Quantifying representative runoff coefficients for small basins: a practical recipe for data-limited conditions

Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: www.tandfonline.com/journals/thsj20

Quantifying representative runoff coefficients for small basins: a practical recipe for data-limited conditions

Giulia Evangelista, Pietro Bogoni, Daniele Ganora & Pierluigi Claps

To cite this article: Giulia Evangelista, Pietro Bogoni, Daniele Ganora & Pierluigi Claps (26 May 2026): Quantifying representative runoff coefficients for small basins: a practical recipe for data-limited conditions, Hydrological Sciences Journal, DOI: 10.1080/02626667.2026.2665734

To link to this article: https://doi.org/10.1080/02626667.2026.2665734

Published online: 26 May 2026.

HYDROLOGICAL SCIENCES JOURNAL
https://doi.org/10.1080/02626667.2026.2665734

Quantifying representative runoff coefficients for small basins: a practical recipe for data-limited conditions

Giulia Evangelista, Pietro Bogoni, Daniele Ganora and Pierluigi Claps

Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Torino, Italy

ABSTRACT
The runoff coefficient is a key hydrological variable, but its dependence on event severity and high variability across events makes selecting a representative value for a catchment challenging. This study presents a practical approach to estimate it when high-resolution time series of rainfall and streamflow data are unavailable. Based on over 1000 historical events recorded in the Hydrological Yearbooks, we estimate annual runoff coefficients empirically for 60 small- to medium-sized catchments in Italy. We develop a novel procedure to unambiguously match maximum discharges with corresponding rainfall depths, addressing situations where the occurrence dates of hydrological extremes were not consistently recorded. We also show that pairing rainfall and discharge maxima by frequency provides equally reasonable estimates of runoff coefficients as pairing by temporal coincidence. A consistency analysis and examination of cases with outlying runoff coefficients finally provide useful insights for basin classification to support, for instance, regional flood frequency analysis.

ARTICLE HISTORY
Received 1 December 2025
Accepted 10 April 2026

EDITOR
S. Archfield

ASSOCIATE EDITOR
M. Grodzka-Åukaszewska

KEYWORDS
runoff coefficient; runoff generation; rainfall-runoff

1 Introduction

Italy is characterized by significant climatic variability, with extreme rainfall events that are often highly localized in both time and space. The vulnerability of settlements to such events may be further exacerbated by the country's complex morphology. An example is the Apennine ridge, where longer watercourses span about 100 km before reaching the sea, with many small, steep catchments facing a high risk of flash floods and debris flows. Notably, recent severe flood events in Italy have predominantly occurred in small watersheds, with contributing areas of less than a few hundred square kilometres (e.g. Fortelli *et al.* 2019, Francipane *et al.* 2021).

Although small catchment hydrology is highly representative of the Italian landscape, understanding the hydrological behaviour of such small watersheds remains an ongoing challenge (e.g. Ross *et al.* 2020, Grimaldi *et al.* 2021). A critical aspect in this regard is the analysis of soil absorption and moisture conditions, which can significantly modulate flood severity, particularly in small catchments (Blöschl 2022). A striking example in Italy is the catastrophic flood that affected the Emilia Romagna and Marche regions in May 2023, where pre-existing high soil saturation levels greatly exacerbated flood peaks, as the soil's capacity to absorb additional rainfall was severely reduced (see, e.g. https://www.cimafoundation.org/en/news/the-italian-floods-of-may-2023-a-scientific-analysis/, last access: 26 October 2025).

A parsimonious way to conceptualize soil absorption is the use of the runoff coefficient (C), defined as the fraction of rainfall that becomes direct runoff during a flood event. Although the runoff coefficient is conceptually straightforward and can serve as a lumped metric for representing runoff production at the catchment scale (Viglione *et al.* 2009), its practical use can be affected by some issues.

When rainfall and discharge records are available, event runoff coefficients are typically estimated through base flow hydrograph separation and event selection procedures (e.g. Merz and Blöschl 2009, Tarasova *et al.* 2018a, Wasko and Guo 2022). These techniques require rainfall and streamflow time series at sub-daily temporal resolution, which may not be consistently available on a large scale. In Italy, limited access to continuous hydrological records has resulted in only a few studies addressing the estimation of empirical event runoff coefficients at regional scales (e.g. Norbiato *et al.* 2009, Del Giudice *et al.* 2014, Massari *et al.* 2023, Rahi *et al.* 2023), while national-scale assessments have been attempted in other countries (e.g. Merz *et al.* 2006, Wasko and Guo 2022, Zheng *et al.* 2023).

For design purposes, C is typically treated as a constant, ranging between 0 and 1, with the upper limit indicating highly impervious surfaces. The challenge, however, lies in determining which value is most suitable for a specific watershed. Engineers often rely on recommended basin-scale C values derived from reference tables, which provide general guidance based on factors such as soil permeability, land cover and topography. However, it is acknowledged that these standardized values may not accurately represent real-world conditions (Machado *et al.* 2022) as, for any given basin, the runoff coefficient is not a fixed value; rather, it is strongly controlled by variables like rainfall intensity and soil moisture condition.

CONTACT Giulia Evangelista giulia.evangelista@polito.it Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca Degli Abruzzi 24, Torino 10129, Italy

In this context, our research proposes a data-driven approach for building reasonable estimates of a catchment-representative runoff coefficient, serving as an alternative to the conventional use of continuous time series in this process. To compare occasionally derived values at the event scale with those used in the design context, we carry out historical large-scale spatial reconstructions of the empirical annual runoff coefficient. While event runoff coefficient characterize specific runoff mechanisms, the one computed on an annual basis aggregates catchment processes over a long-term scale (Saft *et al.* 2015). Therefore, the annual C can be seen as a metric of the catchment's overall rainfall-runoff response and can be used for comparing and classifying watersheds. This would be useful to understand the flood generation mechanisms in basins with different morphological features (Blume *et al.* 2007, Ley *et al.* 2011).

To address this goal, in this work we build a dataset of more than one thousand historical flood runoff coefficients, by retrieving data on hydrological extremes from the Hydrological Yearbooks for 60 selected small to medium-sized watersheds across Italy. Unlike most previous studies that rely on continuous time series, our work explicitly derives runoff coefficient values from rainfall and discharge extremes, an approach rarely reported in the literature, yet more practical, given that data on hydrological extremes are often easier to obtain than continuous records. In our investigation, we evaluate and compare two estimation methods for the runoff coefficient. The first approach is event-based, whereby individual daily maxima of discharge are temporally matched with corresponding daily rainfall depths; the second is frequency-based and involves the statistical matching of rainfall and flood frequency. To enhance the accuracy of the event-matching approach, we develop a dedicated procedure designed to effectively associate in time each annual maximum flood to the value of the rainfall depth that most likely produced it. The standardized procedure that we propose here can help to address a significant data gap in Italy, and potentially elsewhere, i.e. the lack of readily available information on the occurrence date (year/month/day) of hydrological extremes.

We also perform a detailed analysis of case studies that exhibit anomalous values of the reconstructed runoff coefficient, which are either unusually small (below 0.1) or exceed 1. By identifying potential sources of estimation errors, whether stemming from inaccuracies in historical data transcription, methodological choices, or specific morphological and climatic basin properties, researchers can significantly improve the accuracy of their runoff coefficient estimates.

This work serves as a preliminary foundation for subsequent, more detailed analyses, which will make use of continuous time series data from recent years. Such future investigations will allow for a more in-depth understanding of temporal variability of the runoff coefficient, especially regarding its relationship with antecedent rainfall.

2 Methods

The empirical annual runoff coefficient, hereinafter referred to as C, is estimated here as a volumetric ratio:

$$C = \frac{R}{P} \tag{1}$$

where R and P are the total basin-averaged runoff depth (mm) and the total areal rainfall depth (mm), respectively.

According to the scheme discussed in Lapides *et al.* (2021), a first, deterministic, interpretation of C can be given by assuming an event-based matching of R and P. The runoff coefficient estimated in this way is associated with physical processes, as each flood is associated with the rainfall depth that produced it. This interpretation implies that, depending on soil conditions and, consequently, soil infiltration capacity, R and P may be associated with different return periods. In the following sections, the runoff coefficient estimated as described above will be denoted as C_{event}.

A second interpretation is possible by adopting a frequency-based matching of R and P (e.g. Schaake *et al.* 1967, Hawkins 1993, Young *et al.* 2009, Dhakal *et al.* 2012). In this way, a statistical runoff coefficient (hereinafter $C_{frequency}$) is defined as an adjustment coefficient that ensures the same exceedance probability between R and P, where R is the runoff with the same return period of P. This assumption is implicit in the practical use of the rational method, where the runoff coefficient is used as a calibration parameter to force the predicted design flood to have the same return period as the design rainfall.

The following sections present the two estimation methods and discuss their strengths and limitations in relation to data availability.

2.1 Event-based approach

2.1.1 Development of an unambiguous event-based matching of discharges and rainfall depths

To compute C_{event} using Equation (1), one needs to estimate R and P for an unambiguously defined flood event. In this work, R is calculated from annual maxima of daily discharge, sourced from the Italian Flood and Catchment Atlas (Claps *et al.* 2024). Daily maxima are preferred because they are generally more readily available than annual peak discharges and are directly related to the runoff volume of the same day. This direct relationship helps minimize ambiguity when associating runoff volumes with rainfall amounts during a given flood event. Therefore, according to the event-based matching technique, P represents the rainfall depth deemed to have generated the observed value of R.

However, the estimation of C_{event} may suffer from a major drawback. As mentioned in the Introduction, the day and month on which historical extreme flood events occurred are not systematically recorded or readily accessible in Italy, particularly in the case of annual maxima of instantaneous discharge. On the other hand, maxima of daily average discharge are always associated with a date, as they are derived directly from continuous daily runoff time series.

In this work, the date of occurrence (year/month/day) of annual maxima of daily discharge and daily rainfall are manually retrieved from the daily time series in the Hydrological Yearbooks (hereinafter HYs). This process is quite time-consuming, as the two variables are recorded in two separate volumes each year. To retrieve an entire time series of extremes measured by a gauging station, all the yearbooks must be consulted, and they are available only as individual images at http://www.bio.isprambiente.it/annalipdf/ (last access: 23 September 2025).

Besides the time demanding nature of this task, an additional challenge remained unsolved. Even after collecting all the necessary data, associating the maximum discharge with the corresponding rainfall depth proved to be a non-trivial operation. This ambiguity arises due to differences in how rainfall and discharge data are recorded, as also reported by Froidevaux *et al*. (2015) in their rainfall-runoff analysis of basins in Switzerland. According to what stated in the HYs, daily rainfall depths were measured each day at 9 a.m. and referred to the total rainfall amount accumulated over the previous 24 hours. On the other hand, daily discharge measurements were recorded at noon of the same day. Before 1940, daily discharge values corresponded to instantaneous readings at noon; from 1940 onward, they represent the average of measurements taken every 6 hours, or more frequently, every 1 to 2 hours, during periods of significant water level fluctuations. Therefore, when searching for $(R\text{-}P)_{event}$ pairs, two main situations are recognized, as represented in Fig. 1. In the first case (case A, Fig. 1a), the annual maximum daily discharge occurs on the same day as a sufficiently high rainfall depth, at least when compared to the rainfall amounts recorded on other days within the month. In this case, it is reasonable to assume the two can be directly associated. In a second case (case B, Fig. 1b), a plausible value for P is reported in the HYs one day after the occurrence of the annual maximum discharge. This may seem counterintuitive. However, looking at the scheme in Fig. 1 and considering that the times of concentration of the basins analysed are certainly less than one day, due to their small-to medium-sized drainage areas (see section 3), one can conclude that this second case is equally reasonable.

A third condition is also identified, although it is much less common. In this case, annual maximum discharges are recorded one day after a feasible value for P. Given the rainfall recording practice, this would imply that the recorded rainfall actually occurred two days prior to the flood. This introduces a time lag between precipitation and discharge that is not consistent with the typical response times of these small basins, which are generally on the order of a few hours. This third case, therefore, is considered physically unacceptable, and events that fall within this scenario are considered potential instances of data transcription errors. An interesting case in which such an error has been identified is discussed in-depth in section 4.2.1. A procedure to match rainfall and discharges on an event-basis preventing ambiguities is then developed. The process is outlined in the flowchart in Fig. 2.

2.2 Frequency-based approach

The selection of events and related data described in the previous section needs to be structured as a standardized procedure and may be very time-consuming to implement. For this reason, we also consider a frequency-based matching approach for R and P in our analysis. Originally proposed by Schaake *et al*. (1967), this approach has been used in the literature to estimate both the runoff coefficient (e.g. Young *et al*. 2009, Dhakal *et al*. 2012, 2013) and the Curve Number (e.g. Hjelmfelt 1980, Hawkins 1993).

Using this method, *"the rainfall depths and runoff depths are sorted separately, and then realigned on a rank-order basis to*

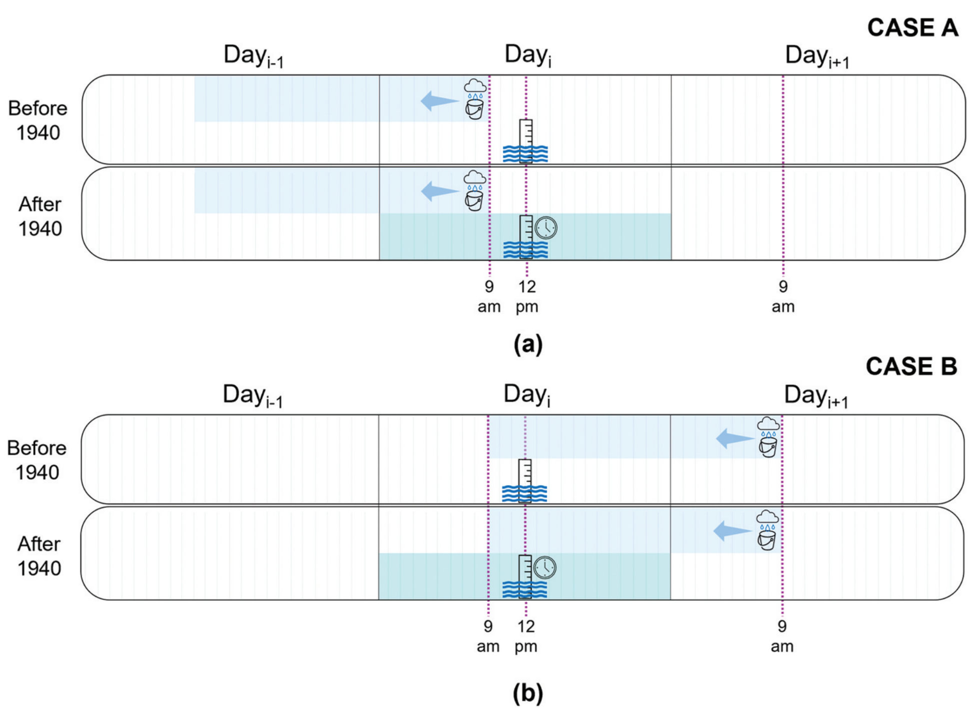

Figure 1. Examples of temporal correspondence between rainfall depths and discharges, as documented in the Hydrological Yearbooks, include: instances where rainfall depth and discharge were recorded on the same day (a), and cases where rainfall depth and discharge were recorded with a one-day delay (b). Additionally, two distinct methodologies for measuring daily discharges are noted, one adopted prior to 1940 and another introduced after 1940.

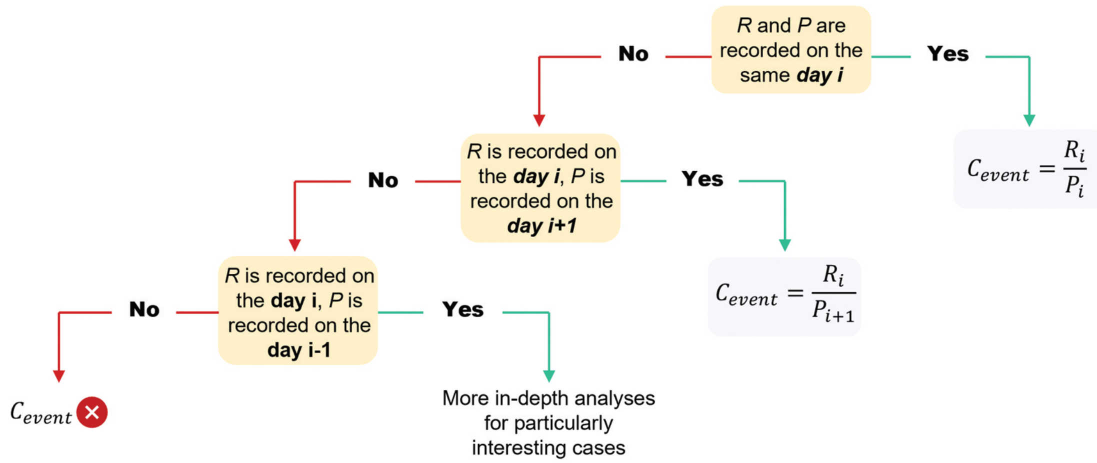

Figure 2. Procedure for establishing the temporal correspondence between runoff (*R*) and rainfall (*P*) depths, and computing the event-based runoff coefficient C_{event}.

form Q-P pairs of equal return periods. The individual runoffs are not necessarily associated with the original causative rainfalls" (Hawkins 1993). By removing the need for event-specific pairing, this approach allows for a simplified, but systematic and rapid estimation of flood runoff coefficients. In fact, the annual maxima of daily discharge (the same used in the previous section) and those of daily rainfall can simply be sorted independently, without reference to specific dates, to form $(R,P)_{frequency}$ pairs, so that $C_{frequency}$ can be computed using Equation (1).

2.3 Estimation of C_{event} and $C_{frequency}$ runoff coefficients

The procedure used to systematically calculate the runoff coefficient *C* is outlined in the steps shown in Fig. 3, where *G*, *R* and *P* stands for "gauge", "runoff depth", and "rainfall depth", respectively. Each step is described below.

Starting point:

For a given catchment and year, we first identify the annual maximum daily discharge from Claps *et al.* (2024) and retrieve the corresponding date of occurrence (year/month/day) from the HYs. Then, we calculate the equivalent runoff depth *R*.

Step 1:

We draw a variable-width buffer around the basin, in order to encompass a minimum of one and a maximum of four active rain gauges within the designated area. We found that including up to four rain gauges strikes a good balance, enhancing the analysis's accuracy without making it overly costly. Any non-active rain gauges for the specified (year/month/day) are excluded from the analysis (e.g. gauge G_i in Fig. 3, step 1).

Step 2:

For the event-based matching procedure (step 2a), once the active rain gauges are selected, we verify that the selected daily rainfall depth is at least a monthly maximum. This check assures us more with the likely temporal match to the extreme discharge. If this check is not met, that rain gauge is excluded from the analysis (e.g. gauge G_k in the figure). For the frequency-based matching procedure (step 2b), we search on the HYs for the date (year/month/day) of each of the annual maximum daily rainfall depth recorded by each rain gauge.

Step 3:

We interpolate local rainfall depths using the weighted inverse distance method to estimate the areal rainfall depth *P*. For the event-based matching procedure (step 3a), we use this areal rainfall *P* to compute the runoff. For the frequency-based matching procedure (step 3b), we calculate a different areal rainfall depth for each day in which each rain gauge records its annual maximum. We use the largest of these areal rainfall values as *P*.

Final step:

We compute the runoff coefficient as the ratio of the total runoff depth and the total rainfall depth.

3 Data and study area

The approach proposed in this research, designed for application over a large area, requires the introduction of some simplifications to ensure consistency and comparability across watersheds. First of all, sufficiently long records of rainfall and discharge are needed to ensure a robust analysis. In addition, elements that may introduce uncertainties in quantifying catchment runoff should be minimized as much as possible. In this regard, possible sources of complexity include snow accumulation and the spatial variability of precipitation within the basin. Furthermore, to ensure feasibility of manually computing of areal rainfall, the analysis is limited to relatively small basins.

To meet these requirements, we introduce the following criteria for basin and record selection: (i) at least 15 years of annual maxima of daily discharge available; (ii) mean elevation lower than 1000 m a.s.l.; (iii) basin area smaller than 100 km^2. Based on these criteria, 45 watersheds across Italy are selected from the Italian Flood and Catchment Atlas (Claps *et al.* 2024).

For comparative purposes, a secondary set of 17 additional watersheds from Claps *et al.* (2024) is selected by relaxing criteria (ii) and (iii). These additional basins are located in northwestern Italy (Piemonte and Valle d'Aosta regions), can cover areas of up to 700 km^2 and may have mean elevations exceeding 1000 meters a.s.l.

The final dataset consists of 62 basins, divided into two distinct groups: Class A, which includes basins meeting the original selection criteria, and Class B, which represents those

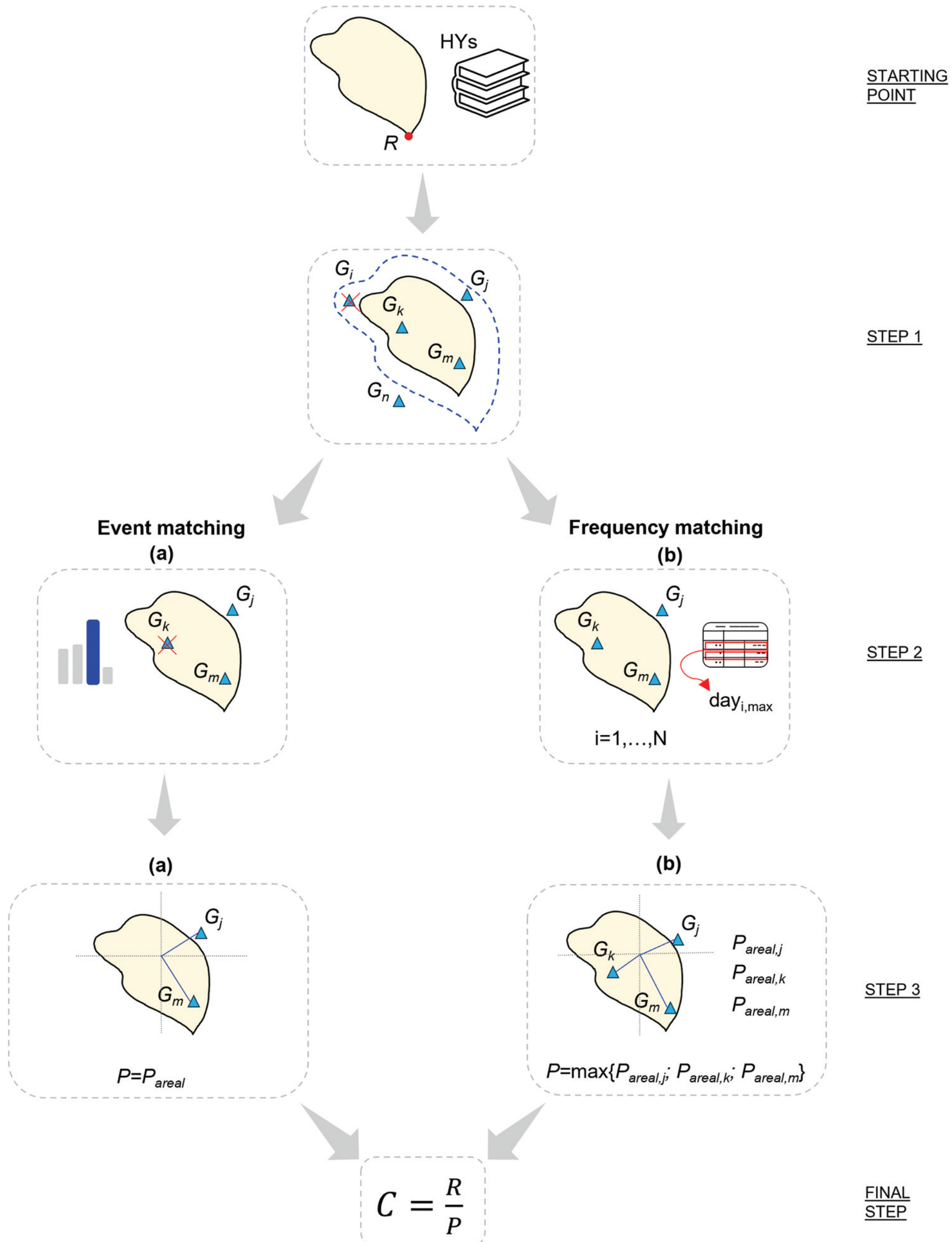

Figure 3. Methodology for estimating runoff coefficients using the event-based (a) and frequency-based (b) approaches.

from the supplementary set. The spatial distribution of these basins is illustrated in Fig. 4a, where the size of each dot corresponds to the basin area. Although the 62 basins do not provide uniform coverage across all Italian regions, they encompass a broad range of climatological diversity throughout Italy, as illustrated in Fig. 4b and 4c, where the variability of catchment-averaged mean annual precipitation (MAP) and mean annual temperature (MAT) is shown.

4 Results and discussion

4.1 Flood runoff coefficients in Italy and robustness of the frequency-based approach

A total of 1113 basin-years are used to reconstruct an equivalent number of flood runoff coefficients for 60 basins. Two of the initial 62 basins were excluded from the dataset due to insufficient data availability, as the final number of years with reconstructed C values, after applying the procedure shown in Fig. 2, was less than 10. The number of available years for each catchment is reported in Fig. 5. The temporal coverage varies across the dataset: the highest number of reconstructed runoff (R) – rainfall (P) pairs for a single catchment is 49, while the median number across all catchments is 17 years. For each catchment-year, two runoff coefficient values are computed: one using the event-based approach (C_{event}) and the other using the frequency-based approach ($C_{frequency}$). In the event-based distribution, approximately 80% of the reconstructed C values, computed following the methodology outlined in Fig. 3, correspond to case A as defined in Fig. 1, indicating that the maximum daily discharge and the associated rainfall depth are recorded on the same calendar day. The remaining 20% fall under Case B.

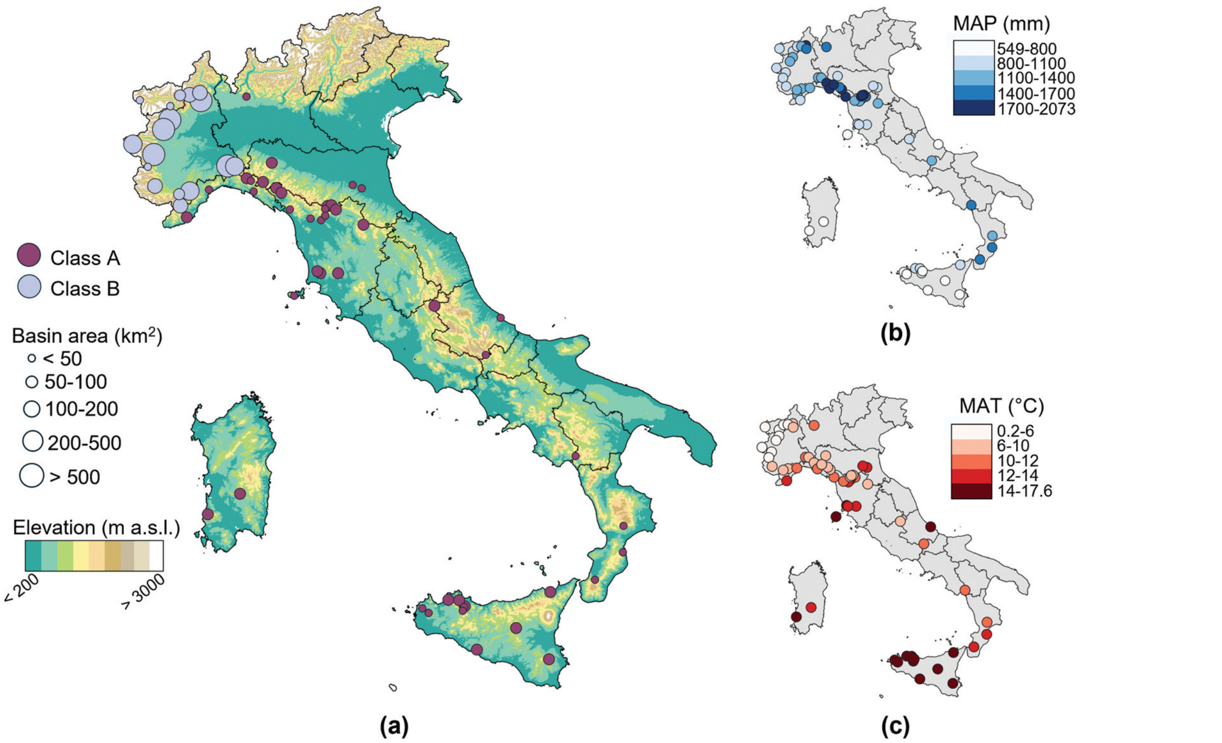

Figure 4. Study area and outlets of basins investigated (a). Catchment-averaged values of mean annual precipitation (MAP, b) and mean annual temperature (MAT, c). MAP and MAT values are sourced from the dataset of Claps *et al.* (2024).

To facilitate the classification of flood runoff coefficients across Italy, catchments were grouped according to the Köppen climate zones (Beck *et al.* 2023). Figure 6 illustrates, in a log-log graph, the ratio of each event's runoff depth (R_i) to that derived from the average of the annual maxima of daily discharge (R_{avg}), together with the corresponding ratio for rainfall (P_i/P_{avg}). The size of each point in the plot represents the reconstructed C_{event} for each basin and event.

Flood runoff coefficients exhibit substantial variability across Italy, spanning the full range from approximately 0 to 1, with generally higher values in cold regions (Df and E) than in temperate ones (Cf and Cs). About 8% of cases exhibit C_{event} values exceeding 1.0, meaning that the outflow volume exceeds the recorded inflow, whereas a smaller subset (almost 3%) shows values below 0.1. In general, coefficients greater than 1 tend to occur when runoff depths are at least comparable to R_{avg}, although this pattern is not consistent across all climate zones. In catchments located within cold regions (Fig. 6c), high and very high runoff coefficients are observed under conditions of relatively low runoff depth and rainfall compared to their respective mean values. This anomalous behaviour is particularly evident in glacial basins, where runoff appears to depend weakly on rainfall inputs. Such patterns likely reflect the influence of snow dynamics and storage–release processes that are not directly related to rainfall amounts.

We acknowledge that some C_{event} values greater than 1 may result from data inconsistencies or measurement uncertainties, as also reported by Beven and Smith (2014). Further discussion of this topic is given in section 4.2 for specific catchments.

Comparisons are then made between the two sets of runoff coefficients reconstructed using the event-based and frequency-based approaches. To explore representative values at the catchment scale, we analyse the 10th percentile, the median and the 90th percentile of the runoff coefficients that caused the maximum daily annual floods in each catchment, estimated using both methods. Results are presented in Fig. 7. Data points are color-coded according to the Köppen climate zone, and the axis ranges vary across panels to better highlight the variability within each plot.

At the 10th percentile (Fig. 7a), the two methods yield closely overlapping results. The coefficient of determination ($R^2 = 0.84$) indicates strong agreement, with all runoff coefficients remaining below 0.75. Within this percentile, catchments located in temperate regions with dry summers (Cs) generally exhibit lower coefficients than those in temperate regions without a dry season (Cf), whereas basins within cold and wet regions (Df and E) tend to cluster around values between 0.2 and 0.3.

For the median and 90th percentiles (Fig. 7b and 7c), C values span nearly the full possible range, from approximately 0.0 to 1.0. A notable number of 90th percentile values exceed 1.0. For both percentiles, $C_{frequency}$ values are consistently lower than the corresponding C_{event} values. This difference is consistent with the concept behind the frequency-based approach, where runoff and rainfall depths are paired by rank rather than by temporal coincidence, effectively reducing atypical values. At the median (Fig. 7b), data points representing cold-climate catchments (Df and E) display a more linear distribution than at the 10th percentile, suggesting a more consistent relationship between the event-based and frequency-based approaches. The overall R^2 is 0.73, indicating a relatively high level of consistency between the two methods. This outcome is particularly encouraging given the considerable effort required to estimate C_{event}s. If representative statistics, such as the median, are approximately consistent between

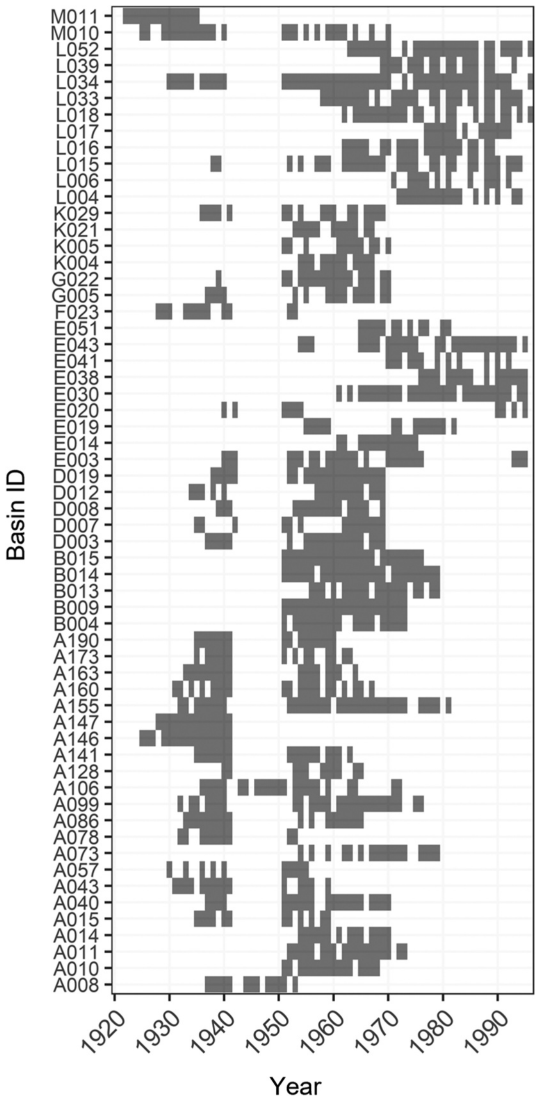

Figure 5. Years with reconstructed runoff coefficients (after quality checks) for 60 catchments across Italy. Basin IDs are those reported in Claps *et al.* (2024).

the two distributions, the frequency-matching approach could provide a more efficient and practical alternative for deriving a basin-representative runoff coefficient, reducing computational demands without compromising accuracy. This method is particularly advantageous because annual maximum discharges are often recorded without dates and this does not affect the computation of $C_{frequency}$. An exception to this general consistency is observed in one catchment (Rutor at Promise, the outlying gray point in Fig. 7b). Results for this Alpine basin, which has an average elevation above 2000 m a.s.l., will be discussed in detail in section 4.2.3.

At the 90th percentile (Fig. 7c), the agreement between the two methods weakens, as indicated by a lower R^2. Deviations from the 1:1 line seem to be particularly pronounced for catchments in cold climates (Df and E), compared with catchments in temperate zones (Cf and Cs).

Table 1 presents the R^2 values describing the relationship between runoff coefficients estimated using the event-based and frequency-based approaches, with results differentiated by climate zone and by percentile of the runoff coefficient distributions. Catchments classified within the glacial climate (E) zone are not included, as only three were identified. The largest differences occur between temperate regions with dry summers (Cs) and cold climates (Df), particularly at the 90th percentile, where the two approaches exhibit stronger agreement for Cs catchments than for those located in the Df zone.

4.2 Addressing anomalous cases to enhance the accuracy of estimates

In this section, we present an overview of cases in which outliers were detected, i.e. very small reconstructed runoff coefficients (lower than 0.1) or values exceeding 1. Previous studies (e.g. Young *et al.* 2009, Dhakal *et al.* 2013) have reported *C* values greater than 1.0, though offering rather general explanations, mostly related to the computational methodology used to estimate C, rainfall characteristics, measurement errors of rainfall and runoff data, or unspecified "*unusual hydrologic factors*" (Dhakal *et al.* 2013). By conducting an in-depth investigation of these anomalous cases, this section aims to provide some elements to exemplify the role of specific factors in producing anomalous runoff coefficient estimates.

4.2.1 Case 1: manual transcription errors in discharge records

This section provides a detailed analysis of the 1951 flood event in the Erro at Sassello river basin, located within the Bormida catchment, in northwestern Italy (Fig. 8a). The historical maximum daily discharge at the Erro at Sassello gauge was recorded on 21 November 1951, at 81.1 m^3/s, representing the highest value in the observed hydrological series (Claps *et al.* 2024). An analysis of precipitation data for the eleven surrounding active rain gauges reveals that the highest rainfall amount occurred on November 20, the day preceding the recorded discharge. It is important to remember that the precipitation value for November 20 represents the cumulative rainfall (expressed in mm) measured from 09:00 on November 19 to 09:00 on November 20. On the other hand, the reported average discharge corresponds to measurements taken every six hours on November 21.

This temporal misalignment reveals a lag of more than 24 hours between precipitation and discharge. Given the relatively small size of the basin (approximately 90 km^2), such a delay seems to be hydrologically inconsistent. However, a systematic shift in the timing of discharge measurements at the Erro at Sassello gauge is evident not only for the 21 November 1951, event but also for the November 8–12 event, when compared to discharges observed at neighbouring gauges (Fig. 8b). Therefore, one possible explanation for this is a transcription error that may have occurred during the manual transfer of data from the station's hydrometrograph records. The operator while copying the recorded discharge values could have introduced this.

4.2.2 Case 2: influence of rain gauge distribution

The Bormida di Mallare catchment is located in the Liguria region, in northwestern Italy, as shown by the small map of Italy in Fig. 8. The following analysis addresses an anomalous runoff coefficient C_{event} reconstructed in this catchment using

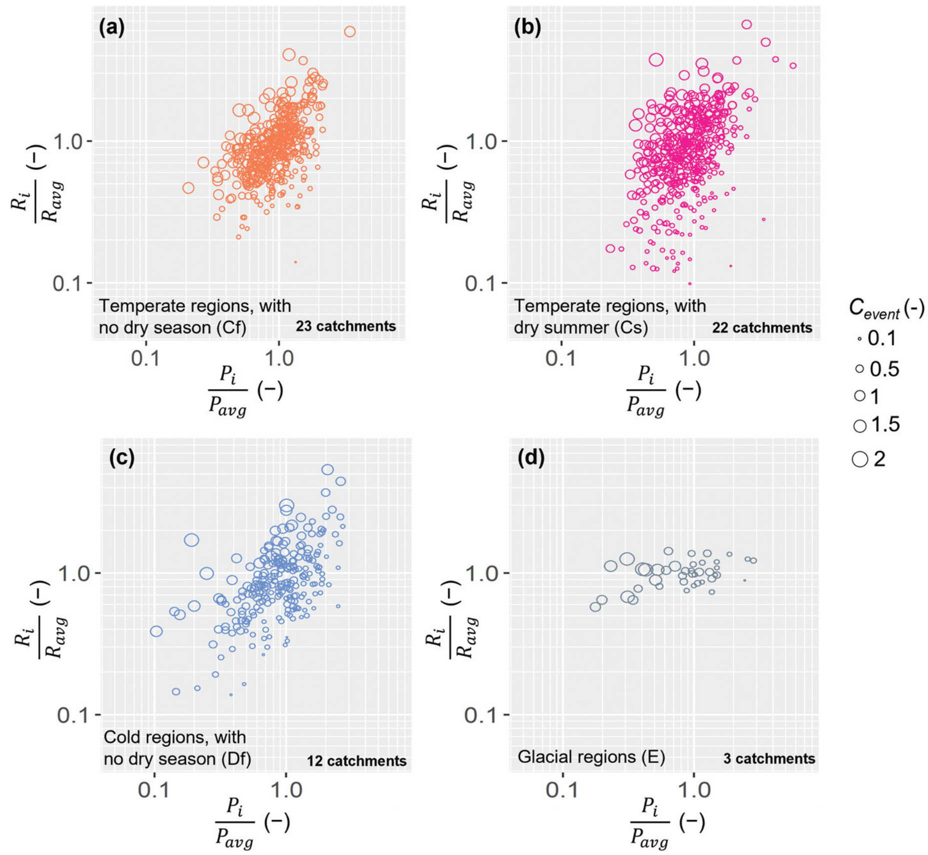

Figure 6. Relationship between normalized rainfall (P_i/P_{avg}) and normalized runoff (R_i/R_{avg}) depths for catchments under different climatic conditions: temperate climate with no dry season (a), temperate climate with dry summer (b), cold climate with no dry season (c), and glacial climate (d). Each point represents an event, and the point size indicates the event coefficient C_{event}.

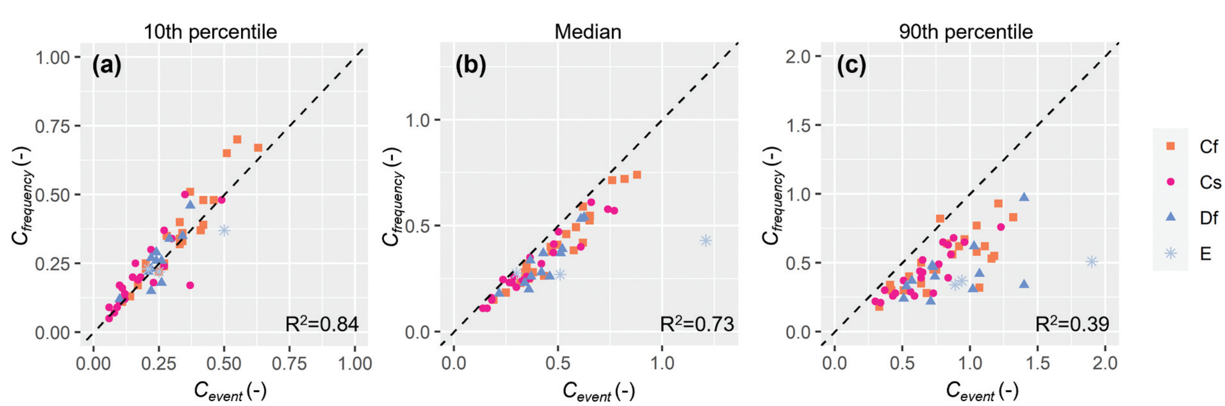

Figure 7. Scatterplots comparing runoff coefficient values reconstructed using the event based approach (C_{event}) and the frequency based approach ($C_{frequency}$): 10^{th} percentiles (a), medians (b) and 90^{th} percentiles (c) of the two distributions. Values are grouped by Köppen climate zone: temperate with no dry season (Cf), temperate with dry summer (Cs), cold with no dry season (Df), glacial (E).

Table 1. Values of the coefficient of determination (R^2) computed between event-based and frequency-based runoff coefficients across different climatic zones and percentiles of the corresponding distributions. Higher values of R^2 indicate stronger linear correlation between the two considered variables.

Climate zone	10^{th} percentile	Median	90^{th} percentile
Temperate with no dry season (Cf)	0.91	0.93	0.53
Temperate with dry summer (Cs)	0.7	0.92	0.71
Cold with no dry season (Df)	0.75	0.82	0.32

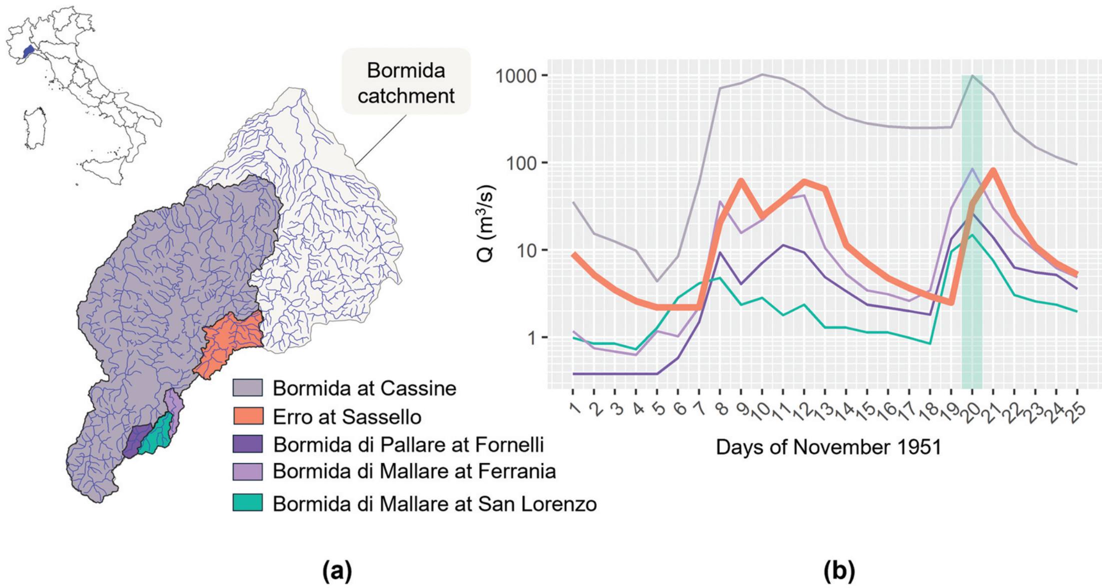

Figure 8. Case study of the 1951 flood event in the Erro at Sassello river basin. The Erro basin and its neighboring catchments (a). Recorded flood hydrographs for the event (b).

Equation (1) for the year 1937, whose value exceeds 3. This unusually high coefficient likely results from an overestimation of the runoff depth or an underestimation of the precipitation input. According to the Hydrological Yearbooks of Parma (1937, Part II), a daily basin-averaged runoff depth of 403 mm, recorded at noon, is attributed to November 1st. Rainfall data from the five active rain gauges within the region are examined for the period spanning October 30 to November 2. Their locations are illustrated in Fig. 9, and the corresponding data are presented in Table 2.

Among them, the Rialto gauge, located approximately 8 km from the catchment centroid, recorded the highest daily rainfall depth (73.3 mm) on November 1st. Alternatively, using data from the two closest rain gauges, i.e. Sella (6.5 km) and Montagna (about 6 km), which both recorded peak rainfall on November 2nd, the interpolated catchment-averaged rainfall depth is 73.4 mm. In both cases, the calculated runoff coefficient remains anomalously high. Hydrometric data from downstream sections of the Bormida river are unavailable for validation or comparative analysis. This anomalous value of the runoff coefficient is most plausibly attributable to underestimated rainfall: as shown in Fig. 9, all active rain gauges during this period were located on the seaward side (right) of the topographic watershed, potentially leading to a substantial underestimation of the spatial rainfall distribution within the basin.

Beven and Smith (2014) describe situations such as that outlined above as instances of "hydrologically inconsistent" data, which contribute to "epistemic sources of uncertainty" in hydrological modelling. Such a discussion can therefore support the identification of these cases and inform strategies for addressing them.

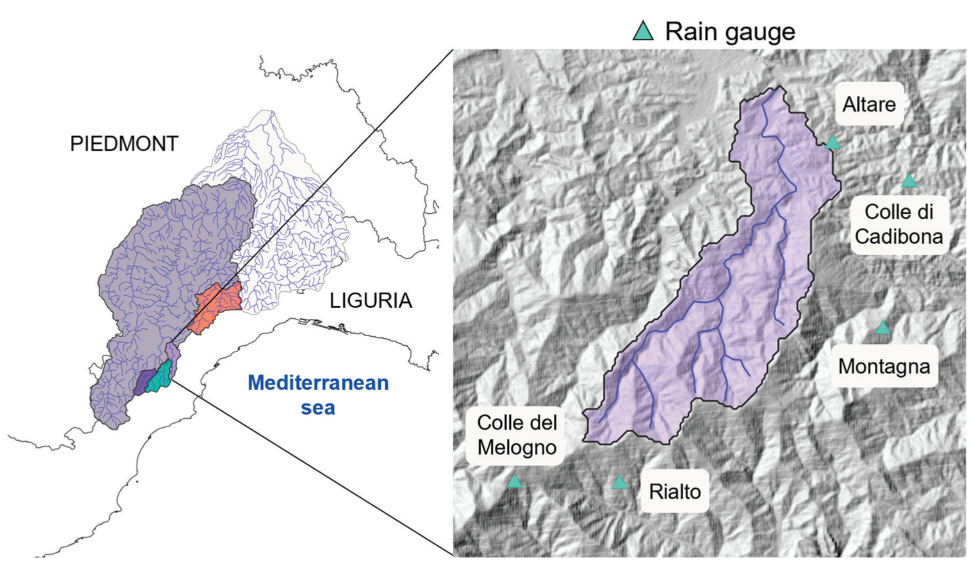

Figure 9. Bormida di Mallare at Ferrania catchment. Distribution of active rain gauges in 1937.

Table 2. Daily rainfall depths (mm) recorded at five rain gauges located near the Bormida di Mallare at Ferrania catchment between October 30 and 2 November 1937. Distances are measured from the centroid of the catchment.

Rain gauge	Distance from basin centroid (m)	Rainfall depth measured on October, 30 (mm)	Rainfall depth measured on October, 31 (mm)	Rainfall depth measured on November, 1 (mm)	Rainfall depth measured on November, 2 (mm)
Sella	6500	0	41.1	26.6	110.1
Colle di Cadibona	7600	24	14	1	32.1
Rialto	8270	5	4.7	73.3	20
Colle Melogno	10700	44	0.2	12.4	73
Montagna	5900	22.2	9.6	2.4	44.2

4.2.3 Case 3: impact of snow melting in Alpine basins

This section presents a detailed analysis of six Alpine basins from the dataset, all having a mean elevation exceeding 2000 m a.s.l. Their main characteristics are summarized in Table 3. All selected basins fall within Class B of Fig. 4 and are located in the Alpine regions of Piemonte and Valle d'Aosta.

To investigate the mechanisms that, in specific events, produced runoff coefficients exceeding 1, the seasonality of annual maxima of daily discharge was analysed. Figure 10a illustrates that nearly 70% of annual maxima of daily discharge occurred between May and July, a period likely influenced by snowmelt processes. Reconstructed runoff coefficients during this period tend to be higher, often exceeding 1. The

Table 3. Mean features of six Alpine catchments from the dataset.

Basin ID	Basin name	Area (km^2)	Mean elevation (m a.s.l.)	Catchment averaged mean annual precipitation (mm)	Number of events
A011	Ayasse at Champorcher	43	2368	1249	19
A057	Dora Riparia at Oulx	260	2164	793	10
A078	Lys at Gressonay Saint Jean	91	2645	970	10
A106	Po at Crissolo	38	2251	1007	23
A128	Rutor at Promise	50	2525	1127	11
A146	Sesia at Campertogno	171	2127	1203	16

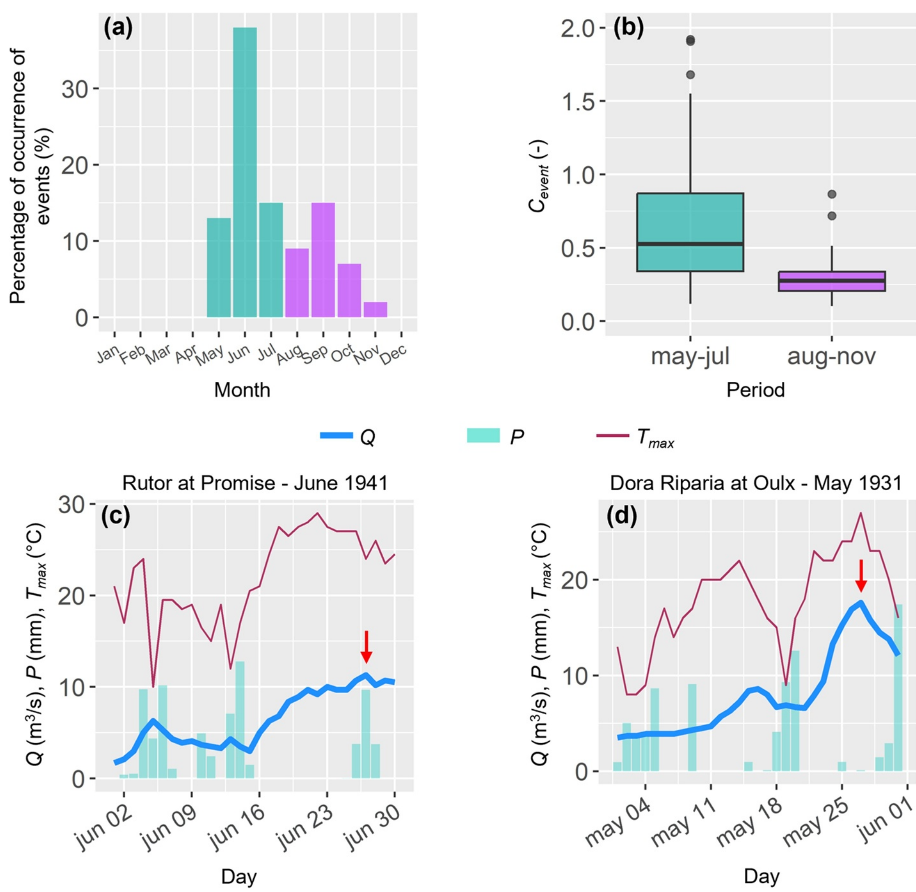

Figure 10. Temporal distribution of annual maxima of daily discharge for the analysed Alpine basins (a). Boxplots of reconstructed runoff coefficients for the same basins, grouped by seasons: May-July and August-November (b). Rutor at Promise, 1941 flood (c). Dora Riparia at Oulx, 1931 flood (d). The light blue bars are the measured daily rainfall depths (*P*), the blue lines represent the observed daily streamflow (*Q*), while the red lines indicate the recorded daily maximum air temperature (T_{max}).

occurrence of snowmelt can be considered a reasonable source of imbalance between rainfall and runoff. Moreover, runoff coefficients of summer events display greater inter-annual variability than those observed during the autumn (Fig. 10b). For the May–July period, the average runoff coefficient is 0.71, with a maximum of 1.92 and a minimum of 0.12. In the August–November period, the average drops to 0.32, with a maximum of 0.87 and a minimum of 0.10. The absence of runoff coefficient values exceeding 1 during the August–November period can be interpreted as supporting evidence for the above assumption.

With reference to mountain basins, Fig. 10c and 10d present two specific case studies: the flood at Rutor at Promise, in 1941 and the flood at Dora Riparia at Oulx, in 1931. Both events occurred during the early summer months of May and June. The graphs illustrate the discharges recorded at the basin outlets during the selected months, along with catchment-averaged rainfall depths and maximum daily temperatures from the nearest weather stations, i.e. Courmayeur for the Rutor basin, and Bardonecchia for the Dora Riparia basin. In the case of the Rutor basin (Fig. 10c), the maximum discharge coincides in time with a rainfall event. However, the gradual rise in temperature during this period suggests that rainfall contributed only marginally to the observed flow, with snowmelt likely representing the dominant source. Under these conditions, the runoff coefficient has approached a value close to 2. In the case of the Dora Riparia catchment (Fig. 10d), the maximum discharge clearly follows the increasing temperature pattern and is not associated with any recorded precipitation. As no rainfall is observed during this period, it is not possible to determine a runoff coefficient value for this event.

5 Discussions

Observed runoff coefficients vary from event to event. According to the Handbook of Hydrology (Maidment 1993; Chapter 9), "*often, the value of C is assumed to increase as average recurrence interval increases, thus allowing for nonlinearity of flood response*". Indeed, higher return periods, which correspond to more intense rainfall events, are associated with greater runoff volumes, due to the finite capacity of the soil to absorb rainfall excess.

As explained in section 2.2, the $C_{frequency}$ values are such that the discharge return period matches the rainfall return period, which is typically assumed in the design procedure. Schaake *et al.* (1967) found this assumption to be approximately correct when compared with observed data, particularly in small catchments. Simulation experiments by Viglione *et al.* (2009) demonstrated that this relationship is influenced by the average soil wetness and may no longer hold in dry systems.

In this study, we do not have sufficient evidence to assess the general validity of this assumption. We also observe that a median runoff coefficient of a distribution may not represent a conservative estimate for design applications, as we are neglecting more extreme values (as shown in Fig. 7). However, since this study does not aim to establish definitive runoff coefficients for design applications, our approach should be seen as a preliminary step toward developing a methodology that can provide reasonable estimates, with a range of variability, in contexts where high-resolution hydrometeorological data are unavailable.

The frequency-matching approach relies on pairing ranked rainfall and discharge values. Therefore, the consistency between the runoff coefficients obtained through this method and those derived from an event-based analysis suggests that the precipitation–streamflow relationship remains largely governed by rainfall amounts, with greater precipitation generally producing higher streamflow. The two methods show stronger agreement in temperate, relatively dry, catchments than in cold and wet ones, with the greatest differences occurring at the 90th percentile of runoff coefficients associated with annual maxima of daily discharge (see Table 1). This finding aligns with the observation that very high runoff coefficients are relatively rare in dry catchments, and their variability is constrained by the limited potential for soil saturation, as also noted by Viglione *et al.* (2009). Consequently, the frequency-matching approach, by smoothing high and very high values through the ranked pairing of rainfall and discharge, yields results that more closely correspond to those obtained from the event-based method.

Once the representativeness of the median $C_{frequency}$ has been assessed, it may be useful to examine its relationship with simple catchment characteristics. Figure 11 shows median values of $C_{frequency}$ as a function of two key basin descriptors, i.e. mean elevation and mean annual precipitation, classified by climate zone. These variables are among the most commonly used spatial controls on runoff coefficients (e.g. Merz and Blöschl 2009, Tarasova *et al.* 2018b, Jahanshahi and Booij 2023, Zheng *et al.* 2023). When all Italian catchments are considered together, no clear pattern emerges (Fig. 11a and 11e). However, grouping catchments by climate zone reveals distinct relationships, particularly for basins in wet regions, both temperate (Cf, Fig. 11b and 11f) and cold (Df, Fig. 11d and 11h). In temperate, wet areas, an increase in the mean catchment elevation generally corresponds to higher precipitation and likely more saturated soils, resulting in higher runoff coefficients. In contrast, in cold, wet areas, increasing elevation corresponds to a greater influence of snow on runoff. Patterns are less pronounced for catchments in drier areas (Cs, Fig. 11c and 11g), consistent with the general understanding that antecedent moisture conditions play a dominant role in such environments, leading to less predictable runoff responses. It is worth noting that nearly 90% of catchments classified as Df, as well as all catchments classified as E, are part of the group identified as Class B in Fig. 4. Without including this additional set of catchments for comparison, the assessment of differences in runoff coefficients across climate zones would have been incomplete.

Median values of runoff coefficients estimated using the frequency-matching approach are also reported by Dhakal *et al.* (2013), who compared them with both event-based runoff coefficients and tabulated values reported in the literature. However, their study does not provide indications on the catchment types or hydrological conditions under which the two methods yield comparable results, leaving

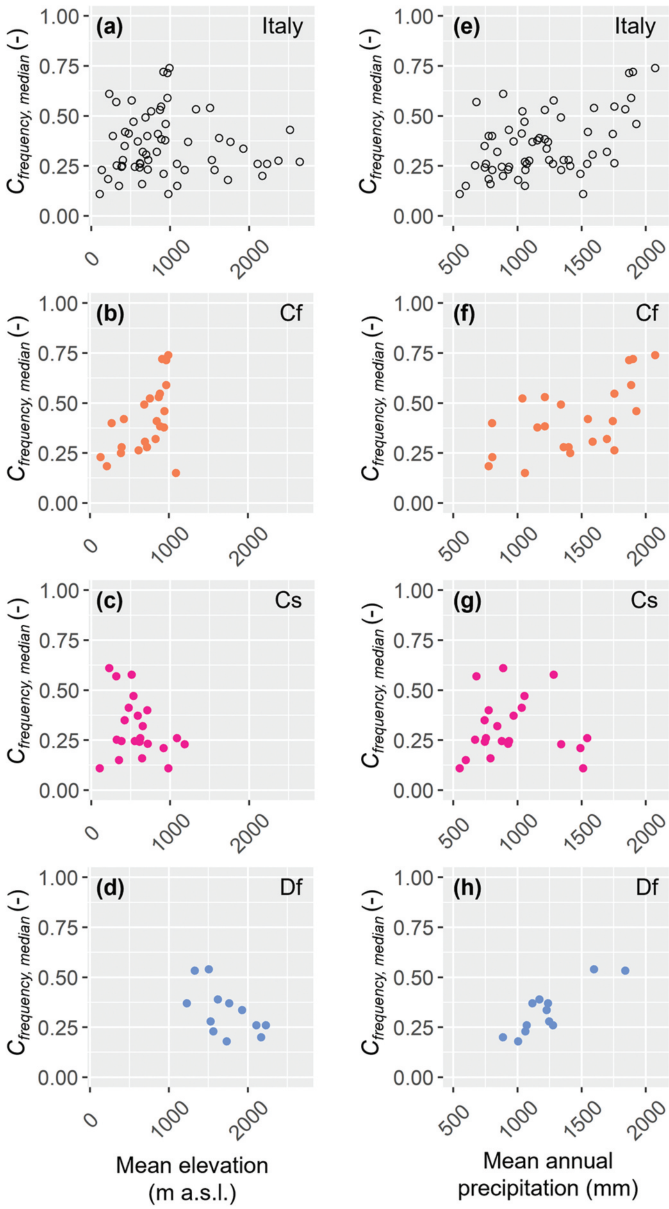

Figure 11. Scatterplots showing the relationship between the median values $C_{frequency}$ and catchment properties, grouped by climate zone: mean catchment elevation (a–d) and catchment-averaged mean annual precipitation (e–h).

some uncertainty regarding the applicability of their conclusions across different climatic and geomorphological contexts.

It is also important to emphasize that the observations presented above should not be interpreted as a conventional or exhaustive analysis of the spatial controls on runoff coefficient variability. Such an analysis would likely require a larger sample of catchments, as well as the inclusion of additional catchment descriptors and hydrological variables.

Our purpose here is twofold. First, we aim to show that the frequency-matching approach, despite being very simple in its application, can capture hydrological variability within the identified catchment classes, as reflected by consistent relationships in the data. Second, our findings provide guidance on estimating catchment-representative runoff coefficients in a straightforward yet robust way, so that they can serve as objective indicators for classifying catchments. Applications that can benefit from this classification are those based on the characteristics of the catchment hydrological response for flood frequency analysis or machine learning algorithms for streamflow prediction. The context-based assessments presented in this paper provide statistical models a better-informed, physically sound parameter, reducing the need for them to infer all relationships from unconstrained inputs.

6 Conclusions

This study presents the reconstruction of over one thousand flood runoff coefficient values for 60 small- to medium-sized catchments in Italy, based on historical hydrological extremes. Two estimation methods, an event-based approach and a frequency-based approach, are employed. One key aspect of the event-based method is the development of a procedure to associate discharge maxima with their corresponding rainfall maxima in time, which was necessary because, in many cases, the dates of occurrence of extremes (day/month) were not recorded. The analysis reveals substantial interannual variability in runoff coefficient values at the event scale. Nevertheless, the catchment-aggregated statistics, in particular the median values, exhibit strong consistency between the two methods. This finding suggests that the frequency-based approach can provide a computationally efficient and operationally feasible alternative for deriving representative runoff coefficients for a given catchment, especially when high-resolution data are unavailable.

Despite its greater complexity and data demands, the event-based method has helped us identify anomalies and data inconsistencies: analyses of anomalous cases highlighted key sources of uncertainty, including manual transcription errors, spatial discrepancies in rain gauge coverage, and snowmelt contributions in Alpine regions. These findings underscore the importance of critical data validation and the need for context-aware interpretation of runoff coefficients, particularly in regions characterized by complex hydrometeorological conditions, such as the mountainous areas of Italy.

Using continuous time series data from recent years, future, more detailed analyses will build upon this preliminary work. These investigations will focus on gaining a deeper understanding of the runoff coefficient's temporal variability, with special consideration of its relationship to antecedent rainfall.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was carried out within the RETURN Extended Partnership and received funding from the European Union Next-Generation EU (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005 - Spoke TS 2). This study was carried out within the «FRISBEE» project – funded by the European Union – Next Generation EU within the PRIN 2022 program (D.D. 104 - 02/02/2022 Ministero dell'Università e della Ricerca). This manuscript reflects only the authors' views and opinions and the Ministry cannot be considered responsible for them.

ORCID

Giulia Evangelista http://orcid.org/0000-0002-6188-6066
Pietro Bogoni http://orcid.org/0009-0004-8282-5715
Daniele Ganora http://orcid.org/0000-0003-0605-6200
Pierluigi Claps http://orcid.org/0000-0002-9624-7408

Data availability statement

Data sources used are listed as follows: basin morphological parameters and annual maxima of daily discharge are available in Claps *et al.* (2024), historical daily rainfall depths at rain gauges are retrieved from the Italian Hydrological Yearbooks (available at http://www.bio.isprambiente.it/annalipdf/).

References

Beck, H.E., *et al.*, 2023. High-resolution (1 km) Köppen-Geiger maps for 1901–2099 based on constrained CMIP6 projections. *Scientific Data*, 10 (1), 724. doi:10.1038/s41597-023-02549-6

Beven, K. and Smith, P., 2014. Concepts of information content and likelihood in parameter calibration for hydrological Simulation models. *Journal of Hydrologic Engineering*, 20 (1). doi:10.1061/(ASCE)HE.1943-5584.0000991

Blume, T., Zehe, E., and Bronstert, A., 2007. Rainfall-runoff response, event-based runoff coefficients and hydrograph separation. *Hydrological Sciences Journal*, 52 (5), 843–862. doi:10.1623/hysj.52.5.843

Blöschl, G., 2022. Three hypotheses on changing river flood hazards. *Hydrology and Earth System Sciences*, 26 (19), 5015–5033. doi:10.5194/hess-26-5015-2022

Claps, P., *et al.*, 2024. FOCA: a new quality-controlled database of floods and catchment descriptors in Italy. *Earth System Sciences Data*, 16 (3), 1503–1522. doi:10.5194/essd-16-1503-2024

Del Giudice, G., Padulano, R., and Rasulo, G., 2014. Spatial prediction of the runoff coefficient in Southern Peninsular Italy for the index flood estimation. *Hydrology Research*, 45 (2), 263–281. doi:10.2166/nh.2013.243

Dhakal, N., *et al.*, 2012. Estimation of volumetric runoff coefficients for Texas watersheds using land-use and rainfall-runoff data. *Journal of Irrigation and Drainage Engineering*, 138 (1), 43–54. doi:10.1061/(ASCE)IR.1943-4774.0000368

Dhakal, N., *et al.*, 2013. Rate-based estimation of the runoff coefficients for selected watersheds in Texas. *Journal of Hydrologic Engineering*, 18 (12), 1571–1580. doi:10.1061/(ASCE)HE.1943-5584.0000753

Fortelli, A., Scafetta, N., and Mazzarella, A., 2019. Nowcasting and real-time monitoring of heavy rainfall events inducing flash-floods: an application to Phlegraean area (Central-Southern Italy). *Natural Hazards*, 97 (622), 861–889. doi:10.1007/s11069-019-03680-7

Francipane, A., *et al.*, 2021. A paradigm of extreme rainfall pluvial floods in complex urban areas: the flood event of 15 July 2020 in Palermo (Italy). *Natural Hazards and Earth System Sciences*, 21 (8), 2563–2580. doi:10.5194/nhess-21-2563-2021

Froidevaux, P., *et al.*, 2015. Flood triggering in Switzerland: the role of daily to monthly preceding precipitation. *Hydrology and Earth System Sciences*, 19 (9), 3903–3924. doi:10.5194/hess-19-3903-2015

Grimaldi, S., *et al.*, 2021. Continuous hydrologic modelling for design simulation in small and ungauged basins: a step forward and some tests for its practical use. *Journal of Hydrology*, 595, 125664. doi:10.1016/j.jhydrol.2020.125664

Hawkins, R.H., 1993. Asymptotic determination of runoff curve numbers from data. *Journal of Irrigation and Drainage Engineering*, 119 (2), 334–345. doi:10.1061/(ASCE)0733-9437(1993)119:2(334)

Hjelmfelt, A.T., 1980. Empirical investigation of curve number technique. *Journal of the Hydraulics Division*, 106 (9), 1471–1476. doi:10.1061/JYCEAJ.0005506

Jahanshahi, A. and Booij, M., 2023. Exploring controls on rainfall–runoff events: spatial dynamics of event runoff coefficients in Iran. *Hydrological Sciences Journal*, 68 (7), 954,966. doi:10.1080/02626667.2023.2193297

Lapides, D.A., Sytsma, A., and Thompson, S., 2021. Implications of distinct methodological interpretations and runoff coefficient usage for rational method predictions. *Journal of the American Water Resources Association*, 57 (6), 859–874. doi:10.1111/1752-1688.12949

Ley, R., *et al.*, 2011. Catchment classification by runoff behaviour with self-organizing maps (SOM). *Hydrology and Earth System Sciences*, 15 (9), 2947–2962. doi:10.5194/hess-15-2947-2011

Machado, R.E., Cardoso, T.O., and Mortene, M.H., 2022. Determination of runoff coefficient (C) in catchments based on analysis of precipitation and flow events. *International Soil and Water Conservation Research*, 10 (2), 208–216. doi:10.1016/j.iswcr.2021.09.001

Maidment, D.R., 1993. *Handbook of Hydrology*. New York: McGraw-Hill.

Massari, C., *et al.*, 2023. On the relation between antecedent basin conditions and runoff coefficient for European floods. *Journal of Hydrology*, 625, 130012. doi:10.1016/j.jhydrol.2023.130012

Merz, R. and Blöschl, G., 2009. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research*, 45 (1), W01405. doi:10.1029/2008WR007163

Merz, R., Blöschl, G., and Parajka, J., 2006. Spatio-temporal variability of event runoff coefficients. *Journal of Hydrology*, 331 (3–4), 591–604. doi:10.1016/j.jhydrol.2006.06.008

Norbiato, D., *et al.*, 2009. Controls on event runoff coefficients in the eastern Italian Alps. *Journal of Hydrology*, 375 (3–4), 312–325. doi:10.1016/j.jhydrol.2009.06.044

Rahi, A., *et al.*, 2023. Unraveling hydroclimatic forces controlling the runoff coefficient trends in central Italy's Upper Tiber Basin. *Journal of Hydrology: Regional Studies*, 50, 101579. doi:10.1016/j.ejrh.2023.101579

Ross, C.A., *et al.*, 2020. Evaluating the Ubiquity of Thresholds in Rainfall-Runoff Response Across Contrasting Environments. *Water Resources Research*, 57 (1), e2020WR027498. doi:10.1029/2020WR027498

Saft, M., *et al.*, 2015. The influence of multiyear drought on the annual rainfall-runoff relationship: an Australian perspective. *Water Resources Research*, 51 (4), 2444–2463. doi:10.1002/2014WR015348

Schaake, J.C., Jr, Geyer, J.C., and Knapp, J.W., 1967. Experimental examination of the rational method. *Journal of the Hydraulics Division*, 93 (6), 353–370. doi:10.1061/JYCEAJ.0001717

Tarasova, L., *et al.*, 2018b. Exploring controls on rainfall-runoff events: 2. Regional patterns and spatial controls of event characteristics in Germany. *Water Resources Research*, 54 (10), 7688–7710. doi:10.1029/2018WR022588

Tarasova, L., *et al.*, 2018a. Exploring controls on rainfall-runoff events: 1. time series- based event separation and temporal dynamics of event runoff response in Germany. *Water Resources Research*, 54 (10), 7711–7732. doi:10.1029/2018WR022587

Viglione, A., Merz, R., and Blöschl, G., 2009. On the role of the runoff coefficient in the mapping of rainfall to flood return periods. *Hydrology and Earth System Sciences*, 13 (5), 577–593. doi:10.5194/hess-13-577-2009

Wasko, C. and Guo, D., 2022. Understanding event runoff coefficient variability across Australia using the hydroEvents R package. *Hydrological Processes*, 36 (4), e14563. doi:10.1002/hyp.14563

Young, C.B., McEnroe, B.M., and Rome, A.C., 2009. Empirical determination of rational method runoff coefficients. *Journal of Hydrologic Engineering*, 14 (12), 1283–1289. doi:10.1061/(ASCE)HE.1943-5584.0000114

Zheng, Y., *et al.*, 2023. Controls on the Spatial and Temporal Patterns of Rainfall-Runoff Event Characteristics—A Large Sample of Catchments Across Great Britain. *Water Resources Research*, 59 (6), e2022WR033226. doi:10.1029/2022WR033226