



OPEN Machine learning models for river flow forecasting in small catchments

Marco Luppichini¹✉, Giada Vailati¹, Lorenzo Fontana² & Monica Bini^{1,3,4,5}

In consideration of ongoing climate changes, it has been necessary to provide new tools capable of mitigating hydrogeological risks. These effects will be more marked in small catchments, where the geological and environmental contexts do not require long warning times to implement risk mitigation measures. In this context, deep learning models can be an effective tool for local authorities to have solid forecasts of outflows and to make correct choices during the alarm phase. In this study, we investigate the use of deep learning models able to forecast hydrometric height in very fast hydrographic basins. The errors of the models are very small and about a few centimetres, with several forecasting hours. The models allow a prediction of extreme events with also 4–6 h (RMSE of about 10–30 cm, with a forecasting time of 6 h) in hydrographic basins characterized by rapid changes in the river flow rates. However, to reduce the uncertainties of the predictions with the increase in forecasting time, the system performs better when using a machine learning model able to provide a confidence interval of the prediction based on the last observed river flow rate. By testing models based on different input datasets, the results indicate that a combination of models can provide a set of predictions allowing for a more comprehensive description of the possible future evolutions of river flows. Once the deep learning models have been trained, their application is purely objective and very rapid, permitting the development of simple software that can be used even by lower skilled individuals.

Keywords River flow prediction, Deep learning, Risk mitigation, Italy

Climate changes due to anthropogenic activities are causing significant alterations in the hydrogeological cycle^{1–3}. The Mediterranean is a hotspot for current global warming^{4–6}, exhibiting variations in river discharge and precipitation across various areas^{7–11}. Future projections for hydrogeological regimes indicate a general decrease in annual precipitation with an increased risk of drought^{12–18}, but also an intensification of extreme and fast events^{8,19–21}. The effects of climate evolution in the Mediterranean will be enhanced by the characteristics of medium and small catchments²². The management of small catchments is challenging on account of their hydrogeological features, which often provide insufficient time for effective emergency intervention. These drawbacks also depend on the current lack of suitable meteorological forecasts at this scale²³.

Italy is no exception to this reality, and it is one of the Mediterranean countries in Europe with the highest annual expenditure for damages caused by hydrogeological events^{24–27}. Italian territorial entities responsible for river risk management currently lack efficient predictive models allowing for timely interventions in small catchments, which are those characterizing a significant part of the peninsula. Strict requirements in terms of knowledge of the territory and of hydrogeological parameters are necessary to build an efficient river flow modelling system based on a physical approach. The definition of several parameters, the often-missing information, and the difficulty in simulating the real natural system may produce strong model uncertainties, which play a predominant role in error quantification^{28–30}. For these reasons, very few tools have so far been able to provide rapid alerts for flood risk mitigation, in particular for these types of catchments. Machine learning models have proven to be a valuable tool for runoff modelling^{31–37}, thanks to the use of neural networks that enable to handle the noise and chaotic nature of time series in prediction problems³⁸. In this respect, some recent works have focused on the development of models in small catchments^{39–41}, which currently represent

¹Department of Earth Sciences, University of Pisa, Via S. Maria, 52, 56126 Pisa, Italy. ²Consorzio di Bonifica 1 Toscana Nord, Via della Migliarina, 64, 55049 Viareggio, Lu, Italy. ³CIRSEC Centro Interdipartimentale Di Ricerca Per Lo Studio Degli Effetti delCambiamento Climatico Dell'Università Di Pisa, Via del Borghetto 80, 56124 Pisa, Italy. ⁴Istituto Nazionale Di Geofisica E Vulcanologia (INGV), Via Vigna Murata 605, 00143 Rome, Italy. ⁵Institute of Geosciences and Earth Resources, IGG-CNR, Via Moruzzi 1, 56124 Pisa, Italy. ✉email: marco.luppichini@dst.unipi.it

a challenge in the research and application of flood risk mitigation. Several authors have applied different deep learning techniques to predict river flows, with promising results^{42–46}. Long short-term memory (LSTM) is one of the most popular, efficient and deeply used learning techniques⁴⁷, widely applied in flood prediction studies^{48–50}.

The study area, located in central Italy (Fig. 1), is particularly vulnerable to climate changes, owing to the presence of the cyclonic area of the Gulf of Genoa⁵¹ and of the steep reliefs orthogonal to the direction of prevailing winds^{52–54}. The area is characterized by the presence of the Apuan Alps, a mountain range with large slopes and a complex karstic system^{53,55} (Fig. 1). The region features a complex river network, a coastal lake (Lake Massaciuccoli), and wetland areas. The plain area is subject to regulations designed to maintain the water table below ground level. The area experiences high annual precipitations (even more than 3000 mm^{56–58}), among the highest in the Mediterranean, and sometimes very intense. Furthermore, owing to the presence of the Apuan Alps, hydrogeological events such as floods, debris flows, and landslides often occur^{52,57,59–61}.

The intricate karstic systems in this area make it difficult to provide efficient tools for predicting river dynamics^{30,55,62}. The steep slopes of the mountains induce rapid runoff times, but their estimation is complicated by the poorly understood karstic systems. The plains are characterized by a significant human presence, with over 300,000 inhabitants, and more than 3 million individuals during the summer, presenting a high population density. In this context, for the future effects of climate change, enhanced by the environmental characteristics of the territory and the substantial presence of human activities, we have developed a group of models based on deep learning approaches that can predict river discharge in a regulated lake and in small watercourses. The set of constructed models diverges from the input data matrix employed for training. We chose to create this ensemble primarily for two purposes: (i) to understand the pros and cons associated with incorporating diverse input data types; and (ii) to have models that react variably (or not) from identical initial conditions, thereby yielding a broader spectrum of scenarios and varying levels of prediction certainty.

Different territorial entities have collaborated to the design and implementation of a computer system providing real-time predictions of watercourses up to 6 h in advance for 8 monitoring points (see stars in Fig. 1). Table 1 shows the main hydraulic characteristics of the water system investigated. Prediction models are limited by the inability to understand how the conditions may evolve in the future. As for the hydraulic models, the primary predictive limitation is rainfall, which is the main time-related input variable. Therefore, machine learning models tend to underestimate flows during the forecasting time, due to data uncertainties and

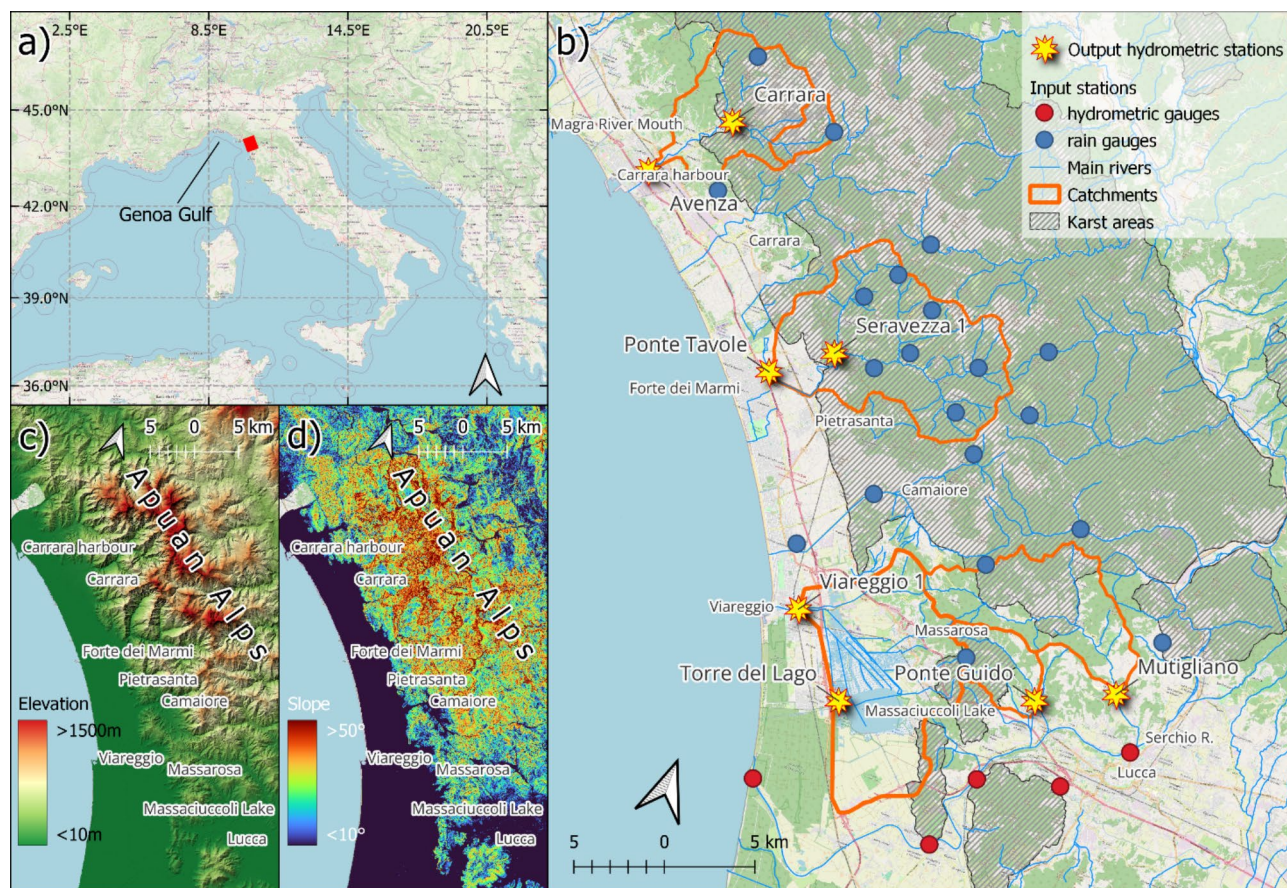


Fig. 1. Study area (a) Input and output gauges; (b) Spatial location of the study area; (c) Morphology of the study area where green indicates an elevation lower than 10 m and red an elevation higher than 1200 m. OpenStreetMap provides basemap.

| Hydraulic system | Station | Median (m) | Average (m) | 25th (m) | 75th (m) | 95th (m) |
|--------------------|----------------|------------|-------------|----------|----------|----------|
| Carrione River | Avenza | 1.17 | 1.09 | 0.91 | 1.29 | 1.49 |
| Carrione River | Carrara | 0.42 | 0.29 | -0.12 | 0.51 | 0.94 |
| Massaciuccoli Lake | Torre del Lago | 0.01 | -0.01 | -0.11 | 0.12 | 0.26 |
| Massaciuccoli Lake | Viareggio 1 | 0.04 | 0.02 | -0.07 | 0.15 | 0.29 |
| Freddana Torrent | Mutigliano | 0.43 | 0.48 | 0.29 | 0.64 | 0.84 |
| Contesora Torrent | Ponte Guido | 0.14 | 0.17 | 0.09 | 0.21 | 0.43 |
| Versilia River | Seravezza 1 | 1.05 | 0.87 | 0.54 | 1.18 | 1.38 |
| Versilia River | Ponte Tavole | 0.7 | 0.63 | 0.29 | 0.95 | 1.23 |

Table 1. Main characteristics of hydrometric time series used as output of the models.

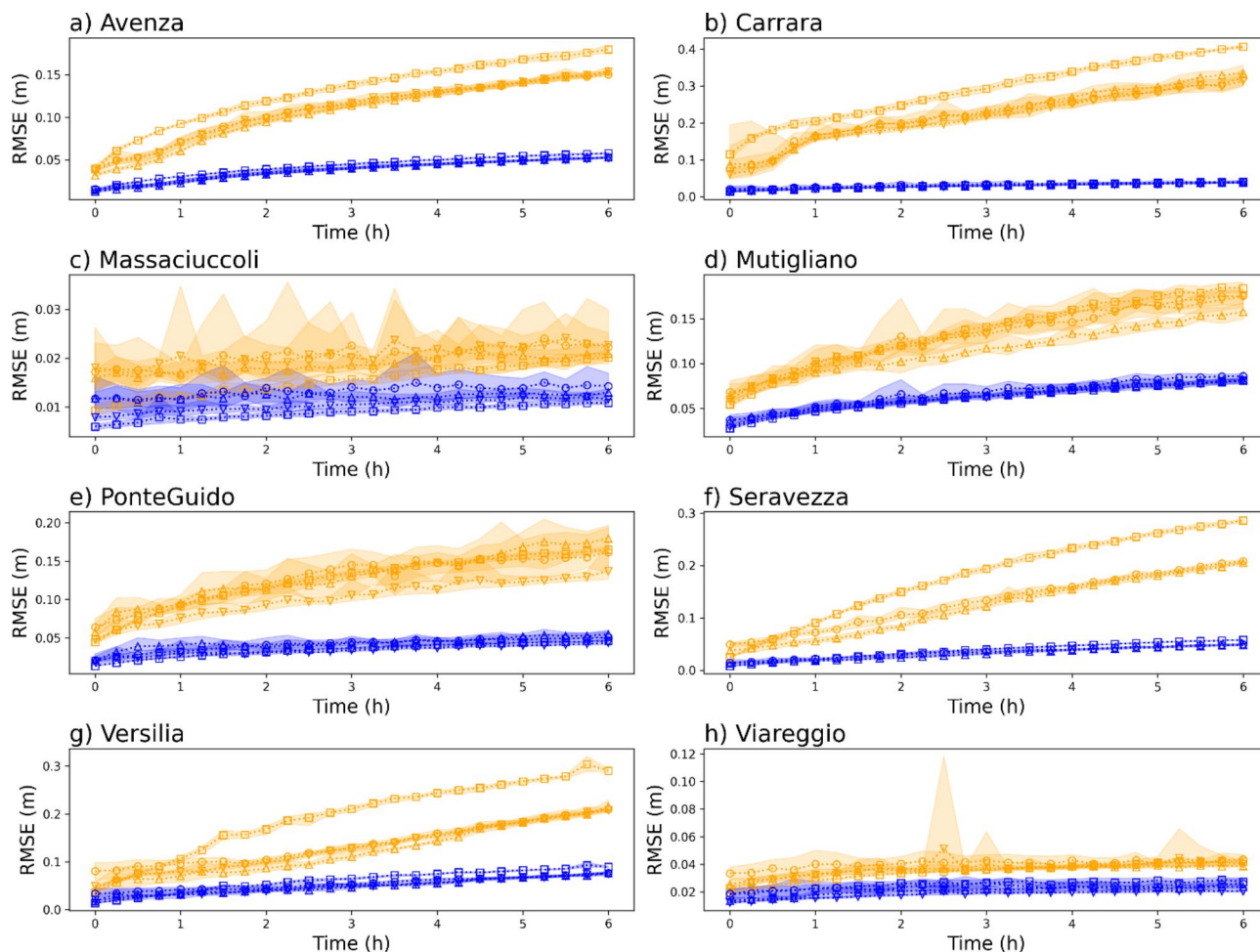


Fig. 2. Root Mean Square Errors (RMSE) of the deep learning models for forecasting time from 0 h to 6 h with a step of 15 min: the lines with ● marker indicate the Standard Model (SM); the lines with ▲ marker indicate the Normalized Model (NM); the lines with ▼ marker indicate the Only Rainfall Model (ORM); the lines with ■ marker indicate the Only Output Model (OOM). The contours indicate a confidential interval of 90%. The blue lines indicate the RMSE using the entire test dataset, while the orange lines indicate the RMSE by analysing the hydrometric height higher than the 95th percentile.

potential biases⁶³. In this work, we experimented an error analysis of the results of predictive models through a combination with additional machine learning models that allow an increase in the lead time for alarms regarding the hydraulic systems studied.

Results and discussion

Figure 2 summarizes the Root Mean Square Errors (RMSE) for the 25 steps of forecasting, for each model and for each output station. The deep learning models developed in this work have small errors for the entire test dataset,

and they focus on events higher than the 95th percentile. We will refer to observations exceeding the 95th percentile as extreme events⁶⁴ by recording errors of only a few centimetres compared to those found in other works^{32,35,63–66}. The results of this work are in agreement with those of other works which used sub-daily river flow data in catchments with a comparable size^{41,67}. However, a direct comparison with other works is not always simple owing to the different measures simulated (hydrometric height or flow rate) and to the different hydraulic condition and behaviour of the rivers. For the Avenza and Carrara (Carriore River), and for the Ponte Tavole and Seravezza (Versilia River) stations the OOM model is the worst, with a clear increase of the performances following the introduction of other gauges. As already said, for the Mutigliano and Ponte Guido stations, the difference between the OOM and the other models is not relevant. However, the NM and the ORM are the best models with the lowest RMSE. For Massaciuccoli and Viareggio (Massaciuccoli Lake system), the OOM model performs similarly to the other ones, highlighting an unclear importance of the introduction of rain and hydrometric gauges as input data. These results indicate the differences in behaviour between the four hydraulic systems studied in this work. For the Carrione and Versilia Rivers, which are characterized by the highest hydraulic ranges, the introduction of more gauges as input to the model increase the possibility of creating accurate models. With the decrease in variability of the system's simulated hydraulic regimes, the importance of the other input gauges decreases. Such behaviour may be observed in the Freddana and Contesora torrents as well as in the Massaciuccoli Lake system, where the difference between the minimum and the maximum regimes is lower than 50 cm (Table 1).

It is worth noting that for step $t = 0$ of each station, and in some cases also for some steps below, the OOM has performances analogous to those of other models. This is an important point because, in the last years, several works (e.g.^{36,46,68,69}), have yielded positive results with regard to deep learning and machine learning models for the prediction of river flows. Complex data input matrices have been used but the role of the time series output on the models has not been understood. The complexity of the neural networks and the high number of parameters estimated during the training phase make it difficult to well understand the input influence on the deep learning models. However, this study underlines the importance of investigating in further detail the real effect of the inputs on the results in these types of models. If the aim of this study had been the construction of models for the analysis of $t = 0$ (making a simple reconstruction of the river flow time series with the river flow data up to step $t - 1$), the use of a network of stations would have resulted useless.

Figures 3 and 4 show some selected applications for each station. These selections derive from an analysis of the 10 highest events in the test dataset (see Supplementary Material for all events). In Fig. 3, we report a case for each station when the models have great performances while, in Fig. 4, we report the cases in which the models have lower performances. No model performs better than the others. The models sometimes produce similar results (see in particular Fig. 3), so that there is a small variability in the predictions; other times there are differences between the predictions (see in particular Fig. 4). This suggests that the combination of models can provide a more exhaustive forecast during the decision-making processes. The hydrological cases of Fig. 4 are affected by a fast variation of the river levels characterized by an increase in river flow in a few tens of minutes. This makes it difficult for the models to provide an efficient prediction, especially with the increase in forecasting time, which induces an underestimation of the real river flow observed. However, for the cases shown in Figs. 3 and 4, the models generally have excellent performances until 2–4 h of forecasting, although in some cases even up to 6 h (see also the Supplementary Material).

The analysis of the Mean Errors (ME; Fig. 5) shows that with the increase of forecasting time, the models present negative errors. In other words, with the increase of the forecasting time, the models always underestimate the flow, and this induces a possible systematic missing alarm, in particular for extreme events (higher than the 95th percentile), as observed during the analysis of Fig. 4. The main factor that can cause an underestimation of the predictions in the highest forecasting times can be the lack of information on the evolution of weather conditions over time. The models can provide a prediction if they know the progressive amount of precipitation falling into the basin and, once the rainfall has started, the models do not know what will be the total amount of precipitation in the future. For this reason, we observe a progressive decrease in the underestimation of the river flows, so that the forecasting time is reduced. In the small hydrographic catchments, we think that the weather predictions cannot solve this limit completely on account of the current small resolution of the weather prediction, with grids greater than 30 km. However, we think that an analysis of the uncertainties can be more efficient to mitigate the likelihood of missing the alarms. The probability of taking wrong decisions is partly linked to the knowledge of the phenomenon, which may result in potentially dubious predictions.

For this reason, the choice of introducing the use of Gradient Boosting Regression (GBR) in this work is key. An example of application of the GBR models to estimate the confidence intervals of the errors on the predictions appears in Fig. 6. The GBR models allow to quantify the uncertainties of the LSTM models on the basis of the last observed hydrometric measures of the output station. The effects of the introduction of the GBR models are crucial. Figures 7 and 8 show the capacity of the models to forecast the alarms for cases higher than the 95th percentile (extreme events). The metrics observable in the figure are True Positive Rate (TPR; Fig. 7) and False Positive Rate (FPR; Fig. 8), which are defined as follows:

$$TPR = \frac{TP}{P}$$

$$FPR = \frac{FP}{N}$$

where P is the number of real cases higher than the 95th percentile in the test dataset; N is the number of real cases lower than the 95th percentile in the test dataset; TP is the number of simulated data that correctly

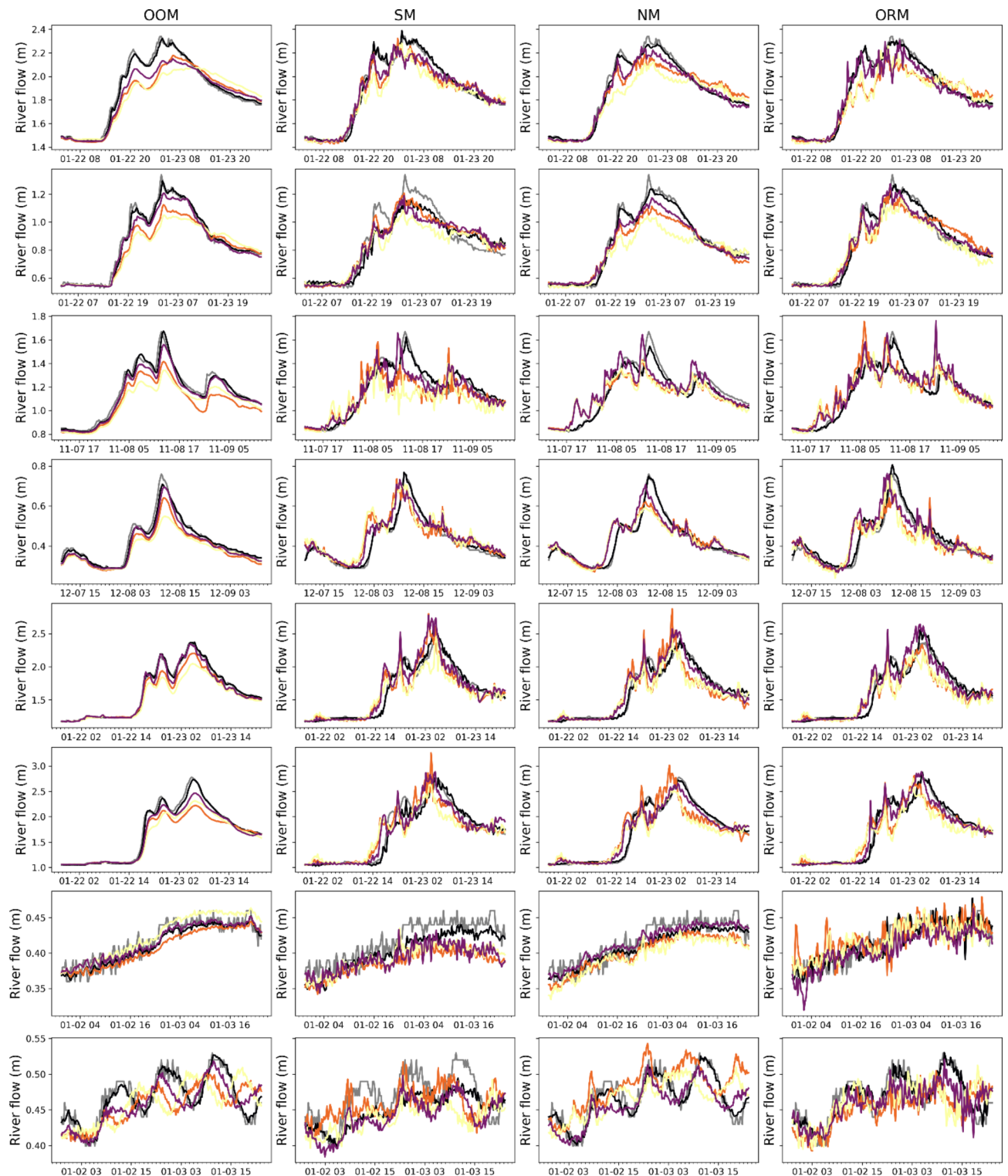


Fig. 3. Selected examples when the models allow good predictions of extreme events. Each row corresponds to a specific station, respectively Avenza, Carrara, Mutigliano, Ponte Guido, Seravezza, Versilia, Massaciuccoli and Viareggio. Each column corresponds to a specific model, respectively: Only Output Model (OOM), Standard Model (SM), Normalized Model (NM), and Only Rainfall Model (ORM). The grey line indicates the observed data, the black line indicates the prediction at 0 h; the violet line indicates the prediction at 2 h; the orange line indicates the prediction at 4 h; and the yellow line indicates the prediction at 6 h.

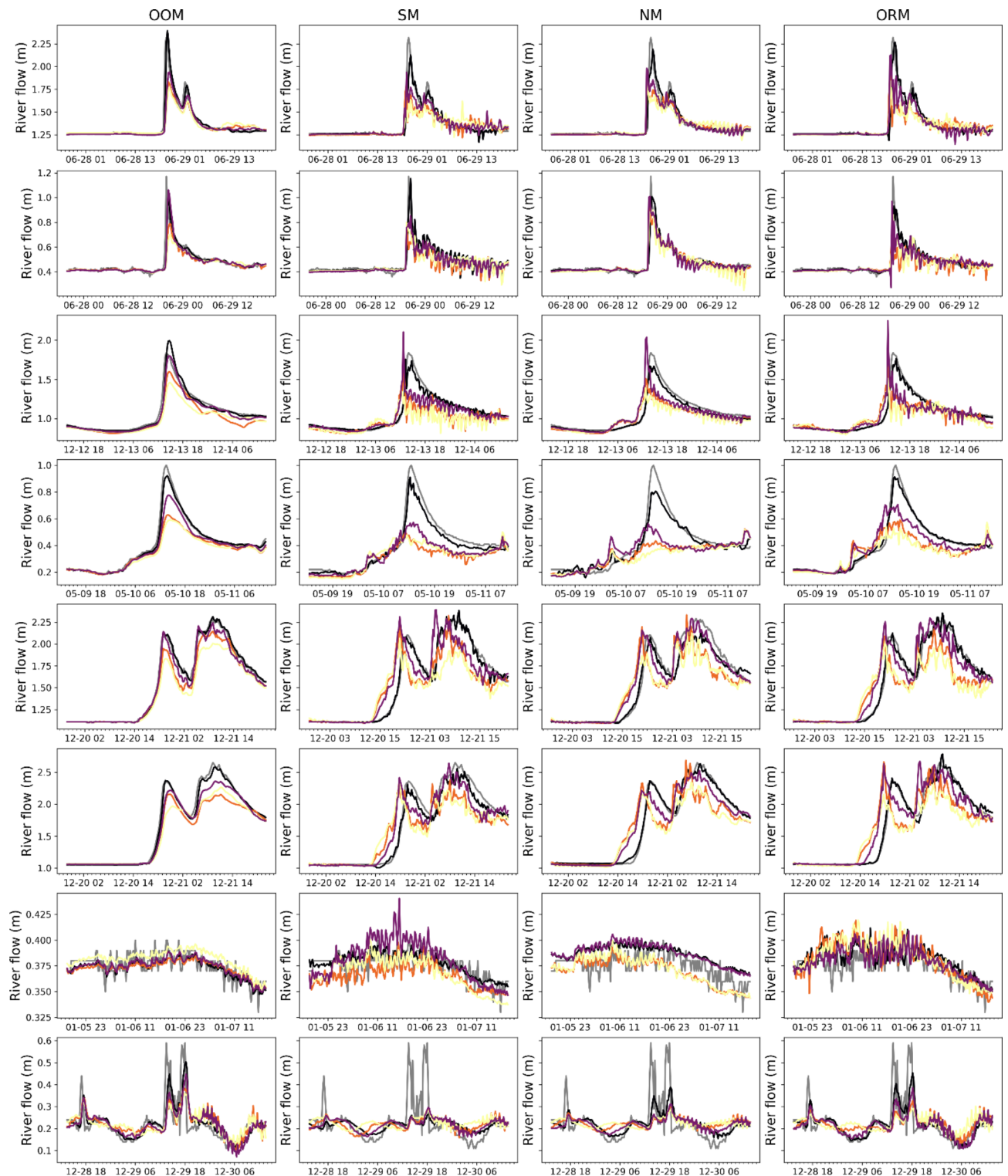


Fig. 4. Selected examples where models do not allow very good predictions of extreme events. Each row corresponds to a specific station, respectively Avenza, Carrara, Mutigliano, Ponte Guido, Seravezza, Versilia, Massaciuccoli and Viareggio. Each column corresponds to a specific model, respectively Only Output Model (OOM), Standard Model (SM), Normalized Model (NM), and Only Rainfall Model (ORM). The grey line indicates the observed data, the black line indicates the prediction at 0 h; the violet line indicates the prediction at 2 h; the orange line indicates the prediction at 4 h; and the yellow line indicates the prediction at 6 h.

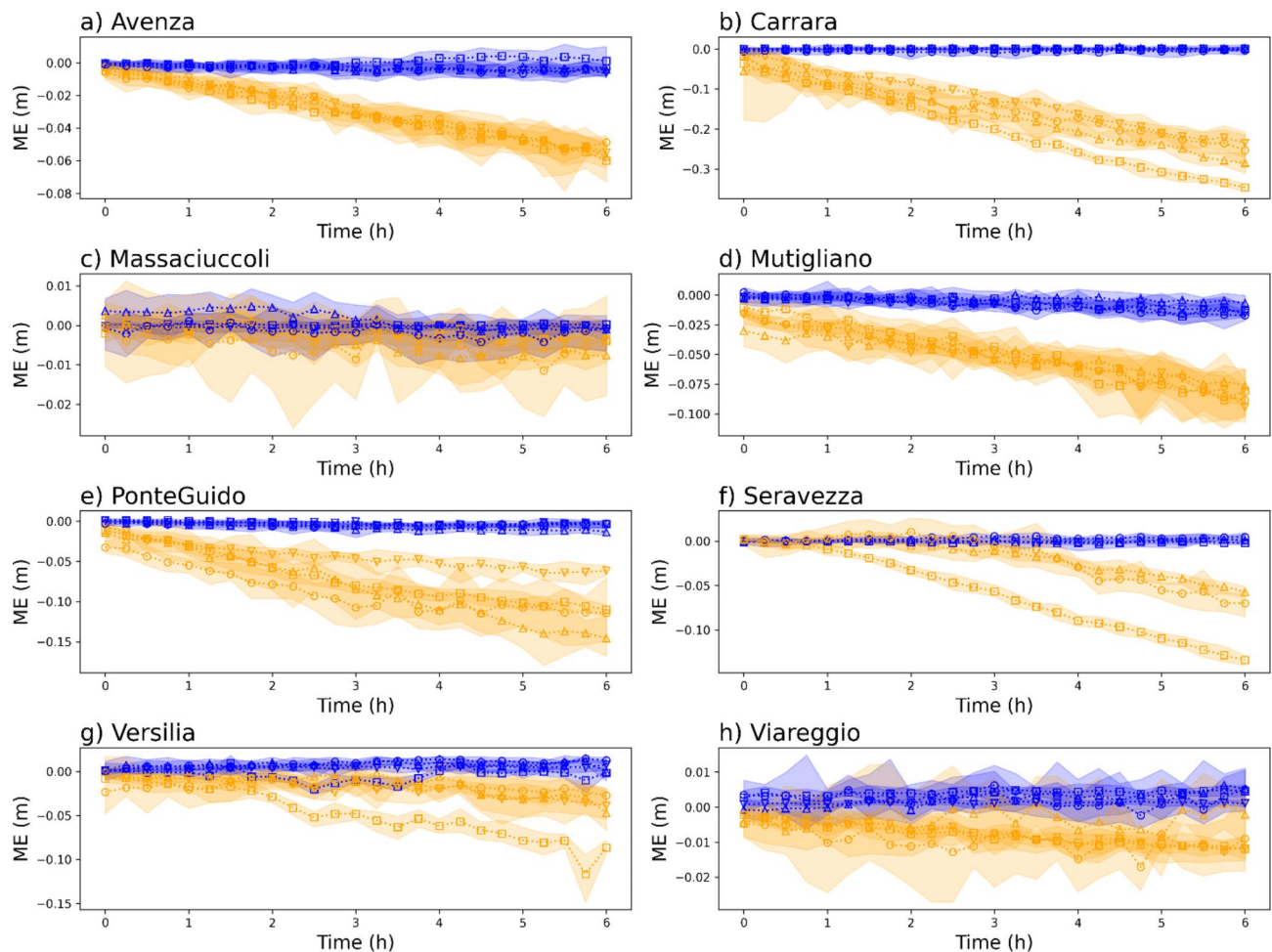


Fig. 5. Mean Errors (ME) of the deep learning models for forecasting time from 0 h to 6 h with a step of 15 min: the lines with ● marker indicate the Standard Model (SM); the lines with ▲ marker indicate the Normalized Model (NM); the lines with ▼ marker indicate the Only Rainfall Model (ORM); the lines with ■ marker indicate the Only Output Model (OOM). The contours indicate a confidential interval of 90%. The blue lines indicate the ME using the entire test dataset, while the orange lines indicate the ME analysing the hydrometric height higher than the 95th percentile.

indicates cases higher than the 95th percentile; and FP is the number of simulated data which wrongly indicates cases higher than the 95th percentile (false alarms). The application of the GBR models must increase the TPR , maintaining the FPR acceptable (Fig. 8). Otherwise, the possibility of predicting real alarms may be hindered by the creation of additional false alarms. The estimation of the uncertainties of the LSTM models can increase the capacity of the models to forecast an alarm (Fig. 7). By analysing the errors of the LSTM models (Figs. 3 and 4), the GBR models provide confidence intervals of only a few centimetres. For example, for the Avenza station, we applied the GBR models by adding an average value of the upper limits of the 50% confidence interval of about 0.01 cm and of the 90% confidence interval of about 0.04 cm. The values for Ponte Guido are about 0.02 cm and 0.04 cm, respectively. These values are very small, but thanks to the nature of the rivers investigated, these very low uncertainties can make the difference between a correct and a missing alarm. Such effects are greater for the smaller hydraulic systems investigated in this work (for example, Mutigliano, Ponte Guido, and Massaciuccoli; Fig. 7). Probably, these hydraulic systems are also affected by the uncertain measurements that are not estimated by the territorial authorities responsible for monitoring the weather network. It is plausible to think that for small channels, lakes, and rivers, the simple error and the instrumental uncertainties used to acquire the data can induce more effects on the results than the hydraulic systems, which are characterized by greater flow rate, and variation from the lower and higher levels. It is thus important to introduce a study of the uncertainties of the deep learning models, so as to provide more efficient forecasting systems like the GBR models proposed in this work. The data-driven characteristics of the deep learning models require an efficient sample and archive of the data, as demonstrated by several authors who found the highest errors in presence of wrong or anomalous data^{70,71}.

The dataset used to train the machine learning models contains, in general, a part of missing values, anomalous data, and redundancy with one or more attributes, which must be eliminated to build a trained generic

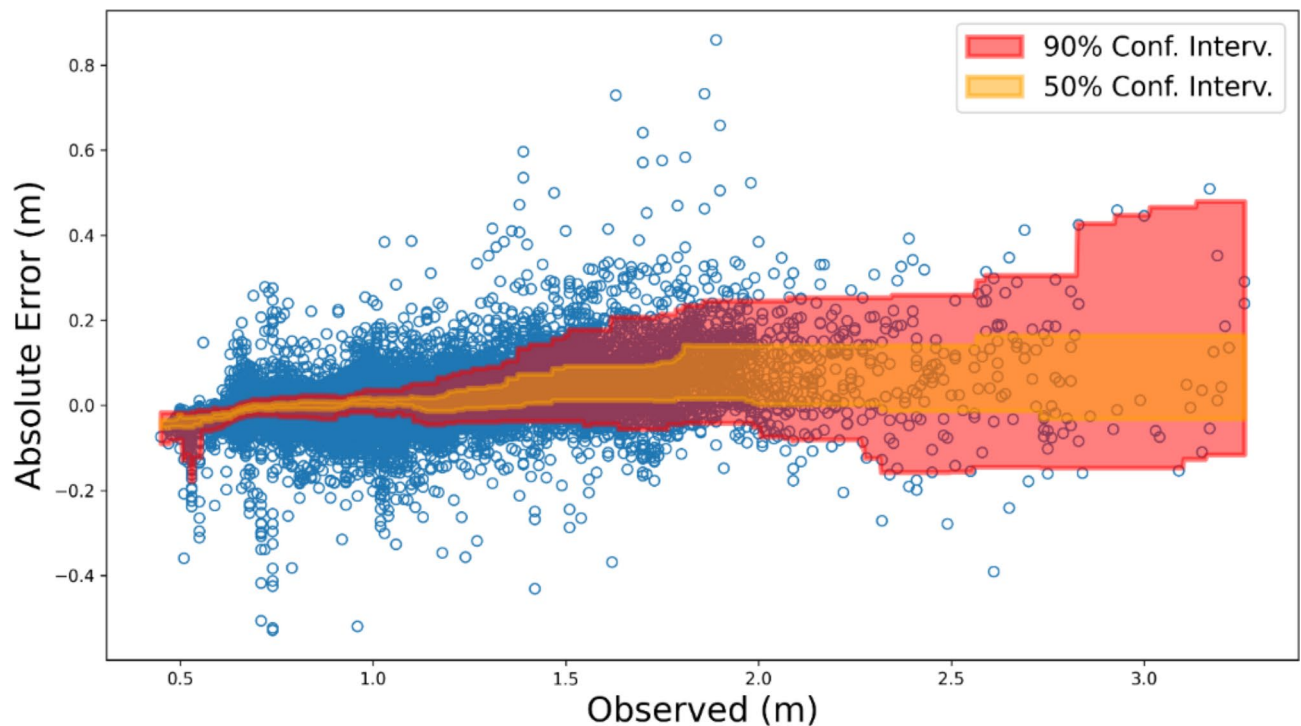


Fig. 6. An example of the applied regression using the Gradient Boosting Regressor (GBR) for the estimation of uncertain predictions. The GBR model is used to estimate the confidential intervals of 50% and 90% based on the observed hydrometric height.

model^{71,72}. However, the automatization of the procedure aimed at solving these issues is very complicated because it depends on several factors such as anomalous or missing data resulting from instrumental anomalies, and their persistence in the time series. Furthermore, several techniques used to solve these types of problems require computational resources and time⁷¹, which in the field of early warning is pressing. In a real application of these models, we cannot exclude the anomalous or missing data provided by a single or a group of stations. For this reason, we need to understand the robustness of these methodologies and the number of anomalous or erroneous data. Figure 9 shows the results obtained when simulating a variable number of not working stations (5–10–20–30%) and comparing the RMSE to the models, without missing data for events higher than the 95th percentile (extreme events). For each percentage, we select a group of stations and we change their input data with a value of 0, simulating missing values from these stations. The selection of the not-working station was applied using a pseudo-random method and the library Random of Python Language. This procedure was replicated using the 10 duplications of the same model and repeating it for five times, by simulating a total of 50 cases for each step of prediction and for each model. The results were encouraging because the variation of also a large number of stations induced a small variation in the results, demonstrating a good capacity of the models to manage the missing values, as demonstrated by⁶³ for a bigger catchment but using an analogous architecture of the neural networks.

After the models have been trained, machine learning techniques do not require any specific parameters, and this simplifies the procedure considerably for the application of these methods by users with poor skills in programming languages or in mathematics or statistics. The machine learning algorithm can be implemented in web and desktop software, which allow the application of the models. In this way, the user requirements are minimal and completely objective, making it possible to obtain the same predictions from different individuals. These possibilities are very important for civil protection and for mitigation of the hydraulic risk, when users need to receive rapid replies for important decisions. The deep learning models are perfect for these aims because the prediction of an event requires very few seconds when using a server machine without great computing capacity. For example, in this work we tested a server machine with 8 Gb of RAM and 6 CPU with a computational capacity of 2.2 GHz, which allowed us to obtain the response of a model (25 t step) in about 20–30 s. Using physical models, this capacity to reply promptly is very rare and in several cases the application of the models needs to determine the specific parameters. This requires further time, slowing down the predictions and then the decisions of the territorial authorities. If these concepts are valid for large river basins, they become even more important in small basins, where response timeliness is in constant struggle with the short flow times characterizing these basins.

Conclusions

The availability of efficient instruments for risk mitigation is crucial also in relation to current global warming and to change in the rainfall regimes, as demonstrated by several authors from different countries. In the field

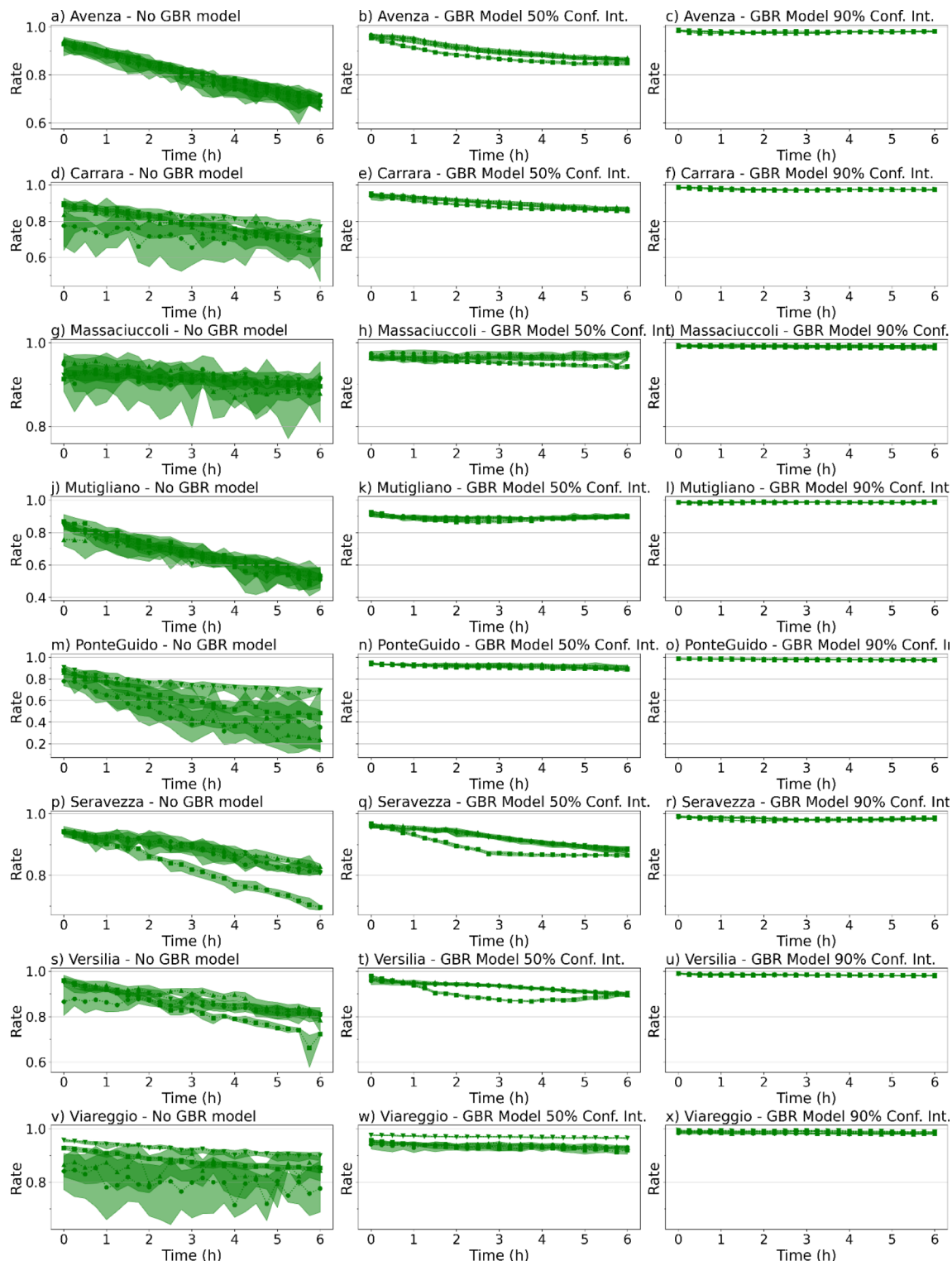


Fig. 7. Sensitivity analysis of the models for events higher than the 95th percentile: True Positive Rate (TPR). The lines with ● marker indicate the Standard Model (SM); the lines with ▲ marker indicate the Normalized Model (NM); the lines with ▼ marker indicate the Only Rainfall Model (ORM); the lines with ■ marker indicate the Only Output Model (OOM).

of hydraulic risk mitigation, we need to solve several challenges to the different geological and environmental characteristics of the catchments. A fundamental characteristic is the size of the catchment that can induce great difficulty in providing instruments able to predict the river flows also amplified by specific geological and climatic conditions. In this work, the study area possesses all these properties since it is characterized by small

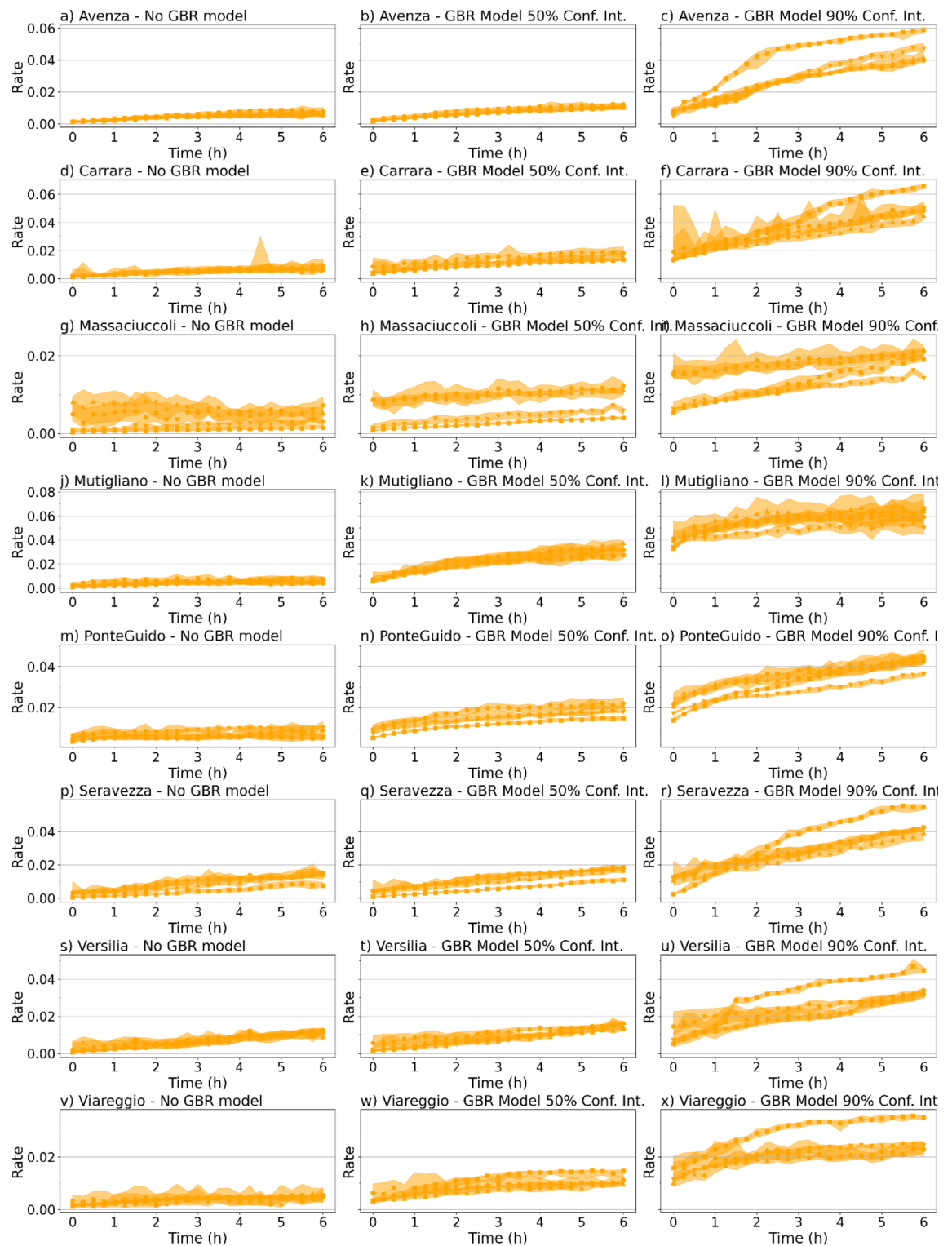


Fig. 8. Sensitivity analysis of the models for events higher than the 95th percentile: False Positive Rate (FPR). The lines with ● marker indicate the Standard Model (SM); the lines with ▲ marker indicate the Normalized Model (NM); the lines with ▼ marker indicate the Only Rainfall Model (ORM); the lines with ■ marker indicate the Only Output Model (OOM).

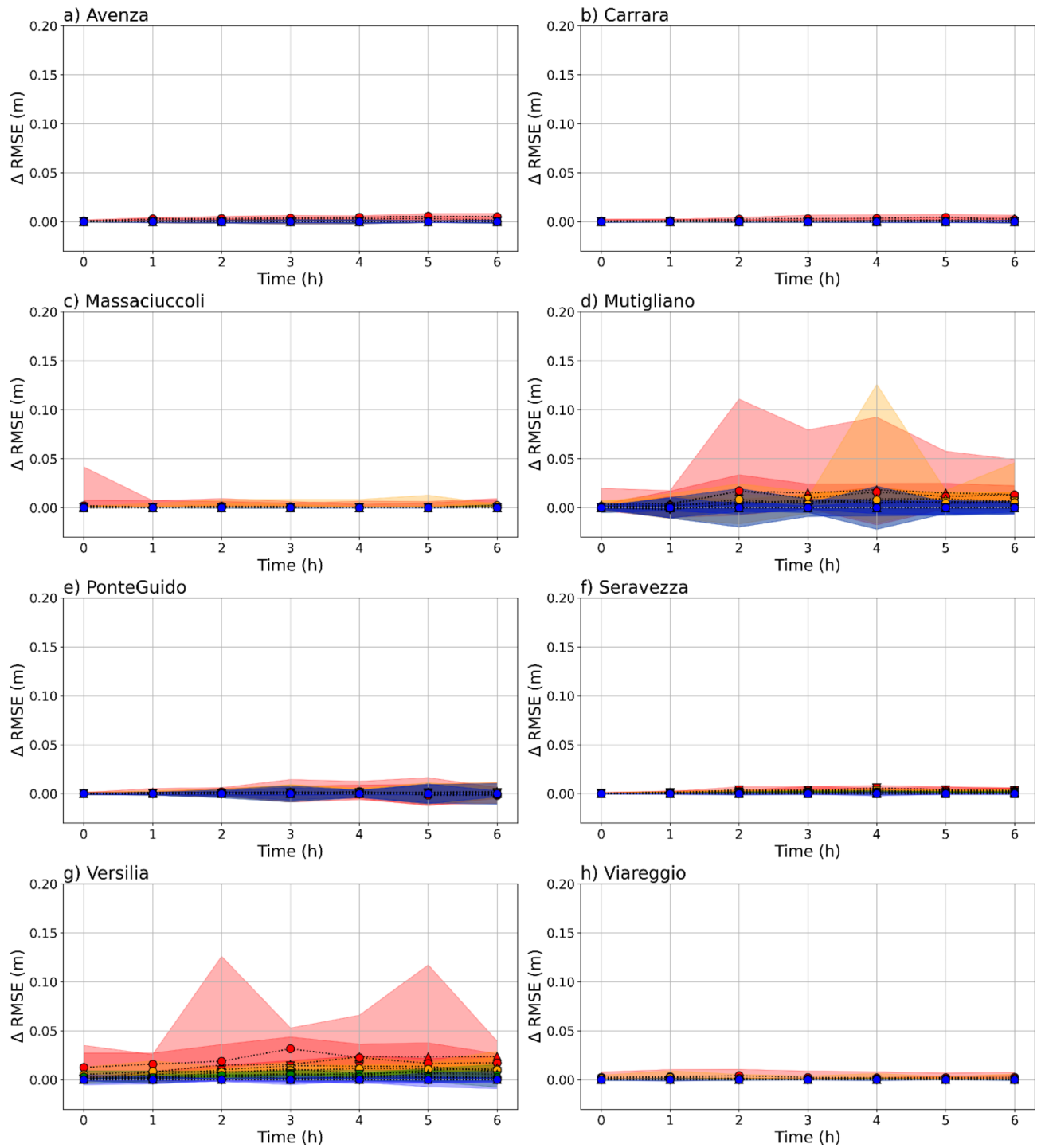


Fig. 9. Susceptibility analysis of the models to the missing values for events higher than the 95th percentile. Blue, green, orange, red bands and markers indicate the results obtained simulating 5%, 10%, 20 and 30%, respectively of the input station as missing values. The lines with ● marker indicate the Standard Model (SM); the lines with ▲ marker indicate the Normalized Model (NM); the lines with ▼ marker indicate the Only Rainfall Model (ORM); the lines with ■ marker indicate the Only Output Model (OOM).

catchments, by a complex karst system, and by high steep mountains. Our work presents an instrument that is able to mitigate hydraulic risk in the study area. We can summarize the main results in the following points:

1. The errors of the models consist in only a few centimetres for the different forecasting times.
2. By using the previous river flow data up to step $t-1$, a regression only of the output time series can be sufficient for reconstruction of the hydrometric time series (forecasting time of 0 h), but also for some forecasting steps.
3. The models tend to underestimate the river flow with an increase of the forecasting time resulting in a systematic possibility of missing the alarms.
4. A blend of various models enables a spectrum of scenarios, enhancing the reliability and certainty of forecasts during critical decision-making.
5. The uncertainty of the models (also due to the sensor measures that are often not estimated) has an important influence on the results for small rivers.
6. The introduction of a statistical method able to estimate uncertainty (in GBR models) allows to increase the capacity of the models to provide alarms on the basis of objective methodologies.
7. The models are sturdy to the presence of the missing values that can affect the results during the real application of these systems.
8. The objective and rapid employment of these systems makes it possible to provide informatic applications that the territorial authorities can use for risk mitigation.

In conclusion, deep learning models can be used to provide instruments for hydraulic mitigation. This will become increasingly important with the development of the Internet of Things (IoT) in the field of weather sensors, which will provide additional data in terms of length of time sensors and of spatial density. Furthermore, the deep learning methodology can be applied to different catchments, with very few modifications. In the future, for a more efficient application of these models in small catchments, researchers will need to further investigate the issue of uncertainty management, try to introduce weather prediction providing different models (using or not artificial intelligence), and try to create different strategies for common and extreme events, which can improve forecasting time.

Methods

In this section, we describe the methodology used in this work. Figure 10 shows the main workflow we employed and that will be described in more detail in the next sections. Step 1 is the analysis and selection of the rainfall and river flow rate time series; Step 2 is the definition of four different input matrices to the deep learning models; Step 3 is the training of the deep learning models based on Long-Short Term Memory; Step 4 is the training of Gradient Boosting Models for estimation of the confidence intervals of the predictions derived from Step 3.

Data and pre-processing

The dataset used in this work is provided by the Servizio Idrologico Regionale (SIR) of the Tuscany Region, the public authority that manages the weather and climatic monitoring system. We selected the best hydrometric and rainfall gauges for this area by identifying 8 hydrometers as output of the DL models (from north to south): Carrara and Avenza (Carrione River); Seravezza 1 and Ponte Tavole (Versilia River); Viareggio 1 and Torre del Lago (Massaciuccoli Lake system); Ponte Guido (Contesora Torrent); and Mutigliano (Freddana Torrent).

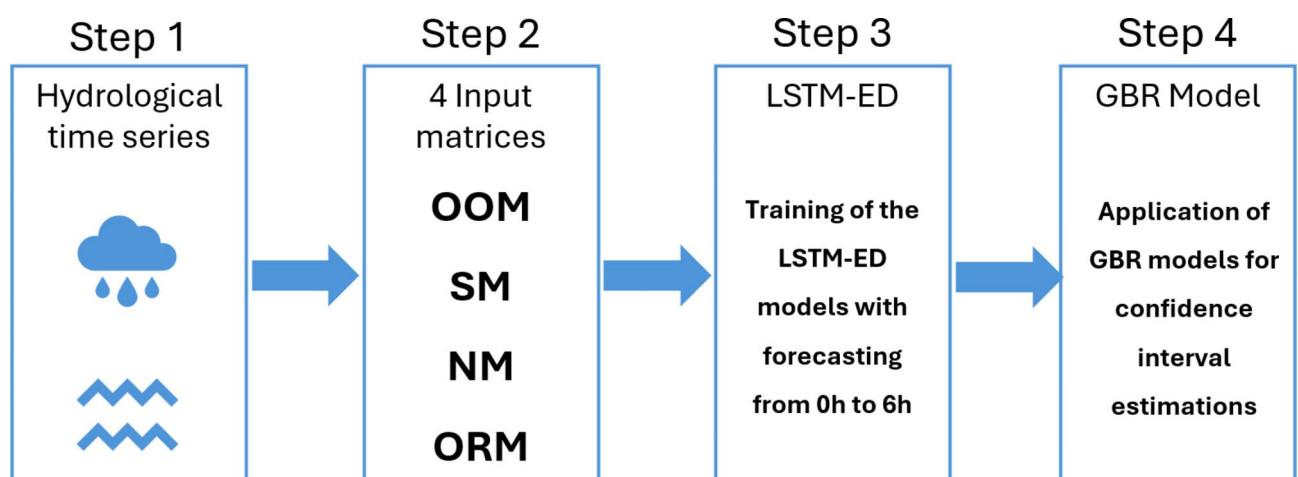


Fig. 10. Workflow of the methodology applied. Step 1: management of the hydrological time series (rainfall and river flow rates). Step 2: definition of four input matrices (Only Output Model; Standard Model; Normalized Model and Only Rainfall Model). Step 3: training of the Long-Short Term Memory Models. Step 4: Application of Gradient Boosting Models for quantification of the confidence intervals of the predictions derived from Step 3.

Figure 1 shows the output gauges (yellow stars), the hydrometric gauges (red points), and the rainfall gauges (blue points) selected and used in this work. The time series used has had a sample frequency of 15 min and a temporal range since 2002. The data, which are sourced from the SIR, lack validation, and undergo no automatic nor manual quality control. Nevertheless, the inherent data-driven nature of machine learning models prompts efforts to address and mitigate issues related to data quality^{70,71}. Consequently, we conducted both manual and graphical checks on the time series utilized. Additionally, the territorial authority employs semi-automatic procedures to validate daily data. Leveraging this process, we excluded sampled data from days in which these procedures had not validated the measurements.

We developed an individual deep learning model for every output hydrometer in order to forecast its 15-minute measurements (H_t). The mathematical representation of the model, applicable to all the sub-basins examined, can be articulated as follows:

$$\hat{H} = f(X_t) = f(H_{t-1}, H_{t-2}, \dots, H_{t-n}, R_{t-1}, R_{t-2}, \dots, R_{t-n})$$

where \hat{H} stands for the predicted hydrometric height at time t ; $H_{t-1}, H_{t-2}, \dots, H_{t-n}$ are the antecedent hydrometric heights (up to $t-1, t-2, \dots, t-n$ time steps); $R_{t-1}, R_{t-2}, \dots, R_{t-n}$ are the antecedent rainfall ($t-1, t-2, \dots, t-n$ time steps). To mitigate the noise inherent in numerous steps and closely spaced measurements, we supplied timestamps for each hour of the preceding period and, subsequently, one timestamp every 4 steps (e.g., $t-0, t-1, t-2, t-3, t-4, t-8, t-12, t-16, \dots, t-96$) up to the 24th hour.

We composed four different input matrices for each station:

- Only Output Model (OOM): this model exploits only the output station by performing a regression, using this time series alone;
- Standard Model (SM): this model is developed using the best rainfall and hydrometric historical series useful for each output station;
- Normalized Model (NM): this model is based on SM. Furthermore, the input data are normalized in a 0–1 range using the maximum and minimum values recorded for each time series;
- Only Rainfall Model (ORM): this model only uses inlet rain gauges.

In theory, each input matrix has advantages (but also disadvantages) that allow a more flexible system for the real application. OOM is the simplest model which makes it scientifically possible to distinguish the next models from a simple regression of the output time series. This is very important because several works have demonstrated the efficiency of complex models without a baseline. Therefore, we can understand the real effect of introducing several data and different datasets. On the other hand, in the application of this system, we need to bear in mind that some gauges cannot work correctly during a flood event. For a nearby area with a similar monitoring system, Luppichini et al. (2022) found that DL models can perform quite well with even a high percentage of missing data (also higher than 25%). In theory, OOM is the model that we can simulate and that in any case always provides us with a prediction model. At the outset, SM is the reference model which, in ideal conditions, should be the best one. The studied hydraulic systems sometimes present a very small range of values if compared to the possible rainfall measurements. In this case, normalization of the values in fixed ranges (for example 0–1) can be useful to train the models and this is represented by NM. The hydrometric data are occasionally influenced by several factors which can produce faulty measurements. For example, in rivers characterized by pebble beds, it is easy to observe that the mainstream of the river is not below the sensor or, in the case of bridges with pillars, the measure can be distorted as a result of the accumulation of wood and of similar materials. In these cases, ORM allows the user to have a model based only on rainfall data. The real strength of this work is the creation of a system based on diverse approaches which can, in principle, perform differently and provide valid tools for prediction under various conditions of use. A summary of the input matrices used in this work is reported in Table 2. For the station of Seravezza1, the SM model coincides with ORM on account of the absence of other significant hydrometers in the catchment of this station.

LSTM regression models

To develop the deep learning models in this study, we primarily utilized the open-source framework Tensorflow⁷³ and the libraries Numpy, Pandas, Scikit-Learn, and Keras⁷⁴ in the Python language v3.9. The architecture of the developed models is based on an encoder-decoder LSTM (LSTM-ED), consisting of two pairs of LSTM nodes. This architecture permits the use of an LSTM able to read the input sequence, one step at a time, and to obtain a fixed-size vector representation in a data structure that occupies a large amount of memory. We then introduced another LSTM to extract the output sequence from that vector⁷⁵. The encoder is composed of two sequential layers (LSTM) of 32 and 16 units, respectively, followed by a repeat vector node. The repeat vector layer repeats the incoming inputs for a specific number of times. The decoder is composed of two LSTM layers of 16 and 32 units respectively, followed by a time-distributed dense node as output of our model.

To evaluate the discrepancy between the predicted and the measured values, we used a loss function for each observation, which allowed us to calculate the cost function. We needed to minimize the cost function by identifying the optimized values for each weight. Through multiple iterations, the optimization algorithm computes the weights that minimize the cost function. In our implementation, we used the Adam optimizer⁷⁶, which is an adaptive learning speed method that computes individual learning rates for several parameters⁷⁶. To halt the training, we employed the specific API of Keras and, specifically, the early stopping method by which the training procedure can be stopped when the monitored metric (namely the value of the cost function) ceases to improve.

| # | Hydraulic system | Station name | Input matrix | N. hydrometric gauges | N. rain gauges |
|----|--------------------|----------------|--------------|-----------------------|----------------|
| 1 | Carrione River | Avenza | SM | 1 | 5 |
| 2 | Carrione River | Avenza | ORM | 0 | 5 |
| 3 | Carrione River | Avenza | OOM | 0 | 0 |
| 4 | Carrione River | Avenza | NM | 1 | 5 |
| 5 | Carrione River | Carrara | SM | 1 | 5 |
| 6 | Carrione River | Carrara | ORM | 0 | 5 |
| 7 | Carrione River | Carrara | OOM | 0 | 0 |
| 8 | Carrione River | Carrara | NM | 1 | 5 |
| 9 | Freddana Torrent | Mutigliano | SM | 2 | 5 |
| 10 | Freddana Torrent | Mutigliano | ORM | 0 | 5 |
| 11 | Freddana Torrent | Mutigliano | OOM | 0 | 0 |
| 12 | Freddana Torrent | Mutigliano | NM | 2 | 5 |
| 13 | Contesora Torrent | Ponte Guido | SM | 2 | 5 |
| 14 | Contesora Torrent | Ponte Guido | ORM | 0 | 5 |
| 15 | Contesora Torrent | Ponte Guido | OOM | 0 | 0 |
| 16 | Contesora Torrent | Ponte Guido | NM | 2 | 5 |
| 17 | Versilia River | Ponte Tavole | SM | 1 | 13 |
| 18 | Versilia River | Ponte Tavole | ORM | 0 | 13 |
| 19 | Versilia River | Ponte Tavole | OOM | 0 | 0 |
| 20 | Versilia River | Ponte Tavole | NM | 1 | 13 |
| 21 | Versilia River | Seravezza 1 | SM | 0 | 13 |
| 22 | Versilia River | Seravezza 1 | OOM | 0 | 0 |
| 23 | Versilia River | Seravezza 1 | NM | 0 | 13 |
| 24 | Massaciuccoli Lake | Torre del Lago | SM | 6 | 10 |
| 25 | Massaciuccoli Lake | Torre del Lago | ORM | 0 | 10 |
| 26 | Massaciuccoli Lake | Torre del Lago | OOM | 0 | 0 |
| 27 | Massaciuccoli Lake | Torre del Lago | NM | 6 | 10 |
| 28 | Massaciuccoli Lake | Viareggio 1 | SM | 6 | 10 |
| 29 | Massaciuccoli Lake | Viareggio 1 | ORM | 0 | 10 |
| 30 | Massaciuccoli Lake | Viareggio 1 | OOM | 0 | 0 |
| 31 | Massaciuccoli Lake | Viareggio 1 | NM | 6 | 10 |

Table 2. Characteristics of the input matrices for each station.

We divided the dataset into three parts with a partition of 60% – 20% – 20% for training, validation, and test dataset respectively. This division of the dataset has been tested in previous studies^{35,66,77,78} and ensures sufficient data for both training and evaluating the model.

The cost function used was the mean square error (MSE) calculated on the validation dataset. This partitioning of the entire dataset helped to minimize the overfitting effect on the training set.

We built a model for each hydrometric station, for each input matrix, and for 25 forecasting steps (from 0 to 6 h every 15 min), totalling 775 models. We repeated the entire procedure ten times, managing to analyse statistically the uncertainty linked to the training procedure based on a study of 7,750 models.

GBR models

The model errors can be analysed using other algorithms that allow to apply prediction uncertainty more efficiently than a simple error metric based on the entire dataset (such as Mean Squared Error or Mean Absolute Error). The combination with a second algorithm able to analyse the residuals obtained from the LSTM-ED model is an innovative approach that can improve the ability to predict a flood event. In this work, we used a Gradient Boosting Regressor (GBR), which is a machine learning algorithm belonging to a group of boosting algorithms that build a predictive model, combining different versions of weaker models⁷⁹. The main idea of Gradient Boosting is to iteratively train models sequentially, where each new model attempts to correct residual errors made by previous models. In other words, each new model focuses on cases that previous models have failed to predict correctly. GBR uses the last observed hydrometric data with the aim of estimating the confidence intervals (50% and 90%) of absolute errors relative to the predictions extracted by the test dataset. The algorithm GBR works with the Scikit-learn library developed in Python language.

Data availability

The authors are available to fully share the raw data which can be requested from the corresponding author.

Received: 15 March 2024; Accepted: 28 October 2024

Published online: 05 November 2024

References

- Dey, P. & Mishra, A. Separating the impacts of climate change and human activities on streamflow: A review of methodologies and critical assumptions. *J. Hydrol. (Amst)*. **548**, 278–290 (2017).
- Wu, P., Christidis, N. & Stott, P. Anthropogenic impact on Earth's hydrological cycle. *Nat. Clim. Change* **3**, 807–810 (2013).
- Milly, P. C. D., Wetherald, R. T., Dunne, K. A. & Delworth, T. L. Increasing risk of great floods in a changing climate. *Nature* **415**, 514–517 (2002).
- Giorgi, F. Climate change hot-spots. *Geophys. Res. Lett.* **33** (2006).
- Giannakopoulos, C. et al. Climatic changes and associated impacts in the Mediterranean resulting from a 2°C global warming. *Glob. Planet. Change* **68**, 209–224 (2009).
- Brogli, R., Kröner, N., Sorland, S. L., Lüthi, D. & Schär, C. The role of hadley circulation and lapse-rate changes for the future European summer climate. *J. Clim.* **32**, 385–404 (2019).
- Lupichini, M., Bini, M., Giannecchini, R. & Zanchetta, G. High-resolution spatial analysis of temperature influence on the rainfall regime and extreme precipitation events in north-central Italy. *Sci. Total Environ.* **880**, 163368 (2023).
- Zittis, G., Bruggeman, A. & Lelieveld, J. Revisiting future extreme precipitation trends in the Mediterranean. *Weather Clim. Extrem.* **34**, 100380 (2021).
- Philandras, C. et al. Long term precipitation trends and variability within the Mediterranean Region. *Nat. Hazards Earth Syst. Sci.* **11**, 3235–3250 (2011).
- Hall, J. & Blöschl, G. Spatial patterns and characteristics of flood seasonality in Europe. *Hydrol. Earth Syst. Sci.* **22**, 3883–3901 (2018).
- Blöschl, G. et al. Changing climate both increases and decreases European river floods. *Nature* **573**, 108–111 (2019).
- Cancelliere, A. & Rossi, G. Droughts in Sicily and comparison of identified droughts in Mediterranean regions. 103–122. https://doi.org/10.1007/978-94-010-0129-8_7 (2003).
- Polemio, M. & Casarano, D. Rainfall and drought in southern Italy (1821–2001). 217–227 (IAHS-AISH Publication, 2004).
- Baronetti, A., Dubreuil, V., Provenzale, A. & Fratianni, S. Future droughts in northern Italy: High-resolution projections using EURO-CORDEX and MED-CORDEX ensembles. *Clim. Change* **172**, 22 (2022).
- José Vidal-Macua, J., Ninyerola, M., Zabala, A., Domingo-Marimon, C. & Pons, X. Factors affecting forest dynamics in the Iberian Peninsula from 1987 to 2012. The role of topography and drought. *Ecol. Manag.* **406**, 290–306 (2017).
- Sousa, P. M. et al. Trends and extremes of drought indices throughout the 20th century in the Mediterranean. *Nat. Hazards Earth Syst. Sci.* **11**, 33–51 (2011).
- Xoplaki, E. et al. 6 - large-scale atmospheric circulation driving extreme climate events in the Mediterranean and its related impacts. in *The Climate of the Mediterranean Region* (ed Lionello, P.) 347–417 (Elsevier, Oxford, 2012). <https://doi.org/10.1016/B978-0-12-416042-2.00006-9>.
- Pal, J. S., Giorgi, F. & Bi, X. Consistency of recent European summer precipitation trends and extremes with future regional climate projections. *Geophys. Res. Lett.* **31**, (2004).
- Barbero, R., Westra, S., Lenderink, G. & Fowler, H. J. Temperature-extreme precipitation scaling: A two-way causality? *Int. J. Climatol.* **38**, e1274–e1279 (2018).
- Pumo, D. & Noto, L. V. Exploring the linkage between dew point temperature and precipitation extremes: A multi-time-scale analysis on a semi-arid Mediterranean region. *Atmos. Res.* **254**, 105508 (2021).
- Pumo, D., Carlino, G., Arnone, E. & Noto, L. V. Relationship between extreme rainfall and surface temperature in Sicily (Italy). *EPIC Ser. Eng.* **3**, 1718–1726 (2018).
- Bertola, M., Viglione, A., Hall, J. & Blöschl, G. Flood trends in Europe: Are changes in small and big floods different? *Hydrol. Earth Syst. Sci. Dis.* 1–23. <https://doi.org/10.5194/hess-2019-523> (2019).
- Mass, C. F. et al. The results of two years of real-time numerical weather prediction over the Pacific Northwest. *Bull. Am. Meteorol. Soc.* **83**, 407–430 (2002).
- Llasat, M. C. et al. High-impact floods and flash floods in Mediterranean countries: The FLASH preliminary database. *Adv. Geosci.* **23**, 47–55 (2010).
- Diodato, N., Ljungqvist, F. C. & Bellocchi, G. A millennium-long reconstruction of damaging hydrological events across Italy. *Sci. Rep.* **9**, 9963 (2019).
- Lupichini, M., Barsanti, M., Giannecchini, R. & Bini, M. Statistical relationships between large-scale circulation patterns and local-scale effects: NAO and rainfall regime in a key area of the Mediterranean basin. *Atmos. Res.* **248**, 105270 (2021).
- Lupichini, M., Bini, M., Barsanti, M., Giannecchini, R. & Zanchetta, G. Seasonal rainfall trends of a key Mediterranean area in relation to large-scale atmospheric circulation: How does current global change affect the rainfall regime? *J. Hydrol. (Amst)*. **612**, 128233 (2022).
- Winter, C. L., Tartakovsky, D. M. & Guadagnini, A. Moment Differential equations for Flow in highly heterogeneous porous media. *Surv. Geophys.* **24**, 81–106 (2003).
- Gómez-Hernández, J. J. Uncertainty in Hydrogeological Modelling. in *Ciba Foundation Symposium 210 - Precision Agriculture: Spatial and Temporal Variability of Environmental Quality* 221–230. <https://doi.org/10.1002/9780470515419.ch14> (2007).
- Lupichini, M. et al. Influence of topographic resolution and accuracy on hydraulic channel flow simulations: Case study of the Versilia River (Italy). *Remote Sens. (Basel)* **11**, (2019).
- Kimura, N., Yoshinaga, I., Sekijima, K., Azechi, I. & Baba, D. Convolutional neural network coupled with a transfer-learning approach for time-series flood predictions. *Water (Basel)* **12**, 96 (2019).
- Sit, M. et al. A comprehensive review of deep learning applications in hydrology and water resources. *Water Sci. Technol.* <https://doi.org/10.2166/wst.2020.369> (2020).
- Ardabili, S., Mosavi, A., Dehghani, M. & Várkonyi-Kóczy, A. R. Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review BT - engineering for sustainable future. in (ed. Várkonyi-Kóczy, A. R.) 52–62 (Springer International Publishing, Cham, 2020).
- Ng, K. W. et al. A review of hybrid deep learning applications for streamflow forecasting. *J. Hydrol. (Amst)*. **625**, 130141 (2023).
- Hu, Y., Yan, L., Hang, T. & Feng, J. *Stream-Flow Forecasting of Small Rivers Based on LSTM*. (2020).
- Kratzert, F., Klotz, D., Brenner, C., Schulz, K. & Herrnegger, M. Rainfall – runoff modelling using long short-term memory (LSTM) networks. 6005–6022 (2018).
- Dibike, Y. B. & Solomatine, D. P. River flow forecasting using artificial neural networks. *Phys. Chem. Earth Part B* **26**, 1–7 (2001).
- Livieris, I. E., Pintelas, E. & Pintelas, P. A CNN-LSTM model for gold price time-series forecasting. *Neural Comput. Appl.* **32**, 17351–17360 (2020).
- Morgenstern, T., Pahnner, S., Mietrach, R. & Schütze, N. Flood forecasting in small catchments using deep learning LSTM networks. *EGU* <https://doi.org/10.5194/egusphere-egu21-15072> (2021).
- Wang, S. & Wang, J. Research on prediction model of mountain flood level in small watershed based on deep learning. in *4th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)* 1024–1027 (2022). <https://doi.org/10.1109/ICMSP55950.2022.9859047>

41. Wu, J. et al. Flash Flood forecasting using support Vector Regression Model in a small mountainous catchment. *Water (Basel)* **11**, (2019).
42. Boulmaiz, T., Guermoui, M. & Boutaghane, H. Impact of training data size on the LSTM performances for rainfall–runoff modeling. *Model. Earth Syst. Environ.* **6**, 2153–2164 (2020).
43. Chattopadhyay, A., Nabizadeh, E. & Hassanzadeh, P. Analog forecasting of extreme-causing weather patterns using deep learning. *J. Adv. Model. Earth Syst.* **12**, e2019MS001958–e2019MS001958 (2020).
44. Marçais, J. & de Dreuz, J. R. Prospective interest of deep learning for hydrological inference. *Groundwater* **55**, 688–692 (2017).
45. Tien Bui, D. et al. A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *Sci. Total Environ.* **701**, 134413 (2020).
46. Van, S. P. et al. Deep learning convolutional neural network in rainfall–runoff modelling. *J. Hydroinform.* **22**, 541–561 (2020).
47. Fawaz, H. I. et al. Deep learning for time series classification: A review To cite this version : HAL Id : hal-02365025 Deep learning for time series classification: A review. (2020).
48. Le, X. H., Ho, H., Lee, G. & Jung, S. Application of long short-term memory (LSTM) neural network for Flood forecasting. *Water (Basel)* **11**, 1387 (2019).
49. Cheng, H., Xie, Z., Wu, L., Yu, Z. & Li, R. Data prediction model in wireless sensor networks based on bidirectional LSTM. *EURASIP J. Wirel. Commun. Netw.* **203** (2019).
50. Nguyen, D. H. & Bae, D. H. Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network. *J. Hydrol. (Amst)*. **584**, 124710 (2020).
51. Trigo, I. F., Bigg, G. R. & Davies, T. D. *Climatology of Cyclogenesis Mechanisms in the Mediterranean*. (2002).
52. Rapetti, C. & Rapetti, F. Levento Pluviometrico Eccezionale Del 19 giugno 1996 in Alta Versilia (Toscana) nel quadro delle precipitazioni delle Alpi Apuane. *Atti Soc. Sci. Nat. Mem. Serie A* **103**, 143–159 (1996).
53. Carmignani, L. & Kligfield, R. Crustal extension in the northern Apennines: The transition from compression to extension in the Alpi Apuane Core Complex. *Tectonics* **9**, 1275–1303 (1990).
54. Carmignani, L., Conti, P., Cornamusi, G. & Pirro, A. Geological map of Tuscany (Italy). *J. Maps* **9**, 487–497 (2013).
55. Baroni, C. et al. Comitato Glaciologico Italiano., Geomorphological and neotectonic map of the Apuan Alps (Tuscany, Italy). in *Geografia Fisica e Dinamica Quaternaria* vol. 38. 201–227 (2015).
56. Rapetti, F. & Vittorini, S. Osservazioni Sul clima del litorale pisano. *Riv Geof Italiana* (1978).
57. Rapetti, F. & Vittorini, S. Le Precipitazioni in Toscana: Osservazioni Sui casi estremi. *Riv Geogr. Ital.* **101**, 47–76 (1994).
58. Fratianni, S. & Acquotta, F. The climate of Italy. in *Landscapes and Landforms of Italy* (eds Soldati, M. & Marchetti, M.) 29–38 (Springer International Publishing, Cham, 2017). https://doi.org/10.1007/978-3-319-26194-2_4.
59. Giannecchini, R. & Avanzi, D. A. Historical research as a tool in estimating hydrogeological hazard in a typical small alpine-like area: The example of the Versilia River basin (Apuan Alps, Italy). *Phys. Chem. Earth Parts A/B/C* **49**, 32–43 (2012).
60. D'Amato Avanzi, G. & Giannecchini, R. Eventi alluvionali e fenomeni franosi nelle Alpi Apuane (Toscana): Primi Risultati Di Un'indagine retrospettiva nel bacino del fiume versilia. **110**, 527–559 (2003).
61. D'Amato Avanzi, G., Giannecchini, R. & Puccinelli, A. The influence of the geological and geomorphological settings on shallow landslides. An example in a temperate climate environment: The June 19, 1996 event in northwestern Tuscany (Italy). *Eng. Geol.* **73**, 215–228 (2004).
62. Piccini, L. Gavorrano., Le aree carsiche della Toscana. in *FST* (2001).
63. Luppichini, M., Barsanti, M., Giannecchini, R. & Bini, M. Deep learning models to predict flood events in fast-flowing watersheds. *Sci. Total Environ.* **813**, 151885 (2022).
64. IPCC. Special report of the intergovernmental panel on climate change managing the risks of extreme events and disasters to advance climate change adaptation. (2013).
65. Aichouri, I. et al. River flow model using artificial neural networks. *Energy Procedia* **74**, 1007–1014 (2015).
66. Xu, W. et al. Using long short-term memory networks for river flow prediction. *Hydrol. Res.* **51**, 1358–1376 (2020).
67. Song, T. et al. Flash Flood forecasting based on long short-term memory networks. *Water (Basel)* **12**, (2020).
68. Liu, D., Jiang, W., Mu, L. & Wang, S. Streamflow prediction using deep learning neural network: Case study of Yangtze River. *IEEE Access*. **8**, 90069–90086 (2020).
69. Hussain, D., Hussain, T., Khan, A., Naqvi, S. & Jamil, A. A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin. *Earth Sci. Inf.* **13**, 1–13 (2020).
70. Lupi, A., Luppichini, M., Barsanti, M., Bini, M. & Giannecchini, R. Machine learning models to complete rainfall time series databases affected by missing or anomalous data. *Earth Sci. Inf.* **16**, 3717–3728 (2023).
71. Hasan, M. K. et al. Missing value imputation affects the performance of machine learning: A review and analysis of the literature (2010–2021). *Inf. Med. Unlocked* **27**, 100799 (2021).
72. Purwar, A. & Singh, S. K. Hybrid prediction model with missing value imputation for medical data. *Expert Syst. Appl.* **42**, 5621–5631 (2015).
73. Abadi, M. et al. TensorFlow: Large-scale machine learning on heterogeneous systems. Preprint at <http://tensorflow.org/> (2015).
74. Chollet, F. & Keras Preprint at <https://github.com/fchollet/keras> (2015).
75. Sutskever, I., Vinyals, O. & Le, Q. V. Sequence to sequence learning with neural networks. (2014).
76. Kingma, D. P., Ba, J. & Adam A. Method for stochastic optimization. Preprint at (2014).
77. Li, W., Kiaghadi, A. & Dawson, C. High temporal resolution rainfall–runoff modeling using long-short-term-memory (LSTM) networks. *Neural Comput. Appl.* <https://doi.org/10.1007/s00521-020-05010-6> (2020).
78. Kao, I. F., Zhou, Y., Chang, L. C. & Chang, F. J. Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting. *J. Hydrol. (Amst)*. **583**, 124631 (2020).
79. Friedman, J. H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 1189–1232 (2001).

Acknowledgements

We thank the Regional Hydrological Service of the Tuscany Region for providing the data used in this study and for allowing the creation of a data passage between different server machines necessary for the application of the models in this territory.

Author contributions

M.L.: idealization, data acquisition, data management, analyses, drafting the manuscript G.V.: data acquisition, analyses L.F.: validation of results, drafting the manuscript M.B.: idealization, validation of results, supervision, drafting the manuscript.

Funding

This research was funded by the project: “Valutazione di scenari dei deflussi superficiali di alcuni selezionati corpi idrici superficiali, basandosi su una analisi dettagliata di banche date meteorologiche e l'applicazione di

tecniche Intelligenza Artificiale”, Consorzio di Bonifica Toscana Nord, (Resp. M. Bini).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-78012-2>.

Correspondence and requests for materials should be addressed to M.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024