

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN  
INGEGNERIA CIVILE, CHIMICA, AMBIENTALE E DEI MATERIALI

Ciclo XXXII

**Settore Concorsuale: 08/A1 Idraulica, Idrologia, Costruzioni idrauliche e marittime**

**Settore Scientifico Disciplinare: ICAR/02 Costruzioni idrauliche e marittime e idrologia**

ASSESSMENT OF FLOOD HAZARD OVER LARGE GEOGRAPHICAL  
AREAS USING DATA-DRIVEN APPROACHES

**Presentata da:** Ricardo Tavares da Costa

**Coordinatore Dottorato**

**Prof. Luca Vittuari**

**Supervisore**

**Dr. Stefano Bagli**

**Co-Supervisore**

**Prof. Attilio Castellarin**

**Esame finale anno 2020**



To my Family and Friends.

Dedico este trabalho à minha família e aos meus amigos. Sem vocês, sem o vosso saber e experiência, não teria logrado tal belo objectivo, não teria sobrevivido às tormentas. Aos que já partiram, a vossa falta é sentida em cada dia que passa. Este ano ficará também ele marcado como aquele em que o meu irmão concluí o seu doutoramento. Que fortúnio.

*"(...) you don't really understand something until you can express it as an algorithm."*

in *The Master Algorithm*, Pedro Domingos, 2015.

*"O fim duma viagem é apenas o começo doutra. É preciso ver o que não foi visto, ver outra vez o que se viu já, ver na Primavera o que se vira no Verão, ver de dia o que se viu de noite, com sol onde primeiramente a chuva caía, ver a seara verde, o fruto maduro, a pedra que mudou de lugar, a sombra que aqui não estava. É preciso voltar aos passos que foram dados, para os repetir, e para traçar caminhos novos ao lado deles. É preciso recomeçar a viagem. Sempre."*

in *Viagem a Portugal*, José Saramago, 1995.



# Abstract

The mapping of flood hazard can be time and resource consuming, but it is essential for assessing flood risk and for designing strategies to avoid consequences and recover faster in the event of flooding. This generally involves setting up complex numerical hydrologic/hydrodynamic models to simulate the flow of water in river channels and over the floodplains. Although such approach is considered standard, it is not always feasible. For example, it is challenging to simulate floods over large areas, produce a number of scenarios, represent flood mechanisms in a detailed way, and make use of all the data that is increasingly available in the field of water resources. Clearly, flood managers could use more options. Machine learning (i.e., algorithms that learn from data, in contrast to physically-based equations) has been seldomly used until now, but is a good candidate because of simplicity, typically faster runtimes and ability to handle large amounts of data. In combination with geographic information systems attractive tools can potentially be created. The combination of nearly instantaneous results with a web-GIS provides the possibility of near-real time analysis using any modern web browser. This thesis seek for additional clues that can help in the answering of the following questions: can data-driven models live to their expectations in flood hazard assessment? to what extent they offer viable alternatives to standard approaches and what are the concrete advantages and limitations? Several aspects of flood hazard assessment are addressed by developing and employing different state-of-the-art data-driven approaches, namely for the estimation and mapping of areas that may be subject to flooding across geographic scales, their downscaling, their extrapolation and regionalisation, or the transfer between catchments based on physical similarity. In each part of the thesis, the viability of selected methods are demonstrated and possible ways to overcome limitations are highlight.



# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Standard methods for flood hazard mapping</b>	<b>5</b>
<b>2 State-of-the-art hydrogeomorphic flood hazard mapping</b>	<b>9</b>
<b>3 Flood hazard: mapping, downscaling and regionalisation</b>	<b>15</b>
3.1 Classification of flood-prone areas . . . . .	15
3.1.1 Geomorphic Flood Index . . . . .	15
3.1.2 Threshold Binary Classification . . . . .	17
3.1.3 Objective function . . . . .	19
3.2 Delineation and downscaling of flood-prone areas . . . . .	21
3.3 Regionalisation of flood-prone areas . . . . .	21
3.3.1 Stepwise Linear Regression . . . . .	22
3.3.2 Random Forest Regression . . . . .	23
<b>4 Performance metrics</b>	<b>25</b>
4.1 Threshold-independent metrics: ROC analysis . . . . .	25
4.2 Threshold-dependent metrics . . . . .	26
4.2.1 Evaluation of classification outcomes . . . . .	26
4.2.2 Evaluation of regression outcomes . . . . .	26
<b>5 Mapping flood-prone areas in Europe</b>	<b>27</b>
5.1 Impact of scale in classification outcomes . . . . .	27
5.1.1 Study area and data . . . . .	29
5.1.2 Results and discussion . . . . .	33
5.2 Regionalisation of classification outcomes . . . . .	45
5.2.1 Study area and data . . . . .	47
5.2.2 Results and discussion . . . . .	54

<b>Conclusions</b>	<b>63</b>
<b>Appendix A Analysis of open-access digital elevation models</b>	<b>67</b>
A.1 Methods . . . . .	68
A.1.1 Digital terrain analysis . . . . .	68
A.1.2 Accuracy assessment . . . . .	70
A.2 Case study and datasets . . . . .	71
A.2.1 Tanaro River Basin . . . . .	71
A.2.2 Digital elevation models datasets . . . . .	72
A.2.3 Benchmark LiDAR dataset . . . . .	75
A.3 Results and discussion . . . . .	75
<b>Appendix B Principal Component Analysis and Regression</b>	<b>81</b>
<b>Appendix C Flood hazard web application</b>	<b>83</b>
C.1 Architecture . . . . .	85
C.2 Back-end . . . . .	85
C.3 Front-end . . . . .	86
C.4 Parallelisation strategy . . . . .	87
C.5 Web application generated flood hazard maps . . . . .	87
<b>Bibliography</b>	<b>91</b>
<b>Acknowledgements</b>	<b>111</b>

# List of Figures

1	Taxonomy of machine learning methods. . . . .	3
3.1	Two components that constitute the Geomorphic Flood Index (GFI). <b>a)</b> D8 flow direction raster for a portion of a catchment. <b>b)</b> The empirically derived stage in each cell under analysis computed as a power law of bank-full depth and upslope contributing area of the hydrologically connected river centreline cell. <b>c)</b> The DEM elevation difference between the cell under analysis and the hydrologically connected river centreline cell. At the lower left, a river basin representation showing the catchment E, in grey, and the river network, in blue. At the lower centre, a cross-section of the river channel and floodplain. . . . .	17
3.2	Workflow involved in the discrete statistical classification of the Geomorphic Flood Index (GFI) for a generic river basin. <b>a)</b> GFI layer. <b>b)</b> Overlay of fixed buffer (green) around the river network centerline. <b>c)</b> Calibration area corresponding to the clipped GFI using the fixed buffer. <b>d)</b> Overlay of benchmark (red) and start of the threshold binary classification. <b>e)</b> Search for optimal $TH$ by means of an objective function. <b>f)</b> Best possible representation of flood-prone areas (black) after classification. . . . .	18
3.3	Relative Operating Characteristic (ROC) curve obtained by plotting the false positive rate (FPR) against the true positive rate (TPR) of each unique $TH$ of the flood descriptor for a generic river basin, in blue, and intersection point with the perpendicular of the line of no-skill (i.e. the diagonal bisector) that corresponds to the point of maximum True Skill Statistic (TSS), in cyan. . . . .	19
5.1	General methodological workflow for classifying a morphological descriptor and assessing its performance in delineating and extrapolating flood-prone areas. . . . .	28

5.2	European digital elevation model (EU-DEM; EEA) for computing the morphological descriptor. . . . .	29
5.3	Flood hazard maps for Europe (Dottori et al., 2016) used as reference in the classification of flood-prone areas. . . . .	31
5.4	Study area consisting of 270 major European river basins obtained from the Catchment Characterisation and Model layer, River and Catchment Database (CCM, de Jager and Vogt 2010). . . . .	32
5.5	Study area consisting of 64 sub-catchments of the Po river, Italy, obtained from the Catchment Characterisation and Model layer, River and Catchment Database (CCM, de Jager and Vogt 2010). . . . .	33
5.6	Spatial distribution of the area under the receiver operating characteristic (AUC) for six return periods, within the classification area set for each major river basin in Europe. . . . .	35
5.7	Spatial distribution of the area under the receiver operating characteristic (AUC) for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy. . . . .	36
5.8	Geographical distribution of the optimal Geomorphic Flood Index (GFI) threshold for six return periods, within the classification area set for each major river basin in Europe. . . . .	37
5.9	Geographical distribution of the optimal Geomorphic Flood Index (GFI) threshold for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy. . . . .	38
5.10	Geographical distribution of the True Skill Score (TSS) for six return periods, within the classification area set for each major river basin in Europe. . . . .	39
5.11	Geographical distribution of the True Skill Score (TSS) for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy. . . . .	40
5.12	Geographical distribution of the modified Pearson's correlation ( $r_\phi$ ) for six return periods, within the classification area set for each major river basin in Europe. . . . .	41
5.13	Geographical distribution of the modified Pearson's correlation ( $r_\phi$ ) for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy. . . . .	42
5.15	Data distribution of the optimal Geomorphic Flood Index (GFI) threshold for six return periods. . . . .	43

5.14	Data distribution for the area under the receiver operating characteristic (AUC), True Skill Score (TSS) and the modified Pearson's correlation coefficient ( $r_\phi$ ) for six return periods. . . . .	44
5.16	Workflow for developing predictive models of envelope flood extents using geomorphic and climatic-hydrologic catchment characteristics. . . . .	47
5.17	The five study river basins in Europe, with drainage divide highlighted in black in the lower left map. The training of the regression models is performed using <b>a)</b> the Thames river basin in the UK, <b>b)</b> the Weser in Germany. <b>c)</b> the upper Rhine in Switzerland, and <b>d)</b> the upper Danube in Austria. Testing is performed using <b>e)</b> the Rhône in France. . . . .	48
5.18	Subdivision into elementary catchments overlaid with a river network representation (the EEA EU-Hydro photo-interpreted river network) and outlets, defined as the intersection of each elementary catchment drainage divide with the river network. . . . .	51
5.19	Data used in the development of estimators and prediction of flood-prone areas for the 10- and 10,000-year return periods, please refer to Table 5.2 for a complete description and units of variables. <b>a)</b> Area Under the relative operating characteristic Curve AUC the modified Pearson's correlation coefficient for discrete dichotomous problems $r_\phi$ and the optimal Geomorphic Flood Index (GFI) thresholds TH obtained from the classification of training catchments (Thames, Weser, upper Rhine and upper Danube river basins). <b>b)</b> Geomorphic and climatic-hydrologic catchment characteristics for the combined training and test catchments (Rhône river basin). <b>c)</b> Correlation matrix between TH and catchment characteristics. . . . .	55
5.20	<b>a)</b> True positive rate (TPR), false discovery rate (FDR) and critical success (C) for the elementary catchments of the Rhône river basin, for the different regression models (rf – random forest; swr – stepwise regression) and for the 10-year and 10,000-year return periods. <b>b)</b> Overlaid samples of flood-prone areas and corresponding envelope flood extents, predicted with the rf model for two return periods. . . . .	60
A.1	Workflow for extracting terrain analysis layers from a digital elevation model, necessary to compute the flood index. . . . .	69

- A.2 Representation of the study area. **A)** Location of the Tanaro river basin in the Piedmont region, NW Italy, with the drainage divide highlighted in red; **B)** digital elevation model (DEM) of the Tanaro river basin and footprint of the light detection and ranging (LiDAR) dataset in blue; **C)** histogram of elevations within the Tanaro river basin. . . . . 72
- A.3 Cumulative frequency curves of terrain descriptors extracted from free digital elevation models (DEM) within the Tanaro river basin, in Italy. **A)** Terrain elevations within the light detection and ranging (LiDAR) footprint; **B)** upslope contributing area; **C)** local slope within the LiDAR footprint; **D)** elevation difference to the nearest stream,  $H$ , within the LiDAR footprint; and **E)** distance to the nearest stream,  $D$ . Logarithmic scales refer to the natural logarithm. . . . . 76
- A.4 Overlay of river networks for 3 distinct regions in the Tanaro river basin, Italy, derived from free digital elevation models (DEM). For each region, the corresponding EU-Hydro photo-interpreted river network is shown on the right. **A)** and **(B)** are sample regions representative of flatter terrain; and **C)** is a sample region representative of mountainous terrain. On the bottom left, the locations of the 3 regions within the Tanaro river basin are marked with red boxes. On the bottom right, the average values of the terrain descriptors are presented. . . . . 78
- A.5 **A)** Map of statistical bias between the SRTM GL3 digital elevation model (DEM) and the light detection and ranging (LiDAR) data in the Tanaro river basin, Italy; **B)** Map of statistical bias between the HydroSHEDS DEM and the LiDAR data; **C)** DEM elevations plotted against LiDAR elevations; **D)** Cumulative frequency of absolute BIAS plotted against the elevation difference to the nearest stream,  $H$ . . . . . 79
- C.1 Client-server diagram showing the implementation of the web application. . 85
- C.2 Usage exemplification of the web application based on the hydrogeomorphic method for delineating flood-prone areas over large scales. **a)** river sub-catchment selection; **b)** reference flood map selection, with associated return period; **c)** classification of morphological descriptor; **d)** visualisation of classification results and performance measures (top-left corner); and, **e)** downscaled extrapolation of the optimal threshold and delineation of flood-prone areas over the corresponding river basin, with dynamic response of performance (top-right corner). . . . . 88

- C.3 Sample overlay of the 100-year return period flood hazard map for Europe (Dottori et al., 2016) on the delineated hydrogeomorphic flood-prone areas using the **a)** optimal flood descriptor threshold and **b)** flood descriptor threshold controlled through slider operation (marked with a yellow circle). 90



# List of Tables

3.1	Example of contingency matrix of possible raster cell values in a binary classification. . . . .	20
5.1	Open-access datasets used in the delineation of flood-prone areas across Europe . . . . .	30
5.2	Summary of geomorphic and climatic-hydrologic catchment characteristics.	52
5.3	Optimization results for the Multiple linear regression (MLR) and the random forest regression model (RF) for simultaneously predicting the 10- and the 10,000-year return period optimal Geomorphic Flood Index (GFI) threshold (TH) using geomorphic and climatic-hydrologic catchment characteristics as explanatory variables. . . . .	58
5.4	Performance of the regression models expressed as true positive rate (TPR), false discovery rate (FDR), critical success (C) and error bias (E), for the 10- and 10,000-year return period flood-prone areas in the Rhône river basin using the RMS flood maps as benchmark. . . . .	62
A.1	Digital elevation models (DEM) currently available free of charge with spatial resolutions below 3 arc seconds. . . . .	73
A.2	Vertical accuracy assessment of the free digital elevation models (DEM) in the Tanaro river basin, Italy, expressed as mean absolute error (MAE), root mean squared error (RMSE), normalized RMSE (NRMSE) and Pearson correlation (PC), with the light detection and ranging data (LiDAR) data used as benchmark. . . . .	80



# Introduction

Floods pose a serious threat to individuals and communities, as shown by disaster data found in the International Disaster Database (EM-DAT) of the Centre for Research of the Epidemiology of Disasters (CRED). A United Nations report on the Human cost of weather related disasters (CRED/UNISDR, 2016) illustrates the dimension of the problem well: 47% of all disasters since 1995 have been floods, affecting a total of 2.3 billion people worldwide. This previous experience paints a grim picture and there is mounting evidence for an increase in frequency and intensity of severe floods due to climate change (Milly et al., 2002; Aerts et al., 2006; Kleinen and Petschel-Held, 2007; Alfieri et al., 2017; Barichivich et al., 2018; Sassi et al., 2019). Emphasised by examples of historical flood disasters (e.g., Barredo, 2007), projected consequences (e.g., Alfieri et al., 2015) represent a serious challenge to flood risk management.

On the other hand, the increase of socio-economic activities in areas that are subject to flooding perseveres in a number of countries, aggravating the exposure of persons and assets (de Moel et al., 2009; Kron et al., 2019). According to the European Environment Agency (EEA, 2016), a significant part of the European population is estimated to be living in, or near to, a floodplain.

Thus, understanding and assessing flood hazard is of great importance to protect lives and to prevent negative socio-economic and environmental impacts. The EU Floods Directive (2007/60/EC) clearly recognises this vital importance and mandates member states to produce flood hazard and risk maps, but much more can be done to help.

In particular, authorities and stakeholders could make use of more prompt and cost-effective alternatives to assess every type of flood consistently across geographic scales (e.g., Leskens et al., 2014). This applies to flood maps – a critical component of risk assessment – but also to information used for the design of protection and emergency measures, spatial planning and awareness raising. In the insurance sector, this knowledge is critical for managing portfolios, risk screening and assessing long-term financial solvency (de Moel et al., 2009).

The larger the geographic area, where management is complex and transboundary coordination is generally required, and the limited the resources of economies and organ-

isations involved, the more relevant the alternatives become. Disaster risk reduction (i.e., mitigation, preparedness, response and recovery) and risk transfer (i.e., insurance) could greatly benefit from innovative tools capable of enriching decision-making processes (e.g., Jongman et al., 2014; Pappenberger et al., 2015; Ward et al., 2015). It is unsatisfactory, and to some degree unjustified, that there is not more to offer, given today’s technological developments and unprecedented quantities of data (e.g., Ma et al., 2015).

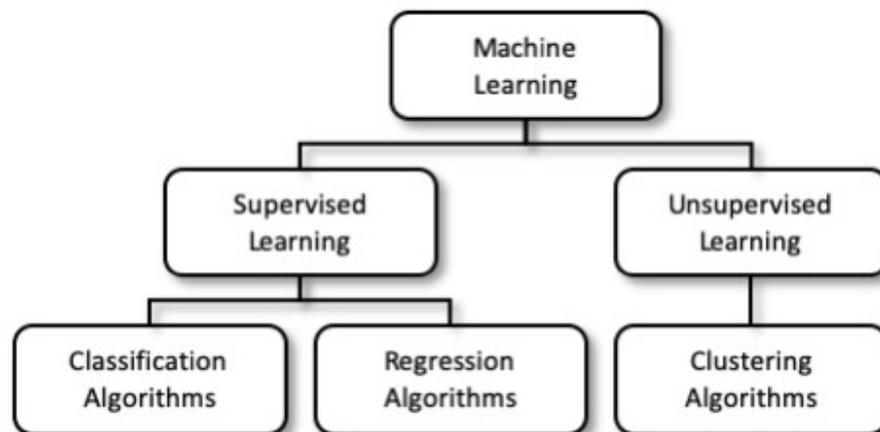
Standard approaches to flood hazard mapping generally involve the combined use of hydrological and hydrodynamic computer models (e.g., Şen and Kahya, 2017), with resulting flood maps typically produced at the reach-scale and in urban settings (Horritt and Bates, 2002). At these local scales, researchers generally devote themselves to increasing the detail of flood models, an example being the move towards very high spatial resolution (Noh et al., 2018). Instead, from regional to global scales, authors tend to focus more on simplification, as parsimonious models are more suitable to be used over larger geographic areas due to their higher computationally efficiency (Yamazaki et al., 2011; Neal et al., 2012, 2011; Pappenberger et al., 2012; Winsemius et al., 2013; Alfieri et al., 2014; Sampson et al., 2015; Dottori et al., 2016; Rebolho et al., 2018; Zheng et al., 2018a).

However, detail (i.e., physics and spatial resolution) and large geographic coverage are two desirable but often competing properties of flood modelling (e.g., Schumann et al., 2014a); in other words, it is hard to setup a flood model with one of these characteristics without compromising the other. Adding to long simulation times, numerical instabilities, limited resources and scarcity of stream gauge observations counter the effort for up-to-date flood information at any location or any time these are requested. Such bottlenecks, have motivated a number of authors to start looking at low-complexity solutions that rely more on data-driven methods (Schumann et al., 2014a; Tang et al., 2018; Giovannettone et al., 2018; Caprario and Finotti, 2019; Zhao et al., 2019).

A new class of methods that take a completely different approach from the one of standard flood modelling schemes has in fact emerged. Machine learning (Mitchell, 1997), whose algorithms instead of explicitly and exhaustively encode the physics of the processes that drive flood hazard, learns information directly from existing data (e.g., historical observations, deterministic model simulations) to infer knowledge, with respect to a specific task and measure of performance, and improve as more examples are added. Comparatively, these methods can be not only faster and cost-effective, as is desirable, but also useful when the problem at hand is hard to describe mathematically, hard to solve computationally or when it involve large amounts of data.

Data is of course central to machine learning and in principle the more quality data is fed to a specific algorithm, carefully setup for a specific task, the more it should learn and the better it should perform – this perhaps explains why big data became such a popular term more or less simultaneously. Given that water resources are no exception to the boom of data experienced in the last decades (see e.g., Chen and Wang, 2018; Levizzani and Cattani, 2019), from remote sensing and photogrammetry to networks of sensors, citizen science and social media, there is a certain expectation that machine learning will open the door to new advances, by leveraging the power of big data, and to the development of innovative tools. The question is, can data-driven models live to their expectations in flood hazard mapping? To what extent do they offer a viable alternative to standard approaches and what are the concrete advantages and limitations? Through the identification of flood hazard patterns using different strategies and case studies, this work seeks to provide additional clues that can help in the answering of these questions.

In summary, machine learning methods can be categorised into at least two major branches, supervised and unsupervised learning (Fig. 1), that are of interest to this thesis.



**Figure 1:** Taxonomy of machine learning methods.

Supervised learning is the branch of machine learning algorithms where predictive modelling is achieved based on evidence. In other words, a supervised learning algorithm infers a function from a training dataset that is able to map each new example to a response. Supervised learning is further subdivided into classification or regression algorithms. The first type of algorithms is specific to problems in the discrete space, where data is in the form of discontinuous sequences – classification algorithms aim to group new examples into categories. For instance, Olthof and Tolszczuk-Leclerc (2018) used a decision tree algorithm to classify satellite imagery into inundated floodplain pixels or

dry land. The second type of algorithms is specific to problems in the continuous space, where data is in the form of continuous sequences – regression algorithms aim to predict response values for new examples. For instance, Garmdareh et al. (2018) tested if a support vector regression algorithm and different types of artificial neural networks could be used to predict peak flood discharges.

Unsupervised learning, on the other hand, seeks patterns in datasets that do not have a response associated to each sample – clustering algorithms aim to group samples in a dataset according to a measure of similarity. Cassalho et al. (2019) tested the ability of different clustering algorithms to group catchments, based on observations from corresponding stream gauges, into hydrologically similar regions for an improved flood frequency analysis and a transfer of flood information between catchments, known as regionalisation (Blöschl and Sivapalan, 1995).

In this thesis several aspects of flood hazard assessment are addressed by employing data-driven approaches. Algorithms used were tailored specifically for the task of estimating and mapping areas that are subject to flooding across geographic scales, their downscaling, their extrapolation and regionalisation.

It brings together all the work that was carried out in the course of the last three years and it organises the material in the following way: Chapter 1 provides an overview of standard methods for flood hazard mapping; Chapter 2 provides an overview of the state-of-the-art in hydrogeomorphic flood hazard mapping; Chapter 3 provides an overview of the specific data-driven methods used for mapping areas subject to flooding and highlights strategies to enhance the transfer of flood extents to other geographic areas; in Chapter 4 different evaluation measures used throughout this thesis are summarised; and, Chapter 5 provides an overview of several case studies where mapping of areas subject to flooding was performed.

This thesis includes additional developments, namely: 1) in Appendix A the analysis of free digital elevation models (DEMs) that are central to the classification and regression problems presented in Chapter 3 and Chapter 5; 2) in Appendix B a brief description of principal component analysis and regression that were also explored in Chapter 3 and Chapter 5 ; and, 3) in Appendix C the description of a web application prototype, based on online geographic information system (GIS) technology, that seeks to integrate in a user-friendly interface the data-driven methods used to assess flood hazard.

# Chapter 1

## Standard methods for flood hazard mapping

Standard approaches to flood hazard mapping can be sub-divided into event-based, semi-continuous or fully continuous approaches according to Grimaldi et al. (2013). In an event-based approach a design hyetograph is estimated from time series of observed rainfall, fed to a hydrological model to produce a corresponding design flood hydrograph, which is then used in a hydrodynamic model to simulate flow both in the stream channel and over the floodplain. In a semi-continuous approach, synthetic time series of rainfall are fed directly to the hydrological model, of which the resulting discharge peaks are analysed in terms of frequency to produce a synthetic design flood hydrograph to be used in a hydrodynamic model. Finally, in a continuous approach, time series of rainfall are transformed by the hydrological model into discharge that are directly fed to the hydrodynamic model, of which the resulting flood maps are analysed in terms flood depth frequency on a cell-by-cell basis to estimate a corresponding return period. Grimaldi et al. (2013), have tested an event-based, a semi-continuous and a fully continuous approach to flood mapping in the Rio Torbido in Italy using a Width Function Instantaneous Unit Hydrograph (WFIUH-1par, Grimaldi et al. 2013) for hydrological modelling and FLO-2D (O'Brien et al. 1993) for hydrodynamic modelling. The authors have showed that the three methods result in significantly different flood maps.

Another example of application, is given by Alfieri et al. (2014) that used the LIS-FLOOD hydrological model (Knijff et al. 2010) to produce a discharge climatology for flood frequency analysis and the estimation of flood hydrographs, which were then input to the hydrodynamic model LISFLOOD-ACC (Bates et al. 2010, Neal et al. 2011) to produce pan-European return period flood hazard maps.

These workflows can be further simplified by performing regional flood frequency anal-

ysis using gauge observations to obtain flood quantiles used to simulate flow in steady state or, alternatively, to estimate a design flood hydrograph to simulate flow dynamically. For example, Sampson et al. (2015) used regional flood frequency analysis to generate extreme flows, derive design flood hydrographs and run the hydrodynamic model LISFLOOD-FP (Neal et al. 2012).

Alternatively, rainfall observations or design hyetographs used in the hydrological models can be substituted by stochastic rainfall scenario generators. For example, Sampson et al. (2014) generated a series of independent synthetic rainfall storms, assumed a spatially uniform rainfall pattern over the catchments of interest in the region of Dublin in Ireland, ran the *Hydrologiska Byråns Vattenbalansavdelning* hydrological model (HBV, Seibert and Vis 2012) to simulate synthetic design flood hydrographs and finally ran the LISFLOOD-FP model to obtain the flood hazard maps for Dublin.

These flood hazard mapping approaches are generally part of catastrophe modelling frameworks, typically used for disaster insurance, that also encompass modules for vulnerability and loss to represent the whole risk chain. This was the case in Sampson et al. (2014), but also Falter et al. (2015, 2016) that coupled a weather generator (Hundecca and Merz 2012) to a hydrological model (SWIM – Soil and Water Integrated Model, Krysanova et al. 1998), a hydrodynamic model based on the inertial formulation of Bates et al. (2010) and a flood loss model (FLEMOps + r – Flood Loss Estimation MOdel for the private sector, Elmer et al. 2012) to estimate risk in a meso-scale catchment, the Mulde, in Germany.

More recently, a number of authors have focused on the integration of hydrological and hydrodynamic models for flood hazard estimation in near-real time and over large-scales, in opposition to static flood maps. For example, O’Loughlin et al. (2020) prepared a hydrological model (HRR – Hillslope River Routing model, Beighley et al. 2011) with precipitation data to obtain estimates of discharge and subsequently run LISFLOOD-FP. The authors applied this approach to the middle reach of the Congo River Basin and compared the resulting flood hazard maps to inundation data inferred from multiple satellites observations. Rajib et al. (2020) used the same approach to simulate flood hazard in the Ohio River Basin in the US, but applied a different hydrological model, the Soil & Water Assessment Tool (SWAT, Arnold and Fohrer 2005); while Hoch et al. (2017) combined the PCRaster GLOBal Water Balance hydrological model (PCR-GLOBWB, van Beek and Bierkens 2008) with the Delft3D (Kernkamp et al. 2011) and the LISFLOOD-FP hydrodynamic models to compare results at the Óbidos River Basin in Brazil.

Although standard approaches to flood hazard estimation share the same principle of coupling a hydrological model to a hydrodynamic model in order to obtain flood maps, the

physical descriptions included in the models, the equations governing the propagation of flow, the numerical schemes used in solving those equations, and how the models are coupled (i.e., transfer information between themselves) can differ substantially. For example the HRR model combines Green and Ampt (1911) infiltration equations, or runoff generation, with kinematic-wave approximation of the shallow water equations, for hillslope flow routing, and diffusive-wave approximation of the shallow water equations, for channel flow routing, while the SWAT model combines the Soil Conservation Service Curve Number infiltration equations, for runoff generation, with a variable storage flow routing (Williams, 1969). The hydrodynamic equations of LISFLOOD-FP neglect the advection term and the stream channel is represented with a subgrid, while Delft3D solves the full shallow water equations and resolves both the stream channel and the floodplain.

Standard methods for flood hazard mapping, however, can be data, resource and knowledge intensive. Hydro-meteorological data required by the models not only as input, but also for calibration, are often sparse or simply inexistent for several catchments around the world; catchment characteristics, specially when they cannot be derived from a DEM, such as land-use and plant growth, can be of insufficient quality and detail; information on man-made structures, such as roads that contribute to imperviousness of urban areas, but also reservoirs that affect runoff routing and flood defences that condition the flow of water, may not be available. In order to get around these challenges, modellers resort to estimates or parametrizations, which can result in added uncertainty at the end of the whole modelling chain. On the other hand, the more process detail is incorporated in each component of the modelling system, the more data and computational resources will probably be needed, the more time it will probably take to execute and the more demanding its setting up and calibration will be. These aspects can all contribute to render standard methods unfeasible or undesirable for some applications and in this context low-complexity data-driven approaches can be good alternative solutions. In the next chapters of this thesis such simplified solutions are comprehensively analysed, and compared to standard approaches, in terms of their ability to map flood hazard, in terms of their concrete advantages and limitations and different applications.



## Chapter 2

# State-of-the-art hydrogeomorphic flood hazard mapping

Simplified data-driven methods for delineating flood-prone areas aim at informing users rapidly in the absence of detailed studies and in data-scarce (e.g., ungauged basins) or resource-limited settings, while enabling large-scale analyses without incurring in high computational time penalties. In some cases, these methods may not depend explicitly on hydrological conditions nor relate directly to event frequency, duration or magnitude, nor to local settings or antecedent conditions. Instead, they can be based solely on causality between historical floods and the floodplain hydraulic geometry (e.g., Bhowmik, 1984).

Leopold and Maddock Jr. (1953) evidenced how the water-surface width, mean depth and mean velocity, at different points of the stream channel in a downstream direction, increase on average with discharge as simple power functions. These relationships hold even for very different river systems and tributaries up to bankfull stage. Dodov and Foufoula-Georgiou (2005) provided evidence of two scaling regimes in maximum annual floods, one below and one above the bankfull flow. The authors confirmed that inundation depth was proportional to bankfull depth at any given scale and that bankfull depth scaled with the contributing area.

Rodda (2005) used historic floods and local knowledge to reason about the patterns (summer rainfall and snow melt) of fluvial flooding in the Czech Republic and generate a set of synthetic flood events. The events were characterised by a growth curve index, given by the ratio of peak flood (maximum mean daily flow) to the median annual flood, and the number of river basins typically affected by each pattern. Flood magnitudes were defined based on the growth curve index that increased in the downstream reach if the two upstream tributaries were flooding with the same magnitude. The authors used a conversion factor obtained from observed data to transform the growth curve index into

peak flow at each gauging station. Using rating curves, flows were converted to water levels and bankfull stage was subtracted, assumed as the water level associated with the 1 to 2 year return period flood. Flood depth was assumed invariant along the reach and was propagated over the floodplain associated to each reach. The authors used a buffer of maximum flooded area and, for each cell within that buffer, the elevation of the river network cell, connected by the flow path of maximum slope, was subtracted from the flood depth at the reach. This approach was found by the authors to represent well the flood extents and some of the errors found were arguably related to DEM resolution and to the presence of ponds.

Based on previous findings, Dodov and Foufoula-Georgiou (2006) used a geomorphic approach to delineate the floodplain within regions of similar climatic and geologic conditions. Inundation depth was estimated at points along the stream network using scaling relations of bankfull depth and a regional proportionality coefficient. Similarly, Nardi et al. (2006) and Nardi et al. (2013) presented a flat-water approach to delineate the floodplain (i.e., simple intersection of a specified flood depth with the surrounding digital topography). They used a variable stage at each stream pixel from a stream-order averaged linear scaling relation. The relation was found by generalising an outlet flood magnitude to each stream pixel, matching it to the discharge obtained from the Manning uniform flow equation, with constant roughness, to estimate the stage.

Rennó et al. (2008), Nobre et al. (2011) and Gharari et al. (2011) found that, similarly to the topographic wetness index by Beven and Kirkby (1979), the Height Above the Nearest Stream (HAND, or incision as defined by Bhownik 1984) previously used to characterise valley-bottom settings (Williams et al., 2000) was correlated to soil water content distribution.

Manfreda et al. (2011) investigated the correlation between the topographic wetness index and flood-prone areas and proposed a tailored version of it by weighing the local upslope contributing area. Flood-prone areas were delineated by searching for the optimal combination of weight and index threshold that would minimise the sum of false negatives and false positives, in relation to a reference flood hazard map. As a result of the discussion of this paper, Cartier and Fuamba (2013) suggested that the authors also tested the downslope index proposed by Hjerdt et al. (2004) that represents a “smoothed estimate of the local slope defined over a length variable with the morphology”, as described in Cartier and Fuamba (2013). The authors found that the modified topographic index outperformed the downslope index in the Arno River Basin.

Degiorgis et al. (2012) proposed the delineation of flood-prone areas from a location where a flood map exists to one where it does not, for example from a training area

to the entire river basin. This extrapolation procedure was achieved by identifying the isoline or contour (optimal threshold) of a chosen flood descriptor that best approximated the areal extent of an existing flood map (benchmark) for a given return period. Flood descriptors can be defined as quantitative layers produced with DEM that classify the digital terrain in such a way that it correlates to the characteristics of hydrological or hydraulic processes; DEMs are digital elevation datasets representing the Earth's surface that are distributed as gridded values representing local terrain elevations (Tavares da Costa et al., 2019b). Among the tested flood descriptors, the elevation difference and the distance to the nearest stream were found to be the best performing flood descriptors.

Jalayer et al. (2014), Risi et al. (2014), De Risi et al. (2014), and Risi et al. (2015) have developed a fast procedure for hazard zoning. The authors relate, through simple linear regression, each flood depth contour, of a given return period flood hazard map, to the topographic wetness index optimal threshold, the value with the highest probability of correct delineation of flood-prone areas given by a maximum likelihood estimate. De Risi et al. (2017) used the same approach but further included information on historical floods through Bayesian updating. With this, the authors were able to delineate and extrapolate flood-prone areas based on the functional relationships and calculate risk.

Rathjens et al. (2016) compared the topographic wetness index, to a relative slope position index, to the method introduced by Nardi et al. (2006) and to the method introduced by Degiorgis et al. (2012) in their ability to delineate floodplains. The authors found that both the slope position and the method introduced by Nardi et al. (2006) performed better in the four selected watersheds in the US, resulting in less under- and overestimation compared to the topographic wetness index and the method introduced by Degiorgis et al. (2012), respectively. On the other hand, McGrath et al. (2018) compared the HAND, as a predictor of flood extent and depth, to the flat-water approach and to a water surface obtained through inverse distance weighting. It was found that the HAND performed best for two study areas in Canada.

Manfreda et al. (2014) compared the linear binary classification, using the modified topographic index and the indices introduced by Degiorgis et al. (2012) to the hydrogeomorphic approach of Nardi et al. (2006). Their study confirmed that the classification using the HAND was reliable for delineating flood-prone areas, something also suggested in a remote sensing application by Westerhoff et al. (2013) using the HAND model and later confirmed by Nobre et al. (2016) and by Zheng et al. (2018a). Manfreda et al. (2015); Samela et al. (2016, 2017) further tested the ability of single and composite indices to classify flood-prone areas. The Geomorphic Flood Index (GFI, Samela et al. 2017), was found to be the best performing and the most consistent index (Manfreda et al., 2015)

and (Samela et al., 2016, 2017). Building upon this, Samela et al. (2017) and Tavares da Costa et al. (2019a) successfully delineated flood-prone areas at the continental scale (pan-European maps) for several return periods by dramatically reducing computational time and costs, opening new possibilities for flood risk assessment and management. In Tavares da Costa et al. (2019a), optimal thresholds of the GFI were also shown to be positively correlated to flood extents associated with specific return periods. Based on the previous findings, Samela et al. (2018) and Tavares da Costa et al. (2019a) developed offline and online tools to automatically delineate flood-prone areas; while Manfreda and Samela (2019) introduced a method to estimate flood depth from the GFI directly.

Clubb et al. (2017) tested the HAND model and the slope in delineating floodplains. The topographic threshold corresponding to the transition between hillslope and floodplain was determined by identifying the value of both the HAND and the slope at which the difference between their probability density function and the Gaussian function was less than 1%.

Jafarzadegan and Merwade (2017) proposed an enhancement to the extrapolation of flood-prone areas by regressing optimal thresholds of the HAND model with catchment characteristics and by introducing a probabilistic threshold classifier to generate probabilistic flood maps (Jafarzadegan et al., 2018; Jafarzadegan and Merwade, 2019).

Afshari et al. (2018) used the HAND model to estimate reach-average rating curves by employing the Manning's equation with hydraulic radius, estimated for each HAND isoline (Zheng et al., 2018a,b). The rating curves allow one to retrieve a flood depth, associated to a HAND isoline, for an input flood quantile and rapidly delineate flood hazard. Godbout et al. (2019) analysed the errors of this approach and realised that it underperformed in short reaches with extreme slopes.

Morrison et al. (2018) utilised the hydrogeomorphic approach of Nardi et al. (2013) to partition a US wetland database into floodplain wetlands. The authors assessed the relationship between the presence of levees and floodplain wetland areas to understand how one affected the other. It was found that floodplain wetland area is linked to maximum Strahler stream order (Strahler, 1964) and that the presence of levees are linked to larger wetlands, particularly at lower Strahler stream order. Based on the same approach, Scheel et al. (2019) assessed the impact of the presence of levees in the delineation of floodplains for the Wabash Basin in the US. The authors found that levee heights captured by the digital terrain decreased the floodplain area. Nardi et al. (2018) investigated the performance of the hydrogeomorphic approach in mapping the floodplain of two river basins in the US. The authors found that results were consistent with reference flood maps, even when scaling parameters varied significantly, and that man-made structures captured by

the DEM were responsible for some of the differences found. Annis et al. (2019) investigated how the power law coefficients in the same hydrogeomorphic approach changed with spatial resolution of DEMs, Strahler stream order and return period and how it performed, using valid ranges of the coefficients, in several river basins of different sizes across Europe. The authors used standard flood hazard maps for different return periods as reference and compared the results to the ones obtained by thresholding the HAND, the constant hydrological flow distance to the stream network, the local slope and the topographic wetness index. Since the power law coefficients showed a strong linear correlation between themselves, the authors kept the constant of the power law fixed and tested only how the exponent varied. Optimal exponent values for the delineation of floodplains increased with an increase in spatial resolution and return period and decreased with Strahler stream order. Furthermore, the best performances were found for smaller return periods and the hydrogeomorphic approach outperformed the results obtained with selected topographic indices. Nardi et al. (2019) delineated the global floodplains by employing the hydrogeomorphic approach with a power law exponent calibrated using a global flood map for the 200-year return period Dottori et al. (2016). A measure-of-fit of the delineated floodplains across Europe was found to support the use of a constant value for the power law coefficient.

Speckhann et al. (2018) have combined the HAND model with frequency analysis of annual maximum water levels to map flood hazard in the Itajaí River Basin in Brazil. The water level corresponding to a specific return period, retrieved from a fitted continuous probability distribution, was first corrected by adding the difference between staff gauge height and the corresponding HAND value before using it to threshold the HAND model. The authors found some sensitivity of the approach to channel initiation and low sensitivity to DEM spatial resolution.

The hydrogeomorphic flood hazard mapping approach has opened the possibility for fast automatic delineations of flood-prone areas without requiring great amounts of computing power, data inputs or even technical knowledge, for example in setting up and calibrating hydrological and hydrodynamic models. It allows for multiple analyses over large-scales, continental to global, which are still a challenge for standard approaches without some degree of simplification. It releases a study of the burden of sophisticated modelling, allowing one to focus, even if in a more coarser way, in other type of questions. For example, Elshorbagy et al. (2017) used the HAND model and the distance from the nearest drainage topographic indices to define hazard classes in Canada. These classes were then combined with exposure assessed using nightlights from satellite imagery to estimate risk. The simplicity of the hydrogeomorphic flood hazard mapping approach

also allows for straightforward integrations in modelling frameworks as shown by Zheng et al. (2018a), or even in web applications as presented in this work and in Tavares da Costa et al. (2019a), which brings a new type of usability and dissemination of products. As a static approach, the hydrogeomorphic flood hazard mapping, as it is, will hardly be able to replicate more than the envelope of major floods (as explained in the following chapters). Therefore, there should not be an expectation of replacement, but instead of complementarity of this approach to standard ones. The hydrogeomorphic flood hazard mapping also has a number of drawbacks, for example the quality and detail of the DEM can largely influence its outcomes, but this is not exclusive of this approach it also affects any hydrological-hydrodynamic simulation. Not only the DEM pixel size, but also vertical accuracy, the representation of man-made structures and of floodplain features can be a cause of unrealistic results, this is briefly explored in one of the Appendices and in Tavares da Costa et al. (2019b). Most importantly, the HAND model or any of the topographic indices based on it, such as the GFI, need to be improved substantially over flat-areas as they present a tendency for overestimation. This is because there are no topographic constraints, in height, that intersect with the flat-water surface earlier.

In this thesis the hydrogeomorphic approach serves as the basis for flood hazard mapping, downscaling and regionalisation. It is used in the following chapters as a starting point for developing novel data-driven methods and for implementing novel technologies, of which their value is demonstrated in several case studies and applications, across spatial scales and with datasets of different sizes.

# Chapter 3

## Flood hazard: mapping, downscaling and regionalisation

### 3.1 Classification of flood-prone areas

The methodology presented in this thesis is based on the prior definition of a composite flood index, the GFI, whose isolines (i.e., represented by unique thresholds values,  $TH$ ) are used to classify benchmark flood extents.

#### 3.1.1 Geomorphic Flood Index

A flood descriptor is a raster layer that is able to identify the propensity to flooding in a given area. It can be obtained by combining different factors into a unique raster layer (e.g., from terrain analysis, land use and land cover, and so forth). In this thesis, the flood descriptor is presented as a combination of hydrogeomorphic factors and is called the Geomorphic Flood Index (GFI) (Samela et al., 2017), which requires several steps to be computed. The GFI is a raster layer estimated from pre-processed terrain analysis layers extracted from a DEM. The computation of the GFI is given by:

$$GFI_{ij} = \ln \left( \frac{h_{ij}}{H_{ij}} \right) \quad (3.1)$$

It is composed of two terms, computed following the steepest downslope path given by a convergent eight direction flow model (abbreviated as D8 flow model). The first term,  $h_{ij}$  (in meters), consists of an empirically derived stage estimated by means of a power law hydraulic scaling relation of bankfull depth and upslope contributing area (Nardi et al., 2006; Dodov and Foufoula-Georgiou, 2006; Manfreda et al., 2015; Samela et al., 2016, 2017). The empirically derived stage in each cell under analysis ( $i, j$ ) is computed using

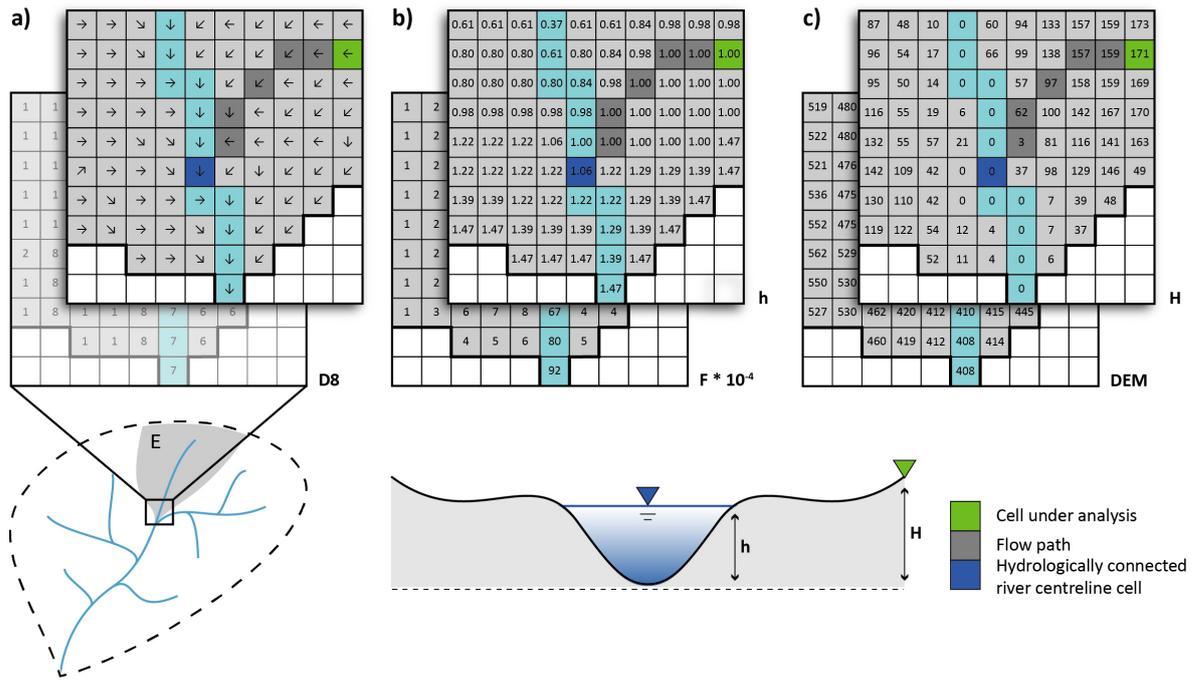
the upslope contributing area specific to the river centreline cell hydrologically connected to cell  $i, j$  and following the D8 flow model:

$$h_{ij} = a \left( A_k^{ch} \right)^n, \quad \text{with } A_k^{ch} = F_k^{ch} \times \text{cellsize} \quad (3.2)$$

For simplicity, the power law constant  $a$  and exponent  $n$  are assumed transferable and kept constant with values of 0.1 and 0.4 (Samela et al., 2017), respectively. The second term of the GFI consists of the HAND (Rennó et al., 2008; Nobre et al., 2016) calculated between the cell under analysis ( $i, j$ ) and the river centreline cell hydrologically connected to cell  $i, j$ , following the D8 flow model:

$$H_{ij} = z_{ij} - z_k^{ch} \quad (3.3)$$

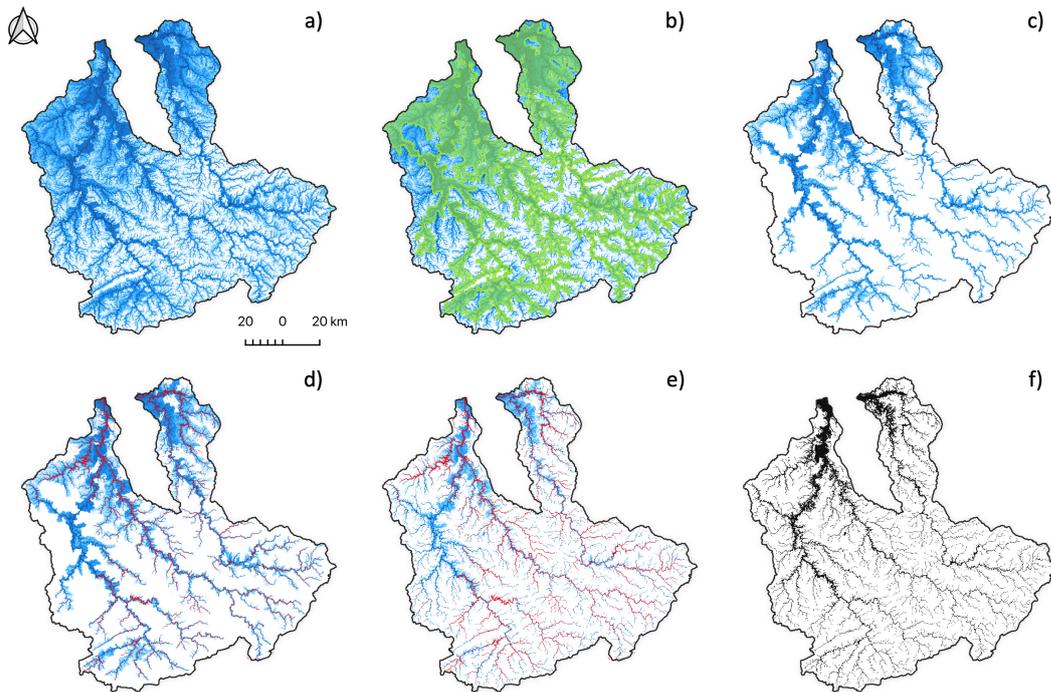
with  $z_{ij}$  the DEM elevation value of the cell under analysis and  $z_k^{ch}$  the DEM elevation value of the hydrologically connected river centreline cell. The GFI is rescaled before use to a range of values lying between 0 and 1, corresponding to low (i.e., away from the river centreline) and high hazard levels (i.e., near the river centreline), respectively. Note that moving away from the river centreline,  $H_{ij}$  increases while the GFI decreases. Scaling is achieved by resorting to the minimum and maximum values of the GFI. The rescaled GFI can effectively be used as a classifier of flood-prone areas (Manfreda et al., 2015; Samela et al., 2016, 2017) and of the extent of the envelope of major floods that is confined to the floodplain, between the active river channel at bankfull and the surrounding marked topography. The GFI computation is summarised in Fig. 3.1.



**Figure 3.1:** Two components that constitute the Geomorphic Flood Index (GFI). **a)** D8 flow direction raster for a portion of a catchment. **b)** The empirically derived stage in each cell under analysis computed as a power law of bankfull depth and upslope contributing area of the hydrologically connected river centreline cell. **c)** The DEM elevation difference between the cell under analysis and the hydrologically connected river centreline cell. At the lower left, a river basin representation showing the catchment E, in grey, and the river network, in blue. At the lower centre, a cross-section of the river channel and floodplain.

### 3.1.2 Threshold Binary Classification

The threshold binary classification, introduced by Degiorgis et al. (2012), is adopted in this study to find the a unique  $TH$  that produce the best possible representation of the benchmark flood maps in terms of flood extent (see Fig. 3.2 for an illustration of the workflow). The resulting unique  $TH$  per catchment of the Thames, Weser, the upper Rhine and the upper Danube river basins are subsequently used as the target variable for training the regression models.



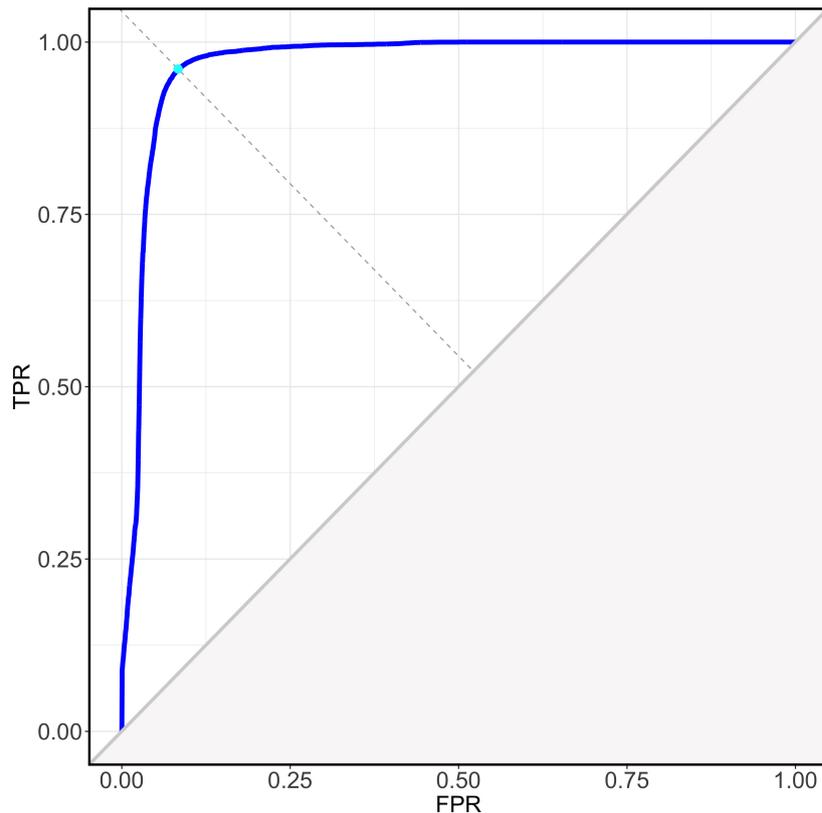
**Figure 3.2:** Workflow involved in the discrete statistical classification of the Geomorphic Flood Index (GFI) for a generic river basin. **a)** GFI layer. **b)** Overlay of fixed buffer (green) around the river network centerline. **c)** Calibration area corresponding to the clipped GFI using the fixed buffer. **d)** Overlay of benchmark (red) and start of the threshold binary classification. **e)** Search for optimal  $TH$  by means of an objective function. **f)** Best possible representation of flood-prone areas (black) after classification.

The binary classifier consists of a mathematical optimisation that outputs the best possible representation of known binary values from a benchmark. The algorithm starts by creating, through image segmentation of the flood descriptor (the GFI), a binary flood mask associated to each  $TH$ , out of a large number of possible values from zero to one, hereinafter called the segmented GFI. The algorithm then searches all these binary cases to find the one that best approximates the benchmark. The optimal case, represented by a unique  $TH$ , is indicated by the maximisation of a specific objective function that expresses the correctness of a representation.

The classifier performs better when binary categories are symmetrically distributed (Kubat et al., 1998), when raster cells of one binary class are not greater in number than the other. Therefore, a portion of the GFI, namely a classification area corresponding to a fixed buffer (ca. 1 km) around the river network centreline of the largest benchmark, is adopted in order to handle class imbalance. The classification area allows one to discard the number of flood-free raster cells in excess.

### 3.1.3 Objective function

The maximisation of the True Skill Statistic ( $TSS$ ) (Peirce, 1884; Youden, 1950; Stephenson, 2000; Baker and Kramer, 2007) is adopted in this study as the classification rule that defines which  $TH$  is optimal to select for each catchment and return period that results in the optimal delineation of flood-prone areas. The  $TSS$  represents the point of maximum forecast value of the classifier (see Fig. 3.3); in other words, it is the point in the Relative Operating Characteristic (ROC, see Chapter 3) that has the maximum perpendicular distance from the line of no-skill (Manzato, 2007), which translates to a good representation of the binary categories in the benchmark. the  $TSS$  can be seen as the probability of making an informed decision with regard to the proportion of correctly predicted classes, assuming false positives to be as undesirable as false negatives. In other words, misclassifying areas that are flood-prone is as serious as misclassifying flood-free areas. In this thesis, the maximum  $TSS$  defines the optimal  $TH$  for a classification area and return period, i.e., the  $TH$  that best captures the reference flood extent. It is determined for all the considered segmentations of the GFI layer during classification.



**Figure 3.3:** Relative Operating Characteristic (ROC) curve obtained by plotting the false positive rate (FPR) against the true positive rate (TPR) of each unique  $TH$  of the flood descriptor for a generic river basin, in blue, and intersection point with the perpendicular of the line of no-skill (i.e. the diagonal bisector) that corresponds to the point of maximum True Skill Statistic ( $TSS$ ), in cyan.

The TSS has been used elsewhere with success by several authors (Stephenson, 2000; Manzato, 2007; Bartholmes et al., 2009; Alfieri et al., 2013) and can also be interpreted as the probability of making an informed decision in terms of the proportion of correct binary categories, assuming for this specific study that the misclassification of flood-prone areas is as undesirable as the misclassification of flood-free areas. The TSS is defined as:

$$TSS = TPR - FPR = \frac{tp}{(tp + fn)} - \frac{fp}{fp + tn} = \frac{tp \times tn - fp \times fn}{(tp + fn)(fp + tn)} \quad (3.4)$$

with  $TPR$ , the true positive rate or the probability of a correct hit (Fawcett, 2006), given by:

$$TPR = 1 - FNR = \frac{tp}{(tp + fn)} \quad (3.5)$$

with  $FNR$  the false negative rate;  $FPR$  the false positive rate or the probability of an incorrect hit (Fawcett, 2006), given by:

$$FPR = 1 - TNR = \frac{fp}{(fp + tn)} \quad (3.6)$$

with  $TNR$  the true negative rate;  $tp$  the number of raster cells marked as flood-prone in both the segmented GFI and the benchmark flood hazard maps;  $fn$  the number of raster cells marked as flood-free in the segmented GFI but marked as flood-prone in the benchmark;  $tn$  the number of raster cells marked as flood-free in both the segmented GFI and the benchmark; and,  $fp$  the number of raster cells marked as flood-prone in the segmented GFI but marked as flood-free in the benchmark. Where the values of  $tp$ ,  $fp$ ,  $fn$  and  $tn$  together constitute the  $2 \times 2$  binary contingency matrix (see Table 3.1), from which several metrics can be derived, such as the  $TSS$ . The  $TSS$  is negative when the segmented GFI has a higher number of  $fp$  and  $fn$  than  $tp$  and  $tn$ ; it is positive when the opposite happens, with  $TSS = 1$  indicating that the segmented GFI perfectly matches the benchmark. The case of  $TSS = 0$  implies that the classifier does not provide any useful information. We note that  $TSS$  maximises the success of the method but not its utility, i.e. benefits and costs of any particular hydrogeomorphic flood hazard delineation for decision-making purposes are not weighed in (Peirce, 1884; Baker and Kramer, 2007).

**Table 3.1:** Example of contingency matrix of possible raster cell values in a binary classification.

	<b>Benchmark flood-prone</b>	<b>Benchmark flood-free</b>
<b>Predicted flood-prone</b>	<b>tp</b> – total number of true positive instances	<b>fp</b> – total number of false positive instances
<b>Predicted flood-free</b>	<b>fn</b> – total number of false negative instances	<b>tn</b> – total number of true negative instances

## 3.2 Delineation and downscaling of flood-prone areas

The delineation of flood-prone areas is achieved by simple image segmentation using the optimal  $TH$  value resulting from the threshold binary classification. In practice, all the values in the flood descriptor layer below the threshold are marked as zero (flood-free) and all the values above are marked as one (flood-prone), resulting in a binary mask of flood-prone areas (see Fig. 3.2f).

On the other hand, the spatial resolution and coverage that is sometimes required for a particular flood hazard assessment often differs from the one promptly available. In these cases, downscaling can be used to enhance the spatial detail and coverage of an existing flood study. Downscaling is achieved by computing a flood descriptor from a DEM with a spatial resolution that is finer than the one of the benchmark flood study. Once the threshold binary classification is performed and the optimal  $TH$  is found, a delineation of flood-prone areas can be undertaken, resulting in a more spatially resolved flood extent and mask of flood-prone areas than the one of the benchmark flood study used. Additionally, in cases where the flood descriptor layer was computed beyond the classification area (see Fig. 3.2b), the delineation will also automatically result in an enhancement of coverage from the main stream to the tributaries (see Fig. 3.1e and f).

## 3.3 Regionalisation of flood-prone areas

The task of estimating flood-prone areas in catchments where flood information is not available can be addressed by regional analysis, also known as statistical regionalisation. Regionalisation consists in transferring flood information to a target catchment based on a measure of similarity, traditionally spatial proximity or physical similarity (Merz and Blöschl, 2005). In this thesis, physical similarity is explored, in particular geomorphic and climatic-hydrologic catchment characteristics (see Chapter 4), to transfer the optimal  $TH$  described in Section 1.1.2 to catchments of interest, predicting like this the flood-prone area specific to a return period.

Multivariate statistical methods can be used to describe the relationship between a unique  $TH$  and a set of explanatory variables. In practice, the unique  $TH$  values resulting from the classification are statistically related to a selection of geomorphic and climatic-hydrologic catchment characteristics, scaled and mean centred before use, by two distinct types of regression models: the stepwise regression and the random forest. The establishing of such relations allows for the prediction of  $TH$ , which represents envelope flood extents and can be used to delineate flood-prone areas, based on physical inputs of any given river basin.

### 3.3.1 Stepwise Linear Regression

Multiple linear regressions (MLR) are well-established models in hydrological sciences, particularly between flood moments and catchment characteristics (Merz and Blöschl, 2005; Haddad et al., 2012). A MLR can be defined in matrix notation as:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \quad (3.7)$$

where  $\mathbf{y}$  is the response vector of size  $n$  corresponding to the total number of catchments in the training set;  $\mathbf{X}$  is the matrix of  $k$  explanatory variables with size  $n \times (k + 1)$ , with  $x_1, x_2, \dots, x_k$  corresponding to the catchment characteristics for a particular return period; and last,  $\beta$  and  $\epsilon$  are the vectors of regression coefficients and residuals, respectively. In vector and matrix notation it can be written as follows:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (3.8)$$

Ideally, catchment characteristics should not be highly correlated to each other or to their linear combination, since multicollinearity may increase the variance of parameter estimates and potentially lead to unreliable results. Therefore, before developing the statistical models, multicollinearity is diagnosed with the variance inflation factor (VIF) that can be defined as:

$$VIF_k = \frac{1}{1 - r_k^2} \quad (3.9)$$

with  $r_k^2$  the coefficient of determination for a regression of the  $k$ -th variable with all other explanatory variables. Multicollinearity may be present when  $VIF > 10$  (Hirsch et al., 1992; Merz and Blöschl, 2005) and therefore variables above such values are considered for elimination prior to model fitting. The problem of estimating the regression coefficients, or the fitting problem, is solved by stepwise analysis with bidirectional elimination (i.e., the sequential addition) and replacement or elimination of explanatory variables based on the relative quality of each competing model. The trade-off between maximum likelihood and explanatory variables, or model's simplicity in this context, is measured by the Akaike's information criterion (AIC) (Akaike, 1974):

$$AIC = -2 \ln(L) + 2k = n \ln \left( \frac{\sum_{i=1}^n \epsilon_i^2}{n} \right) + 2k + const. \quad (3.10)$$

with  $L$  being the maximum likelihood. In practical terms, model selection is based on the minimum possible AIC obtained for competing models (Haddad et al., 2012).

To obtain classical evaluation metrics such as  $R^2$  and the root mean square error (RMSE), a 10-fold cross-validation procedure is used. The cross-validation consists of randomly splitting the dataset in ten equally-sized bins, one of which is retained for testing and the remaining are used for training. The training and testing procedure is repeated ten times, until every single bin has been selected once. Performance results from each of the ten validations are averaged to produce a single final estimation.

### 3.3.2 Random Forest Regression

Random forest (Breiman, 2001) is a rule-based machine learning method that can be used for classification (Wang et al., 2015; Coltin et al., 2016) or regression (Iorgulescu and Beven, 2004; Pappenberger et al., 2006; Prieto et al., 2019). A random forest regression model consists of an ensemble of randomised decision tree-like models  $\{T_1(\mathbf{x}), T_2(\mathbf{x}), \dots, T_j(\mathbf{x})\}$  with size  $j$  that can produce for any input vector of explanatory values  $\mathbf{x}$  an equal number of outputs  $\{\hat{y}_1 = T_1(\mathbf{x}), \hat{y}_2 = T_2(\mathbf{x}), \dots, \hat{y}_j = T_j(\mathbf{x})\}$ , whose average gives the final prediction.

Randomisation is achieved by bootstrapping, or the random draw of  $x_{ik}$  and  $y_i$  samples – each corresponding to a catchment in the original dataset – with replacement, meaning that each drawn sample can be selected multiple times to form  $j$  independent sets  $\theta_j$  that share the same distribution.

Each decision tree in the random forest consists of a top-down recursive partitioning of a specific  $\theta_j$ , using decision rules based on the  $x_{ik}$  variable in  $\theta_j$ . At each tree node, the lowest possible residual sum of squares is used as decision rule for horizontally or vertically partitioning  $\theta_j$  in two subsets with sizes  $l$  and  $r$ :

$$RSS = \sum_{i=1}^l (y_i - \hat{y}_i)^2 + \sum_{i=1}^r (y_i - \hat{y}_i)^2 \quad (3.11)$$

The partitioning procedure continues in each subset until the maximum possible tree size is reached. In practice, nodes are data divide rules and each terminal node corresponds to a best guess of the dependent variable. The tree designation comes from the resulting hierarchy of nodes that represents an acyclic undirected graph of constant decision models.

Some important advantages of the random forest method are that it does not need any specific assumption about the probability distribution (non-parametric), it works well when the relationship between explanatory variables and response is non-linear, as well as when there are high order interactions (Snelder et al., 2013). Furthermore, random

forest is relatively robust against outliers, noise and overfitting (Breiman, 2001) and can handle the problem of multicollinearity well (Cutler et al., 2007).

As opposed to MLR, a chief disadvantage of this method is that it cannot predict target values outside the range of the explanatory variables in the training dataset. Another limitation of the random forest is that it does not provide an easy understanding of the statistical relationships between explanatory variables. Even though, it does provide a simple visualisation of the model structure and of the covariate influence, in contrast to other machine learning methods, such as artificial neural networks (Shortridge et al., 2016).

The random forest regression model used in this study goes through an automatic and distributed optimisation procedure (grid search) of the setup parameters in order to find the best performing model, in terms of both accuracy and computational efficiency. In specific, the optimised parameters are the number of decision trees in the ensemble, the number of sampled variables at each tree node and the maximum depth of each tree. The optimisation of the random forest regression is achieved a priori, using a 10-fold cross-validation to obtain evaluation metrics and compare the multiple models.

# Chapter 4

## Performance metrics

### 4.1 Threshold-independent metrics: ROC analysis

ROC analysis has been used by several authors to distinguish between decision values in a classifier and their trade-offs between costs and benefits (Swets, 1973; Bradley, 1997; Fawcett, 2006; Schumann et al., 2014b). It is considered a threshold-independent performance measure, as points falling along the ROC curve (see Fig. 3.3) represent unique evaluations, in terms of  $TPR$  and  $FPR$ , of a flood descriptor segmentation, corresponding to a unique threshold value, against the benchmark. The top left corner of the ROC space represents the perfect classification, such that  $TPR = 1$  and  $FPR = 0$ ; instead, the diagonal line dividing the ROC space represents the line of no-skill. ROC curves are frequently used in flood model performance assessment (e.g., Tehrany et al., 2013, 2014; Schumann et al., 2014b; Rahmati and Pourghasemi, 2017).

In the specific context of this study, the Area Under the ROC Curve (AUC) summarises the overall discerning capability of the flood descriptor in a single threshold-independent value. As such, irrespective of the threshold, it represents the probability of correctly classifying a randomly chosen raster cell as flood-prone rather than incorrectly classifying it as such (Bradley, 1997; Fawcett, 2006). The AUC can be estimated by trapezoidal rule approximation of the definite integral and may take values from  $AUC = 0.5$ , meaning no discerning capability of the flood descriptor, to  $AUC = 1$ , the perfect classifier. The AUC is used to filter out those cases that are not well suited to serve as a classifier and may therefore impact the formulation of the statistical relationships. AUC has been used by several authors to assess performance of flood models (e.g., Tehrany et al., 2013, 2014; Schumann et al., 2014b; Rahmati and Pourghasemi, 2017).

## 4.2 Threshold-dependent metrics

### 4.2.1 Evaluation of classification outcomes

The modified Pearson's correlation coefficient for discrete dichotomous problems,  $r_\phi$  (Matthews, 1975; Cramér, 1999; Baldi et al., 2000), is used to measure the magnitude of association and direction of the linear relationship between flood-prone areas and a benchmark flood map. It was chosen because it is not sensitive to class imbalance and because it is a more balanced overall evaluation of the threshold binary classification outcomes, due to the full use of the contingency table (see Table 3.1). It can also be seen as the geometric mean of the  $TSS$  and its complementary term and is given by:

$$r_\phi = \sqrt{\chi^{\frac{2}{N}}} = \sqrt{TSS \left( \frac{tp}{tp + fp} - \frac{fn}{fn + tn} \right)} = \frac{tp \times tn - fp \times fn}{\sqrt{tp + fp \times tp + fn \times tn + fn \times tn + fn}} \quad (4.1)$$

where  $\chi^2$  is the Pearson chi-square statistical test (Pearson, 1900), with  $N$  the total sample. As a rule of thumb, it is assumed that  $1 < r_\phi < 0.5$  represent a strong positive degree of association, between  $0.5 < r_\phi < 0.3$  a moderate degree of association,  $0.3 < r_\phi < 0.1$  a weak degree of association and  $0 < r_\phi < 0.1$  a complete absence of association.

### 4.2.2 Evaluation of regression outcomes

The fit between predicted flood-prone areas, obtained through image segmentation of the flood descriptor using the predicted  $TH$  of the regression analysis, and the benchmark, can be evaluated by (Wing et al., 2017): the hit rate, or TPR; the false discovery rate:

$$FDR = \frac{fp}{fp + tp} \quad (4.2)$$

from  $FDR = 0$  (no false alarms) to  $FDR = 1$  (overprediction); the critical success:

$$C = \frac{tp}{tp + fn + fp} \quad (4.3)$$

with values ranging from  $C = 0$  when there is no match between delineated flood-prone areas and benchmark and  $C = 1$  when the match is perfect; and, the error bias:

$$E = \frac{fp}{fn} \quad (4.4)$$

that indicates whether there is a tendency towards underprediction,  $0 \leq E < 1$ , or overprediction,  $E > 1$ , with  $E = 1$  an indication of no bias.

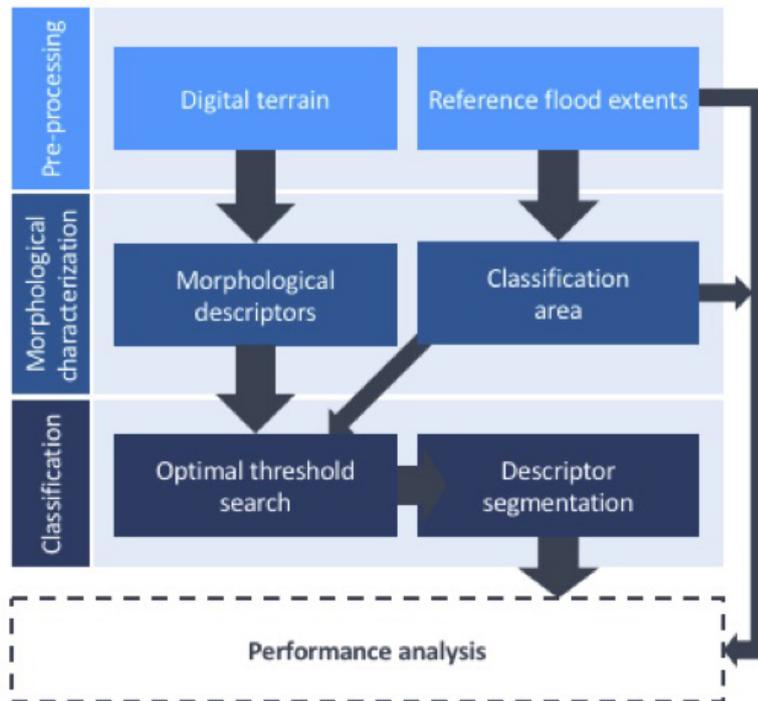
# Chapter 5

## Mapping flood-prone areas in Europe

### 5.1 Impact of scale in classification outcomes

In this section, a pan-European application of a web-based tool, developed in the scope of this work (see Appendix C), is tested and showcased for regional hydrogeomorphic analyses and detailed (i.e., tens of meters) mapping, downscaling and extrapolation of flood-prone areas across geographic scales, namely at the continental scale and at the scale of sub-catchments within a particular river basin.

This study provides a useful indication on the reliability of the selected classification method, based on a flood descriptor and its use as a linear threshold binary classification (Degiorgis et al., 2012). In particular, the flood descriptor and the downscaled and extrapolated flood-prone areas are evaluated, as well as the optimal geographical scale of implementation identified. A complete workflow is presented in Fig. 5.1.



**Figure 5.1:** General methodological workflow for classifying a morphological descriptor and assessing its performance in delineating and extrapolating flood-prone areas.

The methodology starts with the terrain analysis of a specific DEM, i.e., the computation of local slope, flow direction and upslope contributing area using for example using TauDEM utilities, and the transformation of benchmark flood studies into binary masks – in this case, the six flood hazard maps for Europe (Dottori et al., 2016) that correspond to a raster of maximum flood depths per return period (10, 20, 50, 100, 200 and 500 year) were considered.

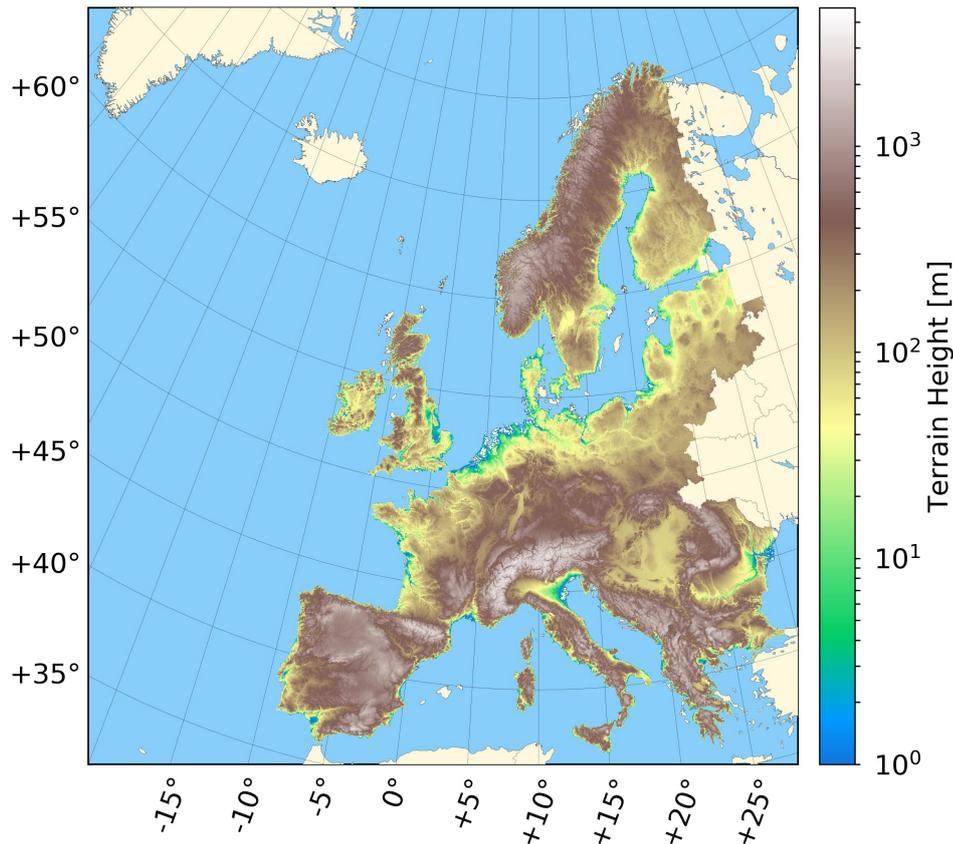
The subsequent morphological characterisation step corresponds to the computation of the selected flood descriptor from pre-processed terrain analysis layers and to the definition of a classification area (roughly a buffer of about 1 *km* around the benchmark), used to avoid some classification shortcomings (Kubat et al., 1998). The threshold binary classification of the morphological descriptor is then performed within the classification area established for each river basin or sub-catchment and for each return period, searching for the optimal threshold  $TH$  that better represents the benchmark. The optimal  $TH$  is then used to segment the morphological descriptor, by flagging all values above or equal to the threshold as flooded, while those below as flood-free, thus obtaining a hydrogeomorphic flood hazard map.

The maps outputted by the web application naturally identify all possible flood-prone areas in a region of interest and according to the DEM, which are often not limited to

the flood extents associated with a particular river reach, as usually found in numerical models, but complete river networks (i.e., main stem and tributaries) and their floodplains. Furthermore, the output maps are obtained with the spatial detail of the DEM used as basis to derive the flood descriptor, that is 25 *m*, instead of the original 100 *m* spatial resolution of the benchmarks, providing another rather important enhancement.

### 5.1.1 Study area and data

Table 5.1 lists all the open-access datasets used in this study. The EU-DEM (version 1.0, EEA, Fig. 5.2) at approximately 25 *m* spatial resolution was selected for computing the flood descriptor. This dataset is freely accessible online through the Copernicus Data and Information Access Services, funded by the European Union. At the easternmost region of Europe, where data was lacking, the EU-DEM was merged with the Shuttle Radar Topography Mission (SRTM – USGS) 1 arc-second near-global digital elevation data at approximately 30 *m* spatial resolution.



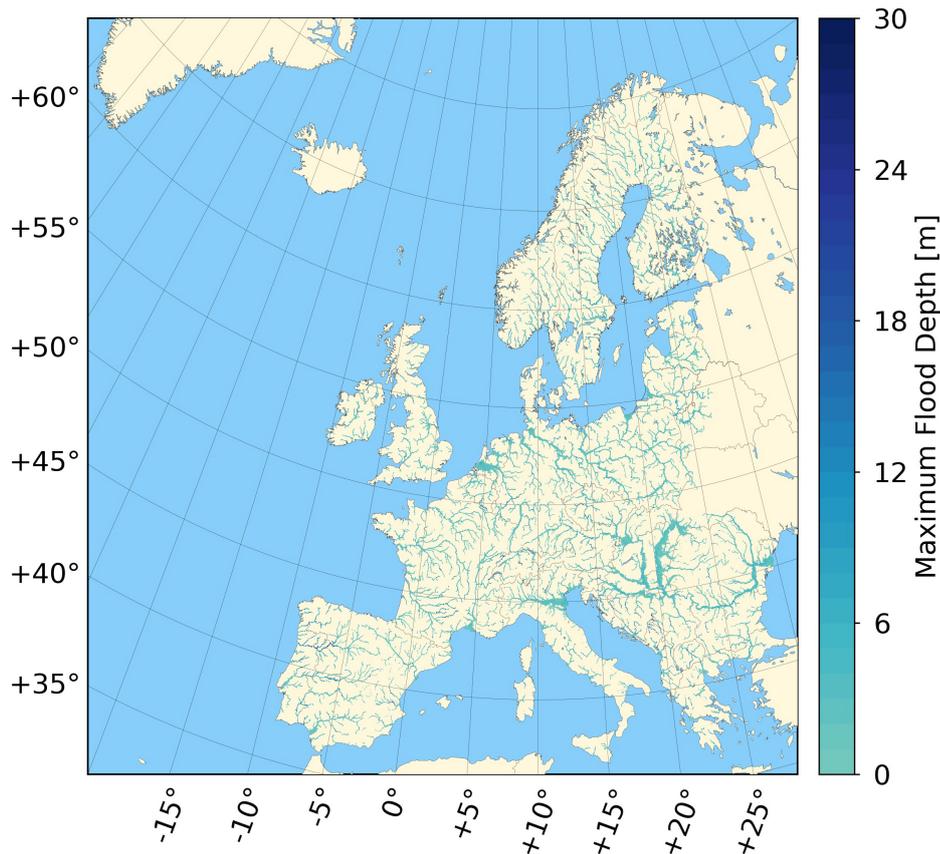
**Figure 5.2:** European digital elevation model (EU-DEM; EEA) for computing the morphological descriptor.

The undefended flood hazard maps for Europe (Dottori et al., 2016) (Fig. 5.3) were selected to derive the benchmark flood extents needed in the methodological workflow.

**Table 5.1:** Open-access datasets used in the delineation of flood-prone areas across Europe

<b>Name</b>	<b>File format</b>	<b>Description</b>	<b>Reference</b>
European Digital Elevation Model (EU-DEM), version 1.0		Terrain elevation data for Europe	EEA
Flood hazard maps for Europe	10-year return period 20-year return period 50-year return period 100-year return period 200-year return period 500-year return period	GeoTIFF raster image Maximum flood depth data for Europe	(Dottori et al., 2016)
Catchment Characterisation and Modelling (CCM) River and Catchment Database, version 2.1	Shapefile geospatial vector	Polygon coverage of selected European river basins	(de Jager and Vogt, 2010)

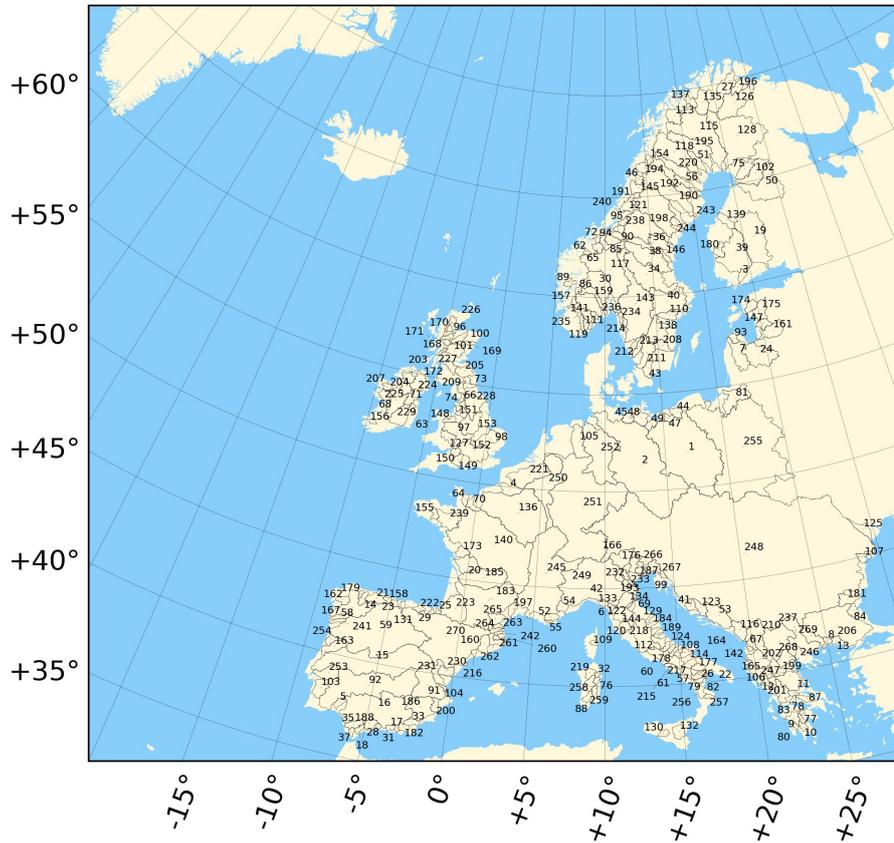
These maps are the only well-known open-access datasets available for the whole Europe. According to Alfieri et al. (2014), they were produced by simulating a European-wide discharge climatology using the LISFLOOD rainfall-runoff model (Knijff et al., 2010), by estimating synthetic design hydrographs from the derived climatology and by simulating floodplain hydrodynamics using the LISFLOOD model using the hydrographs as boundary conditions (Bates et al., 2010; Neal et al., 2012). Hydrodynamic simulations were performed by Alfieri et al. (2014) for every 5 km stretch of the river network (unspecified river network delineation method) and only for river basins with an upstream area greater than 500 km<sup>2</sup>. As highlighted by Alfieri et al. (2014), the flood hazard maps for Europe are affected by a number of uncertainties and limitations, namely space- and time-resolution issues associated with the inputs (e.g., noise in the DEM and its incapability of resolving flood defences, coarse resolution of meteorological inputs and a tendency of the method to overestimate runoff). In fact, the flood hazard map for Europe for the 100-year return period event presents TPRs between 59% and 78% and critical success between 43% and 65% evaluated based on specific national/regional hazard maps.



**Figure 5.3:** Flood hazard maps for Europe (Dottori et al., 2016) used as reference in the classification of flood-prone areas.

The study area is composed of a selection of 270 river basins (Fig. 5.4) from the Eu-

ropean mainland and continental islands that drain into the Atlantic, the Mediterranean, Black, Baltic and North Seas. Europe's major rivers have a length ranging from approximately 324 *km* to 2,860 *km* (excluding the river Volga in Russia) and their river basins are relatively small in area when compared to their counterparts at the global scale, with just 70 European river basins exceeding an area of 10,000 *km*<sup>2</sup>.



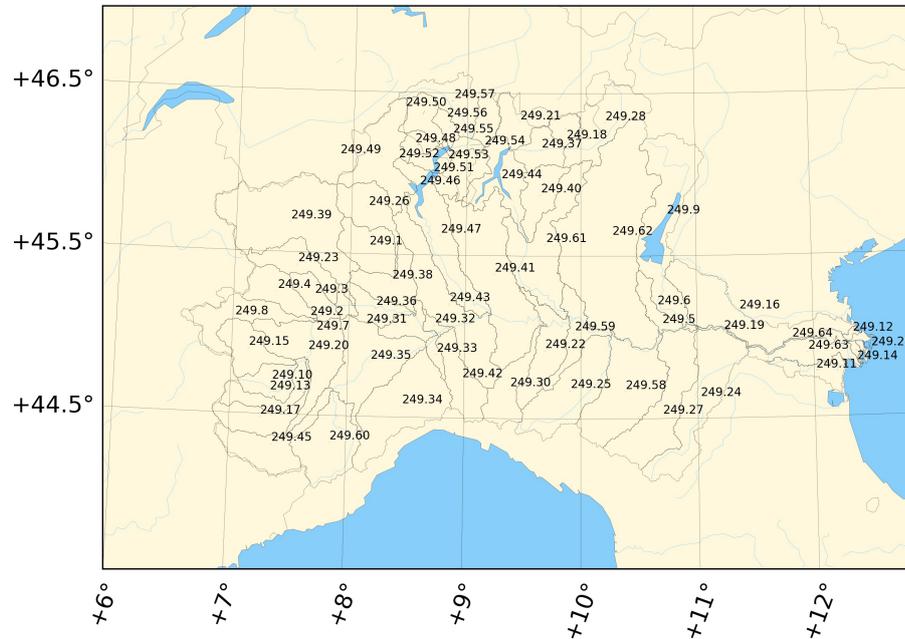
**Figure 5.4:** Study area consisting of 270 major European river basins obtained from the Catchment Characterisation and Model layer, River and Catchment Database (CCM, de Jager and Vogt 2010).

The analyses results are summarised at the river basin scale, where study catchments have been selected by filtering the Catchment Characterisation and Model (CCM) River and Catchment Database, version 2.1 (de Jager and Vogt, 2010). In particular, the study focused on streams with Strahler stream order greater or equal to 5, which is consistent with other large-scale studies of fluvial flood hazard (Muis et al., 2015). River basins corresponding to such streams and intersecting a country from the European Economic Area were selected for the evaluation of flood-prone areas.

Selected river basins are characterised by fairly diverse sizes (ranging from 500 to 750,000 *km*<sup>2</sup>), topography (Western Uplands, North European plain, partially underwater, Central Uplands and Alpine Mountains with altitudes surpassing 4,500 *m*), climate (e.g., Kottek et al. 2006; arid in the eastern Iberian Peninsula to warm temperate in most

central and northern Europe, boreal in the Pyrenees, Alps and Carpathian, with the last two also presenting a polar climate), and different degrees of anthropization.

In order to explore the impact of basin scale on the proposed methodology, 64 additional sub-catchments of the river Po in Italy (Fig. 5.5) have been included in the study (with sizes ranging from 1 to 6,300  $km^2$ ).



**Figure 5.5:** Study area consisting of 64 sub-catchments of the Po river, Italy, obtained from the Catchment Characterisation and Model layer, River and Catchment Database (CCM, de Jager and Vogt 2010).

### 5.1.2 Results and discussion

The skill of the flood descriptor is assessed by means of a ROC analysis and AUC obtained within the pre-defined classification areas and is examined for each major river basin in Europe and return period considered, as well as for the sub-catchments of the river Po in Italy.

Among the 270 river basins analysed herein, some show high discerning capability (i.e., the ROC curve approaches the top left corner and the AUC tends towards unity), while a few underperform, even though they still perform better than random guessing (i.e., curve above the bisecting line of the ROC space and AUC value greater than 0.5). The spatial distribution of AUC values is rather homogeneous with relatively high values in most of the major river basins in Europe (Fig. 5.6). Also, AUC values show a low variability between different return periods. Differences between AUC values become more pronounced within the sub-catchments of the river Po and their spatial distribution

seems to highlight a greater discerning capability of the morphological descriptor GFI where mountainous areas are prevalent (Fig. 5.7).

The classification of the flood descriptor GFI within the pre-defined classification area resulted in an optimal  $TH$  for each European river basin, each Po river sub-catchment, and each return period considered. In this section, the delineated and downscaled hydrogeomorphic flood-prone areas are compared to the benchmark flood hazard maps for Europe, to understand how well the former replicates the latter.

The resulting spatial distribution of optimal  $TH$  obtained after the classification step, maximising the  $TSS$  for all major river basins in Europe (Fig. 5.8) shows low variability. This implies that, at this scale, the optimal  $TH$  is fairly similar across major river basins in Europe. Fig. 5.9 shows much higher heterogeneity in terms of the optimal  $TH$  within the sub-catchments of the river Po: a pattern of higher  $TH$  can be seen in flatter regions and of lower  $TH$  in headwater catchments.

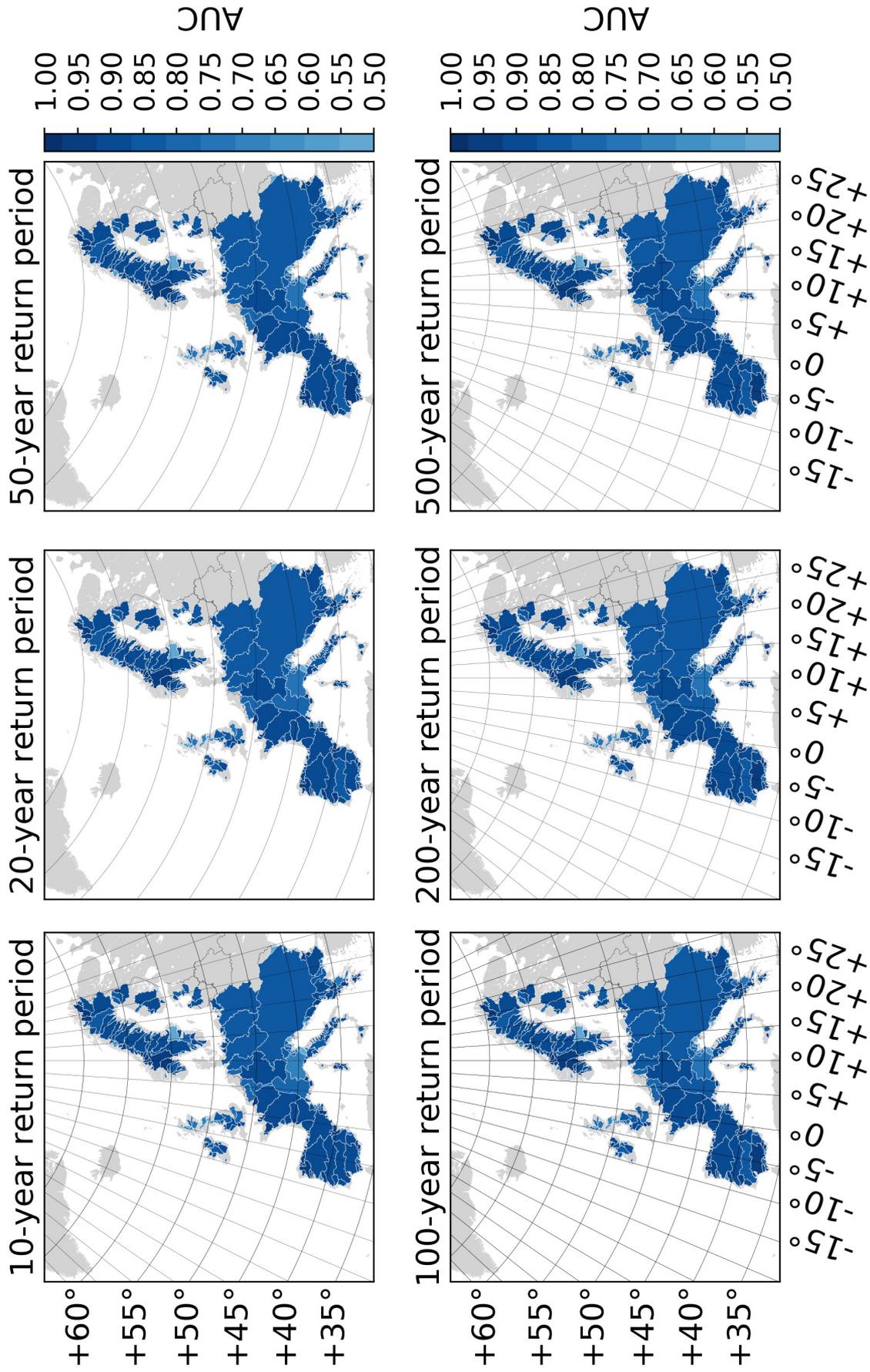
The spatial distribution of the maximum  $TSS$  accounts for the correct identification of the two critical classes, namely flood-prone and flood-free areas, and characterises the success of the classification. In general, a greater discerning capability of the morphological descriptor GFI will result in high  $TSS$  values. Fig. 5.10 shows that the spatial distribution of  $TSS$  values within major river basins in Europe is fairly homogeneous, with low variability across return periods. The variability in the sub-catchments of the river Po is more pronounced (Fig. 5.11) and shows higher  $TSS$  values for headwater catchments and lower  $TSS$  values in sub-catchments in the lower part of the basin.

The  $r_\phi$  is reported here to complement the analysis with a more balanced measure of magnitude and direction of the linear relationship between the hydrogeomorphic delineated flood-prone areas and the benchmark flood hazard maps for Europe.

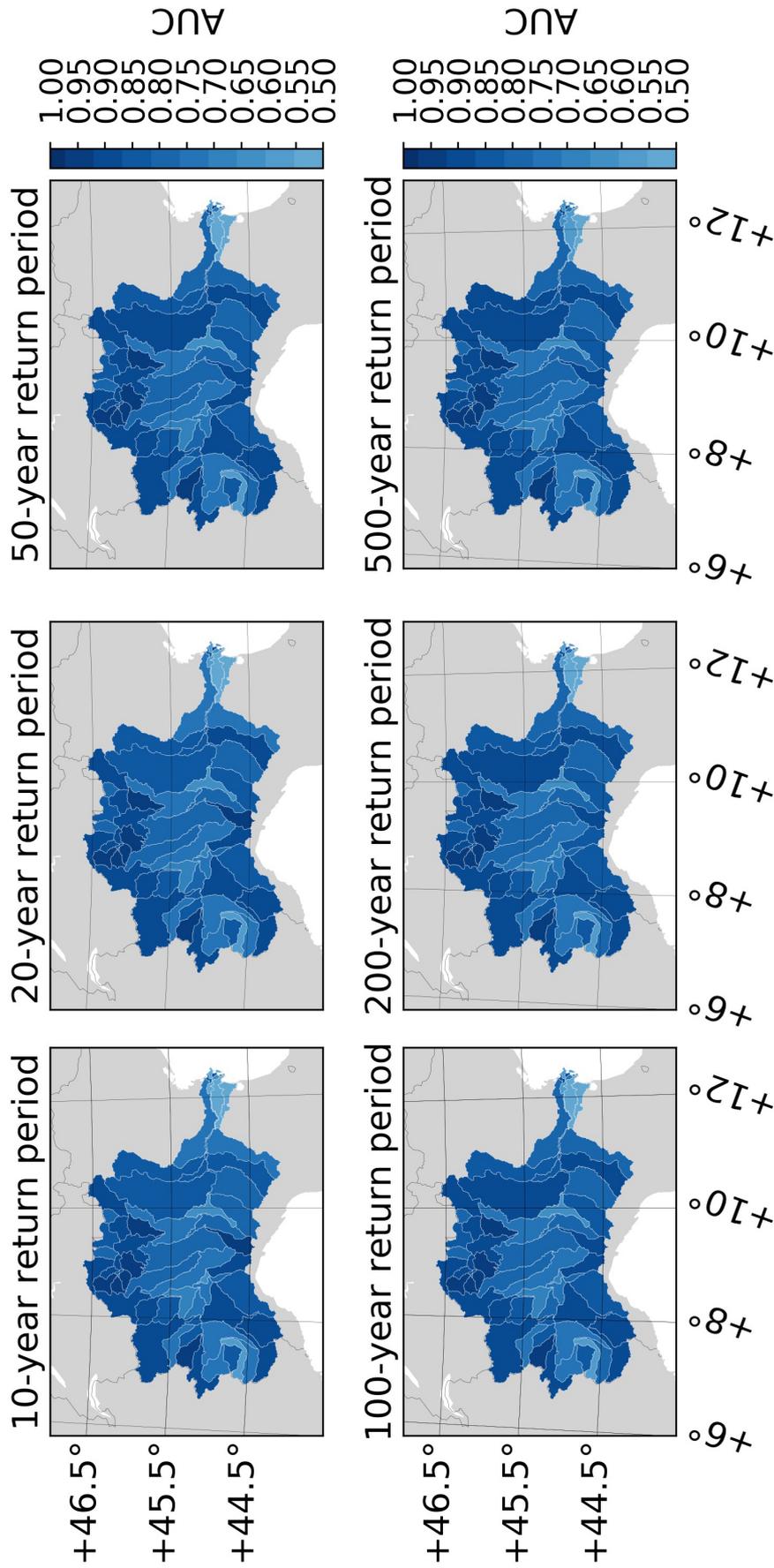
Fig. 5.12 and Fig. 5.13 highlight the correlation of the maps at the river basin and sub-catchment scales. Results show a positive linear correlation between the hydrogeomorphic flood-prone areas and the benchmark. The major European river basins analysed present  $r_\phi$  values around 0.6, with very few basins with values above or below. There is higher heterogeneity at the sub-catchment scale, with most sub-catchments presenting  $r_\phi$  values ranging from 0.4 to 0.6 and very few with values outside this range.

To complement the Geographical distributions presented, and to have a more complete overview of the results, the data distribution is summarised in a set of box plots (Fig. 5.14 and Fig. 5.15). The figures refer to the results obtained by performing the linear binary classification of flood-prone areas within the 270 major river basins of Europe and the 64 sub-catchments of the river Po.

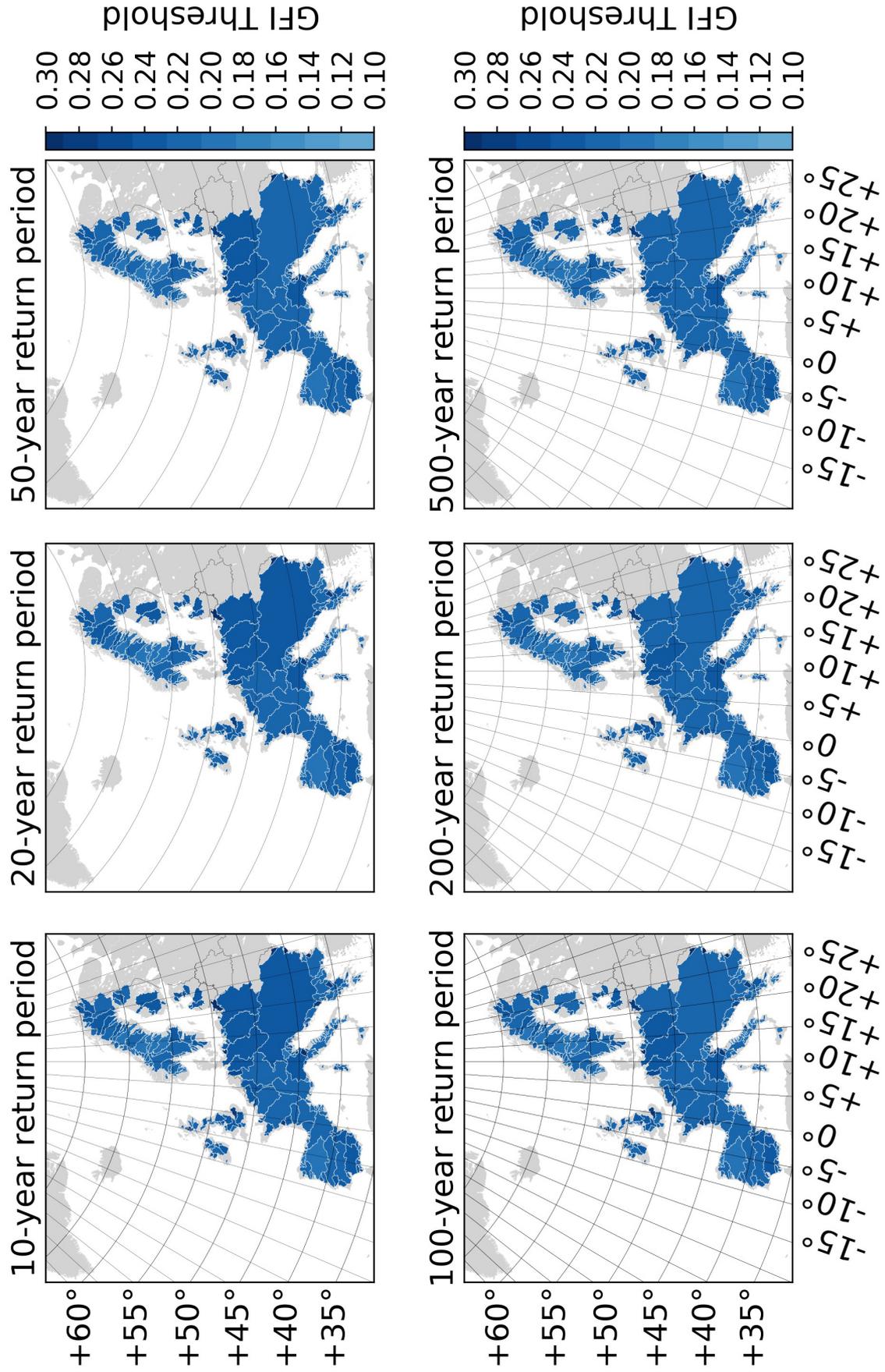
In Fig. 5.14, a lower variability of the AUC (standard deviation equal to c.a. 0.05) in



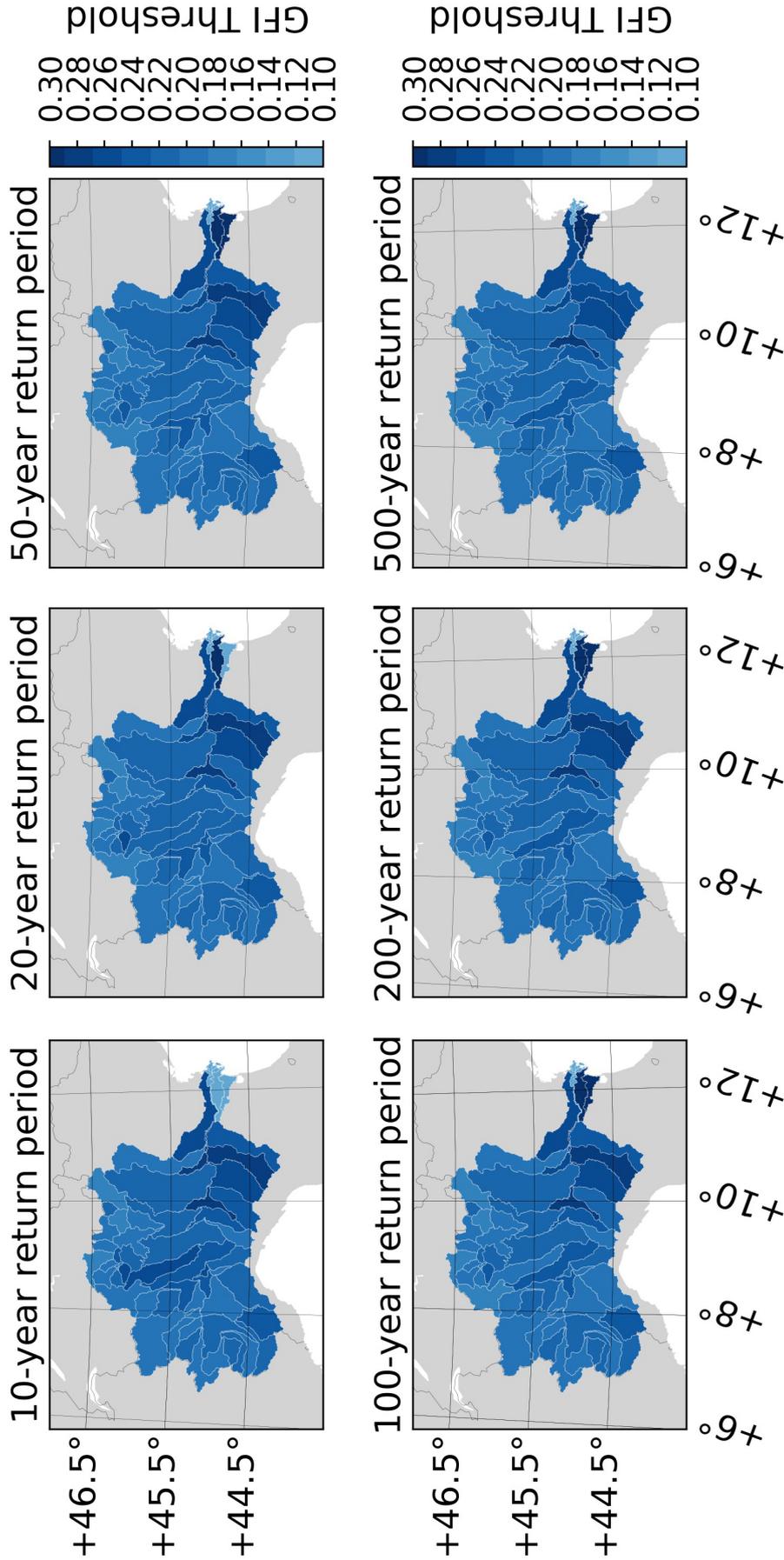
**Figure 5.6:** Spatial distribution of the area under the receiver operating characteristic (AUC) for six return periods, within the classification area set for each major river basin in Europe.



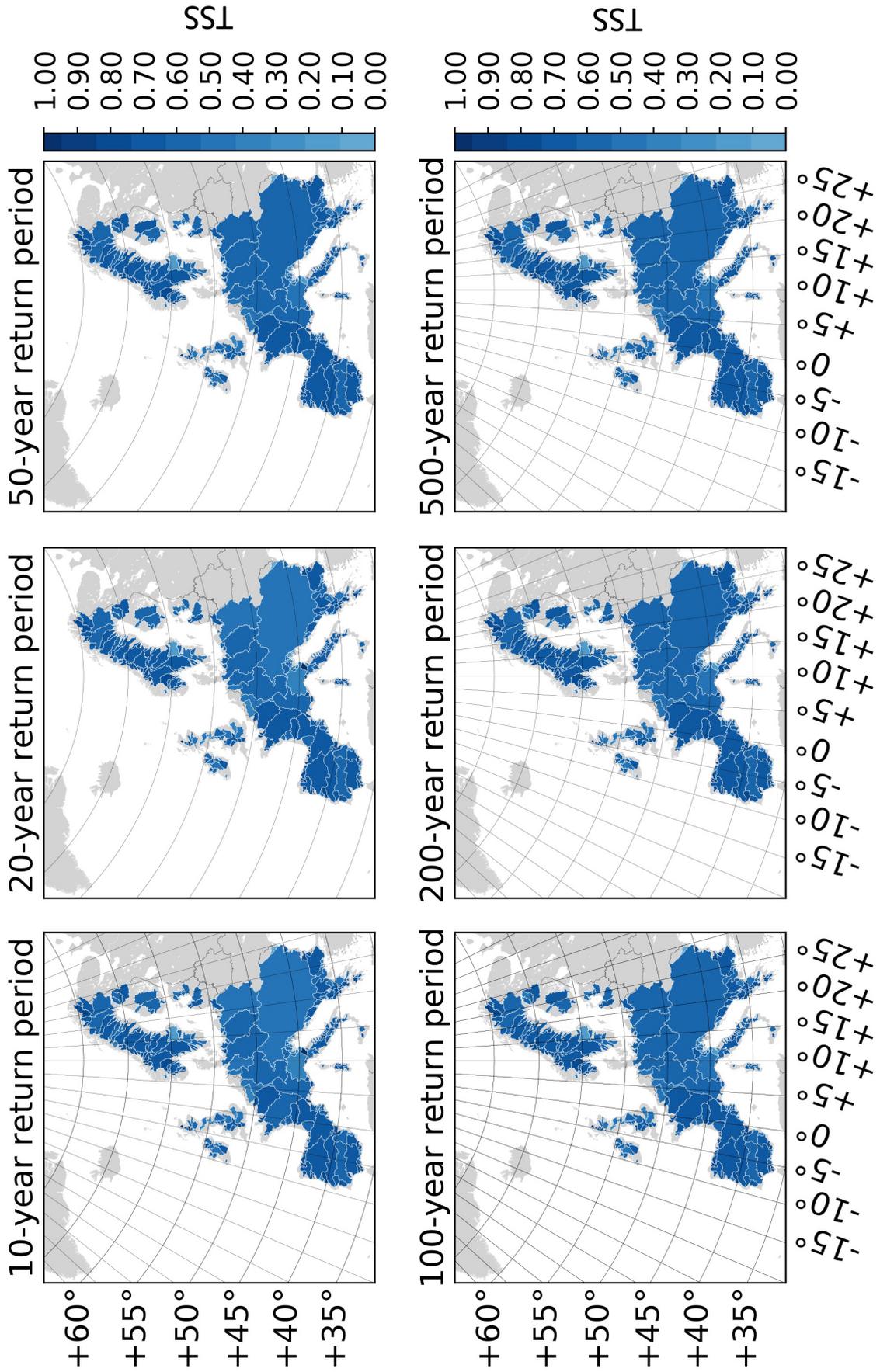
**Figure 5.7:** Spatial distribution of the area under the receiver operating characteristic (AUC) for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy.



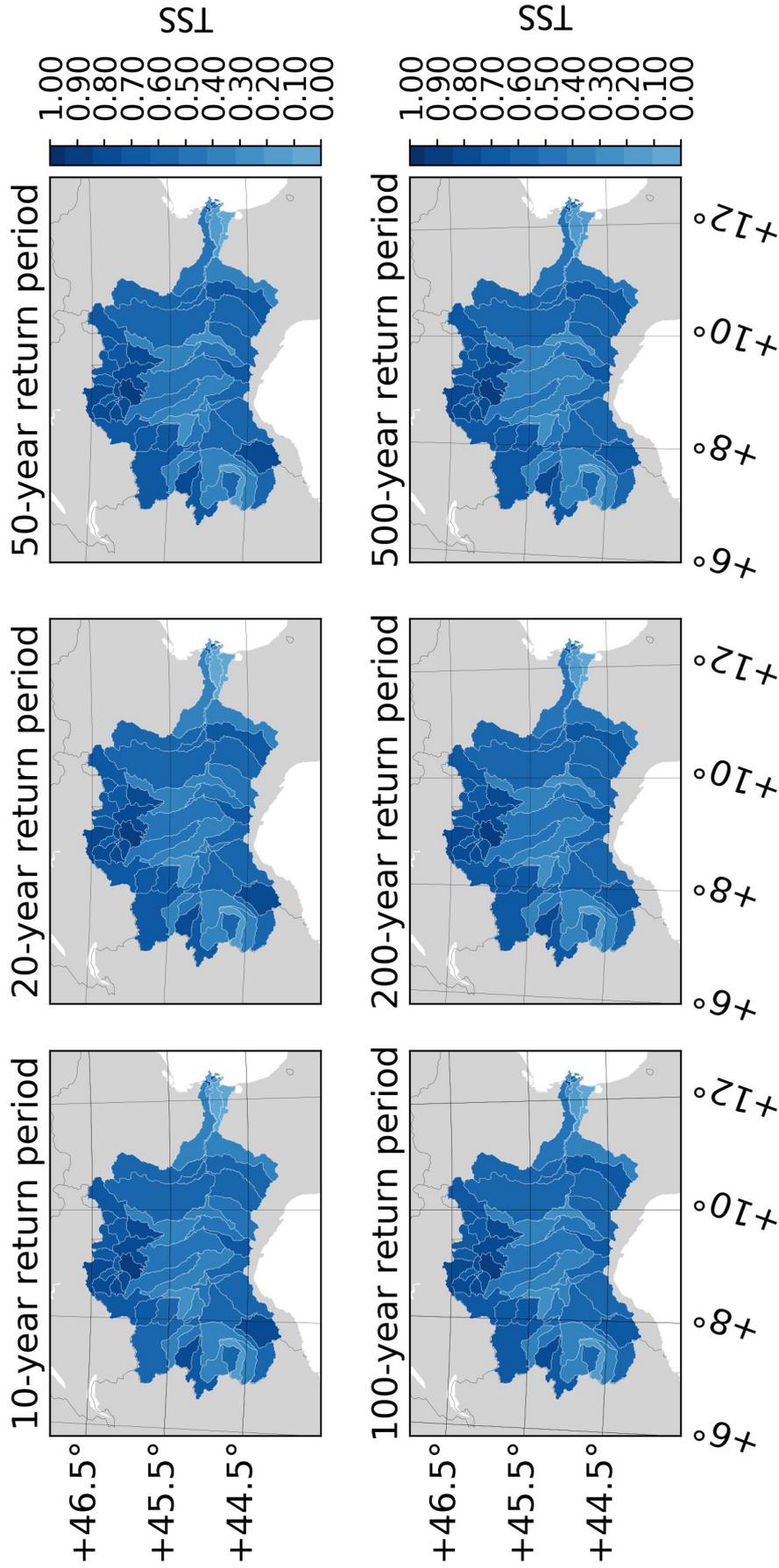
**Figure 5.8:** Geographical distribution of the optimal Geomorphic Flood Index (GFI) threshold for six return periods, within the classification area set for each major river basin in Europe.



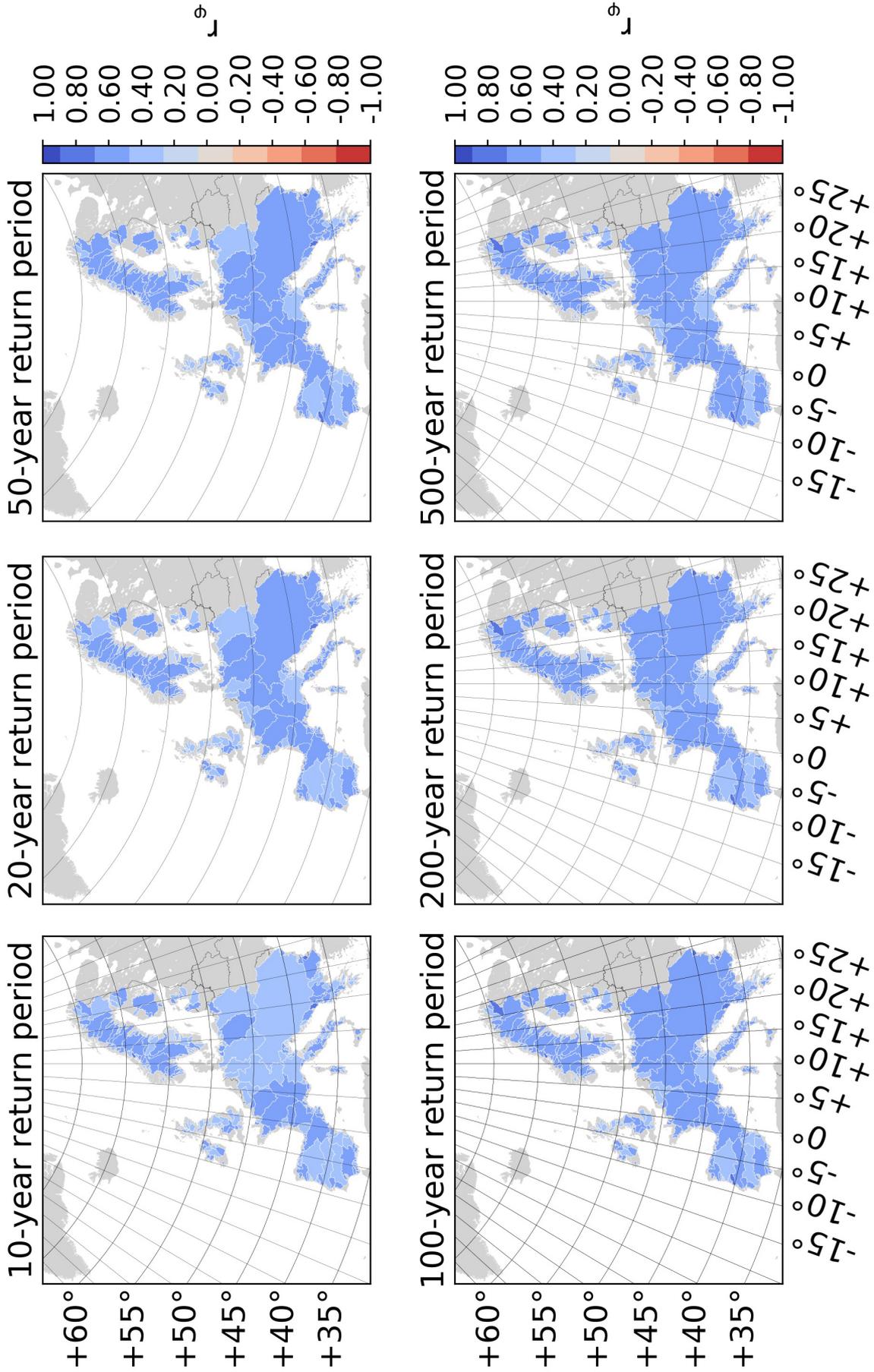
**Figure 5.9:** Geographical distribution of the optimal Geomorphic Flood Index (GFI) threshold for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy.



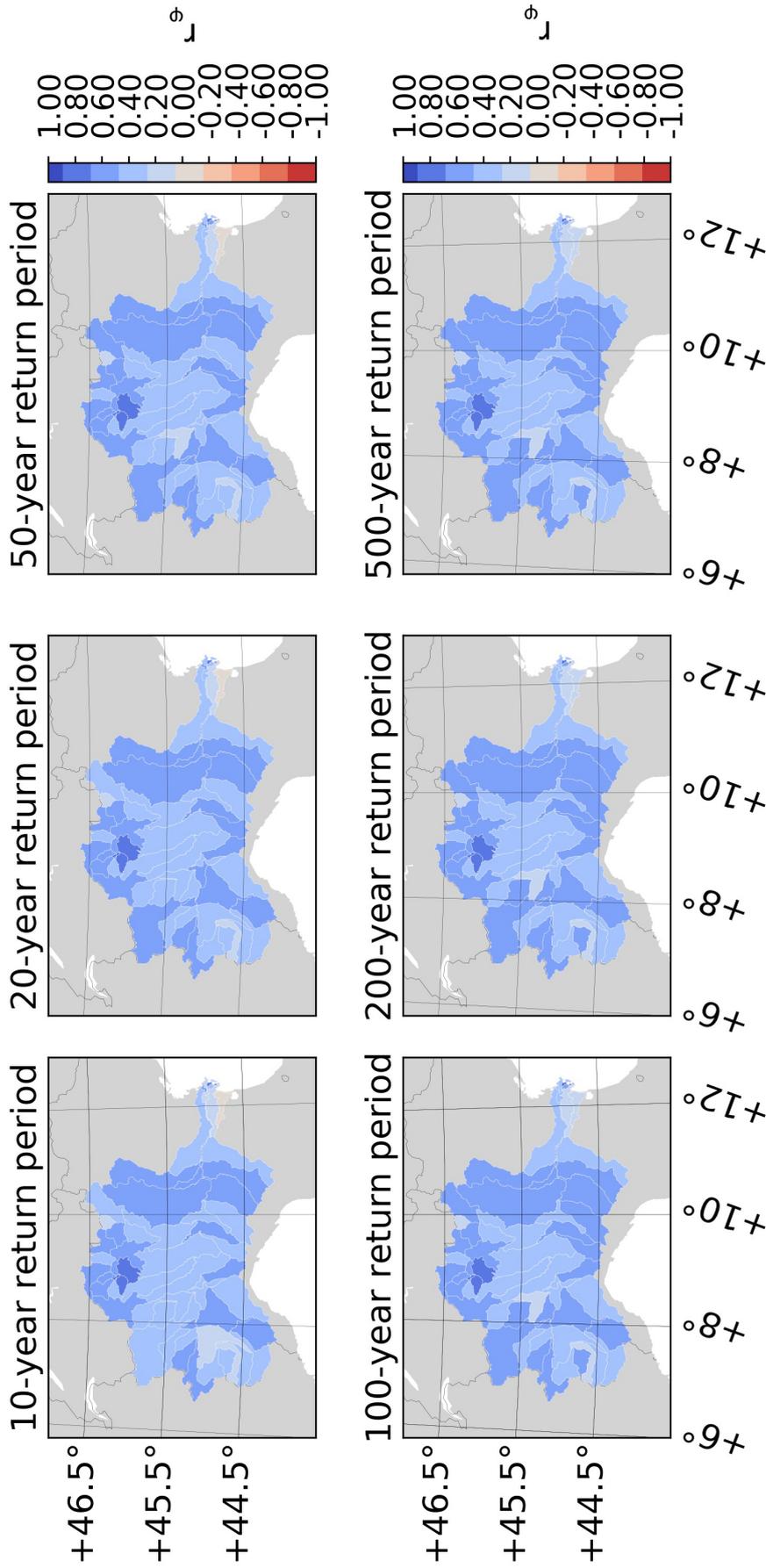
**Figure 5.10:** Geographical distribution of the True Skill Score (TSS) for six return periods, within the classification area set for each major river basin in Europe.



**Figure 5.11:** Geographical distribution of the True Skill Score (TSS) for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy.



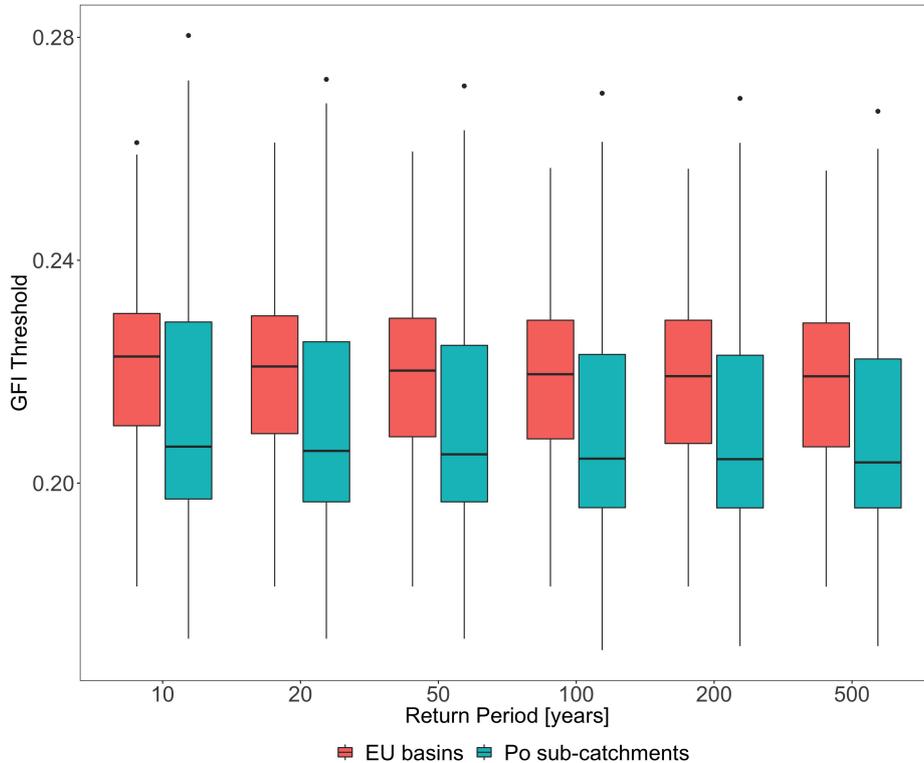
**Figure 5.12:** Geographical distribution of the modified Pearson's correlation ( $r_\phi$ ) for six return periods, within the classification area set for each major river basin in Europe.



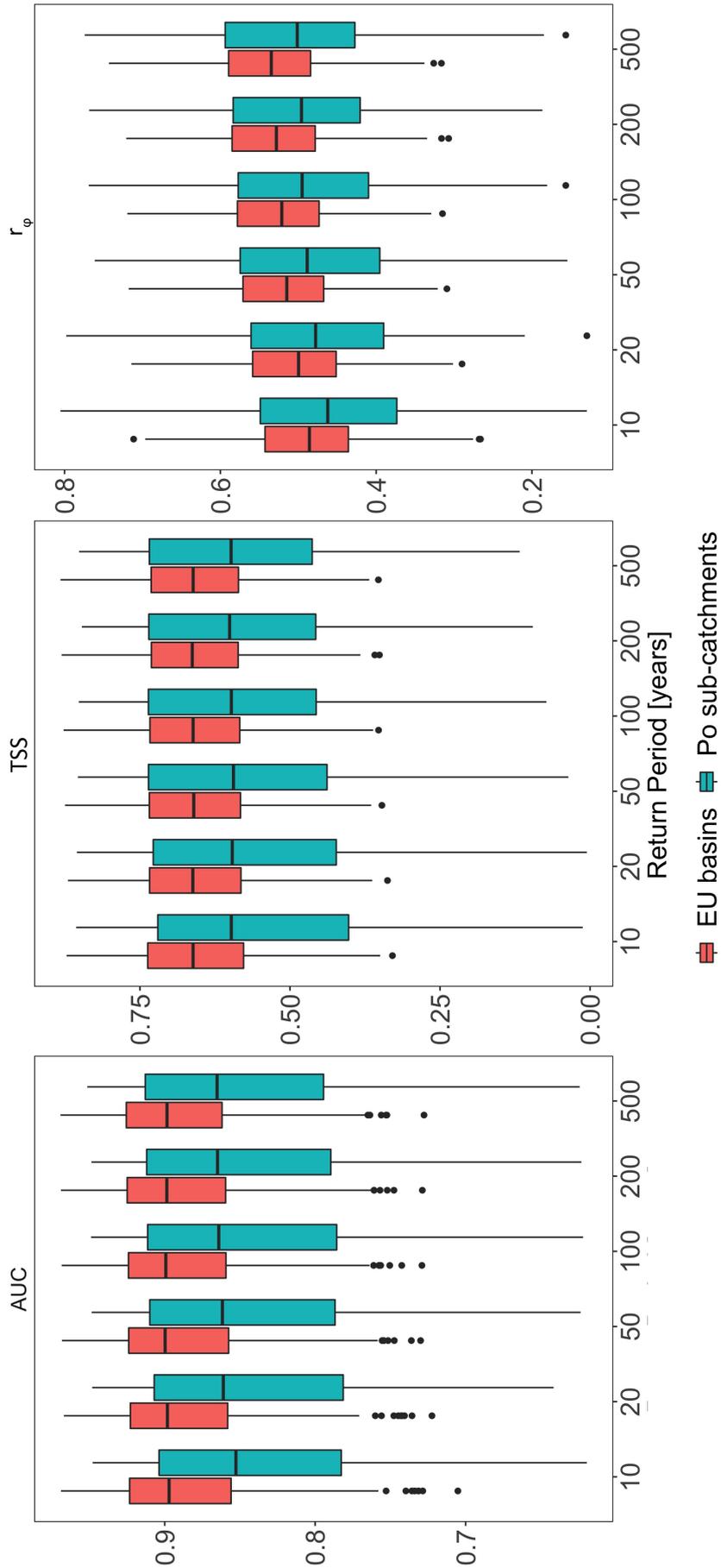
**Figure 5.13:** Geographical distribution of the modified Pearson's correlation ( $r_\phi$ ) for six return periods, within the classification area set for each sub-catchment of the river Po, in Italy.

river basins is observed and remains fairly unchanged across return periods. The AUC at the river basin scale (pan-European level) presents a number of outliers outside the lower bound but higher minima than in sub-catchments of the river Po in Italy. At the sub-catchment scale, there is an increase in variability of the AUC and a slight loss of skill. In general, the AUC mean and median tend to increase with return period at any of the given scales, with average values always above 0.8. The  $TSS$  presents low variability at the river basin scale, while at the sub-catchment scale its variability increases significantly. The minima  $TSS$  are also significantly lower at the sub-catchment scale. In particular, the mean  $TSS$  value is above 0.64 for all considered return periods at the river basin scale, and above 0.55 at the sub-catchment scale. The  $r_\phi$  shows low variability at the river basin scale, taking always positive values with mean above 0.48 at the river basin scale and above 0.46 at the sub-catchment scale.

In Fig. 5.14, the optimal  $TH$  of the flood descriptor GFI present low variability (higher at the sub-catchment scale), with different behaviour in terms of the mean, minima and maxima. The mean GFI  $TH$ , which tend to fall between 0.20 and 0.23, decrease slightly with return period. This is expected, since the higher the return period, the larger the flood-prone areas and the further the discerning edge moves away from the stream network. Few outliers can be observed outside the upper bound in any of the given scales.



**Figure 5.15:** Data distribution of the optimal Geomorphic Flood Index (GFI) threshold for six return periods.



**Figure 5.14:** Data distribution for the area under the receiver operating characteristic (AUC), True Skill Score (TSS) and the modified Pearson's correlation coefficient ( $r_\phi$ ) for six return periods.

In sum, the resulting homogeneous spatial distribution of the optimal GFI  $TH$  that decreases with return period, demonstrates that there is little inter-basin variation and that flood-prone areas actually grow with increasing return periods, as expected. In fact, performances indicate a strong capacity of the GFI to characterise flood-prone areas across return periods, generally achieving high AUC values. It is also observed that high AUC values did not always correspond to high  $r_\phi$  values, as the correlation coefficient takes into account the contingency matrix, because of the additional areas mapped by the hydrogeomorphic web application. Correlation values are nevertheless always positive and relatively high, with few exceptions.

Using the European dataset, small changes in performances and in the objective function  $TSS$  were observed. However, moving on to the sub-catchment scale, higher heterogeneity was found. This may have been due to the fact that over larger river basins different morphological features are smoothed out by the classification and the overall outcomes result similar, while within-basin morphological differences may influence the the method. Moreover, the method becomes more sensitive to the change of return period at the sub-catchment scale. Results suggest that the method should be applied at the basin scale, finding different  $TH$  values within each specific sub-catchment.

These findings open the possibility to explore a detailed procedure for the regionalisation of classification outcomes at a sub-catchment scale, which has the additional advantage of facilitating the parallelisation of computations by distributing tasks performed within sub-catchments among a number of CPUs.

## 5.2 Regionalisation of classification outcomes

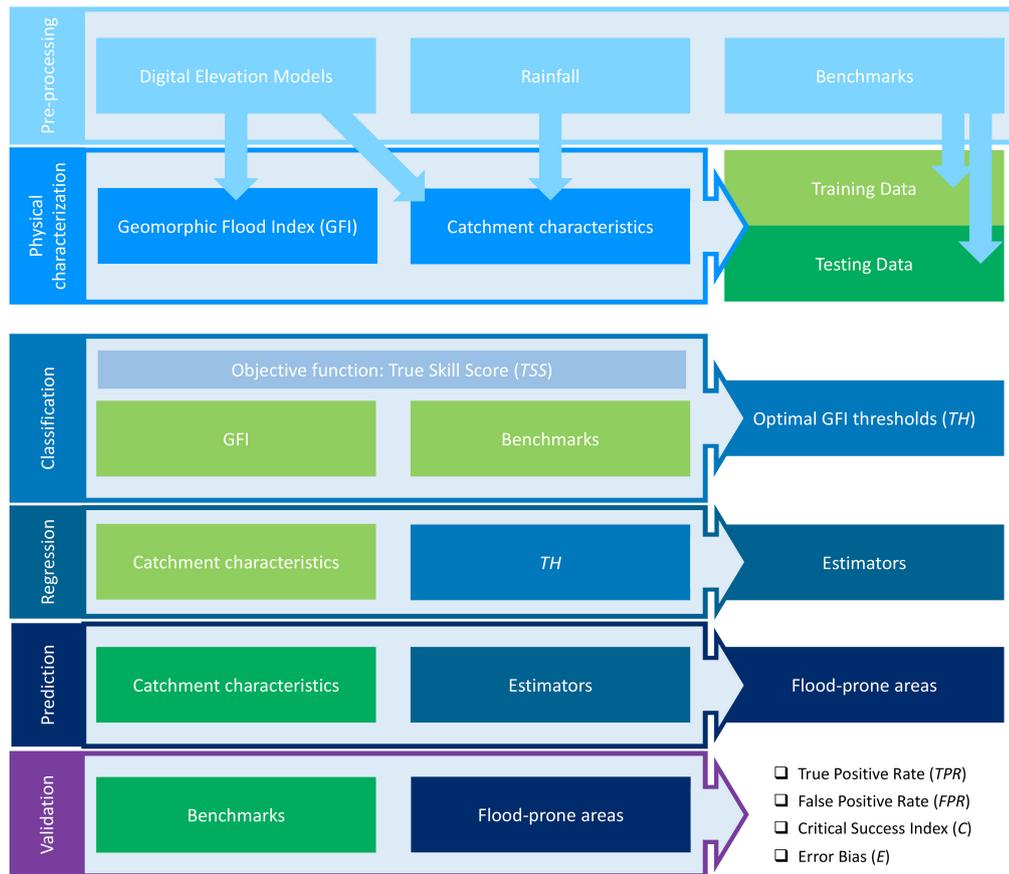
Recently, Jafarzadegan and Merwade (2017) experimented with regression models, obtained by stepwise analysis of climate and catchment characteristics, to delineate the 100-year floodplains in North Carolina, US. The delineation was performed based on a range of thresholds of the HAND, used as the flood descriptor. The authors used the United States Federal Emergency Management Agency's Flood Insurance Rate Maps as a benchmark for validating the results, but pointed out their uncertain and subjective nature. This approach was later extended by Jafarzadegan et al. (2018) and Jafarzadegan and Merwade (2019) to include a probabilistic description of the 100-year floodplains.

Similarly, the data-driven approach employed here complements traditional flood modelling, providing a cost-effective alternative that can exploit big, high-resolution datasets, without limiting the scale of application nor compromising computational speed. Being mostly DEM-based, the approach also relaxes the problem of data-scarcity often found,

enabling the delineation of flood-prone areas in any given catchment (i.e., regionalisation), based solely on statistical relations between flood-prone areas and catchment characteristics. Two major drawbacks found in previous applications are also eliminated by developing such relations: 1) the absolute dependence on a benchmark (Degiorgis et al., 2012); and, 2) the assumption of transferability of the optimal threshold (Samela et al., 2017). The questions posed are: 1) Can functional relationships be established between the GFI and catchment characteristics? 2) If so, can these relationships be used for low-complexity predictive modelling of flood-prone areas and related envelope flood extents?

In this study, estimators of flood-prone areas (linear stepwise regression and random forest regression) were developed for regionalisation based on classification outcomes of a flood index using high-resolution benchmarks and on geomorphic and climatic-hydrologic catchment characteristics. Elementary catchments of four river basins in Europe served as training dataset, while those of another river basin in Europe served as testing dataset. Two return periods were considered in the analysis, the 10- and 10,000-year.

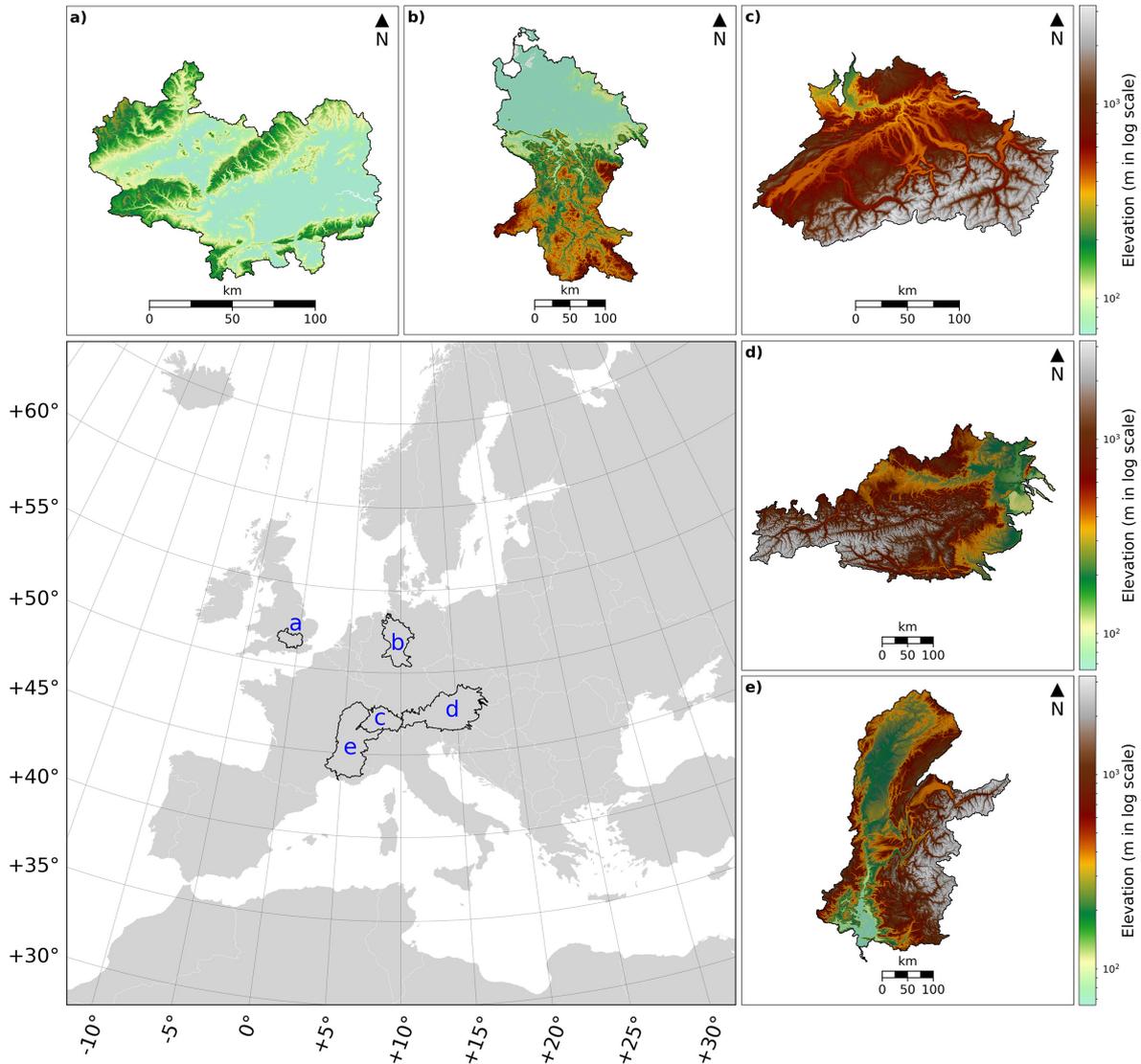
In particular, regression models are established between the derived optimal GFI thresholds – the target variable – and selected geomorphic and climatic-hydrologic catchment characteristics – the explanatory variables. Flood-prone areas in elementary catchments of a distinct major river basin in Europe are delineated by using unique  $TH$  predicted by the established relationships. Each resulting delineation is compared to the benchmark to assess the ability of the method to predict the extent of the envelope of major floods. In Fig. 5.16, the general methodological workflow is presented.



**Figure 5.16:** Workflow for developing predictive models of envelope flood extents using geomorphic and climatic-hydrologic catchment characteristics.

### 5.2.1 Study area and data

Five major river basins in Europe were selected for this study, mostly for their record of major floods and importance in Europe, their locations can be visualised in Fig. 5.17.



**Figure 5.17:** The five study river basins in Europe, with drainage divide highlighted in black in the lower left map. The training of the regression models is performed using a) the Thames river basin in the UK, b) the Weser in Germany. c) the upper Rhine in Switzerland, and d) the upper Danube in Austria. Testing is performed using e) the Rhône in France.

The River Thames in the UK constitutes the longest one in southern England (ca. 350 *km* length). It drains an area (ca. 13,478 *km*<sup>2</sup>) of relatively flat terrain (mean elevation of ca. 100 *m a.s.l.*) to the North Sea. The Thames river basin has a mean annual precipitation (*MAP*) ranging from ca. 610 to 778 *mm yr*<sup>-1</sup> that results in a mean annual runoff (*MAR*) ranging from 100 to 295 *mm yr*<sup>-1</sup>. The Thames is prone to major flooding; the 2013/14 winter floods that the valley sustained are an example of this (Huntingford et al., 2014; Fenn et al., 2016).

The River Weser in Germany has an overall length of ca. 452 *km*. It drains an area of ca. 43,857 *km*<sup>2</sup> with relatively flat terrain (mean elevation of ca. 200 *m a.s.l.*) to the North Sea. The Weser river basin has a *MAP* ranging from ca. 575 to 1,195 *mm yr*<sup>-1</sup>

that results in a *MAR* ranging from 190 to 930  $mm\ yr^{-1}$ . In 2013, the Weser river basin was affected by high flood levels with peak discharges above 50-year return period (Schröter et al., 2015).

The River Rhine (ca. 1,230  $km$  total length) has its source in the Swiss Alps and flows through several major cities in Switzerland, France, Germany and the Netherlands, where it drains to the North Sea. The upper Rhine river basin (drainage area of ca. 32,114  $km^2$ ), the portion of the Rhine river basin considered in this study, has a relatively mountainous terrain (mean elevation of ca. 1,065  $m\ a.s.l.$ ), with *MAP* ranging from ca. 825 to 1,715  $mm\ yr^{-1}$ , and resulting *MAR* ranging from 330 to 2,250  $mm\ yr^{-1}$ . The upper Rhine river basin is prone to major flooding; for example, in 2007, one person lost its life, at least 100 were affected and the country withstood a total estimated damage of more than 312 million EUR (CRED EM-DAT).

Originating in Germany and flowing through major cities (e.g., Vienna, Austria) in 10 different countries before draining to the Black Sea, the River Danube is the second longest river in Europe (ca. 2,850  $km$  length). The upper Danube river basin (drainage area of ca. 97,000  $km^2$ ), section considered in this study, is characterised by a relatively mountainous terrain (mean elevation of ca. 890  $m\ a.s.l.$ ), *MAP* ranging from ca. 460 to 1,785  $mm\ yr^{-1}$ , resulting in *MAR* ranging from ca. 23 to 1,282  $mm\ yr^{-1}$ . The upper Danube river basin is prone to major flooding; for example, in 2013, four persons lost their lives, at least 200 were affected and the country withstood a total estimated damage of more than 893 million EUR (CRED EM-DAT).

The River Rhône in France originates in the Swiss Alps and runs through south-eastern France, where it finally drains to the Mediterranean Sea. The Rhône river basin, with an area of ca. 96,475  $km^2$  has a mean elevation of ca. 785  $m\ a.s.l.$  It is characterised by a *MAP* ranging from ca. 561 to 1,890  $mm\ yr^{-1}$ , resulting in *MAR* ranging from ca. 119 to 1,551  $mm\ yr^{-1}$ . The winter floods of 2003 marked the largest flood in the Rhône river basin since 1856. Consequences arising from this event were severe, with the country withstanding a total estimated damage of ca. 1.130 billion EUR (Arnaud-Fassetta, 2013).

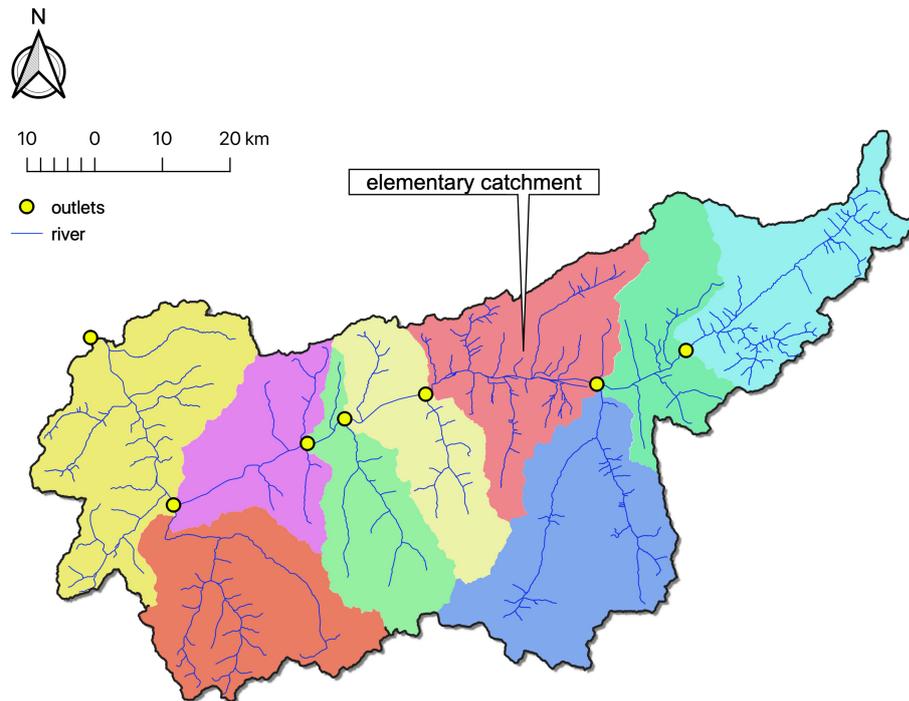
Mean elevation values reported here were estimated from the EEA EU-DEM, *MAP* values from the ECA&D E-OBS and *MAR* values from the UNH/GRDC runoff dataset.

The flood descriptor layer (GFI) and catchment characteristics (Table 5.2) were computed for all selected river basins using a proprietary DEM dataset, the RMS-DEM, at ca. 50  $m$  spatial resolution. The RMS-DEM is well suited for flood inundation modelling, as it does not contain artefacts such as trees, buildings or bridges that can adversely affect the accuracy of the simulations.

The (1) Thames river basin in the UK, the (2) Weser river basin in Germany, the (3) upper Rhine river basin in Switzerland, and the (4) upper Danube river basin in Austria are used for training the regression models, where the unique TH values resulting from a threshold binary classification are used as target variable. The first two river basins are representative of flatter regions and the last two of mountainous regions. The (5) Rhône river basin in France is instead used for testing the regression models, where a unique TH is predicted for each of its elementary catchments, using the statistical relationships established by the regression models.

The benchmark used in the threshold binary classification to find the TH values, and also in the evaluation of the final predictions, is obtained from high-resolution flood hazard maps for Europe, developed by RMS and currently used by global insurance and reinsurance companies. The RMS flood maps were created for several return periods at ca. 50 *m* resolution. They are based on a cascade of sequential modelling components. Rainfall runoff processes are modelled with a semi-distributed, TOPMODEL-based approach (Beven and Kirkby, 1979). Flows are routed through the river network using the Muskingum-Cunge 1D wave propagation method (Cunge, 1969; Georgakakos et al., 1990). Inundation depths and extents used in this study are derived by applying rating curves to river flows in every river segment of 500 *m*, calculating the associated river depth and filling the river cross-section extracted from the DEM for each segment. The maximum flood depths over the floodplain, after propagating the flood wave through the main river channel, represent the flood hazard map for an event. The benchmark is used in the form of a binary mask (raster cells marked as flood-prone or flood-free) that is obtained through image segmentation with a cut-off depth, set to 0.01 *m*.

The data-driven approaches described in previous chapters are performed within elementary catchments of major river basins (i.e., hydrological units defined as the “portion of basin directly drained by a river stretch, between two confluences, or from the headwater to the first confluence” in Castellarin et al. (2018); Fig. 5.18).



**Figure 5.18:** Subdivision into elementary catchments overlaid with a river network representation (the EEA EU-Hydro photo-interpreted river network) and outlets, defined as the intersection of each elementary catchment drainage divide with the river network.

The delineation of elementary catchments accounts for a hierarchical structure that reflects the topology of the river network (Verdin and Verdin, 1999). The main reason for choosing this scale of analysis is the division of each river basin in topographic areas that may contribute significantly to discharge and play a central role in the management of water resources. It also serves the purpose of making computations more manageable through concurrent programming (see Appendix D). Catchments are delineated following the constraint that catchment areas should be less than ca.  $1200 \text{ km}^2$ . The catchment delineation resulted in 83 elementary catchments within the Thames river basin, 171 within the Weser river basin, 109 within the upper Rhine and 286 within the upper Danube. Their merging into a single dataset resulted in a total of 651 elementary catchments, of which, after filtering out issues such as poor classification results, 454 were used for training the regression models. Catchment delineation within the Rhône river basin resulted in 277 elementary catchments that were used for testing the regression models.

Catchment characteristics used in this thesis (Table 5.2) are strictly geomorphic and climatic-hydrologic, as defined by Horton (1932), where soil, geology and vegetation are not taken into account. For consistency, the same high-resolution DEM of the proprietary flood maps is used to extract geomorphic catchment characteristics.

**Table 5.2:** Summary of geomorphic and climatic-hydrologic catchment characteristics.

Characteristics	Description
$A$	Area of elementary catchment ( $km^2$ )
$F$	Flow accumulation at elementary catchment outlet ( - )
$\Delta z$	Relief of elementary catchment ( $m$ )
$S$	Relief-area ratio of elementary catchment ( $m km^{-2}$ )
$L_{ch}$	Total river channel length in elementary catchment ( $km$ )
$\Delta z_{ch}$	Relief of the river channel in elementary catchment ( $m$ )
$S_{ch}$	Relief ratio of the river channel in elementary catchment ( $m km^{-1}$ )
$P_{10}$	10 consecutive days precipitation at elementary catchment scale associated with a 10-year return period ( $mm yr^{-1}$ )
$P_{10k}$	10 consecutive days precipitation at elementary catchment scale, associated with a 10,000-year return period ( $mm yr^{-1}$ )
$MAP$	Mean annual precipitation at elementary catchment ( $mm yr^{-1}$ )
$q_{10}$	Unit discharge at elementary catchment outlet for the P10 statistic ( $m^3 s^{-1} km^{-2}$ )
$q_{10k}$	Unit discharge at elementary catchment outlet for the P10k statistic ( $m^3 s^{-1} km^{-2}$ )
$q_{MAP}$	Unit discharge at elementary catchment outlet for the MAP statistic ( $m^3 s^{-1} km^{-2}$ )

The following single geomorphic catchment characteristics were considered: area of elementary catchment ( $A$ ); flow accumulation at the elementary catchment outlet ( $F$ ), defined as the cumulative sum of raster cells upstream of the outlet; relief of elementary catchment ( $\Delta z$ ), defined as:

$$\Delta z = z_{max} - z_{min} \quad (5.1)$$

the total river channel length in the elementary catchment ( $L_{ch}$ ); and mean river channel fall, or relief, in the elementary catchment ( $\Delta z_{ch}$ ). Composite geomorphic catchment characteristics used in this thesis are representative of the mean declivity. Within each elementary catchment, the relief-area ratio is defined as:

$$S = \frac{\Delta z}{A} \quad (5.2)$$

while the relief ratio of the river channel in the elementary catchment is defined as:

$$S_{ch} = \frac{\Delta z_{ch}}{L_{ch}} \quad (5.3)$$

From a hydrological perspective, declivity has an important relation to surface runoff, to the concentration of rainfall in river channels and to flood magnitude (Horton, 1932). On the other hand, mean river channel declivity gives an estimate of channel storage and time length required by a flood wave to traverse the channel (Horton, 1932), while it also relates to the linear head loss found in the Manning's equation for uniform flow (Manning, 1891). High mean declivities equate to water entering the channel quicker, and thus to higher flooding likelihood. By contrast, gentle sloping channels are slower to route the incoming runoff and have a lower flooding likelihood.

Multi-day rainfall events are an important cause of flooding and therefore critical for the design of structures (Fowler and Kilsby, 2003). In this study, the annual highest 10 consecutive day precipitation is reported, which has a likelihood of occurring or of being exceeded every 10 and 10,000 years on average ( $P_{10}$  and  $P_{10k}$ ), respectively. These statistics are calculated based on the ECA&D E-OBS 0.1 degree regular gridded precipitation dataset (Cornes et al., 2018). Principal component analysis was applied to precipitation anomalies in the dataset for the 1950-2010 period, in order to identify dominant rainfall patterns across Europe. Stochastic precipitation fields are obtained for 50,000 years as linear combinations of empirical orthogonal functions and principal components (Zanardo et al., 2019) (see Appendix B). To complement these statistics, the MAP) calculated by averaging the annual totals obtained from the ECA&D E-OBS dataset is also reported.

Proxies for long-term average runoff are obtained by accumulating precipitation statistics downstream using the hierarchy of connected elementary catchments, which should

reflect the fractal nature of the river network. The general water balance equation for each elementary catchment is given by:

$$P + \Delta Q - E - \Delta W_S = 0 \quad (5.4)$$

with  $P$  the precipitation received at each elementary catchment,  $\Delta Q$  the change in specific runoff,  $E$  the evapotranspiration and  $\Delta W_S$  the change in water storage. If subsurface water flow and evapotranspiration losses are neglected as a simplification for severe rain storms and very humid conditions – meaning that overland flow suffers either from saturation excess or infiltration excess and that evapotranspiration losses are much lower than water entering the elementary catchment – the direct conversion of precipitation into runoff may be assumed dominant at each elementary catchment, water yield tends to 1, and the following equation holds:

$$Q_{out} = P + Q_{in} \quad (5.5)$$

with  $Q_{out}$  the runoff at each elementary catchment outlet and  $Q_{in}$  the runoff from upstream elementary catchments. Equation 5.5 can be further expanded, to cater for the lumped cascading estimation of direct runoff at the elementary catchment:

$$Q_{out} = P_0 \times A_0 + \sum_{i=1}^n (P_i \times A_i) \quad (5.6)$$

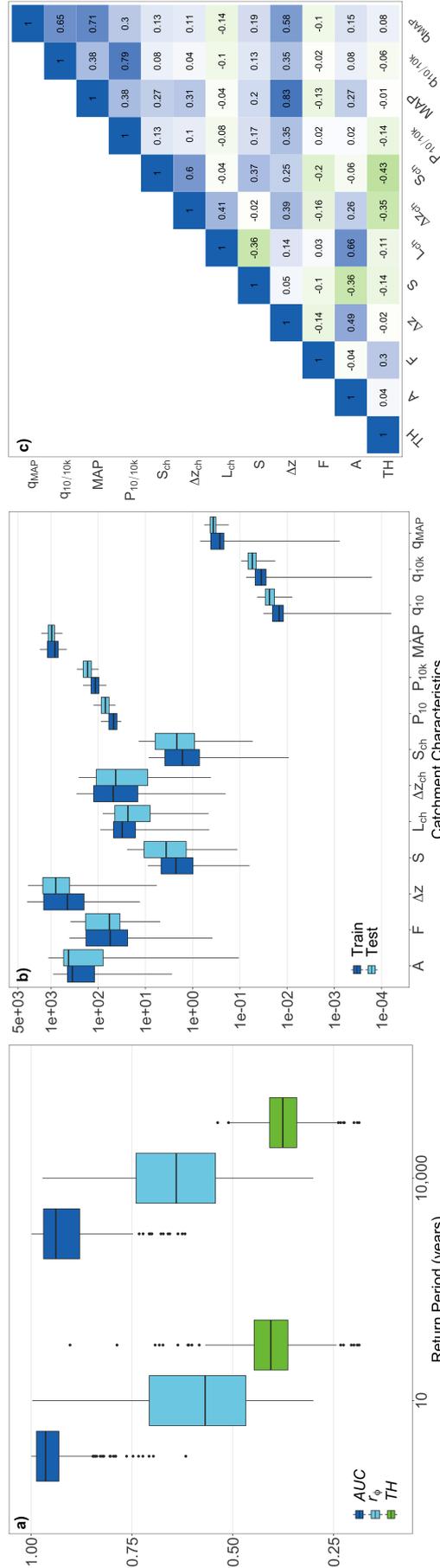
with  $P_0$  corresponding to a unique long-term average precipitation statistic (i.e.,  $P_{10}$ ,  $P_{10k}$  or  $MAP$ ) associated with the elementary catchment under analysis with area  $A_0$ , while  $P_i$  is the unique precipitation statistic associated with the  $n$ -th upstream elementary catchment with area  $A_i$ . Results are reported as unit discharge at each elementary catchment outlet:

$$1_{out} = \frac{Q_{out}}{F \times cellsize} \quad (5.7)$$

with cell size (in  $m^2$ ) equal to the product of pixel length by pixel width, specific to each DEM.

### 5.2.2 Results and discussion

The classification of the GFI layer to obtain the TH was performed using each elementary catchment that constitutes the training set, composed of four major river basins in Europe. In Fig. 5.19, the data used in the development of estimators and prediction of flood-prone areas for the 10- and 10,000-year return periods is presented.



**Figure 5.19:** Data used in the development of estimators and prediction of flood-prone areas for the 10- and 10,000-year return periods, please refer to Table 5.2 for a complete description and units of variables. **a)** Area Under the relative operating characteristic Curve AUC the modified Pearson's correlation coefficient for discrete dichotomous problems  $r_\phi$  and the optimal Geomorphic Flood Index (GFI) thresholds TH obtained from the classification of training catchments (Thames, Weser, upper Rhine and upper Danube river basins). **b)** Geomorphic and climatic-hydrologic catchment characteristics for the combined training and test catchments (Rhône river basin). **c)** Correlation matrix between TH and catchment characteristics.

Average  $AUC$  values of ca. 95% and 91% are found for the 10 and 10,000-year return periods, respectively, which indicates a very high discerning capability of the GFI classifier. These  $AUC$  values translate to a high probability of correctly classifying a raster cell as flood-prone or flood-free. A noticeable number of  $AUC$  outliers in the lower bound of the distribution of outcomes, particularly below the no-skill line, may affect the development of the statistical relationships for the prediction of flood-prone areas. Therefore, values of  $AUC$  below 0.5 were filtered out from the training dataset.

Average  $r_\phi$  values between ca. 60% and ca. 64% are found for the 10 and 10,000-year return periods, respectively, which indicates a strong positive degree of association (i.e., between 1 and 0.5) of the best possible representation of flood-prone areas. This is a necessary pre-condition before any attempt to relate the threshold values and the catchment characteristics. Values of  $r_\phi$  below 0.3 (lower limit of a moderate degree of association) were filtered out from the training dataset.

Values of  $TH$  for the elementary catchments of the training set are found to range between 0.18 and 1, with a mean value of ca. 0.44 and 0.39 for the 10- and 10,000-year return periods, respectively. As expected, there is a tendency towards a value decrease with increasing return period.

For most catchment characteristics more than 50% of test data is contained within the training set interquartile range. Exceptions to this can be found for the precipitation statistics,  $P_{10}$  and  $P_{10k}$ , with median slightly above the test set interquartile range, and for the corresponding unit discharge estimates,  $q_{10}$  and  $q_{10k}$ . The sample variability of the test set is larger than that of the training set for the  $A$ , the  $S$ , the  $L_{ch}$ , and the  $S_{ch}$ . The explanatory variable  $A$  is the only that is noticeably positively skewed, while the  $MAP$  and corresponding  $q_{MAP}$ , are negatively skewed. These differences are expected to negatively impact the final prediction of flood-prone areas as the training set does not represent the test set in the most exhaustive manner.

The correlation matrix in Fig. 5.19c provides an evaluation of the magnitude of association and direction of the linear relationship between explanatory variables and the dependent variable. Correlations between the  $TH$  and the  $F$ , indicate a moderate positive linear relationship. Moreover, a moderate negative linear relationship is revealed between  $TH$ , the  $\delta z_{ch}$  and  $S_{ch}$ . The remaining catchment characteristics reveal weak linear relationships to the  $TH$ .

Furthermore, correlations that exist between different catchment characteristics may be indicative of multicollinearity. Disregarding the correlations between composite explanatory variables and their constituting parts, a moderate to strong positive correlation between precipitation statistics, discharge estimations and  $\delta z$  is noticeable. Also,  $A$  shows

a strong positive correlation with  $\delta z$  and  $\delta\delta z$ . Collinearity between catchment characteristics is an undesirable condition that can negatively impact the quality of the statistical relationships for the prediction of flood-prone areas and need to be addressed before any further step is taken.

Two types of regression models were built from the classification outcomes  $TH$  and catchment characteristics of the training set, namely the stepwise regression and the random forest regression models.

Several data transformations were tried for building different models (log-linear, linear-log and principal component analysis, see Appendix B, which were tested but did not produce any beneficial result). Log-log transformed variables were used as they substantially improved both models' statistical tests and performance metrics.

Several data splits were also tried for building different models. For example, one model for the 10- and another for the 10,000-year return period, were tried but did not yield significantly different results from the ones presented in Table 5.3. Furthermore, the inclusion of specific river basins was tested. Namely, two out of the four training river basins were held out at the time for testing the models built with the remainders. None of the six river basin combinations ( $4!/(2! \times (4-2)!) = 6$ ) significantly improved the overall performance and in most cases holding out specific river basins actually decreased it.

In the stepwise regression, multicollinearity tests point towards strong collinearity between the composite explanatory variables and their individual constituents. Additionally,  $VIF$  values above 2 are found between the combined  $q_{10}$  and  $q_{10k}$  and the  $q_{MAP}$ . Thus, the catchment characteristics  $A$ ,  $\delta z$ ,  $\delta\delta z$ ,  $\delta z_{ch}$  and  $q_{MAP}$  were considered for elimination given the results of the multicollinearity tests and taking into consideration the correlation results presented in Fig. 5.19.

The previous steps were followed by a stepwise selection of explanatory variables based on the  $AIC$ , which reflects the trade-offs between maximum likelihood and model simplicity. A very low  $AIC$  lead to the following equation:

$$TH = 0.1580 - S_{ch}^{0.0631} + F^{0.0345} + S^{0.0023} - P_{(10/10k)}^{0.0774} + MAP^{0.1267} + q_{(10/10k)}^{0.0156} \quad (5.8)$$

As can be seen from Table 5.3 the final optimised linear model is constituted by six of the 11 original explanatory variables and is characterised by a high F-statistic ( $> 3$ ) and very low p-value ( $< 0.01$ ), which indicates a high degree of significance of individual explanatory variables and of the model. From the 10-fold cross-validation procedure, results a  $R^2$  value of ca. 42%, indicating a moderate explanatory power of the model, and a  $RMSE$  of 0.0597.



In the random forest regression, optimisation was performed by an automatic search of the best possible combination of input parameters that lead to the highest possible decrease in  $RMSE$  obtain through cross-validation. This also ensures that overfitting is avoided. The final optimised random forest regression corresponds to a model with 644 trees, three explanatory variables randomly sampled at each tree node and a maximum depth of 25 nodes. As shown in Table 5.3, the random forest regression results have substantially improved the explained variance obtained by the stepwise regression model, from ca. 42% to ca. 64%  $R^2$ , and the  $RMSE$ , which decreased from 0.0598 to 0.0466.

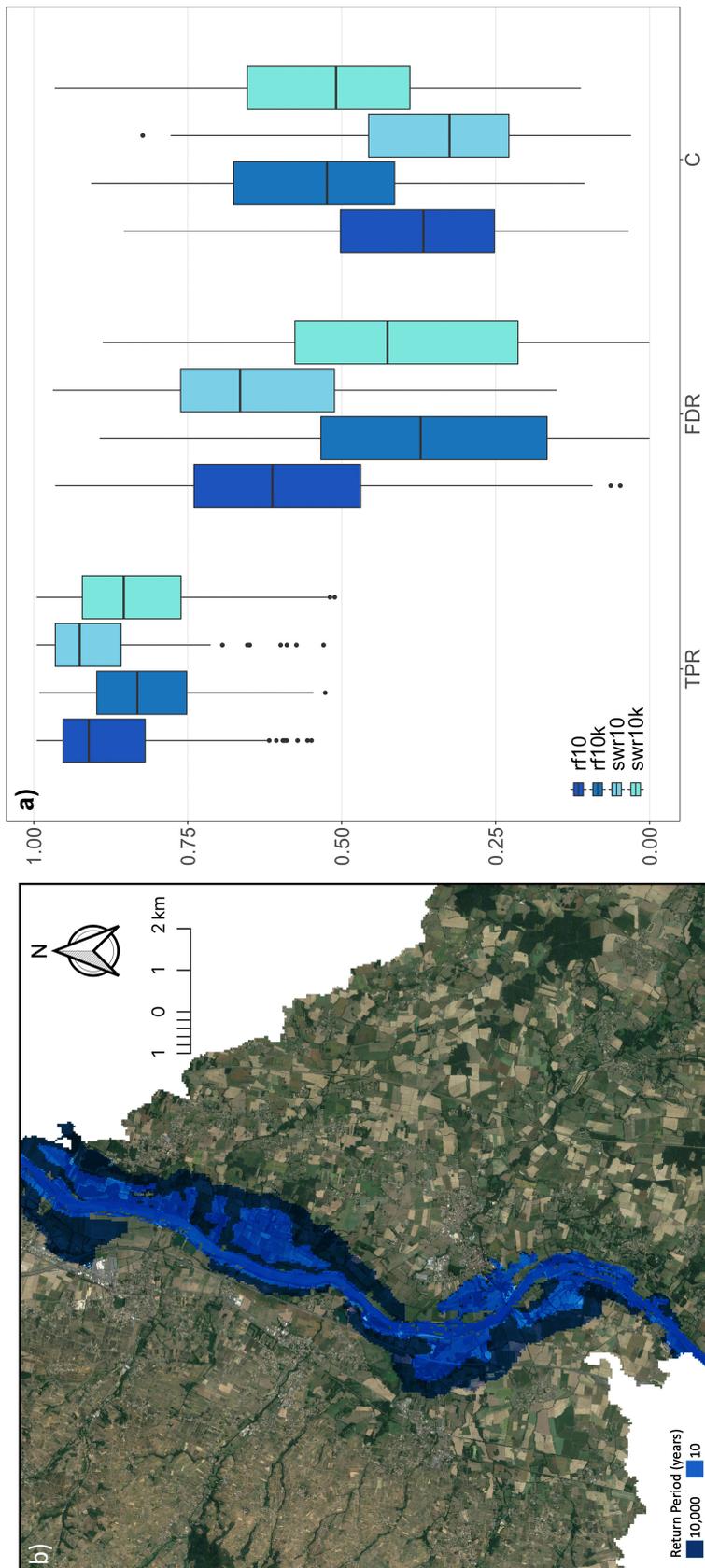
Variable relative importance shows that the  $S_{ch}$  is the explanatory variable with the highest relative importance in both models, ca. 48% and 19%, followed by the  $F$  with ca. 36% and 19%. In the random forest regression model  $\delta z_{ch}$  is also found to have a fairly high variable relative importance with ca. 14%. The remaining catchment characteristics are ranked as relatively less important, or not included at all (stepwise regression). However, it should be noted for the case of the random forest model that as one explanatory variable is randomly selected at a tree node, the importance of other variables is substantially reduced, particularly if there is collinearity. In light of this, variable importance should be interpreted with caution, as explained in (Seibert et al., 2017).

Using the models previously presented, envelope flood extents were predicted based on the physical characteristics of elementary catchments of the Rhône river basin.

Catchment characteristics matching those used for training of the regression models were obtained for the 10- and 10,000-year return periods and used as input. A unique  $TH$  was predicted per elementary catchment, return period and model. The predicted  $TH$  values were used to segment the original GFI raster layer of each corresponding elementary catchment of the Rhône river basin and to delineate the flood-prone areas (see Fig. 5.20a for an example).

By comparing each raster cell of the binary mask of predicted flood-prone/flood-free areas with each corresponding raster cell of the benchmark RMS flood maps, it was possible to obtain a contingency matrix for each model and for each return period considered, from which the performance metrics described in Chapter 3 were computed.

In Fig. 5.20 and Table 5.4, it can be observed that the  $TPR$  is high for the great majority of elementary catchments (average above 80%), similar between models and slightly higher for the 10-year return period. At the same time, the  $FDR$  is high for the 10-year return period (average ca. 63%), moderate for the 10,000-year return period (average ca. 39%), and slightly higher for the stepwise regression. The  $C$  is moderate for the 10-year return period (average ca. 34%), high for the 10,000-year return period



**Figure 5.20:** a) True positive rate (TPR), false discovery rate (FDR) and critical success (C) for the elementary catchments of the Rhône river basin, for the different regression models (rf – stepwise regression) and for the 10-year and 10,000-year return periods. b) Overlaid samples of flood-prone areas and corresponding envelope flood extents, predicted with the rf model for two return periods.

(average ca. 52%), and slightly higher for the random forest. It is also shown by the  $E$  that more than ca. 85% of the flood-prone areas obtained for the elementary catchments of the Rhône river basin suffer from overestimation ( $E > 1$ ). In general, the  $TPR$  decreased with increasing return period, but this seems to be well compensated by a significantly lower number of false alarms, as a higher  $C$  is observed for the 10,000-year return period and the  $E$  decreases.

**Table 5.4:** Performance of the regression models expressed as true positive rate (TPR), false discovery rate (FDR), critical success (C) and error bias (E), for the 10- and 10,000-year return period flood-prone areas in the Rhône river basin using the RMS flood maps as benchmark.

<b>Model run</b>	<b>TPR</b>	<b>FDR</b>	<b>C</b>	<b>E</b>
Stepwise regression 10-year return period	0.8749	0.6534	0.3270	161
Stepwise regression 10,000-year return period	0.8205	0.4026	0.5052	37
Random forest 10-year return period	0.8618	0.6129	0.3597	159
Random forest 10,000-year return period	0.8047	0.3701	0.5268	36

# Conclusions

Integrating big open data with data-driven methods and emerging information and communication technologies in flood risk management may well represent a forward step in supporting robust, evidence-based decision-making. Driven by this idea, the progress in simplified hydrogeomorphic methods for the mapping of flood-prone areas as been reviewed. A methodology for the classification of flood-prone areas based on a particular flood descriptor, the GFI has been adopted and tested using big open datasets for Europe. Additionally, a user-friendly web application, well suited for regional hydrogeomorphic delineation of flood-prone areas, has been implemented, described in detail and used to classify, delineate and downscale six pan-European hydrogeomorphic flood hazard maps at 25-m pixel resolution (one for each return period: 10, 20, 50, 100, 200 and 500 years) and to evaluate their quality within major European basins, relative to the flood hazard maps for Europe (Alfieri et al., 2014; Dottori et al., 2016).

An average efficiency, measured in terms of ROC analysis and AUC of 88.59% was obtained for the flood descriptor. In the analysis focusing on sub-catchments of the river Po in Italy, an average AUC of 84.23% was obtained, in line with results presented in other studies (Manfreda et al., 2015; Samela et al., 2017). The selected hydrogeomorphic method was found to be valid in replicating, downscaling and extrapolating the flood hazard maps for Europe with different return periods. The hydrogeomorphic flood hazard mapping web application presented in this study provides fast, high-resolution delineations of flood-prone areas over large scales at higher spatial resolution than the benchmarks used. The classification method was found to be affected by:

1. pixel size and vertical accuracy of the DEM;
2. processing decisions (e.g., de-noising, smoothing and hydrological conditioning);
3. channel initiation and stream network definition using a convergent flow model;
4. fixed coefficients in the hydraulic scaling relation.

Other factors include differences in the DEMs used to compute the GFI and the reference flood hazard maps. Furthermore, there is a strong possibility that the size of flat areas influences negatively the classification of hydrogeomorphic flood-prone areas.

A GFI computed with inadequate horizontal resolution and vertical accuracy will likely underperform in flat areas because of its strong dependency on the elevation difference, but also on the stream network definition, which is known to be problematic in these areas (e.g., Pan et al., 2012). One could also argue that sub-grid scale features, for example embankments, affect performance, particularly in highly anthropized river basins (e.g., river Po, Italy), but we note that such obstacles are not explicitly taken into consideration neither in the reference flood hazard maps for Europe nor in the GFI computation.

By establishing functional relationships in the form of linear and non-linear regression models between specific isolines of a flood index  $TH$  and geomorphic and climatic-hydrologic characteristics of elementary catchments, the prediction of extents of the envelope of major floods in diverse river basins, gauged or ungauged, and for diverse return period was made possible. This advancement extends the classification approach by relaxing its complete dependence on benchmark flood maps and by providing a physical basis for the transferability of the  $TH$  between catchments or river basins, also making the extrapolation and downscaling goals more realistic.

The classification stage of the methodological workflow in the regionalisation study showed that, overall and in spite of some outliers, the GFI has an overall high discerning capability of flood-prone areas, as shown by the average AUC value above 91% for any of the return periods. At the same time, average value of  $r_\phi$  above 60% for any of the return periods indicates a strong positive degree of association between the GFI delineated flood-prone areas and the benchmark flood maps. These values are significantly higher than the ones reported before, most likely because of the scale of analysis used (elementary catchments) and the quality of datasets (RMS-DEM and RMS flood maps).

The development of estimators of flood-prone areas has shown that in the stepwise regression the VIF and AIC selection of catchment characteristics has been valuable in obtaining statistically significant explanatory variables that improved the explained variance ( $R^2$ ) of the target and the fit of the initial model ( $RMSE$ ).

In comparison to the stepwise regression, the random forest regression proved to be a much more flexible and straightforward approach to setup. The final optimised random forest model could be obtained without any prior selection of catchment characteristics and still substantially increased the  $R^2$  and decreased the  $RMSE$ . Moreover, the improvement of the statistical tests by the random forest model seems to provide some evidence of non-linear behaviour between  $TH$  and catchment characteristics.

When it comes to the predictions of flood-prone areas and envelope flood extents in the elementary catchments of the Rhône river basin, the random forest regression model performed marginally better than the stepwise regression for any of the return periods

considered. Both the stepwise and the random forest regression outputted high *TPR* values, while at the same time moderate to high *FDR* values. This was reflected in the moderate to high critical success *C* obtained and in the *E* values always above 1, indicative of overprediction, especially at lower return periods. As this thesis tries to deal with envelope flood extents, overprediction was already expected. Overall, predicted flood-prone areas better match the benchmark at higher return periods and, particularly for the 10,000-year return period, it is interesting to note that the performance obtained is in line with some modelling results reported by (Wing et al., 2017). A number of additional findings are also worth noting:

1. Although the models proved to be reasonably robust, considering that the sample variability of the training data was limited in comparison to that of the testing, a training of the models with a broader range of values and more degrees of freedom could improve the generalisation properties and prediction capability.
2. The random forest model is not able to predict target values outside the range found in the training dataset and this can be particularly problematic for lower TH values. To account for this feature, different algorithms would need to be considered or modifications to the random forest would need to be implemented.
3. Explicitly including the case where flooding does not occur (e.g., accurate representation of the river at bankfull stage and the corresponding physical climatic-hydrologic characteristics that lead to it) in the models may benefit the analysis. However, such cases should be completely withheld from the performance analysis, as they might artificially influence the performance.
4. The use of a DEM to compute the GFI that was different from the DEM used in the modelling of the benchmark flood maps negatively influenced the results. Caution should be exercised in the selection of the DEM, but also in its processing (e.g., terrain analysis, river network and catchment delineation).
5. Besides what was mentioned above, it would be interesting to see how such methodology would work with benchmark flood extents obtained from remote sensing detection (e.g., Westerhoff et al., 2013; Schumann and Moller, 2015), as well as to take a step further and provide a way to estimate flood depth (e.g., Manfreda and Samela, 2019).

Results presented in this thesis reinforce the idea that data-driven approaches are valuable in the handling of big data and for implementation as online tools, while the web technology is a powerful way to bring researchers, developers and end users together, in order to implement innovative approaches and functionalities in the context of large-scale flood assessment and management.



# Appendix A

## Analysis of open-access digital elevation models

One of the most critical elements in the hydrological response of a river basin is its topography (e.g., Sørensen and Seibert, 2007; Zhang and Montgomery, 1994; Horritt and Bates, 2001). In fact, the great progress experienced in the modelling of both hydrological and hydraulic processes in the last decades cannot be dissociated from the advances in terrain information in the form of DEMs, or digital elevation datasets representing the Earth's surface. Obviously, this popularity might not have been attained had advances in remote sensing not been on a par with those in computational power and software, as well as with the release of other important datasets, land use and land cover, for example. Within the large variety of remote sensing techniques, a focus is here given to the two most disruptive technologies for generating DEMs: synthetic aperture radar (SAR) and light detection and ranging (LiDAR), see Croneborg et al. (2015) for more details.

The open access licensing has allowed DEMs to be distributed online, free of charge to the public. Access to global DEMs before 1996 was either restricted or inexistent and since then the number of publication have significantly increased. This reflects their use in a range of fields, amongst which the modelling of hydrological and hydraulic processes (e.g., Kumar et al., 2000). However, publicly released, freely-available datasets differ in characteristics such as spatial resolution, digital terrain processing decisions and vertical accuracy, which may introduce a range of errors in the modelling of hydrological and hydraulic processes. A number of authors have assessed these errors and conjectured new ways to move forward (e.g., Sanders, 2007; Jarihani et al., 2015; Archer et al., 2018).

Here, a brief investigation of ten DEM dataset (SRTM GL1 and GL3, HydroSHEDS, TINITALY, ASTER GDEM, EU DEM, VFP, ALOS AW3D30, MERIT and the TDX) in terms of the influence of three parameters, i.e., spatial resolution, hydrological reconditioning and vertical accuracy, on four relevant geomorphic terrain descriptors frequently used

to characterise hydrological or hydraulic processes (e.g., Tarboton et al., 1992; Beven and Kirkby, 1979; Montgomery and Dietrich, 1989; Westerhoff et al., 2013; Nobre et al., 2016; Manfreda et al., 2015; Clubb et al., 2017) is presented. Namely, the upslope contributing area, the local slope, the H (see Chapter 2) and the flow path distance to the nearest stream, D, respectively. This is done in terms of their cumulative frequency curves within the Tanaro river basin, in Italy. In addition to investigating the intrinsic characteristics of freely accessible DEMs and given the importance of river networks in the modelling of hydrological and hydraulic processes, DEM-based river networks are delineated and compared to a photo-interpreted river network. Last, the vertical accuracy of each DEM is quantified in relation to LiDAR data.

## A.1 Methods

### A.1.1 Digital terrain analysis

The extraction of terrain descriptors from free DEMs follows a simple workflow using the TauDEM toolbox (see Fig. A.1).

A clipped DEM is first corrected by identifying sinks and by raising cell elevation values to the level of the lowest pour point in the 8 surrounding cells of the structured grid. This is deemed necessary in order to avoid interference with flow routing. From this corrected layer, flow directions from each cell to one of its 8 neighbours are determined by following the D8 flow model and a counter-clockwise coding from 1 (flow to the East) to 8. Using the D8 flow model, the local slope or tangent of the angle of incline,  $\theta$ , is calculated as the drop,  $\Delta y$ , over distance,  $\Delta x$ , between a cell and its neighbours in the flow path:

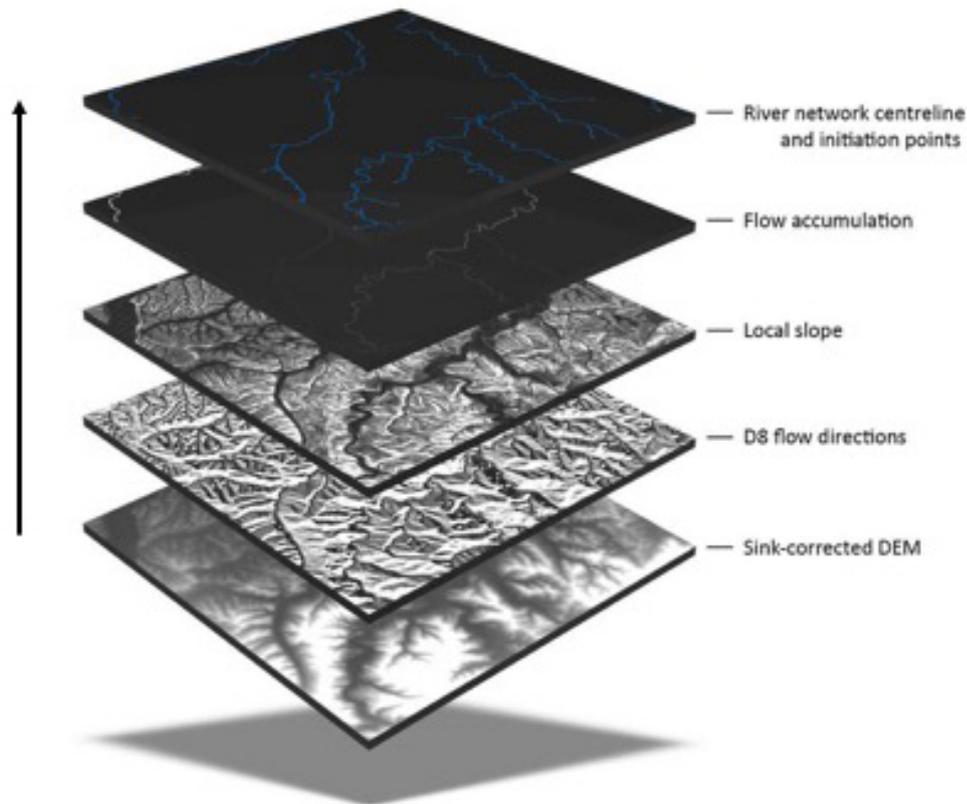
$$\tan\theta = \frac{\Delta y}{\Delta x} \quad (\text{A.1})$$

In turn, the upslope contributing area is obtained by simply accumulating cells following the D8 flow directions.

To delineate river network centrelines from the free DEMs, channel initiation points, or channel heads where water is assumed to begin to flow between banks, are first identified by imposing a threshold of  $10^5 m^2$  (Giannoni et al., 2005) on an area-slope criterion that characterises the transition between transport mechanisms (Montgomery and Dietrich, 1988, 1989), in specific:

$$A_{ij} S_{ij}^k > 10^5 \quad (\text{A.2})$$

where  $A_{ij}$  is the upslope contributing area (in  $m^2$ ) at any given cell, given by the



**Figure A.1:** Workflow for extracting terrain analysis layers from a digital elevation model, necessary to compute the flood index.

product of the flow accumulation and the DEM-specific cell size, and  $S_{ij}$  is the local slope raised to the exponent  $k = 1.75$  (Giannoni et al., 2005) that characterises drainage density changes. Starting at the channel heads, river networks are delineated following the D8 flow directions to the outlet.

With the river network delineated from the free DEMs, the computation of  $H$  and  $D$  was programmed in Python following the description in (Manfreda et al., 2015) and making use of the Geospatial Data Abstraction Library (GDAL) for raster I/O.  $D$  is obtained by counting the number of cells from each position in the DEM raster to the connected river network cell, still following the D8 flow directions. Furthermore, for each unique flow path, adjacent cell counts need to be distinguished from diagonal ones, so that lengths can be obtained by multiplication with the spatial resolution or with the product of spatial resolution and  $\sqrt{2}$ , respectively. The total flow path distance from each location in the raster to the stream is simply the sum of corresponding adjacent and diagonal lengths.

### A.1.2 Accuracy assessment

In order to assess the vertical accuracy of the free DEMs, 3 common error measures were selected to be used with continuous variables, in this case the elevation data. Different error measures are reported in this study as they may complement each other (Chai and Draxler, 2014). The systematic error or statistical bias is defined as the simple difference between DEM,  $\hat{y}_i$ , and LiDAR,  $y_i$ , elevations (here assumed as the truth):

$$BIAS = \hat{y}_i - y_i \quad (A.3)$$

where  $i$  the index of an individual cell in a flattened raster. The MAE is defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (A.4)$$

where  $n$  is the total number of cells. The MAE represents the average absolute difference between DEM and LiDAR elevations and gives an indication of the magnitude of error. Finally, the RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (A.5)$$

where the mean square error is the second moment of the bias. The RMSE also measures the magnitude of error, but with a higher sensitivity to outliers, thus putting stronger emphasis to unfavourable conditions (Chai and Draxler, 2014). Its normalised version that is less sensitive to outliers is given by:

$$NRMSE = 100 * \frac{RMSE}{y_{max} - y_{min}} \quad (A.6)$$

Finally, the linear correlation between  $\hat{y}_i$  and  $y_i$  is also reported and measured using the Pearson's correlation coefficient (PCC) defined as:

$$PCC = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (A.7)$$

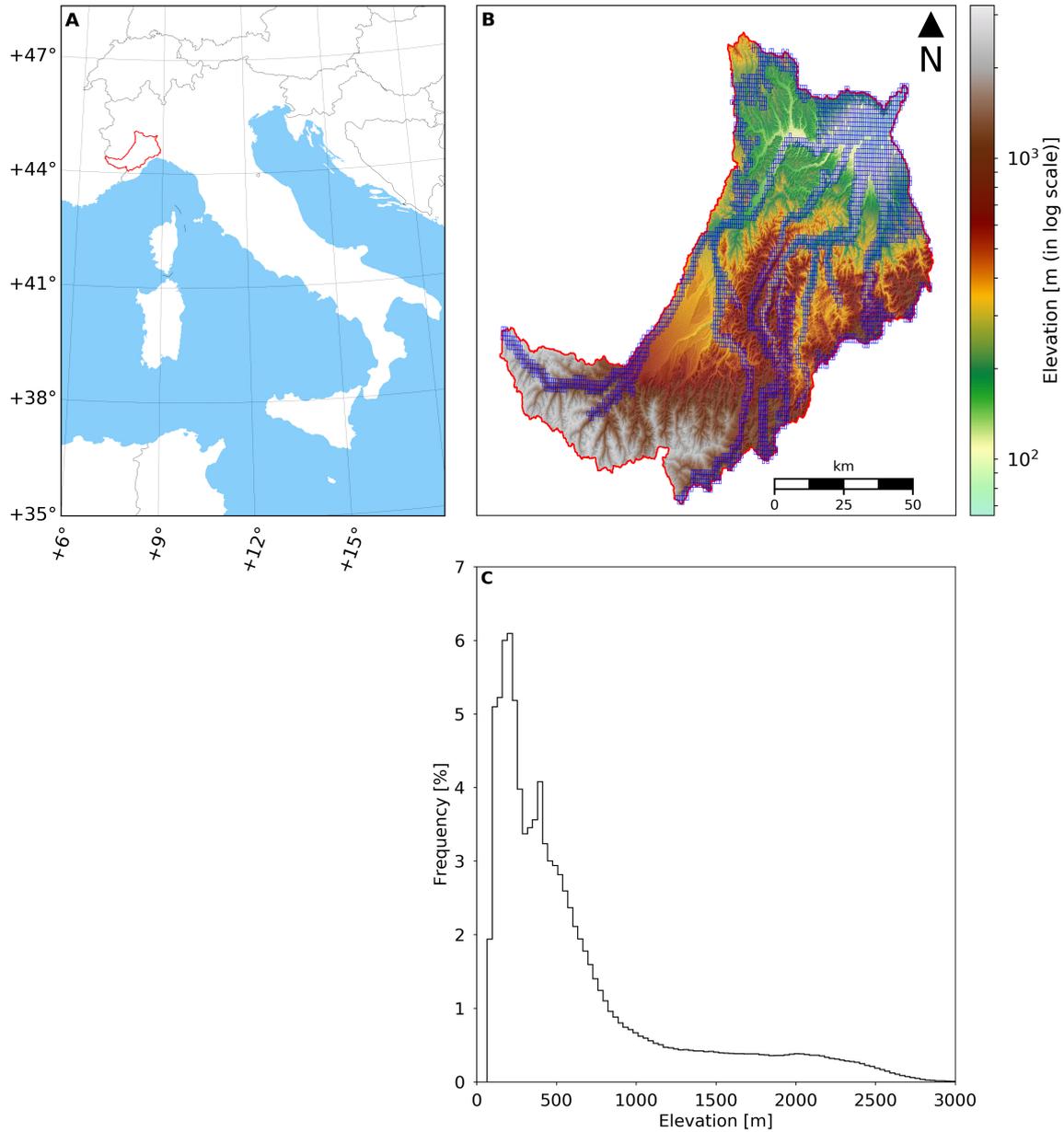
PCC takes values between -1 and 1, with  $PCC = -1$  corresponding to a perfect inverse correlation,  $PCC = 1$  corresponding to a perfect direct correlation and  $PCC = 0$  corresponding to no linear correlation.

## A.2 Case study and datasets

### A.2.1 Tanaro River Basin

With inception in the Ligurian Alps close to France and located in north-western Italy, the Tanaro river is the most significant right-side tributary to the Po River in terms of length (c.a.  $276km$ ) and drainage area (c.a.  $8000km^2$ ), presenting a highly variable discharge (Degiorgis et al., 2012). The Tanaro river basin is characterised by steep mountainous terrain and a nearly flat alluvial region (Fig. A.2). The river itself is highly prone to flooding; indicatively, during the 1994 historical Piedmont flood and landslide, 44 persons lost their lives, 2000 were displaced and a whopping 8.8 billion € in damages were estimated (Luino, 1999).

The Tanaro river basin was chosen as case study due to its peculiar characteristics and history of disastrous events. In this work, the DEMs listed in Table A.1 were clipped with the Tanaro river catchment polygon obtained from the Italian Environmental Agency (ISPRA – *Istituto Superiore per la Protezione e la Ricerca Ambientale*) in shapefile format. The clipped DEMs are used to extract the terrain descriptors within the study area, namely (1) the upslope contributing area, (2) the local slope, (3) the flow path elevation difference to the nearest stream,  $H$ , and (4) the flow path distance to the nearest stream,  $D$ , used for comparison in terms of cumulative frequency curves. In addition, river networks are delineated from the clipped DEMs for visual inspection and comparison with the EU-Hydro photo-interpreted river network (<http://land.copernicus.eu/pan-european/satellite-derived-products/eu-dem>). As a final step, vertical accuracies of the free DEMs are reassessed using LiDAR data.



**Figure A.2:** Representation of the study area. **A)** Location of the Tanaro river basin in the Piedmont region, NW Italy, with the drainage divide highlighted in red; **B)** digital elevation model (DEM) of the Tanaro river basin and footprint of the light detection and ranging (LiDAR) dataset in blue; **C)** histogram of elevations within the Tanaro river basin.

## A.2.2 Digital elevation models datasets

In Table A.1, an overview of some of the most common DEMs with spatial resolutions of the order of 3 arc seconds (c.a. 90m) or less is provided. These datasets are currently in the public domain or available upon request mostly for research or other non-commercial purposes. Table A.1 is organised by ascending order of year of public release. In this work, all datasets in Table A.1 are taken into consideration:

**Table A.1:** Digital elevation models (DEM) currently available free of charge with spatial resolutions below 3 arc seconds.

ID	Name	Entity / Consortium	DEM sources	Spatial reference	Spatial resolution (m)	Vertical accuracy (m)	Type	Year of public release
1	SRTM GL3 DEM	NASA	SRTM, ASTER GDEM topographic maps, LiDAR,	WGS84 EGM96	90	6( <i>MAE</i> )	Global DSM	2003, 2015
2	TINITALY DEM	INGV	GPS data, orthophotos and other	WGS 84	10	6( <i>RMSE</i> )	Italy	2007
3	HydroSHEDS DEM	WWF	SRTM	WGS84 EGM96	90	-	Global hydrologically conditioned DTM	2006 to 2009
4	ASTER GDEM	NASA JPL and METI	-	WGS84 EGM96	30	17(95%)	Global DSM	2009
5	EU-DEM	EEA	SRTM, ASTER GDEM, topographic maps	ETRS89 EVR2000 EGG08	30	7( <i>RMSE</i> )	pan-European DSM	2013
6	VFP DEM	-	SRTM, ASTER GDEM, topographic maps and other	WGS84	90	-	Global DSM	2014
7	SRTM GL1 DEM	NASA	SRTM, ASTER GDEM	WGS84 EGM96	30	6( <i>MAE</i> )	Global DSM	2015
8	AW3D30 DEM	JAXA	-	WGS84 EGM96	30	4( <i>RMSE</i> )	Global DSM	2015
9	MERIT DEM	University of Tokyo	SRTM, ALOS AW3D30, VFP DEM	WGS84 EGM96	90	5( <i>LE90</i> )	Global DTM	2017
10	TDX DEM	DLR	-	WGS84 ellipsoidal	90	10( <i>LE90</i> )	Global DSM	2018

1. The open-access SRTM GL3 released by the United States (US) National Aeronautics and Space Administration (NASA) in 2003 (Rodriguez et al., 2005; Farr et al., 2007).
2. The TINITALY, a seamless DEM for the whole Italian territory developed by Tarquini et al. (2007, 2012) at the Italian National Institute of Geophysics and Volcanology (INGV – *Istituto Nazionale di Geofisica e Vulcanologia*). This DEM is based on the interpolation of heterogeneous data sources, including contour lines and spot heights from Italian topographic maps, global positioning system (GPS) and LiDAR data. The TINITALY is distributed as a non-commercial product available upon request.
3. The open-access HydroSHEDS, a suite of hydrological geo-reference datasets produced and released between 2006 and 2009 by Conservation Science Program of World Wildlife Fund (WWF), within which a seamless hydrologically conditioned global DEM can be found (Lehner et al., 2008). This DEM is derived from the SRTM GL3, but has been hydrologically conditioned using a sequence of automated procedures, namely the deepening of open water surfaces, weeding of coastal zones, stream burning, filtering, moulding of valley courses, sink filling, carving through barriers, and manual corrections.
4. The open-access ASTER GDEM, a dataset produced by a consortium between the NASA Jet Propulsion Laboratory (JPL) and Japan’s Ministry of Economy, Trade and Industry (METI) produced and released in 2009 (Tachikawa et al., 2011).
5. The EU DEM released by the European Environmental Agency (EEA) in 2013, consisting of a seamless open-access pan-European DEM that combines data from the SRTM and ASTER missions with other sources (EEA, 2014).
6. The VFP made available to the public in 2014 as open-access, consisting of a global DEM produced by fusion of SRTM, ASTER GDEM and other elevation data sources.
7. The open-access SRTM GL1 released by NASA in 2018 as an updated version of the SRTM GL3 (Rodriguez et al., 2005; Farr et al., 2007).
8. The open-access AW3D30 global DEM by the Japan Aerospace Exploration Agency (JAXA) publicly available since 2015 (Tadono et al., 2014).
9. The MERIT DEM by Yamazaki et al. (2017), consisting of a seamless global DEM that combines SRTM with AW3D30 and VFP data and has been available upon request since 2017.
10. The open-access TDX (Rizzoli et al., 2017) released by the German Aerospace Centre (DLR – *Deutsches Zentrum für Luft- und Raumfahrt e.V.*) free of charge in 2018.

### A.2.3 Benchmark LiDAR dataset

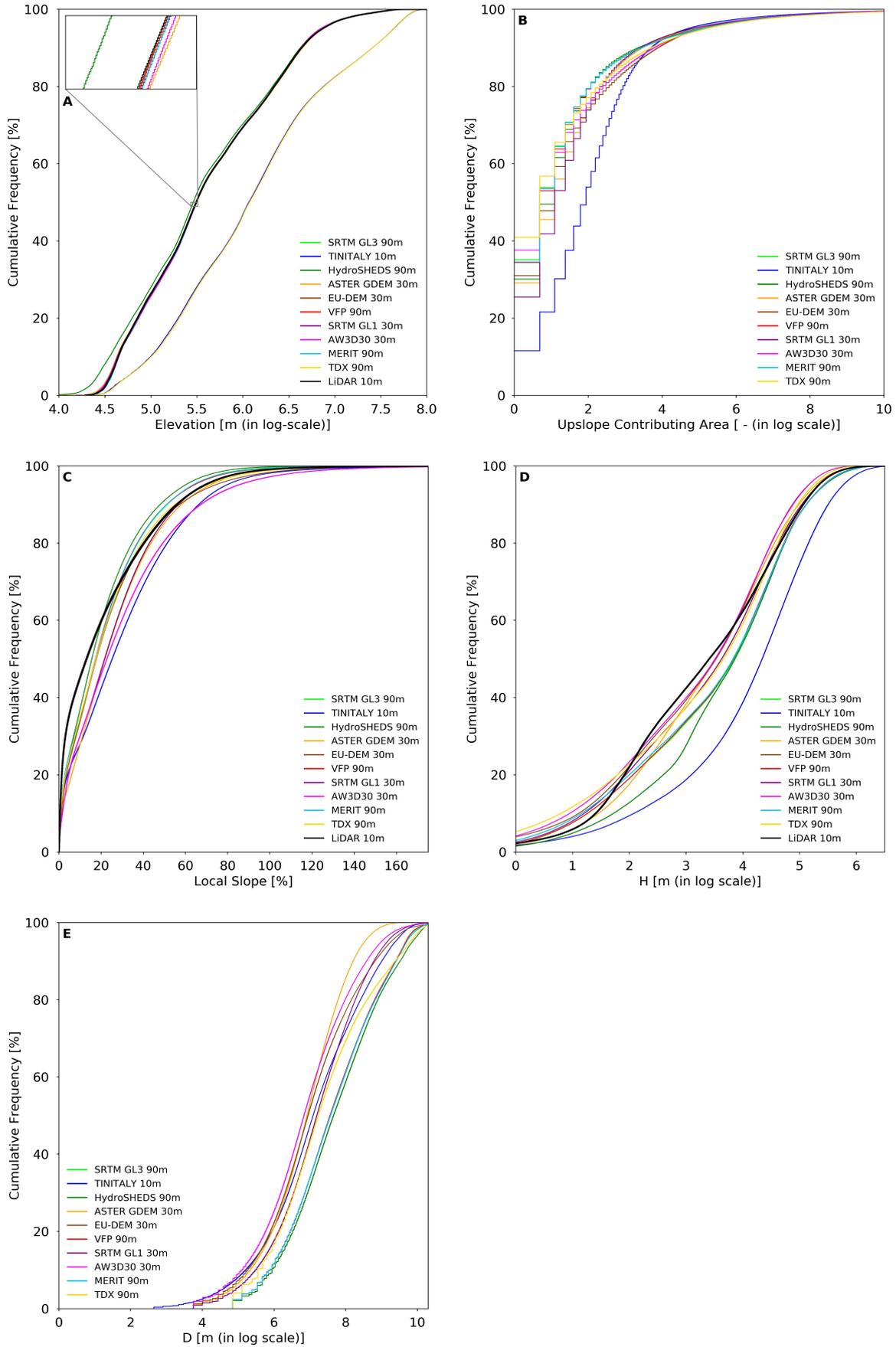
As benchmark for the assessment of vertical accuracies, a LiDAR dataset that partially covers the Tanaro river basin (footprint in Fig. A.2) is used and was obtained from the Italian Ministry of Environment, Land and Sea (*Ministero dell’Ambiente e della Tutela del Territorio e del Mare*). The LiDAR data was resampled to the corresponding spatial resolution of each DEM in Table A.1. Resampling was performed using bilinear filtering (Kirkland, 2010), where the values of the four nearest cells of the original LiDAR to the cell under analysis of the resampled LiDAR, were interpolated using a bilinear function. This non-linear function corresponds to the product of a linear interpolation in the x-direction followed by a linear interpolation in the y-direction of the raster, making it a quadratic interpolator. The LiDAR dataset (with spatial reference WGS84 / ITAL-GE095) is available to the general public upon formal request and upon payment of a processing fee (to visualise its areal coverage, please visit the Italian National Geoportal – *Geoportale Nazionale*).

## A.3 Results and discussion

The cumulative frequency curves of terrain descriptors presented in Fig. A.3 show important differences within the Tanaro river basin and the reader can also refer to Fig. A.4 for the average values of the terrain descriptors. Note that the upslope contributing area and D have not been computed for the LiDAR data, as both indicators proved to be meaningless within the limited extent of the LiDAR footprint.

The EU DEM, the VFP, the SRTM GL1 and GL3, the MERIT, the ASTER GDEM and the AW3D30 DEMs approximate well the LiDAR elevation cumulative frequency. However, this is not necessary or sufficient to assert a faithful representation of geomorphic features. On the other hand, it is also observed that the HydroSHEDS DEM is significantly different from the LiDAR data at lower elevations and that the TDX and TINITALY DEMs is consistently different from LiDAR throughout the entire range of elevations.

For values below  $1 \text{ km}^2$  (c.a. 90% of the basin) DEMs with a finer spatial resolution seem to present lower upslope contributing area values than DEMs with a coarser spatial resolution, while for values above  $1.2 \text{ km}^2$  (c.a. 10% of the basin) DEMs with finer spatial resolutions seem to present larger values than DEMs with coarser ones. The smaller fraction of cells with higher upslope contributing areas can generally be associated with cells belonging to the river network, and channel initiation may occur further upstream in DEMs of finer resolution.



**Figure A.3:** Cumulative frequency curves of terrain descriptors extracted from free digital elevation models (DEM) within the Tanaro river basin, in Italy. **A)** Terrain elevations within the light detection and ranging (LiDAR) footprint; **B)** upslope contributing area; **C)** local slope within the LiDAR footprint; **D)** elevation difference to the nearest stream,  $H$ , within the LiDAR footprint; and **E)** distance to the nearest stream,  $D$ . Logarithmic scales refer to the natural logarithm.

The local slope differs from the LiDAR data to some degree, except in the case of HydroSHEDS, followed closely by the SRTM GL3, the VFP and the MERIT DEM. The best approximations are given by the TDX, the SRTM GL1, the EU DEM and the ASTER GDEM. The average values of local slope show that differences can reach about 10%.

DEMs with spatial resolution of 30 m, in addition to the TDX DEM, give the best approximation of the LiDAR H curve. In particular, the ASTER GDEM perfectly matches the LiDAR cumulative frequency curve at low H values. TINITALY, followed closely by HydroSHEDS at lower H values, presents a significantly different curve from LiDAR, while average values of H are significantly different between every DEM and the LiDAR.

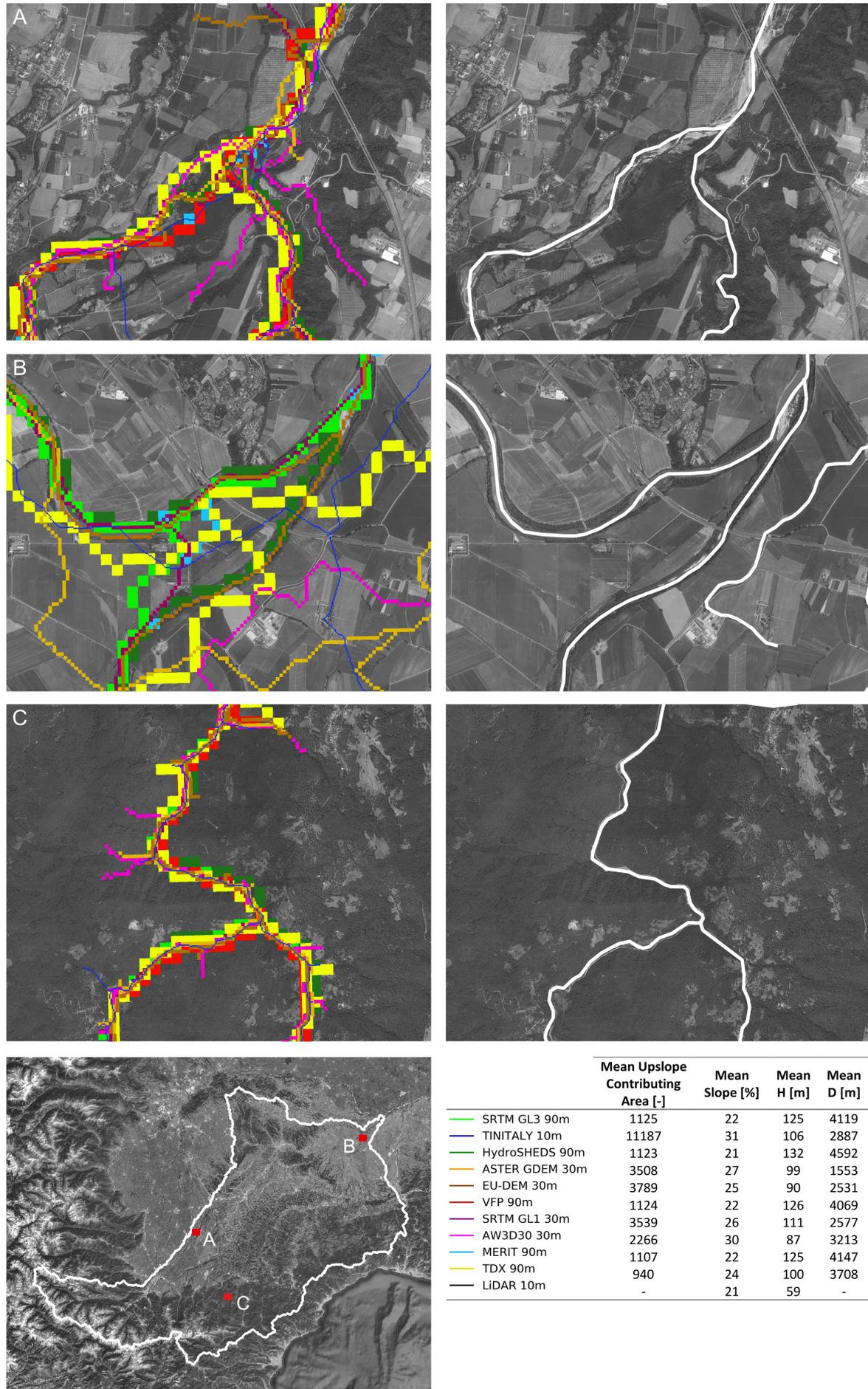
The AW3D30 and the ASTER GDEM present consistently higher cell counts per D contour with respect to the remaining DEMs, while HydroSHEDS and the MERIT DEM present consistently lower cell counts; the SRTM GL1 and the TDX DEM fall more or less between all the other curves.

Regarding the DEM-based river network (Fig. A.4), it is confirmed that extraction can be more problematic in flatter terrain. In particular, the location of river confluences and meanders can be significantly misrepresented. Over flat areas, coarser spatial resolution DEMs tend to better approximate the photo-interpreted river network, while hydrologically reconditioning a DEM (e.g., the HydroSHEDS) also seems to help.

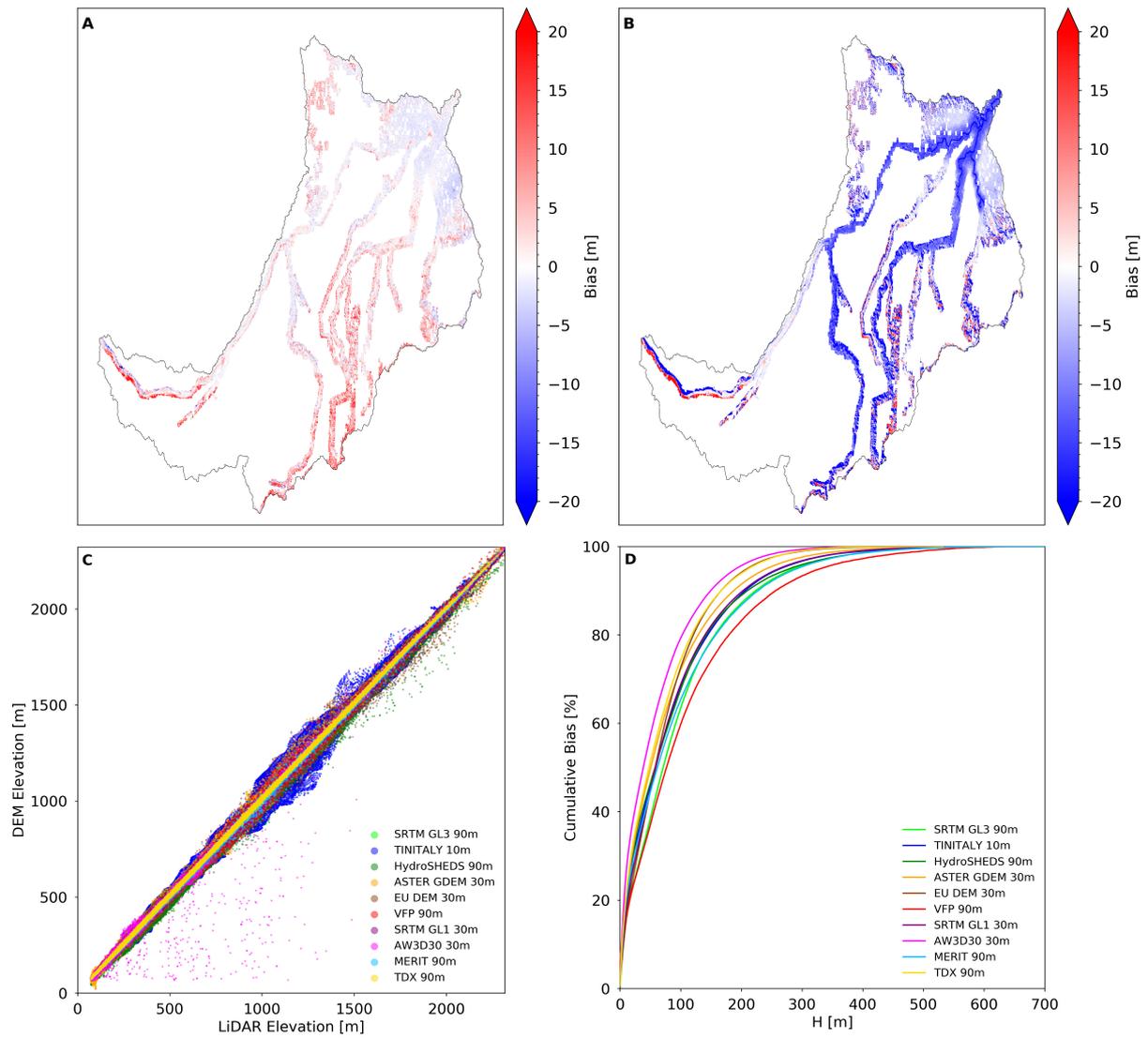
In terms of BIAS (Fig. A.5), a tendency to overestimate elevation values in hillslopes and a tendency to underestimate elevations in floodplains appear to exist, except for the HydroSHEDS DEM that tends to underestimate elevations in a more generalised way. In spite of the BIAS found — less in the case of the SRTM GL3 and VFP DEMs — it is shown that DEMs are highly correlated to the LiDAR data, with the TINITALY, HydroSHEDS and AW3D30 showing some noticeable dispersion.

Vertical accuracy measures were computed from each DEM (Table A.2). The lowest MAE has been obtained by the MERIT DEM, 2.85 m, while the ASTER GDEM and the HydroSHEDS DEM have presented the highest MAE of all the datasets, 11.16 and 6.61 m, respectively. The lowest RMSE, has been presented by the SRTM GL3, 4.83 m, or 0.21% NRMSE, while the ASTER GDEM and the HydroSHEDS DEM have presented, once again, the highest RMSE values among all datasets, 15.78 and 9.39 m, or 0.40 and 0.68% NRMSE, respectively.

Concerning the Tanaro river basin, finer spatial resolutions have not always improved the representation of the morphology. Ideally, a good DEM should aim for an optimal balance between spatial resolution, vertical accuracy and DEM reconditioning. The choice of the EU DEM in Chapter 6 was found to be quite reasonable.



**Figure A.4:** Overlay of river networks for 3 distinct regions in the Tanaro river basin, Italy, derived from free digital elevation models (DEM). For each region, the corresponding EU-



**Figure A.5:** **A)** Map of statistical bias between the SRTM GL3 digital elevation model (DEM) and the light detection and ranging (LiDAR) data in the Tanaro river basin, Italy; **B)** Map of statistical bias between the HydroSHEDS DEM and the LiDAR data; **C)** DEM elevations plotted against LiDAR elevations; **D)** Cumulative frequency of absolute BIAS plotted against the elevation difference to the nearest stream, H.

**Table A.2:** Vertical accuracy assessment of the free digital elevation models (DEM) in the Tanaro river basin, Italy, expressed as mean absolute error (MAE), root mean squared error (RMSE), normalized RMSE (NRMSE) and Pearson correlation (PC), with the light detection and ranging data (LiDAR) data used as benchmark.

DEM	MAE (m)	RMSE (m)	NRMSE (m)	PCC (-)
SRTM GL3 90 m	3.2144	4.8274	0.2083	0.9999
TINITALY 10 m	2.9422	5.5342	0.2330	0.9998
HydroSHEDS 90 m	11.1581	15.7826	0.6811	0.9991
ASTER GDEM 30 m	6.6115	9.3930	0.3972	0.9997
EU-DEM 30 m	4.7818	7.4033	0.3131	0.9997
VFP 90 m	3.9133	7.5296	0.3249	0.9997
SRTM GL1 30 m	3.4330	5.1180	0.2164	0.9999
AW3D30 30 m	3.6570	7.9423	0.3359	0.9997
MERIT 90 m	2.8501	5.0098	0.2162	0.9999
TDX 90 m	3.7389	5.8189	0.2511	0.9999

# Appendix B

## Principal Component Analysis and Regression

Principle Component Analysis (PCA) is a multivariate statistical method often used in water resources research (e.g., Wallis, 1965; Di Prinzio et al., 2011; Salas et al., 2011). Uncorrelated principle components (PCs) are obtained by orthogonal transformation of the original predictors such that the greatest variance is associated with the first PCs. The transformation leads to dimensionality reduction:

$$\mathbf{X} = \mathbf{W}\mathbf{V}^T = \mathbf{W} = \mathbf{X}\mathbf{T} \quad (\text{B.1})$$

where  $\mathbf{X}$  the matrix of  $k$  explanatory variables with size  $(n \times k)$ , excluding the first column,  $\mathbf{W}$  the matrix of PCs  $w_1, w_2, \dots, w_k$  with size  $(n \times k)$ , or scores, and  $\mathbf{V}$  the  $(k \times k)$  symmetric matrix of transformation coefficients  $v_1, v_2, \dots, v_k$ , also known as loadings or rotation matrix, whose columns correspond to the characteristic vectors of the covariance matrix of  $\mathbf{X}$ . The characteristic vectors may be found by:

$$\left( \frac{\mathbf{X}^T \mathbf{X}}{(n-1)} - \lambda_i \mathbf{I} \right) v_i = 0 \quad (\text{B.2})$$

with  $\mathbf{I}$  the  $(k \times k)$  identity matrix and  $\lambda$  the vector of characteristic roots found by solving the characteristic equation, i.e. equating to zero the characteristic polynomial:

$$\left| \frac{\mathbf{X}^T \mathbf{X}}{(n-1)} - \lambda \mathbf{I} \right| = 0 \quad (\text{B.3})$$

The number of PCs, linear combinations of the original covariates with coefficients given by the characteristic vectors, to be used in a principal component regression (PCR) are selected based on the total amount of variance explained. The PCR is fitted by regressing  $\mathbf{y}$  on PCs instead of the raw catchment characteristics:

$$\mathbf{y} = \mathbf{W}\beta + \epsilon \quad (\text{B.4})$$



# Appendix C

## Flood hazard web application

The methodology described in Chapter 2 was implemented as a web application that is currently integrated in the SmartFLOOD platform – a cloud-based system that hosts big data, data-driven models and tools and is connected to flexible computing power (Amazon Web Services, <https://aws.amazon.com>). The concept of the platform was inspired by other existing solutions like the Google Earth Engine (<https://earthengine.google.com/>; Gorelick et al. 2017) that brings massive computational capabilities to planetary-scale geospatial analysis as an integrated platform, the Tethys Platform (<https://www.tethysplatform.org/>; Swain et al. 2016) that implements a development and hosting environment for web applications, and the SWATShare platform (Rajib et al., 2016) that establishes a collaborative environment for hydrology research and education. All of them attempt to organise scientific information and operationalise science.

The web application named Hydrogeomorphic Flood Hazard Mapping has been designed as an innovative tool and is available online to the general public (<http://gecosistema.com/smartflood>). It is currently a demonstrator of the hydrogeomorphic method for flood mapping that makes use of big open-access datasets. The web application was developed to obtain rapid and cost-effective estimates of areas subject to flooding outside their original classification domain, in data-scarce or resource limited settings and across geographical scales. Delineated flood-prone areas are automatically downscaled to pixel size and from the main stream to the tributaries. The quality of the delineation depends on user choices, from the input layers to the selection of areas for classification, downscaling and extrapolation.

In particular, the current version of the web application enables users to:

- run a supervised classification within a user-selected river sub-catchment, based on one of the benchmark flood hazard maps for the six return periods considered, and instantly retrieve the optimal classification outcome and performance;

- perform a DEM-based delineation of flood-prone areas that correspond to a downscaling and extrapolation of selected flood hazard maps;
- visualise the delineated flood-prone areas at 25 m spatial resolution obtained by segmentation of the flood descriptor layer with the optimal classification outcome;
- manually fine-tune the optimal classification outcome via a simple slider, testing on-the-fly its influence on the mapping of flood-prone areas;
- get a dynamic response of performance measures that varies with zoom level and window size.

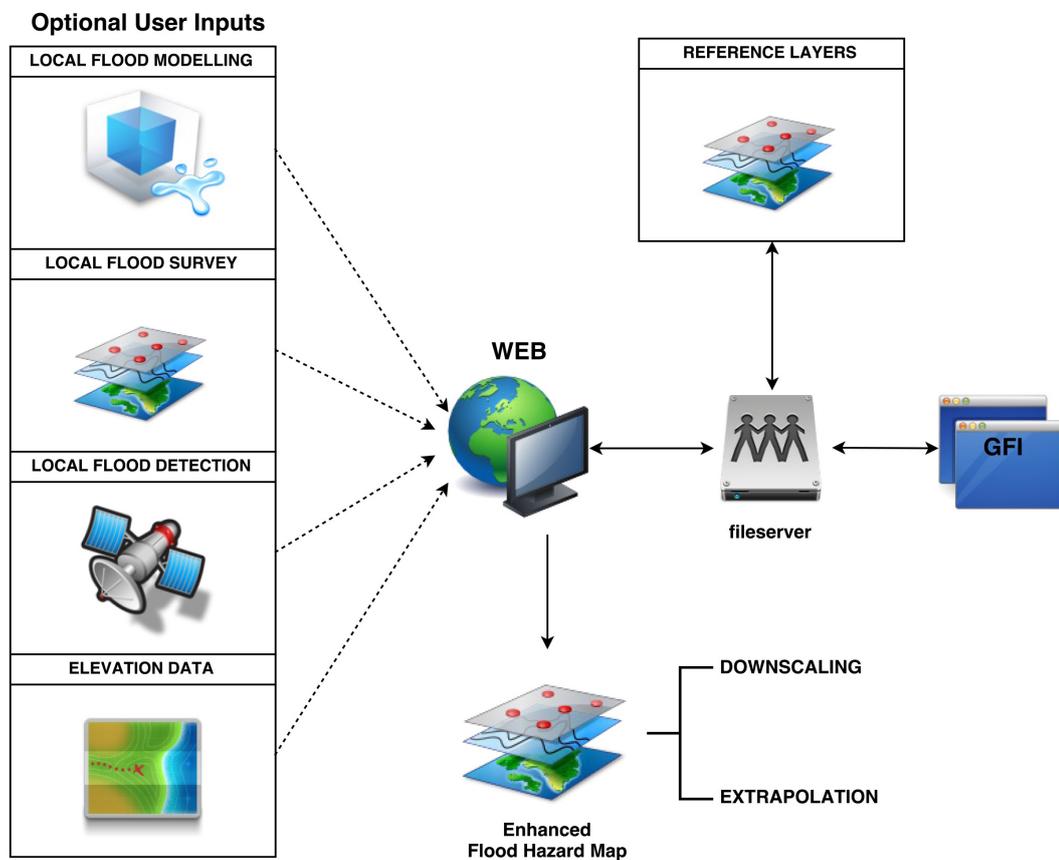
Upload of datasets and download of outputs are currently available for any raster file format supported by the GDAL, visualisation is achieved via HTML5 and JavaScript with the help of OpenLayers and Highcharts libraries, while code sharing is done via a Git-type online version-control system. This prototype is already a step forward in flood hazard assessment and management and its scientific worth can be considered manifold:

- It provides fast and inexpensive estimates of flood-prone areas for specific return periods.
- It provides fast and inexpensive extrapolation and downscaling capabilities of existing flood studies (e.g., modelling results, remote sensing detection, field surveys), particularly in unstudied areas or tributaries, across political borders and data-scarce regions.
- It can be used to test and conduct experiments with different flood descriptors, classification methods and datasets.
- It enables and exploits collaborative and community-based activities (e.g., Citizen and hydrology working group of the International Association of Hydrological Sciences <https://iahs.info/Commissions--W-Groups/Working-Groups/Candhy.do>).

The web application may evolve to a more powerful tool, integrating other data-driven methods and flood information, that attends to users' needs and expectations in supporting evidence-based decision making. Examples of similar tools are the Water Information Forecast Framework (WIFF) that integrates process-based models for waves, tides and surges across geographic scales and predicts floods in coastal areas (Fortunato et al., 2017), the Iowa Flood Information System (Demir and Krajewski, 2013) that provides access to flood maps and other flood-related information using an interactive interface or the FLIRE DSS (Kochilakis et al., 2016) that integrates flood maps with flood warnings, observations and remote sensing data to inform decision makers through a web-based system.

## C.1 Architecture

A cloud-based client-server model was adopted as the web application architecture. The implementation of the web application is illustrated as a network diagram in Fig. C.1. The server host functions as a web and file server for restricted uploading and storing of static layers and results, as well as for their retrieval. Clients and server communicate over the internet via any modern web browser. The web application framework incorporates a Web-GIS front-end made of a combination of HTML5 (a markup language) and OpenLayers (an open source JavaScript library), and a server-side, which combines MapServer (an open source geographic data rendering engine) and Apache (an open-source HTTP server). All core model functions are written in Python.



**Figure C.1:** Client-server diagram showing the implementation of the web application.

## C.2 Back-end

Web application I/O is done using GDAL and is available in any raster file format supported by this library. The core model functions accessible to users consist of:

1. preparing data for supervised classification;

2. linear binary classification of the flood descriptor;
3. computation of classification performance measures.

When the user selects a river sub-catchment and return period in the web user interface, the algorithm searches for the corresponding datasets, constricts the layers to the pre-defined classification area and vectorises them.

The classification stage is performed with the help of scikit-learn, a Python library for machine learning functionalities. A number of thresholds are fixed based on the range of values of the bounded flood descriptor layer. For each threshold, statistics are computed and used to calculate the objective function. The algorithm then solves the optimisation problem by maximising the objective function, through retrieval of the corresponding optimal threshold, the classification outcome. The final flood-prone areas are delineated within the whole river basin (downscaled extrapolation) to which the classification sub-catchment belongs. The delineation is achieved by optimal threshold segmentation of the flood descriptor layer, which discerns flood-prone from flood-free areas. Finally, within the selected classification area, the flood-prone areas are compared to the benchmark flood hazard map. Several performance metrics (including, but not limited to the ones used in this thesis) are computed and made available to the user. It should be noted that the pre-processing and morphological characterisation steps, executed offline to produce the georeferenced static layers, are currently not part of the web application workflow.

### C.3 Front-end

The responsive Web-GIS single page front-end is designed to allow users to visualise and interact with the hydrogeomorphic method at the pan-European scale, using any modern web browser. The front-end is built with HTML5, taking advantage of JavaScript, OpenLayers and Highcharts to draw graphics. OpenLayers provides the client-side mapping utilities to handle geospatial data, while Highcharts provides the additional features that complement the interface. The final results are styled and published by the MapServer on the server-side. Fig. C.2 shows the web user interface as a step-by-step usage example of the web application (video walkthroughs are also provided as supplementary material, MMC 1 and MMC 2). The web application is implemented in the most user-friendly way possible, enabling users to test the methodology without being concerned about the details behind the system. The user requests the web application core functionality through a simple point-and-click application programming interface, selecting the river sub-catchment and return period of the reference desired to execute the linear binary classification of flood-prone areas. Function returns are made visible to

the user through the Web-GIS interface as a combination of interactive maps, summary tables and plots.

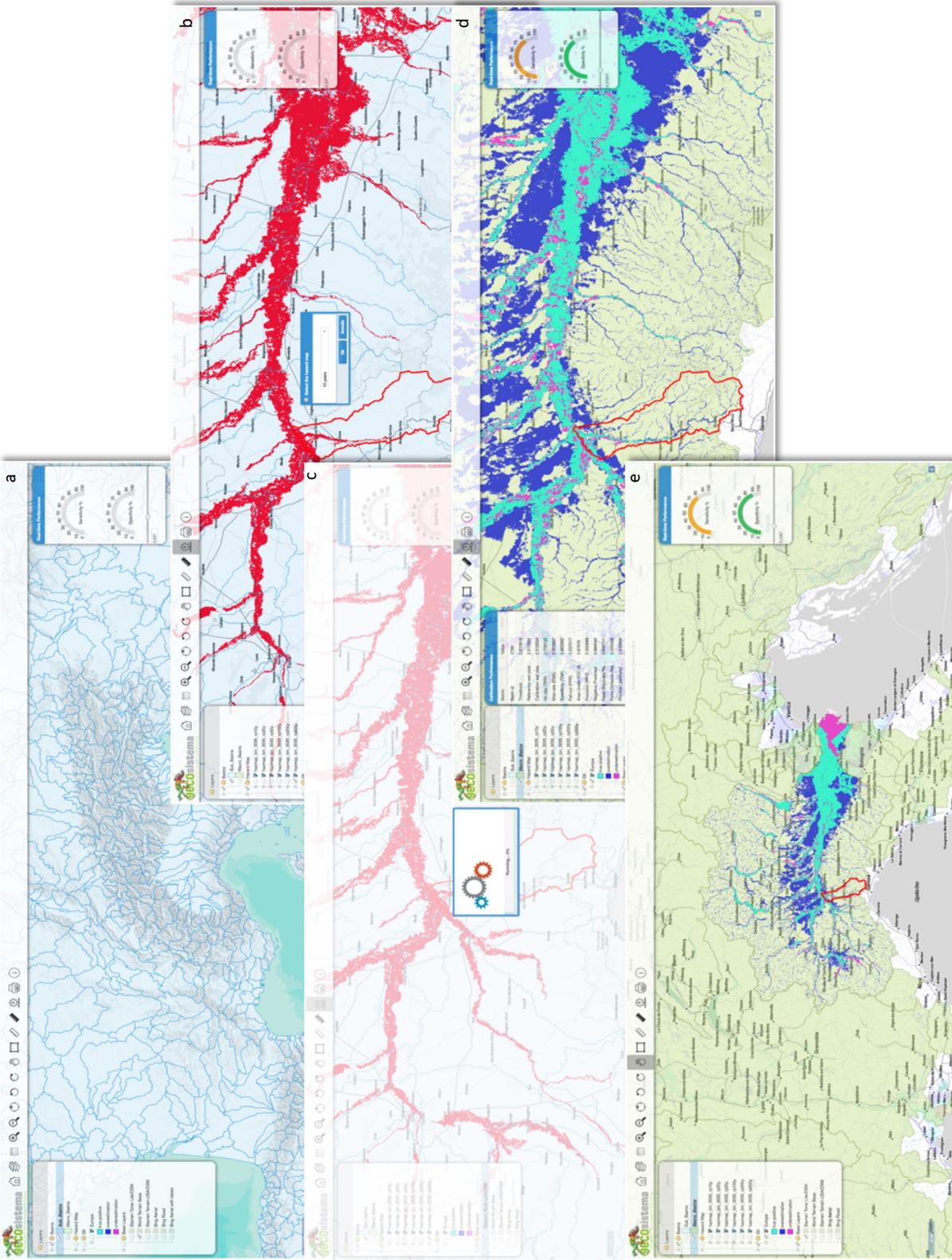
The classification of flood-prone areas results in the maximisation of the method's success, i.e. finding the threshold that corresponds to the optimal operating point in the ROC curve, although not necessarily its utility (Peirce, 1884; Baker and Kramer, 2007). To address this, a slider was implemented to allow the user to adjust on-the-fly the rendered optimal flood-prone areas that resulted from the classification stage, so that subjective benefits and costs of a delineation can be taken into account (Baker and Kramer, 2007). Furthermore, a dynamic response of performance measures that varies with zoom level and window size was implemented based on the pixel count of the two critical classes (i.e., flood-prone and flood-free areas), constricted by the map bounding box. Finally, having static layers in the file server was used to the systems advantage so that it could render the maps faster across scales using pre-built tile caching (MapCache).

## C.4 Parallelisation strategy

The same concurrent programming model is used in different phases of the methodological workflow, from pre-processing and morphological characterisation to classification and validation. An exception was made for terrain analysis, which followed a different parallelisation strategy, used as an out-of-the-box feature of TauDEM utilities. The domain is decomposed in logical units (i.e., hydrological unique river basins and sub-catchments) used as natural geometric domain partitions for parallel computation (data parallelism). Decomposition is achieved using utilities from GDAL to process the geospatial data. Execution is achieved concurrently and asynchronously, distributing queued data across sub-processes spawned from a pool, carrying out exactly the same independent computations on each partition, minimising the total execution time and avoiding overload.

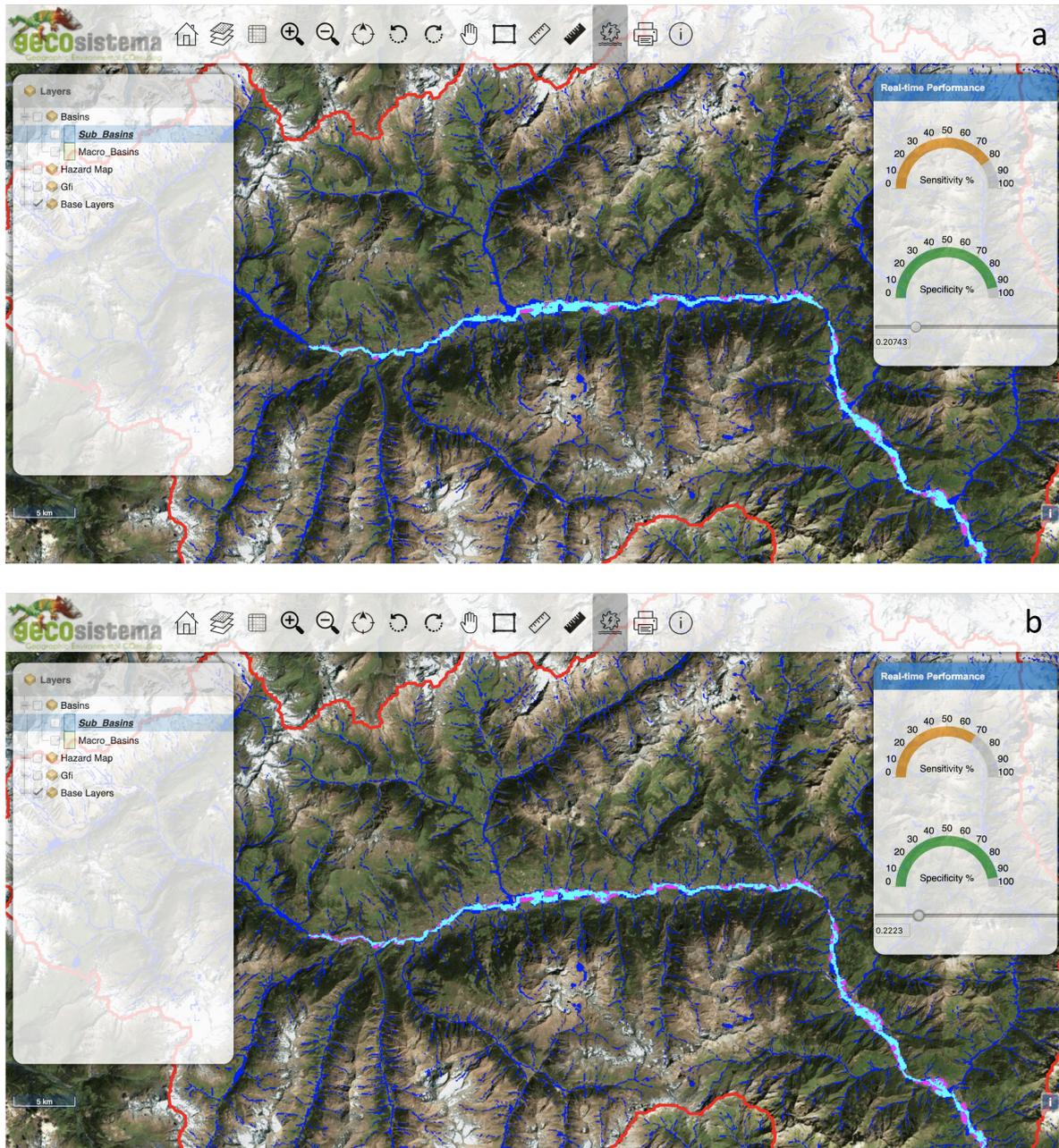
## C.5 Web application generated flood hazard maps

In Fig. C.3, an output of the web application relative to a specific classification is presented. The figure provides a comparison between the hydrogeomorphic flood-prone areas delineated using the web application (merger of cyan and dark blue pixels) and the reference flood extents derived from the 100-year flood hazard map for Europe (merger of cyan and magenta pixels) from Dottori et al. (2016). The example refers to a calibrated threshold value of the flood descriptor (Fig. C.3a) and a slider operation to control the threshold (Fig. C.3b), where the user changes the calibrated threshold and gets an immediate feedback on overestimated (dark blue pixels) and underestimated (magenta



**Figure C.2:** Usage exemplification of the web application based on the hydrogeomorphic method for delineating flood-prone areas over large scales. **a)** river sub-catchment selection; **b)** reference flood map selection, with associated return period; **c)** classification of morphological descriptor; **d)** visualisation of classification results and performance measures (top-left corner); and, **e)** downscaled extrapolation of the optimal threshold and delineation of flood-prone areas over the corresponding river basin, with dynamic response of performance (top-right corner).

pixels) flood-prone areas and how they change visually in the study region or catchment.



**Figure C.3:** Sample overlay of the 100-year return period flood hazard map for Europe (Dottori et al., 2016) on the delineated hydrogeomorphic flood-prone areas using the a) optimal flood descriptor threshold and b) flood descriptor threshold controlled through slider operation (marked with a yellow circle).

# References

- Aerts, J. C. J. H., Renssen, H., Ward, P. J., de Moel, H., Odada, E., Bouwer, L. M., and Goosse, H. (2006). Sensitivity of global river discharges under holocene and future climate conditions. *Geophysical Research Letters*, 33(19). <https://doi.org/10.1029/2006GL027493>.
- Afshari, S., Tavakoly, A. A., Rajib, M. A., Zheng, X., Follum, M. L., Omranian, E., and Fekete, B. M. (2018). Comparison of new generation low-complexity flood inundation mapping tools with a hydrodynamic model. *Journal of Hydrology*, 556:539 – 556. <https://doi.org/10.1016/j.jhydrol.2017.11.036>.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716 – 723. <https://doi.org/10.1109/TAC.1974.11100705>.
- Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K., and Feyen, L. (2017). Global projections of river flood risk in a warmer world. *Earth's Future*, 5(2):171 – 182. <https://doi.org/10.1002/2016EF000485>.
- Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., and Pappenberger, F. (2013). Glofas – global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, 17(3):1161 – 1175. <https://doi.org/10.5194/hess-17-1161-2013>.
- Alfieri, L., Burek, P., Feyen, L., and Forzieri, G. (2015). Global warming increases the frequency of river floods in europe. *Hydrology and Earth System Sciences*, 19(5):2247 – 2260. <https://doi.org/10.5194/hess-19-2247-2015>.
- Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., and Feyen, L. (2014). Advances in pan-european flood hazard mapping. *Hydrological Processes*, 28(13):4067 – 4077. <https://doi.org/10.1002/hyp.9947>.
- Annis, A., Nardi, F., Morrison, R. R., and Castelli, F. (2019). Investigating hydrogeomorphic floodplain mapping performance with varying dtm resolution and stream order. *Hydrological Sciences Journal*, 64(5):525 – 538. <https://doi.org/10.1080/02626667.2019.1591623>.

- Archer, L., Neal, J. C., Bates, P. D., and House, J. I. (2018). Comparing tandem-x data with frequently used dems for flood inundation modeling. *Water Resources Research*, 54(12):43–69. <https://doi.org/10.1029/2018WR023688>.
- Arnaud-Fassetta, G. (2013). Dyke breaching and crevasse-splay sedimentary sequences of the rhône delta, france, caused by extreme riverflood of december 2003. *Geografia Fisica e Dinamica Quaternaria*, pages 7–26. <https://doi.org/DOI10.4461/GFDQ.2013.36.1>.
- Arnold, J. G. and Fohrer, N. (2005). Swat2000: current capabilities and research opportunities in applied watershed modelling. *Hydrological Processes*, 19(3):563–572. <https://doi.org/10.1002/hyp.5611>.
- Baker, S. G. and Kramer, B. S. (2007). Peirce, youden, and receiver operating characteristic curves. *The American Statistician*, 61(4):343–346. <https://doi.org/10.1198/000313007X247643>.
- Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., and Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5):412–424. <https://doi.org/10.1093/bioinformatics/16.5.412>.
- Barichivich, J., Gloor, E., Peylin, P., Brienen, R. J. W., Schöngart, J., Espinoza, J. C., and Pattanayak, K. C. (2018). Recent intensification of amazon flooding extremes driven by strengthened walker circulation. *Science Advances*, 4(9). <https://doi.org/10.1126/sciadv.aat8785>.
- Barredo, J. I. (2007). Major flood disasters in europe: 1950 - 2005. *Natural Hazards*, 42(1):125–148. <https://doi.org/10.1007/s11069-006-9065-2>.
- Bartholmes, J. C., Thielen, J., Ramos, M. H., and Gentilini, S. (2009). The european flood alert system efas – part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrology and Earth System Sciences*, 13(2):141–153. <https://doi.org/10.5194/hess-13-141-2009>.
- Bates, P. D., Horritt, M. S., and Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, 387(1):33–45. <https://doi.org/10.1016/j.jhydrol.2010.03.027>.
- Beighley, R. E., Ray, R. L., He, Y., Lee, H., Schaller, L., Andreadis, K. M., Durand, M., Alsdorf, D. E., and Shum, C. K. (2011). Comparing satellite derived precipitation datasets using the hillslope river routing (hrr) model in the congo river basin. *Hydrological Processes*, 25(20):3216–3229. <https://doi.org/10.1002/hyp.8045>.
- Beven, K. J. and Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology / un modèle à base physique de zone d’appel variable de l’hydrologie du bassin versant. *Hydrological Sciences Bulletin*, 24(1):43–69. <https://doi.org/10.1080/02626667909491834>.

- Bhowmik, N. G. (1984). Hydraulic geometry of floodplains. *Journal of Hydrology*, 68(1):369 – 401. [https://doi.org/10.1016/0022-1694\(84\)90221-X](https://doi.org/10.1016/0022-1694(84)90221-X).
- Blöschl, G. and Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes*, 9(3-4):251–290. <https://doi.org/10.1002/hyp.3360090305>.
- Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145 – 1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5 – 32. <https://doi.org/10.1023/A:1010933404324>.
- Caprario, J. and Finotti, A. R. (2019). Socio-technological tool for mapping susceptibility to urban flooding. *Journal of Hydrology*, 574:1152 – 1163. <https://doi.org/10.1016/j.jhydrol.2019.05.005>.
- Cartier, D. and Fuamba, M. (2013). Discussion of "detection of flood-prone areas using digital elevation models"; by salvatore manfreda, margherita di leo, and aurelia sole. *Journal of Hydrologic Engineering*, 18(3):361–362. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000652](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000652).
- Cassalho, F., Beskow, S., de Mello, C. R., de Moura, M. M., de Oliveira, L. F., and de Aguiar, M. S. (2019). Artificial intelligence for identifying hydrologically homogeneous regions: A state-of-the-art regional flood frequency analysis. *Hydrological Processes*, 33(7):1101–1116. <https://doi.org/10.1002/hyp.13388>.
- Castellarin, A., Persiano, S., Pugliese, A., Aloe, A., Skøien, J. O., and Pistocchi, A. (2018). Prediction of streamflow regimes over large geographical areas: interpolated flow–duration curves for the danube region. *Hydrological Sciences Journal*, 63(6):845 – 861. <https://doi.org/10.1080/02626667.2018.1445855>.
- Chai, T. and Draxler, R. R. (2014). Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7(3):1247 – 1250. <https://doi.org/10.5194/gmd-7-1247-2014>.
- Chen, L. and Wang, L. (2018). Recent advance in earth observation big data for hydrology. *Big Earth Data*, 2(1):86 – 107. <https://doi.org/10.1080/20964471.2018.1435072>.
- Clubb, F. J., Mudd, S. M., Milodowski, D. T., Valters, D. A., Slater, L. J., Hurst, M. D., and Limaye, A. B. (2017). Geomorphometric delineation of floodplains and terraces from objectively defined topographic thresholds. *Earth Surface Dynamics*, 5(3):369 – 385. <https://doi.org/10.5194/esurf-5-369-2017>.

- Coltin, B., McMichael, S., Smith, T., and Fong, T. (2016). Automatic boosted flood mapping from satellite data. *International Journal of Remote Sensing*, 37(5):993 – 1015. <https://doi.org/10.1080/01431161.2016.1145366>.
- Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D. (2018). An ensemble version of the e-obs temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17):9391 – 9409. <https://doi.org/10.1029/2017JD028200>.
- Cramér, H. (1999). *Mathematical Methods of Statistics*, volume Princeton Mathematical Series 9. Princeton University Press.
- CRED/UNISDR (2016). The human cost of weather related disasters 1995-2015. Technical report, Université Catholique de Louvain, Centre for Research of the Epidemiology of Disasters (CRED) and The United Nations Office for Disaster Risk Reduction (UNISDR). [https://www.unisdr.org/2015/docs/climatechange/COP21\\_WeatherDisastersReport\\_2015\\_FINAL.pdf](https://www.unisdr.org/2015/docs/climatechange/COP21_WeatherDisastersReport_2015_FINAL.pdf).
- Croneborg, L., Saito, K., Matera, M., McKeown, D., and van Aardt, J. (2015). A guidance note on how digital elevation models are created and used – includes key definitions, sample terms of reference and how best to plan a dem-mission. Technical report, International Bank for Reconstruction and Development, New York, US. <https://opendri.org/wp-content/uploads/2016/06/Digital-Elevation-Models-10-23-15-web.pdf>.
- Cunge, J. A. (1969). On the subject of a flood propagation computation method (muskingum method). *Journal of Hydraulic Research*, 7(2):205 – 230. <https://doi.org/10.1080/00221686909500264>.
- Cutler, D. R., Edwards Jr., T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., and Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11):2783 – 2792. <https://doi.org/10.1890/07-0539.1>.
- de Jager, A. L. and Vogt, J. V. (2010). Development and demonstration of a structured hydrological feature coding system for europe. *Hydrological Sciences Journal*, 55(5):661 – 675. <https://doi.org/10.1080/02626667.2010.490786>.
- de Moel, H., van Alphen, J., and Aerts, J. C. J. H. (2009). Flood maps in europe - methods, availability and use. *Natural Hazards and Earth System Sciences*, 9(2):289 – 301. <https://doi.org/10.5194/nhess-9-289-2009>.
- De Risi, R., Jalayer, F., De Paola, F., and Giugni, M. (2014). Probabilistic delineation of flood-prone areas based on a digital elevation model and the extent of historical flooding: the case of ouagadougou. *Boletín Geológico y Minero*, 125(3):329 – 340.
- De Risi, R., Jalayer, F., De Paola, F., and Lindley, S. (2017). Delineation of flooding risk hotspots based on digital elevation model, calculated and historical flooding extents:

- the case of ouagadougou. *Stochastic Environmental Research and Risk Assessment*, 32(6):1545 – 1559. <https://doi.org/10.1007/s00477-017-1450-8>.
- Degiorgis, M., Gnecco, G., Gorni, S., Roth, G., Sanguineti, M., and Taramasso, A. C. (2012). Classifiers for the detection of flood-prone areas using remote sensed elevation data. *Journal of Hydrology*, 470 - 471:302 – 315. <https://doi.org/10.1016/j.jhydrol.2012.09.006>.
- Demir, I. and Krajewski, W. F. (2013). Towards an integrated flood information system: Centralized data access, analysis, and visualization. *Environmental Modelling & Software*, 50:77 – 84. <https://doi.org/10.1016/j.envsoft.2013.08.009>.
- Di Prinzio, M., Castellarin, A., and Toth, E. (2011). Data-driven catchment classification: application to the pub problem. *Hydrology and Earth System Sciences*, 15(6):1921 – 1935. <https://doi.org/10.5194/hess-15-1921-2011>.
- Dodov, B. and Fofoula-Georgiou, E. (2005). Fluvial processes and streamflow variability: Interplay in the scale-frequency continuum and implications for scaling. *Water Resources Research*, 41(5). <https://doi.org/10.1029/2004WR003408>.
- Dodov, B. A. and Fofoula-Georgiou, E. (2006). Floodplain morphometry extraction from a high-resolution digital elevation model: a simple algorithm for regional analysis studies. *IEEE Geoscience and Remote Sensing Letters*, 3(3):410 – 413. <https://doi.org/10.1109/LGRS.2006.874161>.
- Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., and Feyen, L. (2016). Development and evaluation of a framework for global flood hazard mapping. *Advances in Water Resources*, 94:87 – 102. <https://doi.org/10.1016/j.advwatres.2016.05.002>.
- EEA (2014). Eu-dem statistical validation. Technical report, European Environment Agency (EEA), Copenhagen, Denmark,. EEA Report.
- EEA (2016). Flood risks and environmental vulnerability: Exploring the synergies between floodplain restoration, water policies and thematic policies. EEA Report 1/2016, European Environment Agency, Copenhagen, Denmark. <https://www.eea.europa.eu/publications/flood-risks-and-environmental-vulnerability>.
- Elmer, F., Hoymann, J., DÜthmann, D., Vorogushyn, S., and Kreibich, H. (2012). Drivers of flood risk change in residential areas. *Natural Hazards and Earth System Sciences*, 12(5):1641 – 1657. <https://doi.org/10.5194/nhess-12-1641-2012>.
- Elshorbagy, A., Bharath, R., Lakhanpal, A., Ceola, S., Montanari, A., and Lindenschmidt, K.-E. (2017). Topography- and nightlight-based national flood risk assessment in canada. *Hydrology and Earth System Sciences*, 21(4):2219 – 2232. <https://doi.org/10.5194/hess-21-2219-2017>.

- Falter, D., Dung, N., Vorogushyn, S., Schröter, K., Hundecha, Y., Kreibich, H., Apel, H., Theisselmann, F., and Merz, B. (2016). Continuous, large-scale simulation model for flood risk assessments: proof-of-concept. *Journal of Flood Risk Management*, 9(1):3 – 21. <https://doi.org/10.1111/jfr3.12105>.
- Falter, D., Schröter, K., Dung, N. V., Vorogushyn, S., Kreibich, H., Hundecha, Y., Apel, H., and Merz, B. (2015). Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. *Journal of Hydrology*, 524:182 – 193. <https://doi.org/10.1016/j.jhydrol.2015.02.021>.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2). <https://doi.org/10.1029/2005RG000183>.
- Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861 – 874. <https://doi.org/10.1016/j.patrec.2005.10.010>.
- Fenn, T., Clarke, C., Burgess-Gamble, L., Harding, E., Ogunyoye, F., Hick, E., Dawks, S., Morris, J., and Chatterton, J. (2016). The costs and impacts of the winter 2013/14 floods in england and wales. *E3S Web Conf.*, 7:05004. <https://doi.org/10.1051/e3sconf/20160705004>.
- Fortunato, A. B., Oliveira, A., Rogeiro, J., da Costa, R. T., Gomes, J. L., Li, K., de Jesus, G., Freire, P., Rilo, A., Mendes, A., Rodrigues, M., and Azevedo, A. (2017). Operational forecast framework applied to extreme sea levels at regional and local scales. *Journal of Operational Oceanography*, 10(1):1 – 15. <https://doi.org/10.1080/1755876X.2016.1255471>.
- Fowler, H. J. and Kilsby, C. G. (2003). A regional frequency analysis of united kingdom extreme rainfall from 1961 to 2000. *International Journal of Climatology*, 23(11):1313 – 1334. <https://doi.org/10.1002/joc.943>.
- Garmdareh, E. S., Vafakhah, M., and Eslamian, S. S. (2018). Regional flood frequency analysis using support vector regression in arid and semi-arid regions of iran. *Hydrological Sciences Journal*, 63(3):426–440. <https://doi.org/10.1080/02626667.2018.1432056>.
- Georgakakos, A. P., Georgakakos, K. P., and Baltas, E. A. (1990). A state-space model for hydrologic river routing. *Water Resources Research*, 26(5):827 – 838. <https://doi.org/10.1029/WR026i005p00827>.
- Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G. (2011). Hydrological landscape classification: investigating the performance of hand based landscape classifications in a central european meso-scale catchment. *Hydrology and Earth System*

- Sciences*, 15(11):3275 – 3291. <https://www.hydrol-earth-syst-sci.net/15/3275/2011/>.
- Giannoni, F., Roth, G., and Rudari, R. (2005). A procedure for drainage network identification from geomorphology and its application to the prediction of the hydrologic response. *Advances in Water Resources*, 28(6):567 – 581. <https://doi.org/10.1016/j.advwatres.2004.11.013>.
- Giovannettone, J., Copenhaver, T., Burns, M., and Choquette, S. (2018). A statistical approach to mapping flood susceptibility in the lower connecticut river valley region. *Water Resources Research*, 54(10):7603 – 7618. <https://doi.org/10.1029/2018WR023018>.
- Godbout, L., Zheng, J. Y., Dey, S., Eyelade, D., Maidment, D., and Passalacqua, P. (2019). Error assessment for height above the nearest drainage inundation mapping. *JAWRA Journal of the American Water Resources Association*, 55(4):952 – 963. <https://doi.org/10.1111/1752-1688.12783>.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18 – 27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Green, W. and Ampt, G. (1911). Studies on soil physics, part 1, the flow of air and water through soils. *Journal of Agricultural Science*, 4:11–24.
- Grimaldi, S., Petroselli, A., Arcangeletti, E., and Nardi, F. (2013). Flood mapping in ungauged basins using fully continuous hydrologic–hydraulic modeling. *Journal of Hydrology*, 487:39 – 47. <https://doi.org/10.1016/j.jhydrol.2013.02.023>.
- Haddad, K., Rahman, A., and Stedinger, J. R. (2012). Regional flood frequency analysis using bayesian generalized least squares: a comparison between quantile and parameter regression techniques. *Hydrological Processes*, 26(7):1008 – 1021. <https://doi.org/10.1002/hyp.8189>.
- Hirsch, R. M., Helsel, D. R., Cohn, T. A., and Gilroy, E. J. (1992). *Statistical analysis of hydrological data*. McGraw-Hill, New York. in: Maidment, R. (Ed.), Handbook of Hydrology.
- Hjerdt, K. N., McDonnell, J. J., Seibert, J., and Rodhe, A. (2004). A new topographic index to quantify downslope controls on local drainage. *Water Resources Research*, 40(5). <https://doi.org/10.1029/2004WR003130>.
- Hoch, J. M., Neal, J. C., Baart, F., van Beek, R., Winsemius, H. C., Bates, P. D., and Bierkens, M. F. P. (2017). Glofrim v1.0 – a globally applicable computational framework for integrated hydrological–hydrodynamic modelling. *Geoscientific Model Development*, 10(10):3913 – 3929. <https://doi.org/10.5194/gmd-10-3913-2017>.

- Horritt, M. and Bates, P. (2001). Effects of spatial resolution on a raster based model of flood flow. *Journal of Hydrology*, 253(1):239 – 249. [https://doi.org/10.1016/S0022-1694\(01\)00490-5](https://doi.org/10.1016/S0022-1694(01)00490-5).
- Horritt, M. and Bates, P. (2002). Evaluation of 1d and 2d numerical models for predicting river flood inundation. *Journal of Hydrology*, 268(1):87 – 99. [https://doi.org/10.1016/S0022-1694\(02\)00121-X](https://doi.org/10.1016/S0022-1694(02)00121-X).
- Horton, R. (1932). Drainage-basin characteristics. *Eos, Transactions American Geophysical Union*, 13(1):350 – 361. <https://doi.org/10.1029/TR013i001p00350>.
- Hundecha, Y. and Merz, B. (2012). Exploring the relationship between changes in climate and floods using a model-based analysis. *Water Resources Research*, 48(4). <https://doi.org/10.1029/2011WR010527>.
- Huntingford, C., Marsh, T., Scaife, A. A., Kendon, E. J., Hannaford, J., Kay, A. L., Lockwood, M., Prudhomme, C., Reynard, N. S., Parry, S., Lowe, J. A., Screen, J. A., Ward, H. C., Roberts, M., Stott, P. A., Bell, V. A., Bailey, M., Jenkins, A., Legg, T., Otto, F. E. L., Massey, N., Schaller, N., Slingo, J., and Allen, M. R. (2014). Potential influences on the united kingdom’s floods of winter 2013/14. *Nature Climate Change*, 4:769 – 777. <https://doi.org/10.1038/nclimate2314>.
- Iorgulescu, I. and Beven, K. J. (2004). Nonparametric direct mapping of rainfall-runoff relationships: An alternative approach to data analysis and modeling? *Water Resources Research*, 40(8). <https://doi.org/10.1029/2004WR003094>.
- Jafarzadegan, K. and Merwade, V. (2017). A dem-based approach for large-scale floodplain mapping in ungauged watersheds. *Journal of Hydrology*, 550:650 – 662. <https://doi.org/10.1016/j.jhydrol.2017.04.053>.
- Jafarzadegan, K. and Merwade, V. (2019). Probabilistic floodplain mapping using hand-based statistical approach. *Geomorphology*, 324:48 – 61. <https://doi.org/10.1016/j.geomorph.2018.09.024>.
- Jafarzadegan, K., Merwade, V., and Saksena, S. (2018). A geomorphic approach to 100-year floodplain mapping for the conterminous united states. *Journal of Hydrology*, 561:43 – 58. <https://doi.org/10.1016/j.jhydrol.2018.03.061>.
- Jalayer, F., De Risi, R., De Paola, F., Giugni, M., Manfredi, G., Gasparini, P., Topa, M. E., Yonas, N., Yeshitela, K., Nebebe, A., Cavan, G., Lindley, S., Printz, A., and Renner, F. (2014). Probabilistic gis-based method for delineation of urban flooding risk hotspots. *Natural Hazards*, 73(2):975 – 1001. <https://doi.org/10.1007/s11069-014-1119-2>.
- Jarihani, A. A., Callow, J. N., McVicar, T. R., Niel, T. G. V., and Larsen, J. R. (2015). Satellite-derived digital elevation model (dem) selection, preparation and correction for hydrodynamic modelling in large, low-gradient and data-sparse catchments. *Journal of*

- Hydrology*, 524:489 – 506. <https://doi.org/10.1016/j.jhydrol.2015.02.049>.
- Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J. C. J. H., Mechler, R., Botzen, W. J. W., Bouwer, L. M., Pflug, G., Rojas, R., and Ward, P. J. (2014). Increasing stress on disaster-risk finance due to large floods. *Nature Climate Change*, 4:264 – 268. <https://doi.org/10.1038/nclimate2124>.
- Kernkamp, H. W. J., Van Dam, A., Stelling, G. S., and de Goede, E. D. (2011). Efficient scheme for the shallow water equations on unstructured grids with application to the continental shelf. *Ocean Dynamics*, 61(8):1175 – 1188. <https://doi.org/10.1007/s10236-011-0423-6>.
- Kirkland, E. J. (2010). *Bilinear Interpolation*, pages 261 – 263. Springer US, Boston, MA. [https://doi.org/10.1007/978-1-4419-6533-2\\_12](https://doi.org/10.1007/978-1-4419-6533-2_12).
- Kleinen, T. and Petschel-Held, G. (2007). Integrated assessment of changes in flooding probabilities due to climate change. *Climatic Change*, 81(3):283 – 312. <https://doi.org/10.1007/s10584-006-9159-6>.
- Knijff, J. M. V. D., Younis, J., and Roo, A. P. J. D. (2010). Lisflood: a gis-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24(2):189 – 212. <https://doi.org/10.1080/13658810802549154>.
- Kochilakis, G., Poursanidis, D., Chrysoulakis, N., Varella, V., Kotroni, V., Eftychidis, G., Lagouvardos, K., Papathanasiou, C., Karavokyros, G., Aivazoglou, M., Makropoulos, C., and Mimikou, M. (2016). A web based dss for the management of floods and wildfires (fire) in urban and periurban areas. *Environmental Modelling & Software*, 86:111 – 115. <https://doi.org/10.1016/j.envsoft.2016.09.016>.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World map of the köppen-geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3):259 – 263. <http://dx.doi.org/10.1127/0941-2948/2006/0130>.
- Kron, W., Eichner, J., and Kundzewicz, Z. W. (2019). Reduction of flood risk in europe – reflections from a reinsurance perspective. *Journal of Hydrology*, 576:197 – 209. <https://doi.org/10.1016/j.jhydrol.2019.06.050>.
- Krysanova, V., Müller-Wohlfeil, D.-I., and Becker, A. (1998). Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecological Modelling*, 106(2):261 – 289. [https://doi.org/10.1016/S0304-3800\(97\)00204-4](https://doi.org/10.1016/S0304-3800(97)00204-4).
- Kubat, M., Holte, R. C., and Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, 30(2):195 – 215. <https://doi.org/10.1023/A:1007452223027>.

- Kumar, P., Verdin, K. L., and Greenlee, S. K. (2000). Basin level statistical properties of topographic index for north america. *Advances in Water Resources*, 23(6):571 – 578. [https://doi.org/10.1016/S0309-1708\(99\)00049-4](https://doi.org/10.1016/S0309-1708(99)00049-4).
- Lehner, B., Verdin, K., and Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, 89(10):93 – 94. <https://doi.org/10.1029/2008EO100001>.
- Leopold, L. B. and Maddock Jr., T. (1953). The hydraulic geometry of stream channels and some physiographic implications. Technical report, U.S. Geological Survey, Washington, D.C. <https://doi.org/10.3133/pp252>.
- Leskens, J., Brugnach, M., Hoekstra, A., and Schuurmans, W. (2014). Why are decisions in flood disaster management so poorly supported by information from flood models? *Environmental Modelling & Software*, 53:53 – 61. <https://doi.org/10.1016/j.envsoft.2013.11.003>.
- Levizzani, V. and Cattani, E. (2019). Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. *Remote Sensing*, 11(19). <https://doi.org/10.3390/rs11192301>.
- Luino, F. (1999). The flood and landslide event of november 4-6 1994 in piedmont region (northwestern italy): Causes and related effects in tanaro valley. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*, 24(2):123 – 129. [https://doi.org/10.1016/S1464-1895\(99\)00007-1](https://doi.org/10.1016/S1464-1895(99)00007-1).
- Ma, Y., Wu, H., Wang, L., Huang, B., Ranjan, R., Zomaya, A., and Jie, W. (2015). Remote sensing big data computing: Challenges and opportunities. *Future Generation Computer Systems*, 51:47 – 60. <https://doi.org/10.1016/j.future.2014.10.029>.
- Manfreda, S., Leo, M. D., and Sole, A. (2011). Detection of flood-prone areas using digital elevation models. *Journal of Hydrologic Engineering*, 16(10):781 – 790. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000367](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000367).
- Manfreda, S., Nardi, F., Samela, C., Grimaldi, S., Taramasso, A. C., Roth, G., and Sole, A. (2014). Investigation on the use of geomorphic approaches for the delineation of flood prone areas. *Journal of Hydrology*, 517:863 – 876. <https://doi.org/10.1016/j.jhydrol.2014.06.009>.
- Manfreda, S. and Samela, C. (2019). A digital elevation model based method for a rapid estimation of flood inundation depth. *Journal of Flood Risk Management*, 12(S1):e12541. <https://doi.org/10.1111/jfr3.12541>.
- Manfreda, S., Samela, C., Gioia, A., Consoli, G. G., Iacobellis, V., Giuzio, L., Cantisani, A., and Sole, A. (2015). Flood-prone areas assessment using linear binary classifiers based on flood maps obtained from 1d and 2d hydraulic models. *Natural Hazards*, 79(2):735 – 754. <https://doi.org/10.1007/s11069-015-1869-5>.

- Manning, R. (1891). *On the flow of water in open channels and pipes*, volume 20. Transactions of the Institution of Civil Engineers of Ireland. 161 - 207.
- Manzato, A. (2007). A note on the maximum peirce skill score. *Weather and Forecasting*, 22(5):1148 – 1154. <https://doi.org/10.1175/WAF1041.1>.
- Matthews, B. (1975). Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 405(2):442 – 451. [https://doi.org/10.1016/0005-2795\(75\)90109-9](https://doi.org/10.1016/0005-2795(75)90109-9).
- McGrath, H., Bourgon, J.-F., Proulx-Bourque, J.-S., Nastev, M., and Abo El Ezz, A. (2018). A comparison of simplified conceptual models for rapid web-based flood inundation mapping. *Natural Hazards*, 93(2):905 – 920. <https://doi.org/10.1007/s11069-018-3331-y>.
- Merz, R. and Blöschl, G. (2005). Flood frequency regionalisation – spatial proximity vs. catchment attributes. *Journal of Hydrology*, 302(1):283 – 306. <https://doi.org/10.1016/j.jhydrol.2004.07.018>.
- Milly, P. C. D., Wetherald, R. T., Dunne, K. A., and Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, 415(6871):514 – 517. <https://doi.org/10.1038/415514a>.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill series in computer science. McGraw-Hill, 1 edition.
- Montgomery, D. R. and Dietrich, W. E. (1988). Where do channels begin? *Nature*, 336(6196):232 – 234. <https://doi.org/10.1038/336232a0>.
- Montgomery, D. R. and Dietrich, W. E. (1989). Source areas, drainage density, and channel initiation. *Water Resources Research*, 25(8):1907 – 1918. <https://doi.org/10.1029/WR025i008p01907>.
- Morrison, R. R., Bray, E., Nardi, F., Annis, A., and Dong, Q. (2018). Spatial relationships of levees and wetland systems within floodplains of the wabash basin, usa. *JAWRA Journal of the American Water Resources Association*, 54(4):934 – 948. <https://doi.org/10.1111/1752-1688.12652>.
- Muis, S., Güneralp, B., Jongman, B., Aerts, J. C., and Ward, P. J. (2015). Flood risk and adaptation strategies under climate change and urban expansion: A probabilistic analysis using global data. *Science of The Total Environment*, 538:445 – 457. <https://doi.org/10.1016/j.scitotenv.2015.08.068>.
- Nardi, F., Annis, A., Di Baldassarre, G., Vivoni, E. R., and Grimaldi, S. (2019). Gfplain250m, a global high-resolution dataset of earth’s floodplains. *Scientific Data*, 6(1):180309. <https://doi.org/10.1038/sdata.2018.309>.

- Nardi, F., Biscarini, C., Di Francesco, S., Manciola, P., and Ubertini, L. (2013). Comparing a large-scale dem-based floodplain delineation algorithm with standard flood maps: the tiber river basin case study. *Irrigation and Drainage*, 62(S2):11–19. <https://doi.org/10.1002/ird.1818>.
- Nardi, F., Morrison, R. R., Annis, A., and Grantham, T. E. (2018). Hydrologic scaling for hydrogeomorphic floodplain mapping: Insights into human-induced floodplain disconnectivity. *River Research and Applications*, 34(7):675 – 685. <https://doi.org/10.1002/rra.3296>.
- Nardi, F., Vivoni, E. R., and Grimaldi, S. (2006). Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. *Water Resources Research*, 42(9). <https://doi.org/10.1029/2005WR004155>.
- Neal, J., Schumann, G., and Bates, P. (2012). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, 48(11). <https://doi.org/10.1029/2012WR012514>.
- Neal, J., Villanueva, I., Wright, N., Willis, T., Fewtrell, T., and Bates, P. (2011). How much physical complexity is needed to model flood inundation? *Hydrological Processes*, 26(15):2264 – 2282. <https://doi.org/10.1002/hyp.8339>.
- Nobre, A., Cuartas, L., Hodnett, M., Rennó, C., Rodrigues, G., Silveira, A., Waterloo, M., and Saleska, S. (2011). Height above the nearest drainage – a hydrologically relevant new terrain model. *Journal of Hydrology*, 404(1):13 – 29. <https://doi.org/10.1016/j.jhydrol.2011.03.051>.
- Nobre, A. D., Cuartas, L. A., Momo, M. R., Severo, D. L., Pinheiro, A., and Nobre, C. A. (2016). Hand contour: a new proxy predictor of inundation extent. *Hydrological Processes*, 30(2):320 – 333. <https://doi.org/10.1002/hyp.10581>.
- Noh, S. J., Lee, J.-H., Lee, S., Kawaike, K., and Seo, D.-J. (2018). Hyper-resolution 1d-2d urban flood modelling using lidar data and hybrid parallelization. *Environmental Modelling & Software*, 103:131 – 145. <https://doi.org/10.1016/j.envsoft.2018.02.008>.
- O’Brien, J., Julien, P., and Fullerton, W. (1993). Two-dimensional water flood and mud-flow simulation. *Journal of Hydraulic Engineering*, 119(2):244 – 261.
- O’Loughlin, F., Neal, J., Schumann, G., Beighley, E., and Bates, P. (2020). A lisflood-fp hydraulic model of the middle reach of the congo. *Journal of Hydrology*, 580:124203. <https://doi.org/10.1016/j.jhydrol.2019.124203>.
- Olthof, I. and Tolszczuk-Leclerc, S. (2018). Comparing landsat and radarsat for current and historical dynamic flood mapping. *Remote Sensing*, 10(5). <https://doi.org/10.3390/rs10050780>.

- Pan, F., Stieglitz, M., and McKane, R. B. (2012). An algorithm for treating flat areas and depressions in digital elevation models using linear interpolation. *Water Resources Research*, 48(6). <http://doi.org/10.1029/2011WR010735>.
- Pappenberger, F., Cloke, H. L., Parker, D. J., Wetterhall, F., Richardson, D. S., and Thielen, J. (2015). The monetary benefit of early flood warnings in europe. *Environmental Science & Policy*, 51:278 – 291. <https://doi.org/10.1016/j.envsci.2015.04.016>.
- Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. L. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade. *Hydrology and Earth System Sciences*, 16(11):4143 – 4156. <https://doi.org/10.5194/hess-16-4143-2012>.
- Pappenberger, F., Iorgulescu, I., and Beven, K. J. (2006). Sensitivity analysis based on regional splits and regression trees (sars-rt). *Environmental Modelling & Software*, 21(7):976 – 990. <https://doi.org/10.1016/j.envsoft.2005.04.010>.
- Pearson, K. (1900). X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157 – 175. <https://doi.org/10.1080/14786440009463897>.
- Peirce, C. S. (1884). The numerical measure of the success of predictions. *Science*, ns-4(93):453 – 454. <https://doi.org/10.1126/science.ns-4.93.453-a>.
- Prieto, C., Le Vine, N., Kavetski, D., García, E., and Medina, R. (2019). Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests. *Water Resources Research*, 55(5):4364 – 4392. <https://doi.org/10.1029/2018WR023254>.
- Rahmati, O. and Pourghasemi, H. R. (2017). Identification of critical flood prone areas in data-scarce and ungauged regions: A comparison of three data mining models. *Water Resources Management*, 31(5):1473 – 1487. <https://doi.org/10.1007/s11269-017-1589-6>.
- Rajib, A., Liu, Z., Merwade, V., Tavakoly, A. A., and Follum, M. L. (2020). Towards a large-scale locally relevant flood inundation modeling framework using swat and lisflood-fp. *Journal of Hydrology*, 581:124406. <https://doi.org/10.1016/j.jhydrol.2019.124406>.
- Rajib, M. A., Merwade, V., Kim, I. L., Zhao, L., Song, C., and Zhe, S. (2016). Swatshare – a web platform for collaborative research and education through online sharing, simulation and visualization of swat models. *Environmental Modelling & Software*, 75:498 – 512. <https://doi.org/10.1016/j.envsoft.2015.10.032>.

- Rathjens, H., Bieger, K., Chaubey, I., Arnold, J. G., Allen, P. M., Srinivasan, R., Bosch, D. D., and Volk, M. (2016). Delineating floodplain and upland areas for hydrologic models: a comparison of methods. *Hydrological Processes*, 30(23):4367 – 4383. <https://doi.org/10.1002/hyp.10918>.
- Rebolho, C., Andréassian, V., and Le Moine, N. (2018). Inundation mapping based on reach-scale effective geometry. *Hydrology and Earth System Sciences*, 22(11):5967 – 5985. <https://doi.org/10.5194/hess-22-5967-2018>.
- Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J. (2008). Hand, a new terrain descriptor using srtm-dem: Mapping terra-firme rainforest environments in amazonia. *Remote Sensing of Environment*, 112(9):3469 – 3481. <https://doi.org/10.1016/j.rse.2008.03.018>.
- Risi, R. D., Jalayer, F., and Paola, F. D. (2015). Meso-scale hazard zoning of potentially flood prone areas. *Journal of Hydrology*, 527:316 – 325. <https://doi.org/10.1016/j.jhydrol.2015.04.070>.
- Risi, R. D., Jalayer, F., Paola, F. D., and Manfredi, G. (2014). *A Semi-Probabilistic GIS-Based Method for Meso-Scale Flood Hazard Zonation*, pages 1445 – 1454. American Society of Civil Engineers. <https://doi.org/10.1061/9780784413609.145>.
- Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Tridon, D. B., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M., Wessel, B., Krieger, G., Zink, M., and Moreira, A. (2017). Generation and performance assessment of the global tandem-x digital elevation model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132:119 – 139. <https://doi.org/10.1016/j.isprsjprs.2017.08.008>.
- Rodda, H. J. E. (2005). The development and application of a flood risk model for the czech republic. *Natural Hazards*, 36(1):207 – 220. <https://doi.org/10.1007/s11069-004-4549-4>.
- Rodriguez, E., Morris, C., Belz, J., Chapin, E., Martin, J., Daffer, W., and Hensley, S. (2005). An assessment of the srtm topographic products. Technical report, Jet Propulsion Laboratory, Pasadena, California. Technical Report JPL D-31639.
- Salas, J. D., Fu, C., and Rajagopalan, B. (2011). Long-range forecasting of colorado streamflows based on hydrologic, atmospheric, and oceanic data. *Journal of Hydrologic Engineering*, 16(6):508 – 520. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000343](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000343).
- Samela, C., Albano, R., Sole, A., and Manfreda, S. (2018). A gis tool for cost-effective delineation of flood-prone areas. *Computers, Environment and Urban Systems*, 70:43 – 52. <https://doi.org/10.1016/j.compenvurbsys.2018.01.013>.
- Samela, C., Manfreda, S., Paola, F. D., Giugni, M., Sole, A., and Fiorentino, M. (2016). Dem-based approaches for the delineation of flood-prone areas in an ungauged basin in

- africa. *Journal of Hydrologic Engineering*, 21(2):06015010. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001272](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001272).
- Samela, C., Troy, T. J., and Manfreda, S. (2017). Geomorphic classifiers for flood-prone areas delineation for data-scarce environments. *Advances in Water Resources*, 102:13 – 28. <https://doi.org/10.1016/j.advwatres.2017.01.007>.
- Sampson, C. C., Fewtrell, T. J., O’Loughlin, F., Pappenberger, F., Bates, P. B., Freer, J. E., and Cloke, H. L. (2014). The impact of uncertain precipitation data on insurance loss estimates using a flood catastrophe model. *Hydrology and Earth System Sciences*, 18(6):2305 – 2324. <https://doi.org/10.5194/hess-18-2305-2014>.
- Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., and Freer, J. E. (2015). A high-resolution global flood hazard model. *Water Resources Research*, 51(9):7358 – 7381. <https://doi.org/10.1002/2015WR016954>.
- Sanders, B. F. (2007). Evaluation of on-line dems for flood inundation modeling. *Advances in Water Resources*, 30(8):1831 – 1843. <https://doi.org/10.1016/j.advwatres.2007.02.005>.
- Sassi, M., Nicotina, L., Pall, P., Stone, D., Hilberts, A., Wehner, M., and Jewson, S. (2019). Impact of climate change on european winter and summer flood losses. *Advances in Water Resources*, 129:165 – 177. <https://doi.org/10.1016/j.advwatres.2019.05.014>.
- Scheel, K., Morrison, R. R., Annis, A., and Nardi, F. (2019). Understanding the large-scale influence of levees on floodplain connectivity using a hydrogeomorphic approach. *JAWRA Journal of the American Water Resources Association*, 55(2):413 – 429. <https://doi.org/10.1111/1752-1688.12717>.
- Schröter, K., Kunz, M., Elmer, F., Mühr, B., and Merz, B. (2015). What made the june 2013 flood in germany an exceptional event? a hydro-meteorological evaluation. *Hydrology and Earth System Sciences*, 19(1):309 – 327. <https://www.hydrol-earth-syst-sci.net/19/309/2015/>.
- Schumann, G. J.-P., Andreadis, K. M., and Bates, P. D. (2014a). Downscaling coarse grid hydrodynamic model simulations over large domains. *Journal of Hydrology*, 508:289 – 298. <https://doi.org/10.1016/j.jhydrol.2013.08.051>.
- Schumann, G. J.-P. and Moller, D. K. (2015). Microwave remote sensing of flood inundation. *Physics and Chemistry of the Earth, Parts A/B/C*, 83 - 84:84 – 95. <https://doi.org/10.1016/j.pce.2015.05.002>.
- Schumann, G. J.-P., Vernieuwe, H., De Baets, B., and Verhoest, N. E. C. (2014b). Roc-based calibration of flood inundation models. *Hydrological Processes*, 28(22):5495 – 5502. <https://doi.org/10.1002/hyp.10019>.

- Seibert, J. and Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, 16(9):3315 – 3325. <https://doi.org/10.5194/hess-16-3315-2012>.
- Seibert, M., Merz, B., and Apel, H. (2017). Seasonal forecasting of hydrological drought in the limpopo basin: a comparison of statistical methods. *Hydrology and Earth System Sciences*, 21(3):1611 – 1629. <https://doi.org/10.5194/hess-21-1611-2017>.
- Şen, O. and Kahya, E. (2017). Determination of flood risk: A case study in the rainiest city of turkey. *Environmental Modelling & Software*, 93:296 – 309. <https://doi.org/10.1016/j.envsoft.2017.03.030>.
- Shortridge, J. E., Guikema, S. D., and Zaitchik, B. F. (2016). Machine learning methods for empirical streamflow simulation: a comparison of model accuracy, interpretability, uncertainty in seasonal watersheds. *Hydrology and Earth System Sciences*, 20(7):2611 – 2628. <https://doi.org/10.5194/hess-20-2611-2016>.
- Snelder, T. H., Datry, T., Lamouroux, N., Larned, S. T., Sauquet, E., Pella, H., and Catalogne, C. (2013). Regionalization of patterns of flow intermittence from gauging station records. *Hydrology and Earth System Sciences*, 17(7):2685 – 2699. <https://doi.org/10.5194/hess-17-2685-2013>.
- Sørensen, R. and Seibert, J. (2007). Effects of dem resolution on the calculation of topographical indices: Twi and its components. *Journal of Hydrology*, 347(1):79 – 89. <https://doi.org/10.1016/j.jhydrol.2007.09.001>.
- Speckhann, G. A., Chaffe, P. L. B., Goerl, R. F., de Abreu, J. J., and Flores, J. A. A. (2018). Flood hazard mapping in southern brazil: a combination of flow frequency analysis and the hand model. *Hydrological Sciences Journal*, 63(1):87 – 100. <https://doi.org/10.1080/02626667.2017.1409896>.
- Stephenson, D. B. (2000). Use of the "odds ratio" for diagnosing forecast skill. *Weather and Forecasting*, 15(2):221 – 232. [https://doi.org/10.1175/1520-0434\(2000\)015<0221:UOTORF>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0221:UOTORF>2.0.CO;2).
- Strahler, A. (1964). *Handbook of Applied Hydrology*. McGraw-Hill, New York.
- Swain, N. R., Christensen, S. D., Snow, A. D., Dolder, H., Espinoza-Dávalos, G., Goharian, E., Jones, N. L., Nelson, E. J., Ames, D. P., and Burian, S. J. (2016). A new open source platform for lowering the barrier for environmental web app development. *Environmental Modelling & Software*, 85:11 – 26. <https://doi.org/10.1016/j.envsoft.2016.08.003>.
- Swets, J. A. (1973). The relative operating characteristic in psychology. *Science*, 182(4116):990 – 1000. <https://doi.org/10.1126/science.182.4116.990>.

- Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D., Oimoen, M., Zhang, Z., Danielson, J., Krieger, T., Curtis, B., Haase, J., Abrams, M., Crippen, R., and Carabajal, C. (2011). Aster global digital elevation model version 2 – summary of validation results. Technical report, Compiled by Dave Meyer on behalf of the NASA Land Processes Distributed Active Archive Center and the Joint Japan-US ASTER Science Team.
- Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., and Iwamoto, H. (2014). Precise global dem generation by alos prism. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-4:71 – 76. <https://doi.org/10.5194/isprsannals-II-4-71-2014>.
- Tang, Z., Zhang, H., Yi, S., and Xiao, Y. (2018). Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis. *Journal of Hydrology*, 558:144 – 158. <https://doi.org/10.1016/j.jhydrol.2018.01.033>.
- Tarboton, D. G., Bras, R. L., and Rodriguez-Iturbe, I. (1992). A physical basis for drainage density. *Geomorphology*, 5(1):59 – 76. [https://doi.org/10.1016/0169-555X\(92\)90058-V](https://doi.org/10.1016/0169-555X(92)90058-V).
- Tarquini, S., Isola, I., Favalli, M., Mazzarini, F., Bisson, M., Pareschi, M. T., and Boschi, E. (2007). Tinality/01: a new triangular irregular network of italy. *Annals of Geophysics*, 50(3):407 – 425. <https://doi.org/10.4401/ag-4424>.
- Tarquini, S., Vinci, S., Favalli, M., Doumaz, F., Fornaciai, A., and Nannipieri, L. (2012). Release of a 10-m-resolution dem for the italian territory: Comparison with global-coverage dems and anaglyph-mode exploration via the web. *Computers & Geosciences*, 38(1):168 – 170. <https://doi.org/10.1016/j.cageo.2011.04.018>.
- Tavares da Costa, R., Manfreda, S., Luzzi, V., Samela, C., Mazzoli, P., Castellarin, A., and Bagli, S. (2019a). A web application for hydrogeomorphic flood hazard mapping. *Environmental Modelling & Software*, 118:172 – 186. <https://doi.org/10.1016/j.envsoft.2019.04.010>.
- Tavares da Costa, R., Mazzoli, P., and Bagli, S. (2019b). Limitations posed by free dems in watershed studies: The case of river tanaro in italy. *Frontiers in Earth Science*, 7:141. <https://doi.org/10.3389/feart.2019.00141>.
- Tehrany, M. S., Pradhan, B., and Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (dt) and a novel ensemble bivariate and multivariate statistical models in gis. *Journal of Hydrology*, 504:69 – 79. <https://doi.org/10.1016/j.jhydrol.2013.09.034>.
- Tehrany, M. S., Pradhan, B., and Jebur, M. N. (2014). Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in gis. *Journal of Hydrology*, 512:332 – 343. <https://doi.org/10.1016/j.jhydrol.2014.03.008>.

- van Beek, L. and Bierkens, M. (2008). The global hydrological model pcr-globwb: Conceptualization, parameterization and verification. Technical report, Utrecht University. <http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf>.
- Verdin, K. and Verdin, J. (1999). A topological system for delineation and codification of the earth's river basins. *Journal of Hydrology*, 218(1):1 – 12. [https://doi.org/10.1016/S0022-1694\(99\)00011-6](https://doi.org/10.1016/S0022-1694(99)00011-6).
- Wallis, J. R. (1965). Multivariate statistical methods in hydrology - a comparison using data of known functional relationship. *Water Resources Research*, 1(4):447 – 461. <https://doi.org/10.1029/WR001i004p00447>.
- Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., and Bai, X. (2015). Flood hazard risk assessment model based on random forest. *Journal of Hydrology*, 527:1130 – 1141. <https://doi.org/10.1016/j.jhydrol.2015.06.008>.
- Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P., De Groeve, T., Muis, S., de Perez, E. C., Rudari, R., Trigg, M. A., and Winsemius, H. C. (2015). Usefulness and limitations of global flood risk models. *Nature Climate Change*, 5:712 – 715. <https://doi.org/10.1038/nclimate2742>.
- Westerhoff, R. S., Kleuskens, M. P. H., Winsemius, H. C., Huizinga, H. J., Brakenridge, G. R., and Bishop, C. (2013). Automated global water mapping based on wide-swath orbital synthetic-aperture radar. *Hydrology and Earth System Sciences*, 17(2):651 – 663. <https://doi.org/10.5194/hess-17-651-2013>.
- Williams, J. (1969). Flood routing with variable travel time or variable storage coefficients. *Transactions of the American Society of Agricultural and Biological Engineers*, 12(1):0100 – 0103. <https://doi.org/10.13031/2013.38772>.
- Williams, W. A., Jensen, M. E., Winne, J. C., and Redmond, R. L. (2000). An automated technique for delineating and characterizing valley-bottom settings. *Environmental Monitoring and Assessment*, 64(1):105 – 114. <https://doi.org/10.1023/A:1006471427421>.
- Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., and Erickson, T. A. (2017). Validation of a 30 m resolution flood hazard model of the conterminous united states. *Water Resources Research*, 53(9):7968 – 7986. <https://doi.org/10.1002/2017WR020917>.
- Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., and Bouwman, A. (2013). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences*, 17(5):1871 – 1892. <https://doi.org/10.5194/hess-17-1871-2013>.
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., and Bates, P. D. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, 44(11):5844 – 5853. <https://doi.org/10.1002/2017GL072874>.

- Yamazaki, D., Kanae, S., Kim, H., and Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, 47(4). <https://doi.org/10.1029/2010WR009726>.
- Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3(1):32 – 35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3).
- Zanardo, S., Nicotina, L., Hilberts, A. G. J., and Jewson, S. P. (2019). Modulation of economic losses from european floods by the north atlantic oscillation. *Geophysical Research Letters*, 46(5):2563 – 2572. <https://doi.org/10.1029/2019GL081956>.
- Zhang, W. and Montgomery, D. R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water Resources Research*, 30(4):1019 – 1028. <https://doi.org/10.1029/93WR03553>.
- Zhao, G., Pang, B., Xu, Z., Peng, D., and Xu, L. (2019). Assessment of urban flood susceptibility using semi-supervised machine learning model. *Science of The Total Environment*, 659:940 – 949. <https://doi.org/10.1016/j.scitotenv.2018.12.217>.
- Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., and Passalacqua, P. (2018a). Geoflood: Large-scale flood inundation mapping based on high-resolution terrain analysis. *Water Resources Research*, 54(12):10013 – 10033. <https://doi.org/10.1029/2018WR023457>.
- Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., and Passalacqua, P. (2018b). River channel geometry and rating curve estimation using height above the nearest drainage. *JAWRA Journal of the American Water Resources Association*, 54(4):785 – 806. <https://doi.org/10.1111/1752-1688.12661>.



# Acknowledgements

This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement No. 676027.

I hereby express my gratitude to my Supervisors, Dr. Stefano Bagli and Prof. Attilio Castellarin, for having selected me to embark on the System-Risk experience and for the opportunity to complete this doctoral thesis. To the German Research Centre for Geosciences for putting the System-Risk project together and for successfully coordinating it. To all the persons I met, befriended, learned from and worked with along the way, my sincere thanks.

I would like to thank GECOsistema Srl., the University of Bologna, Risk Management Solutions Inc., the River Po Basin Authority and the Centre for Water Resource Systems at the Technical University of Vienna for sharing their data, for hosting me and for the thoughtful technical discussions held with their teams. Every single experience mattered and was well worth the effort.

I acknowledge the European Union's Horizon 2020 research and innovation program for the Marie Skłodowska-Curie System-Risk ETN (<https://www.system-risk.eu/>), David Tarboton (Utah State University) for TauDEM utilities, Leo Breiman and Adele Cutler (Utah State University) for the Random Forest code, the Python Software Foundation for the Python programming language, the Open Source Geospatial Foundation for the Geospatial Data Abstraction Library, the R Foundation for Statistical Computing for the R language and R environment, the E-OBS dataset from the EU-FP6 project UERRA (<http://www.uerra.eu/>), the Copernicus Climate Change Service and the data providers in the ECA&D project (<https://www.ecad.eu/>).

This doctoral thesis reflects only the authors' views and the European Union is not liable for any use that may be made of the information contained therein.

