HOKKAIDO UNIVERSITY

Title	Information-theoretic portfolio decision model for optimal flood management
Author(s)	Convertino, Matteo; Annis, Antonio; Nardi, Fernando
Citation	Environmental modelling and software, 119, 258-274 https://doi.org/10.1016/j.envsoft.2019.06.013
Issue Date	2019-09
Doc URL	https://hdl.handle.net/2115/82843
Rights	© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/
Rights(URL)	https://creativecommons.org/licenses/by-nc-nd/4.0/
Type	journal article
File Information	ConvertinoNardi_Portfolio_EnvSoft_May2019_notracks2.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Information-theoretic Portfolio Decision Model for Optimal Flood Management

Matteo Convertino[a,b,c], **Antonio Annis**[d,e], **Fernando Nardi**[d]

[a] Nexus Group, Graduate School of Information Science and Technology, Hokkaido University, Sapporo, JP

[b] Gi-CoRE Station for Big Data & Cybersecurity, Hokkaido University, Sapporo, JP

[c] Department of Electronics and Information Engineering, Hokkaido University, Sapporo, JP

[d] Water Resources Research and Documentation Center (WARREDOC), University for Foreigners of Perugia, Perugia, Italy

[e] Department of Civil and Environmental Engineering (DICEA), University of Florence, Florence, Italy

June 27, 2019

Keywords: river basin management; floods; systemic risk; MaxEnt*; portfolio decision model; MCDA*

Corresponding author: * M. Convertino, Kita 14, Nishi 9, Kita-ku, 11-11, Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido, Japan, 060-0814, phone: +81 011-706-6491, email: matteo@ist.hokudai.ac.jp

Abstract

The increasing impact of flooding urges more effective flood management strategies to guarantee sustainable ecosystem development. Recent catastrophes underline the importance of avoiding local ~~point-scale~~ flood management, but characterizing large scale basin wide approaches for systemic flood risk management. Here we introduce an information-theoretic Portfolio Decision Model (iPDM) for the optimization of a systemic ecosystem value at the basin scale by evaluating all potential flood risk mitigation plans. ~~Different flood management plans, represented by spatial combinations of flood control structures (FCS), are explored.~~ iPDM calculates the ecosystem value predicted by all feasible combinations of flood control structures (FCS) considering environmental, social and economical asset criteria. A multi-criteria decision analytical model evaluates the benefits of all FCS portfolios at the basin scale weighted by stakeholder preferences for assets' criteria as ecosystem services. The risk model is based on a maximum entropy model (MaxEnt) that predicts the flood susceptibility, the risk of floods based on the exceedance probability distribution, and its most important drivers. Information theoretic global sensitivity and uncertainty analysis is used to select the simplest and most accurate model based on a flood return period. A stochastic optimization algorithm optimizes the ecosystem value constrained to the budget available and ~~here expressed as the number of FCS~~ provides Pareto frontiers of optimal FCS plans for any budget level. Pareto optimal solutions maximize FCS diversity and minimize the criticality of floods manifested by the scaling exponent of the Pareto distribution of flood size that links management and hydrogeomorphological patterns. The proposed model is tested on the 17,000 km^2 Tiber river basin in Italy. iPDM allows stakeholders to identify optimal FCS plans in river basins for a comprehensive evaluation of flood effects under future ecosystem trajectories.

1 Introduction

1.1 Flood Protection and Distribution of Floods

Flooding is a major natural hazard affecting some 520 million people every year, claiming the lives of about 25,000 worldwide and causing global economic losses between USD 50 and 60 billion annually. As a result, it is essential that we seek to manage the risk of flooding in an effective and appropriate way. Also in consideration of climate change, the assessment of changing dynamics of flood impacts is a public concern beyond many other climate change effects on populations (Liu et al., 2019). The latest assessment of the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2014) on observed changes and future projections of floods was provided in chapter 3 of the IPCC special report on extremes, often called the "IPCC SREX report1". A summary on projected flooding in this report stated that, overall there is low confidence in projections of changes in fluvial floods. This confidence is low due to limited and uncertain evidence of ecosystem drivers and because of the complexity of these causes at the regional changes that is not necessarily addressed quantitatively speaking. Thus, a model that considers these heterogeneities — such as hydrological, land use/cover change, and socio-political dynamical complexity — and their uncertainty is needed for effective decision making aimed to systemic flood protection.

As for climate change, in general, the major changes are expected into a varied return period (rather than average magnitude) of the events considered, including floods. In particular the return period (or frequency, equivalently) of extreme phenomena is varying because the underlying conditions influencing climate are changing or have been changing for long time in a diverse way. Most people and the literature talks about changes in magnitude of events that is not necessarily affected primarily. Rather, the accumulation of unusually close events — even of small or medium magnitude — can bring to an intensification of extreme events resulting in extreme floods.

Examples are about what happened recently in Europe, and in Rome specifically, where hydrologic changes driven by human induced land development have largely enhanced flooding (Nardi et al., 2015). Extreme flood events of such kind are like the ones in Paris in the spring of 2016 and the 2018 European floods (Paprotny et al., 2018). Floods and other hydrology-related extremes caused economic losses of Euro 453 billion between 1980 and 2017, claiming the lives of more than 115,000 people across Europe (see Zanardo et al. (2019) for recent estimates of European floods). Similar patterns of floods and losses are observed in the Americas (see Quinn et al. (2019) for USA), Asia and Australia. Infrastructures are typically designed for withstanding extreme events with a predetermined return period (such as for

the peak runoff of 100 years, i.e. the runoff expected every 100 years on average) (Volpi et al., 2015). However, recurrence periods are decreasing considering climatic changes; therefore, many extreme floods are occurring with higher frequency than in the past due to the inability of flood control infrastructure to contain them. This is also associated to poorly planned and maintained urban water management systems (i.e., storm water and sewer drainage), originally designed for a 10-30 years frequency scenario, that are now inefficiently managing even frequently expected rainfall events. Yet, in general the probability distribution of events is changing — with a tendency toward a longer power-law distribution of events — rather than an increase in the mean or the variance of these events (Bak et al., 1988; Fonstad and Marcus, 2003; Van De Wiel and Coulthard, 2010; Convertino et al., 2012). This trend has also a tendency to increase the criticality of flood events by decreasing the power-law exponent regulating the flood size; this implies increased persistency of flood phenomena (Bak et al., 1988; Convertino et al., 2012). For this motivation, that leads to high interdependencies between geographical areas beyond administrative boundaries, the development of adaptive and systemic flood management models is needed in order to avoid local fragmentation of decisions that implies large sustained flood risks at the basin scale.

1.2 Current Flood Management Models

Various models have been used to characterize the risk of extreme events to large urban and more localized infrastructure systems (Merz et al., 2010) but very few models have addressed multiscale and multiasset risk assessment problems. Little has been done for ecological systems and far less considering societal impacts on populations affected by floods. Current models include probabilistic risk analysis and decision analysis considered as separate models (Meyer et al., 2009; Merz et al., 2010). Probabilistic models including Bayesian analysis have been developed for situations in which there are sparse data of historical frequencies of extreme events such as extreme floods. Decision analysis has been used in the performance assessment of infrastructure policies and investments, for instance see Meyer et al. (2009) for a multi-criteria decision analytical application for flood risk. There is a considerable literature that addresses the costs and benefits of investment in protecting infrastructure systems from the threat of extreme events. In particular, Bier et al. (1999) discusses several approaches for assessing and managing the risk of extremes, which has inspired to focus on the exceedance probabilities of catastrophic flood losses. These methods can be categorized as "asset management" methods where the assets considered were water infrastructure, and the flood risk was calculated a posteriori based on the likelihood of a flood to occur and asset conditions

(USACE, 2012). Other hydraulics/hydrologic methods have been used to characterized the very local flooding risk without much consideration of infrastructure criticality and feasible multiasset spatial arrangements (Nardi et al., 2015). All these studies are somewhat "risk-centered" in a unidimensional way. In other words, these models do not provide indications of optimal management solutions, integration of risk and decision science models considering socio-environmental and economic outcomes, nor the incorporation of spatial heterogeneities in a spatially explicit way. Lastly, these methods are considering exceedance probability values rather than their distribution in a self-organized criticality perspective (Bak et al., 1988; Fonstad and Marcus, 2003) and therefore missing the systemic probabilistic characterization of the patterns analyzed.

Recently, efforts have explored the applicability of financial analysis models to flood protection. In particular, modern portfolio theory (Markowitz, 1952) was explored for the diversification of flood infrastructure investments. Financial portfolio analysis has been widely used in project selection and choice management (Tasca et al., 2017). There is significant work applying such analysis to the financial risk associated to extreme events (Valverde and Andrews, 2006; Haldane and May, 2011a; Park et al., 2012; Convertino and Valverde, 2013), including the use of metrics like value-at-risk ("VaR") and conditional value-at-risk ("Co-VaR"). Results of these models describe that Markowitz-type diversification (Markowitz, 1952) can generate the highest expected economic return under an acceptable systemic risk threshold for a desired flood protection. Systemic risk is a risk that considers non-linear functional and structural interdependencies among systems' components leading to cascading effects (Haldane and May, 2011b; Battiston and Caldarelli, 2013; Helbing, 2013; Tasca and Battiston, 2014; Battiston et al., 2016; Burkholz et al., 2016). This is a signature of the self-organized criticality of complex systems (Bak et al., 1988). In flood management these interdependent features are for instance stakeholders, type and distribution of flood control structures, river network and landscape features which collectively determine non-linear flood events. As an example, applications to The Netherlands are significative (Aerts et al., 2008) considering their pressing flood security issues. In Aerts et al. (2008), through a systematic combination of four different types of flood protection assets (excluding the "no action" alternative), portfolios of asset types were constructed to reduce the variance of expected losses. The variance reduction was greater with more types of assets compared with portfolios that only contained one asset. These past efforts leave open the question about how such diversification might influence other population endpoints beyond the traditional economic metrics of performance of flood protection. Additionally, the use of exceedance probabilities (epdfs) of extreme losses definable via a multicriteria function rather than just using mean-variance

portfolio approaches is missing.

We believe that there are significant simplifications to using average metrics or mean-variance optimization of flood risk — that often lacks of a spatial component characterization such as in (Zhou et al., 2012) — for designing flood protection plans considering extreme events. First, no study has yet suggested that variance minimization is appropriate as a performance objective for multicriteria projects. In particular, variance may not always be an appropriate component of the objective function in engineering decision making where return period design should be prevalent, for example for coastal protection and flood investment. For engineering design, a mean-variance approach does not represent extreme events because variance of events — whose probability distribution can be a power-law distribution with infinite variance (for the power-law exponent of the flood size) — does not characterize enough the severity of these events. Decision makers are concerned with low-probability catastrophic events along with more frequently occurring and less severe events. Rather than mean-variance optimization, exceedance probabilities and return periods at several levels of systemic loss should be used as metrics of the risk associated to extreme events in hydrologic and ecological engineering.

Yet, this paper will explore how diversification, evaluated using a multicriteria metric, of flood protection alternatives might be effective to reduce the exceedance probabilities of extreme systemic losses under several scenarios of floods. Here we focus on flood management scenarios but any scenario can be considered, for instance related to climate change. It is hypothesized that there are conditions in which diversification could reduce the risk of extreme losses and be effective to address the nonsystematic portion of risks associated with climate change. In other words, good adaptive management can overcome unpredictable extrinsic fluctuations (Convertino et al., 2014b). This risk is non-systematic because extrinsic and potentially unpredictable versus the risk associated to management. The risk model determines a probability distribution of quantified flood losses, in which average, covariances and any other statistical moment for several types of joint flood protection assets are the output of numerical simulations. The spectrum of return periods is selected to measure the risk reduction of selected flood protection portfolios among all feasible ones, including the status quo. Yet, the epdf can be considered as a macro-indicator of flood protection of a basin where neutrally balanced bottom-up basin heterogeneities and top-down systemic risk control are considered by the portfolio model.

1.3 How is the Portfolio Approach Different?

The systemic perspective of the information-theoretic Portfolio Decision Model (iPDM), considering spatial, temporal (aggregated or explicit) and decisional complexities, is the main conceptual difference with respect to traditional models of flood management invoking static and discrete cost-benefit analyses (see for example Zhou et al. (2012)). The underlying idea of the portfolio model is to integrate all information in the least complex way and formulate a decision that is comprehensive in terms of local heterogeneities and outcomes considered for any possible management plan (Convertino and Valverde, 2013). The model leverages any information coming from any data or model, such as hydrological and ecological impact models, and guides further scenarios by exploring all management plans. Figure 1 illustrates the computational steps and information integration of iPDM. The truly novel aspect of the paper is the integration of all models into iPDM via information theoretic analytics. Additionally, iPDM can also guide data collection and define site importance in providing data considering the analysis of information entropy, complexity, sensitivity and diversity. More technically, the main innovative and integrated models of the paper and iPDM are the following.

- **Flood Susceptibility Prediction and Flood Delineation**. The flood susceptibility is determined by the MaxEnt model (Phillips et al., 2006; Phillips and Miroslav, 2008; Convertino et al., 2013b) based on observed flood occurrences dependent on FCS plans, land-cover/-use, and hydrogeomorphological variables. The integrated probability distribution function that is associated to a spatial distribution of floods, constitutes the basis for the calculation of the Systemic Flood Risk (SFR) that also considers criticality and efficiency of FCS; the exceedance probability distribution defines the arrangement and likelihood of floods. The flood delineation method establishes a threshold on the MaxEnt flood susceptibility for calibration based on historical floods. Different thresholds allow stakeholders the selection of a flood size distribution with a desired maximum return period.

- **Multi-Criteria Decision Analytical Model and Pareto Optimization for Flood Plan Design**. The inclusion of a multi criteria decision analytical model (MCDA) for assessing the Systemic Flood Risk (SFR) to minimize; a systemic value to optimize can be formulated by considering social, environmental and economic outcomes related to floods. An optimization model is used to determine Pareto frontiers that define optimal flood control plans for the whole basin. The optimal plan is defined as the one that optimize the systemic value after exploring all potential combinations of flood

infrastructure portfolios; this optimal SFR is not necessarily the absolute minimum if other socio-economical and environmental criteria need to be maximized. Thus, the non-linearity between susceptibility, risk and value is highlighted by the model. If time is considered explicitly, FCS strategies should be selected as optimal combination of single time-point FCS plans in an adaptive management framework (Convertino et al., 2014b).

- **Information theoretic Global Sensitivity and Uncertainty Analyses**. The adoption of an information-theoretic Global Sensitivity and Uncertainty Analyses (iGSUA) (Lüdtke et al., 2008) to characterize all input factors (model parameters and state variables) probabilistically and propagate their uncertainty to outputs (predicted floods). iGSUA also allows to select the most accurate and least complex predictive model within the possible MaxEnt models (where accuracy is about the distance between predictions and data, and complexity is about the number and amount of predictors used in the model) in a broad Occam's razor purview (Grimm et al., 2005) and for a selected return period of floods. GSUA also allows a better probabilistic characterization of the epdf of floods that can be used as reliable macroecological indicator of systemic hydrologic risk.

2 Materials and Methods

2.1 The Tiber River Basin

The Tiber basin (Figure 3 and 4) is the largest river basin in Central Italy (second largest river in Italy), draining towards the Mediterranean Sea and covering a land area of 17,500 km^2 approximately. Various studies have investigated the hydrological behavior of the basin (Fiseha et al., 2013). These studies were conducted considering the hydraulic and hydrologic processes taking place in the basin ecosystem, among others, flood forecasting (see for instance Calenda et al. (2000); Calvo and Savi (2009); Napolitano et al. (2010)), flood routing (Franchini et al., 2011) and soil moisture assessment (Singh, 2009) on selected parts of the Tiber River (Piccolroaz et al., 2015, 2016; Di Baldassarre et al., 2017). The mathematical models used in these studies were of different nature: from process based hydrogeomorphological models to machine learning models such as artificial neural networks. Much attention has been given to the issue of flood risk as the Tiber River passes through many historical places in the region's urban areas including the historical core of the city of Rome (Calenda et al., 2005). Rome is subjected to a non-negligible risk of inundation when extreme floods propagate along the Tiber river (Natale and Savi, 2007). Severe floods, characterized by a

return period slightly lower than 200 years (although 100 years floods are also very severe (Volpi et al., 2015)), can in fact overtop both the left and right river banks and inundate the northern outskirts of Rome, while extreme events, characterized by a return period of about 1,000 years, can submerge large parts of the monumental center of Rome. The overtopping of river banks is predicted to occur for the 200 year design flood event near ponte Milvio, the oldest roman bridge in Rome, in a short reach (200-300 m long) where the levees elevation is slightly depressed causing the potential overtopping of the retaining walls "muraglioni". Extreme events, characterized by a return period of about 1,000 years, can overtop the levees in different segments submerging large parts of the monumental center of Rome. The tributary network of the Tiber is also characterized by elevated flood risk for the significant impact of land use changes related to transportation networks. Channel-road intersections impact the water flow capacity and conveyance during extreme precipitation events causing frequent inundation conditions (Nardi et al., 2015). Despite the large heterogeneities of the basin, most of the flooding studies in Rome are associated to single hydrologic extreme responses rather than focusing on the whole river system flood patterns; yet, the latter are not mapped to the underlying physiographic, climate, and flood control infrastructure features that are potentially responsible for their occurrence. Yet, the Tiber basin provides an ideal setting for testing the proposed portfolio management model considering the magnitude and diversity of flood-related forcing factors, socio-ecological actors and the increasing hydrologic change and urbanization trends that characterize such large basin.

2.2 Stochastic Portfolio Decision Model

In this paper the portfolio model in Convertino and Valverde (2013) is extended to the context of floods in river systems and used via information-theoretic models such as MaxEnt and iGSUA. Figure 1 shows the conceptual diagram of the interlinked models whose connection is related via input-output relationships. The Maximum Entropy Model (MaxEnt) (Dudik et al., 2007; Phillips and Miroslav, 2008; Elith et al., 2010; Merow et al., 2013) based on the maximum entropy principle (Jaynes, 1957, 1988) was previously used in Convertino et al. (2013b) and Convertino et al. (2014a) for the dynamical prediction of landslides and avian species habitat in large scale ecosystems at a variety of different resolutions. Recently the model has been extended to microbial ecosystems (Li and Convertino, 2019). Furthermore, here iPDM is enriched by a fully stochastic characterization of all state variables and model parameters in an information-theoretic framework such as in Lüdtke et al. (2008) (iGSUA is used for this purpose). Such framework describes the variability of the predicted patterns in a

pattern-oriented modeling purview (Grimm et al., 2005) — where in this case the exceedance probability distribution of flooded areas and their geographical distribution are the patterns to predict — driven by the variability of local and systemic river network features (i.e. for instance FCS efficiency and drainage connections). Systemic features are always exerted by the river network that defines the spatial dependencies among different areas resulting into a interconnected flood dynamics even in absence of atmospherical dependencies among rainfall patterns (see e.g. Steinhaeuser et al. (2012) for climate teleconnections). These geomorphic features are well known to influence many river dependent processes such as soil saturation, sediment transport, hydrochory and the spread of passive and reactive agents (Rinaldo et al., 2018). Agents are particles in a modeling sense whose intrinsic and extrinsic dynamics (the latter more deeply dependent on the environment than the former) determines the patterns of interest. In the case of flood patterns local environmental conditions and systemic networks (drainage, flood control structure interdependency, and green infrastructure networks) determine flood patterns. In a more general decision theoretic sense "agents" are stakeholders that have local and global effects on the emergence of floods via their actions and decision making at multiple scales of river basin ecosystems.

The iPDM modeling approach is conceptually and theoretically equivalent to the neutral metacommunity model of Muneepeerakul et al. (2008) and Convertino (2011) that is essentially describing patterns formation of ecosystems resulting from the collective behavior of species ("agents", generally speaking) driven by local and systemic factors. In this context we adopt the same philosophy, but we disregard the dynamics over time. Yet, we focus on the local "generating" function — in this case the drainage area (that is the scaling function governing the runoff; in fact under homogenous conditions the drainage area is the precise surrogate of the runoff) rather than the habitat suitability as in Muneepeerakul et al. (2008) and Convertino (2011) — that is inferred via MaxEnt and the network is considered as a static network. However, network cumulated features are considered, i.e. the drainage area is dependent on the network on each site. Further analyses about the evolution of the network via information-theoretic models for forecasting the dynamics of runoff are worth exploring and undergoing. However, flood management is more concerned about incorporating long term dynamics; therefore, this approach, that neglects the evolution of runoff at short time scales, is very oriented toward long term management and design of water infrastructure in river basins vs. real time forecasting of floods.

2.3 2D MaxEnt Flooding Risk Model

In a iPDM context a payoff function is defined as the balance between total benefits (or systemic value) — dependent on the reduction of systemic flooded area causing socio-environmental damages — and the cost of flood control structures (FCS). This is in analogy to other systemic management models, such as the enhanced adaptive management models in Convertino et al. (2014b), where the difference in payoff is evaluated for different scenarios altering local and systemic conditions and as a function of data quality measured as Value of Information (i.e. the difference in payoff for different data quality states).

More technically, the payoff is a multicriteria function that considers multiple asset arrangements and balances benefits and FCS costs (see Eq. 4). Because cost is considered as a constraint in a Pareto optimization, we prefer to talk about systemic value rather than payoff. Flooded areas are here identified by a layer of occurrences and other numerically determined socio-ecohydrological layers dependent on their importance for flood prediction (land cover/use, flood control structures, drainage area and the topographic index which are socio-ecological, infrastructural and geomorphic variables), which are "rasterized" information provided to the model (Figure 4). The dynamics of the systemic value, whether time is explicitly included in the model, can be described for example by a Langevin-based model (when values are considered) or by a Fokker-Planck model if the probability distribution is considered. Li and Convertino (2019) shows a temporal application of the model where pdfs are propagating macroecological values in a microbial ecosystem and different ecosystem states are detected; the same can be done for flood where transitions occur during the time varying hydrological dynamics of a basin. In this dynamics, criticality occurs when the propagating effects are maximized, that is when cascading effects such as floods or information transmission in general, are maximized. Note that criticality is something desired in biological systems vs. something not desired in the case of floods in an anthropocentric purview. Here, the dynamics is not taken into account explicitly and the systemic value is predicted as a static average pattern dependent on the average flood pattern associated to a return period of 200 years. The 200 years floods were obtained from validated floodplain zoning model grounded on geomorphic theories for large scale identification of floods using digital topographic datasets (Nardi et al., 2006; Nardi et al., 2013, 2015). Floodplain maps are available for the Tiber river basin as well as at the global scale (see Nardi et al. (2019) and `https://github.com/fnardi/GFPLAIN`) using the Global FloodPLAIN (GFPLAIN) geomorphic model (Manfreda et al., 2014; Morrison et al., 2018; Nardi et al., 2018; Annis et al., 2019; Nardi et al., 2019; Scheel et al., 2019).

The probability of a flood at time t (note that here we adopt a macroecological char-

acterization of the probability (Convertino et al., 2013b), i.e. an average susceptibility of a community, but a temporal assessment is possible for time varying conditions) ~~(that for example enter the dynamics in Eq. ??)~~ $P(y = 1|\mathbf{c}(t))$ is calculated as:

$$P(y = 1|\mathbf{c}(t)) = f(\mathbf{c}(t)) \exp^{\eta(\mathbf{c}(t))} \frac{P(y = 1)}{f(\mathbf{c}(t))} , \quad (1)$$

where $f(\mathbf{c})$ is the probability density of covariates $\mathbf{c}$ across the river basin, and $\eta(\mathbf{c}) = \alpha + \rho h(\mathbf{c})$. α is a normalizing constant that ensures that $f_1(c)$ integrates to one (f_1 being the probability density function (pdf) of the flooded area occurrences), and ρ is the constant (Lagrangian multiplier) of the MaxEnt features $h(\mathbf{c})$ (Phillips et al., 2006; Phillips and Miroslav, 2008; Elith et al., 2010; Merow et al., 2013). The Lagrangian multiplier that multiplies all environmental features finds the optimal trade-off between model complexity (defined as the number of environmental variables used as predictors) and model accuracy (that is the distance between predictions and data) (Convertino et al., 2013b; Merow et al., 2013). The set of parameters is identified by minimizing the prediction error between the observed and modeled floods. Features are transformations of the covariates in the covariate space and this allows a faster and more precise computation rather than operating in the geographical space (Elith et al., 2010; Convertino et al., 2013b). Further details of the model can be found in (Convertino et al., 2013b) that was the first application of the model in hydro-geomorphology. The susceptibility of floods in Eq. 1 is calculated for each pixel defined by the Digital Elevation Model (Fig. 4). This susceptibility can be later averaged at the subbasin scale or any scale of interest for the MCDA and Pareto optimization model calculations.

The probability distribution of a flood is recalculated for each FCS plan. FCS plans are matrices that alter the probability of occurrence of a flood. These matrices are populated with the FCS efficiency that is the factor determining the local influence of a FCS to decrease the likelihood of a flood to occur (Eq. 1). This efficiency can be estimated from literature or inferred ex post the calculation of floods dependent only on one FCS typology at a time. Thus, the covariate c and FCS plans allow one to predict the multiplex flood susceptibility where multiple networked layers are interdependent from each other. The FCS matrix is a multilayer matrix in case a diverse set of FCS are used. The collectivity of FCS in terms of their diversity and spatial positioning determines the FSC effectiveness and systemic risk. Section 2.6 describes in details how the systemic risk is calculated.

Predictions of floods may be dependent on the number of observations. MaxEnt predicts the flood susceptibility (i.e. $\sim$ the probability to have a site flooded) of each pixel "nearby" the centers of mass of an observed flood (Fig. 2) as a function of the relationships inferred between the environmental variables and the centers of mass of randomly selected observed

floods ("background points") at the basin scale. For this reason we used a procedure similar to coarse-graining — widely adopted in hydrology (see for example (Convertino et al., 2007)) — to understand how flood predictions are dependent on on the number of observed floods considered as inputs of the model (Figure S2). Note that this is analogous to a coarse graining analysis since "zooming out" preserves the average features of the variables considered but reduces the computational needs related to the data in input. Yet, it is a data reduction approach that aims to preserve the predictive power of the model. Figure S3 shows the probability of flood susceptibility conditional to the selected predictors (also called "response curves" of MaxEnt) and divided by likelihood of susceptibility in the random FCS scenario, in the predictive scale-invariant region of the model that provides the highest information content (Servadio and Convertino, 2018; Li and Convertino, 2019).

2.4 Flood Delineation and Fractal Characterization

The flood delineation method is based on a threshold assigned on the MaxEnt flood susceptibility and calibrated on observed or predicted floods with a certain return period. Such method allows easily the selection of the flood size distribution with a desired maximum return period. Nearest neighbor pixels, according to the von Neumann criteria (or 4-neighborhood), whose flood susceptibility is higher than the calibrated flood susceptibility threshold are categorized as flooded areas (Fig. 2). The validation of the MaxEnt model is based on reproducing the floods with the specified return period associated to a certain susceptibility threshold. Additional explanation of this delineation method, shown in Fig. 2, is contained in Convertino et al. (2013b) in the context of landslides. The fractal distribution of observed and predicted flooded areas is used as a macroecological indicator of hydrological risk. This distribution is captured with a varying degree of accuracy by both the the box counting method (see Supplementary Information) and the "Korcak's law" method that define flooded areas as patches. These areas define the epdf shown in Figure 5. The box-counting method (Figure S1) overestimates the fractal dimension with respect to the Korcak's law method as shown in Convertino et al. (2013c) and does not precisely characterize finite-size effects along the end of the tail of the distribution. Thus, we consider the Korcak's law method to determine the more likely shape and power-law slope of the exceedance distribution of flooded areas that is related to the flood pattern fractal dimension. The probability of exceedance of the flood areas, that can be considered as the equivalent of the patch size in an ecological context (known in literature as "Korcak's law" (Korcak, 1940; Nikora et al., 1999; Convertino et al., 2013c)) is:

$$P(S \geq s) = c\, s^{-\epsilon}\, F\left(\frac{s}{s_c}\right) , \tag{2}$$

where s is the flood area, c is a constant, F is a homogeneity function that depends on a characteristic flood size s_c (dependent on the basin size and hydrological dynamics), and $\epsilon = D/2$ is the scaling exponent of the Pareto epdf considering self-affinity of floods (Korcak, 1940; Mandelbrot, 1982; Convertino et al., 2013c). D is the fractal dimension of floods that is derivable from both the epdf (based on the 1D vector of flood sizes) and the spatial pattern of floods via different methods. Note that the relationship $\epsilon = D/2$ is only valid within the critical regime when system's aggregates are considered as patches as according the Korcak's law (Convertino et al., 2013c). Otherwise for general self-affine systems $\epsilon = 2 - D_f/3 - D_f$ for which the exponent (defining the criticality of the system) is larger the higher the fractal dimension. The true fractal dimension is related to the Hurst exponent H via $D_f = 2 - H$, that manifests the persistency of stochastic processes (Convertino et al., 2013c). The more fractal a process is the higher the fluctuations that occur and the lower H. The probability of exceedance exhibits a power-law behavior, although exponential behavior is observed for small flooded areas. A transition in pdf/epdf is observed for different FCS plans (either different heterogeneous FCS plans and homogeneous FCS plans for one FCS typology at a time). The transition is related to both the spatial distribution of FCS which determines the plan effectiveness, and the FCS local efficiency (see Section 2.6). The probability distribution of the flood-size is used to validate the MaxEnt model and the box-counting estimates on the real occurrences. The fit of the predicted distribution of patches is performed using a Maximum Likelihood Estimation technique (MLE) (Convertino et al., 2013c) . In our case study for the current flood distribution ("status quo") $\epsilon = D/2 = 0.84$ (the absolute value). D is 1.83 from the box counting method, 2 from the space filling counts, and 1.67 from the stability analysis and the Korcak's law. The last value is considered for calculating ϵ because it is calculated directly on the mapped floods and it is verified to be the most reliable estimate (Convertino et al., 2013c) .

2.5 Stochastic MCDA and Pareto Optimization Model

iPDM adopts several foundational concepts drawn from modern portfolio theory (MPT), including the concept of Pareto optimality (Pareto, 1971). Whereas MPT focuses primarily on the task of developing computationally tractable means by which to allocate resources among various expenditures and investments over time, our primary focus is on developing a

holistic, integrated bottom-up/top-down model to flood management, with the use of machine learning models informed by macrophysical processes. One may think that these models, such as MaxEnt, can ~~obviously~~ be replaced by detailed flood mapping hydraulics models at fine resolution; however, it is much less obvious about how to use the information coming from these models solely, for guiding optimal management plans. Here, the purpose is to expose iPDM for the first time and to demonstrate how this model can be used for regional planning and detailed flood mapping with selected information that is also potentially provided by other models whose predictions (e.g. hydraulic mapping of observed floods with a certain return period) can be used to calibrated and validate iPDM and in particular MaxEnt. One of the submodels of iPDM is the MCDA model that defines the necessary information for characterizing ecosystem values.

Specifically, the MCDA model used for evaluating each flood control structure (FCS) plan for all assets, is a linear Multi Attribute Value Theory (MAVT) model (Convertino et al., 2013a). In this paper assets are the areas (or communities) impacted by floods, thus characterized by a flood size; a more detailed characterization of assets can be done if additional information is available. The size of this information is unlimited potentially but only certain information is relevant for the predicted patterns (Servadio and Convertino, 2018). Here a FCS plan is constituted by the current or all other possible sets of FCS attributed to each community of the river network considering their efficiency for flood protection. As we discuss later in more detail our community scale is the subbasin scale.

The MCDA model ranks FCS by scoring them considering a value determined as a linear combination of criteria values and weights. The MCDA model calculates the local value of FCS plans, $V_{i,j}(FCS)$, with respect to asset j in area (community) i. Tables S1-S3 shows all input factors of the MCDA model. In a broader MCDA framework it is better to assess a systemic *value* rather than a systemic *risk* to emphasize the resilience approach that should be considered in landscape management. This approach identifies the most valuable portfolios of alternatives that optimize the systemic value. The systemic value does not focus only on the maximization of the hazard-dependent inverse risk but on the maximization of all ecosystem services given all feasible hazard-controlled scenarios and their social and economic outcomes. We assume implicitly the dependence on time of the predicted values via MaxEnt that characterizes the average pattern of flood areas within a management period defined by the return time. This approach is a macroecological approach that is widely used in ecology and population sciences for characterizing average or persistent patterns. The MCDA value or systemic value (generally speaking the inverse of the systemic risk (SFR) if social and economical values are not considered), that is ultimately the risk of assets at the basin scale

given a FCS plan, is calculated as:

$$V^*_{i,j}(FCS) \;=\; (1-\nu_j(FCS))f_{m(j)}FCS_{m(j),i}V_{i,j}(FCS) \tag{3}$$

where $V_{i,j}(FCS) = \sum_k w_{j,k}\; x_{j,k,i}(FCS)$. $x_{j,k,i}$ are the values of criteria k (in the MCDA model) for asset j and area i. In our case study all criteria are minimized so the optimization problem is actually a minimization problem; however, the MCDA model allows for multiple criteria to minimize and maximize. Yet, the value vs. risk problem formulation is more general than the one presented. The weights $w_{j,k}$ are stakeholder preferences for the asset j and criteria k that are typically independent of the area i. Note that here we use the portfolio considering areas as assets (but more explicitly the portfolio problem is a FCS problem because a FCS can be assigned to each area and we are only characterizing the effects of FCS in each area) but more precisely multiple assets j can exist in the same or different areas i. These assets, protected by a set of FCS "m", can be human populations, infrastructure, and other human resources (or ecosystem services broadly defined) such as agricultural fields with a certain crop yield, and vulnerable species in ecosystems (see for example Convertino and Valverde (2013) for an ecological application). In this paper we just consider the flood area as a criteria for assessing flood risk. However, the systemic risk is not perfectly equivalent to the flood area — both locally (i.e., at the subbasin scale) and at the basin scale — because the former considers criticality and efficiency of FCS (related to their life-cycle status and ability to withstand floods in isolation as a function of their structure, respectively; see Tables S2 and S3) as well as the non-linear predicted flood area determined by the combination of FCS. Here criticality refers to the FCS vulnerability but a more complex multicriteria function can consider a more comprehensive vulnerability for all communities considered. The inverse of the systemic risk is used to calculate the systemic value that ideally considers other ecosystem services. The inverse of the flood area is the local FCS effectiveness that incorporates local FCS criticality and efficiency; in a purely mathematical definition, the convolution of systemic risk and value is equal to one if no other ecosystem services are considered in the formulation of value. The monetary quantification of the risk (or value) is certainly interesting and necessary for practical management. However, here we focus only on proposing the model able to perform such comprehensive assessment of ecosystem services (Table S4 shows for instance economic and social services as an example), also because of the unavailability of other data; thus, we leave the economic evaluation of floods to subsequent applications of the model.

The local value $V_{i,j}(FCS)$ of a FCS plan is adjusted by the probability of success (of flood protection) given by the vulnerability of each asset under a FCS plan ($\nu_j(FCS)$) (as later

described this is more precisely the criticality of a FCS in this case study) and the effectiveness of the FCS plan $f_{m(j)}FCS_{m(j),i}$ (considering all FCS arrangements) that is dependent on the intrinsic efficiency of all FCS locally ($f_{m(j)}$, potentially asset-specific). For example, a flood prone area protects well urban areas but may not protect agricultural areas) and the location with respect to all other FCS on the basin (i.e., $FCS_{m(j),i}$). At the basin scale, the effect of all efficiencies — estimated from FCS' flood protection potential gathered from data — and the spatial location of FCS, constitutes the overall effectiveness of the portfolio set at the basin scale. Here the (unknown) vulnerability is considered related to the structural criticality of each FCS evaluated by experts (Tables S1 and S2); the more critical a FCS is, the higher $\nu_j(FCS)$ and the lower its potential to protect against floods. This structural criticality can be related to infrastructure age, maintenance, and other factors.

The systemic value at the basin scale of the whole FCS plan, $V_T(FCS)$, is calculated as a Euclidian distance where the components of the distance are the local value of areas, dependent on a FCS plan, weighted by stakeholder preferences $w_{i,j}$ for areas and assets. The systemic value that is ultimately the value of a population given a FCS plan is:

$$V_T(FCS) \;=\; \sqrt{\sum_{i=1}^{L}\sum_{j=1}^{J}[V^*_{i,j}(FCS)\; w_{i,j}]^2} \tag{4}$$

Table S4 shows a snapshot of an ideal MCDA model where ecosystem services based on the MaxEnt inferred floods can be assessed for any area and asset. In this paper we calculate V^* as the inverse of the flood size to maximize. The size of floods everywhere in the basin is dependent on the FCS plan that determines the pdf of floods; the pdf is the pattern on which the model is calibrated and validated. The total number of portfolio plans (or FCS plans) is M^L, that is the number of possible FCS for each area raised to a power exponent equal to the number of areas. Each area can be a pixel in the basin dictated by the Digital Elevation Model but any different delineation of communities can be decided. The baseline model that reproduces the 200 years return period pdf of floods (the 200-years floods are 10% larger than the 100 years ones on average) is considered as the "status quo" for the current flood management approach. After exploring all FCS plans the Pareto optimal set is identified as the one that maximizes $V_T(FCS)$ constrained or unconstrained to the available resources (Convertino and Valverde, 2013). Here we explore all potential resources that are identified by the number of potential FCS, one for each community of the drainage network.

The optimization model of iPDM is a linear mixed-integer optimization algorithm that explores all possible combinations of FCS alternatives (for each asset/community one alternative at a time), with their expected value and cost at the local scale. The maximization of

the systemic value (or alternatively the minimization of SFR if economic and social endpoints are not considered) is performed with and without the constraint of the available resources B. These resources are the one available for flood management and here they are taken proportional to the number of FCS considering the unavailability of resource information. Note that in this context optimization coincides with maximization since all criteria are maximized. In principle the number of portfolio combinations is dictated by the number of channelized cells at the lowest scale (i.e., 11,321 for the Tiber basin) and the number of FCS (i.e. 5, including the no action alternative). Specifically, the number of portfolio plans is the product of the number of possible FCS for each community raised to a power exponent equal to the number of communities where each FCS is present. Due to the large number of streams where a FCS can be ideally placed, we decided, more reasonably, to consider the number of subbasins (i.e., 2308 for the extracted network with a threshold of 5 km2) as communities; thus, the total number of portfolio plans is 5^{2308}. Small scales of subbasins allows stakeholders to have a more granular control of floods but large scales may be more effective: this raises the question about which is the optimal scale for flood control; this question is however postponed to further studies. Table S1 reports the FCS considered in this case study and Table S3 reports other potential FCS such as expansion areas or flood prone areas (e.g. floodplains, riparian areas), vegetative buffer strips, and gauging stations that can be evaluated in basin flood management and flood related services. These FCS are green infrastructure that also support other ecosystem services such as promotion of species habitat and dispersal, filtering of chemicals and sediments, and pathogen spreading abatement. In the portfolio constrained case, the cost of the FCS plan, $C(FCS) = \sum_{m=1,M} \sum_{i=1,L} \sum_{j=1,J} C_m(FCS_{m(j),i})$, cannot exceed the resources B. In the case of the Pareto optimization unconstrained by budget, if $V_T(FCS_1) \geq V_T(FCS_2)$ and $C(FCS_1) < C(FCS_2)$, then the portfolio solution FCS_1 dominates FCS_2. Thus, all FCS envisioned by FCS_1 are selected; those are the ones that shift the tail of the pdf of the flooded area to the left the most. In the budget constrained Pareto optimization, if $V_T(FCS_1) \geq V_T(FCS_2)$ then the portfolio combination FCS_1 dominates FCS_2. It can be easily proven that a payoff function $I = V_T(FCS) - C(FCS)$ to maximize can be used to obtain the same Pareto frontiers. Table S6 reports the inputs of the Pareto optimization model and a ideal output.

2.6 Metamodeling for MaxEnt Model Selection and Testing

Metamodeling is the art to detect the model with the highest information power for a stated objective (Saltelli et al., 2004; Gu et al., 2012; Servadio and Convertino, 2018). This is in line

with the information-theoretic principle of optimal model deign (Servadio and Convertino, 2018). The optimal set of input factors is defined as the one that simultaneously minimizes the error $\hat{e}$ in reproducing the observed floods, and maximizes model accuracy that can for instance measured by the "area under the curve" (AUC) (Zweig and Campbell, 1993). The AUC is solely dependent on the reported flood occurrences, while the error is dependent on the delineation of predicted landslides based on the predicted flood susceptibility. Therefore both must be considered when assessing the model and this is one of our strength with respect to classical applications of MaxEnt. The AUC is evaluated considering the receiver operator characteristic curve (ROC) (Phillips et al., 2006; Phillips and Miroslav, 2008). The ROC is a graphical plot of the sensitivity, or true positives (i.e., the percentage of predicted flood occurrences that match the observed ones), versus the "complement function", that is the specificity or false positives (i.e., the percentage of predicted flood occurrences which do not match any observation). Note that this comparison is done at the pixel scale before any flood delineation. The AUC compares the likelihood that a random flood occurrence site has a higher predicted value in the model than a random site where no flood occurs. Thus, the higher the AUC, the better the prediction. In the jackknife test, that is used as the metamodeling design test, each variable is excluded in turn from the MaxEnt model run (on instantiation), and a model is created with the remaining variables. Then, a model is created using each variable in isolation. In addition, a run is created using all available variables. When only one variable is used in the prediction, the AUC measures the absolute importance of the variable in predicting flood patterns (yet, this is $\propto$ the first order sensitivity index (see Section 2.7)). The difference between the AUC with all variables and the AUC for the single-variable prediction is a proxy of the sum of interactions (or interdependencies) between the variable considered in isolation and all others (yet, this is $\propto$ the second order sensitivity index) (see Section 2.7). Thus, the jackknife can be used also as a global sensitivity method for the model (Saltelli et al., 2004) beyond being a model deign test to define which variable to incorporate. After the calibration on the historical flood pattern (for which the error $\hat{e}$ is minimized) it is possible to retain the variables for which the jackknife test shows an AUC greater than 0.5, that is a standard threshold in establishing the importance of the environmental variables considered. AUC greater than 0.5 is meaningful that any prediction is beyond a random pattern prediction where flood are randomly predicted over space. The error $\hat{e}$ is observed to be proportional to $1 - N_p/N_h$, where N_p is the predicted number of flooded pixels over the total number of historical flooded pixels N_h. The calculation of $1 - N_p/N_h$ is conditional to the flood occurrences. Thus, this calculation is conditional to the delineation of floods (Fig. 2).

2.7 Global Sensitivity and Uncertainty Analyses

In a variance decomposition framework (Saltelli et al., 2004; Servadio and Convertino, 2018) the variance of the systemic value, the payoff (i.e. the systemic value minus the FCS plan cost), or of any other selected function (e.g., the flood area) is attributed to the intrinsic local variability of socio-environmental determinants and systemic network determinants. This formulation is considering space explicitly but the variance of the systemic value can lump together spatially defined variables and the variability can be attributed to the intrinsic variability of single variables and their interactions (Servadio and Convertino, 2018; Li and Convertino, 2019).

This variance-based approach has been adopted in Convertino et al. (2013b) for the analysis of fundamental determinants of landslide patterns. In an information theoretic approach, we use a broader application of GSUA via the consideration of the entropy of the payoff that is defined as the necessary and sufficient information contained in its probability distribution and derived from underlying variables. A fully probabilistic information-theoretic GSUA that considers all probability distribution functions (pdfs) of variables and their entropy as in Lüdtke et al. (2008) and Servadio and Convertino (2018); Li and Convertino (2019) more recently is adopted. ~~If the payoff is known, predictors do not need to be considered explicitly.~~ In this context we aim to determine the relative importance of variables used in predicting flood susceptibility. Considering the payoff, its entropy can be seen as the sum of the intrinsic variability conditional to all other areas or variables ($MI(Y_i, Y_j)$, i.e. the mutual information) and the entropy ($H(Y(\mathbf{c}))$) that is how much the variability of the underlying determinants contribute to the total local variability. Thus, when finding the total entropy of the payoff, the information balance equation that defines the total entropy is given by sum of the Shannon entropies of all input variables (lumped or not over space) as considered alone (for the variability of the payoff) as well as the sum of their Transfer Entropies that is assessing the directional variable interdependence (see Li and Convertino (2019) for a biological application of iGSUA). The two quantities assess the variability of the payoff considering local and systemic features.

Thus, the total entropy of the flood susceptibility Y can be written as

$$H(Y) \approx \sum_i H(x_i) + \sum_i \sum_{j \neq i} |TE_i(x_i, x_j)| + \sigma(Y) \tag{5}$$

where x_i denote the $i-s$ variables (covariate c in Eq. 1) that contribute to the payoff Y. In Eq.

5 i denotes a variable rather than an area as in Eq. 4. In this equation, $H(\cdot)$ denotes Shannon entropy, and $TE(\cdot,\cdot)$ denotes Transfer Entropy from the first variable to the second variable (Li and Convertino, 2019). Eq. 5 represents a fundamental concept of Shannon's information theory and forms the general basis of sensitivity analyses (Servadio and Convertino, 2018; Li and Convertino, 2019). The sum of the absolute values of Transfer Entropies is a proxy of the Mutual Information ($MI_i = \sum p(X_j, X_i) \cdot \log_2 \left(\frac{p(X_j,X_i)}{p(X_j)\, p(X_i)} \right)$) (Servadio and Convertino, 2018), thus it considers the whole set of variable interdependencies. Here we do not consider the sign of variable interaction as in Servadio and Convertino (2018) because we focus on absolute flood predictability and because time dynamics of floods is not available explicitly. $\sigma(Y)$ is a noise term that captures the unexplained variability of Y related to variables not considered and discretization factors of the model (e.g. due to numerical solvers and pdf inference algorithms). Eq. 5 can also be extended in space.

Variable first order importance and interaction for reproducing the pdf of floods are then calculated as Mutual Information Indices (MII) (Lüdtke et al., 2008), that are defined by the mutual information normalized by the entropy of the output variable considering one variable or pairs of variables (Servadio and Convertino, 2018). These indices are, respectively: $s_i = \frac{MI_(X_i,Y)}{H(Y)}$ and $s_{ij} = \frac{MI_(X_i,X_j|Y)}{H(Y)}$, where X_i is any covariate and Y is the flooded area or payoff. In this study we use iGSUA for the flood susceptibility that is the fundamental predicted pattern influencing the payoff. The use of the transfer entropy (whether temporal data are available) can give further information about the directionality of the causality between variables (in a predictive sense of the model) in space and time, and the time-lag of their causality.

3 Results and Discussion

The main impact of FCS is assessed considering the flood size (Table S2-S5), although the model can handle any information about losses on assets via the MCDA component of the model (Table S4). The exceedance probability distribution function (epdf) of the predicted flood size is shown in Figure 5. The epdf is computed for the predicted flood size after thresholding the MaxEnt susceptibility to delineate floods. This epdf is computed considering the size of flooded areas defined as the ensemble of pixels with probability higher than 0.85. This flood susceptibility threshold on the flood susceptibility is suitable to reproduce the 200 years return time floods (Figs. 2 and 3). The calibration of the threshold constitutes the validation of the model on the reported floods (Fig. 3). Figure 2 shows the criteria and procedure

used for delineating floods in analogy to the landslide delineation method proposed in Convertino et al. (2013b). Data about the 200 years return time floods were constructed from scenario-based efforts using standard hydrological models and mapping procedures (Nardi et al., 2006; Nardi et al., 2013, 2015). In these hydrologic models, major rivers are integrated with hydrogeomorphic floodplain mapping for extending the flood-prone area information at the basin scale. Other methods for flood delineation have been used by the hydrology community and proposed the use of predictors that are also used in this study such as the topographic index (Manfreda et al., 2014; Manfreda et al., 2014; Samela et al., 2015, 2017). In flood risk assessment, calibration is only performed on low frequency hydrologic data since 30-50 years of observations are available for river flows; longer data series are available for precipitations only. Water infrastructure need to be designed for extreme events which may have never occurred; thus, hydraulics and hydrologic scenario modeling serve the purpose to determine those extreme events which are the inputs of iPDM.

Interestingly, we find that the scaling exponent of the flood size distribution ($\epsilon = D/2 = 0.84$) is about double than the exponent of the runoff or drainage area (i.e., ~ 0.43) (Rodríguez-Iturbe and Rinaldo, 2001); this may shed some light on the relationship between these two random variables (drainage area and flood size) that are physically dependent on the river network structure. The latter forces the propagation of runoff and floods along the scale-free drainage pattern and cross-river elongation features of the two phenomena (floods and runoff) may just be responsible for the different power-law exponents. A power-law exponent that is more than double than the one for floods has been observed for landslides (i.e., ~ 1.92; see Convertino et al. (2013c) and Convertino et al. (2013b)). Hydrologically triggered landslides are of course related to the drainage network and dependent on the same hydroclimatological extremes triggering floods; however, landslides occurs on hillslopes of basins and yet their scaling exponent is higher than floods which dictates their more limited size. The perimeter-area relationship for floods shows an exponent of about 0.75, that leaving aside numerical corrections related to the calculation methods, is in accordance to the estimate of the box-counting. The slope of the epdf of the flooded area provides an exponent equal to 0.78. These differences in the slope of the epdf of the flood areas are only related to the different methods used to calculate such distribution (see Methods section); however, all these methods are linked to each other as shown in Convertino et al. (2013c). The higher scaling exponent of floods shows the lower criticality (in a self-organized criticality perspective (Bak et al., 1988)) and persistency of floods vs. the runoff, as expected. This implies lower probability of floods to occur.

Considering the variability of flood patterns on FCS plans, the epdf is proposed as a

macroindicator of phase transitions in flood dynamics. The no-FCS scenario may correspond to the unaltered ecosystem scenario. This scenario shows an exponent ϵ that is lower than the basin regulated scenario, which implies a higher probability for the same flood size, on average. Moreover, the no-FCS flood size distribution has a much more exponential behavior for small floods that represent the random uncorrelated dynamics of frequent inundation events. It is interesting to see that the regulated basin (e.g. the random and current FCS plans) has a larger maximum flood than the unregulated basin. This is actually confirming the empirical evidence of many river basins where too much suboptimal flood control causes bigger floods. The randomly flood regulated basin scenario is quite worrisome because of the extent of the power-law regime that is much wider and with very large floods; in this case the exponential cutoff for finite size effects in the tail of the distribution highlights that the maximum flood is just limited by the size of the basin but larger floods may occur in very extreme hydrological conditions. In principle, the maximum flood size is weakly constrained by the basin size and more sensitive to extreme hydrometeorological forcings. All other epdfs different than the random epdf are truncated Pareto distribution without finite-size effects; therefore, the maximum flood size is much smaller than the one dictated by the size of the landscape. If we define $\alpha = 1 + \epsilon$ (where ϵ is the slope of the epdf of the flood size), then $\alpha = 1.83$ for the current FCS plan. It is known that when $1 < \alpha < 2$, the first moment of the distribution (the mean or average) is infinite, along with all the higher moments. When $2 < \alpha < 3$, such as for the Pareto optimal FCS plan, the first moment is finite (yet more controllable), but the second (the variance) and higher moments are infinite. Thus, some care needs to be placed in monitoring the optimal FCS plan despite the maximum predicted flood is much smaller than the current regulated state. For the Pareto optimal plan α is 2.2. For the drainage area $\alpha = 1.43$ since $\epsilon = 0.43$; thus, also the drainage area has infinite mean theoretically which highlight the fact that extremely large runoff phenomena occur but they are not necessarily determining floods.

In Figure 6 we show the Pareto optimal plans of FCS, along the Pareto frontier (in green), after considering all potential combinations of FCS tested on the river basin. All potential FCS are placed at the subbasin scale (Fig. S4). Table S7 shows an example of a Pareto optimal solution. Each optimal plan is located along the green Pareto optimal frontier generated by considering the number of FCS as the constraint function. This maximum number is in principle equal to the number of all subbasins if they are all regulated; however this is never the case of course. The maximum number of FCS considered in Fig. 6 is an example based on some more FCS beyond the current number of FCS (see Fig 3 and inset in Fig. 6). The current FCS plan (16 FCS) is shown against the optimal FCS plan (15

FCS) in the Pareto chart. The inset in Fig. 6 is showing over space the newly suggested FCS (according to the Pareto solution), the current FCS that are not envisioned in the Pareto plan ("removed"), and the ones to keep ("maintain"). This result shows the importance of iPDM in detecting the best spatial arrangement of FCS also in consideration of current and future needs of flood protection, which can be also related to socio-ecological services dependent on FCS such as the promotion of species dispersal and the efficient production of hydropower. Roy et al. (2018), Kareiva (2012) and Ziv et al. (2012) are all examples of potential multi ecosystem service problems that iPDM can accommodate. All these needs can be included into the MCDA model and be traded off with the flood protection needs.

The flood susceptibility and epdf over space of the Pareto optimal solution, as well as of all other scenarios considered, is shown in Figure 7. The total number of possible communities where to place a FCS are 2308 that corresponds to the subbasins for the 5 km^2 thresholded stream network (Fig. S4). The corresponding epdf of flooded areas in the Pareto scenario is reported in Figure 5. Note that here the efficiency is identifying the intrinsic structural potential to retain floods of a FCS (Table S3), and the effectiveness is assessing the true "field capacity" of a FCS plan to retain extreme events causing floods. This capacity is learnt based on data of floods and environmental data; note that floods are also the outcomes of the FCS spatial arrangement beyond the socio-environmental factors characterizing a basin. Yet, Pareto solutions are those that maximize the systemic effectiveness of a portfolio (or set) of FCS. We believe that this concept of interconnected effects is often overlooked in water resource management versus traditional approaches to contain floods in specified locations without thinking in details about the effects in downstream areas.

The difference in systemic risk between the no-FCS and the current scenario (first and second plots) is actually really large despite the epdf of these two scenarios is somewhat similar (Figure 5). Such difference is due to the MCDA model that considers FCS efficiency and criticality (Table S2-S5) which have a profound impact in the determination of the systemic risk. If social and economical criteria are also included in the calculation of the systemic risk we would have a large difference with respect to the risk determined solely by the flood size distribution. The formulation of the systemic value in Eq. 3 (that is the inverse of the systemic risk if only floods are considered in the MCDA model) somehow reflects the traditional framework of risk assessment where hazard, vulnerability and exposures are calculated and combined together. In our case the hazard is constituted by the occurrence of a flood (or not), exposure is determined by the predicted flood size dependent on all FCS effectiveness (i.e. the systemic driver), and vulnerability is related to both the specific local criticality and the efficiency of FCS to protect against floods. The It should be noted that

the systemicity of the risk is highly dependent on the joint effect of all FCS arrangements in the basin which alters the susceptibility calculated by Eq. 1. This in turns determines the extent and connectivity of floods (Fig. 7 top plots).

The predicted floods extracted from the flood probability higher than 0.85, the flood susceptibility and the cumulative output of MaxEnt are shown in Figure 7. In MaxEnt the log output is simply a logistic transformation of the raw output, which indicates a relative probability of occurrence of a flood (this is shown in the bottom plots of Figure 7). In the cumulative output (middle plots), the value of a site is the sum of all the raw outputs from the sites with equal or lower values, times 100. This represents the percentage of the potential distribution that is contained within all the sites that are not more suitable than the one considered. In other words, the logistic output (or flood susceptibility output in the bottom plots) provides the (local) probability value of observing a flood in a site (or better a potential part of a flood at the pixel scale), while the cumulative output of the model is equivalent to the exceedance probability distribution that considers the ability to get that value (of flood probability) or higher with respect to all other pixels. In an information-theoretic framework the local probability can be thought as the local response (or "action/local generation") of each site to local environmental features in generating floods while the epdf can be thought as the systemic response (or "reaction/systemic response") of each site considering surrounding sites and their propensity in generating a flood (see Gutiérrez-Roig et al. (2016) for a broader view of such information-theoretical purview). Recently, Gao et al. (2016) formulated a general model for resilience where local and systemic (network) variability are decoupled in assessing the systemic response. Here, we do somewhat the same by decoupling those local flood propensity features dictated by heterogeneous environmental features and network features dictated by the river network and FCS arrangement. In the middle plots of Figure 7 the more red a site is, the higher the contribution of many other sites in determining the flood susceptibility in that point; yet, the more that considered point is connected with all others. Thus, from the epdf in space it is possible to determine clusters of interdependent points that are likely forming floods; the random FCS flood scenario is the one with most of the areas in red which are largely interdependent and giving rise to the power-law distribution of flood size.

It is interesting to note that the Pareto optimal management solution decrease the Pareto criticality of floods. The far right distribution of floods in Fig. 7 shows that floods are limited in size and disconnected which avoids "cascading effects" down the drainage network. In the Pareto optimal solution both flood susceptibility and epdf are much smaller than in any other scenario (middle and bottom plots of Fig. 7). This gives much higher flood protection

to the ecosystem as a whole and to critical cities such as Rome and Perugia located in the South-West and North regions of the basin, respectively (see Fig. 3). The distribution of FCS is shown by the red dots in the top plots of Fig. 7. The Pareto FCS are distributed much more homogeneously than in other scenarios; yet, this underlines the fact that structural distribution of assets is very different than their functional distribution manifested by floods. In terms of flood size and susceptibility the worst scenario is the random FCS plan; the random placement of FSC causes more harm than benefits by likely destroying the natural formation of floods along the drainage network. However, the suboptimal "no FCS" scenario is also relatively worrisome considering the more widespread distribution of medium value flood susceptibility than the random scenario. The random scenario is more extreme in terms of clustered areas at high flood susceptibility (including around the city of Rome) while the "no FCS" scenario is worse for distributed medium flood susceptibility. This is emphasized by the epdf at the basin scale in Fig. 5.

The way in which the normalized systemic value of FCS for flood protection increases with FCS typological diversity is shown in Figure 8. In our case the diversity of FCS (protecting assets) is not very high since the maximum number of diverse assets is five (Table S1 lists all FCS plus the no action alternative). Here we show four groups of assets since two of them (i.e., earth dam with mantle and nucleus, and gravity brick dam and filled with soil spurs; see Table S1) can be grouped in the same structural category ("earth dam"). The cost is here taken as a percentage of each infrastructure type (selected by iPDM at the basin scale) with respect to the feasible total number of FCS in the basin. In real settings the cost function of FCS is determined considering construction and maintenance costs but it can also integrate other monetary functions that are forming the basin payoff function together with the reduced flood risk they contributed to generate.

The reader should note how for high FCS typological diversity the cost of the most expensive assets (concrete dams) is decreasing with respect to low diversity scenarios in favor of other assets. This is likely because this massive high cost FCS are disregarded in favor of other "soft" FCS optimally organized throughout the basin. This high resource optimal diversification pattern is typically seen in other human and natural system portfolio contexts such as in stock markets and ecosystems (Haldane and May, 2011a) where optimality is reached via collective organization (scale free) and diversification. For instance, Zhou et al. (2012) has evidenced this in the context of flood protection. It is known that optimal solutions are more stable and more diverse, as well as less complex considering the balanced mutual interdependencies of assets (May, 2013) and their scale-free distribution of values that make systems more predictable (Li and Convertino, 2019). Li and Convertino (2019)

analyzed the diversity-complexity-stability landscape of microbial ecosystems and derived universal organization rules for complex systems anchored to the Pareto optimality principle. Stability is here considered as ecosystem occurrence in an optimal state as well as the ability to move to better states or to bounce back if perturbed; yet it is a broader view of resilience (Convertino and Valverde, 2019; Li and Convertino, 2019). In ecological literature many papers have speculated the positive "diversity-health" relationship that here we see in the context of flood protection. "Health" is here an ecosystem indicator measured as a continuous function by the systemic value that is the inverse of the number of flooded areas (if other socio-economical outcomes are not considered). The higher the likelihood of floods the higher the health risk considering both morbidity and mortality outcomes of populations. This is particularly true in developing countries, where floods spread infectious and toxic agents, but also in extreme settings such as in Europe and USA where anthropogenic pressure and climate change are exacerbating flood frequency and their effects which can lead to infrastructure and human losses. In a broader sense, the "diversity-resilience" (or "diversity-health") hypothesis formulated for socio-ecological systems is reflected by the findings of Figure 8. In a more engineering sense (with respect to our aforementioned perspective), resilience is a broad concept underlying positive feedback mechanisms leading to optimal and stable (or increasing) system's performance after a sequence of perturbations (Valverde and Convertino, 2019), that can be achieved via activated or automated heterogenous management such as a re-arrangement of system's controls (e.g. FCS in this case).

The difference in systemic value is defined as the Value of Information (VoI) (Convertino et al., 2014b) that is the amount of information a decision maker can be willing to pay before making a decision. In our case study, this information can be economically quantified and measures in more depth as the Δ between the optimal solution and the status quo or any other state being evaluated for the basin in terms of flood protection. It is interesting to note that for a very diversified portfolio the normalized cost of optimal FCS plans is lower. This is likely related to the better distribution of resources allowed by diversely priced FCS. To underline the importance of diverse portfolio approaches, we re-emphasize how the susceptibility of MaxEnt (Fig. 7) is different than the systemic risk (Fig 6) even when social and economical outcomes are not considered. In fact the risk is based on the flood size that is calculated a posteriori the definition of a threshold on the flood susceptibility to represent the 200 year return time floods. Additionally, the systemic risk is not the inverse function of the systemic value because the latter is based on a more complex multicriteria function containing social and economical variables if included (Eq. 4). This underlines the non-linear relationships between susceptibility, risk and value and emphasizes the need of

resilience focused approaches based on long term trajectories of ecosystem values (Valverde and Convertino, 2019; Convertino and Valverde, 2019).

The systemic uncertainty evaluation of the model is shown in Figure 9. In particular the ROC to measure prediction accuracy is show as well as the iGSUA of all variables introduced in the prediction in the form of jackknife test (Phillips et al., 2006; Phillips and Miroslav, 2008; Convertino et al., 2013b) and functional network of interdependency among predictors. The iGSUA results are shown for the optimal version of the model with the highest information-theoretic power (i.e. the highest entropy in Eq. 5). This high entropy state corresponds to the lowest complexity and highest predictive accuracy (see also Li and Convertino (2019) for an application to microbial ecosystems). 89% of the observed floods are reproduced, which translates into a AUC=0.89. All variables whose regularized training gain (that is the variable specific prediction importance) is lower than 0.5 are disregarded from the model. The ROC curve shows that the smallest floods are not reproduced well considering the small values of the sensitivity for small floods; this is a general result of MaxEnt that needs to calibrate on some "event sizes" that are larger than small event sizes. In our case study these are the larger floods that encompass a wide range of environmental variability. However, it is majorly important to represent extreme events better than small ones because they contribute more to forming the extreme risk. The ROC curve shows an exponential increase in sensitivity after small size events. While it is certainly true that small/high frequency events can carry substantial economic losses, big events can be catastrophic events that destroy completely the areas being flooded with cascading outcomes such as diseases and generalized transport of pollutants.

The jackknife test (bottom bar plot in Fig. 9) provides an estimate of the first and second order sensitivity indices (S_i and S_{ij}, respectively) for predicting the observed landslides (see Section 2.7). The functional interaction network shows S_i and S_{ij} proportional to the diameter of nodes and the width of links, respectively. Elevation is the most important predictor of flood susceptibility and it majorly affects the topographic index, as expected, since TI is derived from the elevation information. TI, also known as the wetness index (WI) is a steady state wetness index that is commonly used to quantify topographic (or geomorphological) control on hydrological processes; it is calculated based on the upslope area and the local slope (Convertino et al., 2013b). Ideally, in presence of dynamical information this index can also be dependent on time and contribute to almost real-time forecast of floods The elevation is also affecting in a minor way all other predictors and the directionality of the functional dependence is expected to be from the elevation variable to all others (whether transfer entropy is used to assess that functional variable dependence). Note that in this case

study all these variables are static variables, thus it is not possible to assess the direction of their functional dependency (see Section 2.7). The low importance of drainage directions was expected because floods can occur for any value of drainage directions. Floodplains are low laying flat areas with undefined local flow directions. In fluvial corridors flood waves move by cutting meanders and local flow patterns are progressing downstream along a main direction. As a result flood prone areas are characterized by almost all values of drainage directions independently of the methods used to calculate them (Rodríguez-Iturbe and Rinaldo, 2001; Nardi et al., 2008, 2019).

The proposed iPDM is also suited to evaluate the impact of green infrastructure solutions for flood risk management vs. standard grey infrastructure (Temmerman et al., 2013; Keesstra et al., 2018). Standard FCS are usually adopted for the lack of a quantitative measure of the multi-sectorial (social, economical, and environmental) beneficial effect of non structural and green infrastructure solutions. Additionally, there is currently no model that tells where to locate these green infrastructure in river basins optimally. Nevertheless, the socio-economic benefit of nature-based solutions can be quantitatively evaluated by means of the proposed iPDM which supports innovative decision making towards large scale and long term green infrastructure strategies. A quantitative data-driven decision model could support the much needed cultural change for moving decision making towards river basin hydrology-driven risk management based on diverse FCS and systemic basin scale actions. This would pave the way to a novel flood management paradigm based on the combination of "green" and "grey" solutions (or any other diverse solution) — recognizing the complexity of intertwined natural and human systems — that optimize human and natural resources including life safety and ecosystem productivity.

Lastly, it is crucial to mention that iPDM is open to mental modeling, that is the mapping and quantitative incorporation of the whole decision space of stakeholders involved in the problem (Kolkman et al., 2005; Wood et al., 2012b,a; Convertino et al., 2016). In this case mental modeling would be focused on flood protection and all its dependent ecosystem services of biological, social, and economic nature (Rinaldo et al., 2018). This kind of modeling (Grimm et al., 2005) embraces the socio-hydrology paradigm (embracing sustainability) that is currently relatively widespread in the hydrological community (Blöschl and Montanari, 2010; Viglione et al., 2014; Blöschl et al., 2017). Risk management decisions that are informed by and address decision makers and stakeholder risk perceptions and behavior are in fact essential for effective risk management policy and ultimately environmental planning. Therefore mental modeling is completely fitting in the systemic risk purview of iPDM via the inclusion of the socio-ecological aspects of the problem (e.g. weights in the value of Eqs.

3 and 4) handled beyond the traditional hydraulics and hydrologic models.

4 Conclusions

An information-theoretic Portfolio Decision Model is proposed as an optimal model and operational process to design optimal FCS plans at the basin scale. This paper proposes a case study for the Tiber basin and traditional FCS types but any basin and FCS can be considered by the portfolio model. As for FCS types, for instance, flood-prone riparian areas that embrace a more "engineering with nature" management approach can be considered. Overall, the portfolio approach supports a participatory, transparent, and rigorous decision making proposition about environmental planning that focuses on the optimization of the systemic value of FCS services for the whole basin ecosystem. The model can operationalize detailed flood mapping models based on hydraulic processes and can consider any scenario (such as related to climate and land-use/land-cover change) coupled to all possible management plans. A flood control strategy should consider different FCS plans over time considering the changing basin conditions. For simplicity in this paper we show the average "static" application of iPDM anchored to floods with an extreme return period. The flood susceptibility is proportional to the risk, although in a non-linear way considering the complexity of river basins dictated by land features, geomorphological features, infrastructure heterogeneities, and diverse stakeholder preferences such as for the Tiber basin considered. The truly innovative part of the study is the portfolio framework and model that brings together several submodels aiming to predict characterize, predict and control flood patterns. Beyond the novel introduction of iPDM, the following results are worth mentioning:

- The Pareto management optimal solution is shown to remove critical cascading effects related to propagation of big floods downstream. This can be seen by the "broken" spatial distribution of floods along the network that appeared more isolated one from another. Additionally the Pareto optimal solution decreases consistently the maximum size of the largest flood; yet, it reduces the (Pareto) power-law regime (or "critical" regime) of the flood size distribution. In other words, the optimal Paretian management plan reduces the risk related to the naturally self-organized floods. This Pareto management corresponds to a "neutral" state where local conflicting small scale heterogeneities are well balanced into a maximim systemic value (see Li and Convertino (2019) for a biological application where the neutral state corresponds to the critical state). We propose the shape (and slope in case of a power-law distribution) of the exceedance distribution of flooded areas (that is related to the fractal dimension of floods

in 2D) as a macrohydrological indicator of phase transitions in flood dynamics. The random FCS plan is worse than no-FCS in terms of flood size distribution because FCS produce cascading effects — in terms of connected floods — that can be highly dangerous if the spatial arrangement maximizes flood propagation. The no-FCS solution determines the natural scale-free distribution of floods if no action is taken; this is the critical state that is undesired because persistent and with maximum flood size. When considering criticality and effectiveness of FCS via the MCDA model the systemic risk of the random FCS plan is lower than the no-FCS plan. This risk is different because it considers socio-ecological endpoints beyond the flood size after the calculation of flood susceptibility.

- The flood delineation method, based on the predicted MaxEnt flood susceptibility and the von Neumann criteria on the pixel flood susceptibility, is a novel information-theoretic model that can be used more easily than more complex hydraulic process-oriented models (such as hydrogeomorphic floodplain delineation models or hydrodynamics models), in calculating the flood area and yet the flood risk based on the flood size epdf. The flood size epdf is a geomorphic function indicating phase transitions in flood patterns. MaxEnt is introduced from ecological sciences as a suitable model to predict flood susceptibility; in this context it is found that elevation is the most important predictor of susceptibility. The second most important predictor is the topographic index that is also known as the wetness index and this can be used as a dynamical factor when predicting flood susceptibility over time. Second order predictors are land-cover/-use, drainage area and drainage directions.

- iPDM allows decision makers to create a baseline value of systemic risk, or vice versa of systemic ecosystem value of the basin in terms of flood safety and socio-economical outcomes. Yet, any future investment can be compared to that baseline value via the use of the Pareto frontier that shows the departure of any basin state from the optimal solution. The Pareto frontier also shows the necessary and sufficient investment for flood protection before reaching the plateau where additional investments in resources do not increase the systemic value significantly (or does not decrease the Systemic Flood Risk, equivalently, if only losses in terms of flood area are considered). We show that the normalized systemic value of assets increases when the diversity of FCS increases but the spatial arrangement of FCS must be Pareto optimal distribution. This results shows the non-linearity between structural features of the basin (such as river network structure and FCS arrangement) and functional features (such as flood size

and systemic value distributions). The difference in systemic value can be considered as the Value of Information before reaching a decision about a FCS plan; information that has high socio-economical and environmental value that decision makers should be willing to pay for maintaining or increasing ecosystem value.

Glossary

SFR: Systemic Flood Risk

iPDM: information-theoretic Portfolio Decision Model

FCS: Flood Control Structure

iGSUA: information-theoretic Global Sensitivity and Uncertainty Analyses

MaxEnt: Maximum Entropy

MCDA: Multi-Criteria Decision Analysis

TI: Topographic Index

DEM: Digital Elevation Model

LC/LU: Land Cover/Land Use

AUC: Area Under the Curve

MI: Mutual Information

MII: Mutual Information Indices

Acknowledgments

M.C. acknowledges the funding from the Gi-CoRE Station for Big Data & Cybersecurity at Hokkaido University (Sapporo, JP), the Microsoft AI for Earth Program grant "Bio-Hydro-Geo Dynamics: Mapping Systemic Earth Risk" https://www.microsoft.com/en-us/ai/ai-for-earth-grantsg, and the NSF SRN project ("SRN: Integrated Urban Infrastructure Solutions for Environmentally Sustainable, Healthy, and Livable Cities") http://www.sustainablehealthycities.org. F.N. and A.A. acknowledge the support of the Water Resources Research and Documentation Center (WARREDOC) of Università per Stranieri di Perugia and the financial support of Regione Lazio Grant number A11598 "Research grant Media Valle del fiume Tevere" and the Autorità di Bacino del fiume Tevere Working Group "Flood map PS5 updating for the city of Rome by means of GIS and 2D hydraulic modeling". Dr. James Valverde Jr. (Department of Energy) is highly acknowledged for comments to the paper and the original formulation of the portfolio model with M.C.. All data have been made available (https://github.com/matteoconvertino/FloodTevereMaxEnt). ~~according to the AGU Data Policy and our Reproducible research policy.~~

Authors Contribution

M.C. formulated the original idea of iPDM for flood management, designed the model, performed the calculations and wrote the paper. A.A. helped in structuring the data for the model and provided all the GIS support. F.N. provided the hydrological data and expertise, and contributed to the writing of the paper.

References

Jeroen C.J.H. Aerts, Wouter Botzen, A. van der Veen, Jorg Krywkow, and Saskia Werners. Dealing with Uncertainty in Flood Management Through Diversification. Ecology and Society, 13(1):1–17, 2008. ISSN 1708-3087.

A Annis, F Nardi, RR Morrison, and F Castelli. Investigating hydrogeomorphic floodplain mapping performance with varying dtm resolution and stream order. Hydrological Sciences Journal, (just-accepted), 2019.

Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality. Physical review A, 38 (1):364, 1988.

S Battiston and G Caldarelli. Systemic risk in financial networks. Journal of Financial Management Markets and Institutions, 1:129–154, 2013. doi: https://doi.org/10.12831/75568. URL http://www.simpolproject.eu/sites/default/files/papers/battiston2013systemic.pdf.

S. Battiston, G. Caldarelli, R. M. May, T. Roukny, and J. E. Stiglitz. The price of complexity in financial networks. Proceedings of the National Academy of Science, 113:10031–10036, 2016. doi: 10.1073/pnas.1521573113.

Vicki M. Bier, Yacov Y. Haimes, James H. Lambert, Nicholas C. Matalas, and Rae Zimmerman. A survey of approaches for assessing and managing the risk of extremes. Risk Analysis, 19(1):83–94, 1999. doi: 10.1111/j.1539-6924.1999.tb00391.x.

Günter Blöschl and Alberto Montanari. Climate change impactsthrowing the dice? Hydrological processes, 24(3):374–381, 2010.

Günter Blöschl, Julia Hall, Juraj Parajka, Rui AP Perdigão, Bruno Merz, Berit Arheimer, Giuseppe T Aronica, Ardian Bilibashi, Ognjen Bonacci, Marco Borga, et al. Changing climate shifts timing of european floods. Science, 357(6351):588–590, 2017.

Rebekka Burkholz, Antonios Garas, and Frank Schweitzer. How damage diversification can reduce systemic risk. Phys. Rev. E, 93:042313, 2016. doi: 10.1103/PhysRevE.93.042313. URL https://link.aps.org/doi/10.1103/PhysRevE.93.042313.

G Calenda, M Casaioli, C Cosentino, R Mantovani, and A Speranza. Flood forecasting in the tiber catchment area: a methodological analysis. Annals of Geophysics, —, 2000. doi: http://dx.doi.org/10.4401/ag-3680.

G. Calenda, C.P. Mancini, and E. Volpi. Distribution of the extreme peak floods of the tiber river from the xv century. Advances in Water Resources, 28(6):615 – 625, 2005. ISSN 0309-1708. doi: http://doi.org/10.1016/j.advwatres.2004.09.010.

Benedetto Calvo and Fabrizio Savi. Real-time flood forecasting of the tiber river in rome. Natural Hazards, 50(3):461–477, 2009. doi: 10.1007/s11069-008-9312-9.

M. Convertino. Neutral metacommunity clustering and sar: River basin vs. 2-d landscape biodiversity patterns. Ecological Modelling, 222(11):1863 – 1879, 2011. ISSN 0304-3800. doi: http://dx.doi.org/10.1016/j.ecolmodel.2011.03.015.

M. Convertino, K.M. Baker, J.T. Vogel, C. Lu, B. Suedel, and I. Linkov. Multi-criteria decision analysis to select metrics for design and monitoring of sustainable ecosystem restorations. Ecological Indicators, 26:76 – 86, 2013a. doi: 10.1016/j.ecolind.2012.10.005. URL http://dx.doi.org/10.1016/j.ecolind.2012.10.005.

M. Convertino, A. Troccoli, and F. Catani. Detecting fingerprints of landslide drivers: A maxent model. Journal of Geophysical Research: Solid Earth, 118:1367–1386, 2013b. ISSN 0148-0227. doi: 10.1002/jgrf.20099.

M. Convertino, R. Munoz-Carpena, M.L. Chu-Agor, G.A. Kiker, and I. Linkov. Untangling drivers of species distributions: Global sensitivity and uncertainty analyses of maxent. Environmental Modelling & Software, 51:296 – 309, 2014a. doi: http://dx.doi.org/10.1016/j.envsoft.2013.10.001.

Matteo Convertino and L. James Valverde. Toward a pluralistic conception of resilience. Ecological Indicators, 2019. doi: inpress.

Matteo Convertino and L. James Valverde, Jr. Portfolio decision analysis framework for value-focused ecosystem management. PLoS ONE, 8(6):e65056, 2013. doi: 10.1371/journal.pone.0065056.

Matteo Convertino, Riccardo Rigon, Amos Maritan, Ignacio Rodriguez-Iturbe, and Andrea Rinaldo. Probabilistic structure of the distance between tributaries of given size in river networks. Water Resources Research, 43(11), 11 2007. doi: 10.1029/2007WR006176.

Matteo Convertino, Adam Bockelie, Gregory A Kiker, Rafael Muñoz-Carpena, and Igor Linkov. Shorebird patches as fingerprints of fractal coastline fluctuations due to climate change. Ecological Processes, 1(1):9, 2012.

Matteo Convertino, Filippo Simini, Filippo Catani, Igor Linkov, and Gregory A. Kiker. Power-law of aggregate-size spectra in natural systems. EAI Endorsed Transactions on Complex Systems, 13(2), 2013c. doi: 10.4108/trans.cs.1.2.e2.

Matteo Convertino, Christy M. Foran, Jeffrey M. Keisler, Lynn Scarlett, Andy LoSchiavo, Gregory A. Kiker, and Igor Linkov. Enhanced adaptive management: Integrating decision analysis, scenario analysis and environmental modeling for the everglades. Scientific Reports, 2922, 2014b. doi: doi:10.1038/srep02922.

Matteo Convertino, Rafael Munoz-Carpena, and Carolina Murcia. "Reading the Minds" for Quantitative Sustainability: Assessing Stakeholder Mental Models via Probabilistic Text Analysis, pages 21–38. 2016. doi: 10.1007/978-3-319-25439-5_2. URL https://doi.org/10.1007/978-3-319-25439-5_2.

Giuliano Di Baldassarre, Smeralda Saccà, Giuseppe Tito Aronica, Salvatore Grimaldi, Alessio Ciullo, and Massimiliano Crisci. Human-flood interactions in rome over the past 150 years. Advances in Geosciences, 44:9, 2017.

M. Dudik, S. J. Phillips, and R. E. Schapire. Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. Journal of Machine Learning Research, 8:1217–1260, 2007.

J. Elith, S.J. Phillips, T. Hastie, M. Dudík, Y.E. Chee, and C.J. Yates. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions, -, 2010. doi: 10.1111/j.1472-4642.2010.00725.x.

B. M. Fiseha, S. G. Setegn, A. M. Melesse, E. Volpi, and A. Fiori. Hydrological analysis of the upper tiber river basin, central italy: a watershed modelling approach. Hydrological Processes, 27(16):2339–2351, 2013. doi: 10.1002/hyp.9234.

Mark Fonstad and W Andrew Marcus. Self-organized criticality in riverbank systems. Annals of the Association of American Geographers, 93(2):281–296, 2003.

Marco Franchini, Anna Bernini, Silvia Barbetta, and Tommaso Moramarco. Forecasting discharges at the downstream end of a river reach through two simple muskingum based procedures. Journal of Hydrology, 399:335 – 352, 2011. ISSN 0022-1694. doi: http://doi.org/10.1016/j.jhydrol.2011.01.009.

Jianxi Gao, Baruch Barzel, and Albert-László Barabási. Universal resilience patterns in complex networks. Nature, 530(7590):307–312, 2016. ISSN 0028-0836. doi: 10.1038/nature16948.

Volker Grimm, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M. Mooij, Steven F. Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand, and Donald L. DeAngelis. Pattern-oriented modeling of agent-based complex systems: Lessons from ecology. Science, 310(5750):987–991, 2005. doi: 10.1126/science.1116681. URL http://science.sciencemag.org/content/310/5750/987.

M Gu, K Wiesner, E Rieper, and V Vedral. Quantum mechanics can reduce the complexity of classical models. Nature Communications, —(—):—, 2012. doi: 10.1038/ncomms1761.

M. Gutiérrez-Roig, O. Sagarra, A. Oltra, J. R. B. Palmer, F. Bartumeus, A. Díaz-Guilera, and J. Perelló. Active and reactive behaviour in human mobility: the influence of attraction points on pedestrians. Royal Society Open Science, 3(7), 2016. doi: 10.1098/rsos.160177.

Andrew G. Haldane and Robert M. May. Systemic risk in banking ecosystems. Nature, 469 (7330):351–355, 2011a. ISSN 1476-4687. doi: 10.1038/nature09659.

Andrew G. Haldane and Robert M. May. Systemic risk in banking ecosystems. Nature, 469 (7330):351–355, 2011b. doi: 10.1038/nature09659. URL http://dx.doi.org/10.1038/nature09659.

D. Helbing. Globally networked risks and how to respond. Nature, 497:51–59, 2013. doi: 10.1038/nature12047.

IPCC. Ipcc fifth assessment report. Technical report, Intergovernmental Panel on Climate Change, 2014. http://www.ipcc.ch/report/ar5/index.shtml; Date of access: 04/10/2017.

E. T. Jaynes. Information theory and statistical mechanics. Phys. Rev., 106:620–630, 1957. doi: 10.1103/PhysRev.106.620. URL https://link.aps.org/doi/10.1103/PhysRev.106.620.

E. T. Jaynes. The relation of bayesian and maximum entropy methods. In Maximum-Entropy and Bayesian Methods in Science and Engineering, number 1. Kluwer Academic Publishers, 1988.

Peter M. Kareiva. Dam choices: Analyses for multiple needs. Proceedings of the National Academy of Sciences, 109(15):5553–5554, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1203263109. URL https://www.pnas.org/content/109/15/5553.

Saskia Keesstra, Joao Nunes, Agata Novara, David Finger, David Avelar, Zahra Kalantari, and Artemi Cerdà. The superior effect of nature based solutions in land management for enhancing ecosystem services. Science of the Total Environment, 610:997–1009, 2018.

Marinus J Kolkman, Matthijs Kok, and Anne Van der Veen. Mental model mapping as a new tool to analyse the use of information in decision-making in integrated water management. Physics and Chemistry of the Earth, Parts A/B/C, 30(4):317–332, 2005.

J. Korcak. Deux types fondamentaux de distribution staffstique. Bull. Inst. Int. Star., 30: 295–307, 1940.

Jie Li and Matteo Convertino. Optimal microbiome networks: Macroecology and criticality. Entropy, 21(5), 2019. ISSN 1099-4300. doi: 10.3390/e21050506. URL http://www.mdpi.com/1099-4300/21/5/506.

Yang Liu, Shubhayu Saha, Brendalynn O. Hoppe, and Matteo Convertino. Degrees and dollars health costs associated with suboptimal ambient temperature exposure. Science of The Total Environment, 678:702 – 711, 2019. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2019.04.398. URL http://www.sciencedirect.com/science/article/pii/S0048969719319497.

Niklas Lüdtke, Stefano Panzeri, Martin Brown, David S Broomhead, Joshua Knowles, Marcelo A Montemurro, and Douglas B Kell. Information-theoretic sensitivity analysis: a general method for credit assignment in complex networks. Journal of The Royal Society Interface, 5(19):223–235, 2008. doi: 10.1098/rsif.2007.1079.

B.B. Mandelbrot. The Fractal Geometry of Nature. W.H. Freeman, 1982.

S. Manfreda, F. Nardi, C. Samela, S. Grimaldi, A. C. Taramasso, G. Roth, and A. Sole. Investigation on the use of geomorphic approaches for the delineation of flood prone areas. Journal of Hydrology, 517:863–876, 2014. doi: 10.1016/j.jhydrol.2014.06.009.

Salvatore Manfreda, Fernando Nardi, Caterina Samela, Salvatore Grimaldi, Angela Celeste Taramasso, Giorgio Roth, and Aurelia Sole. Investigation on the use of geomorphic approaches for the delineation of flood prone areas. Journal of hydrology, 517:863–876, 2014.

Harry Markowitz. Portfolio selection. The Journal of Finance, 7(1):77–91, 1952. doi: 10.1111/j.1540-6261.1952.tb01525.x.

Robert M. May. Networks and webs in ecosystems and financial systems. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 371, 2013. ISSN 1364-503X. doi: 10.1098/rsta.2012.0376.

Cory Merow, Matthew J. Smith, and John A. Silander. A practical guide to maxent for modeling species distributions: what it does, and why inputs and settings matter. Ecography, 36(10):1058–1069, 2013. doi: 10.1111/j.1600-0587.2013.07872.x.

B. Merz, H. Kreibich, R. Schwarze, and A. Thieken. Review article "assessment of economic flood damage". Natural Hazards and Earth System Sciences, 10(8):1697–1724, 2010. doi: 10.5194/nhess-10-1697-2010. URL https://www.nat-hazards-earth-syst-sci.net/10/1697/2010/.

Volker Meyer, Sebastian Scheuer, and Dagmar Haase. A multicriteria approach for flood risk mapping exemplified at the mulde river, germany. Natural Hazards, 48(1):17–39, 2009. doi: 10.1007/s11069-008-9244-4. URL https://doi.org/10.1007/s11069-008-9244-4.

Ryan R Morrison, Erin Bray, Fernando Nardi, Antonio Annis, and Quan Dong. 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, 2018.

R. Muneepeerakul, E. Bertuzzo, H. J. Lynch, W. F. Fagan, A. Rinaldo, and I. Rodriguez-Iturbe. Neutral metacommunity models predict fish diversity patterns in Mississippi-Missouri basin. Nature, 453:220–222, 2008. doi: 10.1038/nature06813.

G. Napolitano, L. See, B. Calvo, F. Savi, and A. Heppenstall. A conceptual and neural network model for real-time flood forecasting of the tiber river in rome. Physics and Chemistry of the Earth, Parts A/B/C, 35:187 – 194, 2010. doi: http://doi.org/10.1016/j.pce.2009.12.004.

F. Nardi, E. R. Vivoni, and S. Grimaldi. Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. Water Resources Research, 42, 2006. doi: 10.1029/2005WR004155.

F. Nardi, A. Annis, and C. Biscarini. On the impact of urbanization on flood hydrology of small ungauged basins: the case study of the tiber river tributary network within the city of rome. Journal of Flood Risk Management, —, 2015. doi: 10.1111/jfr3.12186.

F Nardi, A Annis, Giuliano Di Baldassarre, ER Vivoni, and S Grimaldi. Gfplain250m, a global high-resolution dataset of earths floodplains. Scientific data, 6:180309, 2019.

Fernando Nardi, Salvatore Grimaldi, Monia Santini, Andrea Petroselli, and Lucio Ubertini. Hydrogeomorphic properties of simulated drainage patterns using digital elevation models: the flat area issue/propriétés hydro-géomorphologiques de réseaux de drainage simulés à partir de modèles numériques de terrain: la question des zones planes. Hydrological Sciences Journal, 53(6):1176–1193, 2008.

Fernando Nardi, Chiara Biscarini, Silvia Di Francesco, Piergiorgio Manciola, and Lucio Ubertini. 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, 2013.

Fernando Nardi, Ryan R Morrison, Antonio Annis, and Theodore E Grantham. Hydrologic scaling for hydrogeomorphic floodplain mapping: Insights into human-induced floodplain disconnectivity. River research and applications, 34(7):675–685, 2018.

L. Natale and F. Savi. Monte carlo analysis of probability of inundation of rome. Environ. Model. Softw., —(10):1409–1416, 2007. doi: 10.1016/j.envsoft.2006.12.004.

V.I. Nikora, C.P. Pearson, and U. Shankar. Scaling properties in landscape patterns: New Zealand Experience. Landscape Ecology, 1999(14):17–33, 1999.

Dominik Paprotny, Antonia Sebastian, Oswaldo Morales-Nápoles, and Sebastiaan N Jonkman. Trends in flood losses in europe over the past 150 years. Nature communications, 9(1):1985, 2018.

V. Pareto. Manual of political economy. Advances in Water Resources, —, 1971. doi: ISBN978-0-678-00881-2.

J. Park, T. P. Seager, P. S. C. Rao, M. Convertino, and I. Linkov. Integrating risk and resilience approaches to catastrophe management in engineering systems. Risk Analysis, 2012. ISSN 1539-6924. doi: 10.1111/j.1539-6924.2012.01885.x.

S. Phillips, R. Anderson, and R. Schapire. Maximum entropy modeling of species geographic distributions. Ecological Modelling, 190(3-4):231–259, January 2006. doi: 10.1016/j.ecolmodel.2005.03.026.

S. J. Phillips and D. Miroslav. Modeling of species distributions with maxent: new extensions and a comprehensive evaluation. Ecography, 31(2):161–175, April 2008. doi: 10.1111/j.0906-7590.2008.5203.x.

S Piccolroaz, M Di Lazzaro, A Zarlenga, B Majone, A Bellin, and A Fiori. Hydroscape: a multi-scale framework for streamflow routing in large-scale hydrological models. Hydrology and Earth System Sciences Discussion, 12:9055–9090, 2015.

Sebastiano Piccolroaz, Michele Di Lazzaro, Antonio Zarlenga, Bruno Majone, Alberto Bellin, and Aldo Fiori. Hyperstream: a multi-scale framework for streamflow routing in large-scale hydrological model. Hydrology and Earth System Sciences, 20(5):2047–2061, 2016.

N Quinn, Paul D. Bates, Jeffrey Neal, Andy Smith, Oliver Wing, Chris Sampson, James Smith, and Janet Heffernan. The spatial dependence of flood hazard and risk in the united states. Water Resources Research, 02 2019. doi: 10.1029/2018WR024205.

Andrea Rinaldo, Marino Gatto, and Ignacio Rodriguez-Iturbe. River networks as ecological corridors: A coherent ecohydrological perspective. Advances in Water Resources, 112:27 – 58, 2018. ISSN 0309-1708. doi: https://doi.org/10.1016/j.advwatres.2017.10.005.

I. Rodríguez-Iturbe and A. Rinaldo. Fractal River Basins. Cambridge University Press, 2001.

Samuel G. Roy, Emi Uchida, Simone P. de Souza, Ben Blachly, Emma Fox, Kevin Gardner, Arthur J. Gold, Jessica Jansujwicz, Sharon Klein, Bridie McGreavy, Weiwei Mo, Sean M. C. Smith, Emily Vogler, Karen Wilson, Joseph Zydlewski, and David Hart. A multiscale approach to balance trade-offs among dam infrastructure, river restoration, and cost. Proceedings of the National Academy of Sciences, 115(47):12069–12074, 2018. ISSN 0027-8424. doi: 10.1073/pnas.1807437115. URL https://www.pnas.org/content/115/47/12069.

A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto. Sensitivity Analysis in Practice: A guide to Assessing Scientific Models. John Wiley and Sons, 2004.

Caterina Samela, Salvatore Manfreda, Francesco De Paola, Maurizio Giugni, Aurelia Sole, and Mauro Fiorentino. Dem-based approaches for the delineation of flood-prone areas in an ungauged basin in africa. Journal of Hydrologic Engineering, 21(2):06015010, 2015.

Caterina Samela, Tara J Troy, and Salvatore Manfreda. Geomorphic classifiers for flood-prone areas delineation for data-scarce environments. Advances in Water Resources, 102: 13–28, 2017.

Kara Scheel, Ryan R Morrison, Antonio Annis, and Fernando Nardi. Understanding the large-scale influence of levees on floodplain connectivity using a hydrogeomorphic approach. JAWRA Journal of the American Water Resources Association, 2019.

Joseph L. Servadio and Matteo Convertino. Optimal information networks: Application for data-driven integrated health in populations. Science Advances, 4(2), 2018. doi: 10.1126/sciadv.1701088.

L. Brocca F. Melone T. Moramarco V. P. Singh. Assimilation of observed soil moisture data in storm rainfall-runoff modeling. Journal of Hydrologic Engineering, 14(2):153–165, 2009. doi: 10.1061/(ASCE)1084-0699(2009)14:2(153).

Karsten Steinhaeuser, Auroop R. Ganguly, and Nitesh V. Chawla. Multivariate and multiscale dependence in the global climate system revealed through complex networks. Climate Dynamics, 39(3):889–895, 2012. ISSN 1432-0894. doi: 10.1007/s00382-011-1135-9. URL https://doi.org/10.1007/s00382-011-1135-9.

P Tasca and S Battiston. Diversification and financial stability. Technical report, London School of Economics and Political Science, 2014. URL https://www.sg.ethz.ch/wps/pdf/CCSS-11-001.pdf. Systemic Risk Centre Discussion Paper No 10.

Paolo Tasca, Stefano Battiston, and Andrea Deghi. Portfolio diversification and systemic risk in interbank networks. Journal of Economic Dynamics and Control, 82:96 – 124, 2017. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2017.01.013.

Stijn Temmerman, Patrick Meire, Tjeerd J Bouma, Peter MJ Herman, Tom Ysebaert, and Huib J De Vriend. Ecosystem-based coastal defence in the face of global change. Nature, 504(7478):79–83, 2013.

USACE. Asset management portfolio analytics case study. Technical report, U.S. Army Corps of Engineers Asset Management Program, 2012.

L. James Valverde and Matteo Convertino. Insurer resilience in an era of climate change and extreme weather: An econometric analysis. Climate, 7(4), 2019. ISSN 2225-1154. doi: 10.3390/cli7040055. URL http://www.mdpi.com/2225-1154/7/4/55.

LJ Jr Valverde and MW Andrews. Global climate change and extreme weather: An exploration of scientific uncertainty and the economics of insurance. Insurance Information Institute, —, 2006. URL https://www.iii.org/article/global-climate-change-and-extreme-weather-an-exploration-of-scientific-uncertainty-and-the

Marco J Van De Wiel and Tom J Coulthard. Self-organized criticality in river basins: Challenging sedimentary records of environmental change. Geology, 38(1):87–90, 2010.

Alberto Viglione, Giuliano Di Baldassarre, Luigia Brandimarte, Linda Kuil, Gemma Carr, José Luis Salinas, Anna Scolobig, and Günter Blöschl. Insights from socio-hydrology modelling on dealing with flood risk–roles of collective memory, risk-taking attitude and trust. Journal of Hydrology, 518:71–82, 2014.

E Volpi, A Fiori, S Grimaldi, F Lombardo, and D Koutsoyiannis. One hundred years of return period: Strengths and limitations. Water Resources Research, 51(10):8570–8585, 2015.

M Wood, A Bostrom, M Convertino, D Kovacs, and I Linkov. A moment of mental model clarity: response to jones et al. 2011. Ecology and Society, 17(4), 2012a.

M Wood, D Kovacs, A Bostrom, T Bridges, and I Linkov. Flood risk management: Us army corps of engineers and layperson perceptions. Risk Analysis, 32(8):1349–1368, 2012b.

Stefano Zanardo, Ludovico Nicotina, Arno GJ Hilberts, and Stephen P Jewson. Modulation of economic losses from european floods by the north atlantic oscillation. Geophysical Research Letters, 46(5):2563–2572, 2019.

Qian Zhou, James H. Lambert, Christopher W. Karvetski, Jeffrey M. Keisler, and Igor Linkov. Flood protection diversification to reduce probabilities of extreme losses. Risk Analysis, 32(11):1873–1887, 2012. doi: 10.1111/j.1539-6924.2012.01870.x.

Guy Ziv, Eric Baran, So Nam, Ignacio Rodríguez-Iturbe, and Simon A. Levin. Trading-off fish biodiversity, food security, and hydropower in the mekong river basin. Proceedings of the National Academy of Sciences, 109(15):5609–5614, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1201423109. URL https://www.pnas.org/content/109/15/5609.

M.H. Zweig and G. Campbell. Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem., 39(4):561 – 577, 1993.

Figure Captions

Figure 1. Conceptual diagram of the portfolio decision model. Hydro-geomorphological data enter the model as input factors that characterizes them probabilistically via GSUA. The MaxEnt model determines the likelihood of an area to be flooded dependent on a FCS plan. The MaxEnt flood susceptibility is used for extracting floods considering a susceptibility threshold on floods with a certain return period. The MCDA model assesses the systemic risk of all FCS plans over the whole basin, defined by the epdf of flood size, and the Pareto optimization model determines the optimal FCS plan that minimizes the risk of flooding (or conversely maximizes the systemic value considering social and economic criteria) constrained to the resources available. An additional model that can be used for explicitly time dependent flood pattern predictions (not included in the current study) is a Transfer Entropy (TE) model for assessing flood interdependencies over space and time (see Li and Convertino (2019) and Servadio and Convertino (2018) for applications in biology and population health sciences).

Figure 2. Hydro-geomorphological delineation of floods on MaxEnt susceptibility predictions. The plots show the approach of delineating the floods based on the flood susceptibility in analogy to the work done in Convertino et al. (2013b) for landslides. Nearest neighbor pixels, according to the von Neumann criteria (or 4-neighborhood), whose flood susceptibility is higher than the calibrated flood susceptibility threshold are categorized as flooded areas. The calibration of the threshold on the flood susceptibility (0.85) constitutes the validation of the model on the reported 200 years floods.

Figure 3. Hydro-geomorphological delineation of the basin and flood control structures. FCS and previous floods are the flood related input factors (socio-infrastructural variables) of the MaxEnt model beyond the ecohydrological information (Fig. 3). The backbone of the Tiber river network and the basin boundaries are highlighted on the top of the elevation field and the middle plot. The middle plot shows current FCS structures (Table S1) and observed flood locations. The 200 years floods (right plots), that are typically used to design water control structures, are the one used in the MaxEnt model to validate the predictions and these to determine the systemic risk. The extracted river networks (with a threshold on the drainage area equal to 5 km^2) composed by 11,321 streams is shown with the 200 years flood map on the right plot.

Figure 4. Ecohydrological information in input of the MaxEnt model. Drainage

directions, drainage area, river network and topographic index are all derived form the Digital Elevation Model (DEM) at resolution 500×500 m^2. The land-use/-cover is the only information that is not dependent on the geomorphology of the river basin; however, it profoundly determines the likelihood of an area to be flooded. The river network is shown for a threshold on the minimum channelized drainage area of 5 km^2.

Figure 5. Exceedance probability distribution function of the flood size. The epdf is referring to the scenarios represented in Figure 7 and is expressing flood size in m^2. The random, Pareto optimal and no-FCS epdfs correspond to the plausible (network) scenarios where FCS are distributed randomly, optimally (Pareto optimality sensu), and completely not adopted in the basin, respectively. ϵ is the slope of the power-law epdf that is smaller for scenarios with likelihood of larger and more probable floods. ϵ is related to the fractal dimension of flood patterns (Fig. S1). The random FCS plan results to be worse than the no-FCS; however, the MCDA model that considers FCS efficiency and criticality determines a higher risk for the no-FCS plan (Fig. 6); a result that underlines the non-linearity between multi-criteria risk and flooded area based only on flood susceptibility.

Figure 6. Pareto frontiers of optimal FCS plans. The systemic risk is calculated using Eq. 4 that consider flood size and FCS features and floods for any FCS plan. Green dots are along the Pareto frontier of optimal FCS plans while pink dots are all suboptimal FCS plans. All FCS plans are constrained to the resources available (cost) that is the number of FCS in this case in absence of a proper economic valuation. The current FCS plan (16 FCS) is shown against the optimal FCS plan (15 FCS). The random configuration is for the same number of FCS of the current status. The MaxEnt inferred flood susceptibility is proportional to the risk.

Figure 7. Scenario-based predictions of flood suitability and risk based on the exceedance probability. All predictions are based on the flood susceptibility calculated on MaxEnt FCS-based scenarios: no, current, random, and Pareto optimal FCS plans (from left to right). The top plot shows the location of FCS for each scenario and the calculated floods (in black) based on the flood susceptibility (bottom plots) after imposing a threshold of 0.85 (Figure 2) that matches the floods with a return period of 200 years in the current FCS scenario. The exceedance probability for any site in the basin (i.e. a pixel in the Digital Elevation Model) is shown in the middle plots.

Figure 8.

textbfFlood control structure typological diversity, systemic value and normalized cost. The systemic value is here determined as the inverse function of the systemic flood risk and the FCS diversity is the number of diverse water control infrastructure (as listed in the legend) considering the currently implemented typologies (Table S1). The normalized cost is the number of FCS divided by the maximum number of FCS (i.e. the total number of subbasins). This analysis is done considering all subbasins and only the optimal Pareto solutions when changing the number of FCS typologies. The highest systemic value is achieved for the most diverse solution in according to the "diversity-resilience" hypothesis that is observed for many socio-ecological systems.

Figure 9. Prediction accuracy and global sensitivity and uncertainty analyses. The AUC and the jackknife test used for testing the predictions of the model considering the FCS status quo are shown (top and bottom plots). The highest accuracy achievable is for AUC=0.89 that is roughly equivalent to represent 89% of the observed susceptible flood areas with the MaxEnt model. The jackknife test (bottom bar plot) provides an estimate of the first and second order sensitivity indices (S_i and S_{ij}, respectively) for predicting the observed flood susceptibility. The functional interaction network for the predicting variables (Fig 4) shows S_i and S_{ij} that are proportional to the diameter of nodes and the width of links (representing predictor independent importance and mutual interaction with other predictors). The distance between nodes does not have a meaning in this contest.

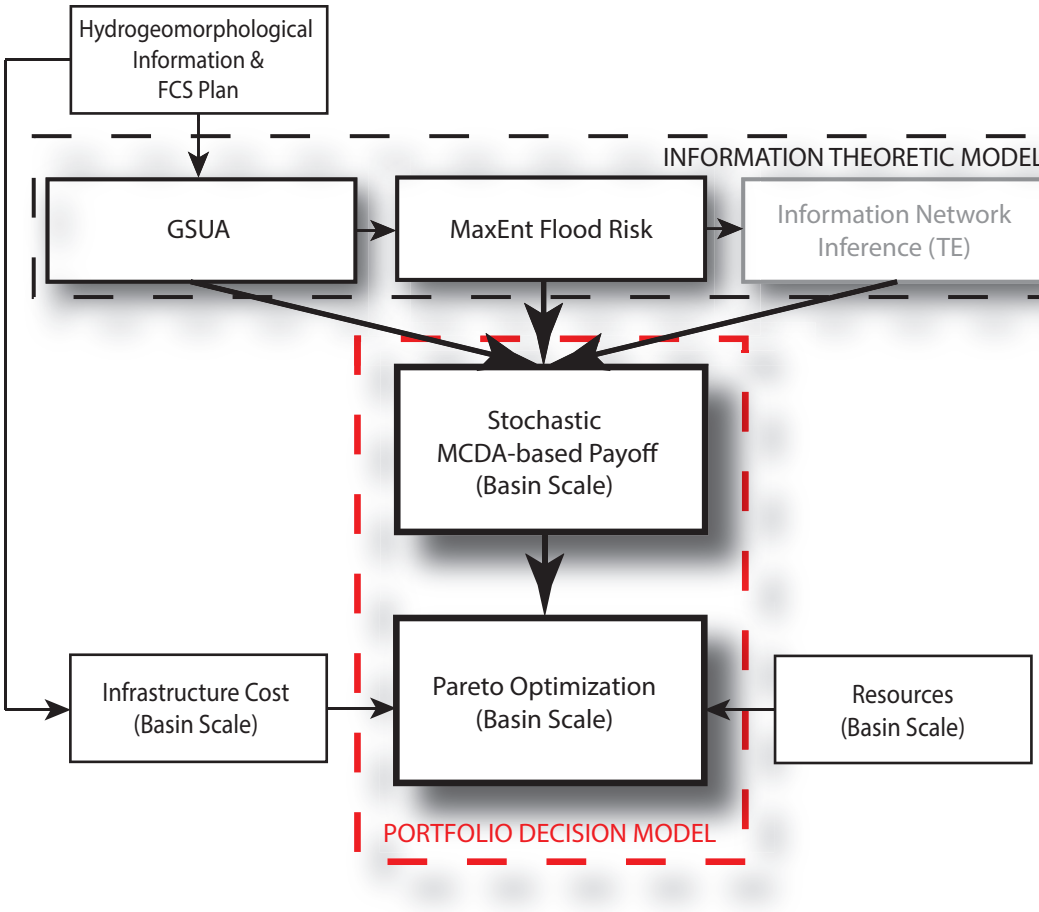

Hydrogeomorphological Information & FCS Plan
INFORMATION THEORETIC MODELS
GSUA
MaxEnt Flood Risk
Information Network Inference (TE)
Stochastic MCDA-based Payoff (Basin Scale)
Infrastructure Cost (Basin Scale)
Pareto Optimization (Basin Scale)
Resources (Basin Scale)
PORTFOLIO DECISION MODEL

Figure 1:

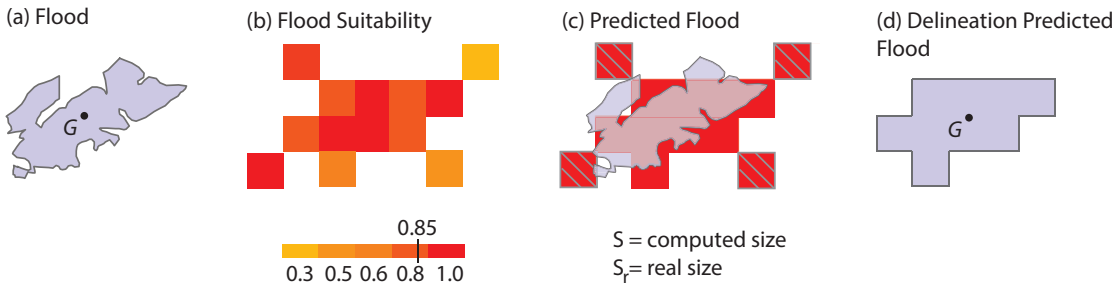

(a) Flood
G
(b) Flood Suitability
0.85
0.3 0.5 0.6 0.8 1.0
(c) Predicted Flood
S = computed size
S_r = real size
(d) Delineation Predicted Flood
G

Figure 2:

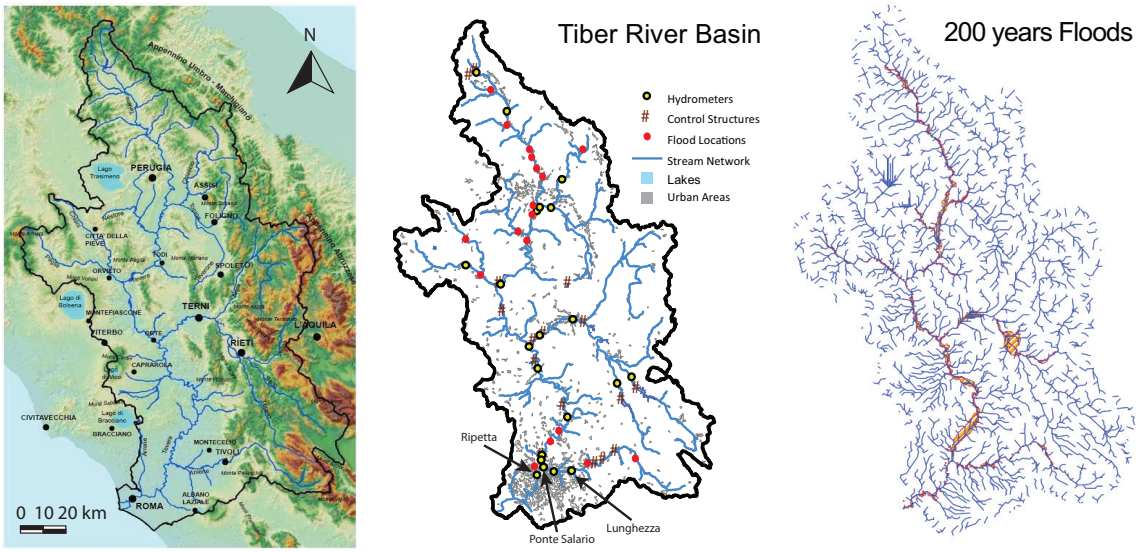

N
PERUGIA
ASSISI
FOLIGNO
CITTA DELLA PIEVE
SPOLETO
ORVIETO
TERNI
MONTEFIASCONE
VITERBO
RIETI
L'AQUILA
CIVITAVECCHIA
BRACCIANO
TIVOLI
ROMA
0 10 20 km
Tiber River Basin
Hydrometers
Control Structures
Flood Locations
Stream Network
Lakes
Urban Areas
Ripetta
Ponte Salario
Lunghezza
200 years Floods

Figure 3:

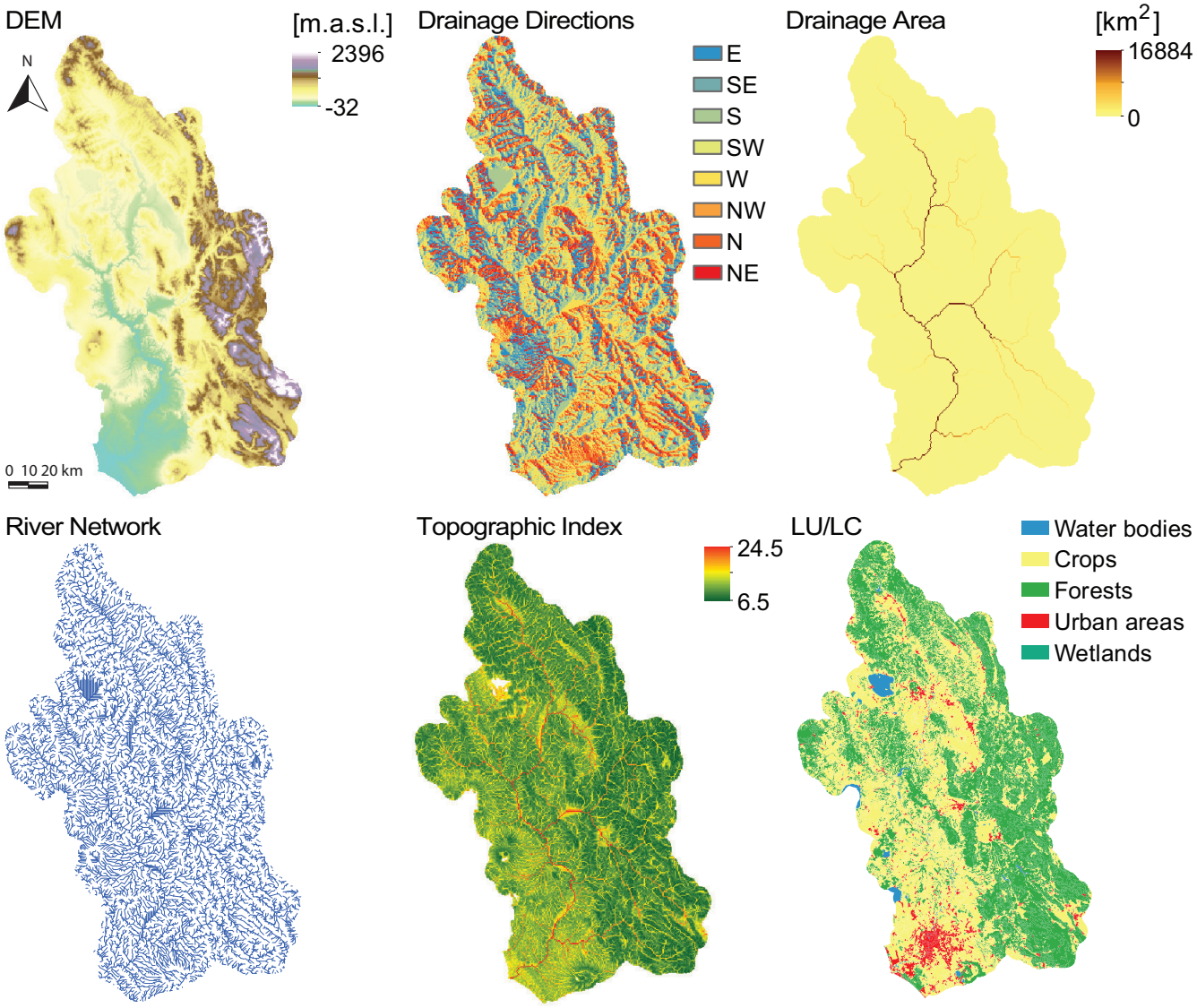

DEM
[m.a.s.l.]
2396
-32
N
0 10 20 km
Drainage Directions
E
SE
S
SW
W
NW
N
NE
Drainage Area
[km2]
16884
0
River Network
Topographic Index
24.5
6.5
LU/LC
Water bodies
Crops
Forests
Urban areas
Wetlands

Figure 4:

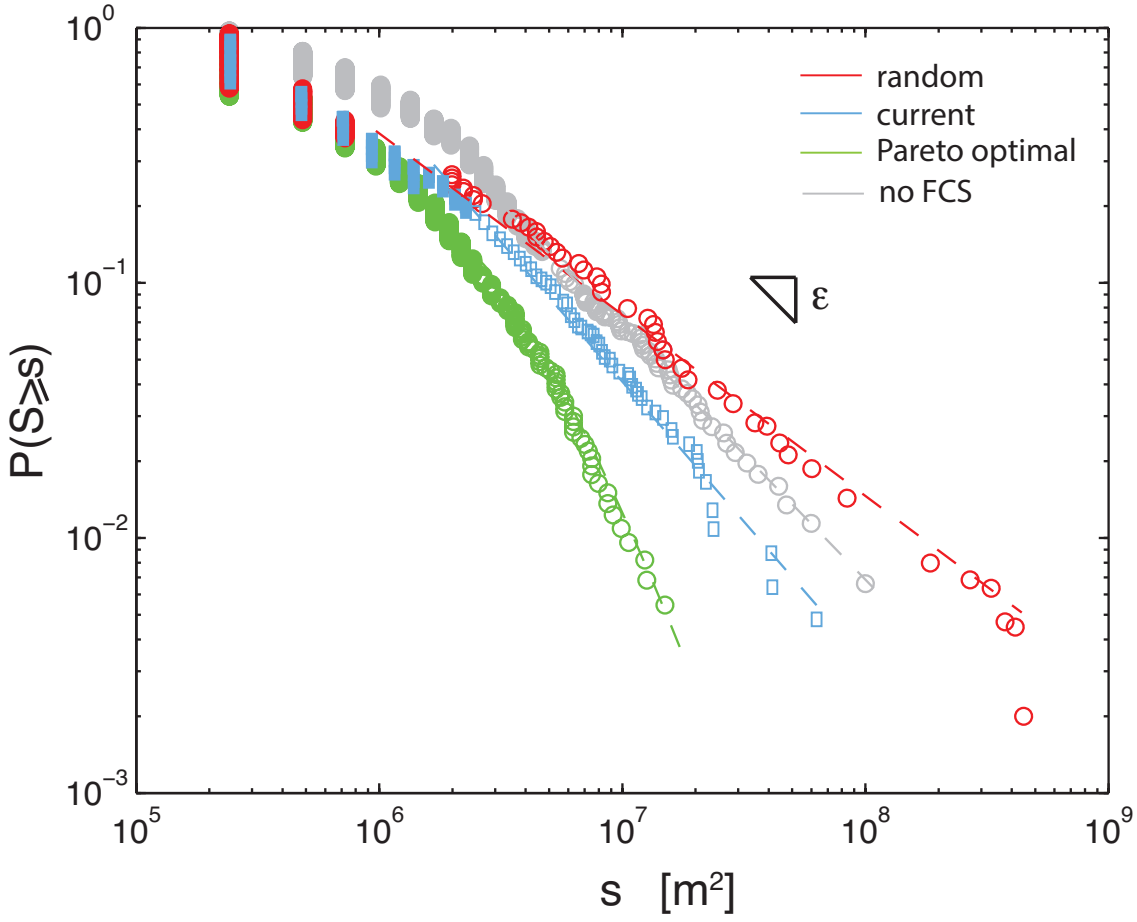

random
current
Pareto optimal
no FCS
ε
P(S≥s)
s [m²]
10^0
10^-1
10^-2
10^-3
10^5
10^6
10^7
10^8
10^9

Figure 5:

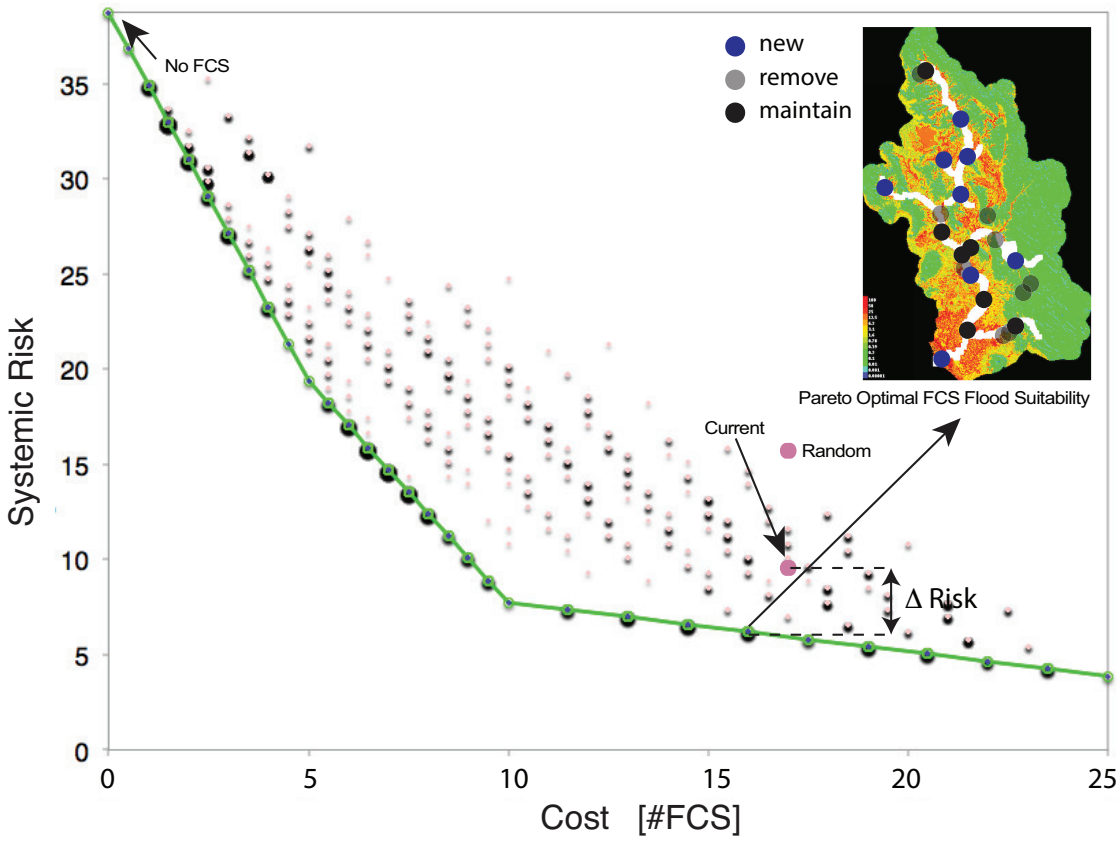

No FCS
new
remove
maintain
Pareto Optimal FCS Flood Suitability
Current
Random
Δ Risk
Systemic Risk
Cost [#FCS]
0
5
10
15
20
25
30
35

Figure 6:

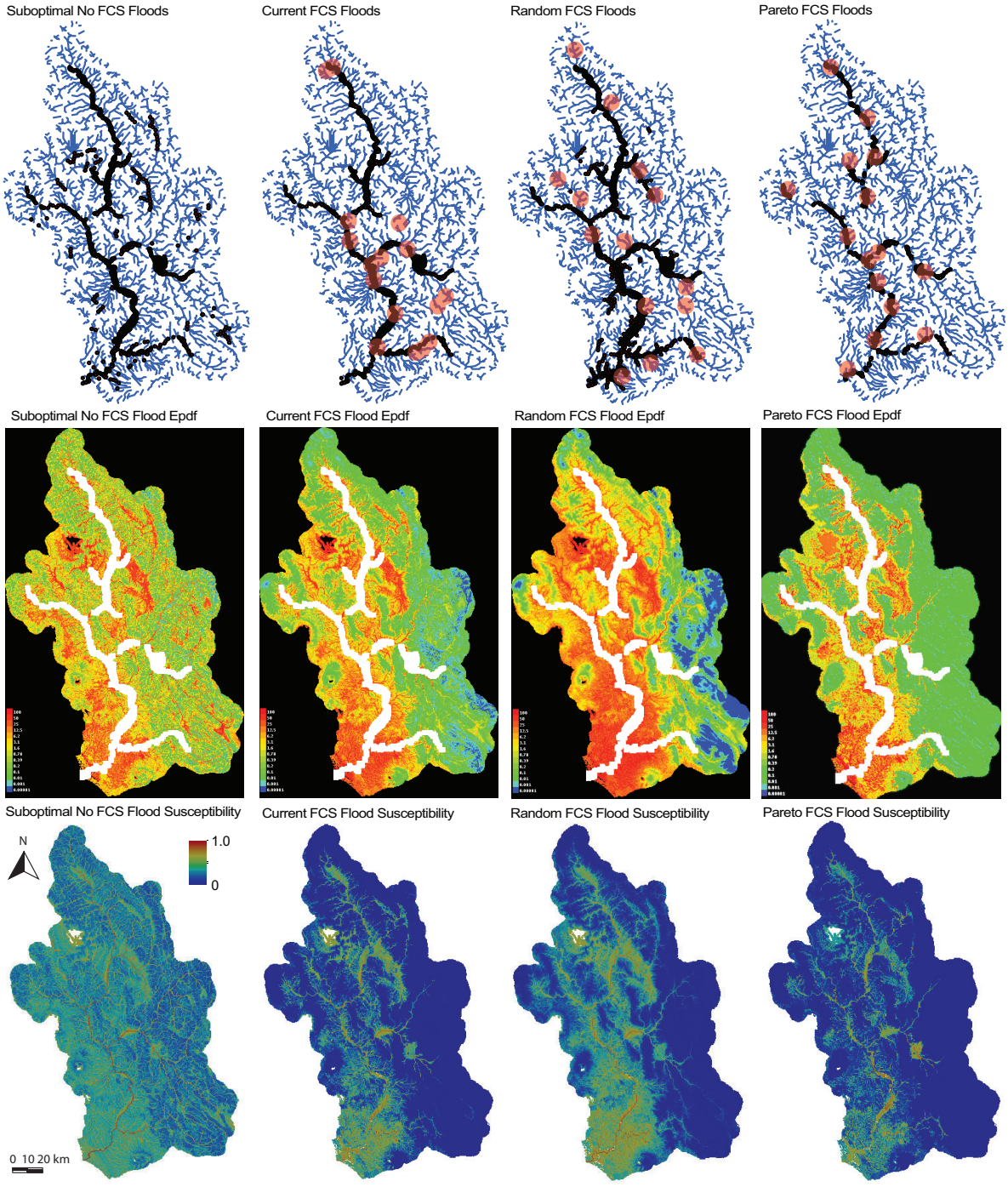

Suboptimal No FCS Floods
Current FCS Floods
Random FCS Floods
Pareto FCS Floods
Suboptimal No FCS Flood Epdf
Current FCS Flood Epdf
Random FCS Flood Epdf
Pareto FCS Flood Epdf
Suboptimal No FCS Flood Susceptibility
Current FCS Flood Susceptibility
Random FCS Flood Susceptibility
Pareto FCS Flood Susceptibility
N
1.0
0
0 10 20 km

Figure 7:

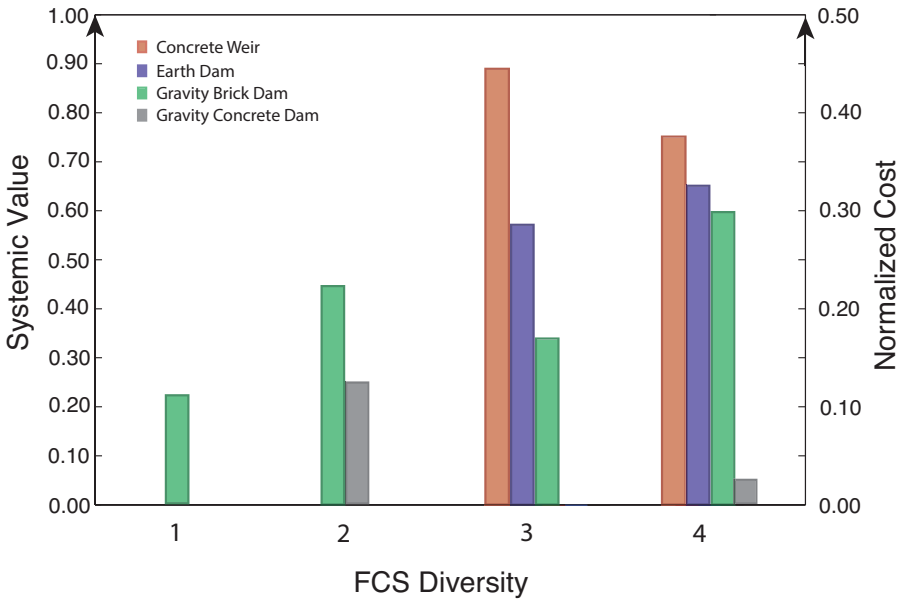

Concrete Weir
Earth Dam
Gravity Brick Dam
Gravity Concrete Dam
Systemic Value
Normalized Cost
FCS Diversity
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4

Figure 8:

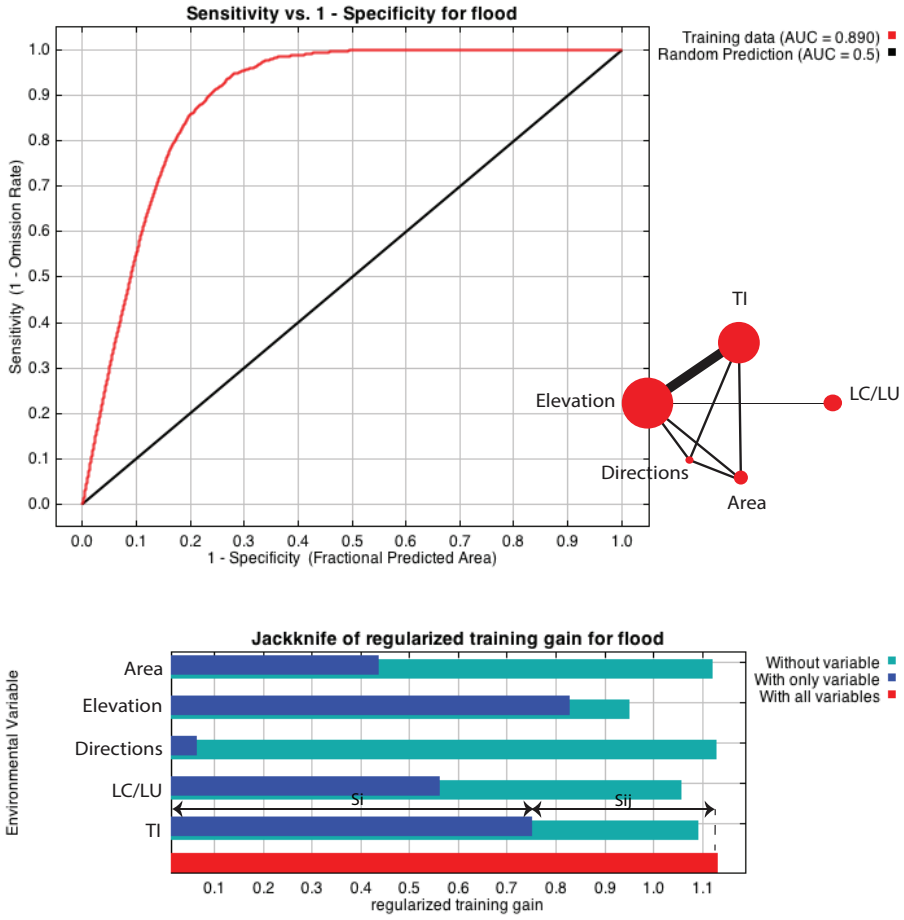

Sensitivity vs. 1 - Specificity for flood
Training data (AUC = 0.890)
Random Prediction (AUC = 0.5)
Sensitivity (1 - Omission Rate)
1 - Specificity (Fractional Predicted Area)
TI
Elevation
LC/LU
Directions
Area
Jackknife of regularized training gain for flood
Without variable
With only variable
With all variables
Environmental Variable
Area
Elevation
Directions
LC/LU
TI
Si
Sij
regularized training gain

Figure 9:

Supplementary Information

S1.1 Methodological Steps and Shared Model

In the following the operational steps to run the model are described as well as indications for the realization of the figures in the manuscript. Figure 1 shows the steps related to input/output relationships of the models. Compositions of plots (made separately in MaxEnt, GIS and Matlab) are made in Adobe Illustrator (individual plots are in the GitHub repository as well as all models, i.e. https://github.com/matteoconvertino/FloodTevereMaxEnt).

- Input factors of the MaxEnt model are created: these files are in the folders `sample` (flood occurrences) and `layers` (predictors or covariates, including FCS) of the GitHub repository of the project. These files are shown in Figures 3 and 4 using GIS visualizations (`TevereHumNat.mxd` is the master file of the GIS project).

- MaxEnt is used for calculations of flood susceptibility and related visualizations; the Github repository contains the outputs folders corresponding to different FCS scenarios. The inset in Figure 6, and plots in Figures 7 and 8 are made in MaxEnt. The flood susceptibility is used to calculated the epdf of flood size. The Matlab code `eprob.m` is used for generating and visualizing the epdf.

- the Portfolio Decision Model `PortfolioModel-tiberValue` is used recursively to predict the flood risk (or systemic value) and generate management solutions dependent on MaxEnt outputs and FCS plans. Figures 6, S2, S3, and Tables S1-S7 are generated using `PortfolioModel-tiberValue`.

- Global Sensitivity and Uncertainty analyses is performed; the Matlab code `GSUAplot.m` is used for reproducing Figure 9. The systemic value is determined (file `ValueDiversity`) as plotted in Figure 8). Other extra analysis can be performed such as the determination of the fractal dimension of flood patterns via the Matlab code `boxcount.m` that is used for the box-counting method; `ReadFlows.m` can be used for reading runoff data in some recorded locations and calculate a dynamics TI.

S1.2 Supplementary Results and Discussion

S1.2.1 Resolution Invariance of Predictions

The analysis of the resolution invariance of predictions, performed with 1, 10, 50, and 90% of background points (observed floods) is reported in Figure S2. The predicted flood suscepti-

bility is highly invariant even with very minimal information (i.e. 1% of background points). Background points are points that are used to infer the conditional pdf of flood occurrence on environmental layers to predict flood susceptibility. This resolution invariance proves the very high computational predictive power of the MaxEnt model. In terms of predictands (predicted pattern features), the absolute value of the predicted flood susceptibility is varying the most rather than the flood susceptibility geographical distribution.

Figure S3 show MaxEnt "response curves" which reveal the probability of an area being flooded conditional to the predictors. Note that this probability is normalized by the by probability of flood susceptibility in the random FCS scenario. The values of this ratio dependent on each predicting variable is shown along the y axis; this can be interpreted as the probability to have a site that is flooded across the basin conditional to the environmental variable considered and with respect the random flood susceptibility scenario. The shape of the curves matter rather than the value of the probability ratio that can be normalized in the same scale. The flood probability is higher for low elevation (from 0 to 500 m), for high topographic index, land use values that correspond to urban and crop areas, and all values of the drainage area (yet there is no higher probability of a flood to occur in areas with high drainage area). The latter results is quite surprising and it means that the flood risk, considering different positions along the river basin — geomorphologically speaking — is the same with the exception of the source streams for which the risk is much lower (as shown by the bottom right plot in Fig. S3). Other factors of course play a joint role in determining the flood susceptibility and at the end the flood risk (Figure 9 shows the interdependencies of these contributing factors). The higher probability of floods for low elevation areas is expected since these areas correspond to floodplains. High values of the flood probability are also expected for high values of the topographic index; a result that pinpoints to flat and wet areas on average.

S1.2.2 Portfolio Decision Model as Population Collective Brain

Beyond the financial origin of the the portfolio idea, the portfolio model has also roots and applications in neuroscience for the analysis of decision making processes and the design of treatment or technology that help individuals with neurological illnesses. In healthy individuals, different parts of the brain are activated in the decision making process to formulate at the end a potentially optimal decision when the accumulation of information is sufficient enough considering the problem at end and decision makers' comfort level. This is the state where the Value of Information (the same contemplated in this paper as the difference in payoff under different scenarios) is the maximum. In the brain multiple criteria are evalu-

ated and these criteria may be interdependent with each other. By focusing on this decision making origin of the portfolio we highlight the importance of adopting the model at the population scale where multiple actors, factors, and decision alternatives exist and need to be considered. Thus, in absence of anything that bring these elements together the portfolio can be considered as the "collective brain" of populations that helps in finding the solution of complex socio-ecological problems such as ecosystem design for managing flood risk. In this broader context the river network is an analog to the structural connectivity network of the brain while the set of criteria (also impacting the decision outcome) is the functional network used in the decision making process. The analogy is made to underline the high importance of iPDM in aiding decision making at the population level; information is complex and iPDM can gather that information and provide the optimal and often non-trivial solution considering all potential solutions available.

Supplementary Table Captions

Captions are reported below each Supplementary Table.

Supplementary Figure Captions

Figure S1. Box-counting relationship and scaling exponent stability. The top plot shows the relationship $N(r) \sim r^D$, where D is the Minkowski-Bouligand dimension that is a good estimate of the fractal dimension (or Hausdorff dimension) related to the power-exponent of epdf of floods ($\epsilon = D/2$ in Eq. 2 for the critical regime when system's aggregates are considered as patches (Convertino et al., 2013c)). The construction of a space-filling curve is based on recursive partitioning of the space into equal parts, or boxes. Theoretically, for pure fractals D is a function of the box-size. When the resolution of the decomposition, i.e. the number of times the partition process is applied, approaches infinity, the sequence of broken-line curves (with nodal points in elementary regions) converges uniformly to a continuous function ("space-filling" box count in red). The blue line is for discrete box counting.

Figure S2. Coarse grained Pareto optimal FCS scenarios. Here coarse graining is applied in relation to the percentage of background points (observed floods) to include in the MaxEnt inference. This is done to assess the stability (invariance) of predictions in relation to the amount of information included into the model. Purple points are the selected background points (among the white points) that are used to learn the relationships between environmental features and floods (see response functions in Fig S3). 1, 10, 50, and 90% of background points (purple points among all white points that are observed floods) are explored.

Figure S3. MaxEnt **response curves for predictors of floods**. The curves show the probability of an area being flooded as a function of elevation, topographic index, land use, and drainage area. These response curves are calculated based on the current FCS scenario. The log of the predicting variables is shown along the y axis.

Figure S4. Subbasins of the extracted drainage network. Subbasins of the extracted river network with a threshold on the drainage area of 5 km^2. These subbasins are used in the IPDM model to select optimal flood control structure plans.

Supplementary Figures

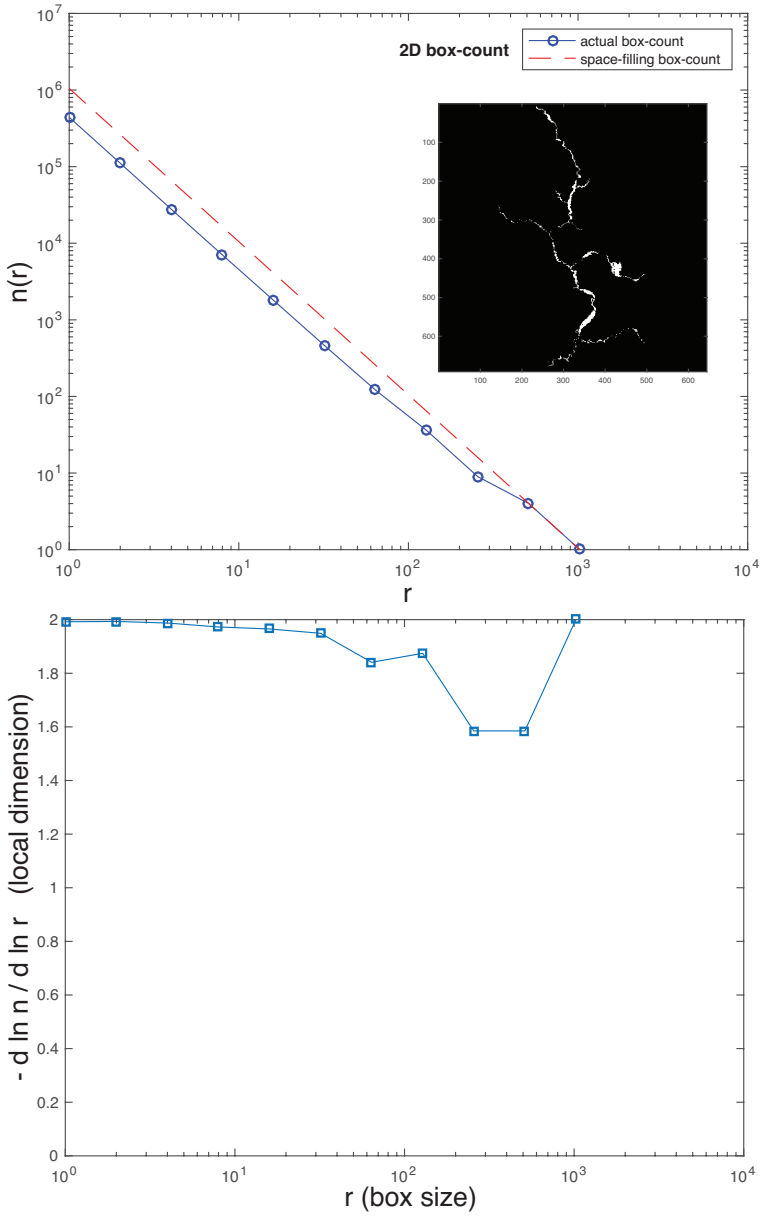

2D box-count
actual box-count
space-filling box-count
n(r)
r
- d ln n / d ln r (local dimension)
r (box size)

Figure S1:

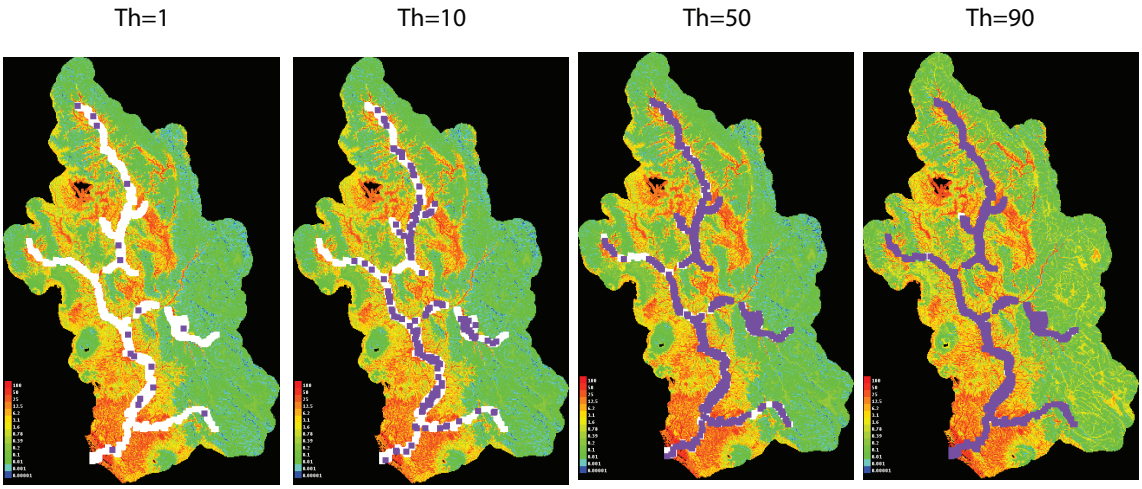


Figure S2:

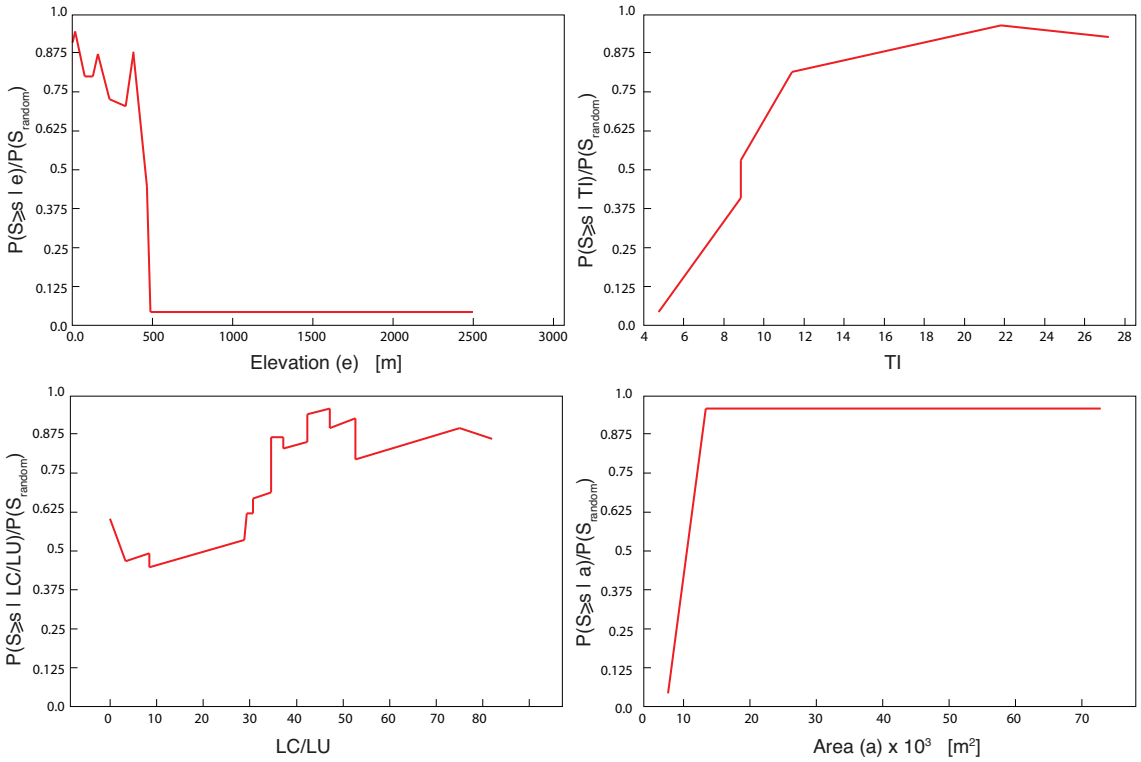

P(S≥s | e)/P(S_random)
Elevation (e) [m]
P(S≥s | TI)/P(S_random)
TI
P(S≥s | LC/LU)/P(S_random)
LC/LU
P(S≥s | a)/P(S_random)
Area (a) x 10³ [m²]

Figure S3:

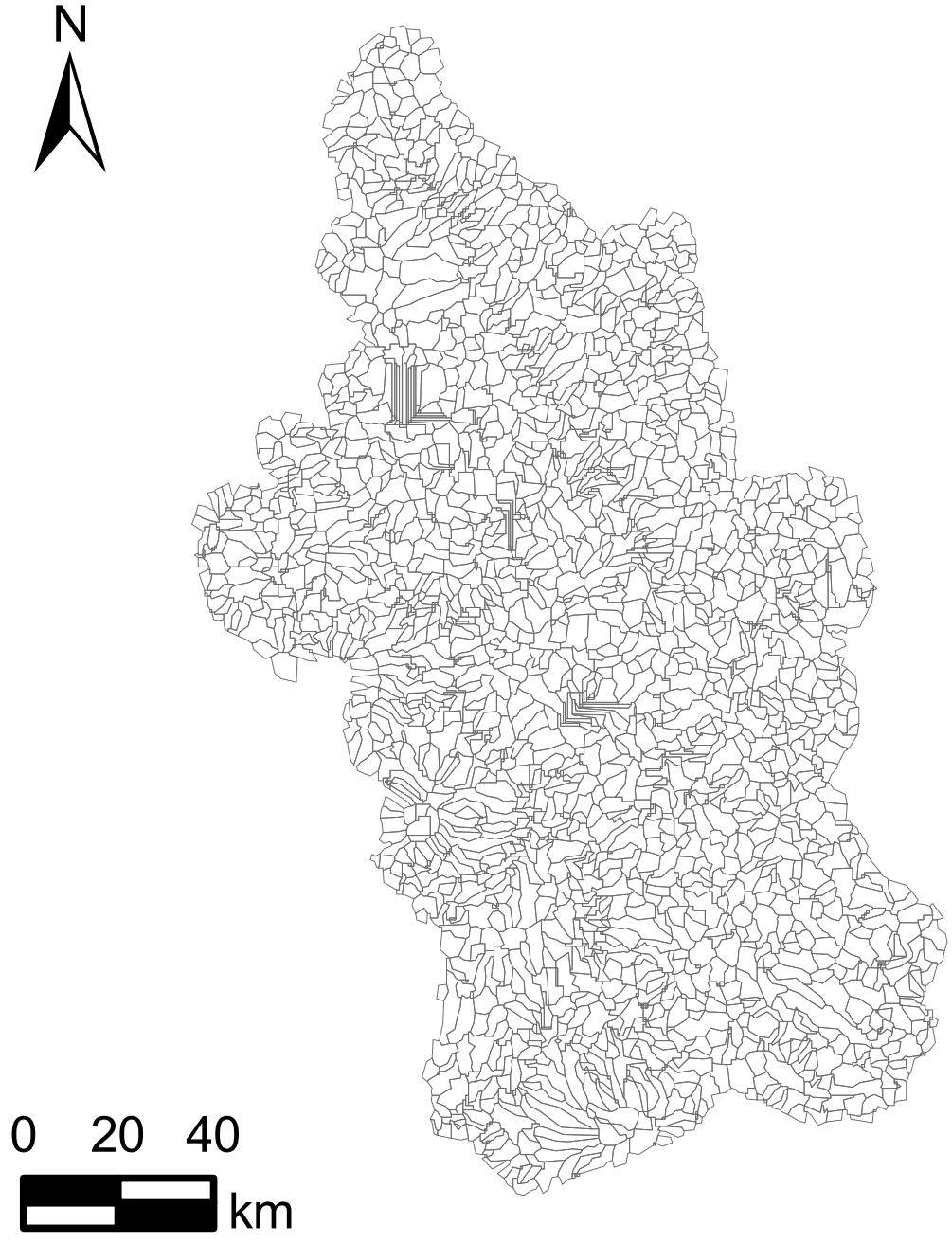


Figure S4: