J. Hydrol. Hydromech., 68, 2020, 2, 155–169
DOI: 10.2478/johh-2020-0008
ISSN 1338-4333



sciendo Journal of Hydrology and Hydromechanics

# Controls on event runoff coefficients and recession coefficients for different runoff generation mechanisms identified by three regression methods

Xiaofei Chen[1*], Juraj Parajka[1, 2], Borbála Széles[1], Peter Strauss[3], Günter Blöschl[1, 2]

[1] TU Wien, Centre for Water Resource Systems, Karlsplatz 13, A-1040, Vienna, Austria. www.waterresources.at
[2] TU Wien, Institute of Hydraulic Engineering and Water Resources Management, Karlsplatz 13, A-1040 Vienna, Austria.
[3] Federal Agency for Water Management, Institute for Land and Water Management Research, A-3252 Petzenkirchen, Austria.
[*] Corresponding author. E-mail: chen@waterresources.at

**Abstract:** The event runoff coefficient (*Rc*) and the recession coefficient (*tc*) are of theoretical importance for understanding catchment response and of practical importance in hydrological design. We analyse 57 event periods in the period 2013 to 2015 in the 66 ha Austrian Hydrological Open Air Laboratory (HOAL), where the seven subcatchments are stratified by runoff generation types into wetlands, tile drainage and natural drainage. Three machine learning algorithms (Random forest (RF), Gradient Boost Decision Tree (GBDT) and Support vector machine (SVM)) are used to estimate *Rc* and *tc* from 22 event based explanatory variables representing precipitation, soil moisture, groundwater level and season. The model performance of the SVM algorithm in estimating *Rc* and *tc* is generally higher than that of the other two methods, measured by the coefficient of determination $R^2$, and the performance for *Rc* is higher than that for *tc*. The relative importance of the explanatory variables for the predictions, assessed by a heatmap, suggests that *Rc* of the tile drainage systems is more strongly controlled by the weather conditions than by the catchment state, while the opposite is true for natural drainage systems. Overall, model performance strongly depends on the runoff generation type.

**Keywords:** Machine learning; Event runoff analyses; Event runoff coefficient; Recession coefficient; Runoff generation.

## INTRODUCTION

The event runoff coefficient and the recession coefficient are important characteristics of hydrologic response at the event scale. Understanding their controls and their variability is essential for predicting runoff in ungauged basins and for extrapolating hydrologic response to extreme events (Blöschl et al., 2013; Sivapalan, 2003; Tarasova et al., 2018ab; Viglione et al., 2009).

The event runoff coefficient *Rc* is defined as the portion of rainfall that becomes direct runoff during an event (Merz et al., 2006). The spatial-temporal variability of *Rc* has been widely studied (e.g. Hayes and Young, 2006; Longobardi et al., 2003; Merz et al., 2006; Merz and Blöschl, 2009; Wainwright and Parsons, 2002). Previous studies show that the magnitude of *Rc* varies between the regions. While regional assessments of meso-scale catchments in Austria or Germany (Merz et al., 2006, Tarasova et al., 2018ab) indicate the median of Rc between 0.18 and 0.43, Rc in small agricultural catchments tends to be lower and varies between 0.03 and 0.10 (Blume et al., 2007; Tachecí et al., 2013) to more than 0.2 over cropland hillslopes in central Iowa (Chen et al., 2019). The controls on *Rc* generally depend on the runoff mechanisms. Precipitation intensity tends to be most important when the infiltration excess mechanism dominates, precipitation depth when the saturation excess mechanism dominates, and soil moisture is important for all mechanisms (Tian et al., 2012). Rodríguez-Blanco et al. (2012) and Palleiro et al. (2014) found that the event runoff coefficient in forested catchments in northwestern Spain depended both on the soil moisture at the start of the event and on rainfall depth, whereas rainfall intensity was less important. They explained this finding by the dominant role of subsurface stormflow in event runoff generation. Based on a comparative study in the eastern Italian Alps, Norbiato et al. (2009) suggested that the effect of antecedent soil moisture on the event runoff coefficient might be largest in catchments with intermediate storage capacities. Besides climate and hydrological conditions, Norbiato et al. (2009) also verified that the 'permeability index' deduced from geology is another considerable control on event runoff coefficient in regions with mean annual precipitation less than 1200 mm. Gottschalk and Weingartner (1998) showed that moderate slopes and low geological permeability in the Swiss midlands basins generally lead to events with low event runoff coefficients.

The recession coefficient, *tc*, is the parameter in a linear function of $\frac{-dQ}{dt} = \frac{1}{tc}Q$ where $\frac{-dQ}{dt}$ is the time derivative of runoff $Q$ (Brutsaert and Nieber, 1977; Tallaksen, 1995). Similarly to the event runoff coefficient, the recession parameter has been widely studied (Biswal and Kumar, 2014; Krakauer and Temimi, 2011; Merz et al., 2006; Patnaik et al., 2015). Krakauer and Temimi (2011) identified soil infiltration capacity and forest cover as important controls on the recession coefficient in small catchments in the United States. Tague and Grant (2004) and Gaál et al. (2012) highlighted the important role of geology for the recession coefficient.

One of the challenges in identifying the controls and predicting *Rc* and *tc* is the typically non-linear nature of the relationships between these two parameters and their controls (Merz and Blöschl, 2009; Krakauer and Temimi, 2011). It is therefore of advantage to use non-linear rather than linear analyses and predictive methods. Machine learning techniques based on regressions are able to capture the non-linearity in the relationship between predictor and predictand (Cánovas-García et al., 2017; Erdal and Karakurt, 2013; Naghibi et al., 2017; Şen and Altunkaynak, 2006). Three widely used methods are random forests (RF), Gradient Boost Decision Trees (GBDT) and Support vector machines (SVM).

RF (Ho, 1995) are a learning method for classification and regression that consists of a number of decision trees at training time and identifies the class that has the greatest number of votes (classification) or the mean prediction (regression) of the individual trees. Random forests have been applied in numerous surface and subsurface hydrological studies (Baudron et al., 2013; Naghibi et al., 2017; Zimmermann et al., 2014).

GBDT make use of the gradient boosting framework to combine decision trees based on the RF algorithm. GBDT are usually composed of hundreds of decision trees with shallow depth. In this algorithm, every decision tree adjusts and modifies the predicted value, finally resulting in the prediction. The trees are trained sequentially, which improves the prediction accuracy but involves a longer training time (Friedman, 2001, 2002). Naghibi et al. (2016) compared boosted regression trees (BRT), classification and regression trees (CART), and RF in producing groundwater spring potential maps according to thirteen hydrological-geological-physiographical factors (Naghibi et al., 2016). Sachdeva et al. (2018) constructed wildfire susceptibility maps by combining evolutionary optimized gradient boosted decision trees, and showed that they outperformed other machine learning models.

SVM are supervised learning models with associated learning algorithms. Given a set of training examples, the training algorithm builds a model that assigns new examples to one of two categories (Basak et al., 2007; Ben-Hur and Weston, 2010; Vapnik et al., 1997). SVM have a number of advantages: they effectively solve the classification and regression problem for high dimensions; they solve various nonlinear classification and regression problems by means of different kernel functions; they are able to generalize well. However, computational costs are very high when the dimension of the mapping kernel function is large, and they are sensitive to missing data. They have been applied to surface flow, evaporation, droughts, soil moisture and groundwater prediction (Asefa et al., 2006; Deka, 2014; Hwang et al., 2012; Maity et al., 2010).

The aim of this paper is (a) to identify factors which control variability of event runoff coefficient (*Rc*) and the recession coefficient (*tc*) in small agricultural catchment; (b) to evaluate the relative importance of the control in different subcatchments representing different runoff generation mechanisms and (c) to compare three regression based machine learning techniques, random forests (RF), Gradient Boost Decision Trees (GBDT) and Support vector machines (SVM), in terms of their ability to estimate the event runoff coefficient and the recession coefficient from their controls.

## STUDY CATCHMENT AND DATA

The study is performed in the Hydrological Open Air Laboratory (HOAL). This is a small experimental catchment situated in Lower Austria (Blöschl et al., 2016; Exner-Kittridge, et al., 2016; Széles et al., 2018) (Figure 1). The land use is mainly agricultural (82%) and part of the catchment is tile drained. Runoff is recorded at a number of flumes within the catchment at a time step of 1 minute. In this study, eight flumes are used (Table 1). MW is the catchment outlet of the HOAL and has a catchment area of 65.8 ha. All the other flumes are nested in this catchment (Figure 1). The runoff generation mechanisms of the catchments drained by the flumes differ. Sys4 represents a piped stream and is considered as "natural subsurface drainage" here. Frau1 and Frau2 are tile drained. A1 and A2 drain wetland areas. Sys2, Sys3 and Sys4 are classified as natural drainage here, although (Széles et al., 2018) classified them as tile drains. The reason for the different classification is that the drainage areas of these systems are not fully covered by underground tile pipes and most of the flow is drained from subsurface without tile pipes. All discharge data are processed to remove outliers caused by instrument malfunction and maintenance and aggregated to an hourly time step. Catchment boundaries are identified by analysing a digital elevation model (DEM) and account for the position of tile pipes (Széles et al., 2018). These are used for estimating specific runoff as the ratio of runoff and catchment area and for estimating catchment precipitation (Table 1).

**Table 1.** Runoff generation mechanism and estimated drainage area of the gauged catchments in the HOAL. Figure 1 shows locations of the gauges. From (Széles et al., 2018).

| Gauge | Runoff generation mechanism | Estimated drainage area (ha) | Tile pipe covered area percentage (%) |
|---|---|---|---|
| A1 | Wetland | 2.1 | 6.5 |
| A2 | Wetland | 1.1 | 11.0 |
| Frau1 | Tile drain | 3.1 | 96.6 |
| Frau2 | Tile drain | 4.8 | 60.8 |
| Sys2 | Natural drainage | 2.4 | 8.7 |
| Sys3 | Natural drainage | 4.3 | 10.7 |
| Sys4 | Natural drainage | 37.4 | 4.3 |
| MW | Outlet | 65.8 | 12.9 |

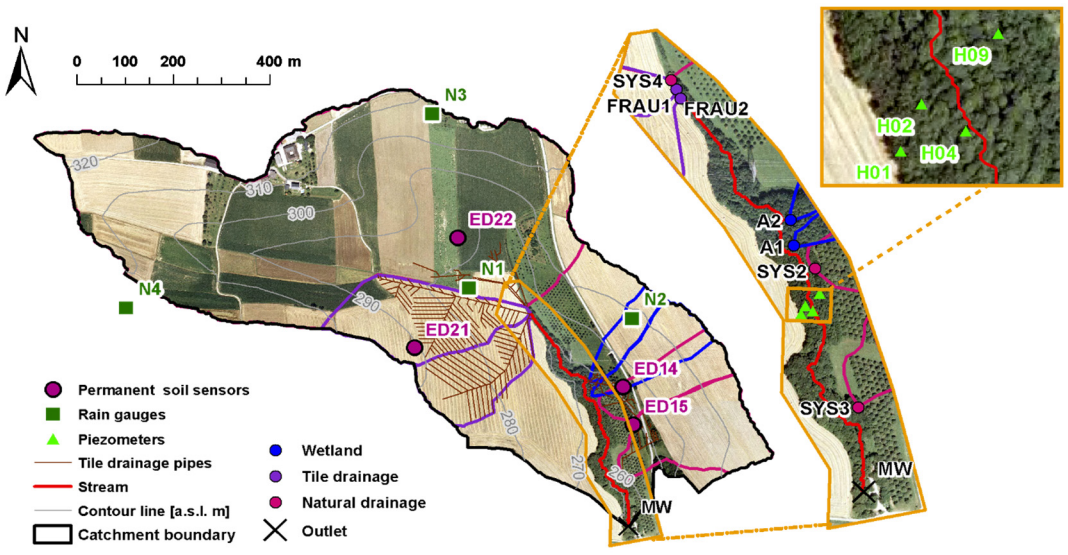


**Fig. 1.** HOAL catchment and its subcatchments. Locations of stream gauges, soil moisture sensors and piezometers used in the analyses.

Additionally, data from four rain gauges (N1-N4), four soil moisture sensors (ED14, ED15, ED21 and ED22) at 5, 10 and 20 cm depth and four groundwater piezometers (H01, H02, H04 and H09) are used (see (Blöschl et al., 2016) for details). Soil moisture data are averaged over depth for extracting soil moisture related variables. Precipitation measurements are averaged by using Thiessen polygon method, and the catchment averages range from 0 to 21.5 mm/h (average: 0.09 mm/h). The higher and smaller precipitation intensities are observed in summer and in winter periods, respectively (Figure 2a). The mean annual precipitation in the study period is decreasing from 937 mm in 2013 to 573 mm in 2015. Comparison of the frequency of daily maximum precipitation from the period 1946 to 2018 indicates that selected study period represents well the frequency of precipitation intensities larger than 30mm/d or 50mm/d. Observed air temperature ranges from –15.2 to 36.3°C during the period from 2013 to 2015 and seasonal dynamics of potential evapotranspiration corresponds to the seasonal dynamics of air

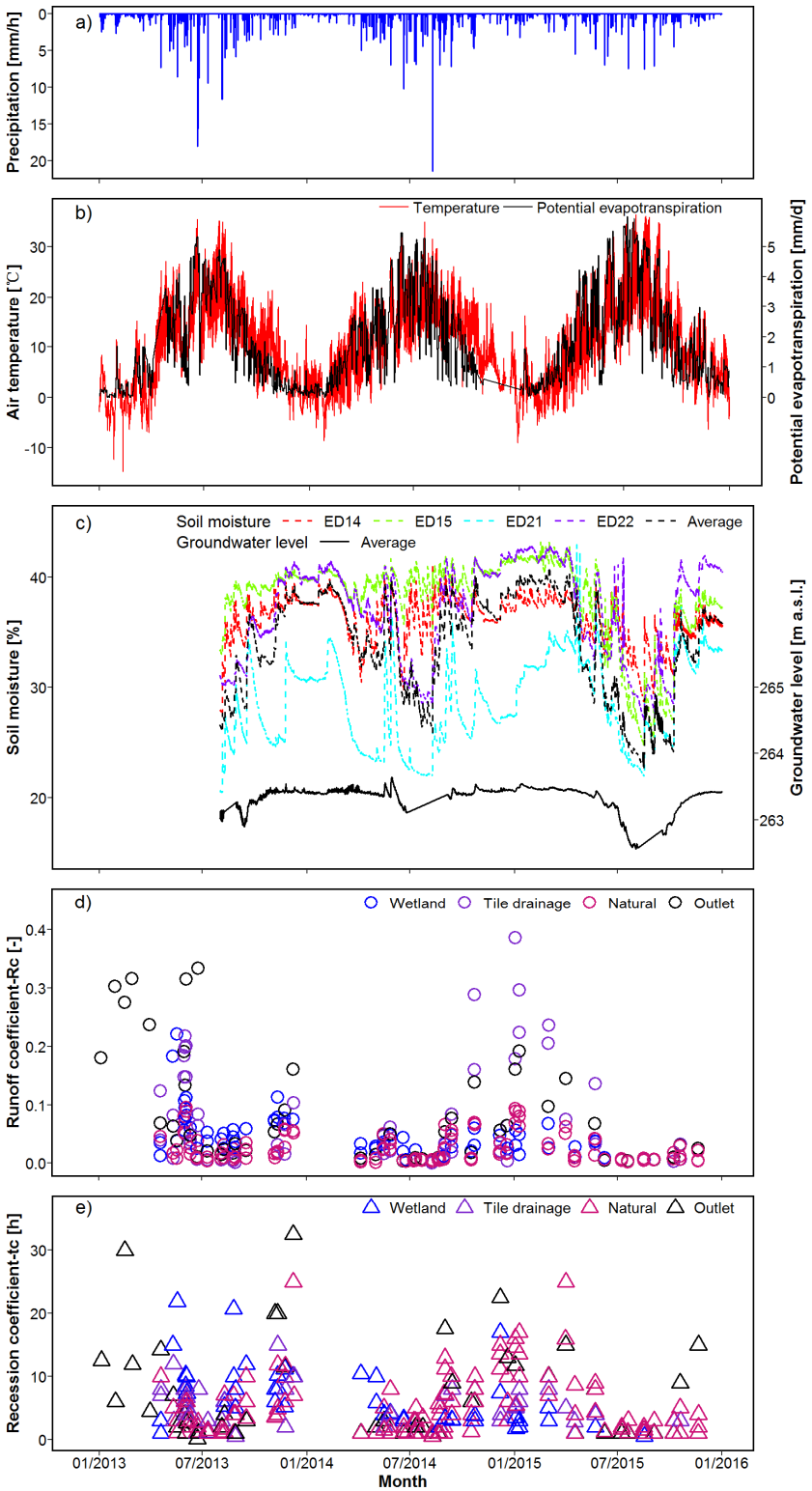


**Fig. 2.** Temporal variability of precipitation (panel a), air temperature and potential evapotranspiration (panel b), soil moisture and groundwater level (panel c), event runoff coefficient (panel d) and recession coefficient (panel e) in HOAL in the period 2013 to 2015.

temperature (Figure 2b). The effect of increased air temperature and potential evaporation in summer is reflected by decreasing soil moisture and groundwater levels (Figure 2c). The seasonal dynamics of event runoff coefficient (*Rc*) and recession coefficient (*tc*) (Figure 2d, e) indicates smaller values of *Rc* and *tc* in summer and and their gradual increase towards winter season.

## METHODS

### Estimation of event runoff coefficients and recession coefficients

The event runoff coefficients and recession coefficients are estimated for each of the eight stream gauges separately using the method of (Merz et al., 2006). The analysis is based on an hourly time step and consists of four steps:

(1) Catchment rainfall estimation: For each of the eight stream gauges, catchment average rainfall is estimated at an hourly time step by spatially interpolating the measurements of four rain gauges using the Thiessen polygon method.

(2) Baseflow separation: Baseflow is estimated for each stream gauge using the Chapman and Maxwell (1996) filter. More details about the filter are given in Merz et al. (2006).

(3) Identification of runoff events: An event peak is identified, if the direct flow is more than double of the baseflow at a certain time step and larger runoff is not observed 5 hours before and after the peak. Around these peaks, the beginning and end times are estimated. An example of an identified event in May of 2014 is shown in Figure 3a. This event is driven by a late spring precipitation event with a peak of 7.0 mm/h, which is larger than 80% of identified events. The event runoff coefficient in a tile drain system Frau2 is about 0.06 higher than at the main catchment outlet MW. Figure 3b shows the dynamics of the ratio of cumulative direct runoff to cumulative precipitation as a function of time during an event for different gauges. The ratio gradually increases and approaches a stable value at the end of the event, which is the *Rc* of that event. There is a large difference between Frau1 and Frau2 due to differences in the controls. A total of 57 event periods are identified at MW outlet (Table 4). At the tributaries slightly fewer are identified, as they do not always respond to rainfall. The number of identified events in individual subcatchments and main outlet are 30 (A1), 38 (A2), 21 (Frau1), 30 (Frau2), 32 (Sys2), 39 (Sys3), 51 (Sys4) and 57 (MW outlet), respectively. This results in a total of 298 event hydrographs from 2013 to 2015 to be further analysed.

(4) Fitting a linear reservoir model: In order to reduce the effect of the selection of the end time of the runoff events, a linear reservoir model is fitted to the direct flow by minimizing the root mean square difference between observed and simulated runoff for each event and stream gauge separately. The event runoff coefficient *Rc* (–) and the recession coefficient *tc* (hrs) are the optimised model parameters.

### Ensemble learning techniques for regression

For the purpose of the study, three learning techniques are used to build regression models. Ensemble learning techniques have been proved in the past to describe and learn various non-linear relationships (Dietterich, 1997). The main advantage of ensemble learning approaches is that they learn in a hierarchical fashion by repeatedly splitting input dataset into separate branches that maximize the information gain of each split. The challenge with their application is, however, that for the selection of an optimal approach, evaluation of performance of different algorithms is recommended (Shen, 2018). Thus the aim of our study is to test three different approaches (i.e random forest, gradient boost decision trees and support vector machines) individually in four different settings represented by four runoff generation systems (wetland, tile drainage, natural and outlet) which are termed as Classified regression model in our study, and in an aggregated system that combines all runoff generation systems termed here as Unclassified regression model.

#### *Random Forests (RF)*

Random forests consist of a number of tree predictors. As the number of trees increases, the mean squared error of out of bag data (*OOBError*) in prediction decreases until it reaches a constant, low level (Ho, 1995). There are two main steps:

(a) Random sampling from the entire database to train a decision tree. The input subsets are different from each other to avoid over-fitting. Out of bag data (OOB) is the remaining subset, which is not used in building a tree. At each node of the trees, the feature that produces the best split in sub-sampling from all features is used for splitting.

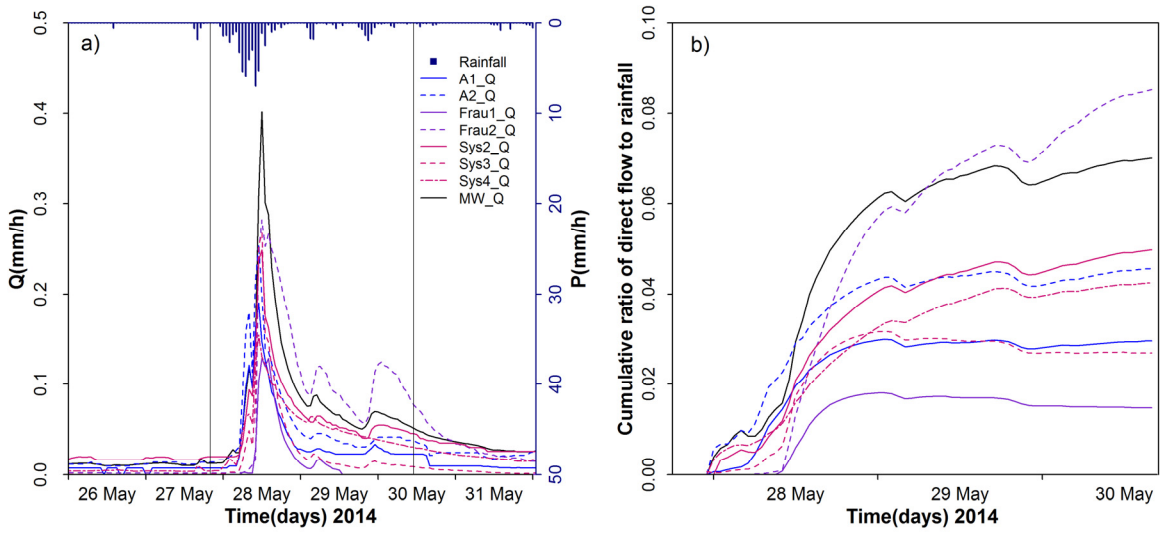


**Fig. 3.** Example event in May 2014. Line colours according to runoff generation mechanisms (wetland runoff is blue; tile drainage is purple; natural discharge is dark pink and outlet flow is black). (a) Hydrographs. Vertical lines indicate start and end of the event. (b) Ratio of cumulative direct flow to cumulative precipitation during the event.

(b) Splitting at each node: Each decision tree is completely split until all samples belong to one class or the leaf node can not be further divided. The feature importance outputs of the RF model with good predictions are used to evaluate the influence of the variables on *Rc* and *tc*.

*Gradient Boost Decision Trees (GBDT)*

A Gradient Boost Decision Tree GBDT regression model is established by integrating multiple decision trees (DTs) in an iterative process (Friedman, 2001, 2002). In every iteration (adding a new tree to the model), model losses are reduced. A Gaussian distribution is chosen as a loss function $L$ for minimizing the squared error,

$$L\left(y, f\left(x\right)\right) = \frac{1}{\sum_i \omega_i} \Sigma \omega_i \left(y_i - f\left(x_i\right)\right)^2 \quad (1)$$

where $\omega_i$ is the weight of sample $i$, $y_i$ are the objectives and $f(x_i)$ are the predictions.

*Support Vector Machines (SVM)*

For a support vector machine (SVM) regression to represent non-linear problems, a proper kernel function needs to be chosen by projecting the data to a high-dimensional space where one can use a linear decision boundary to separate classes (Osuna, 1998). A radial basis function, $K$, is used here which can be expressed for an infinite dimensional space (Horn, 1985) as (2)

$$K\left(u, v\right) = \mathrm{e}^{-\gamma\left|u-v\right|^2} \quad (2)$$

where $u$ and $v$ are two vectors, $\left|u-v\right|^2$ is the squared Euclidean distance of these two vectors, and $\gamma$ influences the number of support vectors. Larger $\gamma$ results in fewer support vectors.

The $\varepsilon$ regression option is adopted here for better controlling model error. Errors beyond the specified $\varepsilon$ are penalized in proportion to $C$, which is the regularization parameter. The $\varepsilon$ insensitive loss function $L_\varepsilon$ proposed by Chapelle and Vapnik (2000) is used according to (3)

$$L_\varepsilon = \begin{cases} 0 & \text{if } \left|y - f\left(x\right)\right| < \varepsilon \\ \left|y - f\left(x\right)\right| - \varepsilon & \text{otherwise} \end{cases} \quad (3)$$

For brevity, only the variable importance of the SVM model, $I_{i,SVM}$ which can reflect the level of variable influences to model objectives (*Rc* and *tc* in this study), are presented in the study. In order to quantify such importance of variables, we take advantage of the Average Absolute Deviation (AAD) from the median in the theory of 1-D sensitivity analysis (Cortez and Embrechts, 2013). $I_{i,SVM}$ is estimated according to the following function

$$\sum_{j=1}^{L} \left|y_j - \tilde{y}\right| / L \quad (4)$$

where $\tilde{y}$ is the baseline of the objectives (median) and $y_j$ is the prediction related to the $j^{\text{th}}$ level of input (totally $L$ = 7 levels).

The detailed procedures of building the SVM regression models are introduced by Cortes and Vapnik (1995).

**Calibration and validation performance of the non-linear regression models**

In order to compare the performance of the three machine learning methods, two types of validations are performed. The first type is the analysis of temporal performance which is evaluated by the coefficients of determination ($R^2$)

$$R^2 = 1 - \frac{\sum_i \left(y_i - f\left(x_i\right)\right)^2}{\sum_i \left(y_i - \bar{y}\right)^2} \quad (5)$$

where $f(x_i)$ is the prediction of *Rc* or *tc* for event $i$, $y_i$ is the observed value, and $\bar{y}$ is the mean of the observed values over all events of a particular runoff generation type. $R^2$ is regarded as 0 when it is negative.

The second type is the validation in space by using leave-one-out cross-validation. Here, the regressions fitted separately to one of the subcatchments representing wetland (A2 station), tile drainage (Frau2 station) and natural drainage (Sys3 and Sys4 stations) are validated in the other stations representing the same runoff generation type, i.e. A1 (wetland), Frau1 (tile drainage), Sys2 (natural drainage). The performance is evaluated by $R^2$, similarly to (5).

## RESULTS

**Evaluation of the observed runoff coefficients and recession time constant and their potential controls**

In order to analyse the potential controls on *Rc* and *tc*, 22 hydrological variables are considered for each stream gauge (Table 2). Event precipitation (VolP), precipitation peak during the event (PeakP) and precipitation duration during the event (DurP) are estimated from the catchment precipitation time series with hourly temporal resolution. Antecedent soil moisture (PreSM), average soil moisture during the event (AverSM), soil moisture at the end of the event (EndSM), soil moisture at the time of peak rainfall (PeakPSM) and soil moisture peak during the event (PeakSM) are estimated from the soil moisture data. Depending on the stream gauge, different soil moisture sensors are used (ED15 for Sys2 and Sys3; ED22 for Sys4; ED21 for Frau1 and Frau2; ED14 for A1 and A2; and the average of all sensors for MW). Soil moisture peak delay to precipitation (DelaySM) is estimated as the time difference between the precipitation peak and the soil moisture peak. Pre-event groundwater level (PreWL), average groundwater level during the event (AverWL), groundwater level at the end of the event (EndWL), groundwater level at the time of peak precipitation (PeakPWL) and peak groundwater level (PeakWL) are estimated from the average groundwater level of all four piezometers, as no separate piezometer data are available for the individual catchments. Groundwater level peak delay to precipitation (DelayWL) is estimated as the time difference between the precipitation peak and the groundwater peak. Potential evapotranspiration (EP) on the day of the event is estimated by the *Penman-Monteith* equation. Additionally, the month of the event occurred (Month), the Normalized Difference Vegetation Index (NDVI) and the runoff generation type (Type: 1-wetland; 2-tile drainage; 3-outlet; 4-natural drainage) are considered as variables. NDVI represents the mean catchment or subcatchment value and it is estimated by linear interpolation between 57 Landsat 7 and 8 scenes available in the period 2013-2015. Finally, drainage area (Area), percentage of piped area (AreaPipes) and forest cover (AreaForest) are also used.

**Table 2.** Event based hydrological variables examined as potential controls on *Rc* and *tc*. While the variables differ between stream gauges, the combined statistics of all stream gauges are shown here.

| Variables | Explanation | Min | 25% | 50% | 75% | Max |
|---|---|---|---|---|---|---|
| VolP | Volume of precipitation during event (mm) | 3.5 | 12.7 | 17.5 | 27.5 | 99.3 |
| PeakP | Precipitation peak during event (mm/hr) | 0.2 | 2.8 | 3.7 | 6.7 | 21.5 |
| DurP | Precipitation duration during event (hrs) | 2.0 | 9.0 | 17.0 | 29.8 | 102.0 |
| PreSM | Antecedent soil moisture (%) | 20.5 | 30.2 | 35.5 | 38.6 | 42.4 |
| PeakSM | Peak of soil moisture during event (%) | 24.9 | 34.2 | 38.2 | 40.0 | 43.2 |
| PeakPSM | Soil moisture at the time of peak precipitation (%) | 20.5 | 30.2 | 36.9 | 39.3 | 43.0 |
| AverSM | Average soil moisture during event (%) | 22.6 | 31.7 | 37.0 | 39.2 | 42.6 |
| EndSM | Soil moisture at the end of event (%) | 23.6 | 33.6 | 37.9 | 39.4 | 42.6 |
| DelaySM | Soil moisture peak delay to precipitation (hrs) (positive value indicates earlier flow peak than soil moisture) | –13.0 | 2.0 | 7.0 | 11.0 | 63.0 |
| PreWL | Pre-event groundwater level (m) | 262.6 | 263.3 | 263.4 | 263.5 | 264.7 |
| PeakWL | Peak groundwater level (m) | 262.7 | 263.3 | 263.5 | 263.5 | 264.9 |
| PeakPWL | Groundwater level at the time of peak precipitation (m) | 262.7 | 263.3 | 263.4 | 263.5 | 264.7 |
| AverWL | Average groundwater level during event (m) | 262.7 | 263.3 | 263.5 | 263.5 | 264.7 |
| EndWL | Groundwater level at the end of event (m) | 262.7 | 263.3 | 263.5 | 263.5 | 264.8 |
| DelayWL | Groundwater level peak delay to precipitation (hrs) (positive value indicates earlier flow peak than groundwater table) | –12.0 | 13.0 | 18.0 | 27.0 | 103.0 |
| EP | Potential evapotranspiration during the day of event (mm/d) | 0.0 | 0.9 | 1.7 | 2.7 | 5.1 |
| Month | Month when event occurred (month) | 1 | 5 | 7 | 9 | 12 |
| NDVI | Normalized Difference Vegetation Index | 0.1 | 0.2 | 0.3 | 0.3 | 0.5 |
| AreaPipes | Piped area relative to drainage area (%) | 4.3 | 6.5 | 10.7 | 12.9 | 96.6 |
| AreaForest | Forest covered area percentage to drainage area (%) | 0.0 | 4.6 | 9.6 | 14.2 | 18.8 |
| Area | Drainage area of subcatchment (ha) | 1.1 | 2.4 | 4.3 | 37.4 | 65.8 |
| Type | Runoff generation type (1-wetland; 2-tile drainage; 3-outlet; 4-natural drainage) | 1 | 2 | 3 | 4 | 4 |

**Table 3.** Statistics of the event runoff coefficients *Rc* and recession coefficients *tc*, including minimum, quartiles and maximum, for the eight stream gauges based on 40 events which can be observed at least at 5 stations.

| | | *Rc* (–) | | | | | *tc* (hrs) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Gauge | Runoff generation mechanism | Min | 25% | 50% | 75% | Max | Min | 25% | 50% | 75% | Max |
| A1 | Wetland | 0.010 | 0.026 | 0.033 | 0.049 | 0.082 | 1.00 | 3.00 | 4.00 | 8.00 | 16.98 |
| A2 | Wetland | 0.006 | 0.028 | 0.048 | 0.068 | 0.222 | 0.50 | 2.00 | 3.89 | 7.57 | 21.90 |
| Frau1 | Tile drain | 0.001 | 0.016 | 0.033 | 0.181 | 0.297 | 0.50 | 3.00 | 4.00 | 6.00 | 10.00 |
| Frau2 | Tile drain | 0.0003 | 0.026 | 0.054 | 0.139 | 0.386 | 1.00 | 3.00 | 5.50 | 8.00 | 15.00 |
| Sys2 | Natural drainage | 0.006 | 0.010 | 0.034 | 0.058 | 0.089 | 1.00 | 4.00 | 5.00 | 7.00 | 17.00 |
| Sys3 | Natural drainage | 0.001 | 0.005 | 0.018 | 0.026 | 0.094 | 0.50 | 1.63 | 4.20 | 7.60 | 14.67 |
| Sys4 | Natural drainage | 0.004 | 0.008 | 0.012 | 0.037 | 0.096 | 1.00 | 1.04 | 3.36 | 8.00 | 25.00 |
| MW | Outlet | 0.004 | 0.014 | 0.048 | 0.071 | 0.316 | 0.10 | 1.00 | 3.00 | 10.00 | 32.56 |
| All gauges | | 0.0003 | 0.012 | 0.032 | 0.063 | 0.386 | 0.10 | 2.00 | 4.00 | 8.00 | 32.56 |

**Table 4.** Number of event hydrographs and number of event periods (in brackets) used for the calibration and validation for the *Rc* and *tc* regressions. The models are fitted for the four runoff generation types (wetland, tile drainage, natural and outlet) termed as Classified regression model, and all together (Unclassified regression model).

| | Number of events for Classified regression model | | | | Number of events for Unclassified regression model |
|---|---|---|---|---|---|
| | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All |
| Calibration | 54 | 40 | 97 | 45 | 252 (40 event periods) |
| Validation | 14 | 11 | 25 | 12 | 46 (17 event periods) |
| Total | 68 | 51 | 122 | 57 | 298 (57 event periods) |

22 hydrological variables and *Rc* and *tc* are normalised to the range between 0 and 1 for all runoff generation systems together.

The statistics and cumulative distributions of the resulting coefficients are respectively showed in Table3 and Figure 4. *Rc* varies from 0.0003 to 0.4. The tile drainage catchment Frau 2 has the highest median (0.054), the natural subsurface drainage Sys4 the lowest (0.012). The tile drainage catchment Frau 2 has the highest median of *tc* (5.5 hrs), the outlet MW the lowest (3.0 hrs) but the 75% quantile is larger than that of the subcatchments (10 hrs). The correlation coefficient of *Rc* and t*c* over all 298 hydrographs is about 0.38. High *Rc* is usually associated with high antecedent flows which may result in slower runoff recession (Patnaik et al., 2015) because of large groundwater contributions (Exner-Kittridge et al., 2016).

**Parameter sensitivities of the three non-linear regression models**

The parameter sensitivities of the regression models are explored based on 252 event hydrographs, when events are observed at a minimum of 5 stations (Figure 5, 6 and 7). Additionally 46 event hydrographs are used for validation (Table 4).

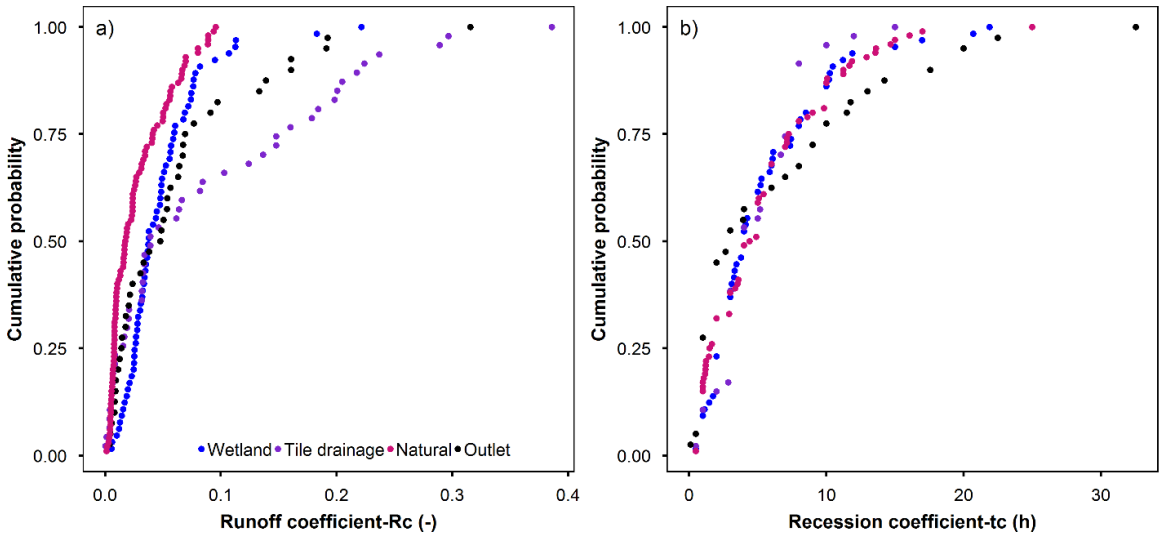


**Fig. 4.** Cumulative distribution of (a) event runoff coefficients and (b) recession coefficients for the four types of runoff generation mechanisms.

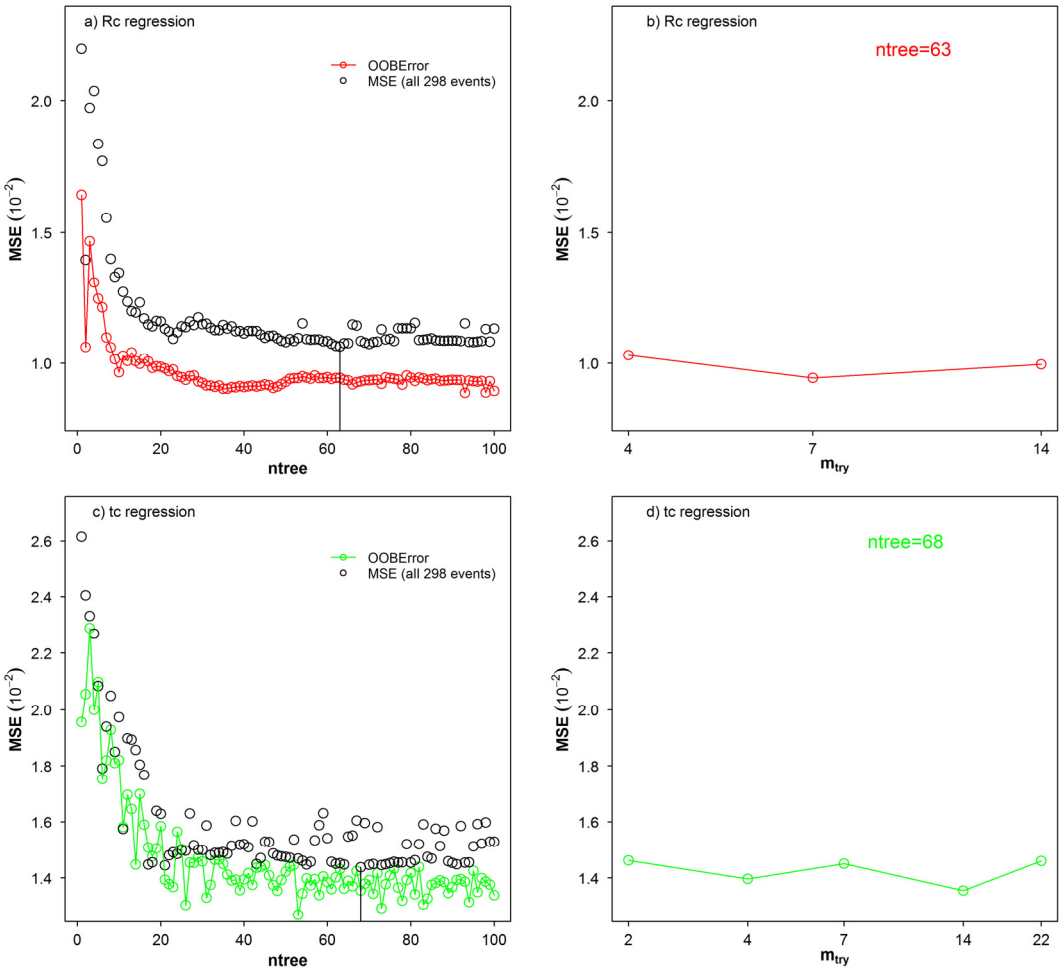


**Fig. 5.** Sensitivity of the model error, MSE, to *ntree* and $m_{try}$ in the RF models for *Rc* and *tc* regressions based on unclassified events (Table 4). The vertical black line indicates the *ntree* chosen, 63 for the *Rc* regression and 68 for the *tc* regression.

*RF model (ntree and $m_{try}$)*

The number of trees in the model, *ntree*, has a controlling function in the model performance (Figure 5 a and c). As *ntree* increases, *OOBError*, i.e. the mean squared error (*MSE*) based on out of bag data (*OOB* data), decreases up to a threshold. *OOBError* is calculated according to (7)

$$OOBError = \frac{\sum_i^M (y'_i - y_i)^2}{M} \quad (7)$$

where $M$ is the number of out of bag data, $y_i$ is $i$th observation value of *OOB* and $y_i^{'}$ is the model prediction of *OOB* based on trees that are not trained by the $i^{th}$ sample.

*OOBError* stabilises at around 0.009 and 0.013 for *Rc* and *tc*, respectively. The model performance would be improved by diversifying regression trees, however, it is difficult to build a new diverse tree if the ensemble size is large. Therefore, *OOBError* tends to approach a constant level (Breiman, 2001).

If the number of variables $m_{try}$ randomly sampled as candidates at each split is close to the total number of variables, regression trees are more likely to be similar, which will reduce the predictive ability. In our case of only 22 variables, the results do not change much by $m_{try}$ (Figure 5 b and d), as suggested by Liaw and Wiener (2002).

Finally, for the unclassified models, *ntree* is set to 63 and 68 for the *Rc* and *tc* regressions, respectively, by minimizing MSE of all 298 events, while $m_{try}$ is automatically optimized. The parameters of the classified models are calibrated in the same way. The final values of *ntree* and $m_{try}$ are listed in Table 5.

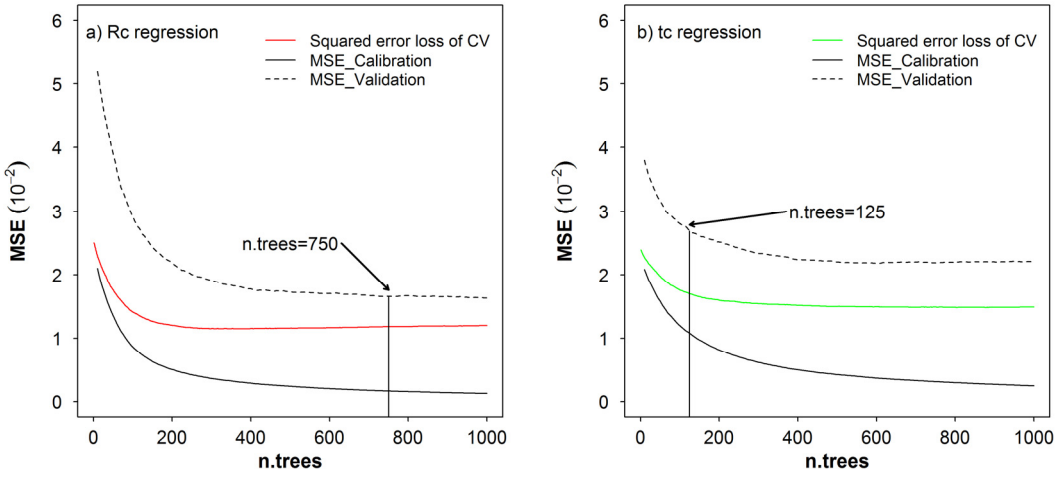


**Fig. 6.** Sensitivity of the model error, MSE, to *n.trees* in the GBDT regression models for *Rc* and *tc* regressions based on unclassified events (Table 4). Vertical black line indicates *n.trees* chosen: 750 for the *Rc* regression and 125 for the *tc* regression.

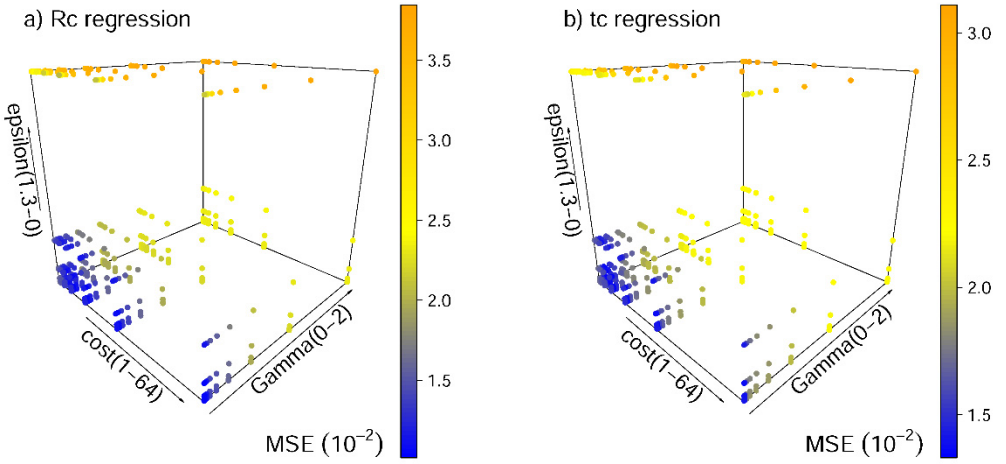


**Fig. 7.** Sensitivity of the model error, MSE, to different combinations of $\gamma$, $\varepsilon$ and $C$ for the SVM models based on unclassified events (Table 4). MSE is the squared error loss of 10 fold cross validation.

**Table 5.** Parameters of *ntree* and $m_{try}$ of the RF models stratified by runoff generation type. The models are fitted for the four runoff generation types (wetland, tile drainage, natural and outlet) termed as Classified regression model, and all together (Unclassified regression model).

| Parameters of Random Forest model | *Rc* | | | | | *tc* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Classified regression model | | | | Unclassified regression model | Classified regression model | | | | Unclassified regression model |
| | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All |
| *Ntree* | 94 | 52 | 28 | 83 | 63 | 86 | 100 | 60 | 58 | 68 |
| $m_{try}$ | 7 | 22 | 4 | 3 | 7 | 4 | 4 | 2 | 4 | 14 |

*GBDT model (n.trees)*

There are four main parameters needed for the GBDT model, *shrinkage, n.minobsinnode, interaction.depth* and *n.trees*. Reducing *shrinkage*, the learning rate, improves the performance while increasing the computational cost, so *shrinkage* is set to a low value of 0.001 (Friedman, 2001, 2002). *n.minobsinnode* is the minimum number of observations in a tree's terminal node, and *interaction.depth* represents the number of splits performed on a tree. Model performance is not sensitive to these parameters, so they were set to 10 and 6, respectively, by minimizing the cross validation errors. *n.trees* is the number of iterations, i.e., the number of trees in the GBDT model. Its sensitivity is analysed in Figure 6. The squared error loss of CV is calculated by averaging MSE of $k = 5$ across folders

$$\text{Squared error loss of CV} = \frac{1}{k}\sum_{i=1}^{k}\frac{\sum_{j=1}^{n}\left(y'_{i,j} - y_{i,j}\right)^2}{n} \quad (8)$$

where $n$ is the number of samples in the $i$ th folder (a total of $k = 5$ folders), $y'_{i,j}$ and $y_{i,j}$ are respectively the model prediction and the objective value of sample $j$ in folder $i$.

The squared error loss of CV in the *Rc* regression is slightly lower than that in the *tc* regression (Figure 6). Increasing *n.trees* provides a greater improvement of performance when *n.trees* is small. The optimum *n.trees* values for the *Rc* and *tc* regressions based on all unclassified events are manually set to *n.trees* = 750 and 125 respectively by minimizing the squared error loss of CV in calibration and by controlling the MSE reduction (e.g. 0.015) between calibration and validation to avoid overfitting. The values of *n.trees* are listed in Table 6.

*SVM model (γ, ε and C)*

When calibrating the Radial Basis Function (RBF) kernel of the SVM models, three parameters are needed, $\gamma$, $\varepsilon$ and $C$. $\gamma$ is the kernel coefficient of the RBF and reducing $\gamma$ will increase the performance and reduce the bias, as the number of support vectors increases. $\varepsilon$ is a parameter in the insensitive-loss function (Eq. 4). Reducing $\varepsilon$ increases the number of support vectors. $C$ is the cost of violating the constraints of the regularization term in the Lagrange formulation. Large $C$ aims at a smaller margin with better prediction but may lead to overfitting.

The squared error loss of 10 fold cross validation (Equation 8 in which $k$ is set to10) and the MSE reduction between calibration and validation is used for optimizing the parameters, thus preventing overfitting. A "grid-search" on $\gamma$, $\varepsilon$ and $C$ is performed using the R tune.svm function. Various triples of values are tried and the one with the best cross-validation accuracy is selected. The test sequences of the parameters are $\gamma = 2^{(-7:2)}$, $\varepsilon = (0.003, 0.01, 0.03, 0.1, 0.3, 1.3)$ and $C = 2^{(0:6)}$ (Hsu et al., 2003). 3D sensitivity plots of the three parameters are shown in Figure 7. Table 7 gives the optimised parameter value for the classified regression models. There are many good combinations of three parameters resulting in the lowest squared error loss of 10 fold cross validation (Figure 7). The final parameter combinations are chosen by both minimizing the squared error loss of CV in the calibration and controlling the MSE reduction.

**Table 6.** Parameter of *n.trees* of the GBDT models stratified by runoff generation type. The models are fitted for the four runoff generation types (wetland, tile drainage, natural and outlet) termed as Classified regression model, and all together (Unclassified regression model).

| Parameters of GBDT model | *Rc* | | | | | *tc* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Classified regression model | | | | Unclassified regression model | Classified regression model | | | | Unclassified regression model |
| | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All |
| *n.trees* | 20 | 665 | 180 | 100 | 750 | 20 | 45 | 35 | 80 | 125 |
| *interaction.depth*=6 *n.minobsinnode* = 10 *shrinkage* = 0.001 *seed*=537 | | | | | | | | | | |

**Table 7.** Parameters of the SVM models stratified by runoff generation type. The models are fitted for the four runoff generation types (wetland, tile drainage, natural and outlet) termed as Classified regression model, and all together (Unclassified regression model).

| Parameters of SVM model | *Rc* | | | | | *tc* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Classified regression model | | | | Unclassified regression model | Classified regression model | | | | Unclassified regression model |
| | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All | Wetland (A1, A2) | Tile drainage (Frau1, Frau2) | Natural (Sys2, Sys3, Sys4) | Outlet (MW) | All |
| $\gamma$ | $2^{-7}$ | $2^{-5}$ | $2^{-7}$ | $2^{-6}$ | $2^{-7}$ | $2^{-6}$ | $2^{-6}$ | $2^{-7}$ | $2^{-5}$ | $2^{-5}$ |
| $\varepsilon$ | 1.3 | 0.1 | 0.1 | 0.03 | 0.01 | 1.3 | 0.1 | 0.3 | 0.1 | 0.03 |
| $C$ | 2 | 2 | $2^5$ | $2^6$ | $2^6$ | 1 | 1 | 1 | 1 | $2^3$ |

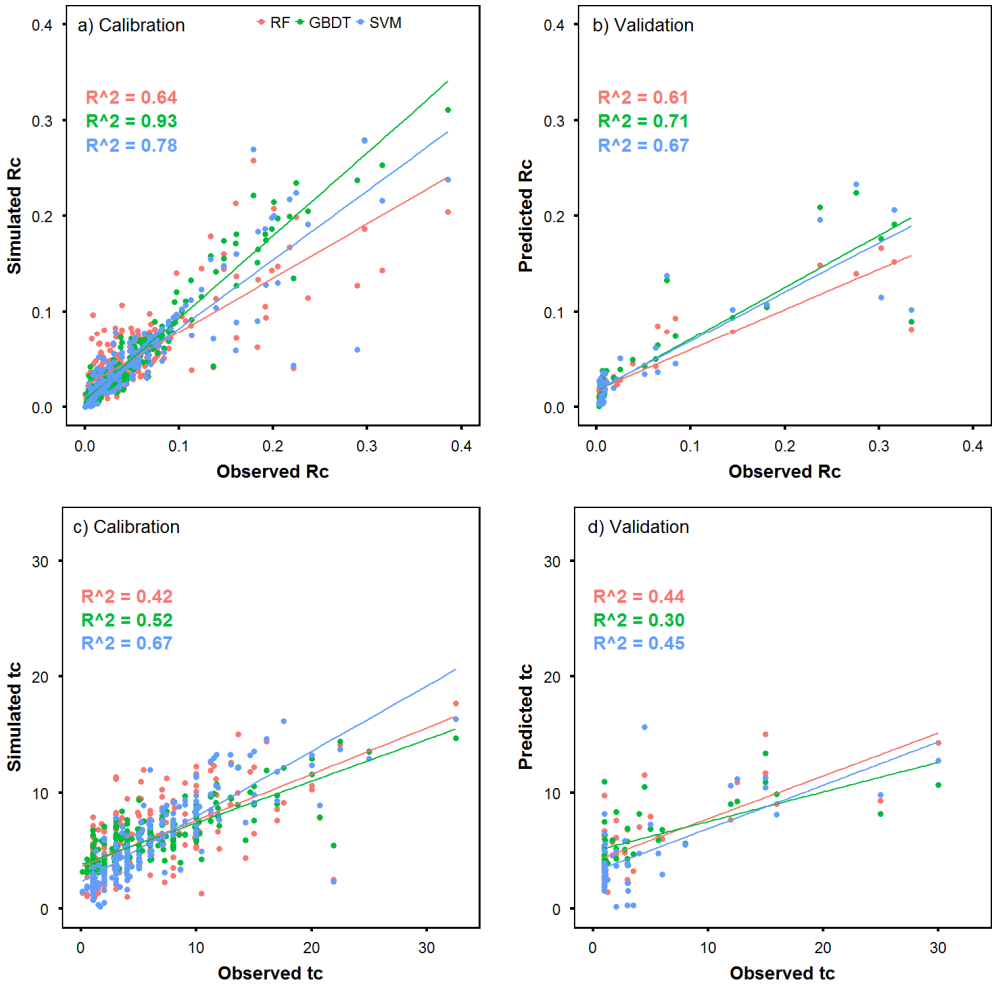


**Fig. 8.** Estimates of the RF, GBDT and SVM regression models plotted against the observations of *Rc* and *tc*. Estimation is based on unclassified events (252 are used for calibration and 46 for validation). Colours indicate the regression models. Lines represent linear relation between estimation and observation. Panels a) and c): calibration. Panels b) and d): validation.

### Temporal calibration and validation performance of the non-linear regression models

Figure 8 shows the estimates of *Rc* and *tc* for the RF, GBDT and SVM models based on unclassified events as in the last column in Table 4. The performances of the *Rc* regressions based on the GBDT and SVM algorithms are better than that of the RF model with $R^2$ of 0.71 and 0.67 in validation, respectively. The performances of the *tc* regressions based on the RF and SVM algorithms are better than that of the GBDT model with $R^2$ of 0.44 and 0.45 in validation, respectively.

Figure 9 compares the calibration and validation performance of the regressions (as in Table 4). The performance of the SVM model is generally higher than that of the others, which may be due to the higher dimensional space of the SVM regression (Asefa et al., 2006). The performance of the *tc* regression models is generally low for all methods. This may be due to the lack of subcatchment groundwater data that might improve the performance, particularly for wetlands and natural drainage runoff generation systems which are in HOAL closely related to groundwater level dynamics. On the other hand, the *Rc* and *tc* regressions of the natural drainage subcatchment and the unclassified events have a somewhat higher performance than the other systems, probably because of the larger number of events that are available. The models highlighted by boxes in the left of the figure have higher values of $R^2$ than 0.6, both in calibration and validation.

### Spatial calibration and validation performance of the non-linear regression models

Table 8 lists the number of events used for calibration and validation in a spatial leave one out mode. The classified regressions of *Rc* and *tc* including wetland, tile drainage and natural subsurface flow, are spatially validated. For the wetland regressions, 38 events from station A2 are used for calibration and 30 events from station A1 for validation; for the tile drainage regressions, 30 events from Frau2 are used for calibration and 21 events from Frau1 for validation; for natural subsurface flow, 90 events in Sys3 and Sys4 are used for calibration and 32 events in Sys2 for validation. Additionally, a total of 252 hydrographs during 40 events that are observed in at least 5 stations, are used for building 8 Leave_one_out models in which the events of one of the subcatchments are used for validation and the rest for calibration. The SVM regression models generally have a better performance (measured by $R^2$) than the other models, which is similar to the temporal validation (Figure 10).

### Linear correlations of *Rc* and *tc* with the explanatory variables

Pearson's correlation coefficient *r* of *Rc* and *tc* with the explanatory variables of Table 2 is evaluated for the entire set of 298 events, i.e. for all stream gauges together (Table 9). *Rc* is positively correlated with DurP, VolP, AreaPipes and ground-

**Table 8.** Number of events used for spatial calibration and validation of the *Rc*/*tc* regressions by runoff generation type. The models are fitted for the four runoff generation types (wetland, tile drainage, natural and outlet) termed as Classified regression model, and all together (Unclassified regression model).

| Number of events | Rc/*tc* regression | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Classified regression model | | | Unclassified regression model | | | | | | | |
| | Wet-land | Tile drainage | Natural | Leave_A1 out | Leave_A2 out | Leave_Frau1_ out | Leave_Frau2_ out | Leave_Sys2_ out | Leave_Sys3_ out | Leave_Sys 4_out | Leave_MW out |
| Calibration | 38 (A2) | 30 (Frau2) | 90 (Sys3, Sys4) | 223 | 216 | 233 | 224 | 223 | 220 | 213 | 212 |
| Validation | 30 (A1) | 21 (Frau1) | 32 (Sys2) | 29 (A1) | 36 (A2) | 19 (Frau1) | 28 (Frau2) | 29 (Sys2) | 32 (Sys3) | 39 (Sys4) | 40 (MW) |
| Total | 68 | 51 | 122 | 252 | 252 | 252 | 252 | 252 | 252 | 252 | 252 |

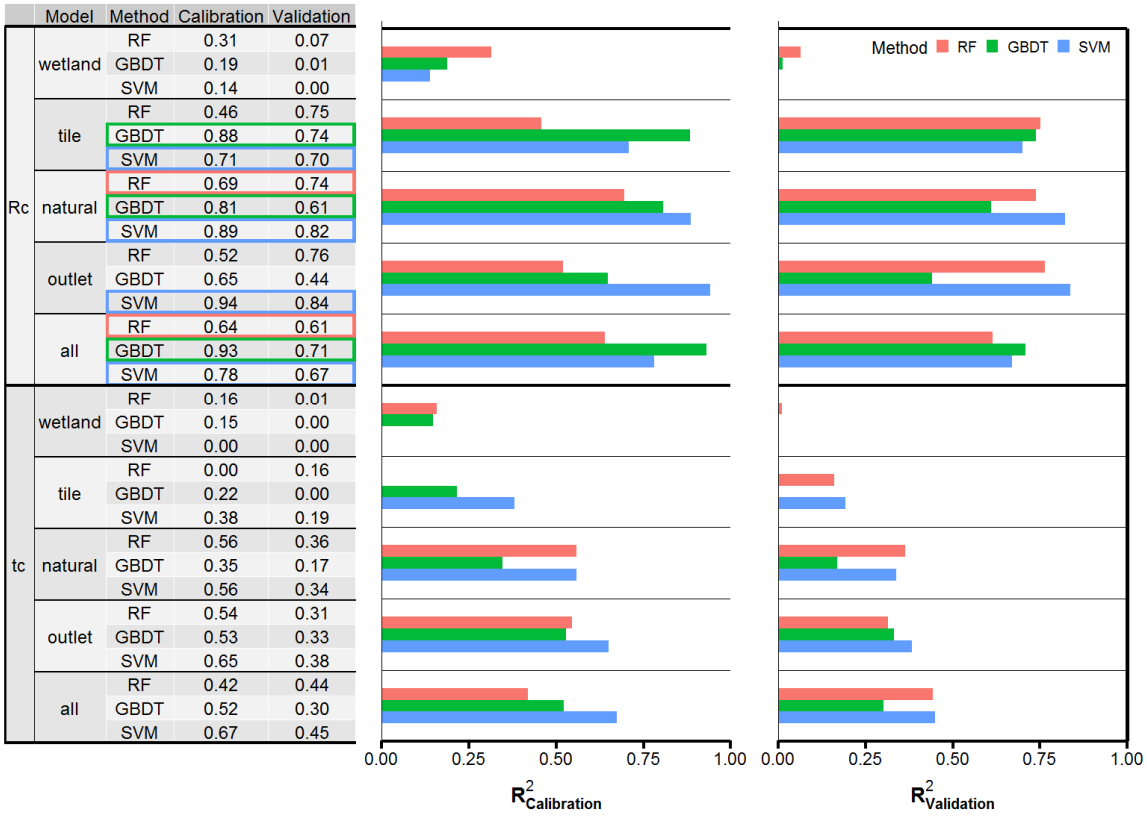


**Fig. 9.** Comparison of $R^2$ (coefficients of determination) of temporal calibration and validation between different regression methods (colours) used to estimate *Rc* and *tc* stratified by runoff generation type. Models with $R^2 > 0.6$ both in calibration and validation are highlighted by boxes at the left.

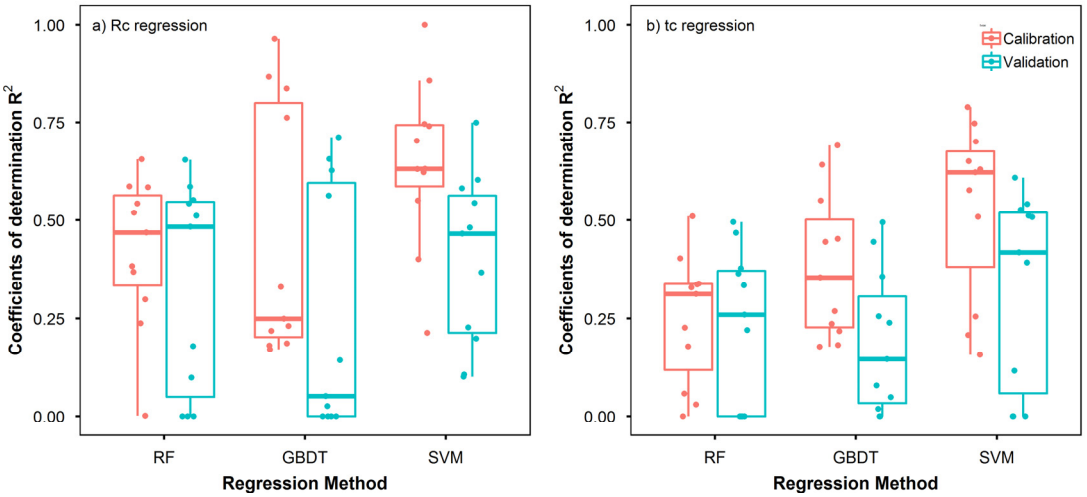


**Fig. 10.** Performance, $R^2$ (coefficients of determination), of the *Rc*/*tc* regression methods (RF, GBDT and SVM) with spatial calibration and validation as in Table 8.

**Table 9.** Correlation coefficient of *Rc* and *tc* with the explanatory variables according to Table 2.

| *R* | VolP | PeakP | DurP | PreSM | PeakSM | PeakPSM | AverSM | EndSM | DelaySM | PreWL | PeakWL |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Rc* | 0.15 | –0.29 | 0.50 | –0.04 | –0.02 | –0.01 | –0.01 | –0.03 | 0.02 | 0.43 | 0.48 |
| *Tc* | –0.04 | –0.40 | 0.38 | 0.30 | 0.29 | 0.32 | 0.31 | 0.28 | –0.01 | 0.19 | 0.14 |
| *R* | PeakPWL | AverWL | EndWL | DelayWL | EP | Month | NDVI | AreaPipes | AreaForest | Area | Type |
| *Rc* | 0.47 | 0.47 | 0.50 | 0.30 | –0.26 | –0.32 | –0.21 | 0.28 | –0.17 | 0.08 | –0.19 |
| *Tc* | 0.16 | 0.15 | 0.14 | 0.33 | –0.33 | 0.05 | –0.22 | –0.04 | 0.05 | 0.06 | –0.002 |

water related variables (PreWL, PeakWL, PeakPWL, AverWL, EndWL and DelayWL). The highest *r* are obtained for EndWL and DurP ($r = 0.5$), suggesting that groundwater and precipitation are the two factors that are most strongly connected to *Rc*. It is also interesting that the rainfall duration (DurP) affects *r* more strongly than event precipitation volume (VolP). It seems that higher groundwater levels lead to more direct flow generated during an event. The groundwater level data reflect the catchment storage conditions, particularly those close to the stream, as most piezometers are close to the stream. The peak rainfall intensity (PeakP) is negatively correlated with *Rc* (but DurP positively), suggesting that saturation excess runoff tends to be more important than infiltration excess runoff.

A clearly positive effect on *tc* can be found of the variables DurP, DelayWL and the soil moisture related variables (PreSM, PeakSM, PeakPSM, AverSM and EndSM). The largest absolute *r* occurs for the precipitation variables (PeakP and DurP). Longer DurP and lower PeakP generally result in slower recessions. Higher rainfall intensities lead to higher peak flows and thinner shapes of the hydrographs with faster recessions. The influence of soil moisture and groundwater on *tc* is somewhat lower than that of precipitation. In summer, higher precipitation intensity together with low soil moisture and groundwater levels generally come out with quick streamflow recessions with low *tc*, while in winter lower precipitation intensity and wetter conditions usually with higher *tc*.

### Importance of explanatory variables for *Rc* and *tc* in the non-linear regression models

The importance of the explanatory variables for estimating *Rc* and *tc* from the non-linear regression models was evaluated on the basis of $I_{i,SVM}$ (section „Ensemble learning techniques for regression“). The heatmap (Figure 11) shows the relative importance of the explanatory variables, where each column represents one model and the colour indicates the importance of the explanatory variables. The temporal non-linear regression models are used for this analysis because of the larger database (Table 4).

For evaluating the effects of multicolinearity to variable importance of regression model, the variance inflation factor (VIF) has been calculated for each variable. VIF of most variables are smaller than 5 apart from soil moisture (PreSM, EndSM, AverSM, PeakSM and PeakPSM) and groundwater

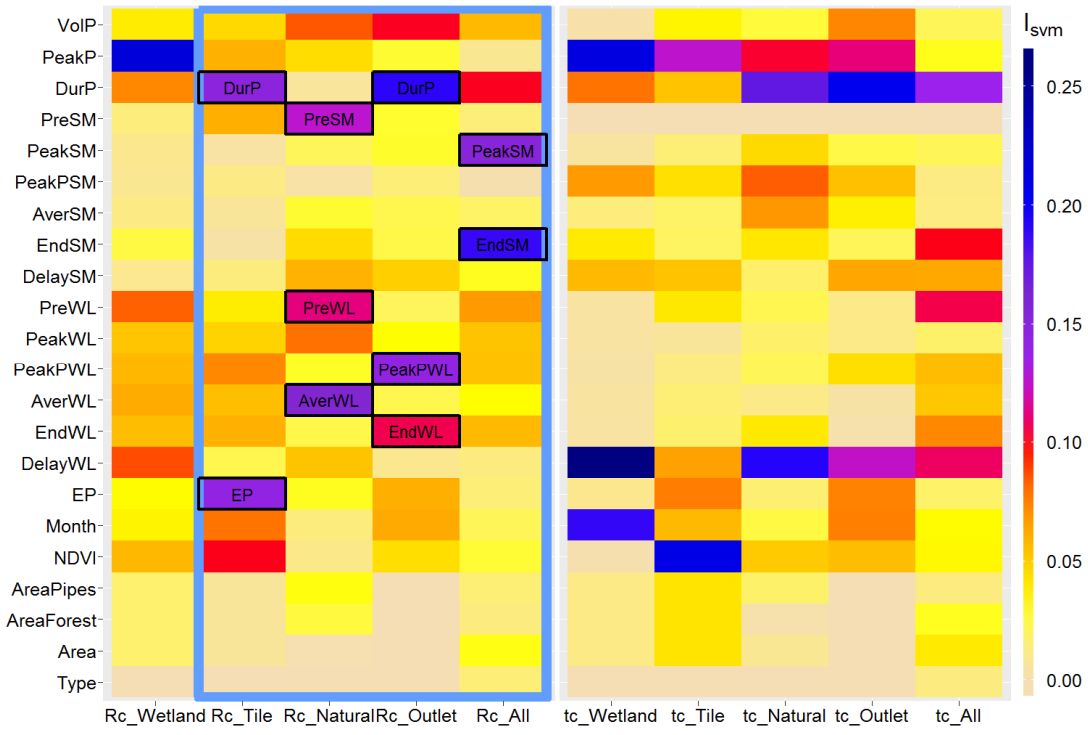


**Fig. 11.** Heatmap of the variable importance $I_{i,SVM}$ for the classified and unclassified models using the SVM regression method. The Y-axis represents 22 variables, the X-axis the models based on different subcatchment types. Left: *Rc*. Right: *tc*. Models with $R^2 > 0.6$ both in calibration and validation (Fig. 9) are highlighted by a blue box, and the variables with the best $I_{i,SVM}$ performance are highlighted by black rectangles. Database as in Table 4.

related variables (PreWL, EndWL, AverWL, PeakWL and PeakPWL). Therefore the difference from five soil moisture/groundwater variables is difficult to be explained from $I_{i,SVM}$. Figure 11 suggests that both antecedent soil moisture (PreSM) and groundwater related variables (PreWL and AverWL) have a relative high importance for the *Rc* regression of natural subsurface flow. This is in line with Tarasova (2018ab) who found that catchment storage plays a considerable role in the prediction of event runoff response. In the wetland drainage area soil moisture is always high, in the tile drainage area soil moisture varies moderately, and in the natural drainage area the seasonal variability of soil moisture is largest, which is reflected in a larger effect on *Rc*. Precipitation duration (DurP) always shows a large influence on *Rc* in the tile drainage system and at the catchment outlet. Potential evaporation (EP) is another important variable for the *Rc* regression in the tile drainage systems. This suggests that *Rc* in the tile drainage systems is more affected by the weather conditions than the subsurface dynamics. Groundwater related variables are more important in the natural system and for the entire catchment. For the unclassified data set, two soil moisture related variables (PeakSM and EndSM) generally play an important role, which means that soil moisture is a good indicator of catchment storage affecting drainage condition during the events.

The performance of the *tc* regressions is not as good as that of the *Rc* regressions ($R^2 < 0.5$ in validation mode). The precipitation variables (DurP and PeakP) and groundwater peak delay to precipitation (DelayWL) are more important than other variables. Exner-Kittridge et al. (2016) has suggested that most of the HOAL baseflow stems from groundwater, so the time when groundwater starts to contribute to the recession compared to the precipitation peak (DelayWL) may have a controlling function on the discharge recession rate.

## DISCUSSION AND CONCLUSIONS

This study addresses two main research questions. The first is the relative performance of three machine learning methods in estimating event runoff coefficients, *Rc*, and recession coefficients, *tc*. We use 22 event based explanatory variables in these methods representing precipitation, soil moisture, groundwater level and season. The regressions are performed for four, classified subcatchment groups (wetland, tile drainage, natural, outlet runoff) and for unclassified regressions using all event hydrographs from all subcatchments. Model performances, measured by the coefficient of determination $R^2$, shows that the SVM algorithm generally gives more accurate *Rc* and *tc* predictions than the other two methods (RF and GBDT). This is due to the fact that the SVM algorithm can transform the 22 dimensional variables to a higher dimensional space and it generally performs better for small sample sizes. The *Rc* regressions using SVM for tile drainage, natural, outlet and unclassified events show good performances with $R^2$ greater than 0.6, but the regression for the wetlands perform less well. The latter is presumably related to the lack of infiltration information in the wetland areas. The best *tc* estimates are obtained by the SVM model for unclassified events with $R^2 = 0.45$ in calibration mode. Overall, the *tc* regressions perform less well than those of *Rc* which is likely related to the more complex nature of subsurface and surface routing as compared to runoff generation. New geophysical observations about the structure and connectivity of different groundwater storages and understanding of the connections between shallow and deep aquifers will allow to improve model performance in the future.

The second research question relates to the most relevant variables for the *Rc* and *tc* regression based on the models described above by different categories of runoff systems. The relative importance of the variables of the SVM model is assessed by a heatmap. It suggests that precipitation duration plays an important role in predicting *Rc* for the tile drainage and outlet systems. This is in line with Merz et al. (2006) who concluded that the spatial patterns of median event runoff coefficients are highly correlated with the spatial patterns of mean annual precipitation in this climatic region. Antecedent soil moisture only affects *Rc* for the natural system, but not for the tile drainage and wetland systems. Potential evapotranspiration (EP) is an important factor in the tile drainage systems. It can therefore be concluded that event based *Rc* of the tile drainage systems is more controlled by weather conditions than by the catchment state, while the opposite is true of the natural drainage systems. This behaviour of the natural drainage systems of the HOAL is similar to the results of Tarasova et al. (2018a; 2018b), who found that the response of lowland catchments with substantial storage is driven by pre-event saturation instead of rainfall properties.

Overall, the paper shows that both the performance of estimating *Rc* and *tc* and the relative importance of explanatory variables depends strongly on the types of the hydrological systems, i.e. the runoff generation mechanism. The paper proposes three machine learning techniques as tools for predicting event based *Rc* and *tc* on the basis of weather and land surface characteristics.

*Acknowledgements.* We thank Matthias and Markus Oismüller for organizing the measurements and managing the data in the Hydrological Open Air Laboratory of Petzenkirchen. Data of soil moisture, groundwater level and weather have been processed by Mariette Vreugdenhil, Lovrenc Pavlin and Patrick Hogan. We acknowledge financial support from the Natural Science Foundation, China, and the Austrian Science Funds (FWF) as part of the Vienna Doctoral Programme on Water Resource Systems (DK W1219-N28).

## REFERENCES

Asefa, T., Kemblowski, M., McKee, M., Khalil, A., 2006. Multi-time scale stream flow predictions: the support vector machines approach. Journal of Hydrology, 318, 1–4, 7–16.

Basak, D., Pal, S., Patranabis, D.C., 2007. Support vector regression. Neural Information Processing-Letters and Reviews, 11, 10, 203–224.

Baudron, P., Alonso-Sarría, F., García-Aróstegui, J.L., Canovas-Garcia, F., Martinez-Vicente, D., Moreno-Brotons, J., 2013. Identifying the origin of groundwater samples in a multi-layer aquifer system with Random Forest classification. Journal of Hydrology, 499, 303–315.

Ben-Hur, A., Weston, J., 2010. A user's guide to support vector machines. In: Carugo, O., Eisenhaber, F. (Eds.): Data Mining Techniques for the Life Sciences. Methods in Molecular Biology (Methods and Protocols), Vol 609. Humana Press, 2010, pp. 223–239.

Biswal, B., Kumar, D.N., 2014. What mainly controls recession flows in river basins? Advances in water resources, 65, 25–33.

Blöschl, G., Blaschke, A.P., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, A., Haas, P., Hogan, P., Kazemi Amiri, A., Oismüller, M., Parajka, J., Silasari, R., Stadler, P., Strauss, P., Vreugdenhil, M., Wagner, W., Zessner, M., 2016. The

Hydrological Open Air laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory. Hydrology and Earth System Sciences, 20, 1, 227.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., Savenije, H., 2013. Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales. Cambridge University Press, Cambridge.

Blume, T., Zehe, E., Bronstert, A., 2007. Rainfall–runoff response, event-based runoff coefficients and hydrograph separation. Hydrological Sciences Journal, 52, 5, 843–862.

Breiman, L., 2001. Random forests. Machine Learning, 45, 1, 5–32.

Brutsaert, W., Nieber, J.L., 1977. Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resources Research, 13, 3, 637–643.

Cánovas-García, F., Alonso-Sarría, F., Gomariz-Castillo, F., Onate-Valdivieso, F., 2017. Modification of the random forest algorithm to avoid statistical dependence problems when classifying remote sensing imagery. Computers & Geosciences, 103, 1–11.

Chapelle, O., Vapnik, V., 2000. Model selection for support vector machines. In: NIPS'99 Proceedings of the 12th International Conference on Neural Information Processing Systems, Denver, CO, November 29 - December 4, 1999, pp. 230–236.

Chapman, T.G., Maxwell, A.I., 1996. Baseflow separation-comparison of numerical methods with tracer experiments. In: Hydrology and Water Resources 23rd Symposium, Hobart, 1996. National Conference Publication – Institution of Engineers Australia NCP, 2(5), pp. 539–546.

Chen, B., Krajewski, W.F., Helmers, M.J., Zhang, Z., 2019. Spatial variability and temporal persistence of event runoff coefficients for cropland hillslopes. Water Resources Research, 55, 2, 1583–1597.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning, 20, 3, 273–297.

Cortez, P., Embrechts, M.J., 2013. Using sensitivity analysis and visualization techniques to open black box data mining models. Information Sciences, 225, 1–17.

Deka, P.C., 2014. Support vector machine applications in the field of hydrology: a review. Applied Soft Computing, 19, 372–386.

Dietterich, T.G., 1997. Machine-learning research. AI Magazine, 18, 4, 97–97.

Erdal, H.I., Karakurt, O., 2013. Advancing monthly streamflow prediction accuracy of CART models using ensemble learning paradigms. Journal of Hydrology, 477, 119–128.

Exner-Kittridge, M., Strauss, P., Blöschl, G., Eder, A., Saracevic, E., Zessner, M., 2016. The seasonal dynamics of the stream sources and input flow paths of water and nitrogen of an Austrian headwater agricultural catchment. Science of the Total Environment, 542, 935–945.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics, 45, 1, 1189–1232.

Friedman, J.H., 2002. Stochastic gradient boosting. Computational Statistics & Data Analysis, 38, 4, 367–378.

Gaál, L., Szolgay, J., Kohnová, S., Parajka, J., Merz, R., Viglione, A., Blöschl, G., 2012. Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. Water Resources Research, 48, W04511.

Gottschalk, L., Weingartner, R., 1998. Distribution of peak flow derived from a distribution of rainfall volume and runoff coefficient, and a unit hydrograph. Journal of Hydrology, 208, 3–4, 148–162.

Hayes, D.C., Young, R.L., 2006. Comparison of peak discharge and runoff characteristic estimates from the rational method to field observations for small basins in Central Virginia. U.S. Department of the Interior, U.S. Geological Survey, Scientific Investigation Reports, 2005-5254.

Ho, T.K., 1995. Random decision forests. In: ICDAR '95 Proceedings of the Third International Conference on Document Analysis and Recognition, 1, IEEE Computer Society Washington, DC, USA, 14–15 August 1995, pp. 278–282.

Horn, R.A., Johnson, C.R., 1985. Matrix Analysis. Cambridge University Press, Cambridge.

Hsu, C.W., Chang, C.C., Lin, C.J., 2003. A practical guide to support vector classification. Technical Report. Department of Computer Science, National Taiwan University, Taipei.

Hwang, S.H., Ham, D.H., Kim, J.H., 2012. Forecasting performance of LS-SVM for nonlinear hydrological time series. KSCE Journal of Civil Engineering, 16, 5, 870–882.

Krakauer, N.Y., Temimi, M., 2011. Stream recession curves and storage variability in small watersheds. Hydrology and Earth System Sciences, 15, 7, 2377–2389.

Liaw, A., Wiener, M., 2002. Classification and regression by random Forest. R News, 2, 3, 18–22.

Longobardi, A., Villani, P., Grayson, R.B., Western, A., 2003. On the relationship between runoff coefficient and catchment initial conditions. In: Proceedings of MODSIM, pp. 867–872.

Maity, R., Bhagwat, P.P., Bhatnagar, A., 2010. Potential of support vector regression for prediction of monthly streamflow using endogenous property. Hydrological Processes, 24, 7, 917–923.

Merz, R., 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.

Merz, R., Blöschl, G., Parajka, J., 2006. Spatio-temporal variability of event runoff coefficients. Journal of Hydrology, 331, 3–4, 591–604.

Naghibi, S.A., Ahmadi, K., Daneshi, A., 2017. Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping. Water Resources Management, 31, 9, 2761–2775.

Naghibi, S.A., Pourghasemi, H.R., Dixon, B., 2016. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environmental Monitoring and Assessment, 188, 1, 44.

Norbiato, D., Borga, M., Merz, R., Blöschl, G., Carton, A., 2009. Controls on event runoff coefficients in the eastern Italian Alps. Journal of Hydrology, 375, 3–4, 312–325.

Osuna, E.E., 1998. Support vector machines: Training and applications. Diss. Massachusetts Institute of Technology.

Palleiro, L., Rodríguez-Blanco, M.L., Taboada-Castro, M.M., Taboada-Castro, M.T., 2014. Hydrological response of a humid agroforestry catchment at different time scales. Hydrological Processes, 28, 4, 1677–1688.

Patnaik, S., Biswal, B., Kumar, D.N., Sivakumar, B., 2015. Effect of catchment characteristics on the relationship between past discharge and the power law recession coefficient. Journal of Hydrology, 528, 321–328.

Rodríguez-Blanco, M.L., Taboada-Castro, M.M., Taboada-Castro, M.T., 2012. Rainfall–runoff response and event-based runoff coefficients in a humid area (northwest Spain). Hydrological Sciences Journal, 57, 3, 445–459.

Sachdeva, S., Bhatia, T., Verma, A.K., 2018. GIS-based evolutionary optimized Gradient Boosted Decision Trees for for-

est fire susceptibility mapping. Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards, Springer. International Society for the Prevention and Mitigation of Natural Hazards, 92, 3, 1399–1418.

Şen, Z., Altunkaynak, A., 2006. A comparative fuzzy logic approach to runoff coefficient and runoff estimation. Hydrological Processes, 20, 9, 1993–2009.

Shen, C., 2018. A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research, 54, 8558–8593. https://doi.org/10.1029/2018WR022643

Sivapalan, M., 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. Hydrological Processes, 17, 15, 3163–3170.

Széles, B., Broer, M., Parajka, J., Hogan, P., Eder, A., Strauss, P., Blöschl, G., 2018. Separation of scales in transpiration effects on low flows: A spatial analysis in the Hydrological Open Air Laboratory. Water Resources Research, 54, https://doi.org/10.1029/2017WR022037.

Tachecí, P., Žlabek, P., Kvitek, T., Peterkova, J., 2013. Analysis of rainfall-runoff events in four subcatchments of the Kopaninský potok (Czech Republic). Bodenkultur, 64, 3–4, 105–111.

Tague, C., Grant, G.E., 2004. A geological framework for interpreting the low-flow regimes of Cascade streams, Willamette River Basin, Oregon. Water Resources Research, 40, 4, W04303.

Tallaksen, L.M., 1995. A review of baseflow recession analysis. Journal of Hydrology, 165, 1–4, 349–370.

Tarasova, L., Basso, S., Poncelet, C., Merz, R., 2018a. Exploring controls on rainfall-runoff events: 2. Regional patterns and spatial controls of event characteristics in Germany. Water Resources Research, 54, https://doi.org/10.1029/2018WR022588

Tarasova, L., Basso, S., Zink, M., Merz, R. 2018b. 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, https://doi.org/10.1029/2018WR022587

Tian, F., Li, H., Sivapalan, M., 2012. Model diagnostic analysis of seasonal switching of runoff generation mechanisms in the Blue River basin, Oklahoma. Journal of Hydrology, 418, 136–149.

Vapnik, V., Golowich, S., Smola, A., 1997. Support vector method for function approximation, regression estimation, and signal processing. In: NIPS'96 Proceedings of the 9th International Conference on Neural Information Processing Systems, Denver, Colorado, 3–5 December 1996, pp. 281–287.

Viglione, A., Merz, R., 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.

Wainwright, J., Parsons, A.J., 2002. The effect of temporal variations in rainfall on scale dependency in runoff coefficients. Water Resources Research, 38, 12, 1271.

Zimmermann, B., Zimmermann, A., Turner, B.L., Francke, T., Elsenbeer, H., 2014. Connectivity of overland flow by drainage network expansion in a rain forest catchment. Water Resources Research, 50, 2, 1457–1473.

Received 9 May 2019
Accepted 30 January 2020