J. Hydrol. Hydromech., 69, 2021, 1, 65–75
DOI: 10.2478/johh-2020-0029
ISSN 1338-4333

sciendo Journal of Hydrology and Hydromechanics

Stepwise prediction of runoff using proxy data in a small agricultural catchment

Borbála Széles[1*], Juraj Parajka[1], Patrick Hogan[1], Rasmiaditya Silasari[1], Lovrenc Pavlin[1], Peter Strauss[2], Günter Blöschl[1]

[1] Institute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Karlsplatz 13/222, 1040 Vienna, Austria.
[2] Federal Agency of Water Management, Institute for Land and Water Management Research, Pollnbergstraße 1, 3252 Petzenkirchen, Austria.
[*] Corresponding author. Tel.: +43 (0)1 58801-22335. E-mail: szeles@waterresources.at

Abstract: In this study, the value of proxy data was explored for calibrating a conceptual hydrologic model for small ungauged basins, i.e. ungauged in terms of runoff. The study site was a 66 ha Austrian experimental catchment dominated by agricultural land use, the Hydrological Open Air Laboratory (HOAL). The three modules of a conceptual, lumped hydrologic model (snow, soil moisture accounting and runoff generation) were calibrated step-by-step using only proxy data, and no runoff observations. Using this stepwise approach, the relative runoff volume errors in the calibration and first and second validation periods were –0.04, 0.19 and 0.17, and the monthly Pearson correlation coefficients were 0.88, 0.71 and 0.64, respectively. By using proxy data, the simulation of state variables improved compared to model calibration in one step using only runoff data. Using snow and soil moisture information for model calibration, the runoff model performance was comparable to the scenario when the model was calibrated using only runoff data. While the runoff simulation performance using only proxy data did not considerably improve compared to a scenario when the model was calibrated on runoff data, the more accurately simulated state variables imply that the process consistency improved.

Keywords: Hydrologic model; Model calibration; Ungauged basins; Experimental catchment.

INTRODUCTION

Runoff is a reflection of the aggregated hydrologic catchment behavior. Therefore, in most cases, runoff observations are used for calibrating hydrologic models. However, in many catchments runoff observations are not available (Blöschl et al., 2013) and therefore, other measurements on the hydrological processes, i.e. proxy data, are used to calibrate the model.

There are only a few studies that used only proxy data for parameter estimation in ungauged basins (Parajka et al., 2013). These studies were mainly focusing on physically based hydrologic models, where model simulations can be explicitly linked to field measurements. For instance, Thyer et al. (2004) evaluated the performance of the distributed hydrology soil vegetation model (DHSVM) using field data. Their study site was a high elevated and forested catchment, where they found that the simulated runoff was most influenced by snowmelt characteristics (Thyer et al., 2004). While Thyer et al. (2004) focused on the micro-meteorological part of the model, in a follow-up study, Kuras et al. (2011) completed this evaluation by testing the subsurface and surface runoff dynamics. Both studies achieved a daily Nash Sutcliffe efficiency for runoff of above 0.75. Kuppel et al. (2018) performed a similar study in the Scottish Highlands with a distributed process-based eco-hydrological model. They found that the model performance was better, when runoff was also used for calibration and they also noted that certain state variables can be only well simulated when the model was calibrated for them.

If in-situ measurements are unavailable, an alternative could be to use remote sensing products for model calibration (López et al., 2017; Nijzink et al., 2018; Silvestro et al., 2015). Nijzink et al. (2018) tested nine remotely sensed products and found that without using runoff data, remotely sensed soil moisture products and the GRACE total water storage anomalies constrained the most the model parameters. López et al. (2017) found that remotely sensed evapotranspiration and remotely sensed soil moisture should be used together, and not independently, to predict runoff. Generally, due to the coarse spatio-temporal resolution of these products, they cannot be used for small catchments. For small catchments, in-situ observations are necessary.

With or without runoff data an efficient way of model calibration is stepwise parameter estimation which reduces the dimensionality of the problem. With runoff data, model parameters can be grouped according to which runoff signatures they influence (e.g. Fenicia et al., 2007; Gelleszun et al., 2017; Hogue et al., 2000), or on which time scales the model parameters are sensitive (e.g. Lu and Li, 2015). Some of the studies also used proxy data to perform a step-by-step model calibration (e.g. Avanzi et al., 2020; Hay et al., 2006; Kuras et al., 2011; Ning et al., 2015), which is very useful in order to understand possible mismatches between model simulations and measurements (Rogger et al., 2012). In a recent study, Széles et al. (2020) proposed a stepwise model calibration approach, where they aimed to calibrate a conceptual hydrologic model according to the simulated processes, such as snow accumulation and snowmelt, soil moisture and evapotranspiration, and runoff generation. They linked the simulated processes with a variety of in-situ field observations. These proxy data and runoff data were together used in their study to estimate the parameters of their conceptual hydrologic model. However, it was not yet clear whether this method could be potentially used in ungauged catchments, to predict runoff without using runoff observations.

The objective of this study was to test whether the stepwise model calibration approach proposed by Széles et al. (2020) could be used for predicting runoff without using runoff observations. Without incorporating runoff in the objective functions, we aimed to test how well we can predict runoff and various state variables of the model on the annual and seasonal time scales. The analysis was performed in the 66 ha Austrian Hydrological Open Air Laboratory (HOAL), where long-term field observations are available (Blöschl et al., 2016).

STUDY AREA AND DATA

Study area

The study site was a 66 ha experimental catchment, the Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, Lower Austria (Figure 1) (Blöschl et al., 2016). The elevation of the catchment ranges between 257 and 323 m above sea level. The stream is approximately 620 m long (Eder et al., 2010, 2014; Széles et al., 2018). The climate is humid. Mean annual (1991–2017) air temperature, precipitation and runoff are 9.6°C, 782 mm/yr and 184 mm/yr, respectively. Air temperature and rainfall amount have a maxima in the summer. Mean monthly runoff tends to peak in winter or early spring. The geology of the catchment consists of Tertiary fine sediments and fractured siltstone of the Molasse zone. The dominant soil types are Cambisols (57%), Kolluvisol (16%) and Planosols (21%) with moderate to low permeability. Gleysols (6%) occur close to the stream (Blöschl et al., 2016). The catchment is dominated by agricultural land use (87% of the catchment area), the rest of the catchment is forested, paved or used as pasture.

Data

In this study, we used the same data presented by Széles et al. (2020). The measurements included precipitation amount, precipitation type, runoff observations and time lapse photographs with one-minute temporal resolution. Air temperature has been measured at 7, 14, 19h until October 2012, since then it has been measured with half hourly time step. The potential evapotranspiration ET_P was calculated with the modified Blaney-Criddle method (Parajka et al., 2003; Schrödter, 1985). Snow depth, soil moisture and actual evapotranspiration have been monitored with half hourly temporal resolution. Groundwater levels have been measured every five minutes. Details on the instruments, their location and spatio-temporal resolution are given in Blöschl et al. (2016) and Széles et al. (2020).

Three time periods were selected for the analysis, a 22-year-long period when only runoff measurements (1991–2012), and a 3-year-long (2013–2015) and a 2-year-long (2016–2017) period when runoff measurements and additional sources of data were available. The 3-year-long period was used for model calibration (Calib), the 22-year-long (Val1) and 2-year-long (Val2) periods for model validation. One year proceeding each period was used as warm-up period. Snow accumulation was simulated with half hourly temporal resolution, while other processes were simulated with daily time step.

METHODOLOGY

Hydrologic model

In this study we used a conceptual hydrologic model, the TUWmodel (Parajka et al., 2007), which follows the structure of the HBV model (Bergström, 1976; Bergström and Lindström, 2015; Lindström et al., 1997). The model has three modules (snow, soil moisture accounting and runoff generation) and 14 free parameters (Merz and Blöschl, 2004; Parajka et al., 2007; Széles et al., 2020). The free parameters according to the three modules and their calibration ranges are shown in Table 1. The ranges were specified based on literature values (Merz et al., 2011; Viglione et al., 2013), except for field capacity *FC*, which was constrained according to a soil survey (Murer et al., 2004).

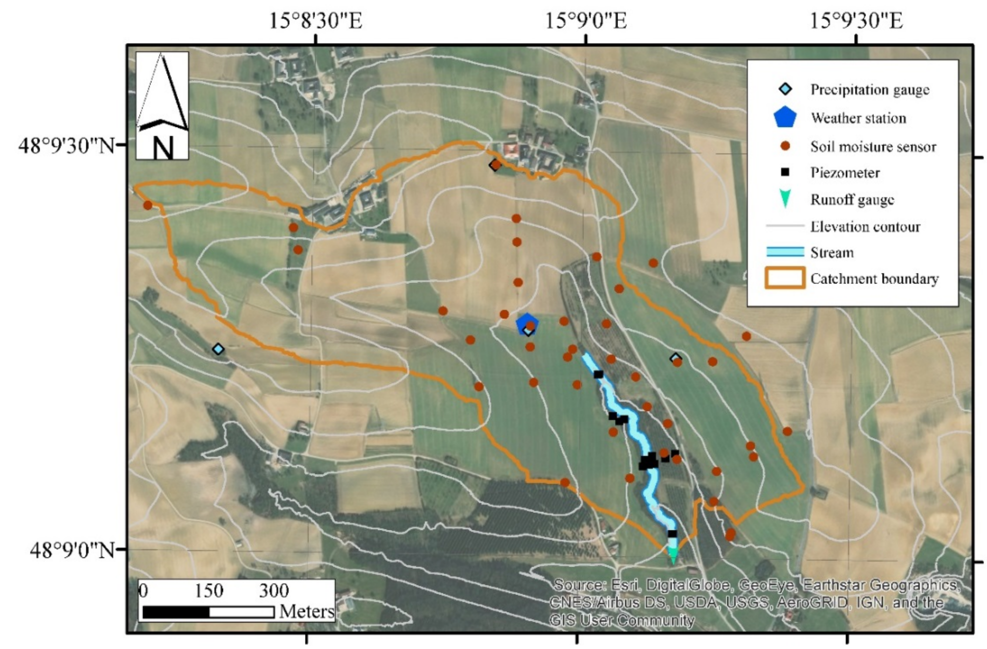

Fig. 1. Study area: Hydrological Open Air Laboratory (HOAL) in Petzenkirchen, Lower Austria and location of weather station (automatic weather station, present weather sensor for precipitation phase measurements, snow sensor, digital camera, one eddy covariance system), precipitation gauges, soil moisture sensors, groundwater level measurements by piezometers, and catchment outlet (source of basemap: Esri et al., 2020).

Table 1. 14 free parameters of the hydrologic model according to the three modules and their calibration range.

Module	Parameter (unit)	Parameter name	Calibration range: min÷max
Snow	*SCF* (–)	Snow correction factor	0.9÷1.5
	DDF (mm/°C/d)	Degree day factor	0.0÷5.0
	T_{wb} (°C)	Wet bulb temperature, i.e. threshold temperature below which precipitation is snow	–3.0÷1.0
	T_m (°C)	Threshold temperature above which melt starts	–2.0÷2.0
Soil moisture accounting	*LPrat* (–)	Limit for potential evapotranspiration	0.0÷1.0
	FC (mm)	Field capacity	0.0÷450.0
	β (–)	Non-linear parameter for runoff production	0.0÷20.0
Runoff generation	k_0 (d)	Storage time for very fast response	0.0÷2.0
	k_1 (d)	Storage time for fast response	2.0÷30.0
	k_2 (d)	Storage time for slow response	30.0÷250.0
	LS_{UZ} (mm)	Threshold storage state for very fast runoff	1.0÷100.0
	c_P (mm/d)	Constant percolation rate	0.0÷8.0
	B_{MAX} (d)	Maximum base at low flows	0.0÷30.0
	c_R (d^2/mm)	Free scaling parameter	0.0÷50.0

In this study, we followed the stepwise calibration approach from Széles et al. (2020) but without using runoff in the optimization steps. The results were compared to a scenario, when the model was calibrated in one step, using only runoff data.

Model calibration without runoff data

The three modules of the model were calibrated step-by-step. The separate steps focused on the rainfall-runoff processes, which were calibrated using field measurements. In this way, the free parameters were step-by-step fixed, according to the modules of the model. The scenarios are listed in Table 2 and the calibrated model parameters for each scenario are shown in Table 3.

Calibration of snow module

First, all the model parameters were calibrated and the temperature threshold parameter (wet bulb temperature T_{wb}) was fixed (Scenario Sim-Snowacc, Table 2) by fitting the modelled phase of the precipitation to the observed one. The precipitation phase was measured by a present weather sensor at the weather station (Figure 1). The number of half hours with false precipitation phase simulations was minimized using the DEoptim R package for parameter optimization (Ardia et al., 2010a, 2010b, 2016; Mullen et al., 2011). The wet bulb temperature parameter was fixed.

In the next step, the remaining three snowmelt parameters (scenario Sim-Snowmelt, Table 2) were fixed. A daily snow cover index was created (showing 1 if there was snow in the catchment, otherwise 0) using 3 types of measurements. First, time lapse photos were checked to decide if there was snow in the catchment. If these were unavailable, daily MODIS Normalized Difference Snow Index images were analyzed (Hall and Riggs, 2016a, 2016b). Finally, if the MODIS images were also unavailable, the snow sensor measurements were examined. The modelled snow cover index was chosen to be 1, if the snow water equivalent exceeded 2 mm. The modelled snow cover index was fitted to the observed one by minimizing the number of days with false snow cover index simulations and using the DEoptim R package for parameter optimization. Out of the 13 calibrated parameters, the three snowmelt parameters were fixed.

Calibration of the soil moisture accounting module

In scenarios Sim-ET+SM (Table 2), the soil moisture accounting module parameters were fixed. For soil moisture, in-situ soil moisture measurements were used (Figure 1). In order to describe the temporal dynamics of soil moisture in the catchment, measurements of all stations over all depths (0.05, 0.10, 0.20 and 0.50 m depths) were averaged. To compare measured and modelled soil moisture, which might be representative for different depths, we compared standardized soil moisture values according to Equation (1)

$$SM_S = \frac{SM - \overline{SM}}{\sigma_{SM}} \tag{1}$$

where SM_S (–) is the simulated standardized soil moisture, SM (mm) is the simulated soil moisture, $\overline{SM}$ (mm) and σ_{SM} (mm) are the average and the standard deviation of the simulated soil moisture. Observed soil moisture was standardized in a similar way to Equation (1). For actual evapotranspiration, average evapotranspiration was calculated over the catchment based on measurements of three eddy covariance stations. According to the land use types, an area weighted evapotranspiration was calculated using the measurements of an eddy covariance system at the weather station (representing grass evapotranspiration) and two mobile systems (representing different crop evapotranspiration). To estimate the evapotranspiration from the riparian forest next to the stream, crop coefficients were introduced. A multi-objective function $Z1$ (–) according to Equation (2) was maximized by optimizing the remaining 10 model parameters with the help of the DEoptim R package. $Z1$ consisted of the daily Nash Sutcliffe efficiency for standardized soil moisture Z_{SM} and evapotranspiration Z_{ET} with different weights

$$Z1 = w_{SM} Z_{SM} + w_{ET} Z_{ET} \tag{2}$$

where w_{SM} (–) is the weight on the soil moisture objective, between 0 and 1. The weight on the evapotranspiration objective w_{ET} (–) is the difference between 1 and w_{SM}. This optimization step was repeated ten times for each w_{SM} weight to check the stability of the optimized model parameters. The results were examined on two time scales, annual and seasonal. On the annual time scale, the volumes of observed and simulated actual evapotranspiration were compared and the relative volume error VE_{ET} (–) was calculated (Criss and Winston, 2008). On the seasonal time scale, monthly average simulated daily actual evapotranspiration and standardized soil moisture were compared and the monthly Pearson correlation coefficient for evapotranspiration $r_{ET,m}$ (–) and standardized soil moisture $r_{SMs,m}$ (–) were calculated. Three main scenarios were chosen and the soil moisture accounting module parameters were fixed according to these. In these scenarios w_{SM} was chosen to be 0, 0.8 (where the relative volume error for ET was the smallest during the second validation period), and 1.0 (scenarios Sim-ET, Sim-ET20-SM80, and Sim-SM, respectively, Table 2).

Calibration of the runoff generation module

In order to optimize the very fast runoff q_0 simulations, we identified saturation excess runoff events according to Silasari et al. (2017). The days with very fast runoff simulations were calibrated to the days when saturation excess runoff events were observed. Storage change in the lower zone dS_{LZ} (mm/month) was calibrated using piezometer measurements (Figure 1). Based on the observed groundwater levels monthly storage change values dS_o (mm/month) were calculated for each piezometer. A catchment average storage change was calculated from spatially interpolated storage change values. The storage change values were standardized, and the simulated monthly average standardized storage change dSs (–) was fitted to the observed one dSs_o (–). A multi-objective function $Z2$ (–) according to Equation (3) was maximized for calibrating the remaining 7 model parameters with the help of the DEoptim R package. $Z2$ consisted of the relative number of days with correctly modelled very fast runoff Z_{OF} (–) and the relative number of months with correctly modelled sign of the standardized storage change Z_{dS} (–) with different weights

$$Z2 = w_{OF} Z_{OF} + w_{dS} Z_{dS} \tag{3}$$

where w_{OF} (–) is the weight on the overland flow OF objective, ranging between 0 and 1. The weight on the storage change objective w_{dS} (–) is the difference between 1 and w_{OF}. This optimization step was repeated 10 times for each w_{OF} weight to check the stability of the optimized model parameters. The modelling results were evaluated on a daily time scale for overland flow, by analyzing Z_{OF} as a function of w_{OF}. The modelling results for storage change simulations were assessed on the monthly time scale, by analyzing Z_{dSs} as a function of w_{OF}. For the selected scenarios (scenarios Sim-ET+G, Sim-SM+G, Sim-ET20+SM80+G, Table 2), w_{OF} was chosen to be 0.5.

The modelling efficiency in terms of simulating runoff was evaluated on annual and monthly time scales. On the annual time scale, the volumes of observed and simulated runoff were compared and the relative volume error for runoff VE_Q (–) was calculated. On the seasonal time scale, monthly average observed and simulated runoff time series were compared and the monthly Pearson correlation coefficient for runoff $r_{Q,m}$ (–) was calculated.

Model calibration with runoff data

Simulation results were compared with a scenario, when only runoff was used for model calibration and the model parameters were estimated in one step (Scenario Sim-R, Table 2). The model was calibrated to observed runoff by minimizing the daily root mean square error between observed and simulated runoff using the DEoptim R package.

Table 2. Scenarios presented in the study.

Scenario name	Details
Sim-Snowacc	Calibration of all model parameters and fixing the temperature threshold parameter (wet bulb temperature T_{wb}) using snow accumulation data
Sim-Snowmelt	Calibration of all model parameters except T_{wb} and fixing the snowmelt parameters (snow correction factor SCF, degree day factor DDF, snowmelt temperature T_m) using snow cover data
Sim-ET+SM	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (field capacity FC, nonlinear parameter for runoff production β, limit for potential evapotranspiration $LPrat$) using evapotranspiration and soil moisture objectives
Sim-ET	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (see scenario Sim-ET+SM) using only evapotranspiration objective
Sim-SM	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (see scenario Sim-ET+SM) using only soil moisture objective
Sim-ET20+SM80	Calibration of soil moisture accounting and runoff generation parameters and fixing the soil moisture accounting module parameters (see scenario Sim-ET+SM) using a combination of w_{ET} = 20% evapotranspiration and w_{SM} = 80% soil moisture objectives
Sim-ET+G	Calibration of runoff generation parameters, the soil moisture accounting module parameters were fixed in scenario Sim-ET
Sim-SM+G	Calibration of runoff generation parameters, the soil moisture accounting module parameters were fixed in scenario Sim-SM
Sim-ET20+SM80+G	Calibration of runoff generation parameters, the soil moisture accounting module parameters were fixed in scenario Sim-ET20+SM80
Sim-R	Calibration of all model parameters in one step using only runoff observations

Table 3. Calibrated model parameters for each scenario. Parameters which were fixed at a certain scenario (according to Table 2) are shown in bold.

Scenario name	Parameters													
	SCF	DDF	T_{wb}	T_m	$LPrat$	FC	β	k_0	k_1	k_2	LS_{UZ}	c_P	B_{MAX}	c_R
Sim-R	**1.3**	**4.3**	**–0.1**	**0.0**	**0.0**	**319.4**	**0.6**	**1.1**	**4.7**	**84.7**	**9.2**	**1.5**	**5.3**	**20.0**
Sim-Snowacc	0.9	2.6	**0.6**	0.4	0.0	587.8	6.5	1.2	20.2	164.6	33.3	3.0	15.1	5.7
Sim-Snowmelt	**1.0**	**3.2**	0.6	**–0.3**	0.7	223.8	6.7	1.9	15.6	88.9	74.4	4.4	29.0	12.6
Sim-ET	1.0	3.2	0.6	–0.3	**1.0**	**480.0**	**1.7**	0.9	9.3	53.0	25.6	7.5	27.2	10.7
Sim-SM	1.0	3.2	0.6	–0.3	**0.9**	**153.6**	**20.0**	1.0	14.3	131.1	47.8	1.2	22.0	15.3
Sim-ET20+SM80	1.0	3.2	0.6	–0.3	**1.0**	**168.7**	**4.6**	0.8	6.1	143.9	17.6	4.2	7.4	6.8
Sim-ET+G	1.0	3.2	0.6	–0.3	1.0	480.0	1.7	**0.1**	**18.8**	**189.2**	**4.2**	**0.2**	11.3	16.0
Sim-SM+G	1.0	3.2	0.6	–0.3	0.9	153.6	20.0	**0.2**	**4.9**	**30.4**	**1.0**	**0.2**	20.1	24.8
Sim-ET20+SM80+G	1.0	3.2	0.6	–0.3	1.0	168.7	4.6	**0.5**	**12.6**	**31.6**	**1.1**	**0.3**	13.3	6.5

Table 4. Performance of snow accumulation and snowmelt simulations for three scenarios (Sim-R, Sim-Snowacc, Sim-Snowmelt) in the calibration and validation periods. Snow simulation efficiency is described by the number of time steps with poor (i.e. when the simulated phase of the precipitation and simulated snow cover index, respectively, mismatched the observed one) snow accumulation and snowmelt simulations relative to the number of time steps with observations. Scenarios are described in Table 2.

Scenario	Relative number of time steps with poor snow accumulation simulations (%)		Scenario	Relative number of time steps with poor snowmelt simulations (%)	
	Calibration period 2013–2015	Validation period 2 2016–2017		Calibration period 2013–2015	Validation period 2 2016–2017
Sim-R	0.45	0.52	Sim-R	4.66	7.25
Sim-Snowacc	0.31	0.40	Sim-Snowmelt	4.38	6.29
Number of half hourly time steps	35626	23972	Number of daily time steps	1095	731

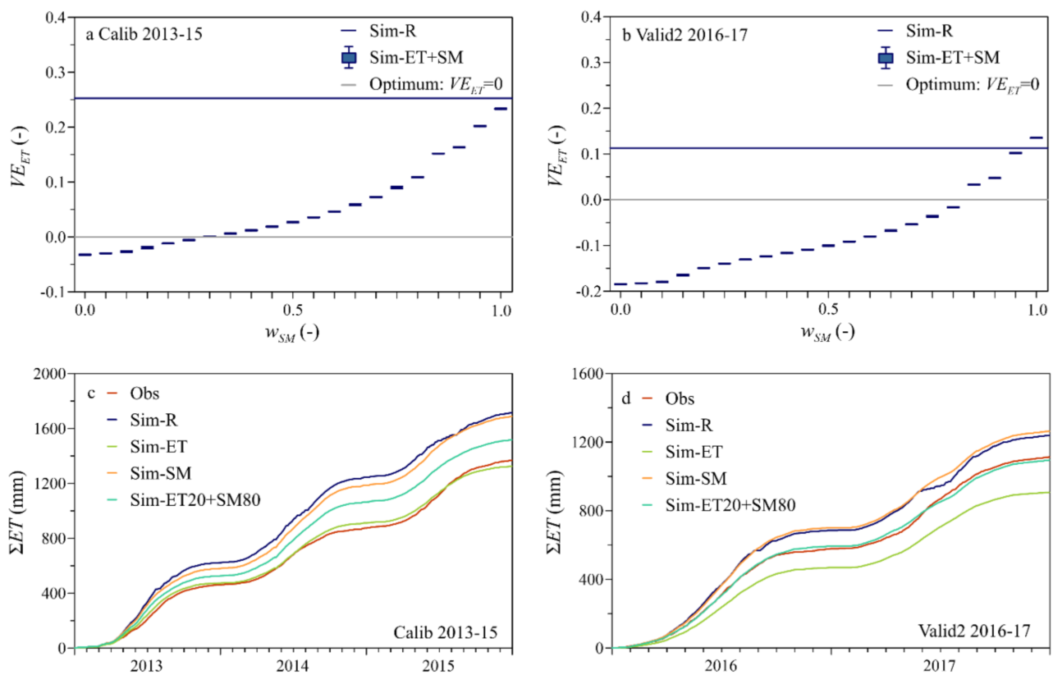

Fig. 2. Performance of model simulations in terms of relative volume error for actual evapotranspiration VE_{ET} as a function weight on soil moisture objective w_{SM} (panels a and b) and cumulative actual evapotranspiration ΣET for 4 scenarios (panels c and d) in the calibration (panels a and c) and validation periods (panels b and d). Scenarios are described in Table 2.

RESULTS

Snow module simulations

Calibrating the wet bulb temperature T_{wb}, to the observations from the present weather sensor (Scenario Sim-Snowacc) gave 0.31 and 0.40% of poor simulation times steps in the calibration and validation periods, respectively. This was slightly better than the simulations that used only runoff for calibration (Scenario Sim-R) (Table 4). Using observations from the present weather sensor, the calibrated wet bulb temperature T_{wb} parameter became closer to 1°C, which was the observed wet bulb temperature in the catchment according to Széles et al. (2020).

Fixing the snowmelt parameters, such as snowmelt temperature T_m, degree day factor DDF and snow correction factor

SCF, to the observed snow cover index (Scenario Sim-Snowmelt) gave 4.38 and 6.29% of poor simulation time steps which, again, was better than the Sim-R scenario (Table 4). By using information on snowmelt for model calibration, the calibrated snow correction factor *SCF* became closer to 1.0, which was closer to the expected value on a lowland catchment (Table 3). In scenario Sim-R, the larger amount of snow as a consequence of the higher *SCF* was then compensated by more intense snowmelt, i.e. higher degree day factor *DDF*, which was not realistic in Austrian flatland catchments (Merz et al., 2011; Sleziak et al., 2020).

Soil moisture accounting module simulations

The relative volume error and the cumulative values of actual evapotranspiration calculated for 4 scenarios (Table 2) are compared in Figure 2. The smallest relative volume errors of actual evapotranspiration were achieved for soil moisture weights of w_{SM} = 0.3 and w_{SM} = 0.8 in the calibration and validation periods based on the average of 10 model runs (Figure 2a and b). The model tended to overestimate evapotranspiration significantly (Sim-ET, Figure 2c and d), except when evapotranspiration was used in the objective function with a higher weight. Compared to the Sim-R scenario, the annual performance of actual evapotranspiration simulations improved for all w_{SM} weights, except for w_{SM} = 1 during validation.

The monthly correlation coefficient for evapotranspiration was above 0.75 both during model calibration and validation (Figure 3). While the monthly correlation coefficient for standardized soil moisture was low if only evapotranspiration was used in the objective function, the monthly correlation for evapotranspiration was almost constant independently from the weight on the evapotranspiration objective (Figure 3a–d). When both evapotranspiration and soil moisture objectives were

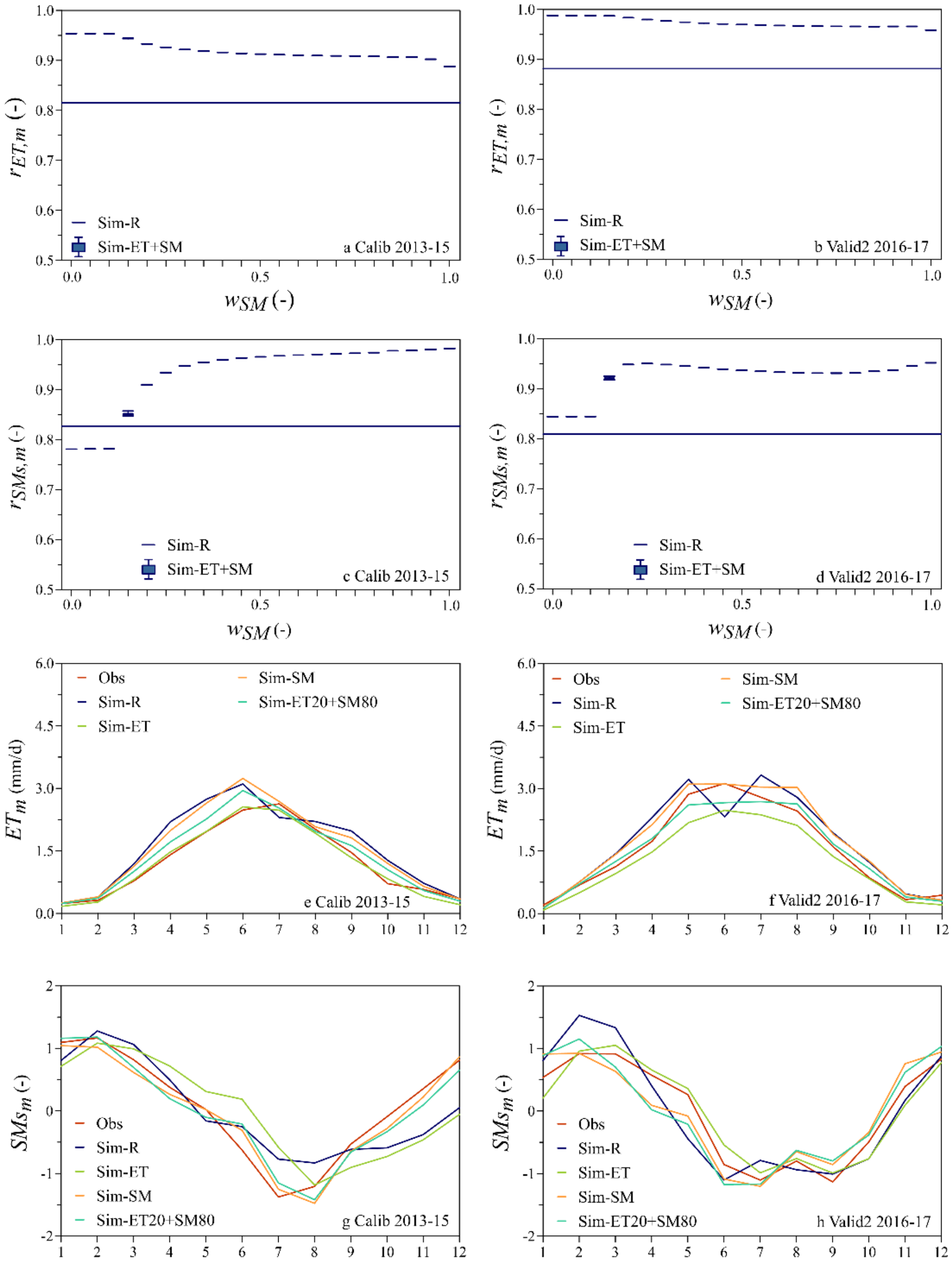

Fig. 3. Performance of model simulations in terms of monthly Pearson correlation coefficients for evapotranspiration (*ET*) and standardized soil moisture (*SMs*) as a function of weight on soil moisture objective w_{SM} (panels a–d) and monthly averages for *ET* and *SMs* for different scenarios according to Table 2 (panels e–h).

involved in the objective function, the proposed approach generally outperformed the Sim-R scenario (Figure 3a–d). Among the analyzed scenarios for the proposed step-by-step model calibration, the monthly average evapotranspiration rates and the standardized soil moisture could be best simulated, when both evapotranspiration and soil moisture were involved in the objective function (Figure 3e–h).

By using information on actual evapotranspiration and/or soil moisture, the limit for potential evapotranspiration *LPrat* became closer to 1.0, i.e. more wetness in the soils was needed to reach the potential value of evapotranspiration, which was more realistic than the close to zero value for scenario Sim-R (Table 3). The nonlinear parameter for runoff production β became larger if actual evapotranspiration and/or soil moisture were used for model calibration, which was again more realistic considering that the catchment is characterized by clayish soils, where the runoff generation can be more nonlinear (Table 3). When only runoff was used for model calibration (Sim-R scenario), the soil often completely dried out in the model, the moisture content became zero, therefore actual evapotranspiration became water limited and the evapotranspiration rates dropped (Figure 3f).

For further analysis, three main scenarios were chosen, when a weight of w_{SM} = 0.0 (Sim-ET), w_{SM} = 1.0 (Sim-SM), and w_{SM} = 0.8 (Sim-ET20-SM80) was used on the soil moisture objective in the objective function when calibrating the soil moisture accounting module.

Runoff generation module simulations

The performance of the runoff generation module simulations in terms of the overland flow Z_{OF} and storage change Z_{dSs} objectives as a function of weight on the overland flow objective are shown in Figure 4. The median of the relative number of days with good overland flow simulations immediately exceeded 0.6 as soon as the weight on the overland flow part in the compound objective function was larger than zero (Figure 4a). The standardized groundwater storage change objective gradually deteriorated as the weight on the overland flow objective increased. Generally, the results outperformed the Sim-R scenario, except for standardized monthly average storage change during validation (Figure 4d).

By using runoff generation information for model calibration, the calibrated very fast storage time k_0 became smaller compared to scenario Sim-R, which was more realistic considering that overland flow events usually last a few hours in the catchment (Table 3). The fast and slow storage times, k_1 and k_2, which are expected to be several months long, increased only for scenario Sim-ET+G (Table 3).

For further analysis, the subsurface module parameters of the three main scenarios (Sim-ET+G, Sim-SM+G and Sim-ET20+SM80+G) were calibrated by choosing the weight on the overland flow objective as w_{OF} = 0.5.

Runoff simulation

The final step was to evaluate the different scenarios in terms of simulating runoff on annual and seasonal time scales (Figure 5 and Table 5). On the annual time scale, the performance of the proposed step-by-step approach was similar to the performance of the traditional Sim-R scenario. The proposed method (scenario Sim-SM) even outperformed the Sim-R scenario during the shorter validation period (Val2, 2016–2017). For all periods, the model performance was the best, when only soil moisture was used in the optimization of the soil moisture accounting module and the runoff generation module was not optimized (Sim-SM).

Similarly, on the seasonal time scale, the proposed method could efficiently model runoff. In terms of the monthly correlation coefficients (Table 5), the Sim-SM scenario performed the best during the validation periods by slightly outperforming the Sim-R scenario.

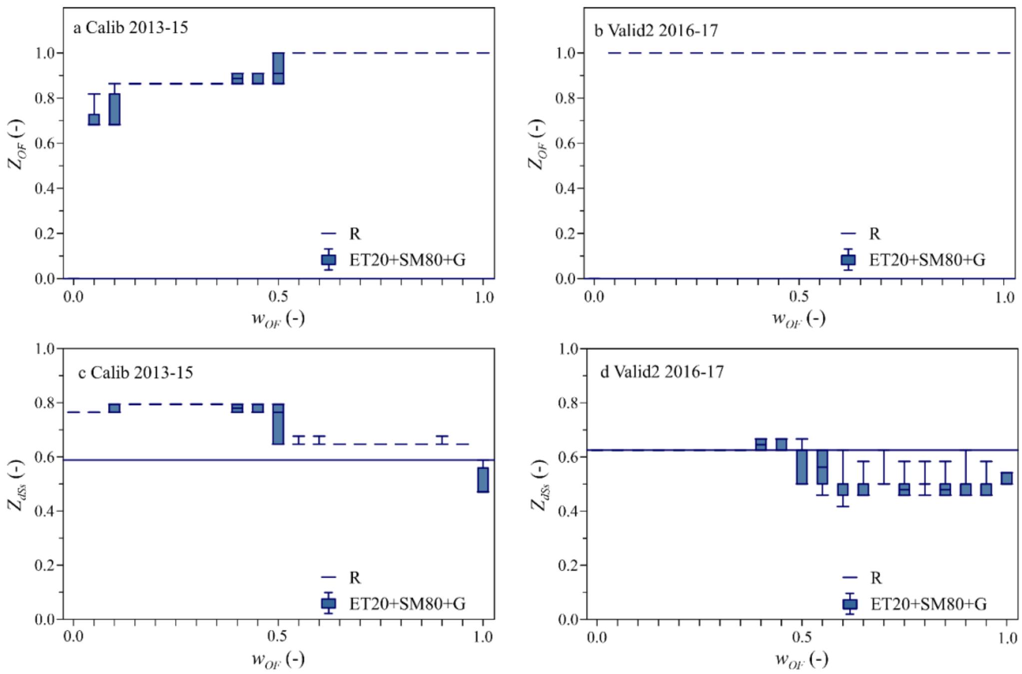

Fig. 4. Performance of model simulations in terms of the relative number of days with good overland flow simulations Z_{OF} (when the model simulated very fast runoff simultaneously with the observed overland flow events) as a function of weight on overland flow part w_{OF} (panels a and b) and the relative number of months with correctly simulated sign of the standardized groundwater storage change as a function of weight on overland flow part w_{OF} (panels c and d) during model calibration (panels a and c) and validation (panels b and d). Scenarios are described in Table 2.

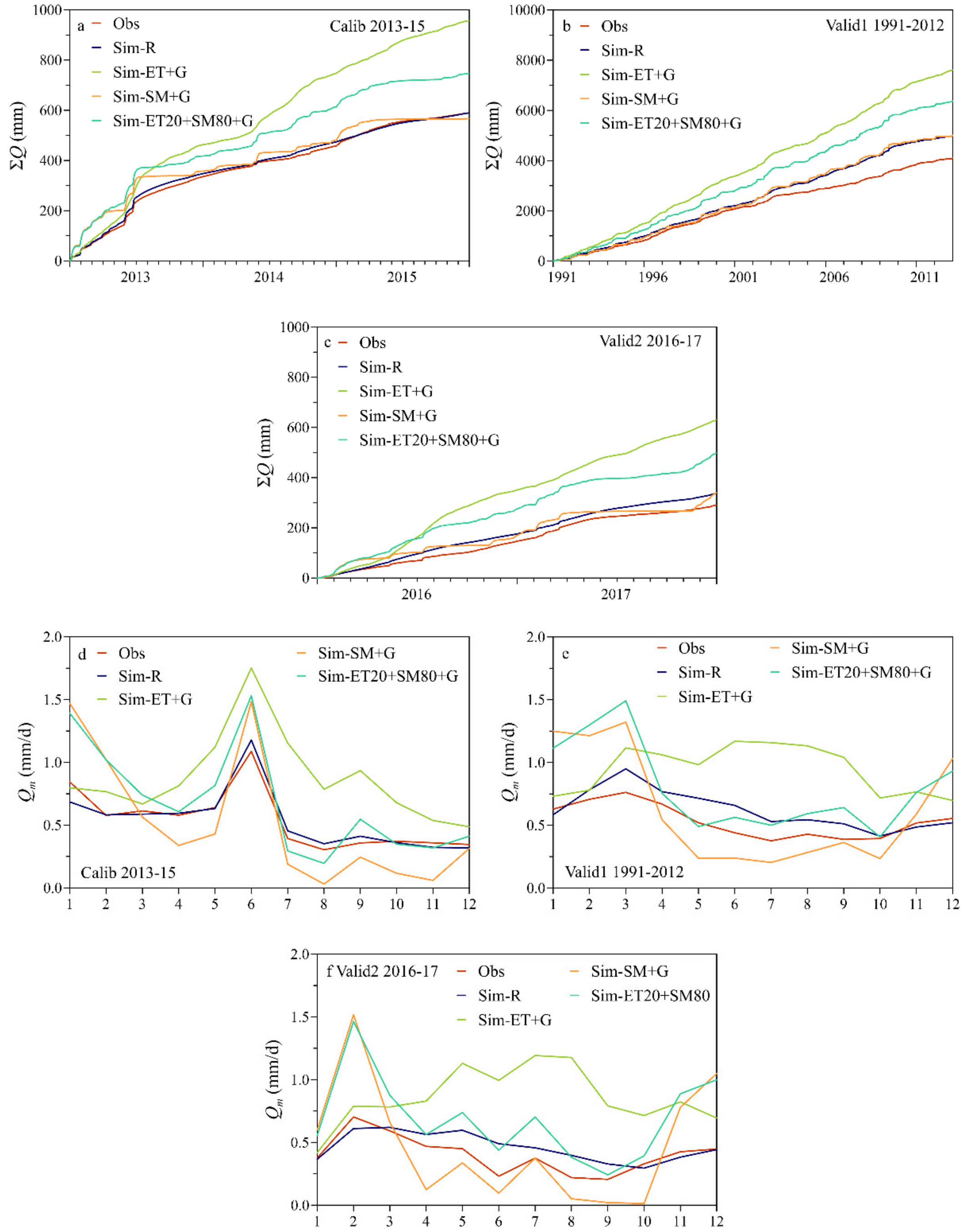

Fig. 5. Performance of model simulations in terms of cumulative runoff (panels a–c) and monthly average runoff Q_m (panels d–f). Scenarios are shown in Table 2.

Table 5. Performance of runoff simulations in terms of relative volume error *VE* and monthly Pearson correlation coefficient r_m for runoff (as in Figure 5) for model calibration and validation periods. Best performing scenario for the proposed approach is shown in bold.

Scenario	Calibration period 2013–2015		Validation period 1 1991–2012		Validation period 2 2016–2017	
	VE (–)	r_m (–)	*VE* (–)	r_m (–)	*VE* (–)	r_m (–)
Sim-R	0.00	0.98	0.18	0.75	0.15	0.75
Sim-Snowacc	–0.97	0.00	–0.71	0.28	–0.99	–0.03
Sim-Snowmelt	–0.23	0.78	–0.03	0.62	–0.31	0.43
Sim-ET	0.65	0.81	0.79	0.53	1.09	–0.02
Sim-SM	**0.00**	**0.93**	**0.19**	**0.76**	**0.00**	**0.79**
Sim-ET20+SM80	0.31	0.96	0.50	0.73	0.50	0.68
Sim-ET+G	0.62	0.82	0.80	0.58	1.15	0.16
Sim-SM+G	–0.04	0.88	0.19	0.71	0.17	0.64
Sim- ET20+SM80+G	0.27	0.91	0.52	0.74	0.70	0.69

DISCUSSION

In this study, we followed the stepwise model calibration approach proposed by Széles et al. (2020) but without using runoff observations. The aim was to test if their method could be efficiently used in a quasi ungauged catchment case, i.e. ungauged in terms of runoff. We aimed to investigate how accurately we could simulate runoff and other state variables, whether there might be a tradeoff between these. Our results showed that additional measurements can help to efficiently predict runoff on the annual and seasonal time scales. This finding suggests that there is room for expanding the usual focus on runoff predictions (Blöschl et al., 2013; Hrachowitz et al., 2013) to other components of the hydrologic cycle. For instance, if snow, soil moisture and actual evapotranspiration observations are available, these can significantly help to constrain hydrologic models and to improve the process consistency, if no runoff measurements exist. This is useful, because this implies that the model simulates runoff well for the right reasons and the model indeed represents reality (Beven and Freer, 2001; Rogger et al., 2012; Savenije, 2001; Viglione et al., 2018). Although a limitation of this method might be that such field observations and well-equipped experimental catchments are rare (Blöschl et al., 2016).

In our study we found that in iterative model calibration, soil moisture measurements were the most important to obtain runoff efficiency comparable to runoff efficiency achieved with a model calibrated to runoff only. This result might be expected considering that the most sensitive parameter of the model was the field capacity in this catchment (Széles et al., 2010), which was influenced by the soil moisture and actual evapotranspiration measurements during model calibration. The soil moisture dynamics in this catchment followed better the runoff dynamics compared to actual evapotranspiration, therefore soil moisture proved to be a better proxy to predict runoff. At the annual and seasonal time scales runoff model performance was similar in the 22-year-long validation period and even slightly larger in the 2-year-long validation period. These results are consistent with the results of Thyer et al. (2004) and Kuras et al. (2011) who reported similarly good results on the annual (relative volume error below 13% and 23%, respectively) and on the daily time scales for a physically-based eco-hydrologic model.

Results from other studies in terms of whether soil moisture and/or evapotranspiration measurements improve runoff simulations are less conclusive. Similarly to our results, Nijzink et al. (2018) also found that soil moisture satellite products were more effective than evaporation products for deriving more constrained parameter distributions. On the other hand, López et al. (2017) showed that estimating runoff was more efficient, if both soil moisture and evapotranspiration satellite products were involved in the model calibration. Bergström and Lindström (2015) and Baroni et al. (2019) pointed out that the relative importance of the measurements is influenced by the time step of the model considering that more water is stored in the unsaturated zone compared to the evapotranspired volumes. Furthermore, the changes in the different processes depend on how much they are decoupled from the atmosphere. For example, storage in the unsaturated and saturated zones changes more slowly than evapotranspiration which is coupled to the atmosphere.

In our study actual evapotranspiration, overland flow and groundwater level measurements did not help much to constrain the conceptual hydrologic model. Possible explanation why this was the case is an apparent mismatch between field observations and the HBV type, soil moisture dependent evapotranspiration calculations. For example, during precipitation events measured evapotranspiration drops to zero, while model simulations increase due to the higher moisture content in the soil. Another issue is the possible overestimation of actual evapotranspiration considering that the Nash Sutcliffe efficiency was used during calibration. This means that actual evapotranspiration was fitted to the higher values and not the lower ones. The third possibility is the difficulty in upscaling evapotranspiration to the catchment scale using point measurements. The aim of the three different eddy covariance stations was to capture the difference between crop types, which could be used for an area-based upscaling. Still, what the model sees as a "catchment average" evapotranspiration rate might be different from the upscaled values. Finally, a fourth explanation might be that water for evapotranspiration especially in the summer months might be extracted from deeper soil layers, and not the layers which are monitored by the soil moisture sensors. Using overland flow and groundwater observations generally did not improve runoff simulations. Although, simulation of runoff generation processes improved, the runoff simulation efficiencies deteriorated. Several studies also found that runoff can no longer be simulated that efficiently, if additional data (e.g. snow, evapotranspiration, groundwater levels) were also used for model calibration besides runoff (e.g. Gui et al., 2019; Parajka et al., 2007; Seibert, 2000). This is the cost of improving model consistency.

CONCLUSIONS

In this study, we calibrated the parameters of a conceptual hydrologic model step-by-step using all the available field observations except runoff. We investigated the value of proxy data for predicting runoff in a small agricultural catchment. Our results showed that:

- Using only snow and soil moisture information for calibration, the runoff model performance was comparable to the scenario when the model was calibrated in one step, using only runoff measurements.
- By using proxy data for model calibration, the simulation of state variables and therefore the process consistency improved, implying that the model represents reality better than the scenario when only runoff was used for model calibration.

Acknowledgements. The authors would like to acknowledge financial support provided by the Austrian Science Funds (FWF) as part of the Vienna Doctoral Program on Water Resource Systems (DK W1219-N28).

REFERENCES

Ardia, D., Ospina Arango, J.D., Giraldo Gomez, N.D., 2010a. Jump-diffusion calibration using differential evolution. Wilmott Magazine, 55, 76–79.

Ardia, D., Boudt, K., Carl, P., Mullen, K.M., Peterson, B.G., 2010b. Differential evolution with 'DEoptim': An application to non-convex portfolio optimization. The R Journal, 3, 1, 27–34.

Ardia, D., Mullen, K.M., Peterson, B.G., Ulrich, J., 2016. 'DEoptim': Differential evolution in 'R'. version 2.2-4.

Avanzi, F., Maurer, T., Glaser, S.D., Bales, R.C., Conklin, M.H., 2020. Information content of spatially distributed ground-based measurements for hydrologic-parameter calibration in mixed rain-snow mountain headwaters. Journal of Hydrology, 582, 124478.

Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., Simmer, C., Attinger, S., 2019. A comprehensive distributed hydrological modeling intercomparison to support process representation and data collection strategies. Water Resources Research, 55, 990–1010.

Bergström, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments. Department of Water Resources Engineering, Lund Institute of Technology, University of Lund, Bulletin Series A, no. 52.

Bergström, S., Lindström, G., 2015. Interpretation of runoff processes in hydrological modelling – experience from the HBV approach. Hydrological Processes, 29, 3535–3545.

Beven, K.J., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. Journal of Hydrology, 249, 1–4, 11–29.

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, United Kingdom, 465 p.

Blöschl, G., Blaschke, A.P., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, 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, 227–255.

Criss, R.E., Winston, W.E., 2008. Do Nash values have value? Discussion and alternate proposals. Hydrological Processes, 22, 2723–2725.

Eder, A., Strauss, P., Kreuger, T., Quinton, J., 2010. Comparative calculation of suspended sediment loads with respect to hysteresis effects. Journal of Hydrology, 389, 1–2, 168–176.

Eder, A., Exner-Kittridge, M., Strauss, P., Blöschl, G., 2014. Re-suspension of bed sediment in a small stream – results from two flushing experiments. Hydrology and Earth System Sciences, 18, 1043–1052.

Esri, DigitalGlobe, GeoEye, i-cubed, USDA FSA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, GIS User Community, 2020. "World Imagery" [basemap]. https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer. (March 31, 2020)

Fenicia, F., Savenije, H.H.G., Matgen, P., Pfister, L., 2007. A comparison of alternative multiobjective calibration strategies for hydrological modeling. Water Resources Research, 43, 3, W03434.

Gelleszun, M., Kreye, P., Meon, G., 2017. Representative parameter estimation for hydrological models using a lexicographic calibration strategy. Journal of Hydrology, 553, 722–734.

Gui, Z., Liu, P., Cheng, L., Guo, S., Wang, H., Zhang, L., 2019. Improving Runoff Prediction Using Remotely Sensed Actual Evapotranspiration during Rainless Periods. Journal of Hydrologic Engineering, 24, 12, 04019050.

Hall, D.K., Riggs, G.A., 2016a. MODIS/Terra Snow Cover Daily L3 Global 500m Grid, Version 6. [January 2013 – December 2017]. Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center. DOI: 10.5067/MODIS/MOD10A1.006 [24 January 2018].

Hall, D.K., Riggs, G.A., 2016b. MODIS/Aqua Snow Cover Daily L3 Global 500m Grid, Version 6. [January 2013 – December 2017]. Boulder, Colorado USA, NASA National Snow and Ice Data Center Distributed Active Archive Center. DOI: 10.5067/MODIS/MYD10A1.006. [24 January 2018].

Hay, L.E., Leavesley, G.H., Clark, M.P., Markstrom, S.L., Viger, R.J., Umemoto, M., 2006. Step wise, multiple objective calibration of a hydrologic model for a snowmelt dominated basin. Journal of the American Water Resources Association, 42, 4, 877–890.

Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., Braatz, D., 2000. A multistep automatic calibration scheme for river forecasting models. Journal of Hydrometeorology, 1, 524–542.

Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., Cudennec, C., 2013. A decade of Predictions in Ungauged Basins (PUB) - a review. Hydrological Sciences Journal, 58, 6, 1198–1255.

Kuppel, S., Tetzlaff, D., Maneta, M.P., Soulsby, C., 2018. What can we learn from multi-data calibration of a process-based ecohydrological model? Environmental Modelling & Software, 101, 301–316.

Kuras, P.K., Alila, Y., Weiler, M., Spittlehouse, D., Winkler, R., 2011. Internal catchment process simulation in a snow-dominated basin: performance evaluation with spatiotemporally variable runoff generation and groundwater dynamics. Hydrological Processes, 25, 3187–3203.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology, 201, 1–4, 272–288.

López, L.P., Sutanudjaja, E.H., Schellekens, J., Sterk, G., Bierkens, M.F.P., 2017. Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. Hydrology and Earth System Sciences, 21, 3125–3144.

Lu, M., Li, X., 2015. Strategy to automatically calibrate parameters of a hydrological model: a multi-step optimization scheme and its application to the Xinanjiang model. Hydrological Research Letters, 9, 4, 69–74.

Merz, R., Blöschl, G., 2004. Regionalisation of catchment model parameters. Journal of Hydrology, 287, 95–123.

Merz, R., Parajka, J., Blöschl, G., 2011. Time stability of catchment model parameters: Implications for climate impact analyses. Water Resources Research, 47, 1015–1031.

Mullen, K., Ardia, D., Gil, D., Windover, D., Cline, J., 2011. 'DEoptim': An R Package for Global Optimization by Differential Evolution. Journal of Statistical Software, 40, 6, 1–26.

Murer, E., Wagenhofer, J., Aigner, F., Cline, J., 2004. Die nutzbare Feldkapazität der mineralischen Böden der landwirtschaftlichen Nutzfläche Österreichs. Schriftenreihe BAW, Band 20, 72–78.

Nijzink, R.C., Almeida, S., Pechlivanidis, I.G., Capell, R., Gustafssons, D., Arheimer, B., Parajka, J., Freer, J., Han, D., Wagener, T., van Nooijen, R.R.P., Savenije, H.H.G., Hrachowitz, M., 2018. Constraining conceptual hydrological models with multiple information sources. Water Resources Research, 54, 8332–8362.

Ning, S., Ishidaira, H., Wang, J., 2015. Calibrating a hydrologic model by step-wise method using GRACE TWS and discharge data. Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering), 71, 4, I_85-I_90.

Parajka, J., Merz, R., Blöschl, G., 2003. Estimation of daily potential evapotranspiration for regional water balance modeling in Austria. In: 11th International Poster Day and Institute of Hydrology Open Day "Transport of Water, Chemicals and Energy in the Soil - Crop Canopy - Atmosphere System", Slovak Academy of Sciences, Bratislava, pp. 299–306.

Parajka, J., Merz, R., Blöschl, G., 2007. Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments. Hydrological Processes, 21, 435–446.

Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M., Blöschl, G., 2013. Comparative assessment of predictions in ungauged basins - Part 1: Runoff hydrograph studies. Hydrology and Earth System Sciences, 17, 1783–1795.

Rogger, M., Kohl, B., Pirkl, H., Viglione, A., Komma, J., Kirnbauer, R., Merz, R., Blöschl, G., 2012. Runoff models and flood frequency statistics for design flood estimation in Austria - Do they tell a consistent story? Journal of Hydrology, 456–457, 30–43.

Savenije, H.H.G., 2001. Equifinality, a blessing in disguise? Hydrological Processes, 15, 2835–2838.

Schrödter, H., 1985. Verdunstung - Anwendungsorientierte Messverfahren und Bestimmungsmethoden. Springer, 186 p. ISBN: 978-3-642-70434-5.

Seibert, J., 2000. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrology and Earth System Sciences, 4, 215–224.

Silasari, R., Parajka, J., Ressl, C., Strauss, P., Blöschl, G., 2017. Potential of time - lapse photography for identifying saturation area dynamics on agricultural hillslopes. Hydrological Processes, 31, 3610–3627.

Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., Boni, G., 2015. Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data. Hydrology and Earth System Sciences, 19, 1727–1751.

Sleziak, P., Szolgay, J., Hlavcova, K., Danko, M., Parajka, J., 2020. The effect of the snow weighting on the temporal stability of hydrologic model efficiency and parameters. Journal of Hydrology, 583, 124639.

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, 9, 6168–6188.

Széles, B., Parajka, J., Hogan, P., Silasari, R., Pavlin, L., Strauss, P., Blöschl, G., 2020. The added value of different data types for calibrating and testing a hydrologic model in a small catchment. Submitted to Water Resources Research.

Thyer, M., Beckers, J., Spittlehouse, D., Alila, Y., Winkler, R., 2004. Diagnosing a distributed hydrologic model for two high-elevation forested catchments based on detailed stand- and basin-scale data. Water Resources Research, 40, W01103.

Viglione, A., Parajka, J., Rogger, M., Salinas, J.L., Laaha, G., Sivapalan, M., Blöschl, G., 2013. Comparative assessment of predictions in ungauged basins – Part 3: Runoff signatures in Austria. Hydrology and Earth System Sciences, 17, 2263–2279.

Viglione, A., Rogger, M., Pirkl, H., Parajka, J., Blöschl, G., 2018. Conceptual model building inspired by field-mapped runoff generation mechanisms. Journal of Hydrology and Hydromechanics, 66, 3, 303–315.

Received 1 April 2020
Accepted 20 July 2020