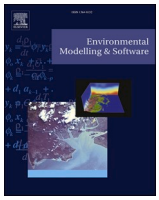
Environmental Modelling and Software 160 (2023) 105582

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



# Near-real-time satellite precipitation data ingestion into peak runoff forecasting models

Paul Muñoz [a,b,*], Gerald Corzo [c], Dimitri Solomatine [c,d,e], Jan Feyen [f], Rolando Célleri [a,b]

[a] Departamento de Recursos Hídricos y Ciencias Ambientales, Universidad de Cuenca, Cuenca, 010150, Ecuador
[b] Facultad de Ingeniería, Universidad de Cuenca, Cuenca, 010150, Ecuador
[c] Hydroinformatics Chair Group, IHE Delft Institute for Water Education, 2611AX, Delft, the Netherlands
[d] Water Resources Section, Delft University of Technology, Mekelweg 5, 2628CD, Delft, the Netherlands
[e] Water Problems Institute of RAS, Gubkina, 3, Moscow, 117971, Russia
[f] Faculty of Bioscience Engineering, Catholic University of Leuven, Leuven, 3001, Belgium

ARTICLE INFO

*Keywords:*
Extreme runoff
Forecasting
PERSIANN
IMERG
Feature engineering
Baseflow separation
Tropical Andes

ABSTRACT

Extreme peak runoff forecasting is still a challenge in hydrology. In fact, the use of traditional physically-based models is limited by the lack of sufficient data and the complexity of the inner hydrological processes. Here, we employ a Machine Learning technique, the Random Forest (RF) together with a combination of Feature Engineering (FE) strategies for adding physical knowledge to RF models and improving their forecasting performances. The FE strategies include precipitation-event classification according to hydrometeorological criteria and separation of flows into baseflow and directflow. We used ~ 3.5 years of hourly precipitation information retrieved from two near-real-time satellite precipitation databases (PERSIANN-CCS and IMERG-ER), and runoff data at the outlet of a 3391-km$^2$ basin located in the tropical Andes of Ecuador. The developed models obtained Nash-Sutcliffe efficiencies varying from 0.86 to 0.59 for lead times between 1 and 6 h. The best performances were obtained for peak runoffs triggered by short-extension precipitation events (<50 km$^2$) where infiltration- or saturation-excess runoff responses are well learned by the RF models. Conversely, the forecasting difficulty is associated with extensive precipitation events. For such conditions, a deeper characterization of the biophysical characteristics of the basin is encouraged for capturing the dynamic of directflow across multiple runoff responses. All in all, the potential to employ near-real-time satellite precipitation and the use of FE strategies for improving RF forecasting provides hydrologists with new tools for real-time runoff forecasting in remote or complex regions.

## 1. Introduction

The use of traditional physically-based models to forecast peak runoffs is either limited or leads to significant uncertainties for regions with complex biophysical characteristics (Clark et al., 2017). This is due to data scarcity and extreme spatio-temporal variability of the main runoff driving forces (precipitation, soil humidity, topographic features, etc.). And even with the increasing availability of Remote Sensing (RS) estimates, it remains mandatory a validation/correction with ground information before RS usage. Moreover, physically-based models suffer from overparameterization issues and intensive computation which prohibits short-term and near-real-time applications (Mosavi et al., 2018; Young, 2002). To overcome these issues, the use of advanced data-driven (black-box) approaches such as Machine Learning (ML) has gained popularity among hydrologists (Bontempi et al., 2012; Chang et al., 2019; Corzo and Solomatine, 2007a; Elshorbagy et al., 2010; Galelli; Castelletti, 2013; Mosavi et al., 2018; Muñoz et al., 2021, 2018; Solomatine et al., 2009; Solomatine and Siek, 2006).

In this sense, the use of ML represents both a challenge and an opportunity. First, the challenge is to select the optimal ML technique for peak runoff forecasting. Different ML methods have been used to explore better performance in conventional hydrological problems. It is common to find new studies using Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Random Forest (RF) (Bhattacharya and Solomatine, 2005; Dibike et al., 2001; Hosseini and Mahjouri, 2016; Mosavi et al., 2018; Muñoz et al., 2021, 2018; Solomatine and Dulal, 2003; Tongal; Booij, 2018; Young et al., 2017). Comparisons between these techniques have demonstrated that RF produces promising results

* Corresponding author. Departamento de Recursos Hídricos y Ciencias Ambientales, Universidad de Cuenca, Cuenca, 010150, Ecuador.
*E-mail address:* paul.munozp@ucuenca.edu.ec (P. Muñoz).

https://doi.org/10.1016/j.envsoft.2022.105582
Received 4 August 2022; Received in revised form 8 November 2022; Accepted 9 November 2022
Available online 18 November 2022

in comparison to SVMs and ANNs, especially because RF is capable to deal with small size samples and complex data structures (Abda et al., 2022; Galelli; Castelletti, 2013; B. Li et al., 2016; Papacharalampous and Tyralis, 2018; Solomatine and Dulal, 2003). Additional advantages of the RF are fewer parameters to calibrate, higher accuracies, robustness, overfitting reduction, and the possibility to interpret results through calculation of estimator importance (Biau and Scornet, 2016; Breiman, 2001; Chen et al., 2004; Contreras et al., 2021; Muñoz et al., 2018, 2021; Orellana-Alvear et al., 2020; Tyralis et al., 2019; Wang et al., 2015). Yet, RF must be implemented in such a way that extreme events can be forecast, e.g., through extreme event-based modeling.

Second, the use of ML techniques is a great opportunity for exploiting readily-available RS satellite imagery, especially for the cases where validation/correction is not possible due to a lack of ground monitoring networks. This is quite useful for complex regions such as the tropical Andes, with inexistent or insufficient ground precipitation networks for characterizing precipitation patterns (Ochoa-Tocachi et al., 2016; Rollenbeck and Bendix, 2011). Among multi-satellite precipitation products, we highlight the NASA Global Precipitation Measurement (GPM), Integrated Multi-satellite Retrievals for GPM (IMERG) (Huffman et al., 2015), and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Hsu et al., 1997). IMERG and PERSIANN products offer quasi-global coverage, free access, and high spatiotemporal resolutions which have yielded a growing body of literature for hydrometeorological applications (Tang et al., 2016). Applications include tracking precipitation anomalies (Nguyen et al., 2014; Sakib et al., 2021), precipitation early-warning systems (Sorooshian et al., 2014), and flood forecasting and mapping (Belabid et al., 2019; Nguyen et al., 2015).

The combination of ML techniques and RS products can be used for understanding key hydrological processes, and for developing operational forecasting systems. For the first case, the forecasting task exploits all available information for characterizing multiple precipitation-runoff relations at different short-term lead times. Then, the gained insight can be used for operational hydrology, for instance, for the deployment of flash flood early warning systems where peak runoff reliability and timing are of utmost importance. In this study we attempt to address the first case, and set the basis for future development focused on operational hydrology. However, the controversy about ML in hydrology has raged unabated during the last decades because, in principle, black-box modeling does not explicitly represent the hydrological processes. And thus, any ML structural interpretation lacks physical knowledge, limiting major performance improvement and the reproducibility of results (Elshorbagy et al., 2010). This controversy motivates current studies to focus on 'grey modeling', i.e., the use of Feature Engineering (FE) strategies for adding physical knowledge of the functioning of the system to enlighten the black-box and improve model performances. In practice, FE strategies can be used for input variable selection, and for processing initial input data either to modify (transform) or create new features for facilitating the data ingestion process of ML algorithms.

Several FE strategies have proven to be successful in hydrological modeling. For example, the use of object-based methods for deriving precipitation attributes from satellite imagery (Davis et al., 2006; Laverde-Barajas et al., 2020, 2019; J. Li et al., 2016; Muñoz et al., 2021; Peña-Barragán et al., 2011; Vogels et al., 2020), runoff separation into subflow components (Corzo and Solomatine, 2007b; Tongal and Booij, 2018a; Willems, 2009), exploitation of topographic characteristics (Huang and Lee, 2021), the addition of stream network information (Akhtar et al., 2009; Huang and Lee, 2021), or various ways of employing hydrological knowledge in choosing input attributes (Moreido et al., 2021).

Within this framework, this study addresses the knowledge gap in developing peak runoff forecasting models using non-validated near-real-time satellite products together with a combination of FE strategies for comprehending the functioning of a catchment for short-term lead times, and for improving the efficiency of RF models. The FE strategies selected are flow separation into baseflow and directflow, and precipitation-runoff event classification according to precipitation attributes derived from satellite imagery. An additional objective is to unravel the influence of multiple precipitation-runoff responses through specialized runoff forecasting of the classified events. The proposed methodology is applied to a 3391-km$^2$ basin located in the southern Andes of Ecuador, and for short-term lead times between 1 and 6 h to account for flash-floods.

## 2. Study area and dataset

### 2.1. Study area

The Jubones basin is located in the tropical Andes of Ecuador, covering an area of 3391 km$^2$ upstream of the Minas-San Francisco (MSF) hydroelectric dam (see Fig. 1). The MSF was constructed and started operating in late 2018. The elevation of the Jubones basin ranges between 1250 and 3920 m above sea level. The climatology of the basin is governed by local topography, the presence of the Andean mountain range, trade winds, and ocean currents from the Pacific Ocean. As a result, the spatial distribution of the climatology is very variable, depicting tropical to semi-arid climate according to the Köppen-Geiger classification (Peel et al., 2007). Mean annual precipitation in the basin ranges from 290 to 925 mm, and the mean annual temperature from 15 to 28 °C (Hasan and Wyseure, 2018).

The dataset comprises ~ 3.5 years of hourly information on two variables, precipitation, and runoff for the period January 2019 to June 2022. Precipitation data were retrieved from two near-real-time databases, the IMERG-Early Run (ER), and the PERSIANN-Cloud Classification System (CCS) products. Data were extracted at the finest temporal resolution and then aggregated to the hourly time step. Apart from inner satellite image processing, the most intuitive difference between both precipitation sources is their spatial resolution. The PERSIANN-CCS presents the highest spatial resolution for the study area (0.04°, ~ 4.4 km), and it is the result of infrared imagery processing and cloud classification using artificial neural networks (Hong et al., 2004). Whereas the IMERG-ER delivers 30-min maps with a spatial resolution of 0.1° (~ 11.1 km) using an approach based on the interpolation of multiple microwave precipitation estimates.

Fig. 2 compares hourly satellite precipitation measured by both satellite products in the Jubones basin, with mean (maximum) annual precipitation depths of 729 (1167) and 1532 (2759) mm, respectively. The mean annual precipitation differences of 803 and 1592 mm for the mean and the maximum precipitation are attributed to the aforementioned reasons. It is also worth noting the difference in the number of pixels (timeseries) obtained with each satellite product, 174 and 30 pixels for the PERSIAN-CCS and the IMERG-ER, respectively.

To date, no ground precipitation gauges are operating in the basin. However, a precipitation comparison can be done with the study of Hasan and Wyseure (2018) in the same basin. In that study, daily historical data for the period 1982–1998 revealed mean annual precipitation ranging from 471 to 1106 mm, which better agrees with the obtained PERSIANN-CCS information. Although it was not possible to perform an hourly validation of the satellite precipitation with ground measurements, this was not a limiting aspect since precipitation is merely an estimator of runoff when ML techniques are employed. Instead, we exploited the spatiotemporal variability of both precipitation signals under the assumption that the overall bias of each of them remains constant for the study area.

On the other hand, hourly runoff data was collected for a hydrological station in the outlet of the basin, i.e., the entrance MSF hydropower dam (see Fig. 1). The runoff data for the period January 2019 to June 2022 were facilitated by the Corporación Eléctrica del Ecuador (CELEC EP, https://www.celec.gob.ec/), the company that manages the MSF hydropower dam.

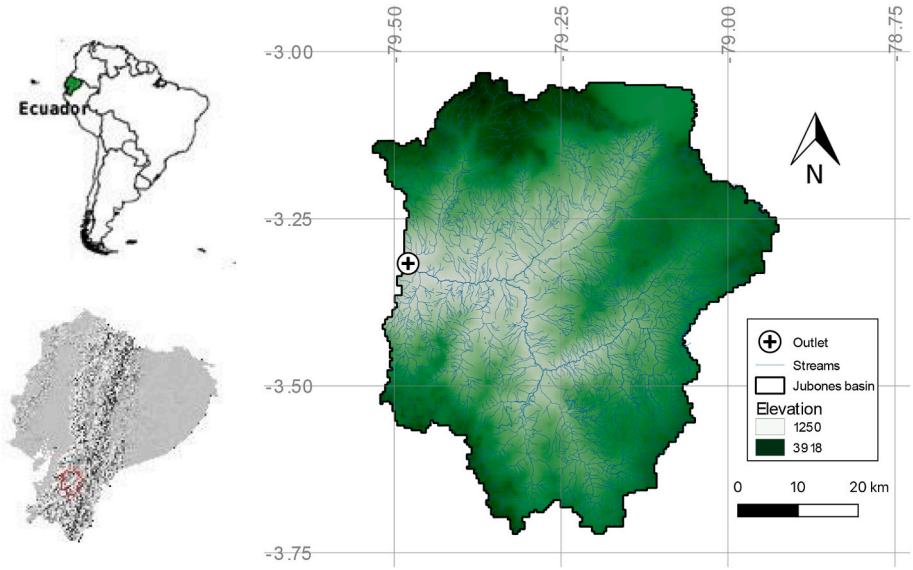



**Fig. 1.** The Jubones basin in the Tropical Andes of Ecuador, South America.Dataset.

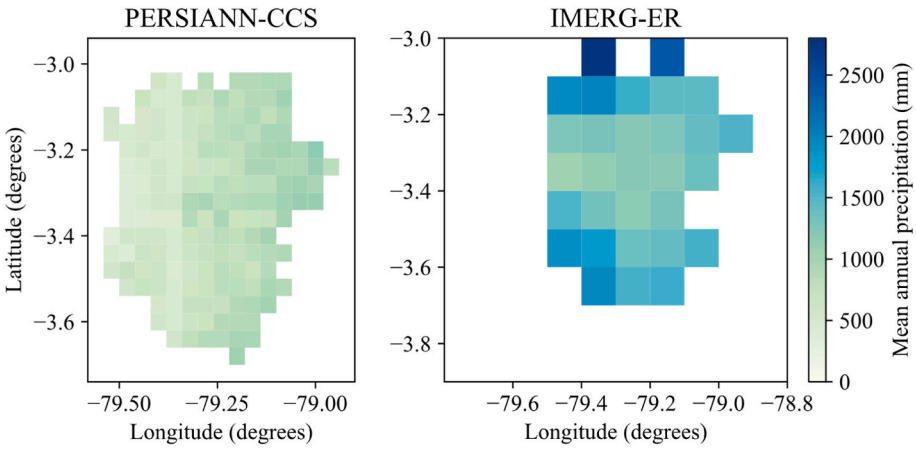



**Fig. 2.** Mean annual precipitation measured by the PERSIANN-CCS and the IMERG-ER satellite products for the study period from January 2019 to June 2022 (Jubones basin, Ecuador).

## 3. Methodology

Fig. 3 summarizes the methodology of this study. First, nearly-independent peak runoff events were selected, and the hourly runoff time series (total flow) was separated into baseflow and directflow series (Fig. 3a). Second, the precipitation imagery associated with peak events was processed using an object-based Connected Component Analysis (CCA) to extract key precipitation object attributes (Fig. 3b). The CCA is applied to the precipitation dataset following a modular approach. This means that the CCA is preferably applied to the finest spatial-resolution product (PERSIANN-CCS), and for the cases when no precipitation is detected by the PERSIANN-CCS, the CCA is applied to the supplementary IMERG-ER database. The precipitation attributes extracted from the CCA serve to classify multiple extreme precipitation-runoff events. Third, for the development of forecasting models, we employed two internal ML sub models, one for baseflow and the other one for directflow, which were summed up to provide the total flow (Fig. 3c). Finally, we contrasted the performances of the developed forecasting models developed for increasing lead times and considering specialized models according to the classification of extreme events. A step-by-step explanation of the proposed methodology is presented in the following subsections.

### 3.1. Determination of nearly-independent peak runoff events and flow separation

Extreme hydrological events were selected from the complete runoff time series by using the WETSPRO time series tool (Willems, 2009). The WETSPRO is based on a peak-over-threshold approach controlled by two parameters to be calibrated, the recession time and a peak height difference of two consecutive runoff events. For the flow separation task, the generalized Chapman filter technique was selected following the recommendations of Willems (2009) and Corzo and Solomatine (2007b). The flow filtering principle is based on a numerical digital filter implemented through a linear reservoir modeling concept. The flow separation method is also available within the WETSPRO tool.

The calibration of the WETSPRO tool was done with the following parameter values. First, an inter-event time of 120 h, i.e., two consecutive events are considered nearly independent when separated by a period of at least 5 days. Secondly, a runoff maximum drop-down ratio of 0.6, which means that runoff, $q$, drops down in between two consecutive events to a ratio $\frac{q_{min}}{q_{max}} < 0.6$). Based on the calibration, 81 nearly-independent peak flow events could be delineated. Fig. 4a shows the obtained hourly baseflow and directflow time series together with the 81 peak flow events depicted as blue dots, while Fig. 4b plots the

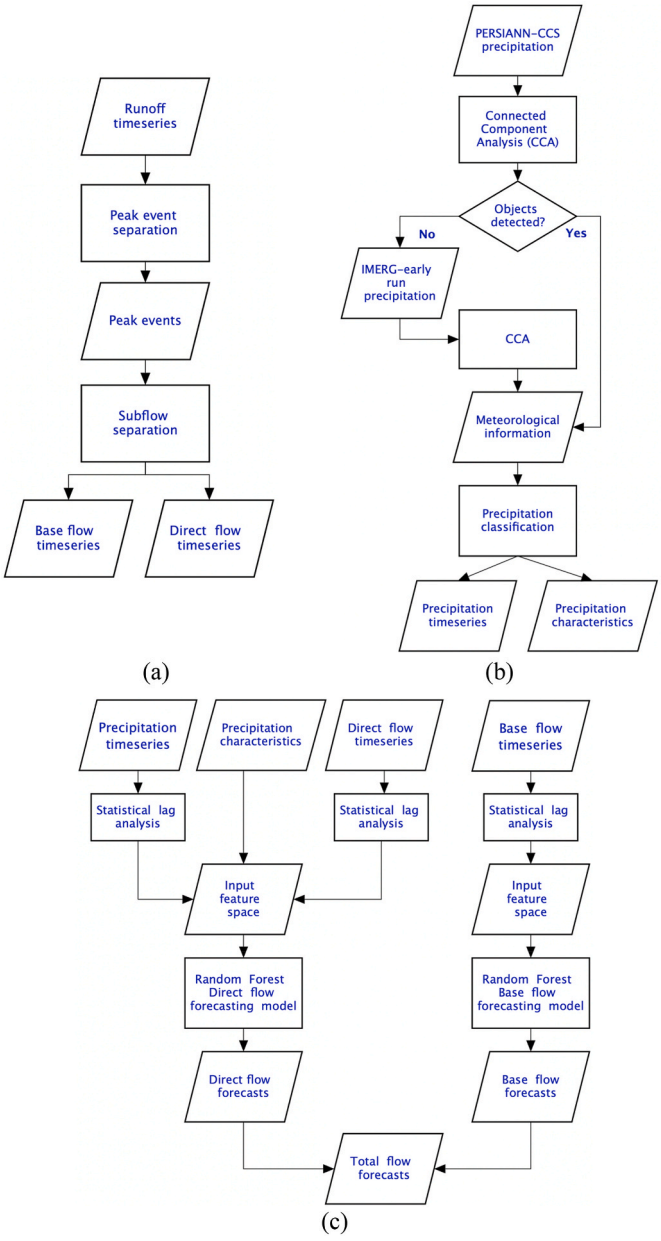



**Fig. 3.** Scheme of the methodology for developing peak runoff forecasting models, (a) extreme peak runoff selection and flow separation, (b) satellite precipitation processing, and (c) forecast modeling approach.

exceedance probability of total flow

### 3.2. Object-based method for deriving precipitation attributes

The precipitation imagery corresponding to the selected peak runoff events was processed using the object-based CCA developed by Laverde-Barajas et al. (2019). The CCA extracts precipitation attributes from RS data through a multidimensional connected component labeling algorithm. The extracted attributes provide a physical description of precipitation events (localization, centroids, area), and meteorological features (duration, volume, maximum intensity, etc.).

Moreover, according to the modular approach for precipitation data acquisition, the IMERG-ER was used as a supplement dataset to the finest spatial resolution, the PERSIANN-CCS. For this, we applied a simple under-sampling technique. It consisted of dispersing the information contained in a pixel into a number of subdivided pixels, i.e., the IMERG-ER cell of size 0.1 × 0.1° was converted into ~ 6.4 cells with a resolution of 0.04 × 0.04°. The CCA was implemented through the scikit-image processing package in Python® version 3.7 (van der Walt et al., 2014). The CCA is fully detailed in Laverde-Barajas et al. (2019), and can be summarized as follows:

i. Precipitation retrieval of the selected peak runoff events, and imagery clipping to the Jubones basin (Fig. 5a).
ii. Detection and localization (latitude, longitude, see Fig. 5b) of precipitation objects. For this, a detection sensitivity threshold is defined to remove noise and keep only clear precipitation objects in the precipitation imagery (Fig. 5c). The detection sensitivity was calibrated on a trial-and-error basis with a precipitation threshold volume of 0.5 mm. This means that precipitation objects associated with a depth of less than 0.5 mm were trimmed-off.
iii. Precipitation object filtering according to size criteria. We defined a minimum object area corresponding to two pixels of the finest-resolution product (~39 $km^2$).
iv. Morphologically closing of the precipitation objects found in step (iii). For this, a dilation-and-erosion algorithm was used to refine precipitation objects (Fig. 5d); dilation expands objects while erosion removes the boundaries of the expansion.
v. Extraction of physical (centroid and extension area) and meteorological attributes (volume of precipitation, maximum intensity, precipitation duration) from the objects refined in step (iv). We defined that two precipitation objects are considered consecutive (i.e., belong to the same event) when the time between their appearance is shorter than 3 h. This threshold was also calibrated on a trial-and-error basis.

### 3.3. Classification of precipitation events associated with extreme hydrologic events

From the CCA, two attributes of each precipitation event were derived: the extension of the precipitation objects (local and spatial extensive) and the duration of the events (short and long). As a result, four precipitation event classes could be defined: i) Local and short-duration extreme events (LSE), ii) Local and long-duration extreme events (LLE), iii) Spatially extensive extreme events (SEE), and iv) Spatially extensive and long-duration extreme events (SLE) (Laverde-Barajas et al., 2019). For instance, for given thresholds of 50 $km^2$ and 7 h, the LLE class will be defined by events with precipitation objects whose areal extension is less than 50 $km^2$, and whose duration is longer than 7 h.

### 3.4. Development of peak runoff forecasting models

The forecasting of peak runoffs was obtained by summing up the forecasts of two internal models, one model for baseflow and one model for directflow. The purpose of separating total flow into baseflow and directflow was to characterize and separately model the different orders of magnitude of hydrological processes (Willems, 2009). The flow separation was done for the base model (considering all extreme hydrological events) as well as for each precipitation event class (see Fig. 3). All the models were developed using the RF algorithm for regression, which is detailed in the following subsection.

For the baseflow model, we assumed a slow (neglectable) response of this subflow to precipitation. As a result, baseflow is assumed to be solely affected by gradual changes in the past baseflow, i.e., fully autoregressive. On the contrary, the quick response of directflow to precipitation was assumed to be influenced by changes in precipitation and directflow.

The input feature space construction for the RF models was conducted following the methodology used by Muñoz et al. (2018). In summary, the input feature space was formed by three elements. First,

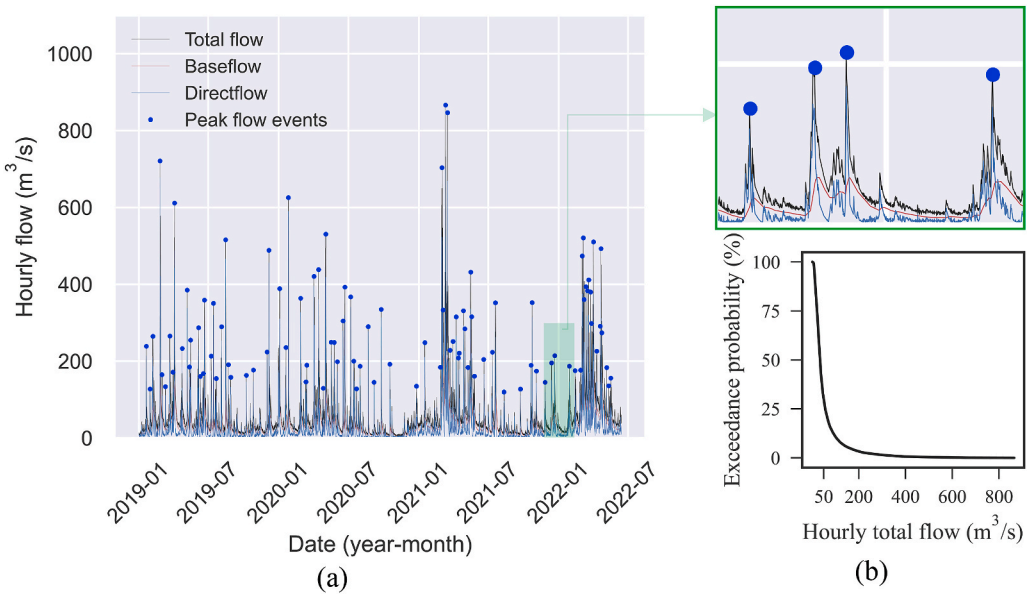



**Fig. 4.** (a) Directflow and baseflow separation from the total flow time series at the outlet of the Jubones basin. Peak flow events selected with the WETSPRO tool are displayed as blue dots. (b) Exceedance probability of total flow for the study period (01/01/2019 to 13/06/2022).

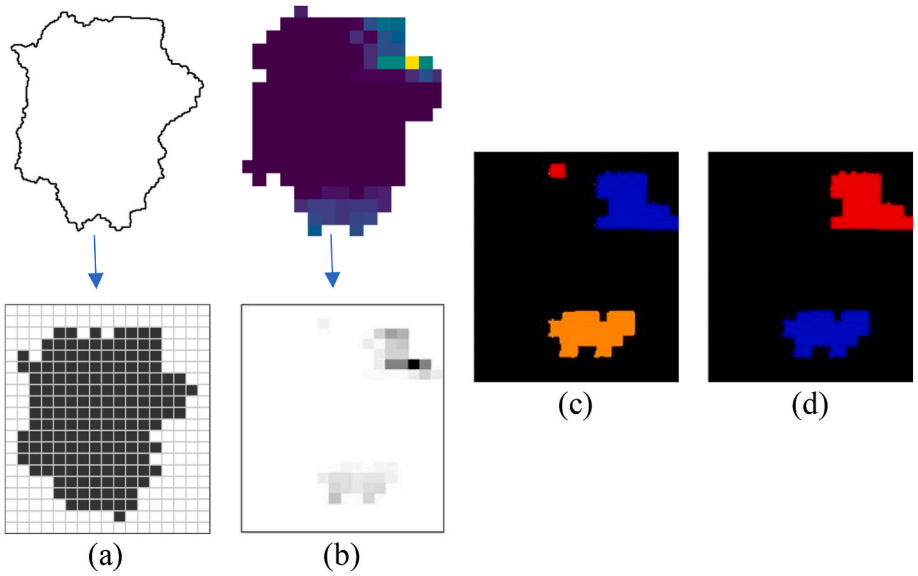



**Fig. 5.** Precipitation identification with an object-based Connected Component Analysis (CCA) Illustration of the PERSIAN-CCS 2021-12-25 05:00 UTC image. (a) Jubones basin boundary, (b) Precipitation identification in mm from the PERSIANN-CCS product, (c) Identification of three precipitation objects with the CCA, and (d) Final identification of two precipitation objects after object size filtering and morphological closing.

hourly precipitation (for each pixel) and runoff timeseries (i.e., baseflow and directflow). Second, two precipitation characteristics from the CCA: total volume and total area of the precipitation objects. And third, past lag information of precipitation (for each pixel) and runoff. The number of precipitation and runoff lags were determined according to statistical correlation analyses: partial- and auto-correlation functions for runoff, and cross-correlation functions for precipitation.

Moreover, we performed a feature selection analysis to reduce the input dimension of the RF models. For this, we used a sensitivity analysis aimed at calculating the relative importance of each feature to the output (Cortez, 2010). This is done by measuring the variance of the output produced by a single feature without considering the interaction between features. The purpose was to retain only features accounting for at least 80% of the total relative importance. The variance produced by a single feature ($V_k$) and its relative importance ($R_k$) can be calculated using equations (1) and (2), respectively.

$$V_k = \frac{\sum_{j=1}^{L} [\hat{y}_{t-k}(j) - \overline{\hat{y}_{t-k}(j)}]^2}{L-1} \quad \text{(equation 1)}$$

$$R_k = \frac{V_k}{\sum_{i=1}^{m} V_i} \cdot 100\,\% \quad \text{(equation 2)}$$

Where, $\hat{y}_{t-k}(j)$ is the model output obtained by holding all $m$ features at their average values except $\hat{y}_{t-k}(j)$, which varies through its entire range with $j \in \{1, \ldots, L\}$ levels.

#### 3.4.1. Random Forest (RF) for regression

RF is a ML technique of supervised learning where multiple decorrelated regression trees are constructed by relating the input feature space with output(s) (Breiman, 2001). A regression tree is a tree-like regression model, with the zero-order regression models (i.e. constants) in the leaves, also proposed by Breiman (2017). For this, a set of hierarchically organized conditions are successively applied to the input feature space. The success of the RF algorithm is attributed to a bagging technique that randomly selects multiple resampled datasets from the input feature space. This assures decorrelation between stochastically formed trees (models).

The RF forecasting models were implemented using the scikit-learn package for ML in Python® version 3.7 (Pedregosa et al., 2011). The full explanation of the RF algorithm is available in Breiman (2001), and can be summarized as:

i. Determination of multiple random bootstrap samples from the input feature space. Each sample is used to form a single regression tree. For the construction of each bootstrap, the out-of-bag (OOB) process is applied with approximately two-thirds of the input feature space. This ensures unbiased estimates of the regression, and more importantly, enables calculation of the generalization error (OOB error) which is comparable as using an independent testing subset of the same size as the training subset (Breiman, 2001).
ii. Splitting of the data for each sample determined in step (i). This process is performed at each node of each tree. Here, a maximum number of features must be defined to perform the best split from the total number of predictors in the feature space. The objective is to avoid overfitting by ensuring the variety and nonexistence of duplicated models.
iii. Growth of all the models formed in step (i) with the splits defined in step (ii). For this, a maximum size must be defined either by setting a hyperparameter controlling the maximum depth or the minimum number of samples expected in the final node. Controlling the maximum size of a tree (pruning) aims at reducing the structural complexity of the model, leading to noise reduction and model parsimony.
iv. Determination of the regression output as the mean response from all regression trees.

The implementation of the RF demands the tuning of several hyperparameters. For hydrological applications, the most influencing hyperparameters are the number of trees (n_trees), the maximum depth for pruning (max_depth), and the maximum number of features to perform the splits (max_features) (Contreras et al., 2021). We obtained the optimal combination of these three hyperparameters based on a random grid search implemented under a 10-fold cross-validation algorithm to prevent overfitting. The Nash-Sutcliffe Efficiency (NSE) coefficient (defined in the following subsection) between simulations and observations was used as a measure of agreement for the training subsets of each model. Table 1 shows the search space (domain) of the selected RF hyperparameters for the optimization task.

**Table 1**
Search space of the RF hyperparameters.

| Hyperparameter | Domain |
|---|---|
| n_trees[a] | 20; 1000; 10 |
| max_features | *n_features* [b], $n\_features^{(1/2)}, \log_2(n\_features)$ |
| max_depth[a] | 40; 800; 10 |

[a] Domain defined by min, max, and increment.
[b] n_features refers to the number of estimators (features) in the input feature space.

#### 3.4.2. Model evaluation

For model evaluation, instead of selecting a fraction of the total number of peak flow events for training/testing purposes, we employed the leave-one-out cross-validation (LOOCV) algorithm (Haddad et al., 2013). This means that each event was treated as an independent testing subset while the remaining events served for training purposes. In the end, the overall performance of a model (NSE) was calculated by averaging the NSE coefficients on the testing subsets when all events were tested separately. This was done since only a few number of events might available after the classification task.

For each event, we simulated the peak runoff inside a 24-h window for capturing the entire hydrograph. To quantify model performance, we used a collection of four metrics following the guidelines proposed by Moriasi et al. (2007). The *NSE* coefficient (Nash and Sutcliffe, 1970) was set as the reference metric for measuring and comparing the overall fit of model simulations to observations. The evaluation was complemented with the Kling-Gupta Efficiency (*KGE*) (Gupta et al., 2009) to account for peak flow underestimations and low flow overestimations, the Percent Bias (*PBIAS*), and the Root Mean Square Error (*RMSE*). Moreover, we contrasted the average NSE coefficients of each model with the correspondent OOB errors (also calculated as NSE coefficients). The corresponding equations are listed in Table 2.

## 4. Results

### 4.1. Object-based CCA and precipitation-event classification

First, CCA results for the 81 nearly-independent peak flow events obtained in section 3.1 showed that for 15 events (19%) there was no clear precipitation signal from the PERSIANN-CCS product. For these 15 cases, we applied the CCA to the IMERG-ER dataset. The utility of the precipitation modular approach (section 3.2) can be seen in the peak hydrological events depicted in Fig. 6. For the event from 2019 to 07-13 20:00 to 2019-07-14 20:00 UTC (Fig. 6a), it seemed evident that the higher resolution of the PERSIANN-CCS product lead to a stronger precipitation-runoff relation when compared to precipitation obtained from the IMERG-ER product. Thus, the precipitation data from the PERSIANN-CCS were used to feed the forecasting models. The opposite was true for the event from 2019 to 10-07 at 16:00 to 2019-10-08 at 16:00 UTC (Fig. 6b), where the PERSIANN-CCS signal was nonexistent for almost 24 h before the runoff peak, whereas there is a significant amount of precipitation from the IMERG-ER product.

From the CCA, we derived duration and extension thresholds of 7 h and 50 km$^2$, respectively (Fig. 7). These thresholds served to classify peak runoff events into the four precipitation classes, 23 events for the LSE class, 24 for the LLE, 25 for the SEE, and 9 for the SLE. Moreover, analysis of the centroid occurrence of the precipitation objects did not reveal any precipitation hotspot in the basin that could be associated

**Table 2**
Performance metrics.

| Metric | Equation | Range | Ideal value |
|---|---|---|---|
| NSE | $1-\frac{\sum_{i=1}^{n}(Q_s(i)-Q_o(i))^2}{\sum_{i=1}^{n}(Q_o(i)-\overline{Q_o})^2}$ | $-\infty, 1$ | 1 |
| KGE | $1-\sqrt{(r-1)^2+(\alpha-1)^2+(\beta-1)^2}$ | $-\infty, 1$ | 1 |
| PBIAS | $\frac{\sum_{i=1}^{n}(Q_s-Q_o)}{\sum_{i=1}^{n}Q_o}$ x 100% | $-100, +\infty$ | 0 |
| RMSE | $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_s-Q_o)^2}$ | $0, +\infty$ | 0 |

Where $n$ is the number of instances, $Q_s(i)$ is the simulated runoff at time $i$, $Q_o(i)$ is observed runoff at time $i$, $\overline{Q_o}$ is the mean observed runoff, $\overline{Q_s}$ is the mean simulated runoff, $r$ is the correlation coefficient between $Q_s$ and $Q_o$, $\alpha=\frac{\sigma_s}{\sigma_o}$ is the variability ratio, $\beta=\frac{\overline{Q_s}}{\overline{Q_o}}$ is the bias ratio, and $\sigma$ stands for the standard deviation.

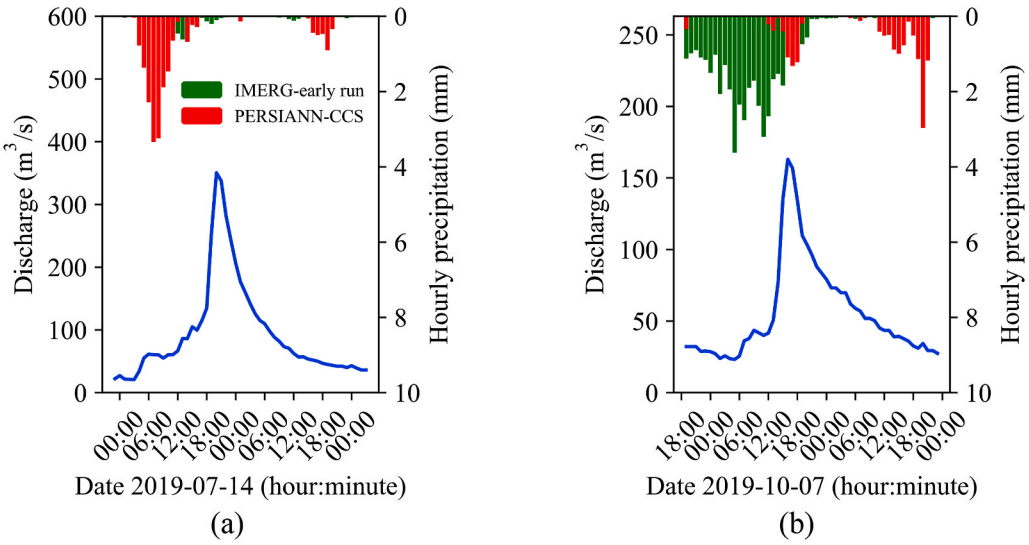



**Fig. 6.** Illustration of the need of the satellite precipitation-retrieval modular approach. (a) Significant amount of precipitation better estimated by the PERSIANN-CCS product for the extreme hydrological event from 2019 to 07-13 18:00 to 2019-07-14 18:00 UTC. (b) Significant amount of precipitation better estimated by the IMERG-ER product for the extreme event 2019-10-07 12:00 to 2019-10-08 12:00 UTC.

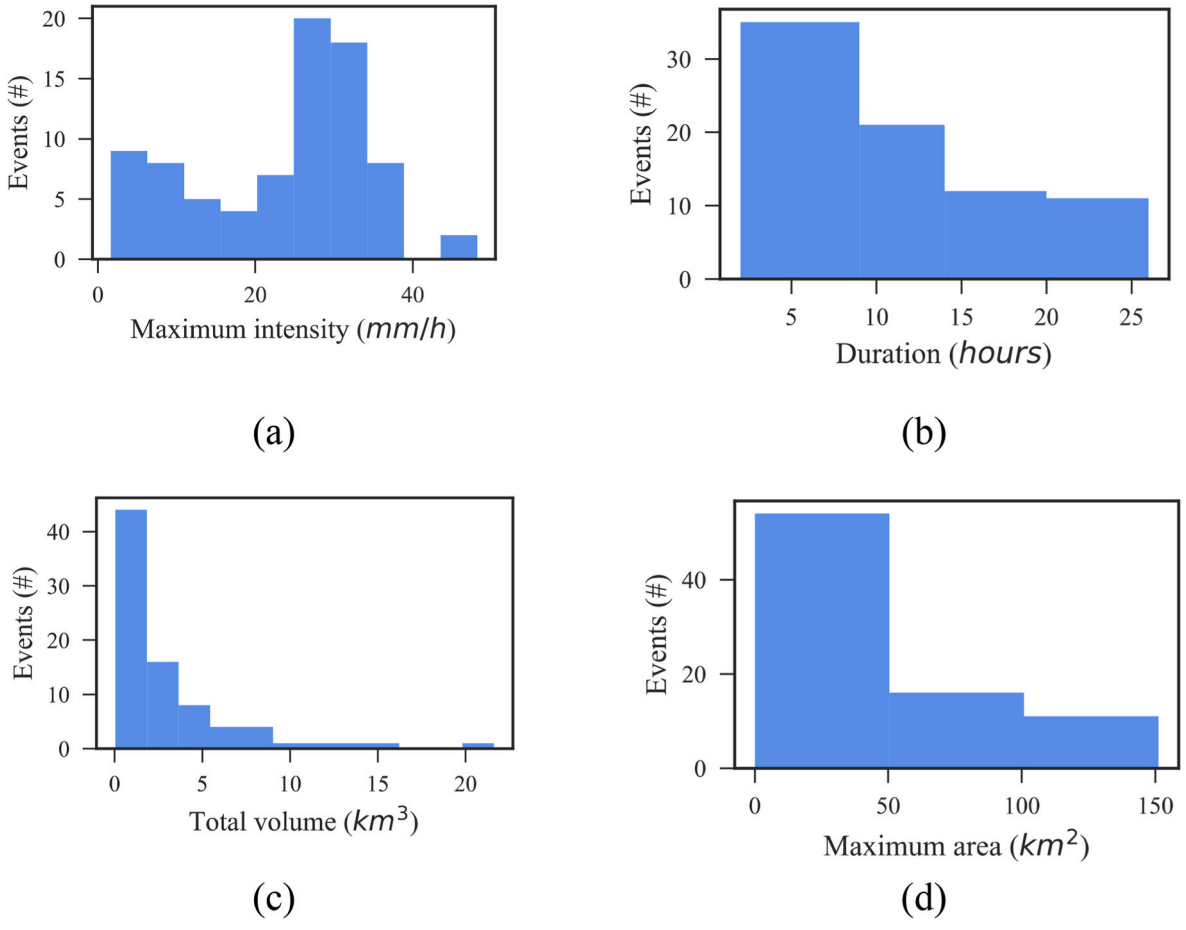



**Fig. 7.** Meteorological precipitation information retrieved from 81 extreme hydrological events: (a) maximum intensity, (b) event duration, (c) total volume, and (d) maximum area.

with peak runoff events (see Fig. 8). There was no evidence that centroid occurrence is driven or can be related to any physical attribute of the Jubones basin (e.g., soil type, land use, elevation, topography, etc.). This might indicate the nonexistence of orographic precipitation enhancement (i.e., cloud formation due to orographic lifting of air masses).

As mentioned in section 3.4, the input feature space to each model with a certain lead time was partly formed with lagged information of

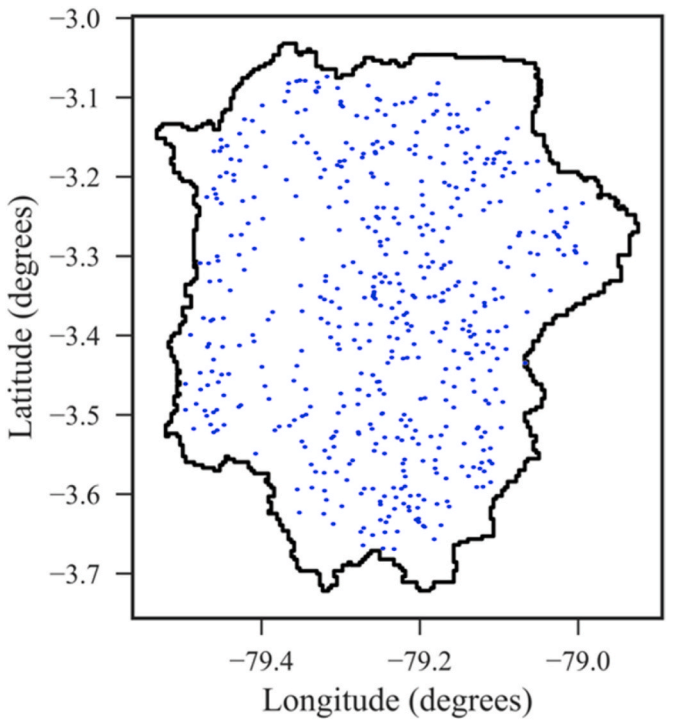



**Fig. 8.** Localization of the centroids of precipitation objects (blue dots) associated with peak runoff events in the Jubones basin.4.2 Development and evaluation of peak runoff forecasting models.

precipitation and runoff. For runoff, results of the partial- and auto-correlation functions suggested using 12 lags (hours), with a 95% confidence level for both correlation values. Similarly, for precipitation, the cross-correlation function determined correlations higher than 0.2 for 13 lags (hours). Cross-correlation results are consistent with the estimated concentration-time of the Jubones basin. The concentration-time was estimated at 11 h by averaging the outputs of the equations of the U. S. Army Corps of Engineers, Johnstone, and Giandotti, being the equations recommended in de Almeida et al. (2014) for the basin extension. For the RF hyperparameterization of each model, we obtained averaged NSE coefficients between simulations and observations always higher than 0.98. Table 3 exemplifies the optimized combination of hyperparameters found for the forecasting models of 1-h lead time.

Concerning model efficiencies, Table 4 presents the averaged performance metrics for the base models, and for the specialized peak runoff forecasting models. In all cases, we present separately the performances for the baseflow, directflow, and the resulting total flow. The color mapping of this table was done on a column-by-column basis. This allows comparing the performances of the models (for a given metric) across lead times, and between the base and the specialized models. The darkest colors represent the best performances.

The first striking result visible in this table is that the base models proved to be satisfactory, with NSE coefficients (LOOCV) for total flow varying from 0.86 to 0.59, for lead times between 1 and 6 h, respectively. We contrasted these values with the NSE coefficients obtained for the validation subset, i.e., OOB errors (results not shown), and we found an overall pattern for the OOB errors to be higher than NSE values using the LOOCV. We found a maximum difference of 0.14 for the 4-h directflow model, whereas for baseflow models the differences were lower (maximum 0.06 for the 1-h models). The higher OOB errors can be attributed to the fact that the validation is performed on a randomly selected one-third of the training subset; thus, considering the 24-h window of each event, it might be possible that most of the scrutinized runoff does not correspond to peak values. As a result, the calculation of NSE coefficients using the LOOCV provides a more severe evaluation of the forecasting models.

**Table 3**
RF hyperparameterization of the forecasting models for the 1-h lead time.

| | | Random Forest hyperparameters | | |
|---|---|---|---|---|
| | Events [#] | n_trees | max_features | max_depth |
| Base model | 81 | 300 | 21 | 200 |
| LSE | 23 | 280 | 9 | 220 |
| LLE | 24 | 250 | 21 | 190 |
| SLE | 25 | 300 | 21 | 160 |
| SEE | 9 | 300 | 9 | 180 |

Moreover, according to the criterion of Singh et al. (2004), the obtained RMSE-values for the base models were also satisfactory for all lead times since their magnitudes were lower than half the standard deviation of measured total flow, 126.2 $m^3 s^{-1}$. The evolution of model performance with lead time is explained by previous thoughts and follows a logical path: the forecast ability of RF decreases with increasing lead time. Moreover, it was also clear that the modeling difficulty came from the modeling of directflow where NSE-values for the base models decayed to 0.36 (6-h lead time). Nevertheless, the satisfactory performance for total flow was a remarkable outcome since the input feature space was derived from non-validated near-real-time satellite estimates using only the object-based CCA as the processing tool.

Once the base models were evaluated, the further analysis focused on the specialized peak runoff forecasting models. First, for total flow, it is apparent from Table 3 that spatially-extensive and short-duration events (SEE) produced the lowest performances (NSE-values across lead times) when compared to the remaining event classes. These results were confirmed by the OOB errors, which followed a similar pattern across event classes. Therefore, it is apparent that SEE are the major source of error for the base models. Here, the hypothesis is that for SEE, precipitation over a mosaic of land uses and soil types produces complex directflow responses that are difficult to be learned by RF regressors. The reason is that small precipitation volumes over extensive areas might be lost before converting into the runoff, especially in non-saturated conditions. Although this issue is strongly linked to land uses, soil types, and the saturation state of the basin, such biophysical information was neither available (soil type and soil moisture) nor updated (land use) for the basin, and could therefore not be used as additional inputs to the forecasting models.

This can be also seen in Fig. 9, where it stands out that overall LSE and LLE perform better than the base model, whereas SLE models perform almost similar but have the advantage of forecasting the highest peak runoffs. For local events, irrespective of their duration (LLE and LSE), the runoff response strongly depends on the land use and soil characteristics where the precipitation occurs. In these cases, the error seems to be absorbed by the RF algorithm by relying on more specific trees (higher values for the max_depth hyperparameter).

Regarding total flow separation for the specialized forecasting models, the NSE-values for the 1-h lead time for baseflow and directflow were comparable for local precipitation events irrespective of their duration (LLE and LSE). On the contrary, for spatially extensive events (SLE and SEE), there was a clear reduction in the ability of RF models to forecast directflow. This issue becomes critical as the lead time increases, where NSE-values for directflow tend to completely deteriorate for the 6-h case (SEE). Overall, for a given lead time, SLE and SEE models depicted the lowest NSE values for directflow, and consequently total flow. Therefore, the considerably greater amount of precipitation input features is rather producing noise to the directflow models. In such cases, the forecasting ability tends to rely more on their autoregressive power and not on what the models can learn from the processed satellite precipitation.

Concerning the remaining performance metrics (KGE, RMSE, and PBIAS), we found patterns similar to NSE, between specialized runoff forecasting models, and across lead times. For instance, for any lead time, the PBIAS for SEE had the highest values between precipitation

**Table 4**
Model efficiencies (LOOCV evaluation framework) for the base and specialized forecasting models across lead times.

| | | Baseflow | | | | Directflow | | | | Total flow | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Lead time** | **Class** | **NSE** | **KGE** | **RMSE** | **PBIAS** | **NSE** | **KGE** | **RMSE** | **PBIAS** | **NSE** | **KGE** | **RMSE** | **PBIAS** |
| **1h** | Base model | 0,86 | 0,70 | 20,50 | 15,40 | 0,88 | 0,85 | 28,60 | 4,50 | 0,86 | 0,81 | 33,55 | 4,85 |
| | LLE | 0,89 | 0,86 | 5,06 | 5,10 | 0,81 | 0,83 | 19,58 | 2,53 | 0,87 | 0,84 | 38,89 | 5,69 |
| | LSE | 0,87 | 0,79 | 5,64 | 11,12 | 0,81 | 0,79 | 15,68 | 10,56 | 0,78 | 0,71 | 29,48 | 9,87 |
| | SLE | 0,85 | 0,80 | 24,56 | 8,79 | 0,70 | 0,72 | 57,46 | 15,40 | 0,70 | 0,74 | 63,27 | 19,80 |
| | SEE | 0,78 | 0,68 | 27,86 | 18,89 | 0,30 | 0,33 | 93,65 | 101,50 | 0,29 | 0,32 | 89,63 | 51,22 |
| **2h** | Base model | 0,82 | 0,75 | 16,58 | 6,52 | 0,75 | 0,76 | 46,28 | 4,50 | 0,82 | 0,75 | 65,50 | 5,87 |
| | LLE | 0,87 | 0,88 | 6,22 | 7,50 | 0,71 | 0,75 | 28,80 | 3,69 | 0,70 | 0,72 | 42,77 | 8,70 |
| | LSE | 0,79 | 0,77 | 6,98 | 5,44 | 0,65 | 0,65 | 33,58 | 8,06 | 0,65 | 0,64 | 55,16 | 10,45 |
| | SLE | 0,68 | 0,77 | 35,56 | 9,02 | 0,58 | 0,59 | 70,05 | 16,06 | 0,49 | 0,53 | 84,50 | 20,50 |
| | SEE | 0,48 | 0,50 | 37,81 | 23,06 | 0,37 | 0,37 | 123,40 | 121,33 | 0,22 | 0,15 | 131,84 | 83,01 |
| **4h** | Base model | 0,78 | 0,69 | 35,40 | 7,80 | 0,63 | 0,70 | 50,40 | 0,48 | 0,69 | 0,70 | 81,41 | 9,87 |
| | LLE | 0,79 | 0,75 | 9,33 | 7,80 | 0,38 | 0,43 | 40,22 | 5,68 | 0,63 | 0,66 | 62,50 | 10,44 |
| | LSE | 0,60 | 0,48 | 23,87 | 11,56 | 0,35 | 0,31 | 70,80 | 12,50 | 0,50 | 0,52 | 66,90 | 15,62 |
| | SLE | 0,49 | 0,46 | 40,52 | 13,25 | 0,19 | 0,22 | 65,80 | 7,95 | 0,26 | 0,27 | 62,50 | 21,60 |
| | SEE | 0,32 | 0,28 | 38,96 | 27,13 | 0,06 | 0,10 | 140,52 | 129,12 | 0,09 | 0,13 | 146,58 | 88,56 |
| **6h** | Base model | 0,67 | 0,71 | 26,50 | 3,56 | 0,40 | 0,38 | 68,50 | 6,89 | 0,59 | 0,55 | 75,34 | 5,11 |
| | LLE | 0,70 | 0,66 | 14,23 | 8,08 | 0,32 | 0,38 | 50,24 | 6,06 | 0,40 | 0,39 | 75,48 | 8,97 |
| | LSE | 0,59 | 0,62 | 26,38 | 15,60 | 0,18 | 0,21 | 85,67 | 4,03 | 0,35 | 0,41 | 125,40 | 16,58 |
| | SLE | 0,45 | 0,42 | 45,60 | 17,69 | 0,14 | 0,15 | 89,96 | 17,36 | 0,29 | 0,28 | 133,68 | 28,60 |
| | SEE | 0,25 | 0,26 | 55,73 | 29,86 | -3,68 | -1,67 | 119,50 | 145,10 | -0,01 | 0,02 | 158,80 | 112,50 |

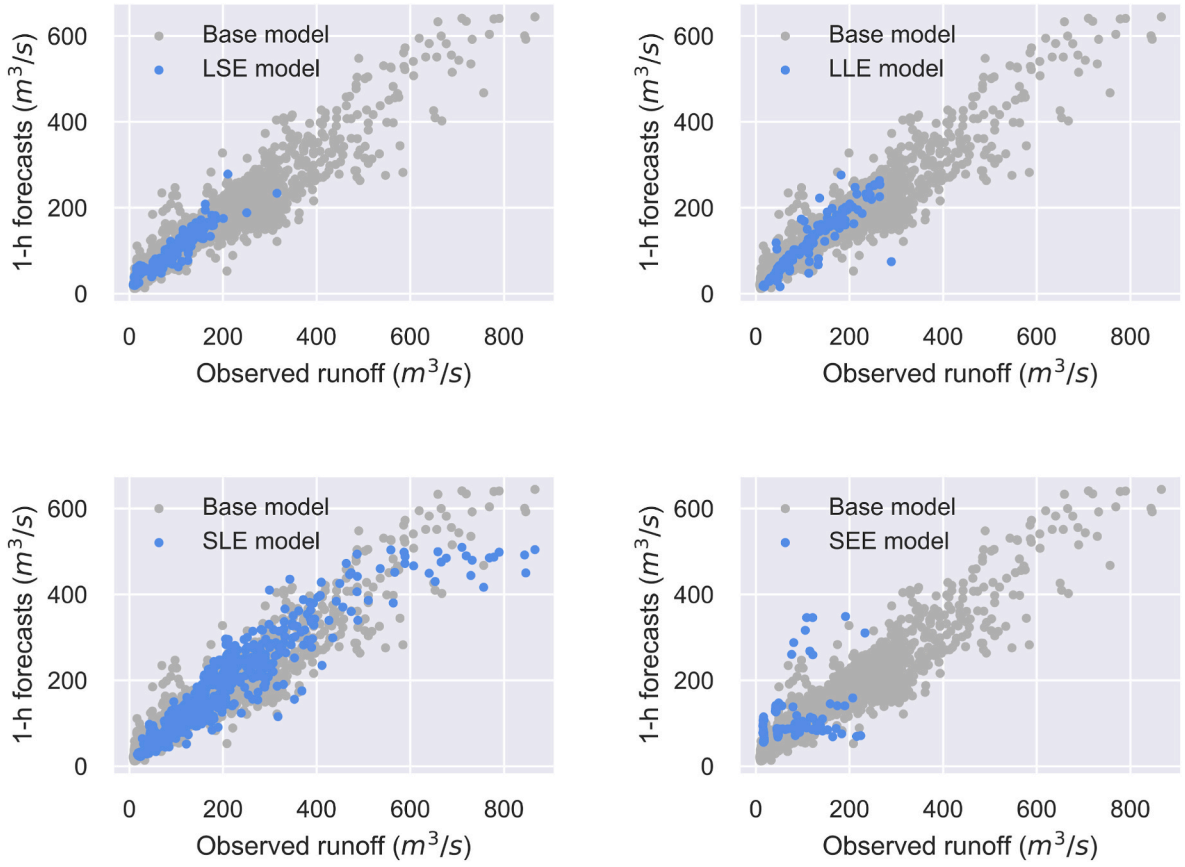



**Fig. 9.** Comparison of the scatter plots of the observed and forecasted runoff for the base model and the specialized models for the 1-h lead time.

event classes, i.e., the highest propensity of forecasted values to be larger than measured total flow. PBIAS for SEE were always higher than 51%, with values up to 113% for the 6-h lead time. Similarly, the KGE metric, which exposes runoff variability to a greater extent than NSE, revealed the lowest efficiencies for SEE for all lead times. Physically, NSE- and KGE-values might be explained by the fact that the precipitation-runoff

correspondence is clear (straightforward) for the cases when either soil saturation is reached or infiltration capacity is exceeded, SLE and LLE, respectively. The straightforward precipitation-runoff relations seem to be well detected by the RF models, especially LLE models, where the highest NSE- and KGE-values were obtained.

## 5. Discussion

In this study, the RF algorithm was used to develop peak runoff forecasting models for a representative basin of the tropical Andes, where physically-based modeling is restricted by the lack of sufficient information. The methodology of this study proposes a solution for exploiting near-real-time satellite precipitation data even without validation with ground precipitation networks.

We developed general base models for lead times between 1 and 6 h to account for flash-floods in the Jubones basin, which although representing a particular solution for the MSF hydropower dam, is useful for planning the operation of other dams under peak extreme runoff conditions. In addition to the base models, we focused on characterizing extreme hydrological events by analyzing satellite precipitation through an object-based CCA. The development of specialized models according to precipitation characteristics (duration and extension) served to identify the hidden strength and weaknesses of the already satisfactory base models.

The performances obtained for the base models (NSE = 0.87) are comparable to the results obtained in other studies using traditional physically-based models, for instance NSE = 0.92 using HEC-RAS (Belabid et al., 2019), NSE = 0.58 using wflow-sbm (Laverde-Barajas et al., 2020), and NSE = 0.94 using the hydrologic-hydraulic HiRes-Flood-UCI model (Nguyen et al., 2015). And although this study did not aim at outperforming physically-based models, a clear advantage of the models hereby developed is the possibility to exploit raw near-real-time satellite precipitation. This possibility is, however, feasible under a modular approach for data acquisition, where a second satellite source is used to overcome detection issues of the primary satellite source. We are aware, however, that further analyses must be performed for choosing not only existing precipitation signals but the satellite source, or even data fusion such as the efforts of Chen et al. (2020) and Xu et al. (2020), presenting the highest correlations with observed runoff.

The framework for unveiling the strengths and weakness of the base models can be replicated to understand the reasons behind unacceptable low performances (e.g., negative NSE), see for instance the study of Belabid et al. (2019). The superiority in performance of the developed local models when compared to spatially extensive events can be explained by the straightforward infiltration- and saturation-excess runoff generation processes in reduced portions of the basin. Conversely, whenever precipitation is extensively distributed within the basin, the forecasting models lower their ability to characterize and learn the specificities of the multiple precipitation-runoff relations. For such cases, the forecasting ability is attributed to the autoregressive dependency in the flow time series. The performances of the specialized forecasting models revealed the need to include information describing the dynamics of antecedent soil saturation during extreme events. This is especially required for initializing the forecast of short-duration precipitation events (SEE and LSE). The antecedent soil saturation state will serve to explain why short-duration and non-extreme precipitation intensities can trigger extreme hydrological events. Given the previous assumption, a future direction would be to include satellite soil moisture observations to improve forecasting efficiencies in ungauged basins, as done in the study of Massari et al. (2018).

We are also aware that the findings of this study were obtained with a relatively shorth-length database when compared to other ML studies; however, the use of the RF together with processing tools, and severe evaluation framework for reducing overfitting served to ensure the use of the models for this study and for the daily operation of the MSF hydropower plant. Particularly, for the Jubones basin, we did not find a pattern or hotspot of the precipitation storms that triggered extreme runoff responses in the study period. This has direct implications for the ability of RF models to recognize patterns and demonstrates the necessity of developing specialized forecasting models according to other precipitation characteristics, and not only in distributed modeling (subbasins), which is commonly employed by traditional physically-based models. Finally, the findings and limitations encountered in this study open the path for future research on exploring additional ML techniques for the modeling of spatially-extensive events, or even model ensemble strategies.

## 6. Conclusions

The main aim of this study was to demonstrate the possibility of using near-real-time satellite sources, the PERSIANN-CCS and IMERG-ER products, for the development of peak runoff forecasting models in a complex tropical basin. A combination of FE strategies was applied to assist data ingestion into RF models and to ensure that the optimal model structures were chosen. The second aim was to gain hydrological knowledge regarding the functioning of the study basin and to unveil the strengths/weaknesses of the developed forecasting models. For this, specialized forecasting models were developed according to hydrometeorological criteria. Moreover, the necessity to use a supplementary satellite source for dealing with satellite detection failures was demonstrated.

A major finding emerging from this research is that an improved representation of near-real-time satellite precipitation enhances the performance of peak runoff forecasting models. On one hand, precipitation classification served to identify the precipitation scenarios misleading the learning process of RF models, while on the other hand total flow separation was applied to attenuate the RF difficulty to forecast total flow by considering the ability and difficulty to forecast baseflow and directflow, respectively. The research also showed that for the Jubones basin, spatially extensive events are the most difficult precipitation scenarios to model without deep characterization of the study area (land use, soil moisture, and topography, among other features). The description of soil saturation conditions might also enhance runoff forecasting associated with local precipitation events, yet their higher efficiencies are attributed to the straightforward infiltration- and saturation-excess runoff generation relations.

It is worth mentioning that since the developed models relied on a data-driven modeling technique (RF), additional verification is strongly recommended as the dataset increases with time, e.g., in a couple of years. To convert the developed models into a fully operational forecasting system, we recommend further investigation on designing strategies for choosing not only existing precipitation signals but the satellite source, or even data fusion presenting the highest correlations with current time runoff. The selection of the optimal precipitation input data can be complemented with uncertainty estimation through sensitivity analyses on the different precipitation satellite sources. All of this, of course, demands exploration of additional near-real-time satellite sources and data fusion techniques, and since we are dealing with hourly forecasts, the computation times must also be congruent with the desired forecast horizons (order of hours).

A natural extension of this work would be to better represent the physical conditions of the basin before and during a precipitation event, not necessarily extreme, which might cause a peak runoff response. It would be also advisable to explore additional ML algorithms, hybridization, and/or model assembling aimed at maximizing the encountered forecasting efficiencies. Overall, the combined feature engineering methodology employed in this study, and the demonstrated potential to use raw near-real-time satellite precipitation sources provide hydrologists with new tools for improving near-real-time runoff forecasting in remote areas.

## CRediT authorship contribution statement

**Paul Muñoz:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft. **Gerald Corzo:** Methodology, Supervision, Writing - review & editing. **Dimitri Solomatine:** Methodology, Supervision, Writing - review & editing. **Jan Feyen:** Supervision, Writing - review & editing. **Rolando Célleri:** Funding acquisition, Methodology, Project administration, Resources, Supervision, 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.

## Acknowledgements

This research was funded by the Vice-rectorate for Research of the University of Cuenca (VIUC) and the Corporación Eléctrica del Ecuador - Unidad de Negocio Sur (CELEC-SUR) through the projects: “Diseño e implementación de un sistema de pronóstico hidrológico para la central hidroeléctrica Minas – San Francisco”, and “Data fusion of remote sensing products and machine learning feature engineering strategies for near-real time runoff forecasting”. Our thanks go to these institutions for their generous funding. The authors would also like to thank the Editors and anonymous Reviewers for their constructive comments that are greatly contributive to enriching the manuscript.

## References

Abda, Z., Zerouali, B., Chettih, M., Guimarães Santos, C.A., de Farias, C.A.S., Elbeltagi, A., 2022. Assessing machine learning models for streamflow estimation: a case study in Oued Sebaou watershed (Northern Algeria). Hydrol. Sci. J. 67 (9), 1328–1341. https://doi.org/10.1080/02626667.2022.2083511.

Akhtar, M.K., Corzo, G.A., van Andel, S.J., Jonoski, A., 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. Hydrol. Earth Syst. Sci. 13, 1607–1618. https://doi.org/10.5194/hess-13-1607-2009.

Belabid, N., Zhao, F., Brocca, L., Huang, Y., Tan, Y., 2019. Near-real-time flood forecasting based on satellite precipitation products. Rem. Sens. 11, 252. https://doi.org/10.3390/rs11030252.

Bhattacharya, B., Solomatine, D.P., 2005. Neural networks and M5 model trees in modelling water level–discharge relationship. Neurocomputing 63, 381–396. https://doi.org/10.1016/j.neucom.2004.04.016.

Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25, 197–227. https://doi.org/10.1007/s11749-016-0481-7.

Bontempi, G., Ben Taieb, S., Le Borgne, Y.-A., 2013. Machine learning strategies for time series forecasting. In: Lecture Notes in Business Information Processing. Springer, Berlin/Heidelberg, Germany, pp. 62–77.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Breiman, L., 2017. Classification and Regression Trees. Routledge, Oxfordshire, UK.

Chang, L.-C., Chang, F.-J., Yang, S.-N., Kao, I.-F., Ku, Y.-Y., Kuo, C.-L., Amin, I.M.Z.b.M., 2019. Building an intelligent hydroinformatics integration platform for regional flood inundation warning systems. Water 11, 9. https://doi.org/10.3390/w11010009.

Chen, C., Liaw, A., Breiman, L., 2004. Using Random Forest to Learn Imbalanced Data, vol. 110. University of California, Berkeley, pp. 1–12.

Chen, S., Xiong, L., Ma, Q., Kim, J., Chen, J., Xu, C., 2020. Improving daily spatial precipitation estimates by merging gauge observation with multiple satellite-based precipitation products based on the geographically weighted ridge regression method. J. Hydrol. (Amst.) 589, 125156. https://doi.org/10.1016/j.jhydrol.2020.125156.

Clark, M.P., Bierkens, M.F.P., Samaniego, L., Woods, R.A., Uijlenhoet, R., Bennett, K.E., Pauwels, V.R.N., Cai, X., Wood, A.W., Peters-Lidard, C.D., 2017. The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism. Hydrol. Earth Syst. Sci. 21, 3427–3440. https://doi.org/10.5194/hess-21-3427-2017.

Contreras, P., Orellana-Alvear, J., Muñoz, P., Bendix, J., Célleri, R., 2021. Influence of random forest hyperparameterization on short-term runoff forecasting in an Andean mountain catchment. Atmosphere 12, 238. https://doi.org/10.3390/atmos12020238.

Cortez, P., 2010. Sensitivity analysis for time lag selection to forecast seasonal time series using neural networks and support vector machines. In: Neural Networks (IJCNN), the 2010 International Joint Conference on 1–8. https://doi.org/10.1109/IJCNN.2010.5596890.

Corzo, G., Solomatine, D., 2007a. Knowledge-based modularization and global optimization of artificial neural network models in hydrological forecasting. Neural Network. 20, 528–536. https://doi.org/10.1016/j.neunet.2007.04.019.

Corzo, G., Solomatine, D., 2007b. Baseflow separation techniques for modular artificial neural network modelling in flow forecasting. Hydrol. Sci. J. 52 (3), 491–507. https://doi.org/10.1623/hysj.52.3.491.

Davis, C., Brown, B., Bullock, R., 2006. Object-based verification of precipitation forecasts. Part I: methodology and application to mesoscale rain areas. Mon. Weather Rev. 134, 1772–1784. https://doi.org/10.1175/MWR3145.1.

de Almeida, I.K., Almeida, A.K., Anache, J.A.A., Steffen, J.L., Alves Sobrinho, T., 2014. Estimation on time of concentration of overland flow in watersheds: a review. Geociencias 33, 661–671.

der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., 2014. scikit-image: image processing in Python. PeerJ 2, e453.

Dibike, Y.B., Velickov, S., Solomatine, D., Abbott, M.B., 2001. Model induction with support vector machines: introduction and applications. J. Comput. Civ. Eng. 15, 208–216. https://doi.org/10.1061/(asce)0887-3801(2001)15:3(208).

Elshorbagy, A., Corzo, G., Srinivasulu, S., Solomatine, D.P., 2010. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: concepts and methodology. Hydrol. Earth Syst. Sci. 14 https://doi.org/10.5194/hess-14-1931-2010.

Galelli, S., Castelletti, A., 2013a. Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling. Hydrol. Earth Syst. Sci. 17, 2669–2684. https://doi.org/10.5194/hess-17-2669-2013.

Gupta, H. v, Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J. Hydrol. (Amst.) 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.

Haddad, K., Rahman, A., A Zaman, M., Shrestha, S., 2013. Applicability of Monte Carlo cross validation technique for model development and validation using generalised least squares regression. J. Hydrol. (Amst.) 482, 119–128. https://doi.org/10.1016/J.JHYDROL.2012.12.041.

Hasan, M.M., Wyseure, G., 2018. Impact of climate change on hydropower generation in Rio Jubones Basin, Ecuador. Water Sci. Eng. 11, 157–166. https://doi.org/10.1016/j.wse.2018.07.002.

Hong, Y., Hsu, K.-L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteorol. 43, 1834–1853. https://doi.org/10.1175/JAM2173.1.

Hosseini, S.M., Mahjouri, N., 2016. Integrating support vector regression and a geomorphologic artificial neural network for daily rainfall-runoff modeling. Appl. Soft Comput. 38, 329–345. https://doi.org/10.1016/j.asoc.2015.09.049.

Hsu, K., Gao, X., Sorooshian, S., Gupta, H. v, 1997. Precipitation estimation from remotely sensed information using artificial neural networks. J. Appl. Meteorol. 36, 1176–1190. https://doi.org/10.1175/1520-0450(1997)036<1176:pefrsi>2.0.co;2.

Huang, P.-C., Lee, K.T., 2021. Influence of topographic features and stream network structure on the spatial distribution of hydrological response. J. Hydrol. (Amst.) 603, 126856. https://doi.org/10.1016/j.jhydrol.2021.126856.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., Yoo, S.-H., 2015. NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). Algorithm Theoretical Basis Document (ATBD) Version 4, 26.

Laverde-Barajas, M., Corzo, G., Bhattacharya, B., Uijlenhoet, R., Solomatine, D.P., 2019. Spatiotemporal analysis of extreme rainfall events using an object-based approach. In: Spatiotemporal Analysis of Extreme Hydrological Events. Elsevier, pp. 95–112. https://doi.org/10.1016/B978-0-12-811689-0.00005-7.

Laverde-Barajas, M., Perez, G.A.C., Chishtie, F., Poortinga, A., Uijlenhoet, R., Solomatine, D.P., 2020. Decomposing satellite-based rainfall errors in flood estimation: hydrological responses using a spatiotemporal object-based verification method. J. Hydrol. (Amst.) 591, 125554. https://doi.org/10.1016/j.jhydrol.2020.125554.

Li, B., Yang, G., Wan, R., Dai, X., Zhang, Y., 2016. Comparison of random forests and other statistical methods for the prediction of lake water level: a case study of the Poyang Lake in China. Nord. Hydrol 47, 69–83. https://doi.org/10.2166/nh.2016.264.

Li, J., Hsu, K.-L., AghaKouchak, A., Sorooshian, S., 2016. Object-based assessment of satellite precipitation products. Rem. Sens. 8, 547. https://doi.org/10.3390/rs8070547.

Massari, C., Camici, S., Ciabatta, L., Brocca, L., 2018. Exploiting satellite-based surface soil moisture for flood forecasting in the mediterranean area: state update versus rainfall correction. Rem. Sens. 10, 292. https://doi.org/10.3390/rs10020292.

Moreido, V., Gartsman, B., Solomatine, D.P., Suchilina, Z., 2021. How well can machine learning models perform without hydrologists? Application of rational feature selection to improve hydrological forecasting. Water 13, 1696. https://doi.org/10.3390/w13121696.

Moriasi, D.N., Arnold, J.G., van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE (Am. Soc. Agric. Biol. Eng.) 50, 885–900.

Mosavi, A., Ozturk, P., Chau, K.W., 2018. Flood prediction using machine learning models: literature review. Water (Switzerland) 10, 1–40. https://doi.org/10.3390/w10111536.

Muñoz, P., Corzo Perez, G.A., Solomatine, D., Feyen, J., Célleri, R., 2021. Use of near-real-time satellite precipitation data and machine learning to improve extreme runoff modeling. Earth Space Sci. Open Arch. 28 https://doi.org/10.1002/essoar.10508861.1.

Muñoz, P., Orellana-Alvear, J., Willems, P., Célleri, R., 2018. Flash-flood forecasting in an andean mountain catchment-development of a step-wise methodology based on the random forest algorithm. Water (Switzerland) 10. https://doi.org/10.3390/w10111519.

Nash, J.E., Sutcliffe, J. v, 1970. River flow forecasting through conceptual models part I — a discussion of principles. J. Hydrol. (Amst.) 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nguyen, P., Sellars, S., Thorstensen, A., Tao, Y., Ashouri, H., Braithwaite, D., Hsu, K., Sorooshian, S., 2014. Satellites track precipitation of super typhoon Haiyan. Eos, Transactions American Geophysical Union 95, 133–135. https://doi.org/10.1002/2014EO160002.

Nguyen, P., Thorstensen, A., Sorooshian, S., Hsu, K., AghaKouchak, A., 2015. Flood forecasting and inundation mapping using HiResFlood-UCI and near-real-time satellite precipitation data: the 2008 Iowa flood. J. Hydrometeorol. 16, 1171–1183. https://doi.org/10.1175/JHM-D-14-0212.1.

Ochoa-Tocachi, B.F., Buytaert, W., de Bièvre, B., Célleri, R., Crespo, P., Villacís, M., Llerena, C.A., Acosta, L., Villazón, M., Guallpa, M., Gil-Ríos, J., Fuentes, P., Olaya, D., Viñas, P., Rojas, G., Arias, S., 2016. Impacts of land use on the hydrological response of tropical Andean catchments. Hydrol. Process. 30, 4074–4089. https://doi.org/10.1002/HYP.10980.

Orellana-Alvear, J., Célleri, R., Rollenbeck, R., Muñoz, P., Contreras, P., Bendix, J., 2020. Assessment of native radar reflectivity and radar rainfall estimates for discharge forecasting in mountain catchments with a random forest model. Rem. Sens. 12, 1986. https://doi.org/10.3390/rs12121986.

Papacharalampous, G.A., Tyralis, H., 2018. Evaluation of random forests and Prophet for daily streamflow forecasting. Adv. Geosci. 45 https://doi.org/10.5194/adgeo-45-201-2018.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11, 1633–1644. https://doi.org/10.5194/HESS-11-1633-2007.

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115, 1301–1316. https://doi.org/10.1016/j.rse.2011.01.009.

Rollenbeck, R., Bendix, J., 2011. Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. Atmos. Res. 99, 277–289. https://doi.org/10.1016/J.ATMOSRES.2010.10.018.

Sakib, S., Ghebreyesus, D., Sharif, H.O., 2021. Performance evaluation of IMERG GPM products during tropical storm imelda. Atmosphere 12, 687. https://doi.org/10.3390/atmos12060687.

Singh, J., Knapp, H. v, Demissie, M., 2004. Hydrologic Modeling of the Iroquois River Watershed Using HSPF and SWAT. ISWS CR 2004-08. Illinois State Water Survey, Champaign, Ill. https://doi.org/10.1111/j.1752-1688.2005.tb03740.x.

Solomatine, D., See, L.M., Abrahart, R.J., 2009. Data-driven modelling: concepts, approaches and experiences. Pract. Hydroinformat. 17–30.

Solomatine, D.P., Dulal, K.N., 2003. Model trees as an alternative to neural networks in rainfall—runoff modelling. Hydrol. Sci. J. 48, 399–411. https://doi.org/10.1623/hysj.48.3.399.45291.

Solomatine, D.P., Siek, M.B., 2006. Modular learning models in forecasting natural phenomena. Neural Network. 19, 215–224. https://doi.org/10.1016/j.neunet.2006.01.008.

Sorooshian, S., Nguyen, P., Sellars, S., Braithwaite, D., AghaKouchak, A., Hsu, K., 2014. Satellite-based remote sensing estimation of precipitation for early warning systems. Extrem. Nat. Hazards. Disasters. Risks. Soc. Implications. 1, 99.

Tang, G., Long, D., Hong, Y., 2016. Systematic anomalies over inland water bodies of High Mountain Asia in TRMM precipitation estimates: No longer a problem for the GPM era? Geosci. Rem. Sens. Lett. IEEE 13 (12), 1762–1766. https://doi.org/10.1109/LGRS.2016.2606769. Dec. 2016.

Tongal, H., Booij, M.J., 2018a. Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. J. Hydrol. (Amst.) 564, 266–282. https://doi.org/10.1016/j.jhydrol.2018.07.004.

Tyralis, H., Papacharalampous, G., Langousis, A., 2019. A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. Water, Switzerland. https://doi.org/10.3390/w11050910.

Vogels, M.F.A., de Jong, S.M., Sterk, G., Wanders, N., Bierkens, M.F.P., Addink, E.A., 2020. An object-based image analysis approach to assess irrigation-water consumption from MODIS products in Ethiopia. Int. J. Appl. Earth Obs. Geoinf. 88, 102067 https://doi.org/10.1016/j.jag.2020.102067.

Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., Bai, X., 2015. Flood hazard risk assessment model based on random forest. J. Hydrol. (Amst.) 527, 1130–1141. https://doi.org/10.1016/j.jhydrol.2015.06.008.

Willems, P., 2009. A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. Environ. Model. Software 24, 311–321. https://doi.org/10.1016/j.envsoft.2008.09.005.

Xu, L., Chen, N., Moradkhani, H., Zhang, X., Hu, C., 2020. Improving global monthly and daily precipitation estimation by fusing gauge observations, remote sensing, and reanalysis data sets. Water Resour. Res. 56, e2019WR026444 https://doi.org/10.1029/2019WR026444.

Young, C.-C., Liu, W.-C., Wu, M.-C., 2017. A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. Appl. Soft Comput. 53, 205–216. https://doi.org/10.1016/j.asoc.2016.12.052.

Young, P.C., 2002. Advances in real–time flood forecasting. Philos. Trans. R. Soc. London, Ser. A: Math. Phys. Eng. Sci. 360, 1433–1450. https://doi.org/10.1098/rsta.2002.1008.