Journal of Hydrology X 24 (2024) 100182

Contents lists available at ScienceDirect

Journal of Hydrology X

journal homepage: www.sciencedirect.com/journal/journal-of-hydrology-x

Optimizing sensor location for the parsimonious design of flood early warning systems

ARTICLE INFO

Keywords:
Flood early warning systems
Feature importance measures
Data driven flood forecasting
Continuous hydrological modelling

ABSTRACT

Flood early warning systems (FEWS) are effective means for saving human lives from the devastating impacts of extreme hydrological events. FEWS relies on hydrologic monitoring networks that are typically expensive and challenging to design. This issue is particularly relevant when identifying the most cost-efficient number, type, and positioning of the sensors for FEWS that may be used to take decisions and alert the population at flood risk. In this study, we focus on a widely recognized FEWS solution to analyze hydrological monitoring and forecasting performances expressed as discharge in various cross-sections of a drainage network. We propose and test a novel framework that aims to maximize FEWS performances while minimizing the number of sections that need instrumentation and suggesting optimal sensor placement to enhance forecasting accuracy. In the selected case study, we demonstrate through feature importance measure that only four sub-basins can achieve the same forecasting performance as the potential twenty-six cross-sections of the local hydrologic monitoring network. The operational dashboard resulting from our proposed framework can assist decision-makers in maximizing the performance and wider adoption of flood early warning systems across geographic and socio-economic scales.

1. Main

The increasing impacts of hydro-meteo extremes cause billions of damages and thousands of deaths every year (Rohde, 2023; Rodell and Li, 2023; Bevere and Remondi, 2022). Nuisance effects of flooding are particularly devastating and uncontrolled in riverine areas that often lack flood risk knowledge and proper awareness of the exposed population (Devitt et al., 2023). Earth observation and satellite-driven monitoring are tackling this knowledge gap, but observing and forecasting flooding from space cannot be the unique solution (Munasinghe et al., 2023). Remote sensing of uplands and small-scale river systems is currently impacted by the size and vegetation cover that avoid accurate observation of water dynamics. Global efforts are fostering to protect our society from floods. Holistic solutions integrating structural and non-structural measures are being financed and engineered to mitigate flood risk (Merz et al., 2021; Jongman, 2018). Flood protection infrastructures (e.g. levees, artificial storage) and nature-based solutions are recognized to be effective but not sufficient alone (Recanatesi and Petroselli, 2020). Non-structural measures, like hazard zoning, increasing resiliency, and awareness of flood-prone areas are fundamental as well as preparing, educating, evacuating, and floodproofing assets at risk.

Among non-structural measures, Flood Early Warning Systems (FEWS) represent a key technology for risk protection and mitigation of hydro-extreme events (Merz et al., 2021). Nonetheless, designing and effectively operationalizing FEWS is a significant scientific, economic, and socio-cultural challenge. From the scientific point of view, the distributed and uncertain nature of hydro-extremes and the complexity of simulating water-urban feature interactions determine serious issues in the engineering and positioning of sensors supporting FEWS. It is not only an issue of too few sensors as related to the complexity of the hydrologic phenomena. It is a more generalized issue of the positioning of available monitoring equipment about the need to simulate rapidly changing distributed hydrodynamic state variables (e.g. discharge, water levels) and maximizing flood forecasting performances. Moreover, limited financial resources exacerbate the scientific-technical issue considering that available economic resources require to balance the need to install more sensors with the effort to develop more sophisticated and accurate flood models. Furthermore, we posit that effective FEWS requires building and adapting to indigenous social and cultural settings. Trust and engagement of flood risk stakeholders, from river basin management entities to local municipalities, is a FEWS challenge for effective protection and awareness of the affected communities. As a result, FEWS implementation and the hydraulic engineering design problem related to monitoring systems across all scales, from coastal areas to mountain domains, challenge flood managers who are called to make important decisions on where and how to position flood sensors. This is a global and socio-economically relevant challenge and is still an open scientific topic to date.

FEWS operations and performances face several limitations related to the availability of hydrologic monitoring observations in terms of spatial coverage and temporal resolution, which hinder their widespread implementation. The existing sensor placement in operational FEWS is often based on subjective or logistic criteria (e.g. urban feature distribution; river segment spacing, bridge position) rather than relying on quantitative statistical frameworks. To date FEWS configurations often lead to either over-monitoring or under-monitoring. As a consequence, FEWS decision makers, driven by a lack of adequate financial resources to cover all river reaches, are unable to produce a parsimonious, yet efficient, warning system. Recent surveys confirm that FEWS at present is failing to protect society from flood risk due to inadequate hydrological network coverage and backup equipment, flood models that are not accurate or unable to produce reliable forecasts, and inappropriate technical skills, resources, and institutional support of

https://doi.org/10.1016/j.hydroa.2024.100182
Received 29 May 2024; Received in revised form 25 July 2024; Accepted 27 July 2024
Available online 28 July 2024

floodplain managers (Perera, 2019). The lack of FEWS is particularly dramatic in developing countries where the vulnerability of local economies and infrastructure is higher and the social impacts of floods more devastating. With such aim, the World Meteorological Organization (WMO) Severe Weather Forecasting Programme (SWFP) has been working since its inception in 2010 to strengthen the capacity of the National Meteorological and Hydrological Services (NMHSs) in developing countries for improved forecasts and warnings of extreme events (World Meteorological Organization, 2010).

In light of these challenges, this paper introduces a novel framework to address FEWS design problems based on enhanced understanding and consideration of distributed statistical specifications of basin response dynamics. The framework offers decision-making support for mitigating or resolving these design challenges and aims to improve the overall performance of FEWS. Importantly, it is applicable in a wide range of basin conditions, from completely ungauged to fully instrumented basins, and can be applied throughout various stages of FEWS development, including initial design, maintenance, or assessment of operational systems.

The proposed framework is based on the following four steps (see Fig. 1). (1) A terrain analysis procedure is used to disaggregate the investigated basin into n sub-basins and identify outlets and confluence nodes. (2) A semi-distributed hydrological model is applied to generate synthetic rainfall and runoff scenarios. A large flood event dataset is simulated, and a sub-sample (with m dimension) is selected. Following this step, m flood events characterized by $n + 1$ hydrographs (n sub-basins plus the whole outlet) are available for training and testing machine learning techniques. (3) The forecast performances and optimal early warning lag times are then evaluated. (4) Feature Importance Measures (FIMs) are applied to the operative dataset, to identify the most influential nodes. These nodes provide the same forecast performances to the whole set of sub-basins, allowing for a parsimonious FEWS design that suggests the minimal number of nodes needed for the machine learning tool to develop timely flood warnings effectively.

As a case study, we select the most challenging option, a medium-sized basin, with a fully ungauged drainage network and without a preexisting FEWS. We illustrate the peculiarities of the proposed framework and the possible outcome available for the decision-makers to optimize the FEWS design, while in the discussion we point out some general possible applications, limitations, and future perspectives.

2. Identifying the most influential sub-basins

The terrain analysis method, characterizing the first step of the proposed framework, serves to the identification of the sub-basins that will be given as input for the semi-distributed hydrological model. Fig. 2 reports three specific moments of the analysis, showing the digital elevation model (Fig. 2a) on which the terrain analysis procedure is applied for tailoring the 21 sub-basins (Fig. 2b) and the 27 nodes (sub-basins outlets and confluence nodes). This first river basin terrain hydrologic analysis, and the resulting sub-basins and nodes characterization, also support the creation of the list of potential optimal monitoring node-basin sub-set that will maximize FEWS performances. The proposed case study applies the procedure by enforcing a minimal sub-basin contributing area equal to 15 km^2 to guarantee an appropriate application of the rainfall-runoff model. In the middle plot of Fig. 2 the raingauges used for calibrating the rainfall simulation model (CoSMoS) are identified. CoSMoS generates seven time series of 1000 years duration at 15 min of time resolution preserving spatial and temporal correlations.

Using the land use input and the synthetic rainfall time series, the COSMO4SUB hydrological model is applied on the 21 sub-basins providing runoff time series (again 1000 years at 15 min of time resolutions). Propagating them in the drainage network, runoff time series are also available at the 6 confluence nodes. Details on the framework application and the resulting dataset characteristics are provided in the Method Section and the Supplementary material (Sections S2 and S3).

The available runoff time series were filtered out to sample a dataset of 5826 flood events at the outlet and identify 1193 events with peak discharge in the range 100–200 m^3/s. This flood peak range corresponds to critical conditions for the selected FEWS case study (see case study description section for further details). Three early warning time lags (defined here as interval ahead forecast) are investigated: 6, 8, and 10 h spanning from a minimum warning time to a maximum time span associated to the time needed by the entire basin to transfer rainfall to the outlet (i.e. maximum flood hydrograph base time). To ensure a robust forecast analysis, avoiding the bias due to the heterogeneity of hydrological response or the partial contribution of sub-basins, only

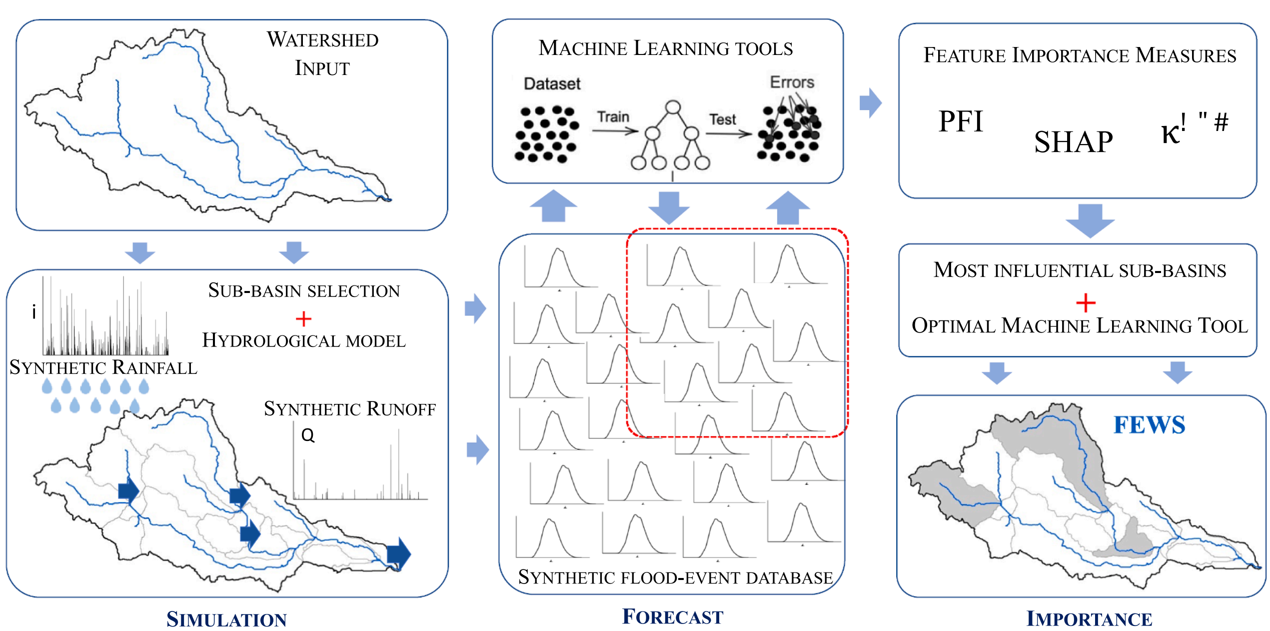

Fig. 1. Proposed framework for the design of parsimonious Flood Early Warning System. From left to right: simulation of synthetic flood event database; forecast with data-driven tools; selection of the most influential sub-basins. Permutation Feature Importance – PFI, Shap feature importance – SHAP, and Derivative-based importance – k^{ALE} are the acronyms of the three FIMs suggested in the framework.

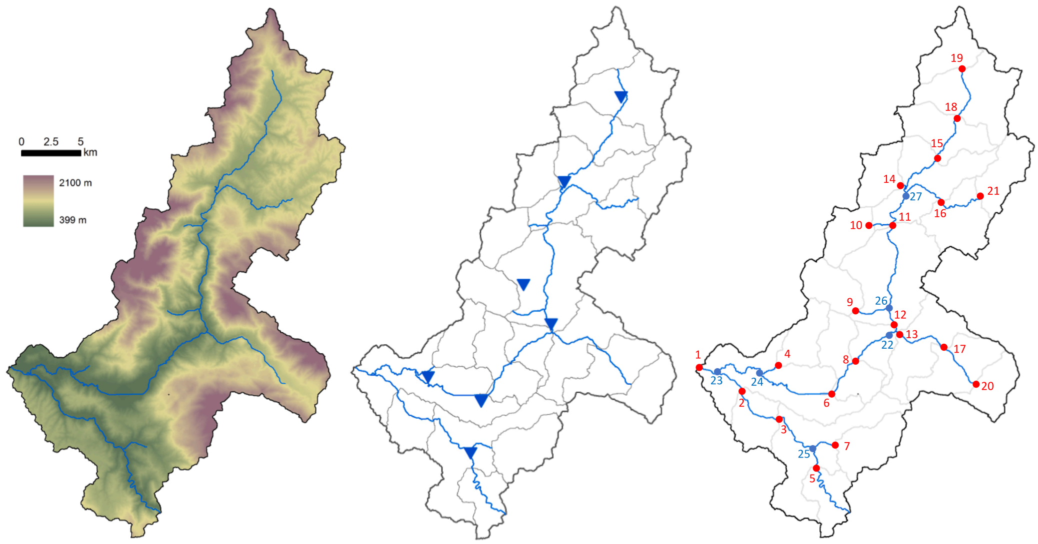

Fig. 2. Step 1 of the proposed framework: sub-basin selection. From the Digital Elevation Model (left panel) the implemented terrain analysis procedure identifies sub-basins (grey contour lines in the central panel), outlets (red points), and confluence nodes (blue points in the right panel). In the central panel, raingauges location is shown (blue triangles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

flood events with a duration greater than 36 h are considered, limiting the dataset to 547 events. Four machine learning (ML) models are employed (Linear model – LM, Least Absolute Shrinkage and Selection Operator Regression – LASSO, Multivariate Adaptive Regression Splines model – MARS, and Random Forest – RF) using 80 % of the dataset as the training set and 20 % as the testing set. Three performance indexes suggest that the RF outperforms the other ML models for each time lag providing a promising relative peak error equal to 7 %. Details on performances, comparisons among benchmarks and forecasted values are provided in Supplementary material (Section S1).

Three FIMs (Permutation Feature Importance (Breiman, 2001) – PFI, SHapley Additive explanations feature importance (Shapley, 1952; Lundberg and Lee, 2017) – SHAP, and derivative-based importance measure (Cappelli, 2023) – κ^{ALE}) are applied to rank the sub-basins

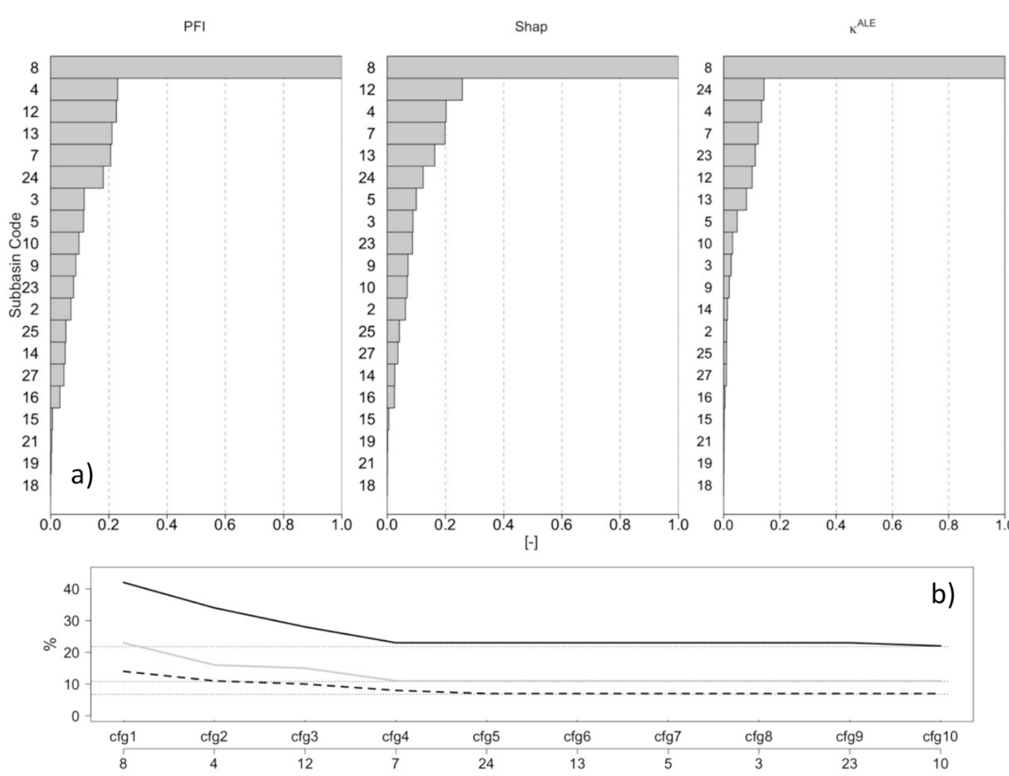

Fig. 3. (a) Estimates of the three feature importance measures (PFI, SHAP, k^{ALE}) for time lag equal to 6 h sorted in descending order. Values are normalized to facilitate the identification of the most important sub-basins. (b) Performance indices (black line: $MARE_{Adj}$; gray line: RPE; dashed line: $BIAS_{Adj}$) for different RF configurations ("cfg"). "cfg1" means that the RF is applied with only the sub-basin 8 as input. "cfg2" means that the RF is applied with the sub-basins 8 and 4 as input, and so on. The sequence of sub-basin codes in abscissa is the ranking average result. Horizontal lines refer to the associated index values estimated using all 26 nodes as input.

according to their influence on the RF using a time lag equal to 6 h. These measures are selected among the available importance measures because they turned out to be the most robust in hydrological applications (Cappelli and Grimaldi, 2023). Indeed, such measures appear to be less sensitive compared to the other measures to non-Gaussian distributions, presence of auto- and cross-correlations, and presence of collinearity. Since this latter could have a significant impact on the final results, we removed the sub-basins (sub-basin codes: 17, 26, 11, 6, 20, 22) whose signal is strongly correlated with that recorded at the outlet ($\rho > 0.9$), keeping only those that are physically disconnected from each other. Fig. 3 shows the estimates of the three FIMs using the RF predictions. To facilitate comparison between FIMs, the estimates are normalized (i.e., the values are within the range of 0 and 1). These estimates are sorted in descending order in graph a) of Fig. 3. This allows us to easily identify the associated importance ranking resulting from each FIM used. Calculating the average ranking we obtain that the sub-basins 8, 4, 12, and 7 are the most influential. As further evidence of their role and effect on the ML model prediction, in graph b) of Fig. 3, we report the behavior of the RF performance (expressed in terms of Mean Absolute Relative Error $MARE_{Adj}$, the Relative Peak Error RPE, and the Bias Adjusted $BIAS_{Adj}$) exhibited using an incremental approach. Specifically, we construct multiple RF configurations (cfg) by exploiting the average ranking obtained. The first configuration includes only the most influential sub-basin; the second includes only the first two most influential sub-basins, and so on. This graph shows that the minimum error (i.e., the best performance) is reached with only the first four sub-basins. Details on the FIMs procedure here adopted and results related to the 8- and 10-hour Lag are reported in the Method Section and Supplementary materials (Section S5).

3. FEWS monitoring decision-maker dashboard

The operational value of the proposed framework was summarized by designing and producing a dashboard to process modelling results into useful information for the decision-maker. The dashboard allows to easily identify the optimal FEWS design strategy with specific regard to the most important subbasins for FEWS optimization. Fig. 4 shows an example of proposed framework results and dashboard infographics. The three basin maps refer to the three different time lags and show (in gray) the most influential nodes and their contributing areas. It is evident, as expected, that by increasing the time lag the most influential nodes propagate upstream. The colored bars suggest the Relative Peak Errors, the Mean Absolute Relative Errors, and the Bias Adjusted values vary as the time lag varies. Moreover, by increasing the early warning time, performance significantly deteriorates. In the present case study, the decision-maker can reach the following conclusions: (a) acceptable forecast performance is guaranteed only for the lowest lag (6 h), (b) for the optimal lag, instrumenting only 3 or 4 cross-sections corresponding to the nodes 8, 4, 12, and 7 allows one to maximize the FEWS performance.

4. Discussion

The proposed framework enables the design of parsimonious FEWS, as it discriminates the role of each sub-basin by identifying the most influential nodes for installing instrumentation. The proposed case study represents the most challenging condition of an ungauged region where monitoring network and runoff data for calibration are not available. However, the results are encouraging. Indeed, the framework can provide useful suggestions for the decision-makers for the FEWS design. However, the usefulness of the proposed procedure is not limited to this challenging situation, but can contribute also to the following cases:

(a) A well-monitored large basin without an existing FEWS. In this case, a more effective semi-distributed or distributed hydrological model can be applied (through calibration) to allow a more realistic synthetic flood event dataset. The latter, besides being crucial for the feature importance analysis and consequently for selecting the most influential nodes, could be adopted for a pre-calibration of the machine learning tool. Indeed, the usual bottleneck of data-driven FEWS forecasting tools is the absence of a large dataset of observed flood events on several basin nodes. In the most favorable cases, the available dataset, although rich, is very heterogeneous concerning the hydrograph peak discharge. This could reduce the ML forecast performance since the ML model is sensitive to outliers (Breiman, 2001). If the hydrological-hydraulic model is well calibrated, the simulated dataset should be enough realistic for calibrating the ML models and the

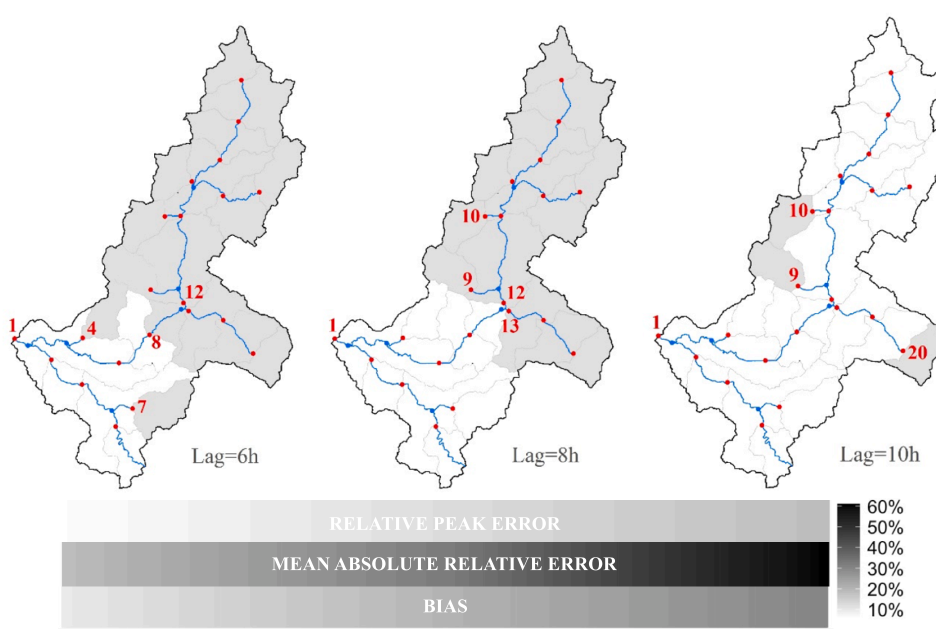

Fig. 4. FEWS monitoring Decision-maker dashboard. In the top panels, identification of the most influential sub-basins is provided for 6-, 8-, and 10-hour time lags. At the bottom, the colored bars correspond to the values of the three performance indices obtained by varying the time lags. Red numbers are the sub-basin codes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

framework could be ready to be operationally applied with the observed data. Such potential will be easily verified in future research by considering several modeling options in fully gauged conditions and comparing the simulated and observed datasets.

(b) A basin with an existing FEWS with a poor monitoring network. This would improve the calibration by making more realistic the simulated dataset. The framework dashboard would provide useful feedback for evaluating whether the current network is poor, how and where it could be improved by adding further instrumented nodes, or simply suggesting their new optimal location.

(c) A basin with an existing FEWS with a rich monitoring network. In such case, useful feedback could be suggested to eventually reduce the number of instrumented nodes and/or relocate some of them to increase the FEWS performance. Moreover, it would be effective to invest more resources in monitoring the most influential sub-basins to ameliorate their rainfall-runoff forecast model that, in cascade, will increase the forecast time lag.

(d) As mentioned in the premises, the proposed framework has also a positive effect on enhancing the knowledge of the basin response. Knowing the most influential sub-basins contributing to the global basin response will help to investigate the physical reasons for exploring some hydro-geomorphometric proxies for such behavior.

5. Methods

5.1. Case study

As a case study, we selected the Velino river basin, in central Italy, a left tributary of the Tiber river, i.e., the river flowing through the city of Rome. The outlet cross section is located in Rieti town, downstream the conjunction with the Salto river. The total contributing area is equal to 518.2 km^2. The geomorphometric analysis useful for the hydrological model implementation refers to the COPERNICUS (Copernicus, 2023) Digital Elevation Model (DEM) at 25 m resolution, while the land cover information is derived from the CORINE (CORINE project, 2000) Program. The investigated area exhibits a topography characterized by rolling hills, valleys, and lakes ranging from 400 m to 2100 m. Land use predominantly encompasses agricultural activities, including vineyards, olive groves, and cereal cultivation, interspersed with settlements. This region also features forested areas and natural reserves, fostering biodiversity conservation and recreational opportunities. As described in the previous Sections, the first framework step concerns the identification of the sub-watersheds that will serve as input for the semi-distributed hydrological model. Based on the methodology described in the next Section, the case study was discretized in 21 sub-basins, characterized by contributing areas ranging between 13 and 518 km^2 (average value equal to 24.6 km^2), by concentration times (estimated using Giandotti's formula (Giandotti, 1934)) between 1 and 5 h (average value equal to 2.1 h), and by Curve Number (CN) values (NRCS look-up tables (McCuen, 1998)) between 60.7 and 64.6 (average value equal to 62.4). See the Section S6 of Supplementary materials for more details of the investigated area.

In order to calibrate the rainfall simulation model, we selected 20 years of contemporaneous rainfall observations at 15-minute time resolutions (obtained from Lazio Region database (Region, 2023). Technical reports available at Autorità di Bacino Distrettuale dell'Italia Centrale (ADBAC (Papalexiou, 2018) suggest that the discharge range of 100–200 m^3/s corresponds to the warning water depth when the inundation is expected to begin at the basin outlet cross-section.

5.2. Multisite Rainfall simulation model using CoSMoS

Multisite rainfall generation poses significant challenges, particularly due to limitations in observational data, especially at fine spatiotemporal scales. These limitations include missing data at various stations and quality issues arising from duplicated values (ties) and mismeasurements. Furthermore, complex cross-correlation properties (such as spatial non-stationarity) and autocorrelation characteristics at each station can create feasibility issues in calibrating a multisite rain model (Papalexiou et al., 2018). These challenges are compounded by seasonal and spatial variability, wherein the probability distribution describing rainfall varies with both season and location. Collectively, these challenges can compromise the accuracy of analyses, introducing biases into the statistical properties of the simulated rainfall. Inaccuracies at fine scales can be magnified at aggregated scales, and inaccuracies at individual stations can have ripple effects at catchment scales, resulting in disparities in total catchment rainfall.

Here, the multisite rainfall simulation at the 15-minute resolution is conducted using the CoSMoS framework (Papalexiou, 2018; Papalexiou et al., 2018; Papalexiou and Serinaldi, 2020; Papalexiou et al., 2021; Papalexiou, 2022; Papalexiou et al., 2023; Papalexiou et al., 2021). Specifically, the mixed-uniform CoSMoS is applied (see reference 25 for details), with an additional adjustment to calibrate daily rainfall at the catchment. The simulation explicitly reproduces the marginal distribution, accounting for the probability of dry, at each station and season, as well as the cross and autocorrelation properties. To streamline the reader's understanding and avoid redundancy in detailing modeling procedures, which are described in the referenced works, we provide a concise overview of the scheme's modeling steps. In implementing this scheme, a set of sequential steps is executed on a monthly basis for each station:

Original Data Evaluation: the original 15-minute precipitation time series are examined to assess distribution and correlation properties.

Distribution Transformation: the observed precipitation time series are transformed to adhere to a mixed-uniform marginal distribution.

Correlation Analysis: the mixed-uniform time series are used to calculate empirical spatial correlations between all pairs of stations and estimate the autocorrelation structure at each station.

Parametric Correlation Description: the mixed-uniform correlations are adjusted through the mixed-uniform correlation transformation function to represent those of a process with Gaussian marginals. Then, parametric spatial and temporal correlation structures are fitted to the estimated empirical Gaussian correlations.

Time Series Generation: time series are generated using a vector autoregressive model, incorporating the estimated Gaussian spatial and temporal correlation structures.

Conversion to 15-Minute Precipitation: the generated Gaussian time series are converted back to 15-minute precipitation by using the marginal distributions describing the original data.

As mentioned earlier, the generated rainfall is adjusted to align with daily rainfall in the catchment. Multisite simulations face challenges in precisely mimicking observed features due to mathematical limitations in model settings (such as positive definite correlation matrices). In any case, observed characteristics such as distributions and correlations do not perfectly mirror the actual process due to random fluctuations. The use of parametric functions to represent probabilities and correlations makes multisite schemes technically feasible but may potentially misrepresent local variations. Therefore, even if the fine-scale simulation reproduces desired characteristics when aggregated over space and time, it might not match the observed process at larger scales. This mismatch can impact hydrologic models, causing more or less intense events to be simulated at the catchment level, resulting in larger or lower floods. To align with the distribution of daily rainfall in the catchment, we aggregate the simulated and observed 15-minute precipitation at the daily scale over the catchment. Then, we identify the probability distributions $G_o(x)$ and $G_s(x)$ for the observed and simulated daily nonzero precipitation over the catchment, respectively. The adjusted daily simulated values $\widetilde{x}_s$ are then given by $\widetilde{x}_s(t) = G_o^{-1}(G_s(x_s(t)))$. Finally, the 15-min simulation is adjusted by multiplying the 15-min values of each day by a corresponding scaling factor $\widetilde{x}_s(t)/x_s(t)$ of each day. Section S2 in the Supplementary Material provides an assessment of the simulated rainfall time series.

5.3. Semi-distributed continuous rainfall-runoff model

The 21 sub-basins and 6 confluence nodes are selected by applying an automatic GIS-based algorithm developed for the present study to generalize the procedure for any river basin. The input DEM has been hydrologically reconditioned using the official digitized stream network released by the regional authority (Lazio Region). Standard terrain analysis algorithms (Soille, 2004; Jenson and Domingue, 1988) are applied to generate the filled DEM, flow direction, flow accumulation, and stream network grids. The algorithm generates different nodes along the DEM-based extracted stream network starting from a user-defined threshold area (*ThrA)* for headwater selection, a minimum distance (*d)* or stream confluence identification, and a minimum contributing area (*minA*) for the sub-basin picking. The *ThrA, d,* and *minA* parameters are here fixed respectively to 15 km^2, 5 km, and 15 km^2. The selected *ThrA* is consistent with the visible stream features from satellite imagery. The *d* and *minA* parameters were considered a compromise among having a minimum number of sub-basins, a significant contribution to the flood wave propagation, and a minimum distance among contiguous outlets consistent with the time step of the flow propagation model. The procedure allows a 20 % tolerance for the sub-basins threshold area to adapt for specific configurations of the stream network (e.g. instead of having one big subbasin along a stream with almost doubling the threshold area, it is better to have two subbasins divided by a node where one of them is slightly smaller than the threshold value).

Once the 21 sub-basins are selected and the rainfall synthetic time series are simulated, a continuous hydrological rainfall-runoff model is applied to generate the runoff time series at each outlet. In the present work, we refer to the COSMO4SUB (Continuous Simulation Model For Small and Ungauged Basin) rainfall-runoff model (Grimaldi et al., 2012; Grimaldi et al., 2021). The original model version consists of four main steps: (1) rainfall scenario simulation, (2) excess rainfall estimation, (3) excess rainfall-runoff transformation, and (4) design simulation strategy; in the present application, we refer only to steps 2 and 3.

Given a sub-basin and the synthetic rainfall time series generated using the raingauge observation nearest to the center of mass of the basin area, step 2 of the COSMO4SUB model is applied. A mixed procedure CN4GA (Curve Number for Green-Ampt (Grimaldi et al., 2013)) allows for estimating the excess rainfall time series using the NRCS-CN method to quantify excess rainfall depth at the event scale and the Green-Ampt equation to distribute this depth within the rainfall event. The Green-Ampt parameters are automatically estimated by the procedure constraining the equation to give the same excess rainfall depth as estimated by NRCS-CN, preserving the procedure as calibration-free. The only parameter needed to be specified by the user is the CN value of the NRCS-CN method, which is automatically estimated using look-up tables linking CN with soil type and land use. An important point to note is that, compared to the event-based procedure, an extra parameter called the separation time (*Ts*) is required. *Ts* is used for identifying isolated rainfall events and indicates the duration of the dry period (or nearly dry period) required for the initial abstraction to become effective again. A previous study (Grimaldi et al., 2021) examined the sensitivity of this parameter (in the range of 18–30 h) using a long synthetic rainfall time series, and it was found that its effect on the design hydrograph is minimal. A reasonable value of 24 h for *Ts*, as also suggested by NRCS-CN implementation, is considered appropriate and it is employed here.

The step 3 of the model involves the excess rainfall-runoff transformation using the WFIUH model, which is a modified version of the traditional IUH (Instantaneous Unit Hydrograph) approach, specifically designed for ungauged conditions (Grimaldi et al., 2012). This approach allows for the optimization of available digital topography information by enforcing hydrogeomorphic processes that represent governing factors of floodplain generation dynamics (Annis et al., 2019; Nardi et al., 2018; Nardi et al., 2019). The WFIUH model uses the travel time distribution of watershed DEM cells as the IUH definition, providing a calibration-free IUH. The flow path, hillslope-channel discrimination, and velocity estimation on hillslopes are easy steps that are useful in quantifying the flow time or travel time distribution. Flow paths have been determined using a classic D8 approach (Jenson and Domingue, 1988). Channel points have been selected (Tarboton et al., 1991) when having a total contributing area greater than 1 km^2. Hillslope velocities have been assigned linking them to local slope and land cover (Grimaldi et al., 2010). The channel velocity has been automatically quantified such that the center of mass of WFIUH corresponds to the basin lag time, which is estimated as 60 % of the basin concentration time (Petroselli and Grimaldi, 2018).

The result of the COSMO4SUB step 3 implementation provides a synthetic runoff time series (1000 years at 15 min of time resolution) in each sub-basin.

Given the runoff information in the 21 sub-basins, a simplified hydraulic propagation model is implemented and applied in order to quantify the runoff time series in the residual 6 confluence nodes, making available the whole hydrological information useful for the FEWS analysis. The combination and propagation of the sub-basin outlet time series is carried out with the Muskingum-Cunge method by imposing a spatial discretization such that the Courant number approximately equals 1 for the mean flow rate at the input node (Ponce, 2014).

5.4. Feature importance measures (FIMs)

FIMs are agnostic tools for any supervised ML model (i.e., linear regression, random forests, gradient boosting, neural networks). ML importance measures are typically defined using the predictive performance of an ML model. These techniques help one to understand the importance of input variables in a predictive ML model, shedding light on the underlying relationships between the input variables and the target variable. In the present work, we employ three ML feature importance measures: Permutation Feature Importance PFI (Breiman, 2001), Shapley Additive explanations feature importance SHAP (Shapley, 1952; Lundberg and Lee, 2017); and derivative-based importance measure κ^{ALE} (Cappelli, 2023).

PFI is the most well-known importance measure in ML. It assesses a feature as important based on the change in the predictive ability of the ML model after that a feature has been randomly permuted.

SHAP is a feature importance measure that relies on the notion of Shapley value (Shapley, 1952). Such a method assigns the importance based on the average absolute contribution of each feature in calculating the predicted value for all data observations.

The third importance measure is defined using the ALE-plot design (Apley and Zhu, 2020). ALE-plots are powerful visualization tools that provide insights into the relationship between a feature and the predicted outcome of an ML model. Cappelli (2023) proposes a derivative-based importance measure κ^{ALE} that quantifies the impact of small changes of a specific feature in the ML model response on average.

Table 1 summarizes feature importance measures used in this work while a detailed description is available in Section S4 of the Supplementary material.

CRediT authorship contribution statement

Salvatore Grimaldi: Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition. **Francesco Cappelli:** Conceptualization, Methodology, Writing – original

Table 1
Equations and references of three feature importance measures applied in this work.

Methods	Definition
Permutation importance measure	$\mathrm{PFI}_j = \mathbb{E}\left[L\left(Y, \widehat{g}\left(X_j^{\pi}, \mathbf{X}_{-j}\right)\right)\right] - \mathbb{E}\left[L\left(Y, \widehat{g}\left(X_j, \mathbf{X}_{-j}\right)\right)\right]^{(11)}$
SHAP importance measure	$\mathrm{SHAP}_j = \frac{1}{n}\sum_{i=i}^{n}\sum_{K\subseteq D\setminus i}\frac{\lvert K\rvert!(\lvert D\rvert - \lvert K\rvert - 1)!}{\lvert D\rvert!}\left[\widehat{g}_K\left(\mathbf{x}_{K\cup\{i\}}\right) - \widehat{g}_K\left(\mathbf{x}_K\right)\right]^{(13)}$
Derivative-based importance measure κ^{ALE}	$\kappa_j^{ALE} = \frac{1}{K}\sum_{k=0}^{K-1}E\left[\frac{\widehat{g}\left(z_j^k, \mathbf{x}_{-j}^{(i)}\right) - \widehat{g}\left(z_j^{k-1}, \mathbf{x}_{-j}^{(i)}\right)}{z_j^k - z_j^{k-1}}\right]^2\frac{\sigma_{X_j}^2}{\sigma_y^2}$ (14)

draft, Writing – review & editing. **Simon Michael Papalexiou:** Conceptualization, Methodology, Writing – review & editing. **Andrea Petroselli:** Methodology, Writing – review & editing. **Fernando Nardi:** Conceptualization, Methodology, Writing – review & editing. **Antonio Annis:** Methodology, Writing – review & editing. **Rodolfo Piscopia:** Methodology, Writing – review & editing. **Flavia Tauro:** Methodology, Writing – review & editing. **Ciro Apollonio:** Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Salvatore Grimaldi carried out this work within the PRIN2022 – DeHySi Project – DEsign HYdrologic Simulation – Ministry of University and Research of Italy – n. 2022NBXJSL, 2023-2025.

Salvatore Grimaldi, Andrea Petroselli, and Flavia Tauro carried out this work within the project ECS 0000024 Rome Technopole, -CUP B83C22002820006, PNRR Mission 4 Component 2 Investment 1.5, funded by the European Union - Next GenerationEU.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.hydroa.2024.100182.

References

Annis, A., Nardi, F., Morrison, R.R., Castelli, F., 2019. Investigating hydrogeomorphic floodplain mapping performance with varying DTM resolution and stream order. Hydrol. Sci. J. 64 (5), 525–538.
Apley, D.W., Zhu, J., 2020. Visualizing the effects of predictor variables in black box supervised learning models. J. R. Stat. Soc. Ser. B: Stat. Methodol. 82 (4), 1059–1086.
Bevere, L., Remondi, F., 2022. Natural catastrophes in 2021: the floodgates are open. Swiss Re Institute. Available online: https://www. swissre.com/institute/research/sigma-research/sigma-2022-01.html (accessed on, 30).
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Cappelli, F., 2023. Interpretability of Machine Learning with Hydrological Applications (Doctoral dissertation, Ph. D. thesis). Università Bocconi.
Cappelli, F., Grimaldi, S., 2023. Feature importance measures for hydrological applications: insights from a virtual experiment. Stoch. Env. Res. Risk A 37 (12), 4921–4939.
Copernicus D-M – Global and European Digital Elevation Model (COP-DEM), 2023. https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model (accessed December 2023).
CORINE project, 2000. CORINE (Coordination of Information on Environment) Database, a key database for European integrated environmental assessment. Programme of the European Commission, European Environmental Agency (EEA). https://land.copernicus.eu/en/products/corine-land-cover (accessed December 2023).
Devitt, L., Neal, J., Coxon, G., Savage, J., Wagener, T., 2023. Flood hazard potential reveals global floodplain settlement patterns. Nat. Commun. 14 (1), 2801.
Giandotti, M., 1934. Previsione delle piene e delle magre dei corsi d'acqua. Istituto Poligrafico Dello Stato 8, 107–117.
Grimaldi, S., Petroselli, A., Alonso, G., Nardi, F., 2010. Flow Time estimation with variable hillslope velocity in ungauged basins. Adv. Water Resour. 33 (10), 1216–1223.
Grimaldi, S., Petroselli, A., Serinaldi, F., 2012. A continuous simulation model for design-hydrograph estimation in small and ungauged watersheds. Hydrol. Sci. J. 57 (6), 1035–1051.
Grimaldi, S., Petroselli, A., Nardi, F., 2012. A parsimonious geomorphological unit hydrograph for rainfall-runoff modelling in small ungauged basins. Hydrol. Sci. J. 57 (1), 73–83.
Grimaldi, S., Petroselli, A., Romano, N., 2013. Curve-number/green-Ampt mixed procedure for streamflow predictions in ungauged basins: parameter sensitivity analysis. Hydrol. Process. 27 (8), 1265–1275.
Grimaldi, S., Nardi, F., Piscopia, R., Petroselli, A., Apollonio, C., 2021. Continuous hydrologic modelling for design simulation in small and ungauged basins: a step forward and some tests for its practical use. J. Hydrol. 595, 125664.
Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. Photogramm. Eng. Remote Sens. 54 (11), 1593–1600.
Jongman, B., 2018. Effective adaptation to rising flood risk. Nat. Commun. 9 (1), 1986.
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Proces. Syst. 30.
McCuen, R.H., 1998. Hydrologic Design and Analysis. Prentice Hall, New Jersey.
Merz, B., Blöschl, G., Vorogushyn, S., Dottori, F., Aerts, J.C., Bates, P., Macdonald, E., 2021. Causes, impacts and patterns of disastrous river floods. Nat. Rev. Earth Environ. 2 (9), 592–609.
Munasinghe, D., Frasson, R.P.d.M., David, C.H., et al., 2023. A multi-sensor approach for increased measurements of floods and their societal impacts from space. Commun. Earth Environ. 4, 462. https://doi.org/10.1038/s43247-023-01129-1.
Nardi, F., Morrison, R.R., Annis, A., Grantham, T.E., 2018. Hydrologic scaling for hydrogeomorphic floodplain mapping: insights into human-induced floodplain disconnectivity. River Res. Appl. 34 (7), 675–685.
Nardi, F., Annis, A., Baldassarre, G.D., Vivoni, E.R., Grimaldi, S., 2019. GFPLAIN250m, a global high-resolution dataset of earth's floodplains. Sci. Data 6, 309.
Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: preserving marginal distributions, correlation structures, and intermittency. Adv. Water Resour. 115, 234–252.
Papalexiou, S.M., 2022. Rainfall generation revisited: introducing CoSMoS-2s and advancing copula-based intermittent time series modeling. Water Resour. Res. 58, e2021WR031641.
Papalexiou, S.M., Serinaldi, F., Strnad, F., Markonis, Y., Shook, K., 2021. CoSMoS: Complete Stochastic Modelling Solution.
Papalexiou, S.M., Serinaldi, F., 2020. Random fields simplified: preserving marginal distributions, correlations, and intermittency, with applications from rainfall to humidity. Water Resour. Res. 56, e2019WR026331.
Papalexiou, S.M., Markonis, Y., Lombardo, F., AghaKouchak, A., Foufoula-Georgiou, E., 2018. Precise temporal disaggregation preserving marginals and correlations (DiPMaC) for stationary and nonstationary processes. Water Resour. Res. 54, 7435–7458.
Papalexiou, S.M., Serinaldi, F., Porcu, E., 2021. Advancing space-time simulation of random fields: from storms to cyclones and beyond. Water Resour. Res. 57, e2020WR029466.
Papalexiou, S.M., Serinaldi, F., Clark, M.P., 2023. Large-domain multisite precipitation generation: operational blueprint and demonstration for 1,000 sites. Water Resour. Res. 59, e2022WR034094.
Perera, et al., 2019. Flood Early Warning Systems: A Review of Benefits, Challenges And Prospects. UNU-INWEH Report Series, Issue 08. United Nations University Institute for Water, Environment and Health, Hamilton, Canada.
Petroselli, A., Grimaldi, S., 2018. Design hydrograph estimation in small and fully ungauged basins: a preliminary assessment of the EBA4SUB framework. J. Flood Risk Manage. 11, S197–S210.
Ponce, V.M., 2014. Fundamentals of open-channel hydraulics. Online textbook.
Recanatesi, F., Petroselli, A., 2020. Land cover change and flood risk in a peri-urban environment of the metropolitan area of Rome (Italy). Water Resour. Manag. 34, 4399–4413.
Lazio Region, 2023. Regional Civil Protection https://temporeale.regione.lazio.it/aegis/map/map2d (accessed December 2023).
Rodell, M., Li, B., 2023. Changing intensity of hydroclimatic extreme events revealed by GRACE and GRACE-FO. Nat. Water 1 (3), 241–248.
Rohde, M.M., 2023. Floods and droughts are intensifying globally. Nat. Water 1 (3), 226–227.
Shapley, L., 1952. A Value for n-Person Games. Princeton University Press, Annals of Mathematics Studies, Study 28, pp. 307–317.
Soille, P., 2004. Optimal removal of spurious pits in grid digital elevation models. Water Resour. Res. 40 (12).
Tarboton, D.G., Bras, R.L., Rodríguez-Iturbe, I., 1991. On the extraction of channel networks from digital elevation data. Hydrol. Process. 5 (1), 81–100.
World Meteorological Organization, 2010. Manual on the Global Data-processing and Forecasting System. World Meteorological Organization.

Salvatore Grimaldi[a,*], Francesco Cappelli[a], Simon Michael Papalexiou[b], Andrea Petroselli[c], Fernando Nardi[d], Antonio Annis[e], Rodolfo Piscopia[f], Flavia Tauro[a], Ciro Apollonio[c]

[a] *DIBAF Department, Università degli Studi della Tuscia, 01100 Viterbo, Italy*

[b] *Department of Civil Engineering, University of Calgary, Calgary, Alberta T2N 1N4, Canada*

[c] *DAFNE Department, Università degli Studi della Tuscia, 01100 Viterbo, Italy*

[d] *DICII Department, Università degli Studi di Roma Tor Vergata, 00133 Roma, Italy*

[e] *WARREDOC, Università per Stranieri di Perugia, 06122 Perugia, Italy*

[f] *Freelance, 00100 Roma, Italy*

[*] Corresponding author.
E-mail address: salvatore.grimaldi@unitus.it (S. Grimaldi).