Beyond Imperviousness: The Role of Antecedent Wetness in Runoff Generation in Urbanized Catchments

Mohamed Saadi, Ludovic Oudin, Pierre Ribstein

Beyond Imperviousness: The Role of Antecedent Wetness in Runoff Generation in Urbanized Catchments

Mohamed Saadi*, Ludovic Oudin, Pierre Ribstein

Sorbonne Université, CNRS, EPHE, UMR METIS, F-75000, Paris, France.

*Corresponding author : *mohamed.saadi@sorbonne-universite.fr*

Abstract

Explaining the response of catchments with a mix of pervious and impervious surfaces requires the inclusion of controls other than catchment imperviousness. In this study, we made use of a large sample of 70,227 events in 852 United States catchments to analyze the interplay of total impervious area (TIA) and antecedent soil moisture (ASM) in determining the event runoff ratio (RR). First, we investigated the importance of TIA as a spatial control compared to other catchment attributes (describing climate and catchment landscape) in explaining the differences between catchments in terms of median event-scale RR. Second, we examined the importance of ASM as a temporal control of the variability of RR in comparison with spatial settings and other event characteristics. Third, events were grouped into 30 classes of TIA and 30 classes of ASM, and linear regression models were used to analyze the evolution of RR depending on each one of the two independent variables. Results showed the following: **(1)** Analysis of the importance of space-varying catchment descriptors highlighted the relevance of TIA as a control of catchment response, whereas ASM showed a stronger ability compared to other event characteristics. **(2)** ASM played a significant role in urbanized catchments as in rural ones. The absence of abrupt changes in the linear model parameters indicates that the impact of urbanization on the relationship between RR and ASM was relatively gradual. **(3)** Differences in RR between rural and intensively urbanized catchments were observable not only for dry pre-event conditions, but also for wet short-term conditions.

Plain language summary

Urbanization causes a substantial increase in surface sealing which leads to dramatic impacts on water cycle. Reliable projections of the impact of urbanization requires a good understanding of the multiple factors that shape the response of a catchment along a wide range of urbanization levels. In catchments with a mix of urban and natural areas, the interplay of total impervious area (TIA) and antecedent soil moisture (ASM) conditions is still not well elucidated. To fill this gap, we attempted to quantify

the relative importance of TIA and ASM in determining the event-scale catchment response using a sample of 852 United States catchments. We found that both controls play an important role in runoff generation. Particularly, ASM should be considered even in catchments with a high level of urban cover. We also found that the impact of total impervious area on runoff ratio was relatively gradual, and became significant when TIA exceeded 5%. This study provides a solid empirical understanding of the behavior of urbanized catchments, which is a step toward more reliable prediction of the impact of urbanization on hydrology for urban planning policies.

1   Introduction

Runoff is an aggregated result of the interaction of numerous hydrological processes within a catchment. Regarded as the catchment response to climatic forcing, it is a complex function of interrelated factors such as climate, morphology, geopedology, and land use. Hierarchizing these factors is necessary in order to understand and predict the behavior of catchments under change. In particular, urbanization and the associated increase in surface imperviousness constitute perhaps one of the most dramatic human-induced changes with regard to landscapes and hydrological processes (Vörösmarty *et al.*, 2010). Yet, quantifying the impact of urbanization on catchment response is still not straightforward due to balancing factors originating from the non-urbanized part of the catchment as well as the dependency on catchment properties.

The event runoff ratio (RR) is a widely used hydrological metric to characterize the catchment response (Blume, Zehe, and Bronstert, 2007). It expresses the fraction of rainfall that is converted into runoff. Numerous factors drive the variability of RR, some of which are variable in time (such as event rainfall depth and intensity, and antecedent soil moisture conditions; see, e.g., Graeff *et al.*, 2012; Hewlett, Fortson, and Cunningham, 1977; Tarasova *et al.*, 2018a), while others reflect the catchment hydrological properties, and thus vary in space from one catchment to another (e.g., climate setting, land use, slope, and soil characteristics; see Addor *et al.*, 2018; Merz and Blöschl, 2009; Tarasova *et al.*, 2018b). In the case of urbanized catchments, RR generally increases with the total impervious area (TIA) of the catchment (Miller and Hess, 2017; Zhou *et al.*, 2017), as a consequence of reduced infiltration due to surface sealing. Although urban-induced imperviousness (measured by TIA) is widely considered as a major control by many predictive regression relationships in urbanized areas (e.g., Survey, 1994), the effect of urbanization on catchment behavior is dependent on the local physical characteristics (Price, 2011). Hierarchizing all involved factors (i.e., soil, climate, topography, and land use) is thus needed to elucidate the relative importance of each one in influencing the hydrological response, particularly that of TIA.

Antecedent soil moisture (ASM) modulates the temporal variability of event-scale RR, and thus it is a widely considered control to explain and predict the hydrological behavior of non-urbanized catchments (Berghuijs *et al.*, 2016). Previous studies attempted to represent the interplay of TIA and ASM in the variability of the RR of urbanized catchments (Ando, Musiake, and Takahasi, 1984; Brun and Band, 2000; Zhang and Shuster, 2014), while the relevance of ASM as a control of RR in highly urbanized catchments is still debated. Smith *et al.* (2013) found that the 5-day antecedent rainfall, used as a proxy for ASM, did not play a significant role in the generation of RR for seven urbanized catchments in the Baltimore region (Unites States). Zhou *et al.* (2017) analyzed the importance of ASM

2

for a sample of 16 catchments located in Charlotte, North Carolina, with TIA ranging between 0.08 and 0.50. They found that the influence of ASM on RR decreased with increasing TIA. Miller and Hess (2017) analyzed the role of ASM for eight catchments located in the United Kingdom, and found that RR was less dependent on ASM in urbanized catchments than in rural ones.

These studies suggest that ASM plays a less important role in the runoff generation as the catchment imperviousness increases. However, they focused on a small number of catchments, with a range of imperviousness that was not rich enough to reach a general conclusion, nor to quantify the explanatory power of ASM in comparison with TIA.

To address these gaps, we propose to analyze the relationship between the catchment response, materialized by RR, and ASM across a wide range of TIA. The novelty of this study is the use of a large sample of 70,227 rainfall-runoff events from 852 urbanized and rural catchments located in the United States (US), which offers a large diversity of urbanization levels resulting in a wide range of TIA. Another key novelty consists in considering both the space-varying factors (defined as catchment attributes, which vary in space from one catchment to another) and the time-varying factors (defined as factors that vary from one event to another, i.e., event rainfall characteristics and ASM) by analyzing their relative importance in explaining the spatiotemporal variability of RR along a continuum of catchment imperviousness.

This paper is structured as follows. In Section 2, we present the catchment sample, and we describe how RR, TIA, and ASM were estimated. We explain how we assessed the importance of TIA (as a space-varying control) and ASM (as a time-varying control) in predicting RR. Then, we detail how the variability of RR was analyzed with regard to ASM and TIA, as well as how the sensitivity of RR to each variable was assessed. The results are presented and discussed in Sections 3 and 4, and Section 5 concludes with the main findings and some perspectives.

2 Dataset and methods

2.1 Catchment sample and spatial characteristics

We used a sample of 852 catchments located in the United States (Saadi, Oudin, and Ribstein, 2019), for which the snow effect and the influence of artificial reservoirs were limited. Their sizes ranged between 1.1 km^2 and 21,300 km^2, with a median value of 321 km^2 (interquartile range: 90 km^2-926 km^2). The majority of these catchments are located in Southeastern United States, and are characterized by a temperate humid subtropical climate (Beck *et al.*, 2018). Their aridity indices varied between 0.48 and 3.4, and a non-negligible number of catchments are located in areas with relatively high aridity indices (Figure 1), indicating the richness of the catchment sample with regard to climatic settings.

The hourly mean areal precipitation depths were extracted from the National Centers for Environmental Prediction's Stage-IV dataset using the `geoknife` R package (Read *et al.*, 2015; Lin and Mitchell, 2005). This dataset was constructed by mosaicking regional multi-sensor (i.e., radar and gages) precipitation analyses into a national product, gridded over the United States at a 4-km resolution. Hourly mean streamflow data were prepared by processing the instantaneous streamflow measurements, which we extracted from the GAGES-II database maintained by the United States Geological Survey (Falcone, 2011) using the `dataRetrieval` R package (Cicco *et al.*, 2018). Daily temperature was extracted
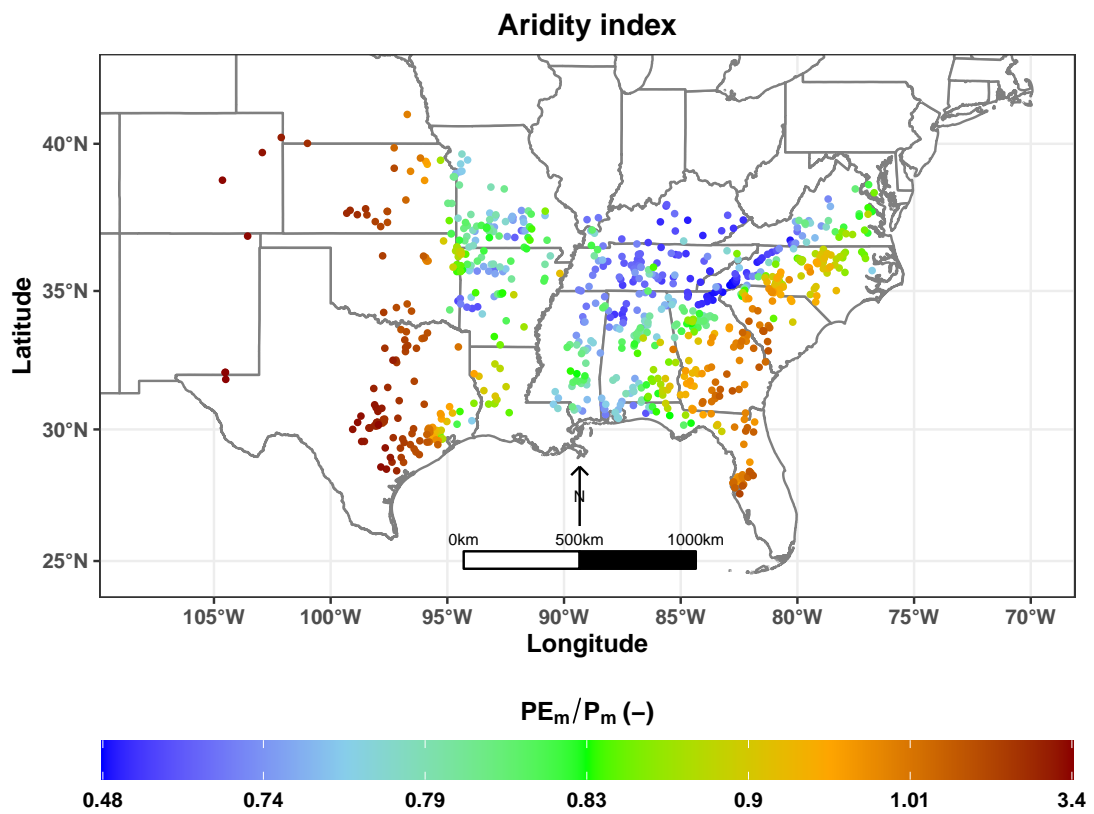
Figure 1. Location and aridity index of the 852 US catchments. Dots represent catchment centroids. Colors indicate aridity indices, computed as the ratio of average annual potential evapotranspiration PE_m (mm/year) to average annual precipitation P_m (mm/year).

from the Daymet product (Thornton *et al.*, 2016), available over the United States at a 1-km resolution. Recorded periods span between 2002 and 2017, with lengths ranging from 8 to 16 years.

Landscape and climate settings of each catchment were characterized using an initial set of 23 descriptors (Table 1). The land cover characteristics were extracted from the Multi-Resolution Land Characteristics Consortium's NLCD database, available for the years 2001, 2006, 2011, and 2016 (e.g., Homer *et al.*, 2015) at a 30-m resolution. These included the mean TIA (computed as the arithmetic mean of percentages of sealed surfaces estimated at each catchment pixel), catchment percent developed (CPD, computed as the proportion of catchment area occupied by urbanization classes), the fraction of catchment area with an imperviousness value higher than 0.8 ($IMP80$), the fraction of forest ($fFOR$), and the fraction of open water (fW). The distribution of imperviousness values of the catchment was characterized using the irregularity of imperviousness ($IrIMP$). A linear interpolation was used to estimate these characteristics for missing years.

Climate conditions were described using six metrics, including long-term mean precipitation (P_m), flashiness of precipitation (FP), hourly precipitation intensity coefficient (P_{99}/P_m), long-term mean potential evapotranspiration (PE_m), flashiness of potential evapotranspiration (FPE), and aridity index (AI). In addition to the catchment area ($Area$), the drainage density (DD), the catchment mean slope ($Slope$), the mean elevation ($Elev_m$), and the median value of compound topographic index (CTI) of catchment pixels were chosen to characterize the morphological settings. The mean content of sand ($Sand$), gravel ($Gravel$), silt ($Silt$), and clay ($Clay$), the mean soil porosity ($Porosity$) and the mean intrinsic permeability ($Perm$) were used as geopedological characteristics. Table 1 summarizes the catchment characteristics, their estimation methods, and their corresponding sources.

Table 1. Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected?[a]
Land use and imperviousness	Mean total impervious area	TIA (–)	Arithmetic mean of imperviousness values (i.e., percentage of developed surface) from catchment pixels		Selected
	Catchment percent developed	CPD (–)	Number of pixels occupied by developed classes divided by the total number of pixels		Rejected, highly correlated with TIA
	Fraction of catchment area with imperviousness value higher than 0.8	$IMP80$ (–)	Number of pixels with an imperviousness value higher than 0.8 divided by the total number of pixels	National Land Cover Database (NLCD) 2001, 2006, 2011, and 2016 (30-m resolution)	Rejected, highly correlated with TIA
	Fraction of forest	$fFOR$ (–)	Number of pixels occupied by forest classes divided by the total number of pixels		Selected
	Fraction of open water	fW (–)	Number of pixels occupied by open water class divided by the total number of pixels		Selected
	Irregularity of imperviousness	$IrIMP$ (–)	$IrIMP = \frac{\sqrt{\sum_j IMP_j^2}}{\sum_j IMP_j}$, where IMP_j is the imperviousness value of the pixel j of the catchment		Selected
Climate	Long-term mean precipitation	P_m (mm/h)	Total precipitation depth of the recorded period, divided by the number of hours		Selected
	Flashiness of precipitation	FP (–)	$FP = \frac{\sum_h |P_h - P_{h-1}|}{\sum_h P_h}$, with P_h the precipitation depth (mm) at hour h (Baker et al., 2004; Holko et al., 2011)	NCEP Stage-IV dataset, 4-km resolution, extracted using the `geoknife` R package (Read et al., 2015)	Selected

Continued on next page...

Table 1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected?[a]
	Hourly precipitation intensity coefficient	P_{99}/P_m (−)	The ratio of the 99th quantile of hourly precipitation P_{99} (mm/h) to long-term mean precipitation P_m (mm/h)		Selected
	Long-term mean potential evapotranspiration	PE_m (mm/h)	Total potential evapotranspiration depth of the recorded period, divided by the number of hours	Estimated using a temperature-based formula (Oudin et al., 2005). Daily temperature was extracted from Daymet dataset, 1-km resolution (Thornton et al., 2016)	Selected
	Flashiness of potential evapotranspiration	FPE (−)	$FPE = \frac{\sum_h |PE_h - PE_{h-1}|}{\sum_h PE_h}$, with PE_h the hourly depth of potential evapotranspiration (mm) at hour h (Baker et al., 2004; Holko et al., 2011)		Rejected, highly correlated with PE_m
	Aridity index	AI (−)	$AI = PE_m/P_m$	PE_m and P_m data sources	Selected
Morphology	Catchment area	$Area$ (km^2)	–	(Falcone, 2011)	Rejected, highly anti-correlated with FP

Continued on next page...

Table 1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected?[a]
	Drainage density	DD (km/km^2)	$DD = \frac{\sum_i L_i}{Area}$, with L_i length of stream i (km) and $Area$ the catchment area (km^2)	The hydrographic networks were extracted from the National Hydrography Dataset (NHD) using the `FedData` R package (Bocinsky, Beaudette, and Chamberlain, 2019)	Selected
	Mean slope	$Slope$ (–)	Arithmetic mean of slope values from catchment pixels		Rejected, highly correlated with $fFOR$
	Mean elevation	$Elev_m$ (m)	Arithmetic mean of elevation values from catchment pixels	Hydrologic Derivatives for Modeling and Applications (HDMA) database (Verdin, 2017)	Rejected, highly anti-correlated with PE_m
	Median compound topographic index	CTI (–)	$CTI = \text{median}\left(\log\left(\frac{A_{s,i}}{\tan(\beta_i)}\right)\right)$, with $A_{s,i}$ the specific contributing area (m) of the catchment pixel i and $\tan(\beta_i)$ its topographic slope		Rejected, highly anticorrelated with $fFOR$
Geopedology	Mean sand content of soil and subsoil geologic units	$Sand$ (%)	Weighted arithmetic mean of sand content values of catchment pixels (30% of topsoil and 70% of subsoil)		Selected
	Mean gravel content of soil and subsoil geologic units	$Gravel$ (%)	Weighted arithmetic mean of gravel content values of catchment pixels (30% of topsoil and 70% of subsoil)	Harmonized World Soil Database (HWSD, Version 1.2; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012)	Selected

Table 1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected?[a]
	Mean silt content of soil and subsoil geologic units	$Silt$ (%)	Weighted arithmetic mean of silt content values of catchment pixels (30% of topsoil and 70% of subsoil)		Selected
	Mean clay content of soil and subsoil geologic units	$Clay$ (%)	Weighted arithmetic mean of clay content values of catchment pixels (30% of topsoil and 70% of subsoil)		Selected
	Mean porosity of the catchment soil and subsoil geologic units	$Porosity$ (–)	Volume of voids divided by the total volume of soil and subsoil, estimated over sets of homogeneous polygons. For each catchment, mean porosity was estimated by area-weighted mean of values extracted from polygons in intersection with the catchment polygon	GLobal HYdrogeology MaPS (GLHYMPS; Gleeson *et al.*, 2014)	Selected
	Mean logarithmic values of intrinsic permeability	$Perm$ (m²)	As for porosity, mean intrinsic permeability was estimated by area-weighted mean of logarithmic values extracted from polygons in intersection with the catchment polygon		Selected

[a] By choosing a Spearman's rank correlation coefficient threshold $|r_{sp}| > 0.7$, variables that were highly correlated with the selected ones were rejected in order not to influence variable importance scores, as explained in Section 2.3.

End of Table 1.

2.2 Time-varying event and catchment characteristics

Many event-separation algorithms are reported in the literature (Mei and Anagnostou, 2015; Merz and Blöschl, 2009; Norbiato *et al.*, 2009). The event-separation method used in this paper is detailed in Appendix A, only the three main steps are mentioned hereafter. First, we proceeded by separating the baseflow from the direct flow, using a combination of the constant-k method (Blume, Zehe, and Bronstert, 2007) and a recursive digital filtering algorithm (Collischonn and Fan, 2013; Eckhardt, 2005; Mei and Anagnostou, 2015). Second, the direct flow time series, which is the difference between the total streamflow and the baseflow, was iteratively screened to look for peak flows in the vicinity of which the start and the end of each event were estimated using direct flow thresholds and a bunch of time characteristics estimated from the cross-correlation of direct flow and rainfall. Third, the event was retained after satisfying additional conditions related to peak flow importance, duration, and total rainfall depth. The event RR was computed as the ratio of direct flow to the total rainfall depth. This event-separation algorithm yielded 70,227 events, and the number of events per catchment ranged between 1 and 615, with a median of 50 events per catchment (interquartile range: 23-104).

For each event, a set of time-varying characteristics summarized in Table 2 were estimated. The first group of characteristics is composed of event rainfall depth P_{tot} (mm), rainfall intensity P_{int} (mm/h), and maximum hourly rainfall depth of the event P_{max} (mm/h). The second group is a set of variables derived to estimate catchment ASM. To this end, four types of proxies were tested:

- The discharge at the start of the event can be considered to reflect the state of catchment moisture (Cappus, 1960; Kirchner, 2009; Tarasova *et al.*, 2018a). Two measures were considered: the total discharge and the baseflow at the start of the event, both normalized by the mean annual streamflow of the catchment.

- The commonly applied Natural Resources Conservation Service's Curve Number (NRCS-CN) method traditionally considers the antecedent precipitation (AP) as a proxy of ASM, namely 5-day AP (Isik *et al.*, 2013; Soulis *et al.*, 2009). For each event, we have estimated a set of AP measures corresponding to different lengths of time windows before the start of the event, including 2, 5, 10, 15, and 20 days, in addition to a specific period Lag (in hours) estimated for each catchment as the delay Δt (in hours) that maximized the correlation between the precipitation time series $P(t)$ and direct flow $Q_d(t + \Delta t)$ (see Appendix A).

- The antecedent precipitations index (API) has been widely used as a proxy of ASM (Berthet *et al.*, 2009; Heggen, 2001; Kohler and Linsley, 1951). It is computed as a weighted sum of the antecedent rainfall depths according to their antecedence in time. As with AP, we considered six APIs including 2, 5, 10, 15, and 20 days as well as Lag hours of antecedent precipitation. The weighting parameter was estimated corresponding to a 10% weight for the earliest considered depth in each API (e.g., the 120th and the 360th hour before the start of each event for 5-day API and 15-day API, respectively).

- Measures based on soil moisture accounting reservoirs ($SMAR$; Berthet *et al.*, 2009; Merz, Blöschl, and Parajka, 2006). We used the $SMAR$ of the hourly hydrological model GR4H (Ficchì, Perrin, and Andréassian, 2019), due to its ability to represent ASM as shown in previous studies (Anctil *et al.*, 2004; Javelle *et al.*, 2010). It enables us to make full use of the precipitation (P) and potential

evapotranspiration (PE) hourly time series to get information on the current hydric state of the catchment. The equations of the $SMAR$ are detailed in Appendix B. The pre-event catchment ASM is taken as the ratio $\frac{S}{S_{max}}$ at the start of the event, where S is the updated $SMAR$ state (mm) and S_{max} (mm) the maximum capacity. This ratio varies between 0 (extremely dry condition, i.e., the $SMAR$ is empty) and 1 (extremely wet condition, i.e., the $SMAR$ is full). S_{max} regulates the $SMAR$ dynamics and memory; setting S_{max} low allows us to target the short-term wetness conditions (i.e., a few days before the event), whereas large values of S_{max} make it possible to target wetness conditions on longer terms. Four values of S_{max} were tested in this study: 10, 50, 100, and 200 mm.

Table 2. Summary of the time-varying event characteristics and antecedent soil moisture (*ASM*) proxies.

Type	Variable name	Notation and unit	Computation	Tested configurations	Selected/rejected?[a]
Event characteristics	Total rainfall depth	P_{tot} (mm)	$P_{tot} = \sum_{h \in event} P_h$, where P_h is the precipitation depth at hour h	–	Selected
	Rainfall intensity	P_{int} (mm/h)	$P_{int} = \frac{P_{tot}}{h_{end}-h_{start}}$, where P_{tot} is the event's total rainfall depth, h_{start} and h_{end} are the hours of event start and end	–	Selected
	Maximum hourly rainfall depth	P_{max} (mm/h)	$P_{max} = \max_{h \in event} P_h$, where P_h is the precipitation depth at hour h	–	Rejected, highly correlated with P_{tot}
Antecedent soil moisture conditions (*ASM*)	Pre-event total flow and baseflow	Q_E (–) and Q_{bE} (–)	The total streamflow and the baseflow at the start of the event, divided by the mean annual streamflow	–	Q_E selected Q_{bE} rejected, highly correlated with Q_E
	Antecedent precipitation	*AP*-D days (mm)	$AP = \sum_{h=1}^{N} P_{h_{start}-h}$, with h_{start} the start of the event, P_h is the precipitation depth at hour h, and $N = 24 \cdot D$ is the period length in hours	$N \in \{48$ h, 120 h, 240 h, 360 h, 480 h, *Lag*}, corresponding to *AP*-02, *AP*-05, *AP*-10, *AP*-15, *AP*-20, and *AP-Lag*	*AP-Lag* selected *AP*-02 rejected, highly correlated with *API*-02 *AP*-05 and *AP*-10 rejected, highly correlated with *API*-10 *AP*-15 and *AP*-20 rejected, highly correlated with *API*-15 and *API*-20 (which in turn are highly correlated with *API*-10)

Continued on next page...

Table 2 continued: Summary of the time-varying event characteristics and antecedent soil moisture (*ASM*) proxies.

Type	Variable name	Notation and unit	Computation	Tested configurations	Selected/rejected?[a]
	Antecedent precipitation index	*API*-D days (mm)	$API = \sum_{h=1}^{N} \alpha^h P_{h_{start}-h}$, with h_{start} the start of the event, P_h is the precipitation depth at hour h, and α chosen as $\alpha^N = 0.1$, where $N = 24 \cdot D$ is the period length in hours	$N \in \{48 \text{ h}, 120 \text{ h}, 240 \text{ h}, 360 \text{ h}, 480 \text{ h}, Lag\}$, corresponding to *API*-02, *API*-05, *API*-10, *API*-15, *API*-20, and *API*-*Lag*	*API*-02 and *API*-10 selected, *API*-05, *API*-15, and *API*-20 rejected, highly correlated with *API*-10, *API*-*Lag* rejected, highly correlated with *AP*-*Lag*
	Soil moisture accounting reservoir	*SMAR*-S_{max} (−)	Presented in Appendix B	$S_{max} \in \{10 \text{ mm}, 50 \text{ mm}, 100 \text{ mm}, 200 \text{ mm}\}$ corresponding to *SMAR*-10, *SMAR*-50, *SMAR*-100, and *SMAR*-200	*SMAR*-50 selected, *SMAR*-10, *SMAR*-100, and *SMAR*-200 rejected, highly correlated with *SMAR*-50

[a] By choosing a Spearman's rank correlation coefficient threshold $|r_{sp}| > 0.7$, variables that were highly correlated with the selected ones were rejected in order not to influence variable importance scores, as explained in Section 2.3.

End of Table 2.

2.3 Relevance of catchment imperviousness and ASM as controls of the spatial and temporal variability of RR

Before investigating the interplay of catchment imperviousness and ASM in determining event RR, we questioned their relevance in comparison with the remaining catchment and event characteristics. First, we ranked the space-varying factors (Table 1) based on their importance as explanatory variables of the evolution of median RR from one catchment to another. This will help to unveil the importance of catchment imperviousness in controlling average catchment hydrological behavior. Second, the time-varying factors (Table 2) were ranked based on their importance in predicting event RR, which helped to reveal the variables that mostly contributed to event-to-event variability of catchment response among spatial settings, event rainfall characteristics, and ASM proxies.

In our case, variable ranking was a difficult task as all space-varying and time-varying factors simultaneously and nonlinearly interact to decide the event-scale RR. To deal with the nonlinear and multidimensional aspects of the problem, we have employed the random forest algorithm (Breiman, 2001), which has been applied by recent studies that exploited large samples of catchments characterized by a relatively large number of descriptors (Addor *et al.*, 2018; Zhang *et al.*, 2018). In addition, random forest reduces the risk of overfitting and robustly handles nonlinear relationships between predictors while allowing for variable importance assessment. These advantages are achieved thanks to a column and row subsampling strategy. A random forest is constructed by growing a number of regression trees. At each tree growing, a number of independent variables is randomly sampled at each split (generally one third of variables), while a subset of data is left out of the training sample (36.8% of cases) in order to assess the accuracy of the constructed random forest, called out-of-bag (OOB) data. OOB data is used to assess the importance of each variable by randomly shuffling the values of the OOB cases specific to each tree. This random shuffling would lead to a decrease in the accuracy of the constructed random forest, and the greater the decrease in accuracy, the more important the variable in question. The decreases in accuracy are averaged over all the trees that included the considered variable, and the obtained mean is normalized by the standard deviation of decreases to compute the variable's Z-score of importance. As an accuracy measure, the mean squared error (MSE) is utilized, and the most important variables are the ones with the highest increases in MSE.

Using this method, simply ranking variables does not allow for deciding which ones are significantly important. To accomplish this, we used the Boruta algorithm (Kursa and Rudnicki, 2010) implemented in the `Boruta R` package (Kursa and Rudnicki, 2020). This algorithm creates copies of all variables, called "shadow variables", by shuffling the values of each one of the original variables. Then, both original and shadow variables are combined to run the random forest algorithm implemented in the `randomForest R` package (Liaw and Wiener, 2002) and obtain variable importance scores. The maximum Z-score of the shadow variables $MZSV$ is used to assess whether a variable is statistically relevant or not. A variable obtaining a Z-score higher than $MZSV$ scores a hit. At each iteration, the ratio of the cumulative number of hits to the total number of trials is used to assess the statistical significance of the variable importance by means of a binomial test. The outcome of this statistical test decides whether the variable is significantly relevant, significantly irrelevant, or without decision.

Boruta algorithm was used to rank both the space-varying and the time-varying variables, and decide whether a variable was significantly relevant or irrelevant in predicting RR (p-value< 0.01). As variable ranking is influenced by the multicollinearity of variables, we identified the ones that showed high

Spearman's rank correlation coefficient (i.e., $|r_{sp}| > 0.7$). Among the space-varying descriptors, we selected the ones that were best correlated (in the sense of Pearson's correlation coefficient) with the 852 median event-scale RR per catchment. This has led to a final selection of 16 out of the initial 23 space-varying descriptors, as detailed in Table 1. Similarly, we selected only the time-varying variables that were not highly correlated (Table 2) by preferring the ones that had the best Pearson's correlation with the 70,227 event RRs. This has led to a rejection of 14 initially considered variables, as detailed in Table 2.

Once the 16 space-varying variables were ranked, the ones that were considered to be relevant by the Boruta algorithm were used to predict the median event-scale RR for each catchment. This predicted median RR, noted $Spatial\ RR$, was combined with the 7 selected time-varying variables (Table 2) to rank them based on their importance in predicting the 70,227 event RRs. This helped to compare the importance of ASM variables with that of event rainfall characteristics and catchment descriptors reflected by $Spatial\ RR$.

2.4 Investigation of the relationship between RR and ASM for different levels of catchment imperviousness

In this section, we present the method used to analyze the relationship between RR and two *a priori* driving characteristics: a time-varying characteristic representing catchment wetness (ASM) and a space-varying characteristic representing the catchment imperviousness (TIA). The 70,227 rainfall-runoff events are grouped into $N_{TIA} \cdot N_{ASM}$ classes, where N_{TIA} is the number of classes of TIA and N_{ASM} the number of classes of the ASM variable (Table 2). For each independent variable (TIA or ASM), the classes were delimited by variable quantiles ($N_{TIA} + 1$ quantiles for TIA, and $N_{ASM} + 1$ quantiles for ASM) extracted from the event dataset. This pre-clustering aims at: **(i)** facilitating the visualization of the RR evolution across TIA and ASM classes, and **(ii)** deriving robust regression models by filtering RR variability among a given class of TIA and ASM.

First, the matrix

$$MRR \;=\; \begin{pmatrix} MRR_{1,1} & ... & MRR_{1,N_{ASM}} \\ ... & MRR_{i,j} & ... \\ MRR_{N_{TIA},1} & ... & MRR_{N_{TIA},N_{ASM}} \end{pmatrix}_{\substack{1 \le i \le N_{TIA} \\ 1 \le j \le N_{ASM}}} \tag{2.1}$$

is constructed, where $MRR_{i,j}$ represents the mean of runoff ratios RR for the class $\{i, j\}$ of events for which TIA values are located between the ith and $(i+1)^{th}$ quantile of TIA (the $(N_{TIA}+1)^{th}$ quantile being the max of TIA of all events) and ASM values are located between the jth and $(j+1)^{th}$ quantile of ASM (the $(N_{ASM}+1)^{th}$ quantile being the max of ASM of all events), with $1 \le i \le N_{TIA}$ and $1 \le j \le N_{ASM}$. This matrix is constructed to qualitatively understand how RR varies along the TIA axis (i.e., urbanization-induced contrasts) and ASM axis (i.e., wetness state-induced contrasts). In particular, $MRR_{1,1}$ and $MRR_{N_{TIA},1}$ represent the mean RR under dry conditions for the least and highest urbanized cases, respectively, whereas $MRR_{1,N_{ASM}}$ and $MRR_{N_{TIA},N_{ASM}}$ represent the mean RR under wet conditions for the least and highest urbanized cases, respectively.

Second, in order to quantify the RR gradient along TIA and ASM axes, a weighted least squares

(WLS) regression model was fitted to each row (i.e., RR as a linear function of ASM for a given class of urbanization) and each column (i.e., RR as a linear function of TIA for a given class of antecedent wetness conditions). The goodness of fit is estimated using the coefficient of determination R^2. In particular, the explanatory power of ASM across all the classes of TIA $R^2(RR, ASM)$ is computed as:

$$
\begin{cases}
R_i^2(RR, ASM) = 1 - \dfrac{\sum_{j=1}^{N_{ASM}} w_{i,j} \cdot \left(MRR_{i,j} - \widetilde{MRR_{i,j}(ASM)}\right)^2}{\sum_{j=1}^{N_{ASM}} w_{i,j} \cdot \left(MRR_{i,j} - \overline{MRR_i}\right)^2} & ; 1 \le i \le N_{TIA} \\
R^2(RR, ASM) = \underset{1 \le i \le N_{TIA}}{\text{median}} R_i^2(RR, ASM)
\end{cases}
\tag{2.2}
$$

where $MRR_{i,j}$ is the element $\{i, j\}$ of matrix MRR (Equation 2.1), $\{\widetilde{MRR_{i,j}(ASM)}\}_{1 \le j \le N_{ASM}}$ are estimations from the regression line $RR = f(ASM)$ for the class i of TIA, and $\overline{MRR_i}$ is the weighted mean of $\{MRR_{i,j}\}_{1 \le j \le N_{ASM}}$ for the class i. The weights $w_{i,j}$ were chosen as the inverse of the variance of RR of each class $\{i, j\}$ to account for the intraclass variability of RR. Similarly, the explanatory power of TIA across the ASM classes is estimated as:

$$
\begin{cases}
R_j^2(RR, TIA) = 1 - \dfrac{\sum_{i=1}^{N_{TIA}} w_{i,j} \cdot \left(MRR_{i,j} - \widetilde{MRR_{i,j}(TIA)}\right)^2}{\sum_{i=1}^{N_{TIA}} w_{i,j} \cdot \left(MRR_{i,j} - \overline{MRR_j}\right)^2} & ; 1 \le j \le N_{ASM} \\
R^2(RR, TIA) = \underset{1 \le j \le N_{ASM}}{\text{median}} R_j^2(RR, TIA)
\end{cases}
\tag{2.3}
$$

where $\{\widetilde{MRR_{i,j}(TIA)}\}_{1 \le i \le N_{TIA}}$ are estimations from the regression line $RR = f(TIA)$ for the class j of ASM, and $\overline{MRR_j}$ is the weighted mean of $\{MRR_{i,j}\}_{1 \le i \le N_{TIA}}$ for the class j.

Then, we analyzed the evolution of the parameters (i.e., slope and intercept) of the regression line $RR = f(ASM)$ along the TIA classes, as well as the evolution of the parameters of the regression line $RR = f(TIA)$ along the ASM classes.

3 Results

## 3.1 Assessing the relevance of catchment imperviousness and ASM as explanatory variables of event-scale RR

The catchment land-use characteristics, namely TIA and fraction of forest ($fFOR$), were the most important variables in explaining the spatial variability of median event RR, as indicated by Figure 2a. Aridity index (AI) obtained almost similar importance scores to those of $fFOR$, followed by mean content of gravel ($Gravel$), coefficient of hourly precipitation intensity (P_{99}/P_m), and mean intrinsic permeability ($Perm$), proving the strong controls of climatic and geopedological conditions over the catchment response. The following group of controls had equivalent importance scores, with mean content of silt ($Silt$) and flashiness of precipitation (FP) slightly ahead of mean content of sand ($Sand$), clay ($Clay$), and long-term averages of precipitation and potential evapotranspiration (P_m and PE_m). The last group of controls was constituted by the drainage density (DD) and the remaining land-use characteristics (fW and $IrIMP$). Nonetheless, all the variables were confirmed significantly important in explaining the spatial variability of RR, illustrating the multi-dependency of RR on catchment attributes.

Spatial RR, the predicted median RR for each one of the 852 catchments by random forest using the space-varying characteristics, proved to be the most relevant feature to explain the spatiotemporal variability of event RR, as shown in Figure 2b. This underlines the importance of the physical properties of the catchment in modulating its event-scale RR. The second most important control after *Spatial RR* was $SMAR$-50, underscoring the high importance of ASM as a primary control of RR ahead of rainfall characteristics, and the usefulness of the $SMAR$ technique as a proxy for ASM in predicting RR. *Spatial RR* and $SMAR$-50 were followed by the pre-event discharge (Q_E) and event characteristics (P_{int} and P_{tot}). A final group of time-varying variables was formed by the remaining ASM proxies (API-02, API-10, and AP-Lag), suggesting their relatively low relevance in explaining RR compared to other event characteristics. Finally, all employed variables were confirmed to be significantly important.

3.2   Analyzing RR variability using TIA and SMAR as a surrogate of ASM

RR is similarly controlled by both the imperviousness level, i.e., TIA, and the short-term ASM, estimated by $SMAR$-50, as shown in Figure 3. In particular, the RR in highly urbanized catchments (i.e., high TIA) was sensitive to changes in $SMAR$-50. Mean RR per each class of TIA and $SMAR$-50 varied between 0.01 and 0.43, the lowest values being attained for the rural cases during the dry period (bottom-left area in Figure 3), whereas the highest values were attained for the most urbanized cases during the wet period (top-right area in Figure 3). Moving from dry to wet conditions, mean RR increased for all catchments regardless of their level of TIA. Similarly, higher values of TIA always resulted in higher values of mean RR for all ASM conditions. Similar patterns were obtained using pre-event discharge measures, AP and API metrics instead of $SMAR$-50 (not shown here).

The obtained WLS linear regression models were satisfactory for both $SMAR$-50 and TIA, with median $R^2(RR, SMAR - 50) = 0.90$ over the 30 classes of TIA (range: 0.78-0.95) and $R^2(RR, TIA) = 0.76$ over the 30 classes of $SMAR$-50 (range: 0.30-0.92). The goodness of fit is illustrated by examples in Figure 4 for the least and most urbanized cases (bottom and top rows of Figure 3) and the lowest and highest ASM conditions (left and right columns of Figure 3). The first two examples (Figures 4a and 4b) illustrate the effect of $SMAR$-50 on mean event RR, where an increase in $SMAR$-50 was accompanied by increasing mean RR similarly for both the least urbanized cases and the most urbanized cases. The last two examples (Figures 4c and 4d) indicate the effect of TIA on mean event RR, where RR got increased with increased TIA. The effect of TIA is visually more pronounced amid dry conditions (i.e., very low $SMAR$-50) than in wet conditions (i.e., very high $SMAR$-50), but in both cases the differences in TIA were accompanied by differences in catchment response.
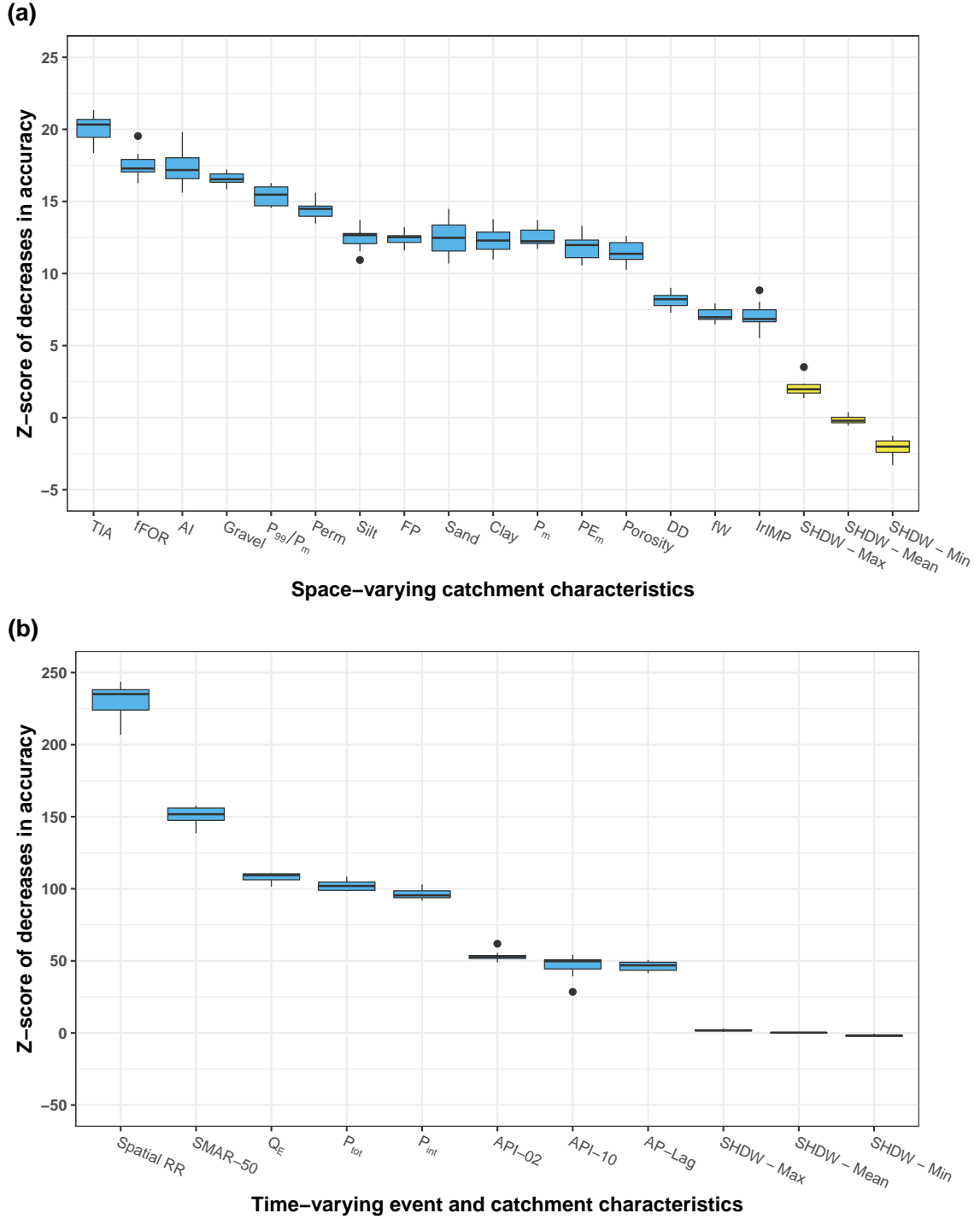
Figure 2. (a) Rankings of space-varying catchment characteristics (Table 1), based on their importance in explaining the variability of the 852 median catchment event-scale runoff ratios, and **(b)** rankings of *Spatial RR* and time-varying event and catchment characteristics (Table 2), based on their importance in explaining the variability of the 70,227 event runoff ratios. Importance is estimated using the Z-score of decreases in accuracy, i.e., of increases in mean squared errors (*MSE*) of the predictions of random forests. *Spatial RR* is the prediction of median catchment runoff ratio estimated by random forest using the space-varying catchment descriptors. *SHDW-Max*, *SHDW-Mean* and *SHDW-Min* are the distributions of maximum, mean and minimum values of Z-score of shadow variables, obtained using the `Boruta` R package.
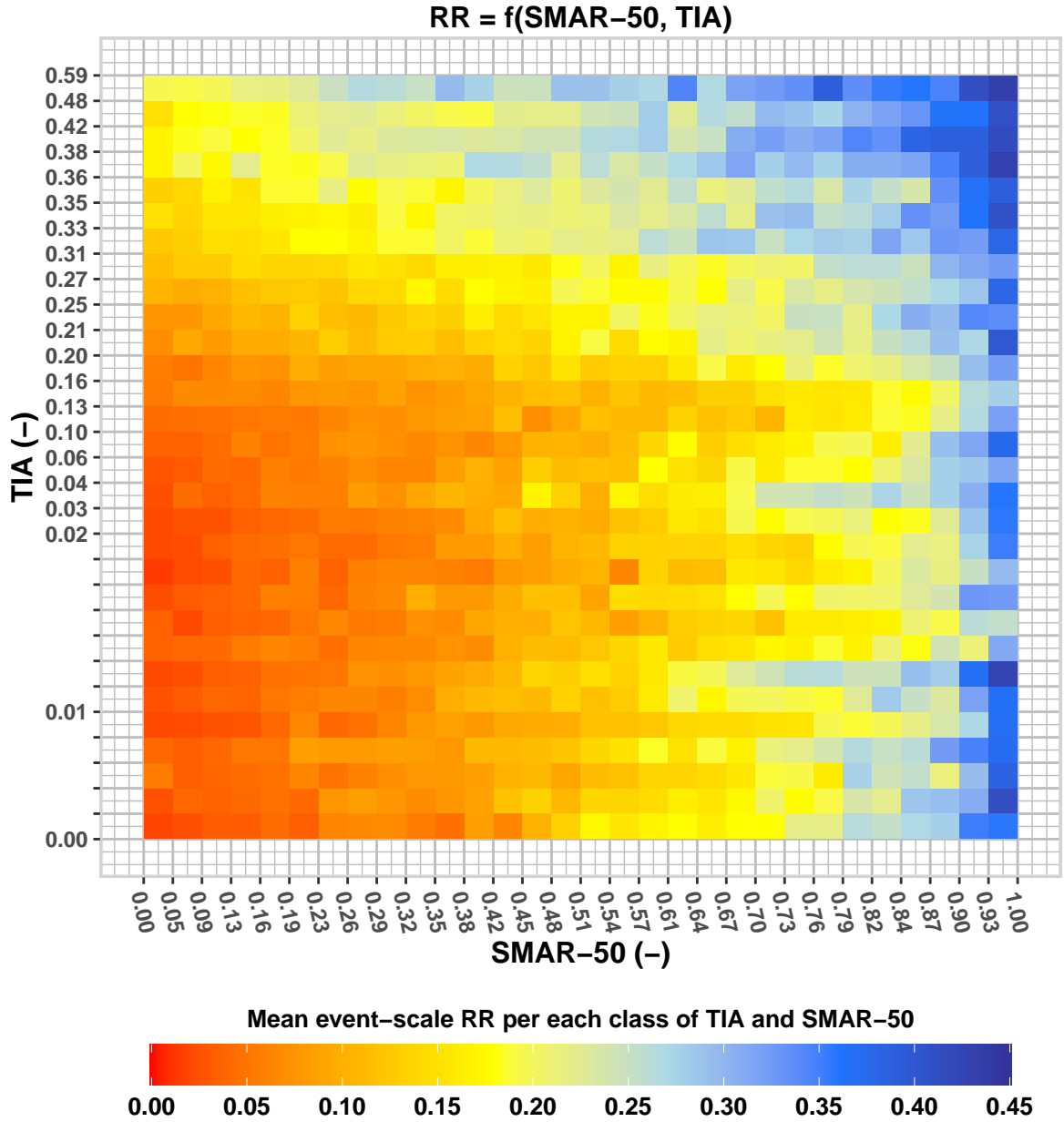
Figure 3. Runoff ratio (RR) dependence on antecedent soil moisture (ASM) conditions, assessed using the level of the soil moisture accounting reservoir ($SMAR$-50, Appendix B), and the catchment total impervious area (TIA). This matrix was constructed by setting $N_{TIA} = N_{SMAR-50} = 30$, yielding 900 classes of $TIA - SMAR$-50. The bottom row shows the evolution of mean RR for rural cases (i.e., TIA close to 0), and the top row shows mean RR for intensively urbanized ones (i.e., TIA between 0.48 and 0.59). The left column shows mean RR for dry conditions (i.e., $SMAR$-50 close to 0), and the right column shows mean RR for wet conditions (i.e., $SMAR$-50 close to 1).
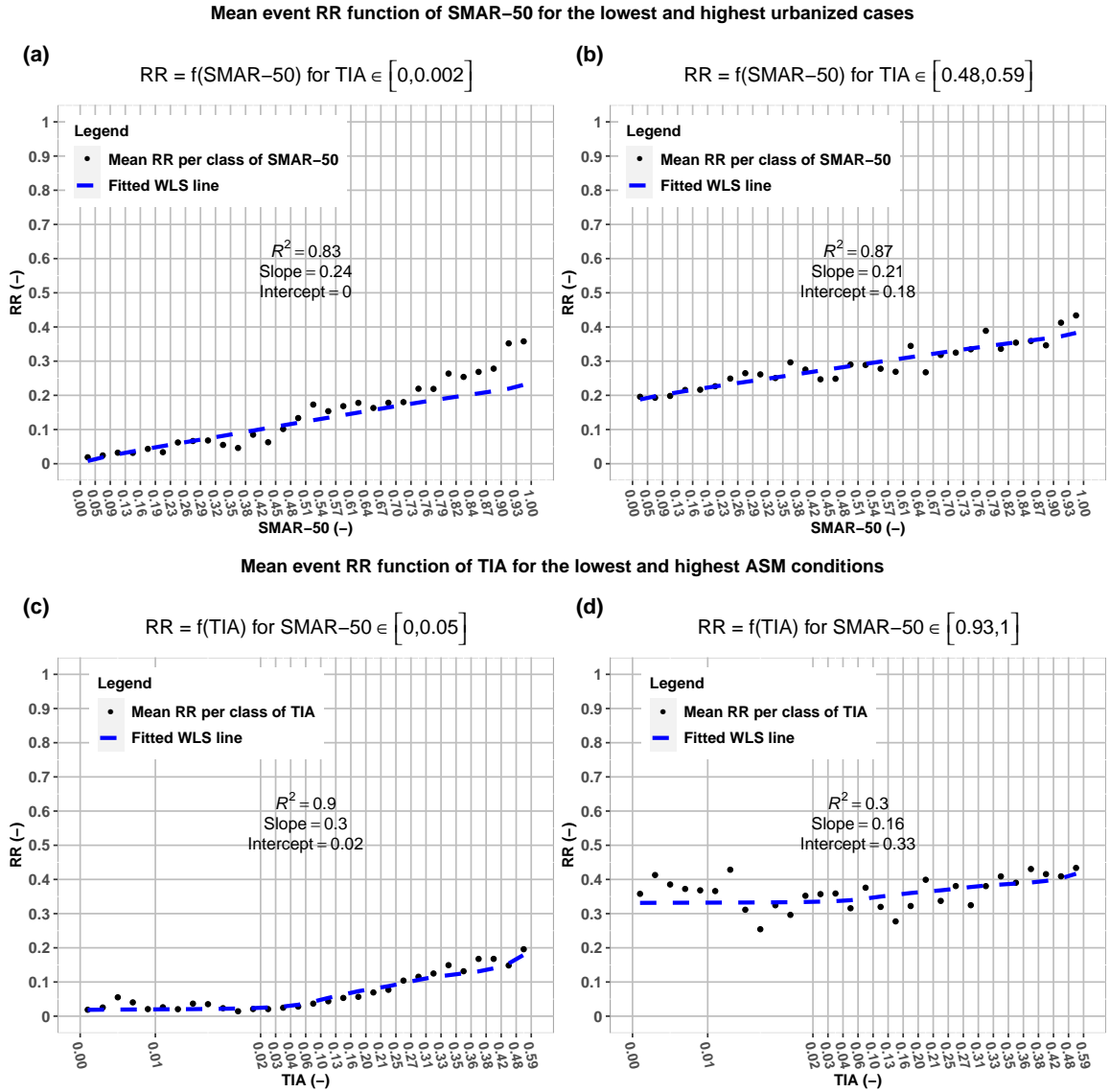
Figure 4. Scatter plots showing the variability of mean runoff ratio (RR) as a function of total impervious area (TIA) and antecedent soil moisture (ASM), assessed using the level of the soil moisture accounting reservoir ($SMAR$-50, Appendix B). The first row shows $RR = f(SMAR-50)$ for **(a)** the lowest TIA level and **(b)** the highest TIA level. The second row shows $RR = f(TIA)$ for **(c)** the driest ASM conditions and **(d)** the wettest ASM conditions. Black dots show mean RR per class of TIA and $SMAR$-50, and blue dashed lines show the fitted weighted least squares (WLS) regression models.

An overview of the evolution of the WLS model parameters suggests that the control of $SMAR$-50 on RR was not influenced by TIA, as manifested by the behavior of the slope of the WLS regression line $RR = f(SMAR - 50)$ in Figure 5a. In general, the slope did not deviate significantly from 0.21 for all classes of TIA. This means that even for the most urbanized catchments, short-term ASM still played a significant role in determining RR as for the rural ones. The differences between rural and urbanized catchments were most reflected in the intercept parameter, which could be interpreted as the value of RR in very dry conditions (i.e., $SMAR$-50 close to 0). The intercept was around zero for catchments with TIA less than 5%, and increased somewhat linearly to reach a value of 0.18 for high levels of TIA. This suggests that in dry conditions, RR is solely controlled by the impervious part of the catchment. Varying intercept with almost constant slope is equivalent to shifting upwardly the line $RR = f(SMAR - 50)$ as TIA gets higher, as illustrated in Figure 5b.

4   Discussion

4.1   Importance of ASM as a control of RR in urbanized catchments

This study analyzed the role of ASM, in addition to TIA, as controls of RR. The limited number of studies that have addressed the role of ASM in urbanized catchments suggested that ASM had a less important control on RR compared with TIA (Miller and Hess, 2017; Smith *et al.*, 2013; Zhou *et al.*, 2017). Our results showed that ASM is in fact an important control to consider for predicting the event RR even in catchments with high levels of TIA. As reported in some studies (Sillanpää and Koivusalo, 2015; Zhou *et al.*, 2017), the differences between rural and urbanized catchments were essentially during the dry period. We showed that this was also the case for very wet short-term ASM (Figures 3, 4, and 5). Although some of the findings are intuitive, they are purely data-based, and their importance lies most in providing means of revising methods in the practice of hydrology that do not account for the effect of ASM in urbanized areas. In addition, the findings confirm previously obtained results concerning the interplay of ASM and imperviousness, namely the logistic function model proposed by Brun and Band (2000), or the scenario-based conclusions of Zhang and Shuster (2014) and Yang *et al.* (2011). Methodologically, this study is a stepping stone to overcome the issue of the "rural-urban dichotomy" (McGrane, 2016), by analyzing the evolution of catchment behavior along a continuum of TIA, which would not have been possible without the use of a large sample of catchments.

Using $SMAR$ as a proxy for ASM (Appendix B) yielded better predictive power than event characteristics (i.e., P_{tot} and P_{int}) or other ASM proxies that were based on pre-event discharge and antecedent rainfall. The relevance of $SMAR$ as a proxy of ASM has been confirmed in other studies (Anctil *et al.*, 2004; Merz and Blöschl, 2009). Moreover, the $SMAR$ allows us to follow the seasonal variability of P and PE, in contrast to API that does not distinguish between hot and cold seasons, and the baseflow that is dependent on local configurations other than climate (e.g., soil type, connectivity of surface water and groundwater). The ranking of ASM proxies in this study is very helpful for event-based techniques applied to estimate runoff in urbanized areas. In particular, the NRCS-CN method would be enhanced by using $SMAR$-50 or API measures in estimating ASM instead of 5-day AP.

The large range of TIA values tested in this study showed that the impact of increasing TIA on the relationship between RR and ASM was more gradual than threshold-like. As illustrated by the evolution
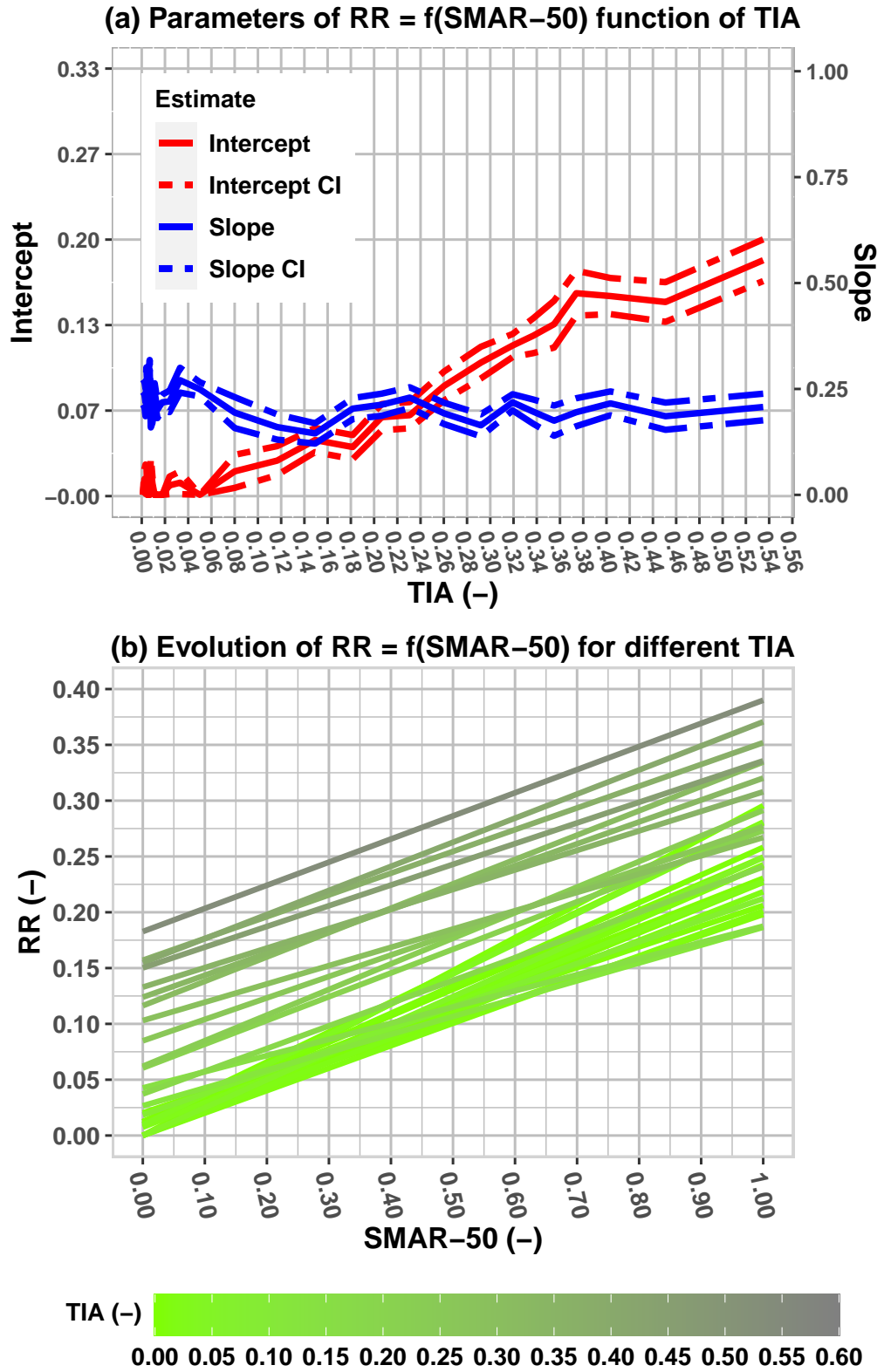
Figure 5. (a) Evolution of the parameters of the WLS regression line representing mean runoff ratio (RR) as a function of $SMAR$-50 for different levels of total impervious area (TIA). Estimates are shown with solid lines, and 95% confidence intervals (CI) are shown with two-dashed lines. Parameter estimates that were not statistically different from zero (p-value threshold at 0.05) were set to zero. **(b)** Evolution of the corresponding lines for different levels of TIA.

of the parameters of the regression line representing RR as a function of $SMAR$-50 (Figure 5a), for different values of TIA, the intercept started to deviate from zero as TIA exceeded 0.05, but no abrupt change was noticed. This advocates the continuousness of the effect of urbanization (see the discussion in Brabec, Schulte, and Richards, 2002), and implies that it would be difficult to detect a significant change in catchment behavior (in response to urbanization) if the absolute change in TIA or the difference between analyzed situations (as in a paired-catchment framework) is less than 10%-5%. The threshold of TIA above which the intercept became significantly different from zero (i.e., 0.05) has been reported in literature reviews (Alley and Veenhuis, 1983; Brabec, Schulte, and Richards, 2002; Salvadore, Bronders, and Batelaan, 2015).

Some would argue that TIA did not exceed 0.6 for the sample used here, leaving a relatively large range of TIA uncovered. When examining the proportion of the catchment that is occupied by urbanization classes, i.e., CPD (Falcone, 2011; Miller and Hess, 2017), this proportion varied between 0% and 100%, with a median value of 29%, meaning that the range of TIA used here covered a large variety of situations of catchment-scale urban sprawl, including catchments completely occupied by urbanization classes. Not reaching a TIA higher than 0.6 is perhaps a result of focusing on the catchment scale, as the minimum catchment size in this study is above 1 km^2, which implies the existence of more natural surfaces (i.e., with low imperviousness values) even for intensively urbanized catchments.

4.2 Relevance of other catchment and event characteristics in explaining the variability of RR

In addition to TIA, $fFOR$ had the highest importance in explaining the variability of median event RR for the catchment set, highlighting the primary control of land-use characteristics on catchment behavior (Merz and Blöschl, 2009; Tarasova *et al.*, 2018b). Climate characteristics (the aridity index and the coefficient of hourly precipitation intensity P_{99}/P_m) and geopedological settings were the next most relevant spatial controls, underscoring the interplay of climate and landscape features in defining the catchment long-term hydrological behavior (Gao *et al.*, 2018; Merz and Blöschl, 2009). Comparing the importance of spatial and temporal controls on the variability of event-scale RR illustrated the primacy of regional settings over ASM and event rainfall characteristics (Figure 2b), as the former defines the long-term behavior of the catchment while the latter shapes the short-term high-frequency variability (Merz and Blöschl, 2009). Event rainfall characteristics have been regularly chosen as explanatory variables (Blume, Zehe, and Bronstert, 2007; Hewlett, Fortson, and Cunningham, 1977; Tarasova *et al.*, 2018a), which has been confirmed by our study results (Figure 2b), but their importance was less than that of regional settings and ASM conditions.

Lastly, it has been shown that not only TIA influences catchment response, but also the spatial arrangement of impervious surfaces (Debbage and Shepherd, 2018; Mejía and Moglen, 2010), which has motivated to search for a more hydrologically relevant imperviousness measure, such as the effective impervious area (EIA) or the landscape fragmentation measures (Alley and Veenhuis, 1983; Ebrahimian, Wilson, and Gulliver, 2016; Oudin *et al.*, 2018). Moreover, not all urbanized cover is impervious to water (Berthier, Andrieu, and Creutin, 2004; Ragab *et al.*, 2003; Ramier, Berthier, and Andrieu, 2011; Redfern *et al.*, 2016), and different types of urban cover are characterized by a wide range of imperviousness. In this respect, we used a second-order measure of imperviousness, i.e., $IrIMP$, which

was a relevant indicator but showed a lesser importance compared to the remaining spatial controls in explaining the spatial variability of event-scale RR (Figure 2a).

5 Conclusion

In this study, we demonstrated the impact of space-varying and time-varying controls on event RR for a large sample of 852 catchments located in the US, characterized by a wide range of TIA. The main finding was that ASM still plays a significant role in shaping RR even for the most intensively urbanized catchments.

As a spatial control, TIA showed the best ability in explaining the spatial variability of catchment response, ahead of the fraction of forest and the remaining climatic, geopedological, and morphological settings. ASM was estimated by the soil moisture accounting reservoir ($SMAR$) of the GR4H model, as it was best able to explain the event-to-event variability of RR compared with event characteristics (P_{tot} and P_{int}), antecedent precipitation (AP and API), and pre-event discharge. Still, regional settings had the highest importance in explaining the variability of RR compared to ASM and event rainfall characteristics.

Using the $SMAR$ as a proxy for ASM, the events were grouped into 900 classes of TIA and ASM. Then, the dependency of RR on TIA and ASM was then quantified by means of linear regression. Mean event-scale RR per each class was found to increase with either increasing TIA or increasing ASM. In the highly urbanized catchments, RR increased as pre-event conditions became wet. Differences in RR were found between rural and intensively urbanized catchments in dry as well as in wet ASM conditions prior to the event. The impact of TIA was less important on the slope of $RR = f(ASM)$, whereas the intercept became significantly different from zero at $TIA > 0.05$.

The fact that the TIA was not uniformly distributed led us to adopt a quantile-based clustering procedure, which would have been less harmful given a wider TIA range. Also, more diverse climatic conditions are missing in this study (including hotter and colder climates; see, e.g., Sillanpää and Koivusalo, 2015), which should be included in order to get an overall understanding of the interplay between climate and urbanization. The analysis was restrained here to mean RR values for each class, and an analysis of intraclass variability would help to synthesize the effect of TIA and ASM on the overall distribution of RR. For process-based modeling, the findings of this study are highly relevant to better represent/reproduce the rainfall-runoff relationship especially for catchments with evolving urbanization. In addition, they offer means to check the physical soundness of commonly applied regression relationships in practical hydrology in order to enhance the reliability of their predictions.

A Event separation method

Many event-separation algorithms have been reported in the literature (Mei and Anagnostou, 2015; Merz and Blöschl, 2009; Norbiato *et al.*, 2009; Tarasova *et al.*, 2018a). The method adopted in this work is empirical, and consists of the following steps.

Step 1: Baseflow separation

Using numerical filtering-based methods is more practical when working on a large sample of catchments (Tarasova *et al.*, 2018a). We used a combination of the constant-k method (Blume, Zehe, and Bronstert, 2007) and a recursive digital filtering algorithm (Collischonn and Fan, 2013; Mei and Anagnostou, 2015; Eckhardt, 2005). We chose this method for its physical soundness: It hypothesizes that the baseflow is associated with the low-frequency component of the total streamflow. Tarasova *et al.* (2018a) compared five methods of baseflow extraction, and the method used herein has behaved acceptably.

Step 2: Estimation of characteristic response times

The direct runoff Q_d, that is considered to be the catchment response to rainfall events, was computed as the difference between the total flow and the baseflow. To constrain the event selection, we determined a set of catchment characteristic response times based on the cross-correlation of the time series of rainfall and direct flow, besides the direct flow auto-correlation.

Three characteristic response times were determined (Figure 6):

- The lag time Lag (in hours), which is the positive shift Δt that maximizes the correction between $Q_d(t + \Delta t)$ and $P(t)$.

- Sometimes, the correlogram of $Q_d(t + \Delta t)$ and $P(t)$ does not show a sharp maximum and can attain significantly high values for $\Delta t > Lag$ or $\Delta t < 0$. A second metric of response time Lag_{max} (in hours) was calculated to account for this issue. The correlogram is scaled by its maximum. Three limits were defined: **(1)** $Limit_1$ corresponded to the peak of the scaled correlogram for $\Delta t < 0$; **(2)** $Limit_2$ was taken as the peak of the scaled correlogram for $\Delta t > dec$, where dec corresponded to the moment when the correlogram reincreased for the first time; and **(3)** an arbitrary limit of $Limit_3 = 0.2$ (i.e., 20% of the maximum value of the correlogram). The maximum of these three limits, noted $LimitPQd$, was then explored in the scaled correlogram, and Lag_{max} was the shift corresponding to this limit.

- The auto-correlogram of Q_d was also investigated to determine the memory Mem (in hours) of the catchment. Normally, the auto-correlogram should decrease for any $\Delta t > 0$. Mem corresponded to the peak of the auto-correlogram for $\Delta t > dec_{Q_d}$, where dec_{Q_d} is the moment at which the auto-correlogram reincreased for the first time. When no reincrease was registered, Mem corresponded to the shift at which an autocorrelation of $LimitQd = 0.2$ was attained for the first time.

Step 3: Estimation of the start of the event

The time series of $Q_d(t)$ were screened to look for a peak flow Q_d^p. Once found, a limit Q_d^{start} function of Q_d^p was determined. Over a time window TW_{start}, the first estimation of the start of the event corresponds to the moment when $Q_d(t) = Q_d^{start}$. The rainfall depth over a time window TWP_{start} before the start should be less than 10% of the rainfall depth between the start and the peak moments.
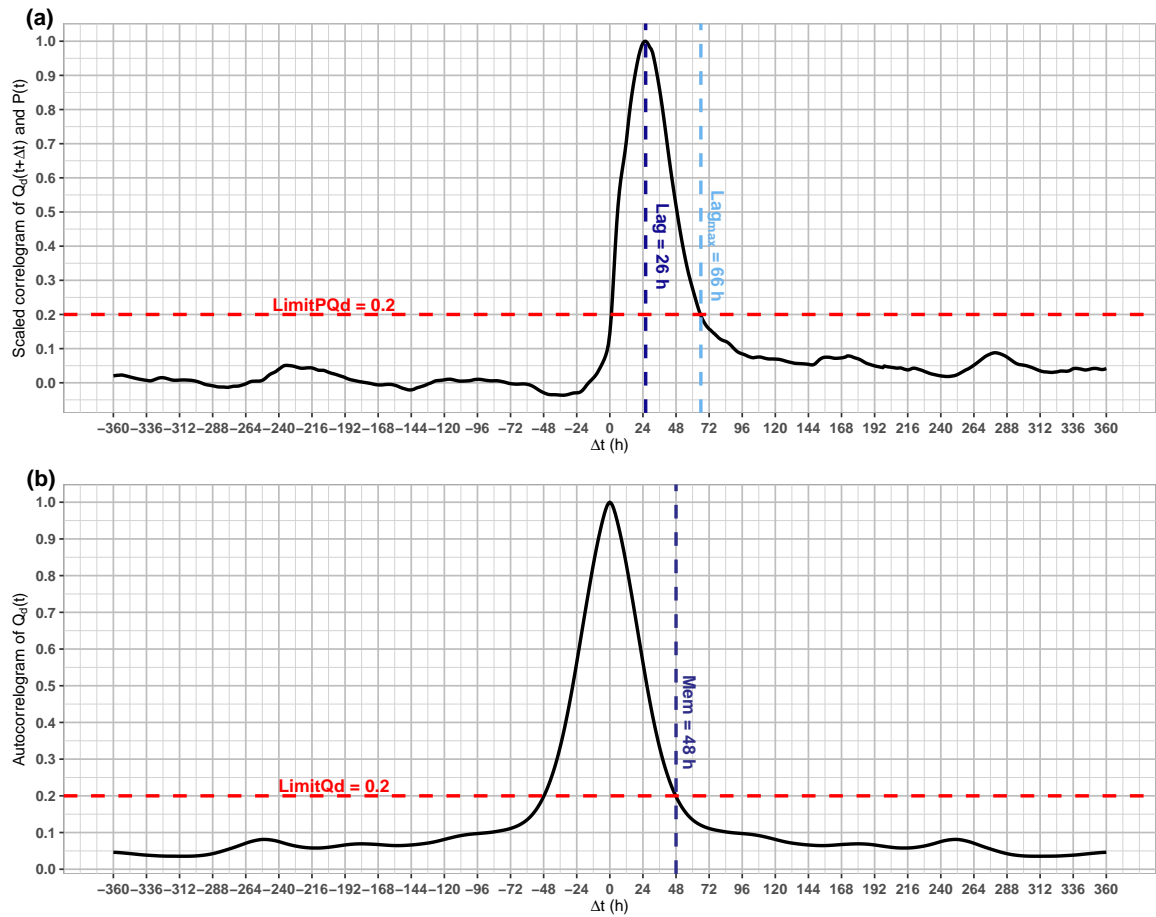
Figure 6. Examples of correlograms computed for the catchment drained by the USGS station 01658000 Mattawoman Creek near Pomonkey, Maryland, which has an area of 143.6 km^2. **(a)** Scaled correlogram of hourly direct discharge Q_d and hourly precipitation P, and **(b)** autocorrelogram of Q_d.

Table 3. Summary of the steps of the event separation method.

Steps	Settings
Start of the event	• $Q_d^{start} = \max\left(\frac{Q_d^p}{10}, 0.05 \cdot (Q_d^p - Q_d^0)\right)$, where Q_d^0 is the minimum of $Q_d(t)$ for $t \in [t_{peak} - Lag - Lag_{max}, t_{peak}]$, where t_{peak} is the peak moment • $TW_{start} = [t_{peak} - Lag - Lag_{max}, t_{peak}]$ • $TWP_{start} = [t_{deb} - Lag_{max}, t_{deb}]$, where t_{deb} is the start moment
End of the event	• $Q_d^{end} = \max\left(\frac{Q_d^p}{10}, 0.05 \cdot (Q_d^p - Q_d^1)\right)$, where Q_d^1 is the minimum of $Q_d(t)$ for $t \in [t_{peak}, t_{peak} + Mem]$, where t_{peak} is the peak moment • $TW_{end} = [t_{peak}, t_{peak} + Mem]$ • $p = 50$
Additional conditions	• The ratio of baseflow to total flow at t_{peak} should be less than BFI • No missing values are allowed • Direct runoff depth should be less than the rainfall depth • The rainfall depth should be higher than 2 mm • Very short or very long events are eliminated (the event duration should be between $0.2 \cdot Mem$ and $1.2 \cdot Mem$)

Step 4: Determination of the end of the event

As in step 3, a limit Q_d^{end} function of Q_d^p is determined. A time window TW_{end} after the peak is screened to identify the moment when $Q_d(t) = Q_d^{end}$. The rainfall depth between the peak and the end should be less than p% of the rainfall depth between the start and the peak.
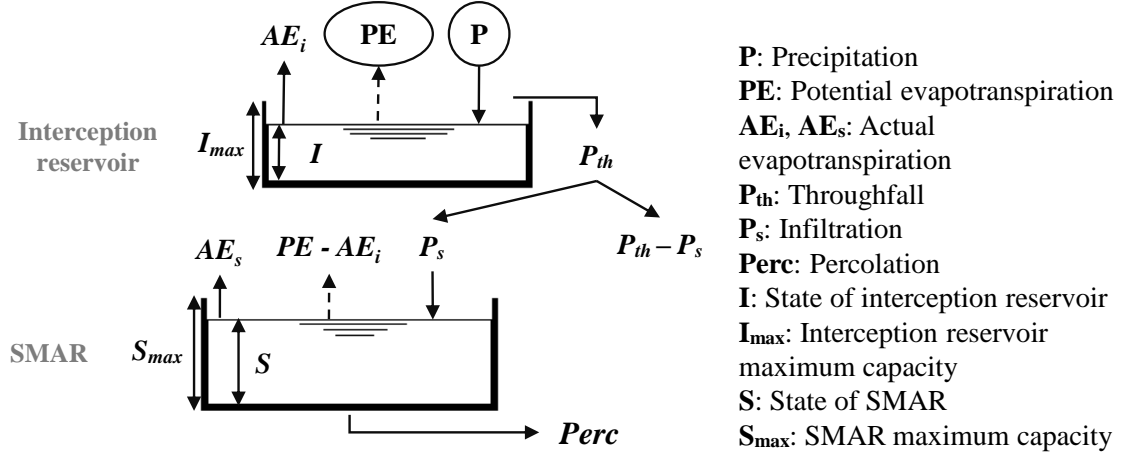
Step 5: Additional conditions

Once the start and the end of an event are determined, the event must satisfy the following conditions: **(1)** At the peak moment, the ratio of the baseflow to the total flow should be less than the whole-period baseflow index of the catchment BFI, estimated from the baseflow separation step; **(2)** The rainfall and direct runoff time series should not contain missing values between the start and the end of the event; **(3)** The sum of the event rainfall depth should be higher than 2 mm; **(4)** The total volume of direct runoff should be less than the total volume of event rainfall (i.e., runoff ratio less than 1); and **(5)** The total duration of the event should be greater than 20% of Mem and less than 120% of Mem.

Finally, event RR is taken as the ratio of the direct runoff depth to the rainfall depth. Table 3 summarizes the steps and the corresponding thresholds.

B   Equations of the soil moisture accounting reservoir ($SMAR$)

As a proxy for antecedent soil moisture (ASM), we used the soil moisture accounting reservoir ($SMAR$) of the GR4H hourly model (Ficchì, Perrin, and Andréassian, [2019]). Detailed equations are given in Figure 7. First, the hourly throughfall P_{th} is estimated using the equations under Step 1 (Figure 7), where P is the hourly precipitation, PE the hourly potential evapotranspiration, and I the interception reservoir state. Second, when PE is not satisfied solely by the interception reservoir content, a depth AE_s is subtracted from the $SMAR$ (Step 2). Third, we estimated the proportion of P_{th} that feeds the $SMAR$, noted P_s, which represents the amount of water that is expected to satisfy the initial losses through soil storage. Fourth, a depth representing the percolation is removed from the $SMAR$. P_s and AE_s are both a function of the state of the reservoir S, and its maximum capacity S_{max}. Finally, the pre-event catchment ASM is taken as the ratio S/S_{max} at the start of the event. This ratio varies between 0 and 1, where a value of 0 indicates a completely dry condition, and 1 indicates a completely wet condition.

P: Precipitation
PE: Potential evapotranspiration
AE$_i$, AE$_s$: Actual evapotranspiration
P$_{th}$: Throughfall
P$_s$: Infiltration
Perc: Percolation
I: State of interception reservoir
I$_{max}$: Interception reservoir maximum capacity
S: State of SMAR
S$_{max}$: SMAR maximum capacity

Step 1: Estimation of throughfall P$_{th}$

   1.1 $AE_i = \min(PE\,;I + P)$

   1.2 $I := \max(0;\, I + P - PE)$

   1.3 $P_{th} = 0$ if $I \leq I_{max}$ / $P_{th} = I - I_{max}$ if $I > I_{max}$

   1.4 $I := I - P_{th}$

Step 2: Subtraction of PE-induced losses

$$AE_s = S * \frac{(2 - \frac{S}{S_{max}})\tanh\left(\frac{PE - AE_i}{S_{max}}\right)}{1 + (1 - \frac{S}{S_{max}})\tanh\left(\frac{PE - AE_i}{S_{max}}\right)} \quad ; \qquad S := S - AE_s$$

Step 3: Feeding the SMAR

$$P_s = S_{max} * \frac{(1 - \left(\frac{S}{S_{max}}\right)^2)\tanh\left(\frac{P_{th}}{S_{max}}\right)}{1 + \frac{S}{S_{max}}\tanh\left(\frac{P_{th}}{S_{max}}\right)} \quad ; \qquad S := S + P_s$$

Step 4: Subtraction of percolation

$$Perc = S * \left(1 - \left(1 + \left(\frac{4}{21}\frac{S}{S_{max}}\right)^4\right)^{-0.25}\right) \quad ; \quad S := S - Perc$$

Figure 7. Description of the soil moisture accounting reservoir ($SMAR$) equations. All variables are in millimeters. The interception reservoir maximum capacity I_{max} was estimated by matching the flux of throughfall at the hourly and daily time steps (Ficchì, Perrin, and Andréassian, 2019). Daily throughfall was estimated by comparing daily P with PE (Perrin, Michel, and Andréassian, 2003). For the 852 catchments used in this study, I_{max} was between 2.00 mm and 4.00 mm with a median value of 2.75 mm. I is initiated at 0 mm whereas S is initiated at $0.5 \cdot S_{max}$.

References

Addor, N., *et al.* 2018. "A ranking of hydrological signatures based on their predictability in space". *Water Resources Research* 54 (11): 8792–8812. ISSN: 1944-7973. doi:10.1029/2018WR022606. Cited on pages 2, 14.

Alley, W. M., and J. E. Veenhuis. 1983. "Effective impervious area in urban runoff modeling". *Journal of Hydraulic Engineering* 109 (2): 313–319. doi:10.1061/(ASCE)0733-9429(1983)109:2(313). Cited on page 23.

Anctil, F., *et al.* 2004. "A soil moisture index as an auxiliary ANN input for stream flow forecasting". *Journal of Hydrology* 286 (1): 155–167. ISSN: 0022-1694. doi:10.1016/j.jhydrol.2003.09.006. Cited on pages 10, 21.

Ando, Y., K. Musiake, and Y. Takahasi. 1984. "Modelling of hydrologic processes in a small urbanized hillslope basin with comments on the effects of urbanization". *Journal of Hydrology*, Global Water: Science and Engineering The Ven Te Chow Memorial Volume, 68 (1): 61–83. ISSN: 0022-1694. doi:10.1016/0022-1694(84)90204-X. Cited on page 2.

Baker, D. B., *et al.* 2004. "A new flashiness index: Characteristics and applications to midwestern rivers and streams". *JAWRA Journal of the American Water Resources Association* 40 (2): 503–522. ISSN: 1752-1688. doi:10.1111/j.1752-1688.2004.tb01046.x. Cited on pages 6, 7.

Beck, H. E., *et al.* 2018. "Present and future Köppen-Geiger climate classification maps at 1-km resolution". *Scientific Data* 5 (1): 1–12. ISSN: 2052-4463. doi:10.1038/sdata.2018.214. Cited on page 3.

Berghuijs, W. R., *et al.* 2016. "Dominant flood generating mechanisms across the United States". *Geophysical Research Letters* 43 (9): 4382–4390. ISSN: 1944-8007. doi:10.1002/2016GL068070. Cited on page 2.

Berthet, L., *et al.* 2009. "How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments". *Hydrology and Earth System Sciences* 13 (6): 819–831. ISSN: 1027-5606. doi:10.5194/hess-13-819-2009. Cited on page 10.

Berthier, E., H. Andrieu, and J. D. Creutin. 2004. "The role of soil in the generation of urban runoff: development and evaluation of a 2D model". *Journal of Hydrology*, Urban Hydrology, 299 (3): 252–266. ISSN: 0022-1694. doi:10.1016/j.jhydrol.2004.08.008. Cited on page 23.

Blume, T., E. Zehe, and A. Bronstert. 2007. "Rainfall—runoff response, event-based runoff coefficients and hydrograph separation". *Hydrological Sciences Journal* 52 (5): 843–862. ISSN: 0262-6667. doi:10.1623/hysj.52.5.843. Cited on pages 2, 10, 23, 25.

Bocinsky, R. K., D. Beaudette, and S. Chamberlain. 2019. *FedData: Functions to automate downloading geospatial data available from several federated data sources*. Visited on 05/21/2019. https://CRAN.R-project.org/package=FedData. Cited on page 8.

Brabec, E., S. Schulte, and P. L. Richards. 2002. "Impervious surfaces and water quality: A review of current literature and its implications for watershed planning". *Journal of Planning Literature* 16 (4): 499–514. ISSN: 0885-4122. doi:10.1177/088541202400903563. Cited on page 23.

Breiman, L. 2001. "Random Forests". *Machine Learning* 45 (1): 5–32. ISSN: 1573-0565. doi:10.1023/A:1010933404324. Cited on page 14.

Brun, S. E., and L. E. Band. 2000. "Simulating runoff behavior in an urbanizing watershed". *Computers, Environment and Urban Systems* 24 (1): 5–22. ISSN: 0198-9715. doi:10.1016/S0198-9715(99)00040-X. Cited on pages 2, 21.

Cappus, P. 1960. "Bassin expérimental d'Alrance. Étude des lois de l'écoulement. Application au calcul et la prévision des débits". *La Houille Blanche*: 493–520. ISSN: 0018-6368, 1958-5551. doi:10.1051/lhb/1960007. Cited on page 10.

Cicco, L. A. D., *et al.* 2018. *dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services*. Reston, VA: U.S. Geological Survey. doi:10.5066/P9X4L3GE. Cited on page 3.

Collischonn, W., and F. M. Fan. 2013. "Defining parameters for Eckhardt's digital baseflow filter". *Hydrological Processes* 27 (18): 2614–2622. ISSN: 1099-1085. doi:10.1002/hyp.9391. Cited on pages 10, 25.

Debbage, N., and J. M. Shepherd. 2018. "The influence of urban development patterns on streamflow characteristics in the Charlanta Megaregion". *Water Resources Research* 54 (5): 3728–3747. ISSN: 1944-7973. doi:10.1029/2017WR021594. Cited on page 23.

Ebrahimian, A., B. N. Wilson, and J. S. Gulliver. 2016. "Improved methods to estimate the effective impervious area in urban catchments using rainfall-runoff data". *Journal of Hydrology* 536:109–118. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2016.02.023`. Cited on page 23.

Eckhardt, K. 2005. "How to construct recursive digital filters for baseflow separation". *Hydrological Processes* 19 (2): 507–515. ISSN: 1099-1085. doi:`10.1002/hyp.5675`. Cited on pages 10, 25.

Falcone, J. A. 2011. *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow*. USGS Unnumbered Series. Reston, VA: U.S. Geological Survey. Visited on 07/23/2019. `http://pubs.er.usgs.gov/publication/70046617`. Cited on pages 3, 7, 23.

FAO/IIASA/ISRIC/ISS-CAS/JRC. 2012. *Harmonized world soil database*. Tech. rep. FAO, Rome, Italy and IIASA, Laxenburg, Austria. Cited on page 8.

Ficchì, A., C. Perrin, and V. Andréassian. 2019. "Hydrological modelling at multiple sub-daily time steps: Model improvement via flux-matching". *Journal of Hydrology* 575:1308–1327. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2019.05.084`. Cited on pages 10, 28, 29.

Gao, H., *et al.* 2018. "Landscape heterogeneity and hydrological processes: a review of landscape-based hydrological models". *Landscape Ecology* 33 (9): 1461–1480. ISSN: 1572-9761. doi:`10.1007/s10980-018-0690-4`. Cited on page 23.

Gleeson, T., *et al.* 2014. "A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity". *Geophysical Research Letters* 41 (11): 3891–3898. ISSN: 0094-8276. doi:`10.1002/2014GL059856`. Cited on page 9.

Graeff, T., *et al.* 2012. "Predicting event response in a nested catchment with generalized linear models and a distributed watershed model". *Hydrological Processes* 26 (24): 3749–3769. ISSN: 1099-1085. doi:`10.1002/hyp.8463`. Cited on page 2.

Heggen, R. J. 2001. "Normalized Antecedent Precipitation Index". *Journal of Hydrologic Engineering* 6 (5): 377–381. ISSN: 1084-0699. doi:`10.1061/(ASCE)1084-0699(2001)6:5(377)`. Cited on page 10.

Hewlett, J. D., J. C. Fortson, and G. B. Cunningham. 1977. "The effect of rainfall intensity on storm flow and peak discharge from forest land". *Water Resources Research* 13 (2): 259–266. ISSN: 1944-7973. doi:`10.1029/WR013i002p00259`. Cited on pages 2, 23.

Holko, L., *et al.* 2011. "Flashiness of mountain streams in Slovakia and Austria". *Journal of Hydrology* 405 (3): 392–401. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2011.05.038`. Cited on pages 6, 7.

Homer, C., *et al.* 2015. "Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a decade of land cover change information". *Photogrammetric Engineering & Remote Sensing* 81 (5): 345–354. ISSN: 0099-1112. doi:`10.1016/S0099-1112(15)30100-2`. Cited on page 5.

Isik, S., *et al.* 2013. "Modeling effects of changing land use/cover on daily streamflow: An Artificial Neural Network and curve number based hybrid approach". *Journal of Hydrology*, Hydrology of peri-urban catchments: processes and modelling, 485:103–112. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2012.08.032`. Cited on page 10.

Javelle, P., *et al.* 2010. "Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations". *Journal of Hydrology*, Flash Floods: Observations and Analysis of Hydrometeorological Controls, 394 (1): 267–274. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2010.03.032`. Cited on page 10.

Kirchner, J. W. 2009. "Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward". *Water Resources Research* 45 (2). ISSN: 1944-7973. doi:`10.1029/2008WR006912`. Cited on page 10.

Kohler, M. A., and R. K. Linsley. 1951. *Predicting the runoff from storm rainfall*. Weather Bureau Research Papers. U.S. Department of Commerce, Weather Bureau. Cited on page 10.

Kursa, M. B., and W. R. Rudnicki. 2020. *Boruta: Wrapper algorithm for all relevant feature selection*. Visited on 10/06/2020. `https://CRAN.R-project.org/package=Boruta`. Cited on page 14.

— . 2010. "Feature selection with the Boruta package". *Journal of Statistical Software* 36 (1): 1–13. ISSN: 1548-7660. doi:`10.18637/jss.v036.i11`. Cited on page 14.

Liaw, A., and M. Wiener. 2002. "Classification and regression by randomForest". *R News* 2 (3): 18–22. ISSN: 1609-3631. Cited on page 14.

Lin, Y., and K. E. Mitchell. 2005. "The NCEP stage II/IV hourly precipitation analyses: Development and applications". In *Proceedings of the 19th conference on Hydrology*, 9–13. Citeseer. Cited on page 3.

McGrane, S. J. 2016. "Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review". *Hydrological Sciences Journal* 61 (13): 2295–2311. ISSN: 0262-6667. doi:`10.1080/02626667.2015.1128084`. Cited on page 21.

Mei, Y., and E. N. Anagnostou. 2015. "A hydrograph separation method based on information from rainfall and runoff records". *Journal of Hydrology* 523:636–649. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2015.01.083`. Cited on pages 10, 24, 25.

Mejía, A. I., and G. E. Moglen. 2010. "Impact of the spatial distribution of imperviousness on the hydrologic response of an urbanizing basin". *Hydrological Processes* 24 (23): 3359–3373. ISSN: 1099-1085. doi:`10.1002/hyp.7755`. Cited on page 23.

Merz, R., G. Blöschl, and J. Parajka. 2006. "Spatio-temporal variability of event runoff coefficients". *Journal of Hydrology* 331 (3): 591–604. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2006.06.008`. Cited on page 10.

Merz, R., and G. Blöschl. 2009. "A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria". *Water Resources Research* 45 (1). ISSN: 1944-7973. doi:`10.1029/2008WR007163`. Cited on pages 2, 10, 21, 23, 24.

Miller, J. D., and T. Hess. 2017. "Urbanisation impacts on storm runoff along a rural-urban gradient". *Journal of Hydrology* 552:474–489. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2017.06.025`. Cited on pages 2, 3, 21, 23.

Norbiato, D., *et al.* 2009. "Controls on event runoff coefficients in the eastern Italian Alps". *Journal of Hydrology* 375 (3): 312–325. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2009.06.044`. Cited on pages 10, 24.

Oudin, L., *et al.* 2018. "Hydrological impacts of urbanization at the catchment scale". *Journal of Hydrology* 559:774–786. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2018.02.064`. Cited on page 23.

Oudin, L., *et al.* 2005. "Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling". *Journal of Hydrology* 303 (1): 290–306. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2004.08.026`. Cited on page 7.

Perrin, C., C. Michel, and V. Andréassian. 2003. "Improvement of a parsimonious model for streamflow simulation". *Journal of Hydrology* 279 (1): 275–289. ISSN: 0022-1694. doi:`10.1016/S0022-1694(03)00225-7`. Cited on page 29.

Price, K. 2011. "Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions: A review". *Progress in Physical Geography: Earth and Environment* 35 (4): 465–492. ISSN: 0309-1333. doi:`10.1177/0309133311402714`. Cited on page 2.

Ragab, R., *et al.* 2003. "Experimental study of water fluxes in a residential area: 2. Road infiltration, runoff and evaporation". *Hydrological Processes* 17 (12): 2423–2437. ISSN: 1099-1085. doi:`10.1002/hyp.1251`. Cited on page 23.

Ramier, D., E. Berthier, and H. Andrieu. 2011. "The hydrological behaviour of urban streets: long-term observations and modelling of runoff losses and rainfall–runoff transformation". *Hydrological Processes* 25 (14): 2161–2178. ISSN: 1099-1085. doi:`10.1002/hyp.7968`. Cited on page 23.

Read, J. S., *et al.* 2015. "geoknife: Reproducible web-processing of large gridded datasets". *Ecography*. doi:`10.1111/ecog.01880`. Cited on pages 3, 6.

Redfern, T. W., *et al.* 2016. "Current understanding of hydrological processes on common urban surfaces". *Progress in Physical Geography: Earth and Environment* 40 (5): 699–713. ISSN: 0309-1333. doi:`10.1177/0309133316652819`. Cited on page 23.

Saadi, M., L. Oudin, and P. Ribstein. 2019. "Random forest ability in regionalizing hourly hydrological model parameters". *Water* 11 (8): 1540. doi:`10.3390/w11081540`. Cited on page 3.

Salvadore, E., J. Bronders, and O. Batelaan. 2015. "Hydrological modelling of urbanized catchments: A review and future directions". *Journal of Hydrology* 529:62–81. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2015.06.028`. Cited on page 23.

Sillanpää, N., and H. Koivusalo. 2015. "Impacts of urban development on runoff event characteristics and unit hydrographs across warm and cold seasons in high latitudes". *Journal of Hydrology* 521:328–340. ISSN: 0022-1694. doi:`10.1016/j.jhydrol.2014.12.008`. Cited on pages 21, 24.

Smith, B. K., *et al.* 2013. "Spectrum of storm event hydrologic response in urban watersheds". *Water Resources Research* 49 (5): 2649–2663. ISSN: 1944-7973. doi:`10.1002/wrcr.20223`. Cited on pages 2, 21.

Soulis, K. X., *et al.* 2009. "Investigation of the direct runoff generation mechanism for the analysis of the SCS-CN method applicability to a partial area experimental watershed". *Hydrology and Earth System Sciences* 13 (5): 605–615. ISSN: 1027-5606. doi:`10.5194/hess-13-605-2009`. Cited on page 10.

Survey, U. G. 1994. *Nationwide summary of US Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993*. Water-Resources Investigations Report 94-4002. Reston, VA. Visited on 10/16/2020. doi:`10.3133/wri944002`. `https://pubs.er.usgs.gov/publication/wri944002`. Cited on page 2.

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. ISSN: 1944-7973. doi:`10.1029/2018WR022587`. Cited on pages 2, 10, 23, 24, 25.

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. ISSN: 1944-7973. doi:`10.1029/2018WR022588`. Cited on pages 2, 23.

Thornton, P. E., *et al.* 2016. "Daymet: Daily surface weather data on a 1-km grid for North America, version 3". *ORNL DAAC*. doi:`10.3334/ORNLDAAC/1328`. Cited on pages 5, 7.

Verdin, K. L. 2017. *Hydrologic Derivatives for Modeling and Analysis—A new global high-resolution database*. USGS Numbered Series 1053. Reston, VA: U.S. Geological Survey. Visited on 10/25/2018. `http://pubs.er.usgs.gov/publication/ds1053`. Cited on page 8.

Vörösmarty, C. J., *et al.* 2010. "Global threats to human water security and river biodiversity". *Nature* 467 (7315): 555–561. ISSN: 1476-4687. doi:`10.1038/nature09440`. Cited on page 2.

Yang, G., *et al.* 2011. "The impact of urban development on hydrologic regime from catchment to basin scales". *Landscape and Urban Planning* 103 (2): 237–247. ISSN: 0169-2046. doi:`10.1016/j.landurbplan.2011.08.003`. Cited on page 21.

Zhang, Y., and W. Shuster. 2014. "Impacts of spatial distribution of impervious areas on runoff response of hillslope catchments: Simulation study". *Journal of Hydrologic Engineering* 19 (6): 1089–1100. ISSN: 1943-5584. doi:`10.1061/(ASCE)HE.1943-5584.0000905`. Cited on pages 2, 21.

Zhang, Y., *et al.* 2018. "Predicting runoff signatures osing regression and hydrological modeling approaches". *Water Resources Research* 54 (10): 7859–7878. ISSN: 1944-7973. doi:`10.1029/2018WR023325`. Cited on page 14.

Zhou, Z., *et al.* 2017. "The complexities of urban flood response: Flood frequency analyses for the Charlotte metropolitan region". *Water Resources Research* 53 (8): 7401–7425. ISSN: 1944-7973. doi:`10.1002/2016WR019997`. Cited on pages 2, 21.