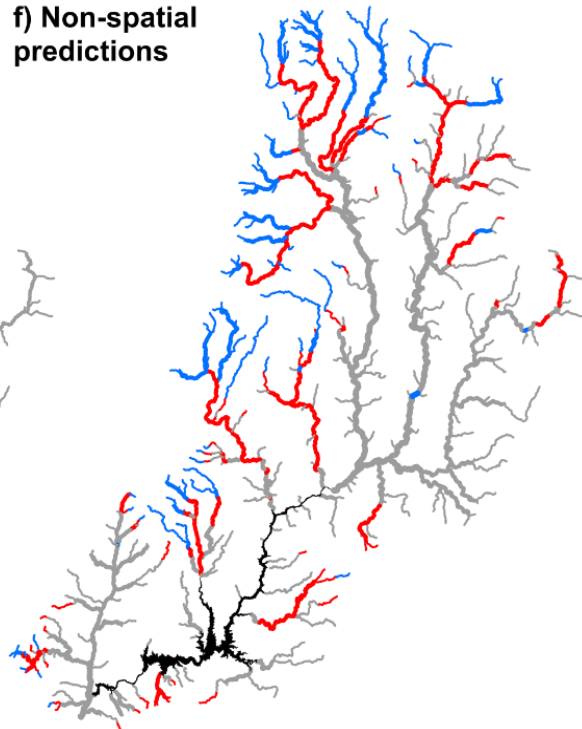
**e) Spatial  
predictions**

**Legend**  
Predictions  
— Unsuitable habitat  
— Lost habitat  
— Suitable habitat

0 5 10 Kilometers

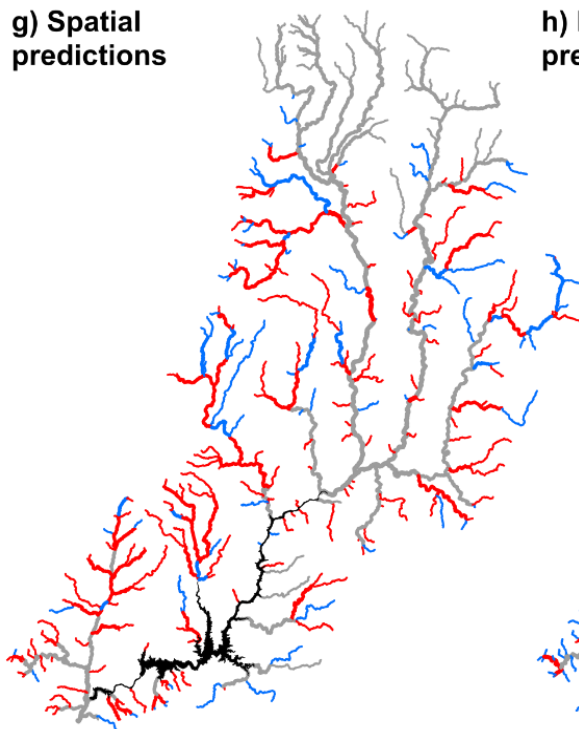


**f) Non-spatial  
predictions**



***Salamandra salamandra***

**g) Spatial  
predictions**



**h) Non-spatial  
predictions**

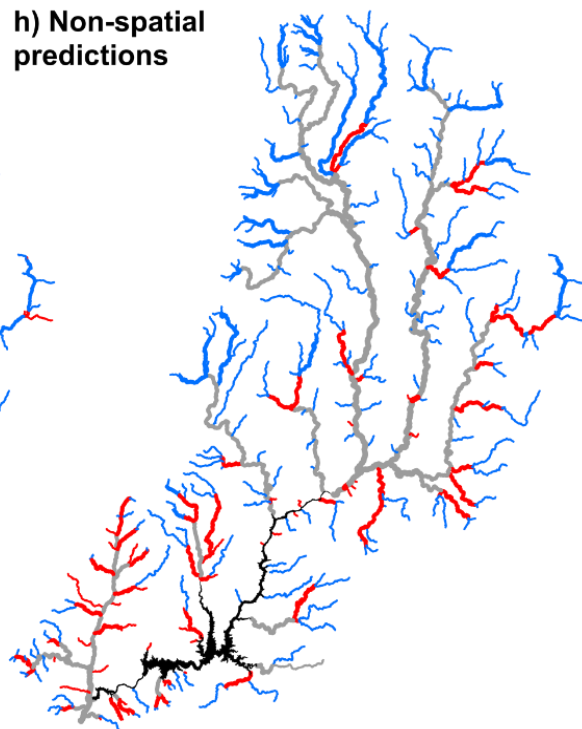


Figure 5.2 – (Cont)

Table 5.2 - Covariance structure selected for the geostatistical models of each species, indicating the percentage of variation accounted by the fixed and each spatial component of the final models. For each spatial component we indicate the function used to specify the covariance structure.

Species	Environment + Biotic	Spatial				Nugget
		Tail Up		Tail Down	Euclidean	
<i>Alytes cisternasii</i>	3.7	-		-	32.2 (Gaussian)	64.1
<i>Alytes obstetricans</i>	1.6	98.4	(Epanech)	-	-	≈ 0.0
<i>Pelophylax perezi</i>	4.4	-		-	31 (Cauchy)	64.6
<i>Rana iberica</i>	9.9	27.5	(Mariah)	30 (Spherical)	-	32.6
<i>Lissotriton boscai</i>	2.8	-		-	≈ 0.0 (Spherical)	97.2
<i>Salamandra salamandra</i>	3.9	17.7	(Spherical)	-	0.28 (Gaussian)	50.9
<i>Triturus marmoratus</i>	1.9	-		-	0.22 (Gaussian)	76.6

Torgegrams suggested the occurrence of spatial dependencies in model residuals among flow-connected sites for *R. iberica*, *L. boscai*, *S. salamandra* and *T. marmoratus*, while spatial dependencies for flow-unconnected sites were apparent for *S. salamandra* and, particularly, *A. cisternasii* (Figure 5.S3). The tail-up component was included in the best covariance structure for *A. obstetricans*, *R. iberica*, and *S. salamandra*, and the Euclidean component was included in the best models for *A. cisternasii*, *P. perezi*, *L. boscai*, *S. salamandra* and *T. marmoratus* (Table 5.2). The tail down component was only included in the best model for *R. iberica*. Overall, the spatial components accounted for a much larger amount of explained variation than the environmental + biotic effects (Table 5.2), and thus the full models showed the best performance in terms of AUC, Coehen's Kappa and TSS (Table 5.). The AUC of the full models was reasonable ( 0.80-0. 91) for *A. cisternasii*, *A. obstetricans*, *R. iberica*, and *S. salamandra*, but it was low for the other species (0.64-0.68) (Table 5.), which were thus not considered to map predicted distributions.

From a total of 1,468 km of waterlines in the Sabor watershed, the maps of predicted distributions based on the geostatistical models combining environmental, biotic and spatial predictors indicated that the species with most potential habitat was *S. salamandra* (64.6% of total stream length), followed by *A. obstetricans* (51.0%), *A. cisternasii* (48.4%) and *R. iberica* (38.9%) (Figure 5.1). *S. salamandra* occurred in lower order streams throughout the basin, while both *Alytes* species were widespread in lower and middle order streams, with *A.*

*cisternasii* occurring primarily in the south and southeast and *A. obstetricans* in the north and northwest (Figure 5.1). *R. iberica* was largely restricted to lower order streams in more mountainous areas of the north and northwest. The distributions of these species correspond to streams with low probability of occurrence of both invasive crayfish, but that were predicted to be progressively colonised under the invasion scenarios (Figure 5.S4 and Figure 5.S5). Expansion of crayfish through the stream network was predicted to decrease the length of potential stream habitat for amphibians, with each species becoming progressively more confined to first and second order streams (Figure 5.2, Table 5.S2). Predictions using spatial models suggested that the length of habitat of *A. cisternasii* will decline up to about 30% due to the expansion of *P. leniusculus*, while a reduction of about 7% was predicted for *A. obstetricans* due to the expansion of *P. clarkii*. The potential habitat for *R. iberica* and *S. salamandra* were expected to decline by about 20%, due to the joint expansion of *P. leniusculus* and *P. clarkii*. Non-spatial models predicted even larger declines for *A. cisternasii* (up to 69.9%), *A. obstetricans* (51.3%) and *R. iberica* (53.9%), but not as much for *S. salamandra* (24.1%) (Figure 5.2, Table 5.S2). For the latter species, there were upstream areas in the far north that were predicted to be occupied by the non-spatial model but not by the spatial model (Figure 5.2), suggesting that environmentally suitable habitats may remain unoccupied due to spatial processes.

## 5.4 Discussion

Our study showed the value of combining geostatistical and biotic interactions modelling to quantify the spatial consequences of biological invasions on native species in dendritic stream networks, and to predict the spatial distribution of ecological refuges under current and future invasion scenarios. Using this approach, we confirmed the strongly negative interactions between invasive crayfish and amphibians (Cruz, Pascoal, et al., 2006; Cruz, Rebelo, et al., 2006; Ficetola et al., 2011; Riley et al., 2005), while advancing previous knowledge by showing that such interactions are causing marked range contractions at the watershed scale in many species, and that this effect may intensify in the future under plausible scenarios of crayfish expansion. Moreover, our results show that stream headwaters (i.e., stream orders 1 and 2; Finn et al., 2011) represent key refuges from crayfish invasion for many amphibian species, as these streams dry out for more or less extended periods during the dry season (Ferreira et al., 2016) and are thus expected to remain largely free from crayfish impacts (Cruz & Rebelo, 2007; Filipe et al., 2017). Overall, our study reinforces the conservation importance of stream headwaters in the Mediterranean region, which are increasingly perceived to play key roles as refuges from biological invasions and other human-mediated disturbances, both for

amphibians (de Vries & Marco, 2017) and other vulnerable species (Quaglietta et al., 2018; Sousa et al., 2019). Our approach may be applied to other aquatic species, with major implications for conservation and management by permitting a better identification of areas acting as ecological refuges under biological invasion.

#### 5.4.1 Study limitations

Although our study had some limitations, it is unlikely that they affected our main conclusions in any significant way. One potential problem is that we sampled streams from first to sixth orders, and it may be argued that the prevalence of amphibian species may be underestimated in larger streams due to lower detectability. Although this can be addressed by modelling occupancy while controlling for detectability (e.g., MacKenzie et al., 2006), this was not possible in our case because occupancy-detection models accounting for hydrological tail-up and tail-down spatial dependencies have yet to be developed. To deal with this problem, we have surveyed only wadable streams and increased the sampling effort in larger streams, which should have contributed to achieve comparable detectability across stream orders. This is supported by the higher prevalence of *P. perezii* and the crayfish *P. clarkii* in higher than lower orders, which suggest that we did not miss species known to occur in larger streams (Cruz, Rebelo, et al., 2006; Filipe et al., 2017). Moreover, species distribution patterns observed in our study were consistent with those of others focusing on stream-dwelling amphibians in Iberia (Cruz, Rebelo, et al., 2006; de Vries & Marco, 2017), thereby suggesting that they were not artefacts shaped by sampling biases. Another potential problem is that amphibian models were based on large scale variables, while ignoring local drivers such as the structure and composition of riparian vegetation, the land uses surrounding streams and water quality (Crawford & Semlitsch, 2007; Guzy et al., 2018; Riley et al., 2005). Missing these variables might have reduced the predictive ability of our models, but we believe that the variables considered are relevant to investigate large scale distribution patterns, as observed for instance in other species modelled in the Sabor watershed (Ferreira et al., 2016; Filipe et al., 2017; Quaglietta et al., 2018). Nevertheless, while our models should be informative to understand broad changes in amphibian distributions in relation to crayfish invasion, they may be less useful to predict whether a given species will be present at any given site, which may be strongly affected by more local environmental conditions.

## 5.4.2 Effects of biotic interaction and environmental effects on amphibian distributions

The negative responses of amphibians to invasive crayfish observed in the Sabor watershed were comparable to those reported elsewhere (Cruz, Pascoal, et al., 2006; Cruz, Rebelo, et al., 2006; Gamradt et al., 1997; Gamradt & Kats, 1996; Girdner et al., 2018; Nyström et al., 2001, 2002). However, impacts of *P. leniusculus* are reported here for the first time in Iberia, though this crayfish was already associated with local amphibian declines in Sweden (Nyström et al., 2001, 2002). We also confirmed that crayfish impacts seem to be particularly strong on urodela (salamander and newts), possibly because their eggs and larvae are highly vulnerable to introduced predators (Cruz, Rebelo, et al., 2006; Gamradt et al., 1997; Gamradt & Kats, 1996; Girdner et al., 2018). Mediterranean amphibians may be especially vulnerable to crayfish because many species are adapted to live in water bodies that dry out in summer and are naturally free from fish and other large predators (e.g., Beja & Alcazar, 2003; Ferreira & Beja, 2013), and thus may be less adapted to cope with crayfish predation than species living in permanent waters (Cruz, Rebelo, et al., 2006; Nunes et al., 2011). This is supported by the lack of negative effects on *P. perezi*, which is known to thrive in permanent waters where predators are abundant (Beja & Alcazar, 2003; Cruz, Rebelo, et al., 2006; Ferreira & Beja, 2013). We also found no negative effects on *B. spinosus*, which is widespread in permanent waters and seems to be less vulnerable to predators due to its toxic eggs and larvae (Cruz & Rebelo, 2005), though a previous study reported negative impacts of *P. clarkii* (Cruz, Rebelo, et al., 2006). In contrast, we found negative impacts on *A. cisternasii*, which is associated with more permanent water bodies and was found previously to be unaffected by crayfish (Cruz, Rebelo, et al., 2006). However, that study was carried out in an area where only *P. clarkii* occurred, while in our study we only found significant negative effects for *P. leniusculus*, suggesting that impacts may differ among crayfish species. *P. leniusculus* also showed strong negative effects on *R. iberica*, possibly because this species occurs primarily in mountainous streams largely unsuitable for *P. clarkii* (Filipe et al., 2017). These results suggest that invasion by multiple crayfish species may be more serious than invasion by a single species, by increasing the types of habitats invaded (Filipe et al., 2017) and the number of species vulnerable to predation.

Although our study showed that many amphibian species were associated with stream headwaters, different species occurred in different areas, probably due to differences in ecological requirements. For instance, environmental models suggested that while both *A. cisternasii* and *A. obstetricans* were mainly found in lower and middle stream orders, the former favoured areas with lower precipitation at low elevation, while the reverse was found for the

latter. This probably explains their largely parapatric distributions in the study area, as observed at broader spatial scales (Reino et al., 2017). Models for *R. iberica* also confirmed their preference for small permanently flowing streams in mountainous areas (Bosch et al., 2006; Rodríguez-Prieto & Fernández-Juricic, 2005). Many of these environmental effects, however, were lost from the best models, or became non-significant, once biotic interactions were included. In some cases, there were also changes in the significant environmental effects, with for instance the selection of permanently flowing waters by *A. obstetricans* only becoming apparent after controlling for the effects of *P. clarkii*. These results support the idea that the two exotic crayfish are key drivers constraining amphibian distributions in our study area, limiting the range of environmental conditions where they can be found.

#### 5.4.3 The role of spatial dependencies across the stream network

Incorporating spatial covariance structure greatly enhanced the distribution models, with both Euclidean and hydrologic distances often included in the best models. The Euclidean component was important for *A. cisternasii*, *P. perezi*, *S. salamandra* and *T. marmoratus*, suggesting that adjacent streams have more similar occupancy status than streams farther apart, which may be due to the dispersal of individuals overland (e.g., Semlitsch, 2008) or similarities regarding unmeasured spatially-structured environmental variables (e.g., land cover/land uses). Euclidean effects may be particularly important in species associated with headwaters, because nearby streams may be flow-unconnected and at long hydrologic distances from each other. The tail-up component was important for *A. obstetricans*, *R. iberica* and *S. salamandra*, which may be a consequence of flow-connected sites having similar environmental conditions, but also of similarities in occupancy status due for instance to downstream drift of larvae or their active swimming upstream. The tail-down component was only important for *R. iberica*, possibly reflecting dispersal movements along the waterlines irrespective of flow, as this is a species strongly attached to riverine habitats during the adult and larval stages, and may have low dispersal ability overland as suggested by the lack of the Euclidean component.

In the full mixed models, spatial dependencies accounted for a far greater proportion of variation in species occurrences than the fixed component, as observed in other studies using geostatistical tools (Filipe et al., 2017; Lois et al., 2015; Lois & Cowley, 2017; Quaglietta et al., 2018). This was probably because the fixed component was specified using variables that are spatially structured, either Euclidean (e.g. precipitation, elevation) or hydrologic (e.g., stream

order), and thus their effects were reduced after considering spatial dependencies. It is worth noting, however, that variables describing biotic interactions generally remained significant in the full models, further emphasizing their importance. The fixed component would probably have had a larger share of the explained variation, probably contributing to the overall predictive power of the models, if we had considered variables describing more local environmental conditions that are known to affect stream-dwelling amphibians (e.g., Crawford & Semlitsch, 2007; Guzy et al., 2018; Riley et al., 2005). Future studies combining drivers operating at landscape and local scales should thus be developed, which would likely improve predictions on potentially favourable areas across the watershed.

#### 5.4.4 Predicting amphibian refuges under crayfish invasion

Despite the limited explanatory power of the spatial distribution models, the mapping of predicted distributions clearly showed that further crayfish expansions will likely result in amphibian range contractions, with populations becoming progressively more encroached in lower order streams. This could be inferred quantitatively for four species with geostatistical models with sufficient predictive ability (*A. cisternasii*, *A. obstetricans*, *R. perezii* and *S. salamandra*), but will probably occur also for the other two species showing negative associations with crayfish occurrence (*L. boscai* and *T. marmoratus*) in ours and other studies (Cruz, Rebelo, et al., 2006). The non-spatial models (i.e., including only environment + biotic interactions) predicted even stronger declines in the availability of potential habitats, particularly for *A. cisternasii*, *A. obstetricans* and *R. iberica*. This is because spatial models anchor model predictions to the current distribution of each species, and thus may be regarded as conservative because they add inertia against abrupt changes in distribution driven by environmental factors (Record et al., 2013). In contrast, the non-spatial models are only driven by changes in crayfish occurrence irrespective of the spatial structure in current amphibian distribution, thereby disregarding possible spatially structured population processes shaping amphibian species distributions (Record et al., 2013). Therefore, we expect that the extent of range contractions will be somewhere in-between the predictions of spatial and non-spatial models, which may thus be substantial for *A. cisternasii* (up to 30.6-69.6%), *A. obstetricans* (6.8%-51.3%) and *R. iberica* (20.6%-53.9%), though only moderate to *S. salamandra* (22.3-24.1%). The consequences of such changes for population persistence should be evaluated in future studies, as it is likely that the risk of local extinctions will be high for small and isolated populations confined to headwater streams.



### 5.4.5 Conservation and management implications

Invasion by alien crayfish is a major cause of concern for amphibian conservation (e.g., Cruz, Pascoal, et al., 2006; Cruz, Rebelo, et al., 2006; Gamradt et al., 1997; Gamradt & Kats, 1996; Girdner et al., 2018; Nyström et al., 2001, 2002). Addressing this problem is challenging, because once established invasive crayfish populations are virtually impossible to eradicate, and thus remediation of ecosystems invaded by crayfish have met very limited success (Gherardi et al., 2011; Stebbing et al., 2014). Furthermore, many invasive crayfish species are still expanding within and across watersheds (e.g., Bernardo et al., 2011; Kouba et al., 2014), and so the problem is likely to get worse in the future. In this context, our study suggests that stream headwaters may be critical for the persistence of many stream-dwelling amphibian species, at least in the Mediterranean region, as they often hold diverse amphibian communities (de Vries & Marco, 2017) and are likely to provide refuges with minimal or no crayfish impacts (Filipe et al., 2017). These headwaters correspond not only to small order streams in mountainous areas, as those inhabited for instance by *A. obstetricans* and *R. perezii*, but also small temporary streams at lower elevation, which seem to be preferred by species such as *A. cisternasii*. Overall, therefore, headwater streams should be regarded as priority targets for conservation, requiring the preservation of habitat conditions compatible with amphibian persistence. Although the ecological requirements of stream-dwelling amphibians are poorly known in the Iberian Peninsula (de Vries & Marco, 2017), it is likely that conservation efforts should target preserving water quality, natural flow regimes, well-developed riparian vegetation and suitable terrestrial habitats (e.g., Crawford & Semlitsch, 2007; Guzy et al., 2018; Riley et al., 2005). Furthermore, efforts should be made to avoid the colonisation of headwater refuges by invasive crayfish such as *P. leniusculus*, which is fast expanding into new areas (Anastácio et al., 2015; Bernardo et al., 2011), and may be able to colonise mountainous headwater streams inhabited by endemic amphibians such as *R. iberica* (Filipe et al., 2017). This would require monitoring crayfish populations in key amphibian refuges, which should be used to trigger careful management programs if the risk of negative impacts become unacceptably high, involving for instance the implementation of control or eradication programs, and/or the introduction of physical barriers to crayfish dispersal (Gherardi et al., 2011; Sousa et al., 2019; Stebbing et al., 2014). Although such conservation measures may require considerable efforts and may only be applicable in some areas, maintaining stream headwaters free of invasive crayfish should have major conservation benefits for a range of endangered species (de Vries & Marco, 2017; Quaglietta et al., 2018; Sousa et al., 2019) and for aquatic biodiversity in general (Finn et al., 2011; Meyer et al., 2007).

## 5.5 Conclusions

Under biological invasion, many native species are becoming confined to refuges where invasive species are still absent or scarce, and thus may hold remnant populations of high conservation value (e.g., Chapman et al., 1996; Grabowski et al., 2009; Habit et al., 2010; Radinger et al., 2019). This study provides a framework to predict the location and environmental drivers of such refuges, using geostatistical tools to model native species responses to exotic species while controlling for environmental effects and spatial dependencies across dendritic stream networks (Filipe et al., 2017; Peterson et al., 2013). This approach is relatively simple and can be used where only snapshot surveys on the occurrence patterns of native and invasive species are available, though it can be easily extended to deal with data on distributional dynamics (Quaglietta et al., 2018) and additional complexities such as climate change (Peterson et al., 2013). This framework may be generally useful to understand the distributional consequences of interactions between native and invasive species, providing information on the location of potential refuges where conservation efforts should concentrate, and on management actions required to enhance the persistence of remnant populations within refuges.

## Acknowledgements

This work was carried out in the scope of the Baixo Sabor Long Term Ecological Research Site. Funding was provided by EDP and by the Portuguese Science and Technology Foundation (FCT). MMF was supported by the FCT PhD grant SFRH/BD/95202/2013, and PB by EDP Biodiversity Chair. We thank collaboration in field work by Ninda Batista, Joana Pinto, Erika Almeida, and Ivo Rosa.

## 5.6 References

- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Anastácio, P. M., Banha, F., Capinha, C., Bernardo, J. M., Costa, A. M., Teixeira, A., & Bruxelas, S. (2015). Indicators of movement and space use for two co-occurring invasive crayfish species. *Ecological Indicators*, 53, 171–181. <https://doi.org/10.1016/j.ecolind.2015.01.019>
- Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16(6), 743–753. <https://doi.org/10.1111/j.1466-8238.2007.00359.x>
- Axelsson, E., Brönmark, C., Sidenmark, J., & Nyström, P. (1997). Crayfish predation on amphibian eggs

- and larvae. *Amphibia-Reptilia*, 18(3), 217–228. <https://doi.org/10.1163/156853897X00107>
- Azevedo, J. C., Moreira, C., Castro, J. P., & Loureiro, C. (2011). Agriculture Abandonment, Land-use Change and Fire Hazard in Mountain Landscapes in Northeastern Portugal. In C. Li, R. Laforzezza, & J. Chen (Eds.), *Landscape Ecology in Forest Management and Conservation: Challenges and Solutions for Global Change* (pp. 329–351). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-12754-0\\_14](https://doi.org/10.1007/978-3-642-12754-0_14)
- Barbosa, A. M., Brown, J. A., Jiménez-Valverde, A., & Real, R. (2016). *modEvA: Model Evaluation and Analysis* (1.3.2). CRAN. <https://cran.r-project.org/package=modEvA>
- Barton, K. (2016). *MuMIn: Multi-Model Inference* (1.43.17). CRAN. <https://cran.r-project.org/web/packages/MuMIn>
- Beja, P., & Alcazar, R. (2003). Conservation of Mediterranean temporary ponds under agricultural intensification: an evaluation using amphibians. *Biological Conservation*, 114(3), 317–326. [https://doi.org/10.1016/S0006-3207\(03\)00051-X](https://doi.org/10.1016/S0006-3207(03)00051-X)
- Bernardo, J. M., Costa, a. M., Bruxelas, S., & Teixeira, a. (2011). Dispersal and coexistence of two non-native crayfish species ( *Pacifastacus leniusculus* and *Procambarus clarkii* ) in NE Portugal over a 10-year period. *Knowledge and Management of Aquatic Ecosystems*, 401, 28. <https://doi.org/10.1051/kmae/2011047>
- Bivand, R., Keitt, T., & Rowlingson, B. (2019). *rgdal: Bindings for the “Geospatial” Data Abstraction Library* (1.4-4). CRAN. <https://cran.r-project.org/package=rgdal>
- Bosch, J., Rincón, P. A., Boyero, L., & Martínez-Solano, I. (2006). Effects of introduced salmonids on a montane population of Iberian frogs. *Conservation Biology*, 20(1), 180–189. <https://doi.org/10.1111/j.1523-1739.2005.00296.x>
- Caetano, M., Marcelino, F., Igreja, C., & Girão, I. (2018). Estatísticas e dinâmicas territoriais em Portugal Continental 1995-2007-2010-2015 com base na Carta de Uso e Ocupação do Solo (COS). <http://mapas.dgterritorio.pt/atom-dgt/pdf-cous/COS2015/COS2015-e-dinamicas-junho-2018.pdf>
- Chapman, L. J., Chapman, C. A., Ogutu-Ohwayo, R., Chandler, M., Kaufman, L., & Keiter, A. E. (1996). Refugia for endangered fishes from an introduced predator in Lake Nabugabo, Uganda. *Conservation Biology*, 10(2), 554–561. <https://doi.org/10.1046/j.1523-1739.1996.10020554.x>
- Clavero, M. (2016). Species substitutions driven by anthropogenic positive feedbacks: Spanish crayfish species as a case study. *Biological Conservation*, 193, 80–85. <https://doi.org/10.1016/j.biocon.2015.11.017>
- Clavero, M., Nores, C., Kubersky-Piredda, S., & Centeno-Cuadros, A. (2016). Interdisciplinarity to reconstruct historical introductions: solving the status of cryptogenic crayfish. *Biological Reviews*, 91(4), 1036–1049. <https://doi.org/10.1111/brv.12205>
- Costa, J. C., Monteiro-Henriques, T., Neto, C., Arsénio, P., & Aguiar, C. (2007). The application of the Habitats Directive in Portugal. *Fitosociologia*, 44(2 SUPPL. 1), 23–28.
- Crawford, J. A., & Semlitsch, R. D. (2007). Estimation of core terrestrial habitat for stream-breeding salamanders and delineation of riparian buffers for protection of biodiversity. *Conservation Biology: The Journal of the Society for Conservation Biology*, 21(1), 152–158. <https://doi.org/10.1111/j.1523-1739.2006.00556.x>
- Crooks, J. A. (2005). Lag times and exotic species: The ecology and management of biological invasions in slow-motion1. *Écoscience*, 12(3), 316–329. <https://doi.org/10.2980/i1195-6860-12-3-316.1>

- Cruz, M. J., Pascoal, S., Tejado, M., & Rebelo, R. (2006). Predation by an exotic crayfish, *Procambarus clarkii*, on natterjack toad, *Bufo calamita*, embryos: its role on the exclusion of this amphibian from its breeding ponds. *Copeia*, 2(2), 274–280. [https://doi.org/10.1643/0045-8511\(2006\)6\[274:PBAECP\]2.0.CO;2](https://doi.org/10.1643/0045-8511(2006)6[274:PBAECP]2.0.CO;2)
- Cruz, M. J., & Rebelo, R. (2005). Vulnerability of Southwest Iberian amphibians to an introduced crayfish, *Procambarus clarkii*. *Amphibia-Reptilia*, 26(3), 293–303. <https://doi.org/10.1163/156853805774408577>
- Cruz, M. J., & Rebelo, R. (2007). Colonization of freshwater habitats by an introduced crayfish, *Procambarus clarkii*, in Southwest Iberian Peninsula. *Hydrobiologia*, 575(1), 191–201. <https://doi.org/10.1007/s10750-006-0376-9>
- Cruz, M. J., Rebelo, R., & Crespo, E. G. (2006). Effects of an introduced crayfish, *Procambarus clarkii*, on the distribution of south-western Iberian amphibians in their breeding habitats. *Ecography*, 29(3), 329–338. <https://doi.org/10.1111/j.2006.0906-7590.04333.x>
- Cruz, M. J., Segurado, P., Sousa, M., & Rebelo, R. (2008). Collapse of the amphibian community of the Paul do Boquilobo Natural Reserve (central Portugal) after the arrival of the exotic American crayfish *Procambarus clarkii*. *Herpetological Journal*, 18(4), 197–204. <http://cat.inist.fr/?aModele=afficheN&cpsidt=21982001>
- Davis, J., Pavlova, A., Thompson, R., & Sunnucks, P. (2013). Evolutionary refugia and ecological refuges: Key concepts for conserving Australian arid zone freshwater biodiversity under climate change. In *Global Change Biology* (Vol. 19, Issue 7, pp. 1970–1984). Wiley-Blackwell. <https://doi.org/10.1111/gcb.12203>
- De Marco, P., Alexandre, J., Diniz-Filho, F., & Bini, L. M. (2008). Spatial analysis improves species distribution modelling during range expansion. *Biol. Lett*, 4, 577–580. <https://doi.org/10.1098/rsbl.2008.0210>
- de Vries, W., & Marco, A. (2017). The importance of fluvial habitats for amphibian conservation in the Mediterranean climate. *Basic and Applied Herpetology*, 31, 5–16. <https://doi.org/10.11160/bah.56>
- Diaz-Paniagua, C. (1992). Variability in timing of larval season in an amphibian community in SW Spain. *Ecography*, 15(3), 267–272. <https://doi.org/10.1111/j.1600-0587.1992.tb00034.x>
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330–342. <https://doi.org/10.1111/j.2041-210X.2010.00036.x>
- ESRI. (2016). *ArcGIS Desktop 10.5* (10.5). Environmental Systems, Research Institute.
- Ferreira, M., & Beja, P. (2013). Mediterranean amphibians and the loss of temporary ponds: Are there alternative breeding habitats? *Biological Conservation*, 165(null), 179–186. <https://doi.org/10.1016/j.biocon.2013.05.029>
- Ferreira, M., Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>
- Ficetola, G. F., Siesa, M. E., Manenti, R., Bottoni, L., De Bernardi, F., & Padoa-Schioppa, E. (2011). Early assessment of the impact of alien species: Differential consequences of an invasive crayfish on adult and larval amphibians. *Diversity and Distributions*, 17(6), 1141–1151. <https://doi.org/10.1111/j.1472-4642.2011.00797.x>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>

- Filipe, A. F., Quaglietta, L., Ferreira, M., Magalhães, M. F., & Beja, P. (2017). Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks. *Biological Invasions*, 19(10), 2899–2912. <https://doi.org/10.1007/s10530-017-1492-3>
- Finn, D. S., Bonada, N., Múrria, C., & Hughes, J. M. (2011). Small but mighty: headwaters are vital to stream network biodiversity at two levels of organization. *Journal of the North American Benthological Society*, 30(4), 963–980. <https://doi.org/10.1899/11-012.1>
- Gamradt, S. C., & Kats, L. B. (1996). Effect of Introduced Crayfish and Mosquitofish on California Newts. *Conservation Biology*, 10(4), 1155–1162. <https://doi.org/10.1046/j.1523-1739.1996.10041155.x>
- Gamradt, S. C., Kats, L. B., & Anzalone, C. B. (1997). Aggression by non-native crayfish deters breeding in California Newts. *Conservation Biology*, 11(3), 793–796. <https://doi.org/10.1046/j.1523-1739.1997.96230.x>
- Gherardi, F., Aquiloni, L., Diéguez-Uribeondo, J., & Tricarico, E. (2011). Managing invasive crayfish: Is there a hope? In *Aquatic Sciences* (Vol. 73, Issue 2, pp. 185–200). SP Birkhäuser Verlag Basel. <https://doi.org/10.1007/s00027-011-0181-z>
- Gil-Sánchez, J. M., & Alba-Tercedor, J. (2002). Ecology of the native and introduced crayfishes *Austropotamobius pallipes* and *Procambarus clarkii* in southern Spain and implications for conservation of the native species. *Biological Conservation*, 105(1), 75–80. [https://doi.org/10.1016/S0006-3207\(01\)00205-1](https://doi.org/10.1016/S0006-3207(01)00205-1)
- Girdner, S. F., Ray, A. M., Buktenica, M. W., Hering, D. K., Mack, J. A., & Umek, J. W. (2018). Replacement of a unique population of newts (*Taricha granulosa mazamae*) by introduced signal crayfish (*Pacifastacus leniusculus*) in Crater Lake, Oregon. *Biological Invasions*, 20(3), 721–740. <https://doi.org/10.1007/s10530-017-1570-6>
- Grabowski, M., Bacela, K., Konopacka, A., & Jazdzewski, K. (2009). Salinity-related distribution of alien amphipods in rivers provides refugia for native species. *Biological Invasions*, 11(9), 2107–2117. <https://doi.org/10.1007/s10530-009-9502-8>
- Gutiérrez-Yurrita, P., M. Martínez, J., Ilhéu, M., Bravo-Utrera, M., Bernardo, J., & Montes, C. (1999). The status of crayfish populations in Spain and Portugal. In F. Gherardi & D. M. Holdich (Eds.), *Crayfish in Europe as Alien Species* (1st ed., Vol. 11, pp. 161–192). Balkema, Rotterdam/Brookfield.
- Guzy, J. C., Eskew, E. A., Halstead, B. J., & Price, S. J. (2018). Influence of damming on anuran species richness in riparian areas: A test of the serial discontinuity concept. *Ecology and Evolution*, 8(4), 2268–2279. <https://doi.org/10.1002/ece3.3750>
- Habit, E., Piedra, P., Ruzzante, D. E., Walde, S. J., Belk, M. C., Cussac, V. E., Gonzalez, J., & Colin, N. (2010). Changes in the distribution of native fishes in response to introduced species and other anthropogenic effects. *Global Ecology and Biogeography*, 19(5), 697–710. <https://doi.org/10.1111/j.1466-8238.2010.00541.x>
- Hoelzer, A. (2003). *Vegetation ecological studies at the lower course of Sabor River (Tras-os-Montes, NE-Portugal)* [University of Bremen]. <http://www.grin.com/en/e-book/28444/vegetation-ecological-studies-at-the-lower-course-of-sabor-river-tras-os-montes>
- Hughes, R. M., Kaufmann, P. R., & Weber, M. H. (2011). National and regional comparisons between Strahler order and stream size. *Journal of the North American Benthological Society*, 30(1), 103–121. <https://doi.org/10.1899/09-174.1>
- Hulme, P. E. (2017). Climate change and biological invasions: evidence, expectations, and response options. *Biological Reviews*, 92(3), 1297–1313. <https://doi.org/10.1111/brv.12282>

- Kerby, J. L., Riley, S. P. D., Kats, L. B., & Wilson, P. (2005). Barriers and flow as limiting factors in the spread of an invasive crayfish (*Procambarus clarkii*) in southern California streams. *Biological Conservation*, 126(3), 402–409. <https://doi.org/10.1016/j.biocon.2005.06.020>
- Kouba, A., Petrušek, A., & Kozák, P. (2014). Continental-wide distribution of crayfish species in Europe: update and maps. *Knowledge and Management of Aquatic Ecosystems*, 413, 05. <https://doi.org/10.1051/kmae/2014007>
- Liu, C., Berry, P. M., Dawson, T. P., & Pearson, R. G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3), 385–393. <https://doi.org/10.1111/j.0906-7590.2005.03957.x>
- Lodge, D. M., Deines, A., Gherardi, F., Yeo, D. C. J., Arcella, T., Baldrige, A. K., Barnes, M. A., Chadderton, W. L., Feder, J. L., Gantz, C. A., Howard, G. W., Jerde, C. L., Peters, B. W., Peters, J. A., Sargent, L. W., Turner, C. R., Wittmann, M. E., & Zeng, Y. (2012). Global Introductions of Crayfishes: Evaluating the Impact of Species Invasions on Ecosystem Services. *Annual Review of Ecology, Evolution, and Systematics*, 43(1), 449–472. <https://doi.org/10.1146/annurev-ecolsys-111511-103919>
- Lois, S., & Cowley, D. E. (2017). Conservation of interacting species in network-constrained environments. *Diversity and Distributions*, 23(11), 1235–1245. <https://doi.org/10.1111/ddi.12623>
- Lois, S., Cowley, D. E., Outeiro, A., San Miguel, E., Amaro, R., & Ondina, P. (2015). Spatial extent of biotic interactions affects species distribution and abundance in river networks: The freshwater pearl mussel and its hosts. *Journal of Biogeography*, 42(2), 229–240. <https://doi.org/10.1111/jbi.12435>
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier Academic Press.
- Meyer, J. L., Strayer, D. L., Wallace, J. B. B., Eggert, S. L., Helfman, G. S., & Leonard, N. E. (2007). The contribution of headwater streams to biodiversity in river networks. *JAWRA Journal of the American Water Resources Association*, 43(1), 86–103. <https://doi.org/10.1111/j.1752-1688.2007.00008.x>
- Murtaugh, P. A. (2009). Performance of several variable-selection methods applied to real ecological data. *Ecology Letters*, 12(10), 1061–1068. <https://doi.org/10.1111/j.1461-0248.2009.01361.x>
- Nunes, A., Andrade, P., Rebelo, R., Guerreiro, C., Almeida, E., & Alves, S. (2011). Antipredator responses of two anurans towards native and exotic predators. *Amphibia-Reptilia*, 32(3), 341–350. <https://doi.org/10.1163/017353711X579849>
- Nyström, P., Birkedal, L., Dahlberg, C., & Brönmark, C. (2002). The declining spadefoot toad *Pelobates fuscus*: Calling site choice and conservation. *Ecography*, 25(4), 488–498. <https://doi.org/10.1034/j.1600-0587.2002.250411.x>
- Nyström, P., Svensson, O., Lardner, B., Brönmark, C., & Granéli, W. (2001). The influence of multiple introduced predators on a littoral pond community. *Ecology*, 82(4), 1023–1039. [https://doi.org/10.1890/0012-9658\(2001\)082\[1023:Tiomip\]2.0.co;2](https://doi.org/10.1890/0012-9658(2001)082[1023:Tiomip]2.0.co;2)
- Peterson, E. E., & Ver Hoef, J. M. (2010). A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91(3), 644–651. <https://doi.org/10.1890/08-1668.1>
- Peterson, E. E., & Ver Hoef, J. M. (2014). Stars: An ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *Journal of Statistical Software*, 56(2), 1–17. <https://doi.org/10.18637/jss.v056.i02>

- Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>
- Quaglietta, L., Paupério, J., Martins, F. M. S., Alves, P. C., & Beja, P. (2018). Recent range contractions in the globally threatened Pyrenean desman highlight the importance of stream headwater refugia. *Animal Conservation*. <https://doi.org/10.1111/acv.12422>
- R Core Team. (2017). R: A Language and Environment for Statistical Computing. Version 3.4.1. (3.4.1). R Foundation for Statistical Computing.
- Radinger, J., Alcaraz-Hernández, J. D., & García-Berthou, E. (2019). Environmental filtering governs the spatial distribution of alien fishes in a large, human-impacted Mediterranean river. *Diversity and Distributions*, 25(5), 701–714. <https://doi.org/10.1111/ddi.12895>
- Rahel, F. J. (2013). Intentional fragmentation as a management strategy in aquatic systems. *BioScience*, 63(5), 362–372. <https://doi.org/10.1525/bio.2013.63.5.9>
- Record, S., Fitzpatrick, M. C., Finley, A. O., Veloz, S., & Ellison, A. M. (2013). Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change. *Global Ecology and Biogeography*, 22(6), 760–771. <https://doi.org/10.1111/geb.12017>
- Reino, L., Ferreira, M., Martínez-Solano, Í., Segurado, P., Xu, C., & Márcia Barbosa, A. (2017). Favourable areas for co-occurrence of parapatric species: niche conservatism and niche divergence in Iberian tree frogs and midwife toads. *Journal of Biogeography*, 44(1), 88–98. <https://doi.org/10.1111/jbi.12850>
- Riley, S. P. D., Busteed, G. T., Kats, L. B., Vandergon, T. L., Lee, L. F. S., Dagit, R. G., Kerby, J. L., Fisher, R. N., & Sauvajot, R. M. (2005). Effects of urbanization on the distribution and abundance of amphibians and invasive species in southern California streams. *Conservation Biology: The Journal of the Society for Conservation Biology*, 19(6), 1894–1907. <https://doi.org/10.1111/j.1523-1739.2005.00295.x>
- Rodríguez-Prieto, I., & Fernández-Juricic, E. (2005). Effects of direct human disturbance on the endemic Iberian frog *Rana iberica* at individual and population levels. *Biological Conservation*, 123(1), 1–9. <https://doi.org/10.1016/j.biocon.2004.10.003>
- Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1(2), 103–113. <https://doi.org/10.1111/j.2041-210x.2010.00012.x>
- Semlitsch, R. D. (2008). Differentiating Migration and Dispersal Processes for Pond-Breeding Amphibians. *Journal of Wildlife Management*, 72(1), 260–267. <https://doi.org/10.2193/2007-082>
- Sousa, R., Nogueira, J. G., Ferreira, A., Carvalho, F., Lopes-Lima, M., Varandas, S., & Teixeira, A. (2019). A tale of shells and claws: The signal crayfish as a threat to the pearl mussel *Margaritifera margaritifera* in Europe. *Science of the Total Environment*, 665, 329–337. <https://doi.org/10.1016/j.scitotenv.2019.02.094>
- Stebbing, P., Longshaw, M., & Scott, A. (2014). Review of methods for the management of non-indigenous crayfish, with particular reference to Great Britain. *Ethology Ecology and Evolution*, 26(2–3), 204–231. <https://doi.org/10.1080/03949370.2014.908326>
- Strayer, D. L. (2010). Alien species in fresh waters: ecological effects, interactions with other stressors, and prospects for the future. *Freshwater Biology*, 55(SUPPL. 1), 152–174. <https://doi.org/10.1111/j.1365-2427.2009.02380.x>

- Twardochleb, L. A., Olden, J. D., & Larson, E. R. (2013). A global meta-analysis of the ecological impacts of nonnative crayfish. *Freshwater Science*, 32(4), 1367–1382. <https://doi.org/10.1899/12-203.1>
- Václavík, T., Kupfer, J. A., & Meentemeyer, R. K. (2012). Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography*, 39(1), 42–55. <https://doi.org/10.1111/j.1365-2699.2011.02589.x>
- Ver Hoef, J. M., & Peterson, E. E. (2010). A Moving Average Approach for Spatial Statistical Models of Stream Networks. *Journal of the American Statistical Association*, 105(489), 6–18. <https://doi.org/10.1198/jasa.2009.ap08248>
- Ver Hoef, J. M., Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *Journal of Statistical Software*, 56(3), 1–45. <https://doi.org/10.18637/jss.v056.i03>
- Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13(4), 449–464. <https://doi.org/10.1007/s10651-006-0022-8>
- Vogt, J., Soille, P., De Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., Bódis, K., Dusart, J., Paracchini, M. L., Haastrop, P., & Bamps, C. (2007). *A pan-European River and Catchment Database*. <https://doi.org/10.2788/35907>
- Walsh, J. R., Carpenter, S. R., & Vander Zanden, M. J. (2016). Invasive species triggers a massive loss of ecosystem services through a trophic cascade. *Proceedings of the National Academy of Sciences*, 113(15), 4081–4085. <https://doi.org/10.1073/pnas.1600366113>
- Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M. C., Normand, S., Öckinger, E., Schmidt, N. M., ... Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88(1), 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>

## Data Accessibility Statement:

The SSN files including information on species presence/absence and on the spatial topology of the stream network and sampling sites are provided in Dryad (<https://doi.org/10.5061/dryad.mw6m905tb>), together with the R scripts and additional files needed to repeat all analysis using the SSN package. The raw data on amphibian occurrences are available from the authors upon request.



## 5.7 Supplementary Material

### Supplementary Methods: Workflow detailing the methodological procedures

# Pre-processing of geographical river basin data:

The catchment of the river Sabor and its river segments were extracted from the Catchment Characterisation and Modelling (CCM) river network database (Vogt et al., 2007), available from the web page: <http://ccm.jrc.ec.europa.eu/>.

In ArcGIS, we clipped the Sabor river basin using the CATCHMENTS shapefile:

We selected the Sabor river catchment polygons using the tool *Select by Attributes* using the expression - "WSO6\_ID" = 442389, and then we exported the selected elements to a new shape file;

We used the new shape file to select the river segments using the tool *Select by Location*, and then we exported the selected elements to a new shape file.

To generate the points for model predictions, we split the stream network in segments with a maximum length of 1000 meters:

We used the *Generate Points Along Lines* tool, which is available in the Data management toolbox under "sampling". We then set to 1000 meters the specific distance between points along the lines;

We used the points from the previous step to split the stream network using the *Split Line at Point tool*, which is in the Data Management toolbox under "Features";

We computed the segments centroids of the split segments using the *Feature to Point* tool, under the same toolbox.

## # Environmental variables

We characterized each point sampled for amphibians and each point used to predict amphibian distributions using four environmental variables: elevation [Alt], total annual precipitation [Prec], Strahler's stream order [SO] and the probability of water presence during the dry season [Water]. We also used two variables describing the potential for biotic interactions between amphibians and either *Procambarus clarkii* [Pclar] or *Pacifastacus leniusculus* [Plen].

We estimated elevation [alt] from a DEM with a 10-meter resolution built from digitalized 1:25,000 topographic maps.

We extracted total annual precipitation [Prec] from WorldClim 2 data set with a 30' ( $\approx$  1km) resolution (Fick & Hijmans, 2017) available at <http://worldclim.org/version2>;

We took Strahler's stream order [SO] from the CCM 2.1 data set.

We computed the probability of each stream segment holding water in the summer [Water] based in 189 site observations in 2012 in previous study (Ferreira et al., 2016).

As biotic variables we used the probability of *P. clarkii* [Pclar] or *P. leniusculus* [Plen] occurring in stream segments, estimated from the geostatistical distribution models computed by Filipe et al. (2017), which were based on electrofishing data collected on 167 sites in 2012.

To extract variables values for each point described in [4]:

For variables in raster formats (Alt and Prec), we used the *Extract Multi Values to Points*, in the Spatial Analyst Tools, under "Extraction";

For variables in vector format (SO, Water, Pclar and Plen), we used the *Spatial Join* tool, in the Analysis Tools toolbox under "Overlay".

Before analysis, the values of all variables were standardised to zero mean and unit standard deviation, by subtracting each value from the overall mean of the observed sites and divide by their standard deviation.

# Simulation of crayfish expansion

We used the geostatistical models of crayfish distribution provided by Filipe et al. (2017) to simulate the potential expansion of *Procambarus clarkii* and *Pacifastacus leniusculus* across the Sabor watershed. Simulations were based on the idea that expansion of a species will occur towards stream reaches where its probability of occurrence is estimated to be higher by the model. In practice, this was implemented by raising the probability of occurrence estimated at a given site assuming increases in the relative risk ( $R_r$ ) of the species occurring at that site. The relative risk was defined as the odds ratio of the probabilities of crayfish occurrence under future and current conditions, where odds are the ratio of the probability of occurrence and the probability of absence. Given a relative risk,  $R_r$ , the new probability of occurrence under an expansion scenario is computed by applying the expression:

$$\frac{\frac{p'_i}{(1-p'_i)}}{\frac{p_i}{(1-p_i)}} = R_r \Leftrightarrow p'_i = \frac{R_r \cdot p_i}{1 + p_i(R_r - 1)}, \text{ for } R_r = \{2,3,5\}$$

where  $p_i$  and  $p'_i$  are the probabilities of crayfish occurrence at present and in the future at site  $i$ .

Using the expression in [12], we simulated for each species three expansion scenarios, corresponding to two, three and five-fold increases of the relative risk ( $R_r$ ). We also considered the isolate and joint effects of expansion of just one or both species, respectively, corresponding to the 16 scenarios provided in Table 5.Sa.

Table 5.Sa: Summary of crayfish expansion scenarios analysed in this study

Relative Risk Scenarios		Pacifastacus leniusculus			
		1x	2x	3x	5x
Procambarus clarkii	1x	C1L1	C1L2	C1L3	C1L5
	2x	C2L1	C2L2	C2L3	C2L5
	3x	C3L1	C3L2	C3L3	C3L5
	5x	C5L1	C5L2	C5L3	C5L5

## # STARS processing

To produce the files to analyse in R with SSN, we used STARS, a toolbox for ArcGIS package available at: [https://www.fs.fed.us/rm/boise/AWAE/projects/SSN\\_STARS/software\\_data.html](https://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/software_data.html).

We followed the tutorial provided by the authors of the STAR toolbox. In each step, we will refer the number the section in the tutorial between brackets. The tutorial is available at: [https://www.fs.fed.us/rm/boise/AWAE/projects/SSN\\_STARS/downloads/STARS/STARS\\_tutorial\\_2.0.7.pdf](https://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/downloads/STARS/STARS_tutorial_2.0.7.pdf).

We imported the Sabor river shapefile produced in step [2.b] into a Landscape Network geodatabase (LSN file) using the *Polyline to Landscape Network* tool under “Pre-processing” (section 7). Because the CCM 2.1 dataset is a topologically correct river network, we did not need to check topological errors (section 8).

We used watershed areas to weight the relative influence of the branching upstream segments. We accumulated the attribute CATCHEMENT of CCM2.1 in the stream network using the *Accumulate Values Downstream* tool under “calculate” (section 12).

To incorporate the observations and all the predictions point sets, we used the *Snap Points to Landscape Network* tool. Because the locations of the observed sites were taken on the field, we had to set a search radius (100 meters) and manually check if the sites were snapped in the right location on the stream network. This was not needed for predictions as these points were generated on the stream network (section 14).

We assigned the accumulated watershed area computed in step [17] to observed and predictions points, using *Watershed Attributes* tool (section 15).

We computed the distance between the outlet and:

each stream segment (*Upstream Distance – Edges* tool);

each point (observed sites and prediction points; *Upstream Distance – Sites* tool).

(section 16)

To compute the spatial weights needed to fit the spatial model, we:

Calculated the segment proportional influence related to accumulated watershed area (*Segment PI* tool) (section 17).

Calculated the additive function in:

Stream segments (*Additive Function – Edges*);

Sites (observed and predictions; Additive Function – Sites).

(section 18)

Because the set of observed sites is different for each species, we repeated the previous process and exported a SSN file for each species. In each SSN file there is a representation the topology of the Sabor stream network, spatial data to build the geostatistical models (observations and variables), and spatial data to make predictions (16 crayfish expansion scenarios). All SSN files are together with the R scripts and associated files (see Tables 5.Sb and 5.Sc) are available in Dryad (<https://doi.org/10.5061/dryad.mw6m905tb>), and they can be used to repeat all analysis in R using the SSN package (section 19).

Table 5.Sb: List of R scripts and SSN files available in Dryad (<https://doi.org/10.5061/dryad.mw6m905tb>) and that can be used to repeat all analysis in R using the SSN package.

Name	Description
Script.R	R script with instructions for model selection and prediction.
start.rdata	R data file with objects to facilitate model selection and prediction (table 3).
IsnAc.ssn	SSN file with site observations of <i>Alytes cisternasii</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnAo.ssn	SSN file with site observations of <i>Alytes obstetricans</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnBs.ssn	SSN file with site observations of <i>Bufo spinosus</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnLb.ssn	SSN file with site observations of <i>Lissonotriton boscai</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnPp.ssn	SSN file with site observations of <i>Pelophylax perezi</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnRi.ssn	SSN file with site observations of <i>Rana iberica</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnSs.ssn	SSN file with site observations of <i>Salamandra</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.
IsnTm.ssn	SSN file with site observations of <i>Triturus marmoratus</i> in the Sabor river basin, variables for model building and selection, and 16 crayfish expansion scenarios.

Table 5.Sc: List of R objects available through Dryad (<https://doi.org/10.5061/dryad.mw6m905tb>) and that can be used together with SSN files and R scripts (Table B) to repeat all analysis in R using the SSN package.

Name	Description
CorMdl	A list with the 180 possible spatial structures that can be fitted in function <code>glmssn</code> of the package <i>SSN</i> .
Sps	A character vector with species abbreviations for use in loops.
Sps.sh	A shorter character vector with species abbreviations for use in loops.

#### # Model building in R using the SSN package

To carry out the statistical analysis we produced an R script that is provided in *Amphibians&Crayfish.zip* available in Dryad (<https://doi.org/10.5061/dryad.mw6m905tb>). This script can run properly, without any editing, if the .zip file is extracted to a drive named “C” in Windows. We inserted in the script a reference to each step of this document.

We fitted the geostatistical models using the SSN package (Ver Hoef et al., 2014). For model selection we used the MuMIn package (Barton, 2016). To handle the spatial data in R we used the `rgdal` package (Bivand et al., 2019). We computed model evaluation metrics using the `modEVA` package (Barbosa et al., 2016).

We imported the SSN file and computed the distance matrix. We also extracted the table with species observations and variables to a spatial data frame object.

We built 3 sets of saturated logistic regression models (full models) with all available variables:

Environmental variables (i.e., Alt, Prec, SO and Water);

Biotic variables (i.e., Pclar and Plen);

Environmental + Biotic.

We fitted all possible combination of variables with the `dredge` function (Barton, 2016) and selected the configuration with the lowest AICc for the 3 sets of models (best models).

For the best Environmental + Biotic model, we extracted the residuals and computed the empirical semivariogram (`Torgegram`).

Using the `glmssn` function (Ver Hoef et al., 2014), we fitted all possible spatial structures to the best Environmental + Biotic model residuals. We retained the spatial structure that minimized the AICc. This step can take about one hour for each species.

Using the `glmssn` function, we built the Environmental, Biotic, and Environmental + Biotic model using the variable configurations selected in step [27]. We computed model residuals using leave-one-out cross-validation procedure (Ver Hoef et al., 2014). With the `AUC` and `threshMeasures` functions (Barbosa et al., 2016) we computed the AUC, Koen's Kappa, and true skill statistics (TSS) using the prevalence as a threshold.

We built the Environmental + Biotic + Spatial model by combining the variable configuration of the Environmental + Biotic model selected in step [27] with spatial structure that best fitted the residuals selected in step [29]. We then computed model residuals using leave-one-out cross-validation procedure (Ver Hoef et al., 2014). With the `AUC` and `threshMeasures` function (Barbosa et al., 2016) we computed the AUC, Koen's Kappa, and true skill statistics (TSS) using the prevalence as a threshold.

#### # Predicting species occurrences under current and simulated crayfish expansion scenarios

We used the models produced in the previous steps, together with the 16 crayfish expansion scenarios (Table S1), to predict the distributions of each amphibian species at present and in the future. For each species and expansion scenario, probabilities of species occurrence produced by the models were transformed into estimated presences/absences using current prevalence as the threshold for species presence.

A first set of estimated distributions was produced using the Environmental + Biotic models built in step [30], thereby excluding the spatial (geostatistical) component (i.e., step [29]). This involved the following procedure:

We imported each of the prediction points produced in step [11] using the function `importPredpts` and created downstream hydrologic distances matrices between points using the `createDistMat` function.

We used the `predict.glmssn` function to predict for each of the set of predictions points.

We transformed the predictions values from the logit scale to the probabilistic scale.

We assigned presence to the stream segment under the current scenario if the predicted value is higher than the observed prevalence calculated on step [33].



Predictions of species distributions were estimated using exactly the same procedure as in [33] using the Environmental + Biotic + Spatial models built in step [31].

We saved an .rdata file for each species. For each species, we manually exported to a point shapefile using the writeOGR function the predicted occurrences for each of the 16 expansion scenario for Spatial and Non-Spatial models (8 species x 2 models x 16 scenarios = 256 shapefiles).

# Mapping predicted species distributions under current and simulated crayfish expansion scenarios

We combined the predictions in the point shapefiles with the splitted stream network we had produced in step [3] using the *Spatial Join* tool, in the Analysis Tools toolbox under “Overlay” on ArcGis.

For the species where the models achieved a good discrimination ability ( $AUC \geq 0.80$ ), we mapped the distribution of amphibians refuges by combining the scenario with no increase in the relative risk of crayfish occurrence (the base scenario, “C1L1”) with the most extreme scenario of crayfish invasion (“C5L5”):

We classified stream segments with predicted absences in both scenarios as unsuitable habitat;

We classified stream segments with predicted presences in both scenarios as suitable habitat held under the most extreme scenario of crayfish invasion;

We classified stream segments with predicted presences in the base scenario (“C1L1”) and predicted absences in the extreme scenario as suitable habitat that will possible be lost under crayfish invasion;

# References cited in the workflow

Barbosa, A. M., Brown, J. A., Jiménez-Valverde, A., & Real, R. (2016). *modEvA: Model Evaluation and Analysis* (1.3.2). CRAN. <https://cran.r-project.org/package=modEvA>

Barton, K. (2016). *MuMIn: Multi-Model Inference* (1.43.17). CRAN. <https://cran.r-project.org/web/packages/MuMIn>

Bivand, R., Keitt, T., & Rowlingson, B. (2019). *rgdal: Bindings for the “Geospatial” Data Abstraction Library* (1.4-4). CRAN. <https://cran.r-project.org/package=rgdal>

Ferreira, M., Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish

distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>

Filipe, A. F., Quaglietta, L., Ferreira, M., Magalhães, M. F., & Beja, P. (2017). Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks. *Biological Invasions*, 19(10), 2899–2912. <https://doi.org/10.1007/s10530-017-1492-3>

Ver Hoef, J. M., Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *Journal of Statistical Software*, 56(3), 1–45. <https://doi.org/10.18637/jss.v056.i03>

Vogt, J., Soille, P., De Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., Bódis, K., Dusart, J., Paracchini, M. L., Haastруп, P., & Bamps, C. (2007). *A pan-European River and Catchment Database*. <https://doi.org/10.2788/35907>

## Supplementary Tables

Table 5.S1 - Amphibian species detected in surveys carried out in the Sabor watershed (NE Portugal) in January-December 2015. For each species we indicate the family, number of sites where it was detected (S), number of sites used in distribution modelling (N) and observed prevalence (S/N). We also report the number of sites where it was detected in each sampling month. Data from months highlighted in grey were used to build the geostatistical distribution models.

Species	S	N	S/N	Month											
				Jan	Feb	Mar	Apr	Jun	Jul	Ago	Set	Oct	Nov	Dec	
<b>ANURA</b>															
<i>Alytes cisternasii</i>	27	168	0.2	2	4	3	5	3	3	0	1	6	6	1	
<i>Alytes obstetricans</i>	23	145	0.2	0	0	1	4	2	11	3	2	1	2	1	
<i>Discoglossus galganoi</i>	1	-	-	0	0	1	0	0	0	0	0	0	0	0	
<i>Bufo spinosus</i>	14	134	0.1	0	0	2	3	5	2	0	3	0	0	0	
<i>Hyla molleri</i>	2	-	-	0	0	0	0	2	1	0	0	0	0	0	
<i>Pelophylax perezi</i>	116	168	0.7	5	14	20	23	24	24	18	19	16	14	5	
<i>Rana iberica</i>	43	168	0.3	3	6	6	8	8	4	7	6	6	12	2	
<b>URODELA</b>															
<i>Lissotriton boscai</i>	19	160	0.1	1	2	0	6	0	4	3	1	0	6	1	
<i>Salamandra salamandra</i>	40	142	0.3	5	5	12	6	2	0	1	0	5	10	8	
<i>Triturus marmoratus</i>	23	152	0.2	0	1	1	2	6	7	6	3	0	0	0	

Table 5.S2 - Percentage decline in the length of streams available for each amphibian species, in relation to expansion scenarios for the crayfish *Procambarus clarkii* (Pc) or *Pacifastacus leniusculus* (Pl), or both. Future distributions were estimated considering two (2x), three (3x), and five-fold (5x) increases in the odds ratio of the probabilities of crayfish occurrence under future and current conditions, where odds are the ratio of the probability of occurrence and the probability of absence (see the main text and Figs. 5.S3 and 5.S4 for details).

Full Model					Environment + Biotic Model				
	PI (1x)	PI (2x)	PI (3x)	PI (5x)		PI (1x)	PI (2x)	PI (3x)	PI (5x)
<i>Alytes cisternasii</i>									
Pc (1x)	0.0	16.3	23.6	30.6	Pc (1x)	0.0	34.7	54.9	69.6
Pc (2x)	0.0	16.3	23.6	30.6	Pc (2x)	0.0	34.7	54.9	69.6
Pc (3x)	0.0	16.3	23.6	30.6	Pc (3x)	0.0	34.7	54.9	69.6
Pc (5x)	0.0	16.3	23.6	30.6	Pc (5x)	0.0	34.7	54.9	69.6
<i>Alytes obstetricans</i>									
Pc (1x)	0.0	0.0	0.0	0.0	Pc (1x)	0.0	0.0	0.0	0.0
Pc (2x)	2.6	2.6	2.6	2.6	Pc (2x)	22.5	22.5	22.5	22.5
Pc (3x)	4.2	4.2	4.2	4.2	Pc (3x)	39.1	39.1	39.1	39.1
Pc (5x)	6.8	6.8	6.8	6.8	Pc (5x)	51.3	51.3	51.3	51.3
<i>Rana iberica</i>									
Pc (1x)	0.0	8.6	15.4	19.6	Pc (1x)	0.0	15.8	32.3	44.4
Pc (2x)	2.3	9.7	16.0	19.7	Pc (2x)	9.5	24.7	39.1	47.5
Pc (3x)	3.7	10.1	16.5	19.9	Pc (3x)	16.0	35.9	43.7	50.0
Pc (5x)	5.5	11.9	17.4	20.6	Pc (5x)	31.2	45.7	49.6	53.9
<i>Salamandra salamandra</i>									
Pc (1x)	0.0	5.3	8.9	14.1	Pc (1x)	0.0	5.0	8.5	18.5
Pc (2x)	3.3	7.2	10.3	15.7	Pc (2x)	0.6	5.7	9.5	20.3
Pc (3x)	5.1	9.4	13.0	18.6	Pc (3x)	2.2	6.8	11.6	22.9
Pc (5x)	10.1	14.1	17.4	22.3	Pc (5x)	4.9	9.4	12.4	24.1

## Supplementary Figures

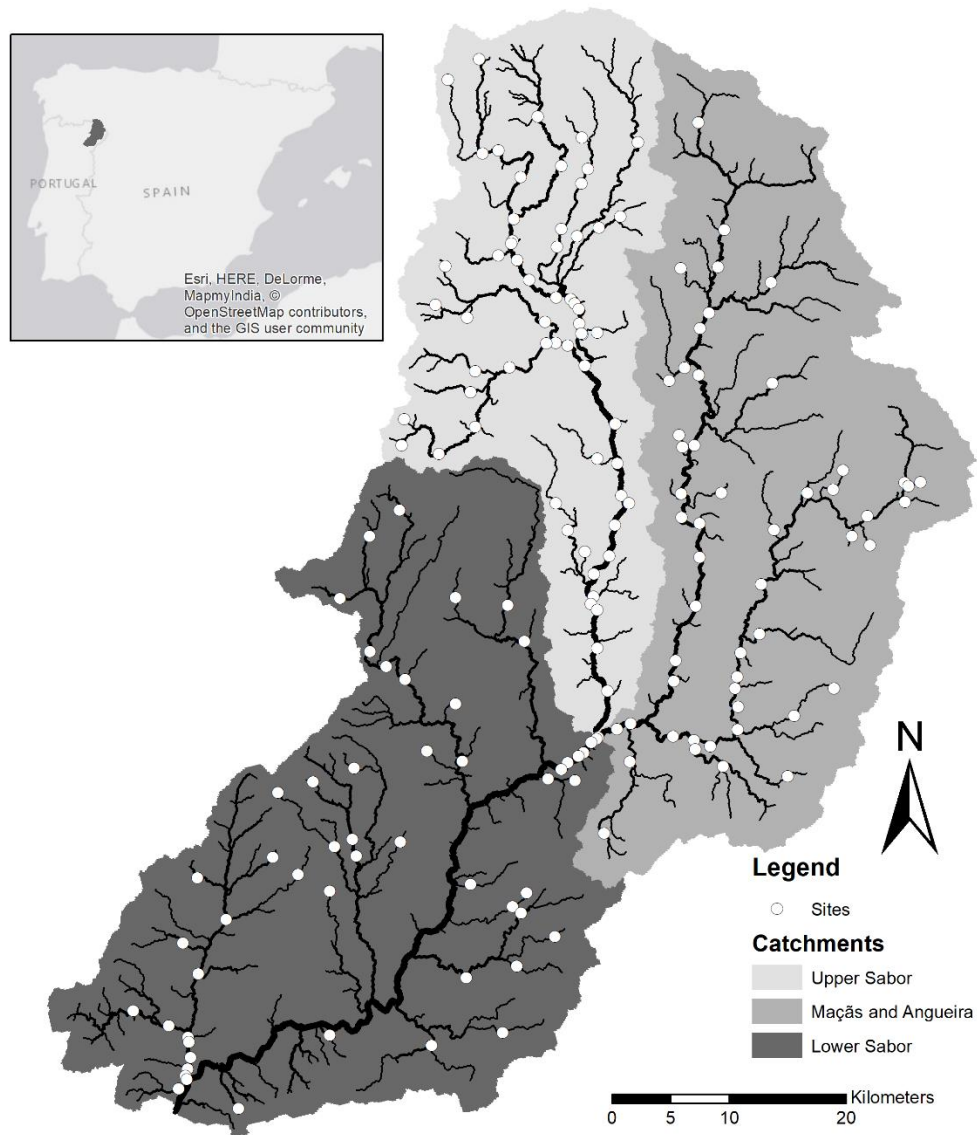


Figure 5.S1 - Map of the study area showing The Sabor watershed (NE Portugal) and the sites sampled for amphibians in January-December 2015. The watershed is divided in three sub-basins that were used to set the stratified sampling design.

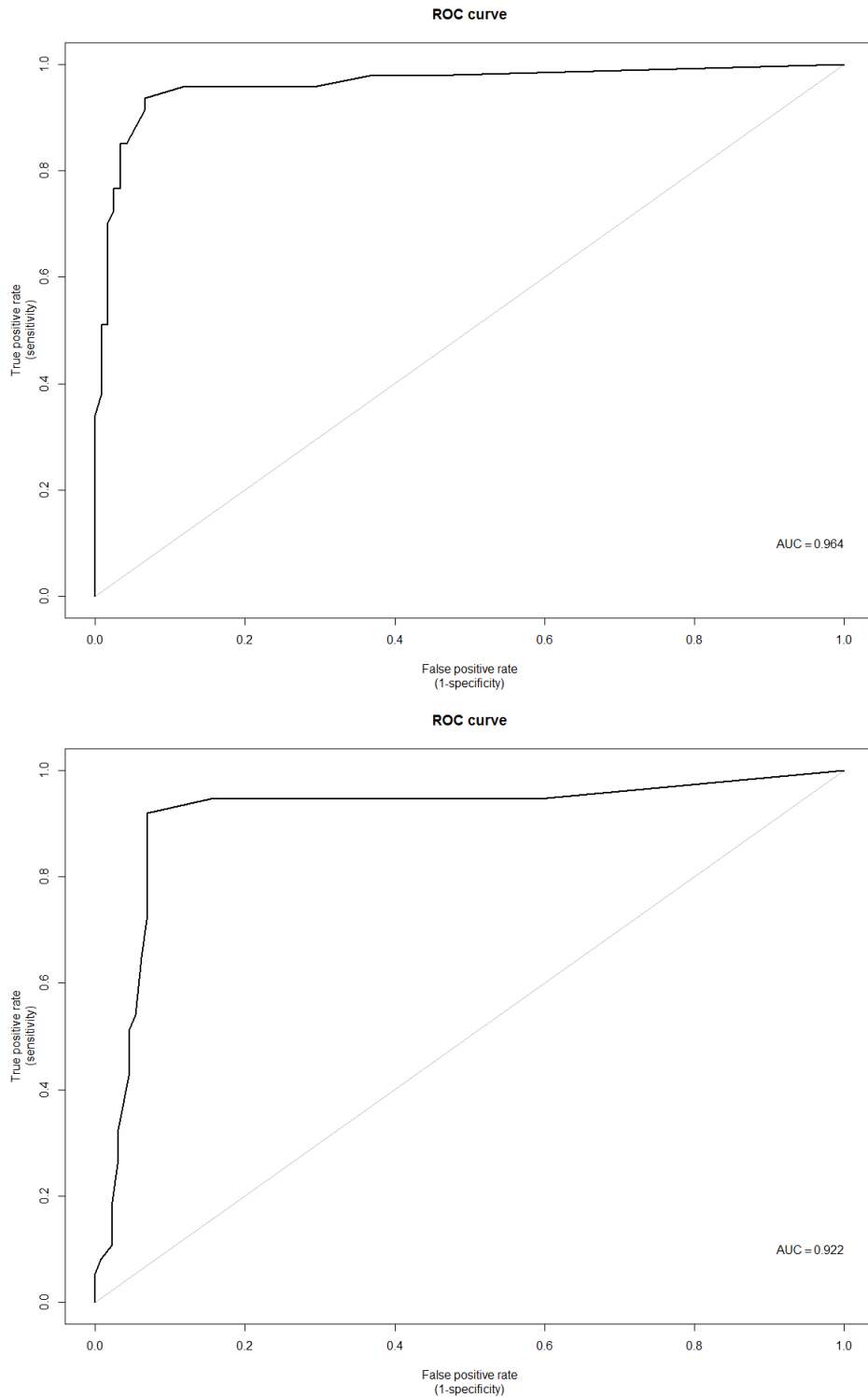


Figure 5.S2 - Receiver operating characteristic (ROC) curve and the area under the curve (AUC) values for the geostatistical model of Filipe et al. (2017) for *Procambarus clarkii* and *Pacifastacus leniusculus*, respectively, using observations recorded collected by the authors during the transects sampled in 2015. Model performance was computed using the modEvA package in R (Barbosa et al. 2018).

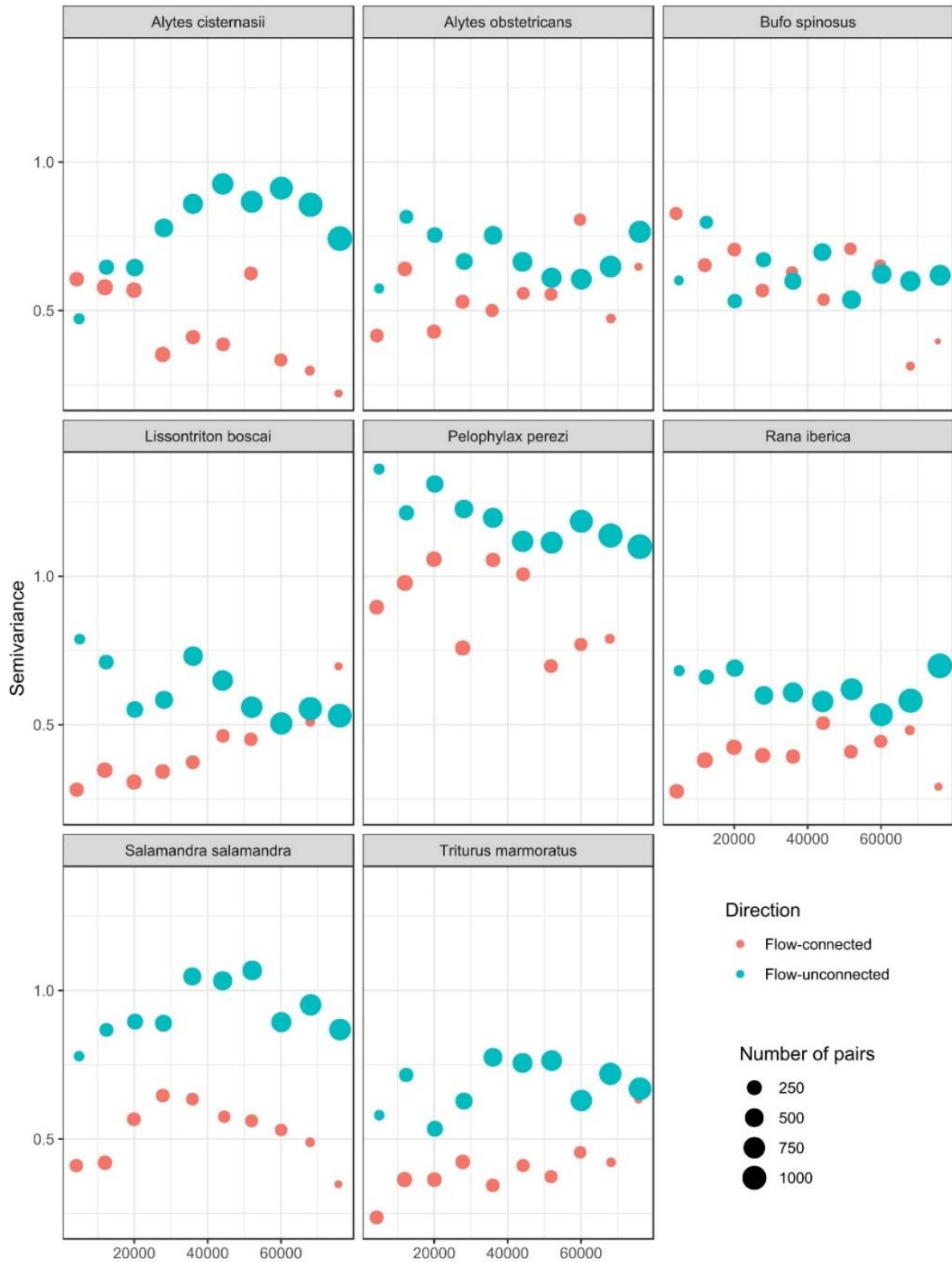


Figure 5.S3 - Torgegrams showing spatial dependencies in the residuals of models relating the probability of amphibian occurrence to variables reflecting environmental factors and biotic interactions (Table 1). Spatial dependencies are quantified as the semivariance in model residuals between either flow connected or flow unconnected sites, in relation to the hydrologic distances separating such sites. An increase in semivariance with lag distance is a signal of spatial autocorrelation.

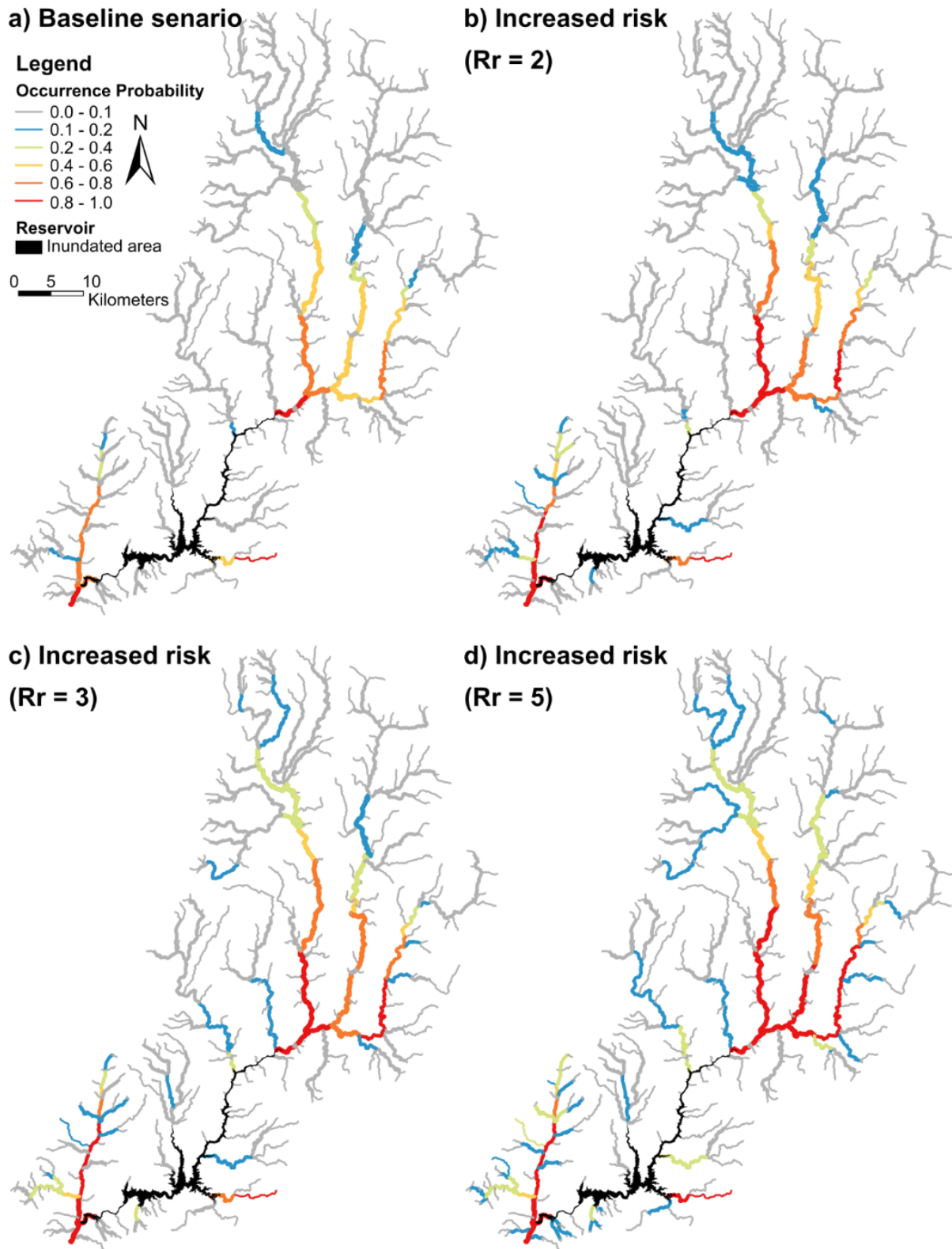


Figure 5.S4 - Predicted distribution of *Procamburus clarkii* across the Sabor watershed (NE Portugal) at present and under different invasion scenarios, estimated using the geostatistical model of Filipe et al. (2017). Future distributions were estimated considering two, three, and five-fold increases in the relative ( $R_r$ ) risk of crayfish occurrence at each stream segment in relation to the baseline scenario corresponding to the predicted distribution of each species in 2012 (Filipe et al., 2017). See the main text for details.



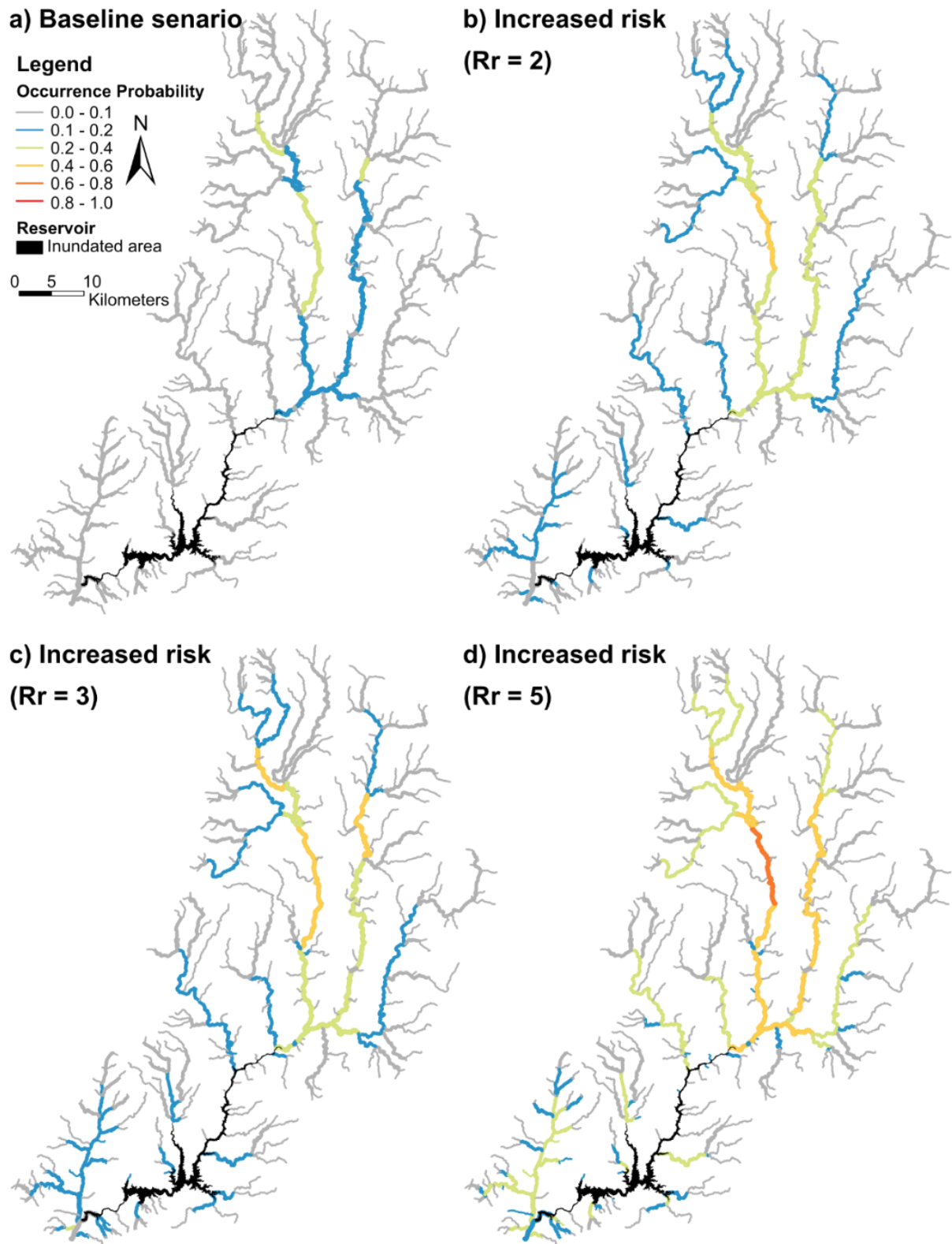


Figure 5.S5 - Predicted distribution of *Pacifastacus leniusculus* across the Sabor watershed (NE Portugal) at present and under different invasion scenarios, estimated using the geostatistical model of Filipe et al. (2017). Future distributions were estimated considering two, three, and five-fold increases in the relative risk of crayfish occurrence at each stream segment in relation to the baseline scenario corresponding to the predicted distribution of each species in 2012 (Filipe et al., 2017). See the main text for details.



## Chapter 6 - Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change





## Chapter 6 - Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change

**Mário Mota-Ferreira**, Ana Filipa Filipe, Maria Filomena Magalhães, Sara Carona, Pedro Beja

This chapter was published in the journal *Diversity and Distributions* and should be cite as:

**Mota-Ferreira, M.**, Filipe, A. F., Magalhães, M. F., Carona, S., & Beja, P. (2021). Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change. *Diversity and Distributions*, 27(2), 313–326. <https://doi.org/10.1111/ddi.13189>

### Keywords

Community dynamics, community trajectory analysis, dendritic networks, exotic species, geostatistics, hydropower generation, stream ecology

### Abstract

**Aim:** Understanding temporal changes in aquatic communities is essential to address the freshwater biodiversity crisis. In particular, it is important to understand the patterns and drivers of spatial variation in local community dynamics, generalising temporal trends from discrete locations to entire landscapes that are the main focus of management. Here we present a framework for producing spatially continuous views of community dynamics, focusing on stream fish affected by hydropower development.

**Location:** River Sabor, NE Portugal

**Methods:** We sampled stream fish at thirty sites between 2012 and 2019. Community trajectory analysis was used to quantify the directionality and velocity of community change, and the geometric resemblance of community trajectories between sites. Geostatistical models for

stream networks were used to relate metrics describing community dynamics to environmental variables, while controlling for Euclidean and hydrologic spatial dependencies, and to map spatial variation in community dynamics across the watershed.

**Results:** Trajectories in multivariate space underlined strong temporal dynamics, with local communities deviating and returning to previous states, but without evidence for directional changes. Accordingly, directionality values were low and not consistently affected by environmental variables. The velocity of community change varied markedly across the watershed and it was strongly affected by stream order and elevation, with faster changes observed in lowland streams draining into hydroelectric reservoirs and with a high proportion of exotic species. Pairwise distances between community trajectories were strongly related to hydrologic and environmental distances between sites.

**Main conclusions:** Local stream fish communities were in a loose equilibrium across the watershed, but they fluctuated at a faster rate closer to a hydroelectric reservoir. Integrating community trajectory analysis and geostatistical modelling provides a relatively simple framework to understand how, where and why temporal community dynamics vary across dendritic stream networks, and to visualise spatial patterns of community change over time in relation to anthropogenic impacts.

## 6.1 Introduction

In freshwater ecosystems, biological communities are changing worldwide due to multiple anthropogenic pressures (Albert et al., 2020; Reid et al., 2019). These changes involve for instance species extinctions, defaunation and taxonomic homogenization, which tend to occur faster at smaller spatial scales, but then scale-up to entire watersheds, regions and even continents (Magalhães et al., 2007; Matthews & Marsh-Matthews, 2016, 2017; Villéger et al., 2011; Zbinden, 2020). Therefore, much effort has been devoted to understanding where, how and why local freshwater biological communities change over time, usually through studies conducted for extended periods at a number of discrete locations (e.g., Baranov et al., 2020; Bêche et al., 2009; Erős et al., 2020; Jourdan et al., 2018; Matthews & Marsh-Matthews, 2016). However, relatively few studies have investigated the patterns and drivers of spatial variation in the temporal dynamics of local community, though this would be important to generalize patterns from particular sites to landscapes and regions that are the main focus of management (Erős & Lowe, 2019; Hugueny et al., 2010; Schlosser, 1991; Wiley et al., 1997). Therefore, it is essential to provide conservation and water agencies with spatially continuous views of community

dynamics, thereby contributing to assess anthropogenic impacts and to prioritize management action (Cid et al., 2020; Fausch et al., 2002).

Changes in local biological communities may result from extinctions or colonizations affecting species richness and composition, and from variations in species abundances (Grossman et al., 1990; Magalhães et al., 2007; Matthews & Marsh-Matthews, 2017). Moreover, changes may follow distinct temporal patterns, being for instance gradual or saltatory, or reflecting variations around loose equilibria, shifts between alternative stable states or gradual directional transitions away from initial community structures (Collins, 2000; DeAngelis et al., 1985; Matthews & Marsh-Matthews, 2017). These changes can be quantified using relatively simple metrics, such as Kendall's coefficient of concordance to estimate constancy in species rank abundances, or the coefficient of variation to estimate variability in species abundances (Grossman et al., 1990). However, these measures do not reveal patterns in temporal change, which have often been inferred by visual examination of trajectories in a chosen space of community resemblance (Magalhães et al., 2007; Matthews et al., 2013). Based on this idea, De Cáceres et al. (2019) developed an analytical framework for describing and comparing community trajectories, with their geometric properties providing information on the level and type of community dynamics. For instance, the lengths and speed of trajectories can be used to quantify how much the community changes over time and whether changes are gradual or abrupt, while direction can quantify whether changes are directional or not (De Cáceres et al., 2019). Moreover, the temporal convergence/divergence of trajectories at pairs of sites can be used to quantify whether communities are varying in synchrony or converging/diverging from each other (De Cáceres et al., 2019). All these metrics can be used to model spatial patterns in community dynamics.

To generalize a community dynamics metric obtained at discrete locations to a spatially continuous surface, it is necessary to find variables that (i) are correlated with variation in that metric and (ii) can be easily mapped at the landscape scale. This may be difficult when community dynamics reflect mainly idiosyncratic variations in local conditions (Erős & Lowe, 2019; Matthews & Marsh-Matthews, 2017), which are hard to extrapolate at larger spatial scales. However, generalization is possible when community dynamics are affected by large scale gradients, such as for instance the upstream-downstream gradients in rivers, or gradients related to sources of human disturbance (Gavioli et al., 2019; Gorman & Karr, 1978; Milardi et al., 2019; Schlosser, 1987). Moreover, local dynamics can be influenced by spatial connectivity associated for instance with the topology of stream networks, anthropogenic barriers and habitat

fragmentation (Cañedo-Argüelles et al., 2020; Crabot et al., 2020; Erős & Lowe, 2019; Hugueny et al., 2010), as it affects meta-community mass effects mediated by dispersal (Heino et al., 2015; Tonkin et al., 2018), as well as the spread of invasive species (Filipe et al., 2017; Gavioli et al., 2019; Milardi et al., 2019; Mota-Ferreira & Beja, 2020). Therefore, spatial modelling of community dynamics requires establishing relations with environmental variables predicting variation in dynamics metrics, and accounting for spatial variables reflecting the effects of connectivity. In the case of rivers, recently developed geostatistical models provide a convenient framework to undertake such modelling exercise, as they account for the complex topology of spatial relations in dendritic networks (Peterson et al., 2013), integrating (Euclidean) spatial dependencies that occur overland, as well as (hydrological) spatial dependencies along the river network and the effects of flow connection (Peterson & Ver Hoef, 2010; Ver Hoef & Peterson, 2010).

Here we combine community trajectory analysis (De Cáceres et al., 2019) and geostatistical modelling (Peterson et al., 2013) to understand and map spatial patterns of community temporal dynamics in dendritic stream networks. We focused on stream fish communities (*sensu* Matthews & Marsh-Matthews, 2017) in a watershed where a hydroelectric development was built and started operating during the study (Jackson, 2011; Santos et al., 2017). We expected communities to be more unstable and eventually undergoing directional changes in lotic reaches close to hydroelectric reservoirs, mainly due to the spread of exotic species (Santos et al., 2017). To test this idea, we used data from stream fish monitoring carried out at 30 sites, encompassing from the construction phase (2012-2014), through the filling of the reservoir (2014-15), to the operation phase (2015-19), and we (i) described community variation in terms of species composition, richness and abundances; (ii) quantified dynamics in terms of the velocity and directionality of community change; (iii) modelled community trajectory metrics in relation large scale ecological gradients (e.g., stream order, elevation), interannual variation in local environmental conditions (e.g., water flow and depth), and the prevalence of exotic species; (iv) investigated spatial patterns in community change in relation to environmental and spatial factors; and (v) developed predictive geostatistical models accounting for large scale ecological gradients and spatial dependencies to produce spatially continuous maps of community dynamics.



## 6.2 Methods

### 6.2.1 Study area

The study was conducted in NE Portugal, in the River Sabor watershed (N41°09'–42°00', W7°15'–6°15'; Figure 1), encompassing a wide range of variation in elevation (100–1500 m above sea level), annual precipitation (443–1163 mm), and mean annual temperature (6.9–15.6 °C). Climate is Mediterranean, with precipitation largely concentrated in October–March, and virtually none in hot summer months (June–August). Flow regime is highly seasonal, with most headwater streams drying out or being reduced to pools in summer, while the main watercourse and the largest tributaries are permanent. Two hydroelectric dams (Feiticeiro: 181 ha; Baixo Sabor: 2 820 ha) located near the mouth of the River Sabor started to be built in 2009, with the main reservoir filling in autumn/winter of 2014/2015 (Jackson, 2011; Santos et al., 2017). The Sabor watershed and its fish communities are more thoroughly described by Ferreira et al. (2016) and Santos et al. (2017).

### 6.2.2 Fish sampling

As part of a preliminary fish survey (Ferreira et al., 2016), a total of 184 sites thoroughly covering the Sabor watershed were visited in the summer of 2012, of which 30 were selected for long term monitoring (Figure 6.1). Sites were stream reaches 50-m long and  $10.0 \pm 5.9$  m [Mean $\pm$ SD] (range: 1.5 – 30.0 m) wide, located at a nearest distance of  $11,162 \pm 4,430$  m (6,186 – 26,204 m) from each other, and representing the main ecological gradients across the watershed (Ferreira et al., 2016). A relatively small sampling reach was chosen because (i) we were interested in investigating how local community dynamics varied across the watershed, (ii) communities are more dynamic at finer spatial grains (Zbinden, 2020), and (iii) previous studies demonstrated this reach length to be adequate for capturing responses of Mediterranean stream fish communities to environmental fluctuations (Magalhães et al., 2007). Sites were sampled annually from 2012 to 2019 in June-July (Table 6.S1), when reduced water flows favoured sampling efficiency, but before the peak summer drought when harsh conditions might cause high fish mortality. Each reach was electrofished by the same operator (MMF) using procedures detailed in Ferreira et al. (2016), with consistent effort and methods at each site over the years. Reaches were electrofished for 15 to 25 minutes, with longer surveys in wider and deeper streams to enhance detectability of all species in the local fish communities (Ferreira et al., 2016). Fish were identified to species level, measured for total length, and returned alive to the stream. Sampling was conducted under license from the Instituto da Conservação da Natureza e Florestas, which required individuals of exotic species to be euthanized.

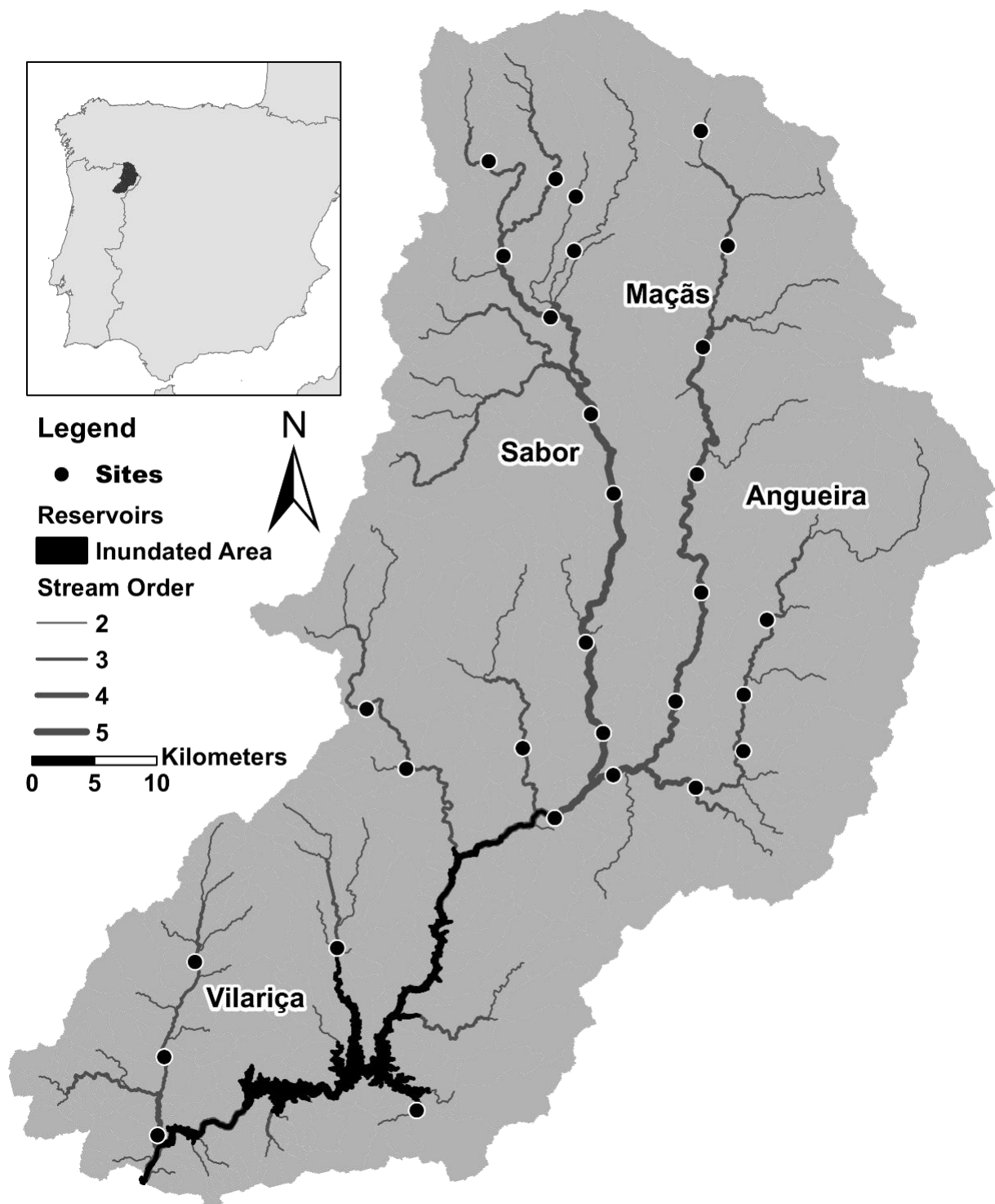


Figure 6.1 - Map of sites sampled for fish in 2012-2019 in the Sabor watershed (NE Portugal).

### 6.2.3 Environmental and spatial variables

Each sampling site was characterised using five large scale variables retrieved from digital maps (mean elevation [Alt], maximum slope [Slope], altitudinal gradient [Alt\_Grad], Strahler's stream order [SO], and annual precipitation [Prec]), which have already been used in predictive species distribution modelling in the Sabor watershed (Ferreira et al., 2016; Filipe et al., 2017; Mota-Ferreira & Beja, 2020; Quaglietta et al., 2018). These variables were extracted using the CCM 2.1 database (Vogt et al., 2007), following procedures detailed in Table 6.S2 in Supporting Information. We only used precipitation extracted from WORLDCLIM (Hijmans et al., 2005), because climatic variables tend to be highly intercorrelated and correlated with physiography (e.g., elevation). Sites were also characterized at the local scale, using the coefficients of variation (CV) of stream width [Width\_cv] and depth [Depth\_cv], and water velocity [Vel\_cv], which were estimated in the field following procedures detailed in Table 6.S2. We used CV rather than mean values, because we wanted to evaluate how community dynamics was affected by interannual variation in local conditions, and because mean values were correlated with landscape variables such as stream order. To account for the possibility of biological invasions increasing community variability (Erős et al., 2020), we computed the proportion of exotic fish species at each site [Exot] (Table 6.S2).

Spatial data necessary to account for spatial autocorrelation (see below) were obtained in a GIS using the stream network extracted from CCM2.1, and the layer of sampling sites. Estimates included the Euclidean and hydrologic distances (total and downstream hydrologic distances) between each pair of sites (Peterson & Ver Hoef, 2010). To deal with confluences in tail-up models (see below), we also estimated catchment areas to weight the relative influence of the branching upstream segments (e.g. Peterson & Ver Hoef, 2010). Spatial estimates were made using the package "riverdist" (Tyers, 2017) in R, and the Spatial Tools for the Analysis of River Systems (STARS) toolbox version 2.0.0 (Peterson & Ver Hoef, 2014) for ArcGIS 10.2 (ESRI, 2016).

### 6.2.4 Community trajectory analysis and modelling

Community temporal variation was quantified using the Community Trajectory Analysis (CTA) framework (De Cáceres et al., 2019), implemented in "vegclust" (De Cáceres et al., 2010). Preliminary analyses were carried out considering community variation in terms of either species composition (i.e., presence/absence) or abundances, considering or not size (i.e., length) structure (De Cáceres et al., 2013). Results were very similar across analyses, so we present only the results based on fish abundances, without size structure. We used counts of each

species at each site and year instead of species densities, because the later can be driven by changes in fish numbers, habitat area, or both (Magalhães et al., 2007). In all analysis, we excluded fish <5 cm, often corresponding to young of the year (yoy), because they were poorly sampled and their recruitment to the fishing gear might be strongly influenced by the timing of sampling in relation to fish spawning.

CTA considers community dynamics as trajectories in a given space of community resemblance, using trajectories as objects to be analysed and compared based on their geometry (De Cáceres et al., 2019). Analysis started by using the Bray-Curtis dissimilarity index (Legendre & Legendre, 2012) to compute the distance matrix among the states (observations) of the community. Using this matrix, we performed a Principal Coordinates Analysis (PCoA; Legendre & Legendre, 2012), and we projected the trajectory of each site in the biplot formed by the first two axis retrieved by the PCoA. We then computed (i) the mean velocity of community change (i.e., the mean length of trajectories between consecutive years divided by the number of observation years minus one); (ii) the directionality of each trajectory; and (iii) the pairwise distances between community trajectories. The mean velocity measured how fast the community changed during the study period, while directionality was used to assess whether the community changed over time following a directional pattern or otherwise showed cyclic or random patterns. The pairwise dissimilarity in trajectories was used to understand whether spatial patterns in community temporal variation were associated to spatial environmental patterns and spatial autocorrelation.

To quantify the environmental factors driving temporal community change, we first used generalized linear models (GLM) with Gaussian errors and identity link, to relate the mean velocity and directionality of trajectories to environmental variables. Pairwise scatterplots were visualized to check for potential outliers and influential points. We found that the single site of order 6 was a potential influential point, and so it was combined with order 5 sites in a single category. We then screened the pairwise relationships between dependent and predictor variables, considering both linear and non-linear relations using orthogonal polynomials of second degree. More complex relations (i.e., higher order polynomials) were not considered because of relatively small sample sizes. In subsequent multivariate model building, we considered for each predictor either the linear or polynomial terms that provided the best fit to the data, judged considering the adjusted R-squared and the Akaike Information Criteria corrected for small sample sizes (AICc). We then screened all combinations of predictors for each dependent variable and, in each case, we retained as best model the combination of predictors minimizing AICc (Murtaugh, 2009).

We also modelled variation in pairwise distances between community trajectories as a function of environmental and spatial distances between sites. First, we computed environmental

pairwise distances considering all standardized environmental variables, which were summarized using the symmetric multidimensional scaling implemented in “smacof” (De Leeuw & Mair, 2009). Then, we computed Mantel correlograms to assess the scale of spatial dependencies in community trajectories and environmental conditions (Legendre et al., 2015), using “vegan” (Oksanen et al., 2012). Finally, we computed a multiple linear regression on distance matrices (MRM; Lichstein, 2007) with “ecodist” (Goslee & Urban, 2007) and 100 000 permutations, using combinations of environmental and spatial distances between sites, and retaining the model with the largest  $R^2$ . MRM was used, despite the pitfalls of Mantel-based approaches, because these are still considered adequate to analyse dissimilarity matrices, and interpretation was made considering potential problems such as inflated Type I errors (Legendre et al., 2015). All analysis were performed using R software (R Core Team, 2019).

### 6.2.5 Geostatistical modelling and mapping

We used geostatistical modelling to relate variables describing community dynamics to both environmental and spatial predictors, considering the spatial structure of dendritic stream networks (Peterson et al., 2013; Peterson & Ver Hoef, 2010; Ver Hoef et al., 2006; Ver Hoef & Peterson, 2010). These geostatistical models are similar to conventional linear mixed models, with specification in random errors of spatial dependencies as functions of either straight-line distances (Euclidean model) between sites, hydrologic distances between sites connected by the water flow (tail-up model), or hydrologic distances irrespective of water flow connection (tail-down model). The fixed component corresponded to the best linear models (GLM) developed in previous analysis for the mean velocity and directionality. The random component was specified considering the full autocovariance structure, which provides the greatest flexibility for representing multiple types of autocorrelation simultaneously (Ver Hoef & Peterson, 2010). To select the best autocovariance function for each spatial component, we tested all combinations of functions for the models including the three spatial components and selected the one minimising AICc.

To map spatial variation in community dynamics, we projected the mean velocity and directionality of community change predicted from the geostatistical models on the stream network of the entire Sabor watershed. First, we divided the stream network into segments of a maximum length of 1000 meters using ArcGIS desktop (ESRI, 2016), and we extracted the value of environmental variables from the centroid of each segment. We then predicted the values of the metrics in each segment using universal kriging within the ‘SSN’ package (Ver Hoef et al., 2014).

## 6.3 Results

### 6.3.1 Overall assemblage patterns

From 2012 to 2019, we sampled 16,733 fishes with total length >5cm from 14 species, of which six were native and eight were exotic (Table 6.). Nearly 90% of the fish caught belonged to just four native (*Luciobarbus bocagei*, *Squalius carolitertii*, *Pseudochondrostoma duriense*, *Squalius alburnoides*) and two exotic (*Lepomis gibbosus*, *Gobio lozanoi*) species with individual catches >5%. Another four species accounted for >1% of catches each (*Achondrostoma* sp., *Salmo trutta*, *Cobitis paludica*, and *Alburnus alburnus*), while the remaining four species (*Gambusia holbrooki*, *Micropterus salmoides*, *Rutilus rutilus*, *Carassius auratus*) were scarce. There was wide variation across sites in total species richness ( $3.9 \pm 1.7$  [mean $\pm$  SD], range: 2-9) and in mean catches per year ( $7.0 \pm 6.3$  to  $183.5 \pm 123.4$ ) (Table 6.S1).

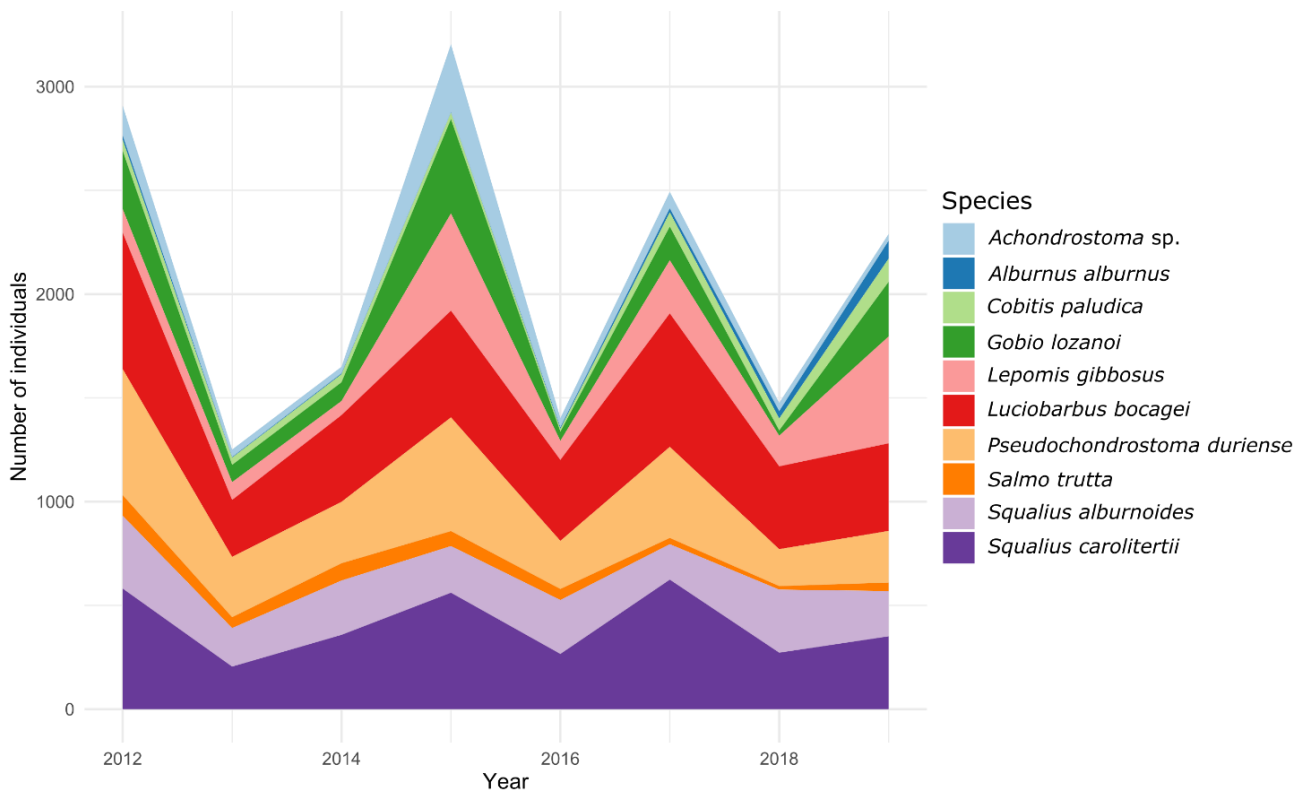


Figure 6.2 - Temporal variation in annual fish catches (2012-2019) in the Sabor watershed (NE Portugal). From the 14 species recorded, four were captured in very small numbers ( $N < 50$  individuals each) and are not represented (see Table 6.1).

Table 6.1 - Fish species recorded through electrofishing in the River Sabor watershed (NE Portugal) between 2012 and 2019. For each species we indicate its status in the region (native [N] vs. exotic [E] status, the total, mean ( $\pm$  SD), range and coefficient of variation (CV) of number of individuals > 5 cm collected per year, the percentage of sites where it was detected (% Sites) and the percentage of years (% Years) when it was detected.

Family/Species	Status	Counts			CV	% Sites (n=30)	%Years (n=8)
		Total	Mean $\pm$ SD (Range)				
<b>Cobitidae</b>							
<i>Cobitis paludica</i>	E <sup>a</sup>	412	51.5 $\pm$ 29.9 (14-112)	0.6	33.30%	100.00%	
<b>Cyprinidae</b>							
<i>Achondrostoma</i> sp. <sup>b</sup>	N	738	92.2 $\pm$ 102.5 (32-328)	1.1	33.30%	100.00%	
<i>Alburnus alburnus</i>	E	176	22 $\pm$ 28.2 (2-86)	1.3	26.70%	100.00%	
<i>Carassius auratus</i>	E	1	0.1 $\pm$ 0.3 (0-1)	2.8	3.30%	12.50%	
<i>Gobio lozanoi</i>	E	1,401	175.1 $\pm$ 147.4 (25-454)	0.8	43.30%	100.00%	
<i>Luciobarbus bocagei</i>	N	3,717	464.6 $\pm$ 131.7 (274-657)	0.3	80.00%	100.00%	
<i>Pseudochondrostoma duriense</i>	N	2,841	355.1 $\pm$ 157.4 (178-608)	0.4	80.80%	100.00%	
<i>Rutilus rutilus</i>	E	2	0.2 $\pm$ 0.7 (0-2)	2.8	3.30%	12.50%	
<i>Squalius alburnoides</i>	N	1,977	247.1 $\pm$ 60 (171-351)	0.2	73.30%	100.00%	
<i>Squalius carolitertii</i>	N	3,218	402.2 $\pm$ 162.7 (205-624)	0.4	90.00%	100.00%	
<b>Salmonidae</b>							
<i>Salmo trutta</i>	N	447	55.8 $\pm$ 27.4 (17-99)	0.5	33.30%	100.00%	
<b>Centrarchidae</b>							
<i>Lepomis gibbosus</i>	E	1,748	218.5 $\pm$ 178.9 (68-515)	0.8	56.60%	100.00%	
<i>Micropterus salmoides</i>	E	5	0.6 $\pm$ 1.7 (0-5)	2.8	6.70%	12.50%	
<b>Poeciliidae</b>							
<i>Gambusia holbrooki</i>	E	50	6.2 $\pm$ 8 (0-21)	1.3	16.70%	75.00%	
<b>TOTAL CATCHES</b>		16,733	2,091.6 $\pm$ 744.3 (1259-3206)	0.4			
<b>SPECIES RICHNESS</b>		14	11.1 $\pm$ 0.6 (10-12)	0.1			

<sup>a</sup> The species is native in to the Iberian Peninsula, but it was considered exotic in our study area based on Doadrio et al. (2011); but see Sánchez-Hernández et al. (2018)

<sup>b</sup> The species was listed as *Achondrostoma arcasii* by Ferreira et al. (2016) but probably belongs to an undescribed species (Robalo et al., 2006)

### 6.3.2 Changes in species richness and fish catches

The annual number of species recorded in the Sabor watershed remained essentially constant at 10-12 species, but the number of individuals captured (i.e., fish catches) varied widely (Table 6., Figure 6.2). Variability over time, as measured by the coefficient of variation, was highest ( $CV > 1$ ) for rare species (Mean catch per year  $< 100$  individuals) such as *C. auratus*, *R. rutilus*, *G. holbrooki*, *A. alburnus* and *Achondrostoma* sp. Pearson correlations with sampling year indicated no consistent temporal trends in species richness ( $r = 0.05$ ,  $P = 0.915$ ), overall catches ( $r = -0.10$ ,  $P = 0.822$ ) and catches of native ( $r = -0.36$ ,  $P = 0.389$ ) and exotic species ( $r = 0.35$ ,  $P = 0.400$ ). Temporal trends in catches of individual species were negative for *S. trutta* ( $r = -0.80$ ,  $P = 0.016$ ), positive for *A. alburnus* ( $r = 0.69$ ,  $P = 0.054$ ) and non-significant for all the other ( $-0.577 < r < 0.579$ ,  $P > 0.10$ ).

### 6.3.4 Community trajectories

The fish community trajectories represented in the PCoA biplot indicated major variations across sites, albeit without obvious temporal patterns (Figure 6.3). There was no visual evidence for directional changes, with communities deviating and later returning to previous states. This was supported by the small directionality values ( $0.33 \pm 0.04$ , which varied little (0.25-0.43) across the watershed. In univariate analysis for directionality, the best relations were mostly linear, but there were quadratic relations for elevation, stream order and the CV of water depth (Supplementary Table 6.S3). However, all relations were weak and statistically non-significant. The best multivariate model (Supplementary Table 6.S4), included a weak U-shaped effect of elevation, suggesting slightly higher directionality at the lowest and highest elevations, and showed a slightly increase in directionality along with slope (Table 6., Figure 6.4a).

The velocity of community change at each site varied between 0.29 and 0.89 ( $0.54 \pm 0.15$ ). In univariate analysis, the best relations were always linear, except for the quadratic relation with stream order (Supplementary Table 6.S3). Some univariate effects were statistically significant, with mean velocity declining linearly with elevation, and showing quadratic relations with stream order, with faster changes in 3<sup>rd</sup> orders and smaller in 2<sup>nd</sup>, 4<sup>th</sup> and particularly 5/6<sup>th</sup> order streams, and with the proportion of exotics, with increases up to about 0.8 and levelling off or slightly declining thereafter. The best multivariate model accounted for 61% of variation, showing faster changes at lower elevation irrespective of stream order, while at any given elevation the velocity of change increased from 5<sup>th</sup> to 3<sup>rd</sup> orders, while declining again slightly in 2<sup>nd</sup> order streams (Table 6., Figure 6.4b; Supplementary Table 6.S5).



Mantel correlograms showed that pairwise distances between community trajectories were significantly related to Euclidean distances up to about 15 km, and to hydrologic distances up to about 30 km (Figure 6.S1). Environmental and spatial distances were also related to each other, albeit weakly (Figure 6.S2). The best MRM model ( $F= 137.018$ ,  $P < 0.001$ ) accounted for ~39% of variation in community trajectories between sites, underlining significant effects of environmental (coefficient =  $1.6 \times 10^{-6}$ ;  $P < 0.001$ ) and hydrologic distances (coefficient = 0.051,  $P < 0.001$ ).

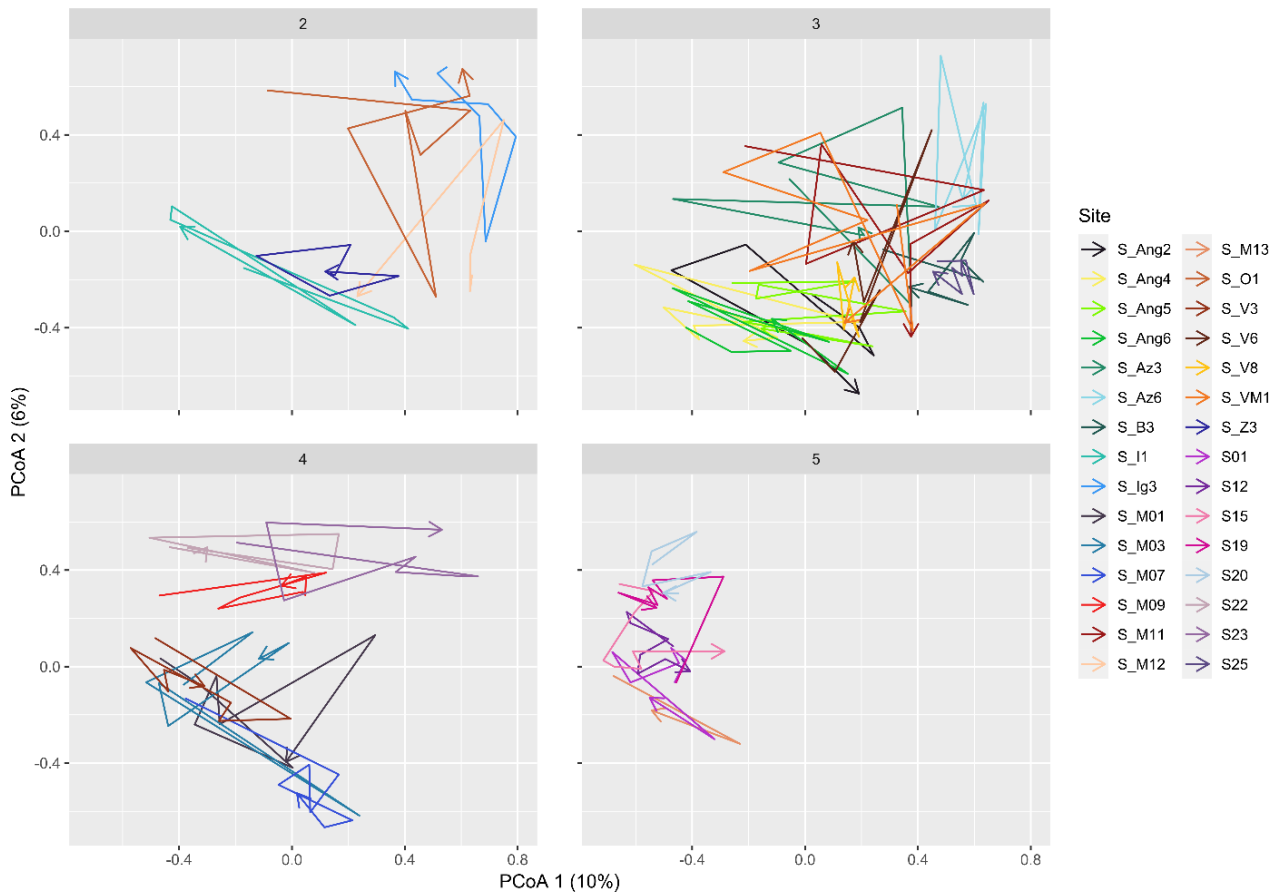


Figure 6.3 - Temporal fish community trajectories at 30 sites sampled in the Sabor watershed (NE Portugal) in the period 2012-2019, represented in the first two axis of a Principal Coordinate Analysis. PCoA was performed on a distance matrix computed with Bray-Curtis index on the matrix of fish catches per species, site, and year. Trajectories are indicated using arrows and, for clarity, sites sampled at second (2); third (3); fourth (4); and fifth and sixth (5) stream orders are represented separately. Different colours in each panel represent trajectories observed at different sampling sites.

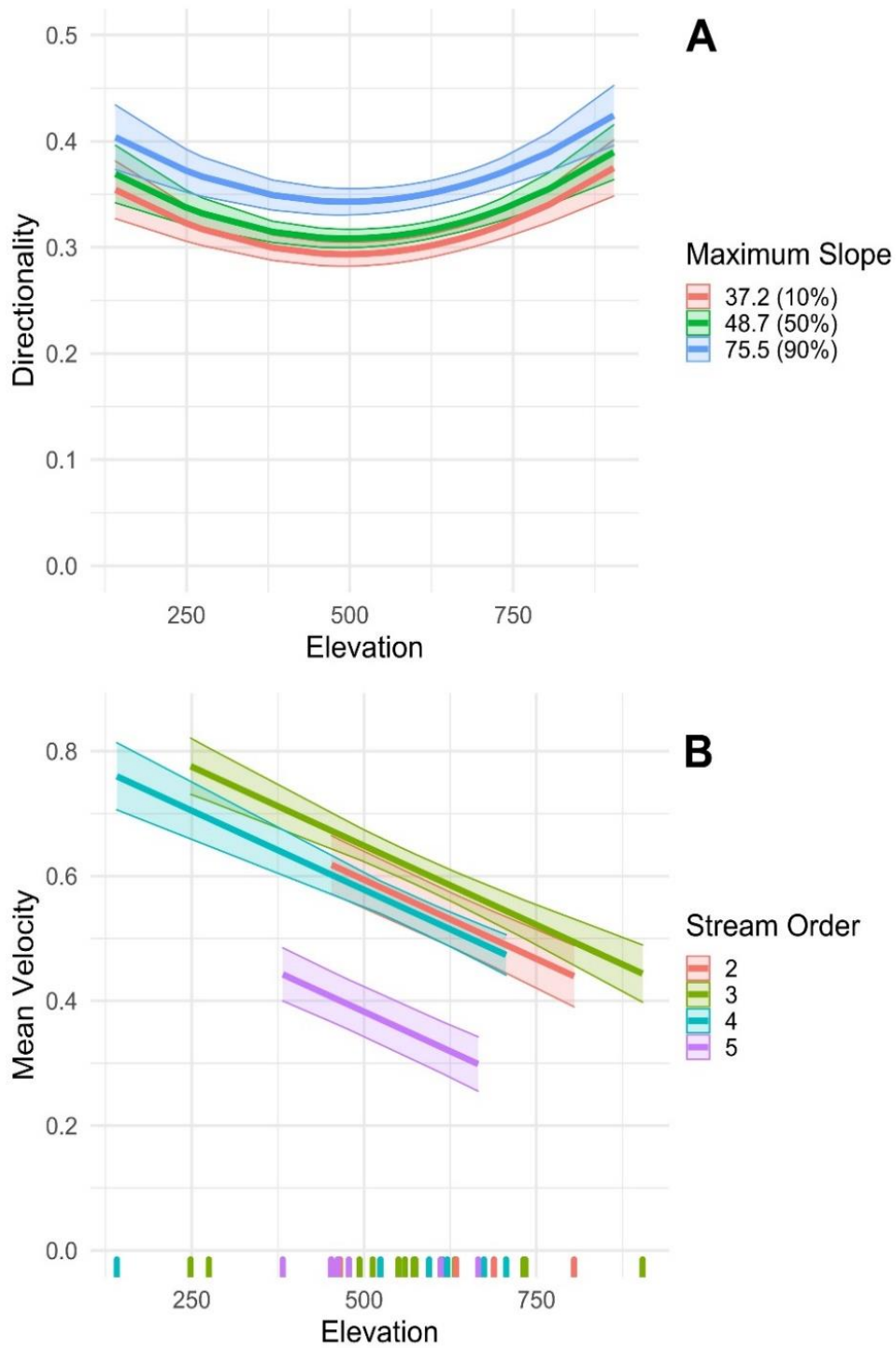


Figure 6.4 - Trend lines ( $\pm$  standard errors) describing the relations inferred from models (Table 2) relating the directionality of community change to elevation for three levels of maximum slope (percentiles 10%, 50% and 90%) (A), and the mean velocity of community change to elevation for each stream order (B).

Table 6.2 - Summary results of the best AICc models explaining variation in mean velocity of community change, directionality and pairwise distances between community trajectories. For the intercept and each variable in each model we provide the regression coefficient, the standard error of the coefficient estimate, and the corresponding *t*- and *P*-values.

	Estimates	Std. Error	<i>t</i>	<i>P</i>
<b>Directionality</b>	(R <sup>2</sup> = 0.37, F <sub>3,26</sub> = 5.064, P < 0.001)			
Intercept	0.364	0.05	7.254	<0.001***
Elevation	-4.8x10 <sup>-4</sup>	1.8x10 <sup>-4</sup>	-2.605	0.015*
Elevation <sup>2</sup>	4.8x10 <sup>-7</sup>	1.7x10 <sup>-7</sup>	2.814	0.009**
Slope	1.3x10 <sup>-3</sup>	4.4x10 <sup>-4</sup>	2.942	0.007**
<b>Mean velocity</b>	(R <sup>2</sup> = 0.61, F <sub>3,26</sub> = 13.584, P < 0.001)			
(Intercept)	0.364	0.248	1.467	0.154
Elevation	-5.1x10 <sup>-4</sup>	1.2x10 <sup>-4</sup>	-4.313	<0.001***
Stream Order	0.367	0.139	2.646	0.014*
Stream Order <sup>2</sup>	-0.062	0.019	-3.242	0.003**

Table 6.3 - Partitioning of the proportion of explained variation by geostatistical linear mixed models relating community directionality and mean velocity of change to environmental variables while accounting for spatial random effects. Spatial components: TU = Tail-up; TD = Tail-down; Eucl = Euclidean.

Community Dynamics metrics	Spatial			Environmental	Nugget
	TU	TD	EUCL		
Directionality	Residual standard error: 0.119; Generalized R <sup>2</sup> : 0.41				
Proportion	0.06	≈ 0	0.54	0.41	≈ 0
Function	Linear with Sill	Mariah	Gaussian		
Range (km)	28.3	71.6	367.4		
Sill	1.4x10 <sup>-3</sup>	6.8x10 <sup>-8</sup>	1.3x10 <sup>-2</sup>		
Mean velocity	Residual standard error: 0.256; Generalized R <sup>2</sup> : 0.27				
Proportion	0.08	0.65	≈ 0	0.27	≈ 0
Function	Spherical	Epanech	Exponential		
Range (km)	16.5	490.7	201.2		
Sill	7.1x10 <sup>-3</sup>	5.8x10 <sup>-2</sup>	4.9x10 <sup>-7</sup>		

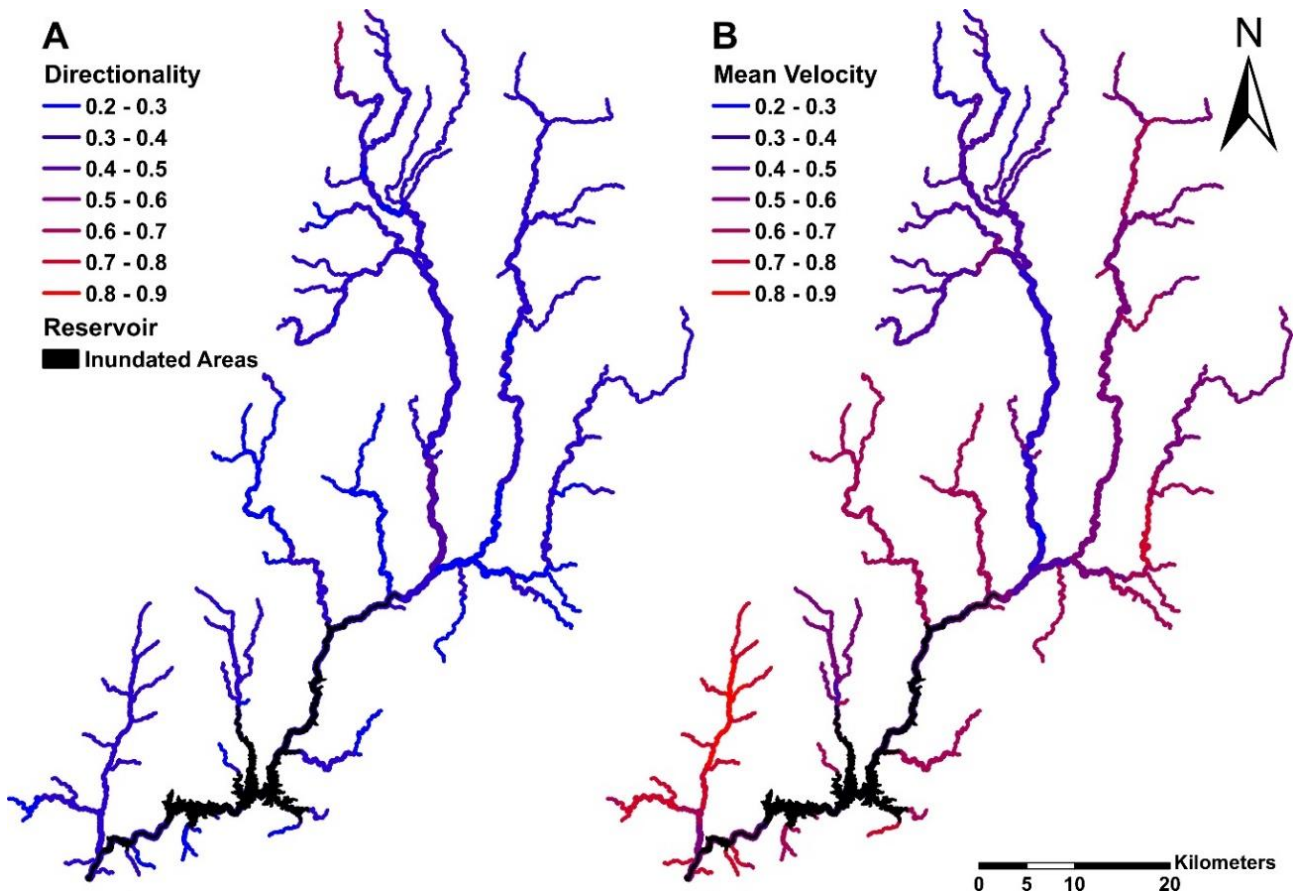


Figure 6.5 - Maps showing variation across the Sabor watershed of the directionality (A) and mean velocity (B) of fish community change, based on geostatistical models with fixed and random components described in Tables 2 and 3.

### 6.3.5 Spatial variation in community dynamics

The geostatistical model explained about 40% of variation in community directionality (Table 6.3). Most explained variation was accounted for by the environmental model (0.41), Euclidean effects (0.54) and, to a much lesser extent, the tail-up effects (0.06). The long range of the Euclidean structure (367.4 km) suggested a large-scale gradient across the watershed, while the range of the tail-up model (28.3 km) pointed out spatial dependencies between flow-connected sites over relatively small distances. The spatial projection of model predictions produced a map showing that directionality was always low, with minor spatial variation across the watershed (Figure 6.5a).

Regarding the mean velocity of community change, the geostatistical model explained close to 30% of its variation across the watershed (Table 6.3), most of which was accounted for by the environmental model (0.27), tail-down effects (0.65) and, to a much lesser extent, tail-up effects (0.08). The long tail-down range (490.7 km) suggested large-scale spatial dependencies along

the hydrologic network irrespective of flow connection, while the short range of the tail-up model (16.5 km) pointed out spatial dependencies over short distances between flow-connected sites. The spatial projection of model predictions showed wide variation in the velocity of community change across the watershed, with highest values in small streams at lower elevation and draining directly into the hydroelectric infrastructure. In the main course of the Sabor, mean velocity was higher immediately upstream of the main reservoir and declined further upstream. The two main tributaries of the Sabor (Maçãs and Angueira), showed a much higher velocity of community change than the main river (Figure 6.5b).

## 6.4 Discussion

Our study shows that combining robust descriptors of community change with state-of-the-art geostatistical modelling contributes to understanding and predicting where, how, and why the dynamics of local communities vary across stream networks. We found that local fish communities varied widely over the years, but there was no evidence for directional changes, pointing out a state of loose equilibrium (*sensu* Matthews et al., 2013) across the watershed. However, there was much spatial variation in the velocity of community changes, which were strongly influenced by environmental gradients associated with elevation and stream order. The spatial patterns of local community dynamics appeared to be affected also by stream network topology, given the strong influence of hydrological spatial dependencies. Mapping of community dynamics highlighted faster changes in lowland streams affected by hydroelectric development and exotic species. Overall, our framework helps to generalise community dynamics from discrete locations to entire watersheds, providing spatial information needed for freshwater ecosystem assessment and management (Cid et al., 2020; Fausch et al., 2002).

### 6.4.1 Temporal drivers of local community change

As in other Mediterranean-type streams (Bêche et al., 2009; Magalhães et al., 2007), fish communities in the Sabor watershed were highly dynamic, with temporal changes involving mainly fluctuations in species abundances, while much less variation was found in species composition and richness. The high dynamism observed was probably affected by the scale at which the study was conducted, as communities tend to show far more marked changes at the local than at the watershed or regional scales (Magalhães et al., 2007; Zbinden, 2020). However, community dynamics were probably also driven by strong environmental fluctuations during the study period, particularly the occurrence of extreme droughts that strongly affect stream fish survival and recruitment (Lennox et al., 2019; Magalhães et al., 2003, 2007;

Matthews & Marsh-Matthews, 2017). High fish mortality probably occurred during severe droughts in the summers of 2012, 2015 and 2017 (Parente et al., 2019), eventually causing the major declines in fish abundances in 2013, 2016 and 2018, respectively. Moreover, fish communities take time to recover after environmental extremes (Bêche et al., 2009; Matthews et al., 2013; Resh et al., 2013), which may have contributed further to the variability observed. The spread of exotic species may also have affected community dynamics, due to temporal changes in their own prevalence and abundance, but also due to eventual negative effects on native species (Erős et al., 2020; Gavioli et al., 2019; Milardi et al., 2019; Zanden et al., 2015). This is supported by faster changes in sites with higher proportion of exotic species, and by the high coefficients of variation in the abundance of exotic species when compared to most native species. Finally, random sampling variation may have contributed to the fluctuations observed, though this was minimized through preliminary testing and optimization of sampling methodologies (Ferreira et al., 2016), and by having the same operator applying a constant sampling effort at every site in every year. Therefore, we are confident that the temporal patterns observed are unlikely to result from methodological artefacts. Overall, the local community dynamics recorded in our study seem to be comparable to that reported elsewhere for Mediterranean-type stream fish (Bêche et al., 2009; Magalhães et al., 2007), but also other aquatic organisms (Bêche et al., 2009; Crabot et al., 2020), suggesting that the patterns observed may apply to other study systems.

#### 6.4.2 Patterns and environmental drivers of spatial variation in community dynamics

Despite the temporal changes observed, there was no evidence for directional community dynamics at any sampling site, suggesting that local fish communities were in a loose equilibrium (Matthews et al., 2013; Matthews & Marsh-Matthews, 2016). This view was supported by consistently small directionality values estimated at sampling sites across the watershed, as well as by the community trajectories in multivariate space, which deviated and returned to previous states, following either gradual or abrupt steps (De Cáceres et al., 2019). Moreover, the significant relations between directionality and environmental variables were weak and difficult to interpret ecologically, possibly reflecting spurious correlations. Comparable patterns have already been described in stream fish communities (Matthews et al., 2013; Matthews & Marsh-Matthews, 2016), including in Mediterranean streams facing extreme droughts and floods (Magalhães et al., 2007), and are in line with the view that communities often exist in a dynamic temporal equilibrium (Collins, 2000; DeAngelis et al., 1985). However,

the patterns may be considered unexpected given the major anthropogenic changes affecting the Sabor watershed, including the building of a large hydroelectric infrastructure and the increasing spread of exotic species, including fish (Santos et al., 2017) and crayfish (Filipe et al., 2017). This may be because the study period was too short to capture community trends, as most exotic species were already present in the watershed at the start of our study (Ferreira et al., 2016), and probably not enough time elapsed for the new dams causing major shifts in species composition in nearby lotic areas. This is supported by observations regarding three exotic species potentially spreading from reservoirs (Ribeiro & Veríssimo, 2014; Vinyoles et al., 2007), one of which has been steadily increasing in abundance since 2015 (*A. alburnus*), while two others only started to be detected towards the end of the study period (*M. salmoides* and *R. rutilus*). Directional changes may thus happen in the future, most likely associated with increasing prevalence and abundance of exotic species (Gavioli et al., 2019; Milardi et al., 2019).

The velocity of local community changes varied widely across the watershed. This spatial variation was strongly related to structural landscape features such as stream order and elevation, while factors associated with interannual variability in local environmental conditions did not show measurable influences. The effect of stream order might be expected, as it reflects strong longitudinal gradients along rivers in for instance water discharge, and habitat size and heterogeneity (Hughes et al., 2011), which are strong drivers of ecological processes (Vannote et al., 1980) and the distribution of organisms (Harrel et al., 1967; Paller, 1994; Platts, 1979), including in the study area (Ferreira et al., 2016; Filipe et al., 2017; Mota-Ferreira & Beja, 2020; Quaglietta et al., 2018). We found that changes were fast in 2<sup>nd</sup> order streams, still a little faster in 3<sup>rd</sup> order streams, and then velocity declined in larger order streams. This is in line with studies suggesting higher temporal changes in fish communities in headwaters than further downstream (Schlosser, 1987), though other studies suggest that longitudinal gradients in stream fish community dynamics may vary across watersheds depending on local environmental conditions (Matthews & Marsh-Matthews, 2017). The later was supported by our study, because the observed joint effect of stream order and elevation implied that there were communities in lower order streams at high elevation that varied slower than those in higher order streams at low elevation. Reasons for these patterns are not completely clear, but it is noteworthy that lower orders at high elevation correspond to cold water mountain streams with species poor communities dominated by *S. trutta*, which may be relatively stable over time. In contrast, lower orders at low elevation generally correspond to warm water streams with richer communities dominated by cyprinids and exotics, which during the dry summer months are often reduced to a series of disconnected pools, and thus where fish communities may vary widely from year to year in association with droughts and floods (Bêche et al., 2009; Magalhães et al., 2007). Other possibility is that elevation acted as a surrogate for increasing human disturbance in the

lowlands driving higher variability in fish communities, which may be mediated by the increasing prevalence and abundance of exotic species (Erős et al., 2020; Gavioli et al., 2019; Gorman & Karr, 1978; Milardi et al., 2019). This idea is supported by the positive relation observed between the proportion of exotic fish and community variability, and by the inverse relation between the prevalence of exotic crayfish and elevation also found in the watershed (Filipe et al., 2017). However, in multivariate models the effect of elevation was retained but not that of exotic species, possibly because the former may capture the effect of the later, and account in addition for unmeasured ecological processes driving community dynamics. Overall, the velocity of community change appeared to be mainly associated with large scale proxies, possibly reflecting spatial gradients in more local ecological processes such as biological invasions, which would require further clarification.

### 6.4.3 Spatial dependencies in community temporal dynamics

The patterns of community temporal change were also related to spatial dependencies, with more similar community trajectories in sites closer to each other, either overland (Euclidean) or along the waterlines (hydrologic). The effects of Euclidean distances were significant up to about 15 km, and possibly reflected similarity between sites in environmental conditions influencing community trajectories. For instance, sites close to each other are likely to be more similar than those farther apart in environmental conditions driven for instance by elevation, which was related to the velocity and, to a lesser degree, the directionality of community change. Hydrologic spatial dependencies were significant up to about 30 km and may be a consequence of similarities between sites associated for instance to stream order and elevation, and to unmeasured spatially structured environmental factors (Legendre & Legendre, 2012). In addition, however, hydrologic spatial dependencies were probably also influenced by mass effects (Heino et al., 2015), with fish dispersal among neighbouring locations homogenizing species composition and synchronizing population fluctuations (Erős & Lowe, 2019; Hugueny et al., 2010; Tonkin et al., 2018). For instance, fish dispersal from larger streams to headwaters may contribute to reduce community fluctuations in the later (Matthews & Marsh-Matthews, 2017). Also, dispersal of exotic fish species across the watershed may contribute to biotic homogenization, and to similarities in community fluctuations in sites nearby (Gavioli et al., 2019; Milardi et al., 2019).

The geostatistical models further supported the importance of spatial effects and allowed a finer examination of their contribution to community dynamics. In the case of directionality, there were only marked Euclidean effects, with a long range, suggesting influences driven by large scale spatial gradients overland. This effect should be interpreted with care, given the low values of



directionality and its low variability across the watershed. Regarding the velocity of community change, spatial dependencies also explained a large share of variability, with a strong contribution of the tail-down model and a minor contribution of the tail-up model. This is in line with the idea that the tail-down model captures spatial dependencies associated with organisms that can move actively both up and downstream, while the tail-up model mainly reflects spatial dependencies resulting from the passive drift of materials or organisms downstream (Peterson et al., 2013). This, together with the observation that tail-down effects were strong even after accounting for the variable reflecting spatially structured variation in environmental conditions along the stream network (i.e., stream order), further suggest that spatial variation in the velocity of community change was affected by movement of individuals along the waterlines. Overall, our results support the idea that fish community dynamics is strongly affected by spatial dependencies and the topology of the stream network (Erős & Lowe, 2019; Hugueny et al., 2010; Tonkin et al., 2018).

#### 6.4.4 Mapping community temporal dynamics to guide management

Mapping of community dynamics highlighted areas across the watershed where larger changes seem to be occurring, some of which may be associated to anthropogenic pressures. In the case of directionality, mapping showed little variation across the watershed, suggesting that at least until now the construction and operation of the Baixo Sabor Hydroelectric Infrastructure did not disrupt the loose equilibrium of fish communities, as there was no evidence for streams closer to the reservoirs showing more directional changes than streams elsewhere in the watershed. It should be noted, however, that our study only encompassed four years after the filling of the larger reservoir, and so it cannot be ruled out that directional changes will become apparent in the long term. The velocity of community change varied across the watershed, with some evidence for faster changes occurring in streams draining into the reservoirs, and in the Sabor river immediately upstream of the reservoir. This suggests that the presence of large reservoirs may be increasing fish community instability in surrounding lotic environments, through for instance the spread of exotic species (Santos et al., 2017). Notwithstanding, the highest velocity of community change was found in a small watershed (Vilariça) that does not drain into the Baixo Sabor reservoirs. This watershed is affected by a number of anthropogenic pressures, draining into another large dam downstream of Baixo Sabor, flowing through an area of intensive agriculture, and being subject to habitat management interventions (Boavida et al., 2018), all of which may have contribute to fast community changes. Overall, the spatially continuous mapping of temporal community dynamics provided a visual representation of the

type and spatial extent of anthropogenic impacts on stream fish communities, which would have been more difficult to perceive otherwise.

## 6.5 Conclusions

Freshwater biological communities are rapidly changing worldwide due to direct and indirect anthropogenic pressures, making it critical to understand how, where and why such changes are occurring (Albert et al., 2020; Reid et al., 2019). Our study combining community trajectory analysis (De Cáceres et al., 2019) and geostatistical modelling (Peterson et al., 2013) contributes to address these issues, by offering a relatively simple and flexible framework to spatially generalize data on community dynamics collected at discrete sampling locations. Using this framework, we were able to show that local dynamics were affected by larger scale processes operating within the stream network, including both environmental gradients and spatial processes mediated by network topology (Erős & Lowe, 2019). Moreover, we produced maps that helped visualizing community changes across the stream network, and that highlighted the effects of a new hydroelectric development in nearby lotic systems. We suggest that our framework may be widely useful to freshwater ecologists aiming to understand spatial variation in local community dynamics under anthropogenic change, while providing a tool for managers to make spatially continuous predictions of community temporal dynamics that can be used in bioassessment and mitigation of anthropogenic impacts on freshwater ecosystems (Cid et al., 2020; Fausch et al., 2002).

## Acknowledgements

This work was funded by EDP and the Portuguese Science and Technology Foundation (FCT). MMF and SC were supported by FCT PhD grants SFRH/BD/95202/2013 and PD/BD/52596/2014, respectively, PB by EDP Biodiversity Chair, and AFF by FRESHING project funded by FCT and COMPETE (PTDC/AAGMAA/2261/2014 – POCI-01-0145-FEDER-356016824). We thank collaboration in fieldwork by Pedro Henrique, Rita Severino, Sara Ivone, Sergio Henrique, Duarte Prata, João Lopes, João Guilherme, José Barradas, Rui Cereja, Giorgio Pace, Ivo Rosa, Miguel Pinto, William Agudelo, Guilherme Buzzo, Filipe Sá, Lorenzo Quaglieta, and Nuno Pereira.

## 6.6 References

- Albert, J. S., Destouni, G., Duke-Sylvester, S. M., Magurran, A. E., Oberdorff, T., Reis, R. E., Winemiller, K. O., & Ripple, W. J. (2020). Scientists' warning to humanity on the freshwater biodiversity crisis. *Ambio*. <https://doi.org/10.1007/s13280-020-01318-8>
- Baranov, V., Jourdan, J., Pilotto, F., Wagner, R., & Haase, P. (2020). Complex and nonlinear climate-driven changes in freshwater insect communities over 42 years. *Conservation Biology*, *0*(0), 1–11. <https://doi.org/10.1111/cobi.13477>
- Bêche, L. A., Connors, P. G., Resh, V. H., & Merenlender, A. M. (2009). Resilience of fishes and invertebrates to prolonged drought in two California streams. *Ecography*, *32*(5), 778–788. <https://doi.org/10.1111/j.1600-0587.2009.05612.x>
- Boavida, I., Jesus, J. B., Pereira, V., Santos, C., Lopes, M., & Cortes, R. M. V. (2018). Fulfilling spawning flow requirements for potamodromous cyprinids in a restored river segment. *Science of the Total Environment*, *635*, 567–575. <https://doi.org/10.1016/j.scitotenv.2018.04.167>
- Cañedo-Argüelles, M., Gutiérrez-Cánovas, C., Acosta, R., Castro-López, D., Cid, N., Fortuño, P., Munné, A., Múrria, C., Pimentão, A. R., Sarremejane, R., Soria, M., Tarrats, P., Verkaik, I., Prat, N., & Bonada, N. (2020). As time goes by: 20 years of changes in the aquatic macroinvertebrate metacommunity of Mediterranean river networks. *Journal of Biogeography*, *April*, 1–14. <https://doi.org/10.1111/jbi.13913>
- Cid, N., Bonada, N., Heino, J., Cañedo-Argüelles, M., Crabot, J., Sarremejane, R., Soininen, J., Stubbington, R., & Datry, T. (2020). A Metacommunity Approach to Improve Biological Assessments in Highly Dynamic Freshwater Ecosystems. *BioScience*, *70*(5), 427–438. <https://doi.org/10.1093/biosci/biaa033>
- Collins, S. L. (2000). Disturbance frequency and community stability in native tallgrass prairie. *American Naturalist*, *155*(3), 311–325. <https://doi.org/10.1086/303326>
- Crabot, J., Heino, J., Launay, B., & Datry, T. (2020). Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. *Ecography*, *43*(4), 620–635. <https://doi.org/10.1111/ecog.04835>
- De Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wiser, S. K., Fortin, M.-J., Condit, R., & Hubbell, S. (2019). Trajectory analysis in community ecology. *Ecological Monographs*, *89*(2), e01350. <https://doi.org/10.1002/ecm.1350>
- De Cáceres, M., Font, X., & Oliva, F. (2010). The management of vegetation classifications with fuzzy clustering. *Journal of Vegetation Science*, *21*(6), 1138–1151. <https://doi.org/10.1111/j.1654-1103.2010.01211.x>
- De Cáceres, M., Legendre, P., & He, F. (2013). Dissimilarity measurements and the size structure of ecological communities. *Methods in Ecology and Evolution*, *4*(12), 1167–1177. <https://doi.org/10.1111/2041-210X.12116>
- De Leeuw, J., & Mair, P. (2009). Multidimensional Scaling Using Majorization: SMACOF in R. *Journal of Statistical Software*, *31*(3), 1–30. <https://doi.org/10.18637/jss.v031.i03>
- DeAngelis, D. L., Waterhouse, J. C., Post, W. M., & O'Neill, R. V. (1985). Ecological modelling and disturbance evaluation. *Ecological Modelling*, *29*(1–4), 399–419. [https://doi.org/10.1016/0304-3800\(85\)90063-8](https://doi.org/10.1016/0304-3800(85)90063-8)
- Doadrio, I., Perea, S., Garzón-Heydt, P., & González, J. L. (2011). *Ictiofauna continental española. Bases para su seguimiento*. (DG Medio Natural y Política Forestal (ed.)). Ministerio de Medio Ambiente y Medio

Rural y Marino, Centro de Publicaciones.

- Erős, T., Comte, L., Filipe, A. F., Ruhi, A., Tedesco, P. A., Brose, U., Fortin, M. J., Giam, X., Irving, K., Jacquet, C., Larsen, S., Sharma, S., & Olden, J. D. (2020). Effects of nonnative species on the stability of riverine fish communities. *Ecography*, 1–11. <https://doi.org/10.1111/ecog.04985>
- Erős, T., & Lowe, W. H. (2019). The Landscape Ecology of Rivers: from Patch-Based to Spatial Network Analyses. *Current Landscape Ecology Reports*, 4(4), 103–112. <https://doi.org/10.1007/s40823-019-00044-6>
- ESRI. (2016). *ArcGIS Desktop 10.5* (10.5). Environmental Systems, Research Institute.
- Fausch, K. D., Torgersen, C. E., Baxter, C. V., & Li, H. W. (2002). Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Fishes. *BioScience*, 52(6), 483. [https://doi.org/10.1641/0006-3568\(2002\)052\[0483:lrbtjg\]2.0.co;2](https://doi.org/10.1641/0006-3568(2002)052[0483:lrbtjg]2.0.co;2)
- Ferreira, M., Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>
- Filipe, A. F., Quaglietta, L., Ferreira, M., Magalhães, M. F., & Beja, P. (2017). Geostatistical distribution modelling of two invasive crayfish across dendritic stream networks. *Biological Invasions*, 19(10), 2899–2912. <https://doi.org/10.1007/s10530-017-1492-3>
- Gavioli, A., Milardi, M., Castaldelli, G., Fano, E. A., & Soininen, J. (2019). Diversity patterns of native and exotic fish species suggest homogenization processes, but partly fail to highlight extinction threats. *Diversity and Distributions*, 25(6), 983–994. <https://doi.org/10.1111/ddi.12904>
- Gorman, O. T., & Karr, J. R. (1978). Habitat Structure and Stream Fish Communities. *Ecology*, 59(3), 507–515. <https://doi.org/10.2307/1936581>
- Goslee, S. C., & Urban, D. L. (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software*, 22(7), 1–19. <https://doi.org/10.18637/jss.v022.i07>
- Grossman, G. D., Dowd, J. F., & Crawford, M. (1990). Assemblage stability in stream fishes: A review. *Environmental Management*, 14(5), 661–671. <https://doi.org/10.1007/BF02394716>
- Harrel, R. C., Davis, B. J., & Dorris, T. C. (1967). Stream Order and Species Diversity of Fishes in an Intermittent Oklahoma Stream. *American Midland Naturalist*, 78(2), 428. <https://doi.org/10.2307/2485240>
- Heino, J., Melo, A. S., Siqueira, T., Soininen, J., Valanko, S., & Bini, L. M. (2015). Metacommunity organisation, spatial extent and dispersal in aquatic systems: Patterns, processes and prospects. *Freshwater Biology*, 60(5), 845–869. <https://doi.org/10.1111/fwb.12533>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/joc.1276>
- Hughes, R. M., Kaufmann, P. R., & Weber, M. H. (2011). National and regional comparisons between Strahler order and stream size. *Journal of the North American Benthological Society*, 30(1), 103–121. <https://doi.org/10.1899/09-174.1>
- Hugueny, B., Oberdorff, T., & Tedesco, P. A. (2010). Community Ecology of River Fishes: A Large-Scale Perspective. In K. B. Gido & D. A. Jackson (Eds.), *Community Ecology of Stream Fishes: Concepts, Approaches, and Techniques* (pp. 29–62). American Fisheries Society.
- Jackson, A. L. R. R. (2011). Renewable energy vs. biodiversity: Policy conflicts and the future of nature

- conservation. *Global Environmental Change*, 21(4), 1195–1208. <https://doi.org/10.1016/J.GLOENVCHA.2011.07.001>
- Jourdan, J., O'Hara, R. B., Bottarin, R., Huttunen, K. L., Kuemmerlen, M., Monteith, D., Muotka, T., Ozoliņš, D., Paavola, R., Pilotto, F., Springe, G., Skuja, A., Sundermann, A., Tonkin, J. D., & Haase, P. (2018). Effects of changing climate on European stream invertebrate communities: A long-term data analysis. *Science of the Total Environment*, 621, 588–599. <https://doi.org/10.1016/j.scitotenv.2017.11.242>
- Legendre, P., Fortin, M., & Borcard, D. (2015). Should the Mantel test be used in spatial analysis? *Methods in Ecology and Evolution*, 6(11), 1239–1247. <https://doi.org/10.1111/2041-210X.12425>
- Legendre, P., & Legendre, L. F. J. (2012). *Numerical Ecology* (3rd ed.). Elsevier.
- Lennox, R. J., Crook, D. A., Moyle, P. B., Struthers, D. P., & Cooke, S. J. (2019). Toward a better understanding of freshwater fish responses to an increasingly drought-stricken world. In *Reviews in Fish Biology and Fisheries* (Vol. 29, Issue 1, pp. 71–92). Springer International Publishing. <https://doi.org/10.1007/s11160-018-09545-9>
- Lichstein, J. W. (2007). Multiple regression on distance matrices: A multivariate spatial analysis tool. *Plant Ecology*, 188(2), 117–131. <https://doi.org/10.1007/s11258-006-9126-3>
- Magalhães, M. F., Beja, P., Schlosser, I. J., & Collares-Pereira, M. J. (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, 52(8), 1494–1510. <https://doi.org/10.1111/j.1365-2427.2007.01781.x>
- Magalhães, M. F., Schlosser, I. J., & Collares-Pereira, M. J. (2003). The role of life history in the relationship between population dynamics and environmental variability in two Mediterranean stream fishes. *Journal of Fish Biology*, 63(2), 300–317. <https://doi.org/10.1046/j.1095-8649.2003.00148.x>
- Matthews, W. J., & Marsh-Matthews, E. (2016). Dynamics of an upland stream fish community over 40 years: Trajectories and support for the loose equilibrium concept. *Ecology*, 97(3), 706–719. <https://doi.org/10.1890/14-2179.1>
- Matthews, W. J., & Marsh-Matthews, E. (2017). Stream fish community dynamics: A critical synthesis. In *Stream Fish Community Dynamics: A Critical Synthesis*. JHU Press.
- Matthews, W. J., Marsh-Matthews, E., Cashner, R. C., & Gelwick, F. (2013). Disturbance and trajectory of change in a stream fish community over four decades. *Oecologia*, 173(3), 955–969. <https://doi.org/10.1007/s00442-013-2646-3>
- Milardi, M., Gavioli, A., Soininen, J., & Castaldelli, G. (2019). Exotic species invasions undermine regional functional diversity of freshwater fish. *Scientific Reports*, 9(1), 17921. <https://doi.org/10.1038/s41598-019-54210-1>
- Mota-Ferreira, M., & Beja, P. (2020). Combining geostatistical and biotic interaction modeling to predict amphibian refuges under crayfish invasion across dendritic stream networks. *Diversity and Distributions*, 26(6), 699–714. <https://doi.org/10.1111/ddi.13047>
- Murtaugh, P. A. (2009). Performance of several variable-selection methods applied to real ecological data. *Ecology Letters*, 12(10), 1061–1068. <https://doi.org/10.1111/j.1461-0248.2009.01361.x>
- Oksanen, A. J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Hara, R. B. O., Simpson, G. L., Solymos, P., & Stevens, M. H. H. (2012). vegan: Community Ecology Package. In CRAN (p. 263). CRAN. <https://doi.org/10.4135/9781412971874.n145>
- Paller, M. H. (1994). Relationships between Fish Assemblage Structure and Stream Order in South

- Carolina Coastal Plain Streams. *Transactions of the American Fisheries Society*, 123(2), 150–161. [https://doi.org/10.1577/1548-8659\(1994\)123<0150:rbfasa>2.3.co;2](https://doi.org/10.1577/1548-8659(1994)123<0150:rbfasa>2.3.co;2)
- Parente, J., Amraoui, M., Menezes, I., & Pereira, M. G. (2019). Drought in Portugal: Current regime, comparison of indices and impacts on extreme wildfires. *Science of the Total Environment*, 685, 150–173. <https://doi.org/10.1016/j.scitotenv.2019.05.298>
- Peterson, E. E., & Ver Hoef, J. M. (2010). A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91(3), 644–651. <https://doi.org/10.1890/08-1668.1>
- Peterson, E. E., & Ver Hoef, J. M. (2014). Stars: An ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *Journal of Statistical Software*, 56(2), 1–17. <https://doi.org/10.18637/jss.v056.i02>
- Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>
- Platts, W. S. (1979). Relationships among Stream Order, Fish Populations, and Aquatic Geomorphology in an Idaho River Drainage. *Fisheries*, 4(2), 5–9. [https://doi.org/10.1577/1548-8446\(1979\)004<0005:rasofp>2.0.co;2](https://doi.org/10.1577/1548-8446(1979)004<0005:rasofp>2.0.co;2)
- Quaglietta, L., Paupério, J., Martins, F. M. S., Alves, P. C., & Beja, P. (2018). Recent range contractions in the globally threatened Pyrenean desman highlight the importance of stream headwater refugia. *Animal Conservation*. <https://doi.org/10.1111/acv.12422>
- R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation Statistical Computing* (3.6.1). R Foundation for Statistical Computing. <https://doi.org/ISBN 3-900051-07-0>
- Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J. J., Kidd, K. A., MacCormack, T. J., Olden, J. D., Ormerod, S. J., Smol, J. P., Taylor, W. W., Tockner, K., Vermaire, J. C., Dudgeon, D., & Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, 94(3), 849–873. <https://doi.org/10.1111/brv.12480>
- Resh, V. H., Bêche, L. A., Lawrence, J. E., Mazor, R. D., McElravy, E. P., O'Dowd, A. P., Rudnick, D., & Carlson, S. M. (2013). Long-term population and community patterns of benthic macroinvertebrates and fishes in Northern California Mediterranean-climate streams. *Hydrobiologia*, 719(1), 93–118. <https://doi.org/10.1007/s10750-012-1373-9>
- Ribeiro, F., & Veríssimo, A. (2014). Full westward expansion of *Rutilus rutilus* (Linnaeus, 1758) in the Iberian Peninsula. *Journal of Applied Ichthyology*, 30(3), 540–542. <https://doi.org/10.1111/jai.12418>
- Sánchez-Hernández, J., Vieira-Lanero, R., Barca, S., Silva, S., Lago, L., Gómez, P., Cobo, M. D. C., & Cobo, F. (2018). An update on the distribution of *Cobitis paludica* (De Buen, 1930) in the NW Iberian Peninsula. *Limnetica*, 37(2), 181–185. <https://doi.org/10.23818/limn.37.15>
- Santos, R. M. B., Sanches Fernandes, L. F., Cortes, R. M. V., Varandas, S. G. P., Jesus, J. J. B., & Pacheco, F. A. L. (2017). Integrative assessment of river damming impacts on aquatic fauna in a Portuguese reservoir. *Science of the Total Environment*, 601–602, 1108–1118. <https://doi.org/10.1016/j.scitotenv.2017.05.255>
- Schlosser, I. J. (1987). A conceptual framework for fish communities in small warmwater streams. In W. J. Matthews & D. C. Heins (Eds.), *Community and evolutionary ecology of North American stream fishes* (pp. 17–28). Island Press.
- Schlosser, I. J. (1991). Fish Ecology : A Landscape Perspective. *BioScience*, 41(10), 704–712.

- Tonkin, J. D., Altermatt, F., Finn, D. S., Heino, J., Olden, J. D., Pauls, S. U., & Lytle, D. A. (2018). The role of dispersal in river network metacommunities: Patterns, processes, and pathways. *Freshwater Biology*, 63(1), 141–163. <https://doi.org/10.1111/fwb.13037>
- Tyers, M. (2017). *riverdist: River Network Distance Computation and Applications* (R package version 0.15.0). CRAN. <https://cran.r-project.org/package=riverdist>
- Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(1), 130–137. <https://doi.org/10.1139/f80-017>
- Ver Hoef, J. M., & Peterson, E. E. (2010). A Moving Average Approach for Spatial Statistical Models of Stream Networks. *Journal of the American Statistical Association*, 105(489), 6–18. <https://doi.org/10.1198/jasa.2009.ap08248>
- Ver Hoef, J. M., Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *Journal of Statistical Software*, 56(3), 1–45. <https://doi.org/10.18637/jss.v056.i03>
- Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13(4), 449–464. <https://doi.org/10.1007/s10651-006-0022-8>
- Villéger, S., Blanchet, S., Beauchard, O., Oberdorff, T., & Brosse, S. (2011). Homogenization patterns of the world's freshwater fish faunas. *Proceedings of the National Academy of Sciences of the United States of America*, 108(44), 18003–18008. <https://doi.org/10.1073/pnas.1107614108>
- Vinyoles, D., Robalo, J., Sostoa, A., Almodóvar, A., Elvira, B., Nicola, G., Fernández-Delgado, C., Santos, C., Doadrio, I., Sardá-Palomera, F., & Almada, V. (2007). Spread of the alien bleak *Alburnus alburnus* (Linnaeus, 1758) (Actinopterygii cyprinidae) in the Iberian Peninsula: The role of reservoirs. *Graellsia*, 63(1), 101–110. <https://doi.org/10.3989/graellsia.2007.v63.i1.84>
- Vogt, J. V., Soille, P., Jager, A. De, Rimavičiūtė, E., Mehl, W., Haastrup, P., Paracchini, M. M. L., Dusart, J., Bódis, K., Foisneau, S., & Bamps, C. (2007). Developing a pan-European Database of Drainage Networks and Catchment Boundaries from a 100 Metre DEM. *Proceedings 10th AGILE Int. Conference on Geographic Information Science, May, 8–11*. [http://ams.jrc.it/publications/pdfs/JVogt\\_etal\\_AGILE2007.pdf](http://ams.jrc.it/publications/pdfs/JVogt_etal_AGILE2007.pdf)
- Wiley, M. J., Kohler, S. L., & Seelbach, P. W. (1997). Reconciling landscape and local views of aquatic communities: Lessons from Michigan trout streams. *Freshwater Biology*, 37(1), 133–148. <https://doi.org/10.1046/j.1365-2427.1997.00152.x>
- Zanden, M. J. Vander, Lapointe, N. W. R., & Marchetti, M. P. (2015). Non-indigenous fishes and their role in freshwater fish imperilment. In G. P. Closs, M. Krkosek, & J. D. Olden (Eds.), *Conservation of Freshwater Fishes* (pp. 238–269). Cambridge University Press. <https://doi.org/10.1017/CBO9781139627085.009>
- Zbinden, Z. D. (2020). Temporal dynamics of stream fish assemblages and the role of spatial scale in quantifying change. *Ecology and Evolution*, 10(2), 952–961. <https://doi.org/10.1002/ece3.5954>

## 6.7 Supplementary Material

### Supplementary Tables

Table 6.S1 - Summary statistics of fish communities sampled at 30 sites in the Sabor watershed, between 2012-2019. For each site we indicate the number of years when sampling was conducted, the species richness and the mean catches per year (total, native and exotic species).

Sites	Years	Species richness			Catches		
		Total	Natives	Exotics	Mean total (SD)	Mean natives (SD)	Mean exotics (SD)
<b>S_Ang2</b>	7	5	4	1	40.8 (48.7)	31.4 (34.7)	9.4 (16.9)
<b>S_Ang4</b>	8	6	4	2	76.1 (50.1)	49.5 (41.5)	26.6 (24.2)
<b>S_Ang5</b>	8	7	4	3	52.6 (44.2)	43.6 (40.7)	9 (15.4)
<b>S_Ang6</b>	8	7	4	3	120.6 (112.1)	58.3 (56.4)	62.4 (80.2)
<b>S_Az3</b>	8	6	4	2	50.5 (52)	32 (31.9)	18.5 (22.1)
<b>S_Az6</b>	8	3	2	1	10.9 (9.9)	10.3 (9.8)	0.6 (1.1)
<b>S_B3</b>	8	3	3	0	23.3 (15.6)	23.3 (15.6)	0.0 (0.0)
<b>S_I1</b>	7	5	3	2	65.1 (70.9)	53 (57.6)	12.1 (23.2)
<b>S_Ig3</b>	8	4	4	0	37.8 (16.9)	37.8 (16.9)	0.0 (0.0)
<b>S_M01</b>	8	7	5	2	33.5 (17.5)	28.1 (12.6)	5.4 (7.2)
<b>S_M03</b>	8	6	4	2	60.3 (32.7)	45.9 (22.3)	14.4 (16.5)
<b>S_M07</b>	8	7	6	1	70.6 (69.6)	11.9 (17.6)	58.8 (72.4)
<b>S_M09</b>	8	7	6	1	109.6 (57.9)	101.5 (58.8)	8.1 (7.3)
<b>S_M11</b>	8	6	6	0	52.1 (79.2)	52.1 (79.2)	0.0 (0.0)
<b>S_M12</b>	7	3	3	0	7.0 (6.3)	7.0 (6.3)	0.0 (0.0)
<b>S_M13</b>	3	7	3	4	26.6 (40.0)	20.0 (31.5)	6.6 (10.5)
<b>S_O1</b>	8	4	4	0	34.1 (33.3)	34.1 (33.3)	0.0 (0.0)
<b>S_V3</b>	8	9	3	6	130.0 (86.8)	63.0 (54.0)	67 (57.2)
<b>S_V6</b>	8	8	3	5	10.5 (8.3)	3.9 (5.5)	6.6 (5.1)
<b>S_V8</b>	6	5	3	2	63.9 (164.6)	33.3 (83.0)	30.6 (81.8)
<b>S_VM1</b>	8	5	4	1	54.8 (60.2)	53.1 (57.4)	1.6 (3.4)
<b>S_Z3</b>	8	6	4	2	28.6 (13.8)	24.9 (12.9)	3.8 (4.8)
<b>S01</b>	8	7	3	4	100.4 (30.5)	68.4 (23.4)	32 (30.3)
<b>S12</b>	8	8	4	4	183.5 (123.4)	144.0 (96.1)	39.5 (36.2)
<b>S15</b>	8	8	4	4	136.9 (60.0)	127.6 (61.5)	9.3 (5.8)
<b>S19</b>	8	6	4	2	176.9 (102.1)	176.8 (102.3)	0.1 (0.4)
<b>S20</b>	8	6	4	2	143.1 (45.3)	142.6 (45.3)	0.5 (1.1)
<b>S22</b>	8	6	5	1	125.3 (59.7)	125.3 (59.7)	0.0 (0.0)
<b>S23</b>	8	5	5	0	49.1 (18.6)	49.1 (18.6)	0.0 (0.0)
<b>S25</b>	8	2	2	0	17.3 (13.5)	17.3 (13.5)	0.0 (0.0)
<b>Total</b>	8	14	6	8	2091.6 (744.3)	1668.8 (530.6)	422.9 (319.8)



Table 6.S2 - Description and summary statistics (mean  $\pm$  SD, range) of environmental variables used to characterize the 30 stream reaches sampled for fish in the Sabor watershed, between 2012 and 2019. Variables with coefficient of variation (Width\_cv, Depth\_cv, and Vel\_cv) were computed by dividing the standard deviation by the mean of the annual measurements.

Variable (abbreviation, unit)	Estimation	Mean (SD)	Range
<b>Large scale variables</b>			
Altitudinal gradient (Alt_Grad, %)	The altitudinal gradient measured as percentage difference of the elevation of the upstream and downstream nodes of the river segment taken from a 100-meters digital elevation model, built using the CCM 2.1 geodatabase (Vogt et al., 2007)	0.9 (1.0)	0.1 - 5.5
Elevation (Alt, m)	The average of the elevation of the catchment of the segment taken from a 100-meters digital elevation model, built using the CCM 2.1 geodatabase (Vogt et al., 2007)	558.5 (159.5)	141.3 - 904.0
Slope (Slope, %)	The maximum slope (in percentage) present in the catchment of the segment computed from a 100-meters digital elevation model, built using the CCM 2.1 geodatabase (Vogt et al., 2007)	53.6 (15.2)	27.7 - 84.0
Stream order (SO)	The segment Strahler's stream order, taken from the CCM 2.1 geodatabase (Vogt et al., 2007)	3.5 (1.0)	2.0 - 5.0
Precipitation (Prec, mm)	The long-term average annual precipitation of the catchment of the segment taken from the WORLDCLIM 2 database (Fick & Hijmans, 2017) and the CCM 2.1 geodatabase (Vogt et al., 2007)	679.3 (55.5)	597.4 - 806.0
<b>Local variables</b>			
Coefficient of variation of stream width (Width_cv)	The wetted width at a site was computed each year by averaging measurements taken at four equally spaced transversal segments along the 50-m sampling reach. Annual values were then averaged across years, and the interannual coefficient of variation (SD/Mean) of stream width was estimated.	0.22 (0.12)	0.06 - 0.58
Coefficient of variation of stream depth (Depth_cv)	Stream depth at a site was computed each year by averaging measurements taken at 25%, 50%, and 75% of the length of four equally spaced transversal segments along the 50-m sampling reach. Only one measurement was taken in segments <3m wide. Annual values were then averaged across years, and the interannual coefficient of variation (SD/Mean) of stream depth was estimated.	0.29 (0.12)	0.07 - 0.64
Coefficient of variation of water velocity (Vel_cv)	Water velocity at a site was computed each year by averaging measurements taken at 25%, 50%, and 75% of the length of four equally spaced transversal segments along the 50-m sampling reach. Only one measurement was taken in segments <3m wide. Annual values were then averaged across years, and the interannual coefficient of variation (SD/Mean) of stream depth was estimated.	1.38 (0.55)	0.68 - 2.82
<b>Exotic species</b>			
Proportion of exotic species (Exot)	Ratio between the total number of individuals from exotic species captured during the study at a site and the total number of individuals of all species captured at that site.	0.22 (0.24)	0.00 - 0.83

Table 6.S3 - Summary results of univariate analysis screening for the presence of linear (L;  $y = b_0 + b_1X$ ) or quadratic (Q;  $y = b_0 + b_1X + b_2X^2$ ) relations between variables describing fish community dynamics (directionality and mean velocity) and environmental variables. Analysis involved generalized linear models with Gaussian errors and identity link. Quadratic relations were evaluated using orthogonal polynomials of second degree. For each test we provide the F- statistics and the P-value for the overall significance of the model, the adjusted R-squared ( $R^2$ ) and the AICc. The relations used to build the multivariate models are shaded in grey. Abbreviations are given in Table 6.S2.

	Directionality				Mean Velocity			
	F-statistic	P	R <sup>2</sup>	AICc	F-statistic	P	R <sup>2</sup>	AICc
<b>Landscape variables</b>								
Alt_Grad (L)	0.14	0.72	-0.03	-99.05	0.40	0.53	-0.02	-22.97
Alt_Grad (Q)	0.40	0.67	-0.04	-97.79	0.98	0.39	0.00	-22.64
Elev (L)	0.26	0.62	-0.03	-99.18	5.07	0.03	0.12	-27.53
Elev (Q)	2.89	0.07	0.12	-102.72	2.51	0.10	0.09	-25.66
Prec (L)	0.18	0.67	-0.03	-99.10	0.00	0.95	-0.04	-22.55
Prec (Q)	0.60	0.56	-0.03	-98.21	0.04	0.96	-0.07	-20.63
SO (L)	1.01	0.32	0.00	-99.97	7.05	0.01	0.17	-29.28
SO (Q)	1.42	0.26	0.03	-99.90	6.10	0.01	0.26	-31.72
Slope (L)	2.70	0.11	0.06	-101.67	2.58	0.12	0.05	-25.19
Slope (Q)	1.30	0.29	0.02	-99.67	1.24	0.30	0.02	-23.19
<b>Local variables</b>								
Depth_cv (L)	0.39	0.54	-0.02	-99.32	0.09	0.77	-0.03	-22.64
Depth_cv (Q)	1.29	0.29	0.02	-99.63	0.05	0.96	-0.07	-20.64
Vel_cv (L)	2.63	0.12	0.05	-101.60	2.55	0.12	0.05	-25.16
Vel_cv (Q)	1.28	0.30	0.02	-99.62	1.84	0.18	0.05	-24.37
Width_cv (L)	0.13	0.72	-0.03	-99.05	1.06	0.31	0.00	-23.66
Width_cv (Q)	0.35	0.71	-0.05	-97.68	0.88	0.43	-0.01	-22.44
<b>Exotic species</b>								
Exot (L)	0.83	0.37	-0.01	-99.78	9.04	0.01	0.22	-30.94
Exot (Q)	0.41	0.67	-0.04	-97.81	5.48	0.01	0.24	-30.76

Table 6.S4 - Ranking of alternative candidate models relating the directionality of community change to environmental variables. For each alternative model we indicate the variables included (Abbreviations provided in Table 6.S2), the degrees of freedom of the model (df), its log-likelihood (logLik), the Akaike Information Criteria corrected for small sample sizes (AICc), the difference between the AICc of the model and the AICc of the best model (delta), and the Akaike weight (weight). Only models with Delta AIC < 4 are shown.

Model	df	logLik	AICc	delta	weight
Alt (Q) + Slope	5	57.72	-102.90	0.00	0.06
Alt (Q) + SO(Q)	6	59.01	-102.40	0.56	0.05
Slope + Vel_cv	4	55.85	-102.10	0.84	0.04
Alt (Q) + Slope + Vel_cv	6	58.72	-101.80	1.15	0.04
Alt (Q)	4	55.36	-101.10	1.81	0.03
Alt_Grad + Alt (Q) + Slope	6	58.38	-101.10	1.82	0.03
Alt (Q) + Prec + Slope	6	58.37	-101.10	1.84	0.02
Slope	3	53.84	-100.70	2.19	0.02
Vel_cv	3	53.80	-100.70	2.26	0.02
Alt (Q) + Slope + Exot	6	58.06	-100.50	2.47	0.02
NULL	2	52.45	-100.50	2.47	0.02
Alt (Q) + Prec + Slope + Vel_cv	7	59.75	-100.40	2.54	0.02
Alt (Q) + SO(Q) + Prec	7	59.62	-100.10	2.79	0.02
Alt_Grad + Slope + Vel_cv	5	56.26	-100.00	2.92	0.01
Alt_Grad + Alt (Q) + Slope + Vel_cv	7	59.55	-100.00	2.93	0.01
Alt (Q) + Prec	5	56.24	-100.00	2.95	0.01
Alt (Q) + Slope + Width_cv	6	57.79	-99.90	3.00	0.01
Alt (Q) + SO (Q) + Width_cv	7	59.45	-99.80	3.13	0.01
Alt_Grad + Alt (Q) + SO (Q)	7	59.26	-99.40	3.50	0.01
Alt (Q) + SO (Q) + Slope	7	59.24	-99.40	3.54	0.01
Slope + Width_cv + Vel_cv	5	55.93	-99.40	3.58	0.01

Model	df	logLik	AICc	delta	weight
Alt (Q) + Vel_cv	5	55.91	-99.30	3.62	0.01
Slope + Vel_cv + Exot	5	55.90	-99.30	3.63	0.01
Prec + Slope + Vel_cv	5	55.87	-99.20	3.70	0.01
Alt (Q) + Exot	5	55.78	-99.10	3.88	0.01
Alt (Q) + Prec + Slope + Exot	7	59.07	-99.00	3.89	0.01
Alt (Q) + SO (Q) + Exot	7	59.05	-99.00	3.93	0.01
Alt_Grad + Alt (Q) + Slope + Exot	7	59.04	-99.00	3.95	0.01
Alt (Q) + Slope + Depth_cv (Q)	7	59.04	-99.00	3.95	0.01
Alt (Q) + SO (Q) + Vel_cv	7	59.03	-99.00	3.97	0.01

Table 6.S5 - Ranking of alternative candidate models relating the mean velocity of community change to environmental variables. For each alternative model we indicate the variables included (Abbreviations provided in Table 6.S2), the degrees of freedom of the model (df), its log-likelihood (logLik), the Akaike Information Criteria corrected for small sample sizes (AICc), the difference between the AICc of the model and the AICc of the best model (delta), and the Akaike weight (weight). Only models with Delta AIC < 4 are shown.

Model	df	logLik	AICc	delta	weight
Alt + SO (Q)	5	25.55	-38.60	0.00	0.10
Alt + SO (Q) + Slope	6	27.09	-38.50	0.07	0.10
SO (Q) + Exot (Q)	6	26.94	-38.20	0.38	0.09
Alt + SO (Q) + Prec	6	26.24	-36.80	1.77	0.04
Alt + SO (Q) + Vel_cv	6	26.14	-36.60	1.96	0.04
Alt + SO (Q) + Exot (Q)	7	27.83	-36.60	2.04	0.04
Alt + SO (Q) + Prec + Slope	7	27.82	-36.60	2.05	0.04
Alt + SO (Q) + Width_cv	6	25.79	-35.90	2.68	0.03
Alt_Grad + Alt + SO (Q)	6	25.71	-35.80	2.83	0.03
Alt + SO (Q) + Depth_cv	6	25.64	-35.60	2.96	0.02
Alt_Grad + SO (Q) + Exot (Q)	7	27.23	-35.40	3.23	0.02
Alt + SO (Q) + Slope + Width_cv	7	27.21	-35.30	3.27	0.02
Alt + SO (Q) + Slope + Vel_cv	7	27.17	-35.30	3.34	0.02
Alt + SO (Q) + Slope + Depth_cv	7	27.12	-35.20	3.45	0.02
SO (Q) + Slope + Exot (Q)	7	27.11	-35.10	3.48	0.02
Alt_Grad + Alt + SO (Q) + Slope	7	27.10	-35.10	3.50	0.02
SO (Q) + Width_cv + Exot (Q)	7	27.01	-34.90	3.66	0.02
SO (Q) + Prec + Exot (Q)	7	26.94	-34.80	3.82	0.02
SO (Q) + Vel_cv + Exot (Q)	7	26.94	-34.80	3.82	0.02
SO (Q) + Depth_cv + Exot (Q)	7	26.94	-34.80	3.82	0.02

## Supplementary Figures

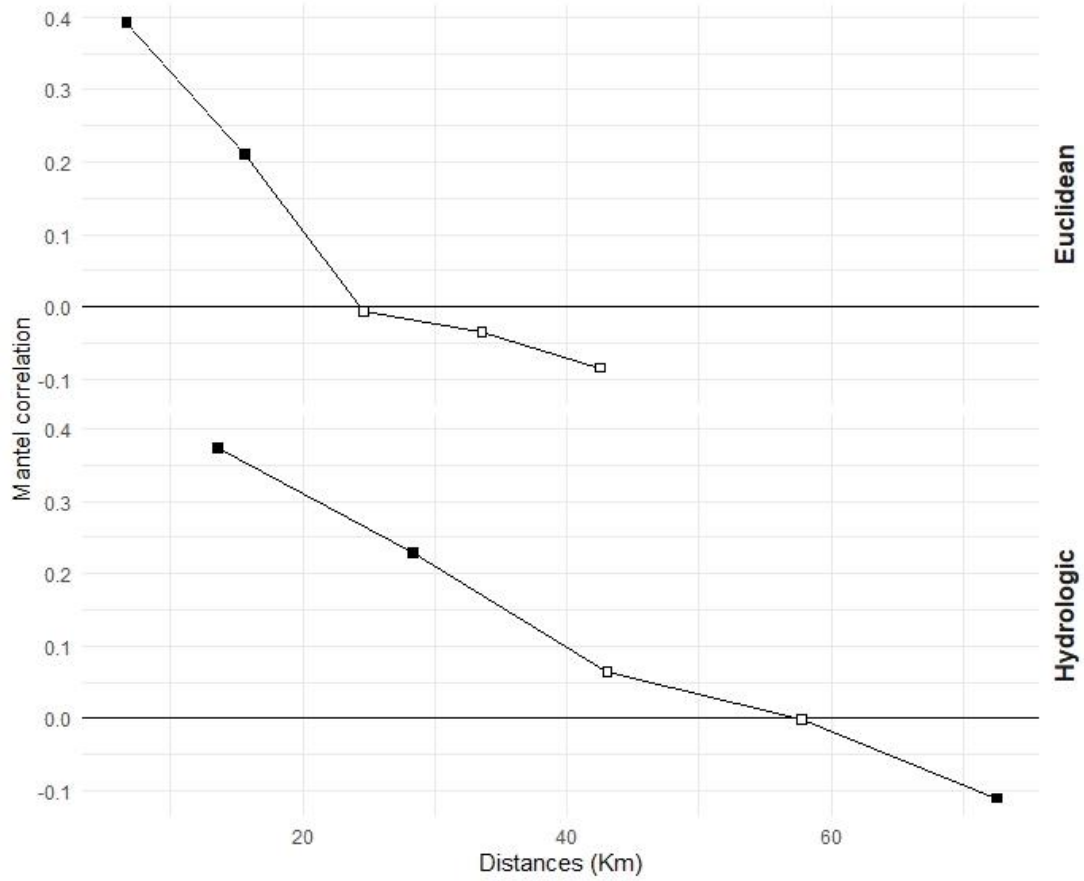


Figure 6.S1 - Mantel correlograms showing variations in the correlations between pairwise distances in community trajectories and spatial distances, considering either Euclidean or hydrologic distances. Black squares indicate statistically significant correlations at  $P < 0.05$ .

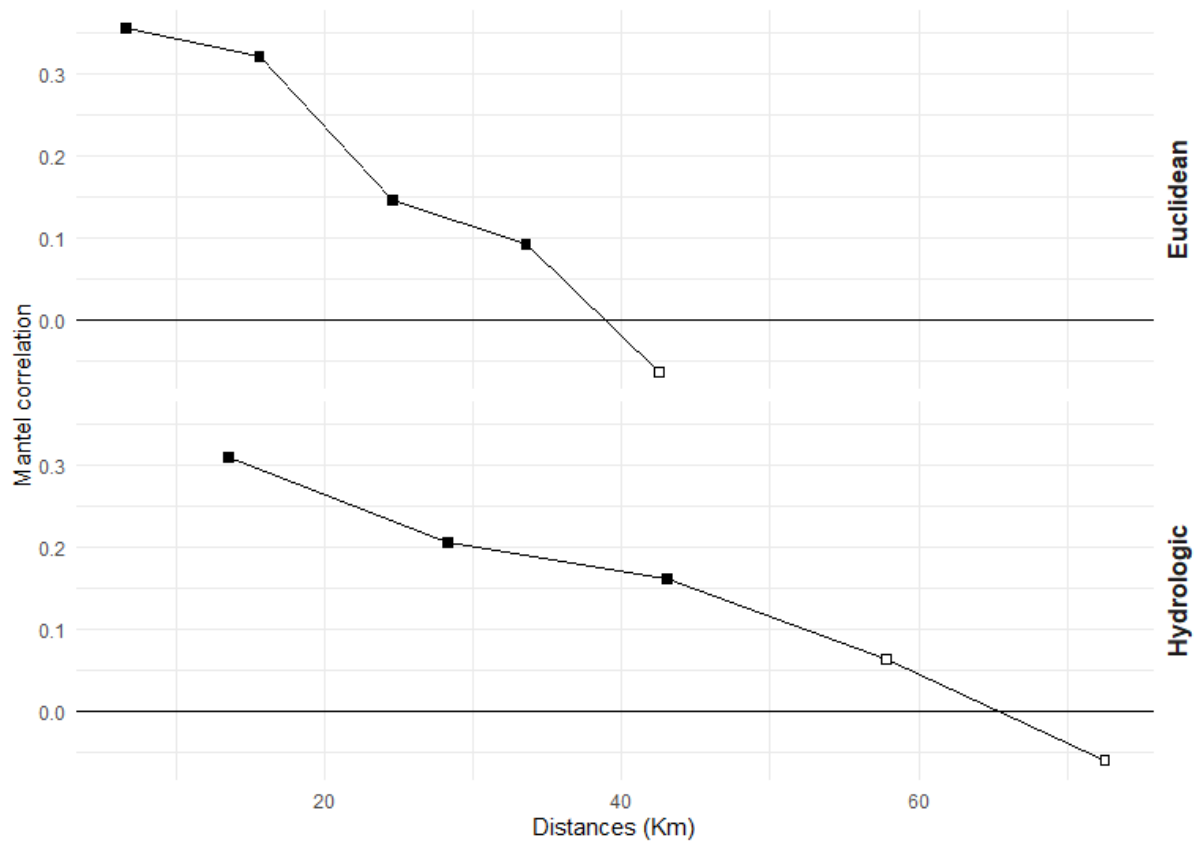


Figure 6.S2 - Mantel correlograms showing variations in the correlations between pairwise environmental and spatial distances between sites, considering either Euclidean or hydrologic distances. Black squares indicate statistically significant correlations at  $P < 0.05$ .

## Supplementary references

- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Vogt, J., Soille, P., De Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., Bódis, K., Dusart, J., Paracchini, M. L., Haastруп, P., & Bamps, C. (2007). *A pan-European River and Catchment Database*. <https://doi.org/10.2788/35907>





## Chapter 7 – General discussion





## Chapter 7 – General discussion

### 7.1 In a nutshell

This thesis advances the development of methodologies to model biodiversity patterns and processes in stream networks, at both species and community levels. Each methodology was developed as an answer to a particular challenge that arises when analysing data collected on stream-dwelling organisms. In **chapter 3**, we presented a novel framework to deal with imperfect detectability by using Time-to-first-detection instead of repeated surveys. A novelty of this framework is that time to detection doesn't have to be known precisely, requiring only a lower and an upper bound. To account for spatial autocorrelation (SAC) in data from stream networks, we presented two approaches. In **chapter 3**, we include two autocovariance terms in the hierarchical model to account for SAC in the probability of stream fish presence conditional to the probability of the stream segment holding water. In **chapters 4, 5, and 6**, we used geostatistical models specifically designed for taking into account the spatial structure of stream networks. In these models, SAC is partitioned into three components: one component based on Euclidean distances, and two based on hydrological distances, of which one for sites connected by water flow, and the other irrespectively of flow. In **chapter 4**, we presented a framework to model invasive species in stream networks, especially in the early stages of invasion where spatial processes may be more relevant than environmental filters. In **chapter 5**, we used the output of **chapter 4** to assess the distribution of refuge habitats for species displaced by the invasive species. In **chapter 6**, we developed a novel framework to analyse the temporal dynamics of species communities with geomodelling, thereby producing a continuous spatial view of communities dynamics across an entire river network. Although the approaches and frameworks described in this thesis were developed to answer particular questions when analysing data for our study systems, these tools can have a broader scope in their application. They can be useful for researchers aiming to study the biodiversity patterns and processes in stream networks, at both species and community levels, and also to provide crucial data for an informed planning and management actions.

## 7.2 Challenges of spatial distribution modelling in dendritic networks

Species distribution models have become very popular methods to assess the spatial distribution and habitat requirements of species for different applications, from pure academic purposes to conservation and management. While the framework is similar across the terrestrial, marine and freshwater realms, constructing a link between species records and environmental conditions, each realm has its unique challenges (Domisch et al., 2015). Domisch et al. (2015) identified three challenges inherent to modelling species distribution in streams networks: (1) the spatial configuration in terms of the hierarchical structure of catchments and dendritic stream networks; (2) obtaining relevant and spatially continuous environmental predictors along the stream network; and (3) species detectability and thus the challenge of obtaining freshwater species occurrence data along the stream network. In this thesis, we investigate mainly the role of the stream network structure and the impact of imperfect detectability, but also take into account the quality of the predictors.

The sampling of streams and rivers is complex as the organism are rarely available for direct observation (Domisch et al., 2015; Ferreira et al., 2016). The failure to detect a species when present is a very common source of bias for SDM's (Domisch et al., 2015; Guillera-Aroita et al., 2014; Lobo et al., 2010), and may result from several factors like species abundances and traits, habitat characteristics, and sampling site selection and methodology (Einoder et al., 2018; Ferreira & Beja, 2013; Mazerolle et al., 2007). Ignoring or assuming a constant imperfect detectability, even in standard methodologies like electrofishing, may result in spurious conclusions even for little variation in the detection probability (Gwinn et al., 2016; Hangsleben et al., 2013). Although there are methods to account for imperfect detectability (e.g.: MacKenzie et al., 2002), these methods are based on repeated sampling and so they are often disregarded by freshwater researchers as resource intensive (Gwinn et al., 2016 but see Beesley et al., 2014; Dextrase et al., 2014; Som et al., 2018; Sullivan et al., 2019). Time-to-first-detection models, as the model we developed for **chapter 3**, are a cost-effective alternative by modelling the observations as a continuous process (Ferreira et al., 2016; Garrard et al., 2008, 2013).

The hierarchical model we developed to model fish distribution in **chapter 3** included three levels: we used Bayesian equivalent to autologistic regressions (Augustin et al., 1996; Bardos et al., 2015; Besag, 1974; Gumpertz et al., 1997; Hoeting et al., 2000) to model water presence in the stream channel that would allow fish dwelling, and the probability of fish presence (occupancy) conditional on the water presence. Using the interval censored time-to-first-detection, we modelled time to detection of the species, conditional on presence, regarding local variables. The model structure is summarised in fig 7.1. The water and the fish presence was

modelled in regard to landscape variables, and to account for spatial autocorrelation we included an autoregression term in both. To construct this term, we used hydrological distances to define the network of neighbours using a cut-off threshold that would remove spatial correlation, assessed through Moran's I correlograms (Legendre & Legendre, 2012). This approach is straightforward and can be used to account for SAC in Bayesian hierarchical modelling applied to stream networks, but it can be computationally intense, particularly the covariance matrix inversion and model extrapolation (Ver Hoef et al., 2018).

### Hierarchical model:

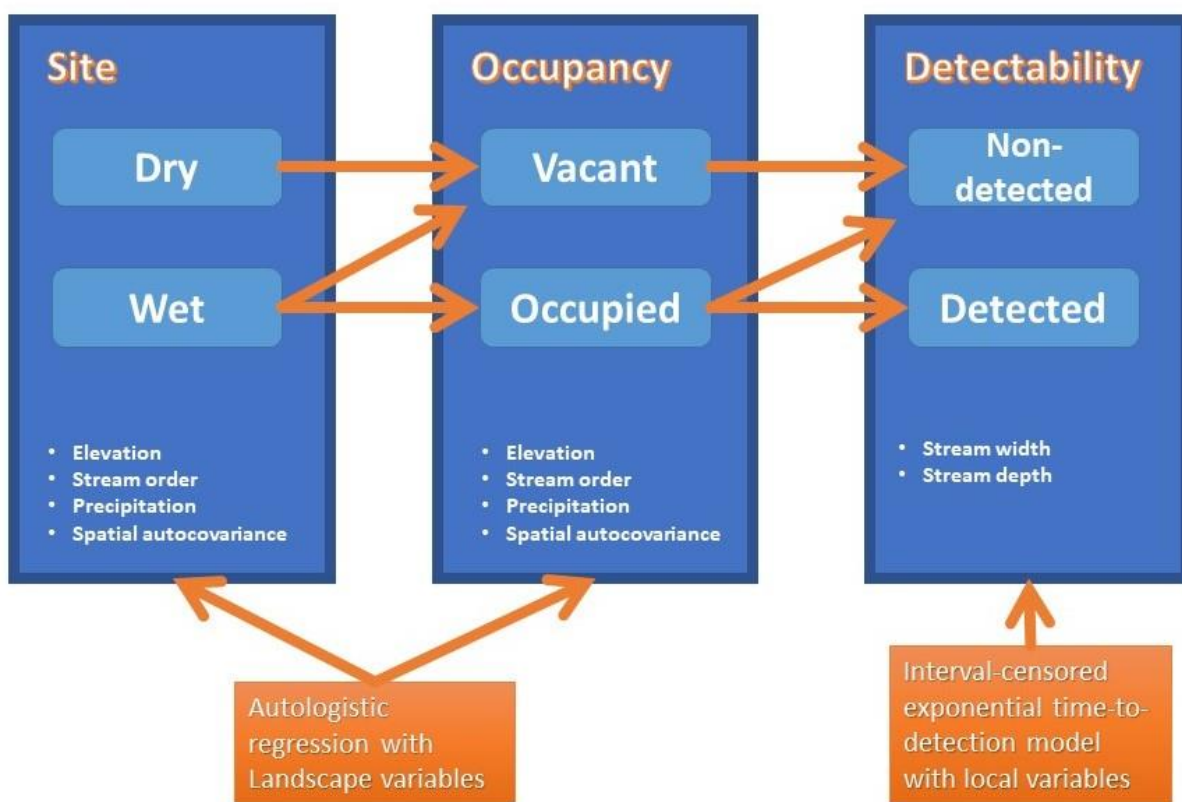


Figure 7.1 - Structure for the hierarchical model developed for chapter 3

Besides the computation burden of autologistic models, Bayesian hierarchical modelling can be a daunting task for beginners and non-statisticians, albeit it does allow great flexibility of modelling processes and patterns (Kéry, 2010; Kéry & Royle, 2016). For the more common modelling applications in rivers, Ver Hoef et al. (2006) proposed a new class of geostatistical models: spatial stream network models (SSNM). SSNM are a framework to deal with SAC in stream networks that take into account the dendritic structure and the direction of flow (Peterson & Ver Hoef, 2010; Ver Hoef & Peterson, 2010). Similar to traditional glm, SSNM relates a given

response variable to co-variables, but in the random component includes up to three autocovariate functions to account for spatial dependencies based on Euclidean and hydrologic distances among sites, both with and without flow connection (Isaak et al., 2014; Peterson & Ver Hoef, 2010; Ver Hoef & Peterson, 2010). These models were included in the package SSN (Ver Hoef et al., 2014) for the popular statistical program R (R Core Team, 2019), and can be used for most types of stream survey data (Isaak et al., 2014). In previous studies, SSNM were used to model water physic-chemistry properties (Garreta et al., 2009), fish density (Isaak et al., 2017), the distribution of desmans (Quaglietta et al., 2018), and the genetic diversity of willow trees (Rodríguez-González et al., 2019), among other applications. In this thesis, we used SSNM to model the distribution of invasive crayfish (**chapter 4**) and the distribution of refuge habitat for amphibians (**chapter 5**), as well as the spatial variation in the temporal dynamics of fish communities (**chapter 6**).

### 7.3 Modelling expanding distributions of invasive species and their consequences

SDM's are an essential tool in the study of invasive biology (Franklin, 2013; Guisan et al., 2013; Rodríguez et al., 2007; Srivastava et al., 2019). SDM's have been used to identify areas under the risk of biological invasions (Capinha et al., 2011; Hortal et al., 2010; Jiménez-Valverde et al., 2011); predict the distribution of invasive species in newly invaded areas (Capinha & Anastácio, 2011); assess the factors associated with colonization success (Capinha et al., 2012); measure the impact on native species (Ficetola et al., 2011); and select possible areas for conservation actions (Préau et al., 2020).

The assumption of equilibrium between organisms and their environment is a standard working postulate in SDM's that is seldom met, particularly for species that are expanding their range like invasive species (De Marco et al., 2008; Guisan & Thuiller, 2005; Václavík & Meentemeyer, 2012). SDM's calibrated with records from initials stages of an invasion are likely to yield less accurate and robust predictions and under-represent the potential invaded range (Václavík & Meentemeyer, 2012).

For aquatic, or semi-aquatic, species invading river systems, the dendritic structure of the stream network will constrain the patterns of the expansion from initial founder populations, with individuals dispersing up and downstream and progressively colonizing favourable habitats across the stream network (Bernardo et al., 2011; Bronnenhuber et al., 2011; Hein et al., 2011). To model these species, geostatistical models based on Euclidean distances alone are not

enough, and the spatial structure of the stream network should be taken into account (Peterson & Ver Hoef, 2010; Ver Hoef & Peterson, 2010).

The spatial stream network models (SSNM), proposed by Ver Hoef et al. (2006) as a framework to deal with SAC in stream networks that take into account the dendritic structure and the direction of flow are an obvious choice to model the distribution of invasive species in river systems. In **chapter 4**, we describe the distribution of two invasive crayfish using SSNM's. The spatial components greatly improved model performance, evidencing that the distribution of these invasive crayfish was more of a product of spatial process than environmental filtering.

Biotic interactions play an important role in shaping the species distributions, and failing to account for them may affect the power and sensibility of SDM's (Araújo & Luoto, 2007; Wisz et al., 2013). There are several tools proposed to account for and to incorporate biotic interactions into SDM's, but a straightforward approach is to use the distribution of a species as a predictor in the distribution of another (Wisz et al., 2013). This approach makes no assumptions about the nature of the biotic relation, just a suspicion of an ecological link between species is required.

As invasive species spread through a landscape, they cause several impacts to native species, mainly through competition and predation/herbivory, that may result in species extinctions, but the most common outcome is a modification in the species range (Bellard et al., 2016; Genovesi et al., 2015; Pyšek et al., 2017). At least at the landscape level, the persistence of native species may be assured by the presence of ecological refuges<sup>2</sup>, patches of habitat where biota can retreat, persist, and eventually expand following more favourable conditions (Davis et al., 2013). These refuges may be patches of habitat that are unsuitable for invasive species or areas where invasive species have not reached due to distance, physical barriers or lagging in the expansion. Identifying the distribution and the environmental drivers of these refuges is of conservation concern, particularly for species threatened by invasive species.

In **chapter 5**, to predict the distribution of amphibian ecological refuges under the invasion by two species of crayfish, we produced SDM's where we included, as predictor variables, the probability of presence of the two invasive crayfish computed in **chapter 4**. We used the outputs of the models from **chapter 4** instead of actual observations to facilitate the spatial extrapolation of the models to an entire watershed. We found that the refuges of amphibians are located

---

<sup>2</sup> Not to be confounded with evolutionary refugia. Evolutionary refugia and ecological refuges differ mainly in the spatial-temporal scale, while ecological refuges operate a landscape scale and in a short time frame, typically up to a few decades, the evolutionary refugia operates at continental and millennium scale (Davis et al., 2013; Keppel et al., 2012)

mainly in the headwaters and that under plausible expansion scenarios of the crayfish species, these refuges are likely to contract in the future.

The remediation of freshwater ecosystems invaded by alien species, particularly crayfish, is difficult, expensive and has met limited success (Gherardi et al., 2011; Stebbing et al., 2014). As such, is essential to identify the distribution of refuge habitats, focus on their conservation, monitoring the spread of the alien species, and plan and execute management actions to prevent the invasion of refuges.

## 7.4 Modelling temporally dynamic ecological processes in stream networks

Ecological communities change through space and time in response to a variety of biotic and abiotic factors, but this variability may not translate to instability (Collins, 2000). In the absence of disturbance, communities can alternate, stochastically, between stable states or vary around a loose equilibrium without a discernible direction of change (Collins, 2000; Jones et al., 2017; Matthews et al., 2013; Renner et al., 2014), and can be stable even at geological scale (Blanco et al., 2021). Nonetheless, communities may change to new states of equilibrium, slowly if tracking subtle changes in the environment, or shift abruptly when environmental drivers approach certain thresholds (Bestelmeyer et al., 2011; Drake et al., 2020; Scheffer et al., 2001).

To understand the dynamic in the ecological communities and detect shifts in the ecological states, it is crucial to establish long term monitoring programs (Matthews & Marsh-Matthews, 2017). As such, much effort has been devoted to understanding how biological communities change over time, particularly in freshwater ecosystems, usually through studies conducted for extended periods at discrete locations (e.g. Baranov et al., 2020; Bêche et al., 2009; Erős et al., 2020; Jourdan et al., 2018; Matthews & Marsh-Matthews, 2016). To use this accumulated knowledge in conservation and management, it is necessary to develop metrics that reflect the temporal dynamics of the communities and extrapolate these metrics to the landscape level, which in river systems may not be straightforward.

Several metrics have been proposed to quantify temporal changes in biological communities, like Kendall's coefficient of concordance to estimate constancy in species rank abundances or the coefficient of variation to estimate variability in species abundances (Grossman et al., 1990). Yet, these metrics reflect only partly, and yet, they do not reflect the temporal patterns of communities' dynamics. These patterns have been, traditionally, inferred by visual inspection of the plotted trajectories of the communities in a chosen resemblance space (e.g. Magalhães et al., 2007; Matthews et al., 2013). Borrowing techniques from trajectory analysis, De Cáceres et



al. (2019) proposed an analytical framework to describe and compare the trajectories of communities using their geometric properties. For instance, length and speed of the trajectories reflects how fast and how much is the community changing, and if the change is more or less progressive or abrupt. On the other hand, the angles between trajectory segments and overall directionality of the trajectory, quantifies if the change in the communities are directional or not (De Cáceres et al., 2019).

In **chapter 6**, we used the community trajectory analysis framework to study stream fish data collected in 30 locations on the Sabor river basin, between 2012 and 2019. We computed the mean velocity and the overall directionality of change of the fish community, and used the SSNM framework (Peterson et al., 2013) to relate these metrics to environmental drivers and extrapolate the community dynamics to the entire watershed. We did not find any evidence of directionality in the temporal change of the fish communities in the Sabor watershed, supporting the hypothesis that these communities exist in a loose equilibrium. However, the rate of change in the communities of stream fishes varied with stream order, and it was higher at sites with lower elevation. After we extrapolated and plotted this metric, we were able to identify that faster changes were associated to streams draining into the hydroelectric reservoir. These streams are probably under increased stress from the reservoir, due to alterations of the flow regime and/or expansion of alien species from the reservoir (R. M. B. Santos et al., 2017).

For the conservation and management of riverine ecosystems, it is essential a continuous view of the entire landscape (“Riverscape”) to understand how the processes among species and their habitats interact at different scales (Fausch et al., 2002). The frameworks such as the one we developed in chapter 6 are useful to extrapolate data collected at discrete locations into a continuous spatial view of the community dynamics, which can be used by managers and conservationist to identify streams under anthropogenic pressures and to plan mitigation actions.

## 7.5 Stream network biodiversity modelling as an essential tool for conservation

There is an increasing demand for species distribution models to support conservation decision making (Guisan et al., 2013). SDM's can be useful in several stages of the decision process, e.g. identify conservation issues, help define objectives, evaluating consequences of possible alternative actions, and assessing trade-offs between benefits and costs of such actions (Guisan et al., 2013). SDM's should be of particular usefulness to generalize our observations made on discrete parts of streams to a continuous view of the watershed wide patterns, and thus

enhancing our understanding of the functioning of riverine ecosystems as a continuous system (Fausch et al., 2002; Vannote et al., 1980).

Failure to detect a species when a species is present, is, arguably, the most common source of bias in datasets used to produce SDM's (Guillera-Arroita, 2017). Not accounting for imperfect detectability affects model performance (Lahoz-Monfort et al., 2014), may compromise the range estimation (Comte & Grenouillet, 2013; Kéry et al., 2013) and thus affect our ability to understand the environmental drivers of the species distribution (Guillera-Arroita et al., 2014; MacKenzie et al., 2006). There have been several applications of occupation-detection models to the conservation of river ecosystems, e.g.; evaluating the potential of sites to receive re-introduction of endangered fish species (Dextrase et al., 2014; Lamothe et al., 2019), determining environmental flows for fish (Gwinn et al., 2016), detecting invasive species (Britton et al., 2011), determine environmental drivers for crayfish (Magoulick et al., 2017), estimating the effort needed to survey freshwater mussels (Reid, 2016), and electrofishing efficiency (Reid & Haxton, 2017).

Spatial autocorrelation (SAC) can be defined as the property of the majority of ecological datasets where two observations taken at nearby locations are more similar (positive autocorrelation) or less similar (negative correlation) than expected to occur by random processes (Legendre, 1993). As most statistical techniques for producing SDM's assume independence among observations, SAC can be a source of bias increasing the probability of type 1 error (Diniz-Filho et al., 2003; Legendre, 1993; Record et al., 2013; Václavík & Meentemeyer, 2012), i.e. with SAC it's possible to be more confident in a presence and that presence may be false. While there are several tools to account for SAC in SDM's (Dormann et al., 2007), the dendritic nature rivers calls for more applied solutions such as the class of geostatistical models known as SSNM (Peterson et al., 2013; Peterson & Ver Hoef, 2010). SDM's that account for SAC should be of particular relevance when dealing with species whose distribution is more related to spatial processes than environmental filters such as invasive species.

Biotic interactions can be another source of mismatch between the real and the predicted distribution of a species. Biotic interaction can shape distributions through several processes, e.g. exclusion due to competition (Hewitt, 1999), apex predator induced cascading effects (Estes et al., 2011), emerging diseases introduced by exotic species (Gherardi, 2007) or global commerce (Fisher & Garner, 2020), and these interactions can shape the distribution of species at different scales (Araújo & Luoto, 2007; Wisz et al., 2013). There are several techniques to incorporate biotic interactions (review by Wisz et al., 2013), but a straightforward approach, when a link between species is known or suspected, is to include the distribution patterns of a

species as a predictor in the SDM of other species (Mota-Ferreira & Beja, 2020; Wisz et al., 2013). This approach has been used particularly to assess how species will adapt to climate changes (Araújo & Luoto, 2007; Preston et al., 2008; Schweiger et al., 2010), and it has been demonstrated to increase model performance (Meier et al., 2010).

Management of stream networks is usually planned at the river basin scale, and as such it is important to develop frameworks that allow the extrapolation of the community dynamics observed at discrete segments of rivers to a continuous spatial view of the entire river basin (Mota-Ferreira et al., 2021). This is particularly relevant because current biomonitoring methods usually assume that communities are stable and are the result of environmental filters with no regard to stochastic and dispersion processes (Cid et al., 2020). The community trajectory analysis framework (CTA, De Cáceres et al., 2019) allows for researchers and managers that are monitoring communities that may persist in a state of loose equilibrium (Collins, 2000; Matthews et al., 2013; Matthews & Marsh-Matthews, 2017) to assess the impact of environmental disturbances and/or anthropogenic stressors.

## 7.6 Conclusions and future research

In this thesis, we explored several methodological approaches to modelling the distribution of organisms in stream networks. Although we used particular faunal groups, the approaches developed in this thesis can have a broader application to other groups of organisms associated with river systems, particularly if the distribution is heavily dependent on spatial processes. The methods presented in this thesis should be useful for researchers and managers dealing with data collected in discrete locations along streams, and aiming to obtain a general view of biodiversity patterns and dynamics across a stream network.

Like every scientific work, during the elaboration of this thesis many questions could be further asked and are worth pursuing in the future. The modelling of other groups of organisms with particular relevance in the riverine ecosystems could also be addressed with some of the tools we developed for this thesis. There are also extensions and modifications to our approaches that might be considered.

The interval censored time-to-first-detection model that we developed in **chapter 3** is a particularly flexible tool. When used with exponential survival model and constant interval size, the interval censored time-to-first-detection model is the mathematical equivalent of the removal sampling design framework in occupation detection modelling (MacKenzie et al., 2006 and check the proof in appendix 3.S2 of **chapter 3**). But the survival exponential model is the simplest of the parametric survival models and assumes a constant rate of events (e. g.

detections in time-to-first-detection models), a more complex model may be fitted in this framework (e.g. Weibull model, Kleinbaum & Klein, 2012). More, the interval censored time-to-first-detection model may be used with uneven sampling times/intervals providing a framework to account for imperfect detectability when is impossible to ascertain the precise time of detection of the target species. Another application of time-to-first-detection models could be to model the amount of any positive quantity needed to observe some event of interest, e. g. distance-to-detection in transects or the amount of water volume need to be filtered for eDNA monitoring.

One logical extension of the SSNM framework is to combine it with occupation-detection modelling. Such model will have to be implemented in a software that allows the construction of tailored models such as the programs for Bayesian inference like WinBUGS (Lunn et al., 2000) or Jags (Plummer, 2003), because the other software packages used to implement such models, as the stand-alone Presence (Hines, 2006) and Unmarked for R (Fiske & Chandler, 2011) for occupancy-detection and the SSN package (Ver Hoef et al., 2014) for SSNM models, currently do not have this option. While the occupancy-detection models have been extensively adapted into the Bayesian framework (e.g. Ferreira et al., 2016; Ferreira & Beja, 2013; R. A. L. Santos et al., 2018), the SSNM adaptation is less straightforward due to its complexity and computational intensity of inverting large covariance matrices (Santos-Fernandez et al., 2021; Ver Hoef et al., 2018). To our knowledge, the first adaptation of the SSNM was proposed by Santos-Fernandez et al. (2021) in a pre-print. Combining the two frameworks should be feasible although the resulting model would be very complex, with many latent parameters to be estimated and would require a significant amount of high quality data to avoid identifiability issues.

Euclidean distances alone are not appropriated to build SDM's for species associated with river ecosystems due to the dendritic nature of stream networks (Peterson et al., 2007; Peterson & Ver Hoef, 2010; Ver Hoef & Peterson, 2010). In this thesis, we used two approaches to deal with SAC in stream networks, but there are more. Blanchet et al. (2008) proposed a modification of Moran's eigenvector maps to model species distributions with SAC resulting from an asymmetric process along a gradient or a directed network such as a stream network called asymmetric eigenvector maps. This method would also have the potential to be combined with occupancy-detection modelling in a Bayesian hierarchical model.

In **chapter 5**, we identified stream segments that can act as refuges for amphibian species under an invasion by exotic crayfish. The next step would be a systematic selection and prioritization of the most important refuges to ensure the resilience of the amphibian populations in the Sabor river basin (Margules & Pressey, 2000). Such selection should take into account the connectivity

among populations, and the structure and seasonality of the stream network (Hermoso et al., 2012, 2013). Bernardo et al. (2011) reported the expansion speed for the invasive crayfish in the Sabor basin. Confirming these values would allow us to create a time frame for managers to plan and execute actions to halt crayfish expansion and conserve the remaining amphibians' refuges.

Mediterranean amphibian species are known to use the aquatic habitat at different periods (Diaz-Paniagua, 1992; Ferreira & Beja, 2013; Jakob et al., 2003; Richter-boix et al., 2006), with some species also using the aquatic habitats in different periods at different locations (Caetano & Castanet, 1993; Caetano & Leclair Jr., 1999). With the possibility of the amphibian species varying the usage of the stream network through space and time, accounting for the phenology in SDM's would be of interest to understand the patterns and drivers of amphibian aquatic usage.

In **chapter 6**, we used the velocity of change and overall directionality, some geometric properties of the trajectory of the Sabor fish communities, to infer about the stress posed by a large infrastructure. Under the community trajectory analysis (De Cáceres et al., 2019) it would be possible to derive other properties like the variation in the speed of change in communities (aka acceleration). We found that communities more close to the hydroelectric dam near the mouth of the Sabor river presented a higher rate of change than communities located further. Using acceleration, we could assess if the rate of change is increasing or even if it varied after the filling of the dam. Following a BACI logic (Smith, 2014), estimating acceleration in the change of communities could be an important assessment in putative impacts.

In **chapter 5** we assumed a putative negative biotic interaction between the invasive crayfish and amphibians that was largely supported by the literature, and in **chapter 6** we describe the dynamics of the community but refrained to infer about the nature of the interactions among species. Biotic interactions are important drivers of species distribution (Wisiz et al., 2013), but it is challenging to infer biotic interactions from observational data alone (Blanchet et al., 2020; Dormann et al., 2018). There are several tools to account for biotic interactions, most of them implies a priori knowledge of the nature of the interaction and usually are restrain among pair of species and symmetric interactions (Wisiz et al., 2013). Porto & Beja (2021) proposed a new framework with a simultaneous system of equations for species modelling while accounting for asymmetric interaction networks in a community. The occurrence of each species is modelled as a function of measured and unmeasured environmental factors and the potential occurrence of every other species in the community (Porto & Beja, 2021). This framework would help understand the dynamics of freshwater species and combined with an appropriated framework, make predictions about species co-occurrence across the watershed.

In summary, the distribution of species in stream networks is a result of multiple processes, from environmental filters to spatial processes at work at different scales (Domisch et al., 2015). These properties and the nature of stream dwelling species make the application of SDM's in riverine ecosystems particularly challenging (Olden et al., 2010). The main aim of this thesis was to tackle some of these challenges and to contribute to the understanding of the drivers that govern the distribution of species in stream networks, which is key knowledge for the conservation of these ecosystems (Geist, 2011). It is essential that we regard the riverine ecosystems as a continuous system (Fausch et al., 2002; Vannote et al., 1980) and not be limited by our observations made at discrete locations. We encourage researchers and managers to regard SDM's as part of the scientific process that structures our observations, knowledge, and assumptions, using an abstraction of reality with a formal description of the elements and their relations to enhance our understanding of these particular ecosystems (Wang & Grant, 2019a, 2019b).

## 7.7 References

- Araújo, M. B., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16(6), 743–753. <https://doi.org/10.1111/j.1466-8238.2007.00359.x>
- Augustin, N. H., Muggleston, M. A., & Buckland, S. T. (1996). An Autologistic Model for the Spatial Distribution of Wildlife. *The Journal of Applied Ecology*, 33(2), 339. <https://doi.org/10.2307/2404755>
- Baranov, V., Jourdan, J., Pilotto, F., Wagner, R., & Haase, P. (2020). Complex and nonlinear climate-driven changes in freshwater insect communities over 42 years. *Conservation Biology*, 0(0), 1–11. <https://doi.org/10.1111/cobi.13477>
- Bardos, D. C., Guillera-Aroita, G., & Wintle, B. A. (2015). Valid auto-models for spatially autocorrelated occupancy and abundance data. *Methods in Ecology and Evolution*, 6(10), 1137–1149. <https://doi.org/10.1111/2041-210X.12402>
- Bêche, L. A., Connors, P. G., Resh, V. H., & Merenlender, A. M. (2009). Resilience of fishes and invertebrates to prolonged drought in two California streams. *Ecography*, 32(5), 778–788. <https://doi.org/10.1111/j.1600-0587.2009.05612.x>
- Beesley, L. S., Gwinn, D. C., Price, A., King, A. J., Gawne, B., Koehn, J. D., & Nielsen, D. L. (2014). Juvenile fish response to wetland inundation: how antecedent conditions can inform environmental flow policies for native fish. *Journal of Applied Ecology*, 51(6), 1613–1621. <https://doi.org/10.1111/1365-2664.12342>
- Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology Letters*, 12(4). <https://doi.org/10.1098/rsbl.2015.0623>
- Bernardo, J. M., Costa, a. M., Bruxelas, S., & Teixeira, a. (2011). Dispersal and coexistence of two non-native crayfish species ( *Pacifastacus leniusculus* and *Procambarus clarkii* ) in NE Portugal over a 10-year period. *Knowledge and Management of Aquatic Ecosystems*, 401, 28. <https://doi.org/10.1051/kmae/2011047>

- Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society - Series B (Methodological)*, 36(2), 192–236. <http://www.jstor.org/stable/2984812>
- Bestelmeyer, B. T., Ellison, A. M., Fraser, W. R., Gorman, K. B., Holbrook, S. J., Laney, C. M., Ohman, M. D., Peters, D. P. C., Pillsbury, F. C., Rassweiler, A., Schmitt, R. J., & Sharma, S. (2011). Analysis of abrupt transitions in ecological systems. *Ecosphere*, 2(12), art129. <https://doi.org/10.1890/es11-00216.1>
- Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23(7), 1050–1063. <https://doi.org/10.1111/ele.13525>
- Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Modelling directional spatial processes in ecological data. *Ecological Modelling*, 215(4), 325–336. <https://doi.org/10.1016/j.ecolmodel.2008.04.001>
- Blanco, F., Calatayud, J., Martín-Perea, D. M., Domingo, M. S., Menéndez, I., Müller, J., Fernández, M. H., & Cantalapiedra, J. L. (2021). Punctuated ecological equilibrium in mammal communities over evolutionary time scales. *Science*, 372(6539), 300–303. <https://doi.org/10.1126/science.abd5110>
- Britton, J. R., Pegg, J., & Gozlan, R. E. (2011). Quantifying imperfect detection in an invasive pest fish and the implications for conservation management. *Biological Conservation*, 144(9), 2177–2181. <https://doi.org/10.1016/j.biocon.2011.05.008>
- Bronnenhuber, J. E., Dufour, B. A., Higgs, D. M., & Heath, D. D. (2011). Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous round goby, *Neogobius melanostomus*, in Great Lakes tributaries. *Molecular Ecology*, 20(9), 1845–1859. <https://doi.org/10.1111/j.1365-294X.2011.05030.x>
- Caetano, M. H., & Castanet, J. (1993). Variability and microevolutionary patterns in *Triturus marmoratus* from Portugal: age, size, longevity and individual growth. *Amphibia-Reptilia*, 14(2), 117–129. <https://doi.org/10.1163/156853893X00291>
- Caetano, M. H., & Leclair Jr., R. (1999). Comparative Phenology and Demography of *Triturus boscai* from Portugal. *Journal of Herpetology*, 33(2), 192. <https://doi.org/10.2307/1565714>
- Capinha, C., & Anastácio, P. (2011). Assessing the environmental requirements of invaders using ensembles of distribution models. *Diversity and Distributions*, 17(1), 13–24. <https://doi.org/10.1111/j.1472-4642.2010.00727.x>
- Capinha, C., Brotons, L., & Anastácio, P. (2012). Geographical variability in propagule pressure and climatic suitability explain the European distribution of two highly invasive crayfish. *Journal of Biogeography*, n/a-n/a. <https://doi.org/10.1111/jbi.12025>
- Capinha, C., Leung, B., & Anastácio, P. (2011). Predicting worldwide invasiveness for four major problematic decapods: an evaluation of using different calibration sets. *Ecography*, 34(3), 448–459. <https://doi.org/10.1111/j.1600-0587.2010.06369.x>
- Cid, N., Bonada, N., Heino, J., Cañedo-Argüelles, M., Crabot, J., Sarremejane, R., Soininen, J., Stubbington, R., & Datry, T. (2020). A Metacommunity Approach to Improve Biological Assessments in Highly Dynamic Freshwater Ecosystems. *BioScience*, 70(5), 427–438. <https://doi.org/10.1093/biosci/biaa033>
- Collins, S. L. (2000). Disturbance frequency and community stability in native tallgrass prairie. *American Naturalist*, 155(3), 311–325. <https://doi.org/10.1086/303326>
- Comte, L., & Grenouillet, G. (2013). Species distribution modelling and imperfect detection: comparing occupancy versus consensus methods. *Diversity and Distributions*, 19(8), 996–1007. <https://doi.org/10.1111/ddi.12078>

- Davis, J., Pavlova, A., Thompson, R., & Sunnucks, P. (2013). Evolutionary refugia and ecological refuges: Key concepts for conserving Australian arid zone freshwater biodiversity under climate change. In *Global Change Biology* (Vol. 19, Issue 7, pp. 1970–1984). Wiley-Blackwell. <https://doi.org/10.1111/gcb.12203>
- De Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wisser, S. K., Fortin, M.-J., Condit, R., & Hubbell, S. (2019). Trajectory analysis in community ecology. *Ecological Monographs*, 89(2), e01350. <https://doi.org/10.1002/ecm.1350>
- De Marco, P., Diniz-Filho, J. A. F., & Bini, L. M. (2008). Spatial analysis improves species distribution modelling during range expansion. *Biology Letters*, 4(5), 577–580. <https://doi.org/10.1098/rsbl.2008.0210>
- Dextrase, A. J., Mandrak, N. E., & Schaefer, J. A. (2014). Modelling occupancy of an imperilled stream fish at multiple scales while accounting for imperfect detection: implications for conservation. *Freshwater Biology*, 59(9), 1799–1815. <https://doi.org/10.1111/fwb.12384>
- Diaz-Paniagua, C. (1992). Variability in timing of larval season in an amphibian community in SW Spain. *Ecography*, 15(3), 267–272. <https://doi.org/10.1111/j.1600-0587.1992.tb00034.x>
- Diniz-Filho, J. A. F., Bini, L. M., & Hawkins, B. A. (2003). Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, 12(1), 53–64. <https://doi.org/10.1046/j.1466-822X.2003.00322.x>
- Domisch, S., Jähnig, S. C., Simaika, J. P., Kuemmerlen, M., & Stoll, S. (2015). Application of species distribution models in stream ecosystems: The challenges of spatial and temporal scale, environmental predictors and species occurrence data. *Fundamental and Applied Limnology*, 186(1–2), 45–61. <https://doi.org/10.1127/fal/2015/0627>
- Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke, H., Moretti, M. D., Pagel, J., Pinkert, S., Schleuning, M., Schmidt, S. I., Sheppard, C. S., Steinbauer, M. J., Zeuss, D., & Kraan, C. (2018). Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. *Global Ecology and Biogeography*, 27(9), 1004–1016. <https://doi.org/10.1111/geb.12759>
- Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, D. W., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr, F. M., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5), 609–628. <https://doi.org/10.1111/j.2007.0906-7590.05171.x>
- Drake, J. M., O'Regan, S. M., Dakos, V., Kéfi, S., & Rohani, P. (2020). Alternative stable states, tipping points, and early warning signals of ecological transitions. In K. S. McCann & G. Gellner (Eds.), *Theoretical Ecology* (pp. 263–284). Oxford University Press. <https://doi.org/10.1093/oso/9780198824282.003.0015>
- Einoder, L. D., Southwell, D. M., Lahoz-Monfort, J. J., Gillespie, G. R., Fisher, A., & Wintle, B. A. (2018). Occupancy and detectability modelling of vertebrates in northern Australia using multiple sampling methods. *PLoS ONE*, 13(9). <https://doi.org/10.1371/journal.pone.0203304>
- Erős, T., Comte, L., Filipe, A. F., Ruhi, A., Tedesco, P. A., Brose, U., Fortin, M. J., Giam, X., Irving, K., Jacquet, C., Larsen, S., Sharma, S., & Olden, J. D. (2020). Effects of nonnative species on the stability of riverine fish communities. *Ecography*, 1–11. <https://doi.org/10.1111/ecog.04985>
- Estes, J. A., Terborgh, J., Brashares, J. S., Power, M. E., Berger, J., Bond, W. J., Carpenter, S. R., Essington, T. E., Holt, R. D., Jackson, J. B. C., Marquis, R. J., Oksanen, L., Oksanen, T., Paine, R. T.,



- Pikitch, E. K., Ripple, W. J., Sandin, S. A., Scheffer, M., Schoener, T. W., ... Soulé, M. E. (2011). Trophic Downgrading of Planet Earth. *Science*, 333(July), 301–307. <https://doi.org/10.1126/science.1205106>
- Fausch, K. D., Torgersen, C. E., Baxter, C. V., & Li, H. W. (2002). Landscapes to Riverscapes: Bridging the Gap between Research and Conservation of Stream Fishes. *BioScience*, 52(6), 483. [https://doi.org/10.1641/0006-3568\(2002\)052\[0483:lrbtg\]2.0.co;2](https://doi.org/10.1641/0006-3568(2002)052[0483:lrbtg]2.0.co;2)
- Ferreira, M., & Beja, P. (2013). Mediterranean amphibians and the loss of temporary ponds: Are there alternative breeding habitats? *Biological Conservation*, 165(null), 179–186. <https://doi.org/10.1016/j.biocon.2013.05.029>
- Ferreira, M., Filipe, A. F., Bardos, D. C., Magalhães, M. F., & Beja, P. (2016). Modeling stream fish distributions using interval-censored detection times. *Ecology and Evolution*, 6(15), 5530–5541. <https://doi.org/10.1002/ece3.2295>
- Ficetola, G. F., Siesa, M. E., Manenti, R., Bottoni, L., De Bernardi, F., & Padoa-Schioppa, E. (2011). Early assessment of the impact of alien species: Differential consequences of an invasive crayfish on adult and larval amphibians. *Diversity and Distributions*, 17(6), 1141–1151. <https://doi.org/10.1111/j.1472-4642.2011.00797.x>
- Fisher, M. C., & Garner, T. W. J. (2020). Chytrid fungi and global amphibian declines. *Nature Reviews Microbiology*, 18(6), 332–343. <https://doi.org/10.1038/s41579-020-0335-x>
- Fiske, I., & Chandler, R. B. (2011). unmarked : an R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, 43(10), 1–23. <https://doi.org/10.18637/jss.v043.i10>
- Franklin, J. (2013). Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions*, 19(10), 1217–1223. <https://doi.org/10.1111/ddi.12125>
- Garrard, G. E., Bekessy, S. A., McCarthy, M. A., & Wintle, B. A. (2008). When have we looked hard enough? A novel method for setting minimum survey effort protocols for flora surveys. *Austral Ecology*, 33(8), 986–998. <https://doi.org/10.1111/j.1442-9993.2008.01869.x>
- Garrard, G. E., McCarthy, M. a., Williams, N. S. G., Bekessy, S. a., & Wintle, B. a. (2013). A general model of detectability using species traits. *Methods in Ecology and Evolution*, 4(1), 45–52. <https://doi.org/10.1111/j.2041-210x.2012.00257.x>
- Garreta, V., Monestiez, P., & Ver Hoef, J. M. (2009). Spatial modelling and prediction on river networks: up model, down model or hybrid? *Environmetrics*, 21(5), n/a-n/a. <https://doi.org/10.1002/env.995>
- Geist, J. (2011). Integrative freshwater ecology and biodiversity conservation. *Ecological Indicators*, 11(6), 1507–1516. <https://doi.org/10.1016/j.ecolind.2011.04.002>
- Genovesi, P., Carnevali, L., & Scalera, R. (2015). *The impact of invasive alien species on native species in Europe - Technical report for the European Commission* (Issue April 2016). <http://www.isprambiente.gov.it/files/notizie-ispra/notizie-2015/the-impact-of-invasive-alien-species-on-native-threatened-species-in-europe/the-impact-of-invasive-alien-species-on-native-threatened-species-in-europe>
- Gherardi, F. (2007). Understanding the impact of invasive crayfish. In F. Gherardi (Ed.), *Biological invaders in inland waters: Profiles, distribution, and threats* (pp. 507–542). Springer Netherlands. [https://doi.org/10.1007/978-1-4020-6029-8\\_28](https://doi.org/10.1007/978-1-4020-6029-8_28)
- Gherardi, F., Aquiloni, L., Diéguez-Urbeondo, J., & Tricarico, E. (2011). Managing invasive crayfish: Is there a hope? In *Aquatic Sciences* (Vol. 73, Issue 2, pp. 185–200). SP Birkhäuser Verlag Basel. <https://doi.org/10.1007/s00027-011-0181-z>

- Grossman, G. D., Dowd, J. F., & Crawford, M. (1990). Assemblage stability in stream fishes: A review. *Environmental Management*, 14(5), 661–671. <https://doi.org/10.1007/BF02394716>
- Guillera-Aroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*, 40(2), 281–295. <https://doi.org/10.1111/ecog.02445>
- Guillera-Aroita, G., Lahoz-Monfort, J. J., MacKenzie, D. I., Wintle, B. A., & McCarthy, M. A. (2014). Ignoring Imperfect Detection in Biological Surveys Is Dangerous: A Response to ‘Fitting and Interpreting Occupancy Models’’. *PLoS ONE*, 9(7), e99571. <https://doi.org/10.1371/journal.pone.0099571>
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9), 993–1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16(12), 1424–1435. <https://doi.org/10.1111/ele.12189>
- Gumpertz, M. L., Graham, J. M., & Ristaino, J. B. (1997). Autologistic Model of Spatial Pattern of Phytophthora Epidemic in Bell Pepper: Effects of Soil Variables on Disease Presence. *Journal of Agricultural, Biological, and Environmental Statistics*, 2(2), 131. <https://doi.org/10.2307/1400400>
- Gwinn, D. C., Beesley, L. S., Close, P., Gawne, B., & Davies, P. M. (2016). Imperfect detection and the determination of environmental flows for fish: Challenges, implications and solutions. *Freshwater Biology*, 61(1), 172–180. <https://doi.org/10.1111/fwb.12684>
- Hangsleben, M. A., Allen, M. S., & Gwinn, D. C. (2013). Evaluation of Electrofishing Catch per Unit Effort for Indexing Fish Abundance in Florida Lakes. *Transactions of the American Fisheries Society*, 142(1), 247–256. <https://doi.org/10.1080/00028487.2012.730106>
- Hein, C. L., Öhlund, G., & Englund, G. (2011). Dispersal through stream networks: modelling climate-driven range expansions of fishes. *Diversity and Distributions*, 17(4), 641–651. <https://doi.org/10.1111/j.1472-4642.2011.00776.x>
- Hermoso, V., Kennard, M. J., & Linke, S. (2012). Integrating multidirectional connectivity requirements in systematic conservation planning for freshwater systems. *Diversity and Distributions*, 18(5), 448–458. <https://doi.org/10.1111/j.1472-4642.2011.00879.x>
- Hermoso, V., Ward, D. P., & Kennard, M. J. (2013). Prioritizing refugia for freshwater biodiversity conservation in highly seasonal ecosystems. *Diversity and Distributions*, n/a-n/a. <https://doi.org/10.1111/ddi.12082>
- Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68(1–2), 87–112. <https://doi.org/10.1111/j.1095-8312.1999.tb01160.x>
- Hines, J. E. (2006). *PRESENCE- Software to estimate patch occupancy and related parameters* (13.11). USGS-PWRC. <https://www.mbr-pwrc.usgs.gov/software/presence.html>
- Hoeting, J. A., Leecaster, M., & Bowden, D. (2000). An Improved Model for Spatially Correlated Binary Responses. *Journal of Agricultural, Biological, and Environmental Statistics*, 5(1), 102. <https://doi.org/10.2307/1400634>
- Hortal, J., Borges, P. A. V., Jiménez-Valverde, A., de Azevedo, E. B., & Silva, L. (2010). Assessing the areas under risk of invasion within islands through potential distribution modelling: The case of *Pittosporum undulatum* in São Miguel, Azores. *Journal for Nature Conservation*, 18(4), 247–257.

<https://doi.org/10.1016/j.jnc.2009.11.002>

- Isaak, D. J., Peterson, E. E., Ver Hoef, J. M., Wenger, S. J., Falke, J. A., Torgersen, C. E., Sowder, C., Steel, E. A., Fortin, M.-J., Jordan, C. E., Ruesch, A. S., Som, N., & Monestiez, P. (2014). Applications of spatial statistical network models to stream data. *Wiley Interdisciplinary Reviews: Water*, 1(3), 277–294. <https://doi.org/10.1002/wat2.1023>
- Isaak, D. J., Ver Hoef, J. M., Peterson, E. E., Horan, D. L., & Nagel, D. E. (2017). Scalable population estimates using spatial-stream-network (SSN) models, fish density surveys, and national geospatial database frameworks for streams. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(2), 147–156. <https://doi.org/10.1139/cjfas-2016-0247>
- Jakob, C., Poizat, G., Veith, M., Seitz, A., & Crivelli, A. J. (2003). Breeding phenology and larval distribution of amphibians in a Mediterranean pond network with unpredictable hydrology. *Hydrobiologia*, 499, 51–61. <http://www.springerlink.com/index/LH7047P06242454P.pdf>
- Jiménez-Valverde, A., Peterson, A. T., Soberón, J., Overton, J. M., Aragón, P., & Lobo, J. M. (2011). Use of niche models in invasive species risk assessments. *Biological Invasions*, 13(12), 2785–2797. <https://doi.org/10.1007/s10530-011-9963-4>
- Jones, S. K., Ripplinger, J., & Collins, S. L. (2017). Species reordering, not changes in richness, drives long-term dynamics in grassland communities. In *Ecology Letters* (Vol. 20, Issue 12, pp. 1556–1565). <https://doi.org/10.1111/ele.12864>
- Jourdan, J., O'Hara, R. B., Bottarin, R., Huttunen, K. L., Kuemmerlen, M., Monteith, D., Muotka, T., Ozoliņš, D., Paavola, R., Pilotto, F., Springe, G., Skuja, A., Sundermann, A., Tonkin, J. D., & Haase, P. (2018). Effects of changing climate on European stream invertebrate communities: A long-term data analysis. *Science of the Total Environment*, 621, 588–599. <https://doi.org/10.1016/j.scitotenv.2017.11.242>
- Keppel, G., Van Niel, K. P., Wardell-Johnson, G. W., Yates, C. J., Byrne, M., Mucina, L., Schut, A. G. T., Hopper, S. D., & Franklin, S. E. (2012). Refugia: Identifying and understanding safe havens for biodiversity under climate change. In *Global Ecology and Biogeography* (Vol. 21, Issue 4, pp. 393–404). John Wiley & Sons, Ltd (10.1111). <https://doi.org/10.1111/j.1466-8238.2011.00686.x>
- Kéry, M. (2010). Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses (1st ed.). Elsevier.
- Kéry, M., Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2013). Analysing and mapping species range dynamics using occupancy models. *Journal of Biogeography*, 40(8), 1463–1474. <https://doi.org/10.1111/jbi.12087>
- Kéry, M., & Royle, J. A. (2016). Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS. Volume 1, Prelude and static models.
- Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis*. Springer New York. <https://doi.org/10.1007/978-1-4419-6646-9>
- Lahoz-Monfort, J. J., Guillera-Arroita, G., & Wintle, B. a. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, 23(4), 504–515. <https://doi.org/10.1111/geb.12138>
- Lamothe, K. A., Dextrase, A. J., & Drake, D. A. R. (2019). Characterizing species co-occurrence patterns of imperfectly detected stream fishes to inform species reintroduction efforts. *Conservation Biology*, 33(6), 1392–1403. <https://doi.org/10.1111/cobi.13320>
- Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6), 1659–1673.

<https://doi.org/10.2307/1939924>

- Legendre, P., & Legendre, L. F. J. (2012). *Numerical Ecology* (3rd ed.). Elsevier.
- Lobo, J. M., Jiménez-Valverde, A., & Hortal, J. (2010). The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, *33*(1), 103–114. <https://doi.org/10.1111/j.1600-0587.2009.06039.x>
- Lunn, D. J. J. D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, *10*(4), 325–337. <http://www.springerlink.com/index/M832242U211K4437.pdf>
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*(8), 2248–2255. [https://doi.org/10.1890/0012-9658\(2002\)083\[2248:ESORWD\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2)
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier Academic Press.
- Magalhães, M. F., Beja, P., Schlosser, I. J., & Collares-Pereira, M. J. (2007). Effects of multi-year droughts on fish assemblages of seasonally drying Mediterranean streams. *Freshwater Biology*, *52*(8), 1494–1510. <https://doi.org/10.1111/j.1365-2427.2007.01781.x>
- Magoulick, D. D., DiStefano, R. J., Imhoff, E. M., Nolen, M. S., & Wagner, B. K. (2017). Landscape- and local-scale habitat influences on occupancy and detection probability of stream-dwelling crayfish: implications for conservation. *Hydrobiologia*, *799*(1), 217–231. <https://doi.org/10.1007/s10750-017-3215-2>
- Margules, C. R. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, *405*(6783), 243–253. <https://doi.org/10.1038/35012251>
- Matthews, W. J., & Marsh-Matthews, E. (2016). Dynamics of an upland stream fish community over 40 years: Trajectories and support for the loose equilibrium concept. *Ecology*, *97*(3), 706–719. <https://doi.org/10.1890/14-2179.1>
- Matthews, W. J., & Marsh-Matthews, E. (2017). Stream fish community dynamics: A critical synthesis. In *Stream Fish Community Dynamics: A Critical Synthesis*. JHU Press.
- Matthews, W. J., Marsh-Matthews, E., Cashner, R. C., & Gelwick, F. (2013). Disturbance and trajectory of change in a stream fish community over four decades. *Oecologia*, *173*(3), 955–969. <https://doi.org/10.1007/s00442-013-2646-3>
- Mazerolle, M. J. M. J., Bailey, L. L., Kendall, W. L., Andrew Royle, J., Converse, S. J. S. J., & Nichols, J. D. (2007). Making great leaps forward: accounting for detectability in herpetological field studies. *Journal of Herpetology*, *41*(4), 672–689. <https://doi.org/10.1670/07-061.1>
- Meier, E. S., Kienast, F., Pearman, P. B., Svenning, J.-C., Thuiller, W., Araújo, M. B., Guisan, A., & Zimmermann, N. E. (2010). Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, *33*(6), 1038–1048. <https://doi.org/10.1111/j.1600-0587.2010.06229.x>
- Mota-Ferreira, M., & Beja, P. (2020). Combining geostatistical and biotic interaction modeling to predict amphibian refuges under crayfish invasion across dendritic stream networks. *Diversity and Distributions*, *26*(6), 699–714. <https://doi.org/10.1111/ddi.13047>
- Mota-Ferreira, M., Filipe, A. F., Magalhães, M. F., Carona, S., & Beja, P. (2021). Spatial modelling of temporal dynamics in stream fish communities under anthropogenic change. *Diversity and Distributions*,

27(2), 313–326. <https://doi.org/10.1111/ddi.13189>

- Olden, J. D., Kennard, M. J., Leprieur, F., Tedesco, P. A., Winemiller, K. O., & García-Berthou, E. (2010). Conservation biogeography of freshwater fishes: Recent progress and future challenges. *Diversity and Distributions*, 16(3), 496–513. <https://doi.org/10.1111/j.1472-4642.2010.00655.x>
- Peterson, E. E., Theobald, D. M., & Ver Hoef, J. M. (2007). Geostatistical modelling on stream networks: Developing valid covariance matrices based on hydrologic distance and stream flow. *Freshwater Biology*, 52(2), 267–279. <https://doi.org/10.1111/j.1365-2427.2006.01686.x>
- Peterson, E. E., & Ver Hoef, J. M. (2010). A mixed-model moving-average approach to geostatistical modeling in stream networks. *Ecology*, 91(3), 644–651. <https://doi.org/10.1890/08-1668.1>
- Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>
- Plummer, M. (2003). JAGS : A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling JAGS : Just Another Gibbs Sampler. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Dsc. <https://doi.org/10.1.1.13.3406>
- Porto, M., & Beja, P. (2021). Simultaneous equations modelling of communities with interacting species networks. *BioRxiv*. <https://doi.org/10.1101/2021.03.07.434253>
- Préau, C., Nadeau, I., Sellier, Y., Isselin-Nondedeu, F., Bertrand, R., Collas, M., Capinha, C., & Grandjean, F. (2020). Niche modelling to guide conservation actions in France for the endangered crayfish *Austroptamobius pallipes* in relation to the invasive *Pacifastacus leniusculus*. *Freshwater Biology*, 65(2), 304–315. <https://doi.org/10.1111/fwb.13422>
- Preston, K. L., Rotenberry, J. T., Redak, R. A., & Allen, M. F. (2008). Habitat shifts of endangered species under altered climate conditions: Importance of biotic interactions. *Global Change Biology*, 14(11), 2501–2515. <https://doi.org/10.1111/j.1365-2486.2008.01671.x>
- Pyšek, P., Blackburn, T. M., García-berthou, E., Perglová, I., & Rabitsch, W. (2017). Displacement and Local Extinction of Native and Endemic Species. In M. Vilà & P. E. Hulme (Eds.), *Impact of Biological Invasions on Ecosystem Services*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-45121-3>
- Quaglietta, L., Paupério, J., Martins, F. M. S., Alves, P. C., & Beja, P. (2018). Recent range contractions in the globally threatened Pyrenean desman highlight the importance of stream headwater refugia. *Animal Conservation*. <https://doi.org/10.1111/acv.12422>
- R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation Statistical Computing* (3.6.1). R Foundation for Statistical Computing. <https://doi.org/ISBN 3-900051-07-0>
- Record, S., Fitzpatrick, M. C., Finley, A. O., Veloz, S., & Ellison, A. M. (2013). Should species distribution models account for spatial autocorrelation? A test of model projections across eight millennia of climate change. *Global Ecology and Biogeography*, 22(6), 760–771. <https://doi.org/10.1111/geb.12017>
- Reid, S. M. (2016). Search effort and imperfect detection: Influence on timed-search mussel (*Bivalvia*: Unionidae) surveys in Canadian rivers. *Knowledge and Management of Aquatic Ecosystems*, 417, 17. <https://doi.org/10.1051/kmae/2016004>
- Reid, S. M., & Haxton, T. J. (2017). Backpack electrofishing effort and imperfect detection: Influence on riverine fish inventories and monitoring. *Journal of Applied Ichthyology*, 33(6), 1083–1091. <https://doi.org/10.1111/jai.13463>

- Renner, S. C., Gossner, M. M., Kahl, T., Kalko, E. K. V., Weisser, W. W., Fischer, M., & Allan, E. (2014). Temporal Changes in Randomness of Bird Communities across Central Europe. *PLoS ONE*, *9*(11), e112347. <https://doi.org/10.1371/journal.pone.0112347>
- Richter-boix, A., Llorente, G. A., & Montori, A. (2006). Breeding phenology of an amphibian community in a Mediterranean area. *Amphibia-Reptilia*, *27*, 549–559. [www.brill.nl/amre](http://www.brill.nl/amre)
- Rodríguez-González, P. M., García, C., Albuquerque, A., Monteiro-Henriques, T., Faria, C., Guimarães, J. B., Mendonça, D., Simões, F., Ferreira, M. T., Mendes, A., Matos, J., & Almeida, M. H. (2019). A spatial stream-network approach assists in managing the remnant genetic diversity of riparian forests. *Scientific Reports*, *9*(1), 6741. <https://doi.org/10.1038/s41598-019-43132-7>
- Rodríguez, J. P., Brotons, L., Bustamante, J., & Seoane, J. (2007). The application of predictive modelling of species distribution to biodiversity conservation. *Diversity and Distributions*, *13*(3), 243–251. <https://doi.org/10.1111/j.1472-4642.2007.00356.x>
- Santos-Fernandez, E., Hoef, J. M. Ver, Peterson, E. E., McGree, J., Isaak, D., & Mengersen, K. (2021). *Bayesian spatio-temporal models for stream networks*. *March*. <https://doi.org/10.13140/RG.2.2.17707.98089>
- Santos, R. A. L., Mota-Ferreira, M., Aguiar, L. M. S., & Ascensão, F. (2018). Predicting wildlife road-crossing probability from roadkill data using occupancy-detection models. *Science of The Total Environment*, *642*, 629–637. <https://doi.org/10.1016/j.scitotenv.2018.06.107>
- Santos, R. M. B., Sanches Fernandes, L. F., Cortes, R. M. V., Varandas, S. G. P., Jesus, J. J. B., & Pacheco, F. A. L. (2017). Integrative assessment of river damming impacts on aquatic fauna in a Portuguese reservoir. *Science of the Total Environment*, *601–602*, 1108–1118. <https://doi.org/10.1016/j.scitotenv.2017.05.255>
- Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591–596. <https://doi.org/10.1038/35098000>
- Schweiger, O., Biesmeijer, J. C., Bommarco, R., Hickler, T., Hulme, P. E., Klotz, S., Kühn, I., Moora, M., Nielsen, A., Ohlemüller, R., Petanidou, T., Potts, S. G., Pyšek, P., Stout, J. C., Sykes, M. T., Tscheulin, T., Vilà, M., Walther, G. R., Westphal, C., ... Settele, J. (2010). Multiple stressors on biotic interactions: How climate change and alien species interact to affect pollination. In *Biological Reviews* (Vol. 85, Issue 4, pp. 777–795). <https://doi.org/10.1111/j.1469-185X.2010.00125.x>
- Smith, E. P. (2014). BACI Design. In *Wiley StatsRef: Statistics Reference Online* (pp. 1–12). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118445112.stat07659>
- Som, N. A., Perry, R. W., Jones, E. C., De Juilio, K., Petros, P., Pinnix, W. D., & Rupert, D. L. (2018). N-mix for fish: estimating riverine salmonid habitat selection via N-mixture models. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*(7), 1048–1058. <https://doi.org/10.1139/cjfas-2017-0027>
- Srivastava, V., Lafond, V., & Griess, V. C. (2019). Species distribution models (SDM): Applications, benefits and challenges in invasive species management. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, *14*(020). <https://doi.org/10.1079/PAVSNNR201914020>
- Stebbing, P., Longshaw, M., & Scott, A. (2014). Review of methods for the management of non-indigenous crayfish, with particular reference to Great Britain. *Ethology Ecology and Evolution*, *26*(2–3), 204–231. <https://doi.org/10.1080/03949370.2014.908326>
- Sullivan, C. J., Weber, M. J., Pierce, C. L., & Camacho, C. A. (2019). Influence of river discharge on grass carp occupancy dynamics in south-eastern Iowa rivers. *River Research and Applications*, *35*(1), 60–67. <https://doi.org/10.1002/rra.3385>

- Václavík, T., & Meentemeyer, R. K. (2012). Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, 18(1), 73–83. <https://doi.org/10.1111/j.1472-4642.2011.00854.x>
- Vannote, R. L., Minshall, G. W., Cummins, K. W., Sedell, J. R., & Cushing, C. E. (1980). The river continuum concept. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(1), 130–137. <https://doi.org/10.1139/f80-017>
- Ver Hoef, J. M., & Peterson, E. E. (2010). A Moving Average Approach for Spatial Statistical Models of Stream Networks. *Journal of the American Statistical Association*, 105(489), 6–18. <https://doi.org/10.1198/jasa.2009.ap08248>
- Ver Hoef, J. M., Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R Package for Spatial Statistical Modeling on Stream Networks. *Journal of Statistical Software*, 56(3), 1–45. <https://doi.org/10.18637/jss.v056.i03>
- Ver Hoef, J. M., Peterson, E. E., Hooten, M. B., Hanks, E. M., & Fortin, M. J. (2018). Spatial autoregressive models for statistical inference from ecological data. *Ecological Monographs*, 88(1), 36–59. <https://doi.org/10.1002/ecm.1283>
- Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13(4), 449–464. <https://doi.org/10.1007/s10651-006-0022-8>
- Wang, H.-H., & Grant, W. E. (2019a). Systems, models, and modeling. In H.-H. Wang & W. E. Grant (Eds.), *Ecological Modeling: An Introduction to the Art and Science of Modeling Ecological Systems* (Vol. 31, pp. 3–20). <https://doi.org/10.1016/B978-0-444-64163-2.00001-3>
- Wang, H.-H., & Grant, W. E. (2019b). What good is modeling? In H.-H. Wang & W. E. Grant (Eds.), *Ecological Modeling: An Introduction to the Art and Science of Modeling Ecological Systems* (Vol. 31, Issue 4, pp. 215–224). <https://doi.org/10.1016/B978-0-444-64163-2.00011-6>
- Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M. C., Normand, S., Öckinger, E., Schmidt, N. M., ... Svenning, J. C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88(1), 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>