



# The value of satellite soil moisture and snow cover data for the transfer of hydrological model parameters to ungauged sites

Rui Tong<sup>1,2</sup>, Juraj Parajka<sup>1,2</sup>, Borbála Széles<sup>1,2</sup>, Isabella Greimeister-Pfeil<sup>1,3</sup>, Mariette Vreugdenhil<sup>3</sup>, Jürgen Komma<sup>2</sup>, Peter Valent<sup>2,4</sup>, and Günter Blöschl<sup>1,2</sup>

<sup>1</sup>Centre for Water Resource Systems, TU Wien, Vienna 1040, Austria

<sup>2</sup>Institute of Hydraulic Engineering and Water Resources Management, TU Wien, Vienna 1040, Austria

<sup>3</sup>Department of Geodesy and Geoinformation, TU Wien, Vienna 1040, Austria

<sup>4</sup>Department of Land and Water Resources Management, Slovak University of Technology in Bratislava, Bratislava 810 05, Slovakia

**Correspondence:** Rui Tong (tong@hydro.tuwien.ac.at)

Received: 2 April 2021 – Discussion started: 23 April 2021

Revised: 16 February 2022 – Accepted: 16 March 2022 – Published: 7 April 2022

**Abstract.** The recent advances in remote sensing provide opportunities for estimating the parameters of conceptual hydrologic models more reliably. However, the question of whether and to what extent the use of satellite data in model calibration may assist in transferring model parameters to ungauged catchments has not been fully resolved. The aim of this study is to evaluate the efficiency of different methods for transferring model parameters obtained by multiple-objective calibrations to ungauged sites and to assess the model performance in terms of runoff, soil moisture, and snow cover predictions relative to existing regionalization approaches. The model parameters are calibrated to daily runoff, satellite soil moisture (Advanced Scatterometer – ASCAT), and snow cover (Moderate Resolution Imaging Spectroradiometer – MODIS) data. The assessment is based on 213 catchments situated in different physiographic and climate zones of Austria. For the transfer of model parameters, eight methods (global and local variants of arithmetic mean, regression, spatial proximity, and similarity) are examined in two periods, i.e., the period in which the model is calibrated (2000–2010) and an independent validation period (2010–2014). The predictive accuracy is evaluated by the leave-one-out cross-validation. The results show that the method by which the model is calibrated in the gauged catchment has a larger impact on runoff prediction accuracy in the ungauged catchments than the choice of the parameter transfer method. The best transfer methods are global and local similarity and the kriging approach. The performance

of the transfer methods differs between lowland and alpine catchments. While the soil moisture and snow cover prediction efficiencies are higher in lowland catchments, the runoff prediction efficiency is higher in alpine catchments. A comparison of the model transfer methods, based on parameters calibrated to runoff, snow cover, and soil moisture with those based on parameters calibrated to runoff, only indicates that the former outperforms the latter in terms of simulating soil moisture and snow cover. The performance of simulating runoff is similar, and the accuracy depends mainly on the weight given to the runoff objective in the multiple-objective calibrations.

## 1 Introduction

The prediction of runoff hydrographs is one of the main scientific and practical hydrological applications. Understanding and representing the processes and their interactions during runoff generation and their sensitivity to the controls are some of the unsolved problems in hydrology (Blöschl et al., 2019). It is also essential for many purposes of societal relevance, such as engineering design, water resources management, hydropower operation, or risk assessment (Sachs and McArthur, 2005; Blöschl and Montanari, 2010; Kovacs et al., 2012; Parajka et al., 2013). However, in most places of interest, no runoff data are directly observed, so the runoff predic-

tion needs to be based on additional information observed in the region or transferred from other gauged catchments.

Numerous studies have explored and evaluated methods for the prediction of runoff in ungauged sites (He et al., 2011; Hrachowitz et al., 2013; Blöschl et al., 2013). The most frequently used method consists of applying a hydrological model driven by model parameters derived in catchments with runoff observations. The review and synthesis of studies and applications of hydrological models in ungauged sites presented in Parajka et al. (2013), and recently in Guo et al. (2021), indicate that there is a variety of transfer approaches tested in different parts of the world, covering different climates, altitudes, or landscape settings. The main outcome of these studies is that the runoff predictions in ungauged catchments tend to be more accurate in humid than in arid regions and more accurate in large than in small catchments (Parajka et al., 2013). However, the selection and performance of methods differ between studies, and there is no general recommendation for the choice of the approach. There seems to be a consensus that the spatial proximity and similarity methods perform better in humid regions (Guo et al., 2021). In arid catchments, the similarity and parameter regression methods tend to be applied more frequently and perform slightly better (Parajka et al., 2013; Yang et al., 2017, 2020).

Recent studies have explored the role and impact of the gauge density on the efficiency of the methods (Parajka et al., 2015; Lebecherel et al., 2016; Neri et al., 2020), as well as the advanced definition of similarity measures based on spatial patterns (Li and Zhang, 2017; Beck et al., 2020; NARBONDO et al., 2020) or catchment response characteristics (Tegegne and Kim, 2018). Along with these investigations, one of the recent focuses in hydrological modeling evaluates the use of observations in addition to runoff (Bouaziz et al., 2021). Multiple-objective calibrations can help constrain hydrological models, reduce uncertainty, and improve hydrological predictions (Efstratiadis and Koutsoyiannis, 2010; Rakovec et al., 2016; Dembélé et al., 2020). Most of the previous studies investigated the potential of calibrating hydrological model parameters by using different runoff signatures or using some additional hydrological characteristics such as snow cover, cover, soil moisture, evaporation, groundwater level, or combinations thereof (please see the review in Tong et al., 2021). However, only a few studies have investigated the application of multiple-objective approaches for evaluating the transfer of model parameters to ungauged sites. Parajka et al. (2005, 2006) examined the value of snow depth observation and scatterometer soil moisture measurements for improving the hydrological simulations in ungauged catchments and showed that the use of the scatterometer resulted in more consistent patterns of soil moisture estimates, but it did not improve the runoff model efficiencies in ungauged catchments. Zhang et al. (2020) recently used remotely sensed evapotranspiration to calibrate the hydrological model parameters without streamflow observations and found that the streamflow-free calibration could be sat-

isfactory for monthly and mean annual runoff simulation in the humid catchments. Huang et al. (2020) verified the effectiveness of this approach for the regionalization with spatial proximity method in ungauged basins. They reported that using the bias-corrected remotely sensed evapotranspiration has great potential in estimating daily and monthly runoff. However, the role and impact of using additional data for the prediction of daily hydrographs in ungauged sites are still not well understood. Such an assessment is essential, particularly for projecting the impacts of changing climate or land use conditions on the hydrological cycle, in general, and, specifically, on runoff generation. The aim of this study is to investigate the value of using model parameters obtained by multiple-objective calibrations for daily hydrological predictions in ungauged sites. Specifically, we test the efficiency of different methods for the transfer of such model parameters to ungauged sites and evaluate the model performance to observations of runoff and remotely sensed estimates of soil moisture and snow cover. We extend the results of Tong et al. (2021), who assessed the value of Metop ASCAT (Advanced Scatterometer) and MODIS (Moderate Resolution Imaging Spectroradiometer) satellites for the calibration of conceptual hydrological models and to test to what extent it improves the performance of existing regionalization approaches. The analysis and comparison of model simulations of runoff, soil moisture, and snow cover are performed for a large sample of catchments situated in different physiographic and climate zones of Austria, so this allows the evaluation of the potential of the application of remotely sensed data at the regional scale. The effect of multi-objective calibration with the use of satellite soil moisture and snow cover data on model parameter transferability is hence assessed.

## 2 Method

### 2.1 Hydrological model

The transfer of the model parameters is evaluated for a conceptual hydrological model (TUWmodel; Viglione and Parajka, 2020). The TUWmodel is a variant of the HBV model (Bergström, 1992; Parajka et al., 2007) implemented in the R environment (Astagneau et al., 2021). This study uses a semi-distributed version in which the inputs and outputs of the model are processed for elevation zones of 200 m but model parameters are assumed to be lumped in each catchment. Such a setting allows one to effectively account for the spatial variability in rainfall and melt processes, particularly in alpine regions, and to keep the number of parameters small, which simplifies their transfer to ungauged sites. For comparison with soil moisture satellite data, we used a single soil (root zone) layer version of TUWmodel. It simulates changes in snow, root zone, and groundwater storages in each elevation zone. The model runs on a daily time step and combines the following three routines: snow routine, soil mois-

ture routine, and river flow routing routine. The snow routine uses a degree–day concept to reflect snow accumulation and melt with a degree–day factor and a threshold melt temperature. The snowfall part of the precipitation and snow accumulation is calculated by using snow and rain threshold temperatures. The soil moisture routine represents the changes in the soil moisture state of the root zone due to evapotranspiration and runoff generation.

$$S_{SM,i} = S_{SM,i-1} + (PR + M - E_A) \cdot \Delta t, \quad (1)$$

where  $S_{SM}$  is the root zone soil moisture, which controls runoff generation and actual evaporation  $E_A$ ,  $PR$  is rain,  $M$  is snowmelt, and  $\Delta t$  is the time step (1 d). The contribution  $\Delta S_{UZ}$  of rain and snowmelt to runoff is calculated by an explicit scheme, as a function of the  $S_{SM}$ , using a nonlinear relationship controlled by two model parameters, namely maximum soil storage  $FC$  and a nonlinearity parameter  $BETA$  controlling characteristics of runoff generation, as follows:

$$\Delta S_{UZ} = (\theta)^{BETA} (PR + M) \cdot \Delta t \quad (2)$$

$$\theta = \frac{S_{SM}}{FC}. \quad (3)$$

Actual evapotranspiration is estimated from potential evaporation (model input) and a model parameter representing the soil moisture state above which the actual equals potential evaporation. For a comparison with satellite soil moisture estimates, simulated root zone soil moisture is scaled by the maximum soil moisture storage  $FC$  model parameter and hence represents the relative root zone soil moisture  $\theta$ .

The runoff routing module represents routing on the hillslopes and river flow routing in the stream. The runoff response function consists of two reservoirs representing the upper and lower storage zones. The outflows from the reservoirs in each elevation zone are summed up and routed by a triangular transfer function.

The model involves 15 model parameters. They are automatically calibrated using the multi-objective calibration strategy of Tong et al. (2021). The joint objective function  $O_F$  consists of weighting three individual parts related to runoff ( $O_Q$ ), soil moisture ( $O_{SM}$ ), and snow cover ( $O_{SC}$ ).

$$O_F = w_Q \cdot O_Q + w_{SM} \cdot O_{SM} + w_{SC} \cdot O_{SC}, \quad (4)$$

where  $w_Q$  is the weight to the runoff objective function, and  $w_{SM}$  and  $w_{SC}$  are the weights to the soil moisture and snow cover objectives, respectively. Multiple-objective approaches for each regionalization method are examined for 11 runoff weights  $w_Q$  (ranging from 0.0 to 1.0 with a step of 0.1). The combinations of the weights are taken from Tong et al. (2021), where the weights of soil moisture and snow are equal, and the total sum of all weights is 1.0. More details about the combinations of the objective functions are presented in Tong et al. (2021).

The runoff objective function  $O_Q$  emphasizes both high and low flows (Parajka and Blöschl, 2008) and is described by a compound objective function combining two variants of Nash–Sutcliffe coefficients, i.e., it is estimated from observed and logarithmic transformed river flow values, as follows (Nash and Sutcliffe, 1970):

$$O_Q = 0.5 \cdot NSE + 0.5 \cdot NSE_{\log} \quad (5)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2} \quad (6)$$

$$NSE_{\log} = 1 - \frac{\sum_{i=1}^n (\log(Q_{obs,i}) - \log(Q_{sim,i}))^2}{\sum_{i=1}^n (\log(Q_{obs,i}) - \log(\bar{Q}_{obs}))^2}, \quad (7)$$

where  $Q_{obs,i}$  and  $Q_{sim,i}$  are the river flow observation and simulation of day  $i$ .

The measure of soil moisture agreement ( $O_{SM}$ ) is determined by the Pearson correlation coefficient ( $O_{SM}$ ) between the satellite soil water index (SWI), which estimates the saturation degree of the root zone (Wagner et al., 1999b; Silvestro et al., 2015), and the simulated relative soil moisture in each elevation zone, as in the following:

$$O_{SM} = \frac{\sum_{i=1}^{N_{days}} \sum_{j=1}^{N_{zones}} ((\theta_{sim,i,j} - \bar{\theta}_{sim})(\theta_{obs,i,j} - \bar{\theta}_{obs}))}{\sqrt{\sum_{i=1}^{N_{days}} \sum_{j=1}^{N_{zones}} ((\theta_{sim,i,j} - \bar{\theta}_{sim})^2 (\theta_{obs,i,j} - \bar{\theta}_{obs})^2)}}, \quad (8)$$

where  $\theta_{sim,i,j}$  is the simulated relative soil moisture from the hydrological model, and  $\theta_{obs,i,j}$  is the averaged value of the observed soil water index (SWI) from ASCAT pixels for the day  $i$  and elevation zone  $j$ .  $\bar{\theta}_{sim}$  and  $\bar{\theta}_{obs}$  are the mean values of the simulations and observations for the days and elevation zones which are not masked in the ASCAT SWI product due to presence of snow or frozen ground. The rationale behind selecting the Pearson correlation as a measure of agreement is that it assesses the spatial and temporal correspondence of the satellite soil moisture (assumed to be the ground truth) and simulated root zone soil moisture time series. At the spatial resolution of the original ASCAT dataset (ca. 12.5 km), the satellite estimates of root zone soil moisture reflect the regional rainfall and melt processes patterns more and are thus more closely related to altitudinal zonality than to morphometric characteristics of terrain that operate at smaller scales. The calculation of  $O_{SM}$  from soil moisture averages for elevation zones thus allows the representation of the agreement in regional and seasonal soil moisture patterns. The choice of a correlation coefficient has the advantage of not being sensitive to the units. In a preliminary analysis, we tested different ways for calculation of the  $O_{SM}$  as combining soil moisture estimated from different elevation zones better describes the agreement than the correlation between soil moisture estimates averaged at the catchments scale (see Fig. S3 in the Supplement). Particularly in the alpine regions, the correlation calculated from the catchment averages masks the spatial variability in the agreement between ASCAT and hydrologic root zone soil moisture estimates. A similar approach

has also been used in previous studies (e.g., Parajka et al., 2006; Gruber et al., 2020; Beck et al., 2021). A more detailed description of the calculation of the soil moisture agreement is presented in the Supplement.

The snow cover objective function  $O_{SC}$  minimizes the sum of the snow overestimation  $S_O$  and underestimation  $S_U$  errors as follows (Parajka and Blöschl, 2008):

$$O_{SC} = 1 - (S_O + S_U). \quad (9)$$

The snow overestimation error shows the percentage of the condition if the snow is simulated from the model, but the snow cover is not retrieved by the satellite (MODIS).

$$S_O = \frac{1}{\sum_{i=1}^{N_{\text{days}}} \sum_{j=1}^{N_{\text{zones}}} A_{i,j}} \times \sum_{i=1}^{N_{\text{days}}} \sum_{j=1}^{N_{\text{zones}}} A_{i,j} \cap (\text{SWE}_{i,j} > \xi_{\text{SWE}}) \cap (\text{SCA}_{i,j} = 0), \quad (10)$$

where  $A_{i,j}$  is the area of zone  $j$  that is cloud free, according to the MODIS observation on day  $i$ .  $\text{SWE}_{i,j}$  is the simulated snow water equivalent in elevation zone  $j$  greater than a threshold ( $\xi_{\text{SWE}}$ ) 10 mm,  $\text{SCA}_{i,j}$  is the MODIS snow covered area within this zone, and  $N_{\text{days}}$  is the number of days with cloud cover that is less than a threshold ( $\xi_C$ ) 50 %.

The snow underestimation error indicates the percentage of the condition if no snow is simulated, but snow cover retrieved by MODIS is over a threshold ( $\xi_{\text{SCA}}$ ) of 25 % in the zone, as in the following:

$$S_U = \frac{1}{\sum_{i=1}^{N_{\text{days}}} \sum_{j=1}^{N_{\text{zones}}} A_{i,j}} \times \sum_{i=1}^{N_{\text{days}}} \sum_{j=1}^{N_{\text{zones}}} A_{i,j} \cap (\text{SWE}_{i,j} = 0) \cap (\text{SCA}_{i,j} > \xi_{\text{SCA}}). \quad (11)$$

The thresholds of  $\xi_{\text{SWE}}$ ,  $\xi_C$ , and  $\xi_{\text{SCA}}$  were determined by the sensitivity analysis of Parajka and Blöschl (2008).

## 2.2 Transfer of model parameters to ungauged sites

The hydrological predictions at the ungauged sites are, in this study, based on model simulations driven by model parameters transferred from the gauged locations where the model has been calibrated, i.e., the sites with runoff observations. The model performances in terms of runoff, soil moisture, and snow with the calibrated parameters in the catchments assumed to be gauged can refer to Tong et al. (2021). For the transfer (i.e., regionalization), four groups of methods are evaluated (Table 1). The first group estimates the model parameters as the arithmetic mean of all calibrated values in the study region (termed the global mean) or, alternatively, as the arithmetic mean of the model parameters within a radius of 50 km from the catchment of interest (termed the local mean). This arithmetic mean regionalization approach as-

**Table 1.** The list of model parameter transfer methods tested in the study.

Group	Transfer method	Abbr.
Arithmetic mean	Global mean	GM
	Local mean	LM
Regression	Global regression	GR
	Local regression	LR
Spatial proximity	Inverse distance weighting	ID
	Kriging	KR
Similarity	Global combination of physiographic attributes	GS
	Local combination of physiographic attributes	LS

sumes the similarity of all catchments within a specified radius, where differences in the parameter values are caused only by random factors.

In the second group, the model parameters for ungauged sites are independently estimated from linear regressions between calibrated model parameters and catchment attributes. Similarly, as in the first group, two approaches are tested. The global multiple linear regression uses attributes and model parameters from all gauged catchments. The local multiple linear regression is applied within a 50 km search radius from an ungauged site. In all cases, the regression coefficients are estimated by the ordinary least squares method. For consistency with previous studies, a set of three catchment attributes associated with the largest multiple correlation coefficient for each ungauged site and each model parameter is used. To avoid multicollinearity, the variance inflation factor (Hirsch et al., 1992) is examined. For the transfer of model parameters, such a regression model is used, which has the largest correlation coefficient for the inflation factor less than 10.

The third group of transfer methods is based on the spatial proximity (or spatial distance) between the ungauged and the gauged catchments. The spatial distance between the two catchments is characterized by the distance between the respective catchment centroids. We test the following two methods of this group: the inverse distance weighting and the ordinary kriging. In both methods, individual parameters from several donor catchments are independently interpolated to a centroid of the ungauged catchment and then combined and used in the hydrological model. The power parameter in inverse distance interpolation is set to two. The ordinary kriging method is based on a fixed exponential variogram with a nugget of 10 % of the observed variance, a sill equal to the variance, and a range of 60 km. The test calculations in previous studies (Merz and Blöschl, 2004; Parajka et al., 2005) showed that this setting is consistent with the

empirical variograms of most of the calibrated model parameters.

The fourth group of methods is based on the similarity between catchments with runoff observations and ungauged sites. The main idea of the similarity group of methods is to find, for an ungauged site, a donor catchment that is most similar in terms of certain catchment attributes. The entire collection of model parameters calibrated for a donor catchment is then transferred to the ungauged site. The similarity is defined by a similarity index  $\Phi$  as follows (Burn and Boorman, 1993; Merz and Blöschl, 2004):

$$\Phi = \sum_{i=1}^k \frac{|X_i^G - X_i^U|}{\Delta X_i}, \quad (12)$$

where  $X^G$  represents a vector of the normalized catchment attributes of the gauged (donor) catchments,  $X^U$  are the normalized attributes of the ungauged catchment, and  $\Delta X$  is the normalized range of attributes. In previous studies (e.g., Parajka et al., 2005) a large number of catchment attributes, and their combinations, have been tested in the study region. Based on the results and preliminary analyses (not shown here) for this study, we selected the approach with the best results. This variant is based on an a priori defined combination of the following catchment attributes: mean catchment elevation, stream network density, lake attenuation index and areal proportion of porous aquifers, land use, soils, and geologic units. Similar to other approaches, in this approach, we also examine two variants. While the global similarity combination uses all study catchments for the estimation of the similarity, the local similarity combination estimates the similarity only within a 50 km radius around the ungauged site.

### 2.3 Evaluation of the prediction accuracy

The performance and efficiency of parameter transfer methods are evaluated by the leave-one-out cross-validation. Each catchment with observed runoff is considered, in turn, as ungauged, and the transfer methods are used to estimate the parameter sets from other gauged catchments. The hydrological model is applied to simulate daily runoff, soil moisture, and snow cover in the ungauged catchment. These simulations are then compared with the observations. The accuracy is quantified by three objectives  $O_Q$ ,  $O_{SM}$ , and  $O_{SC}$  in two periods, i.e., in the period used for model calibration (2000–2010 and 2007–2010 for  $O_{SM}$ ) and in an independent validation period (2010–2014). The efficiencies of the transferred model parameters are estimated for 11 different calibration variants (i.e., the weight given to runoff) and compared to the efficiency obtained by a transfer of model parameters calibrated to runoff only.

## 3 Data

### 3.1 Study region

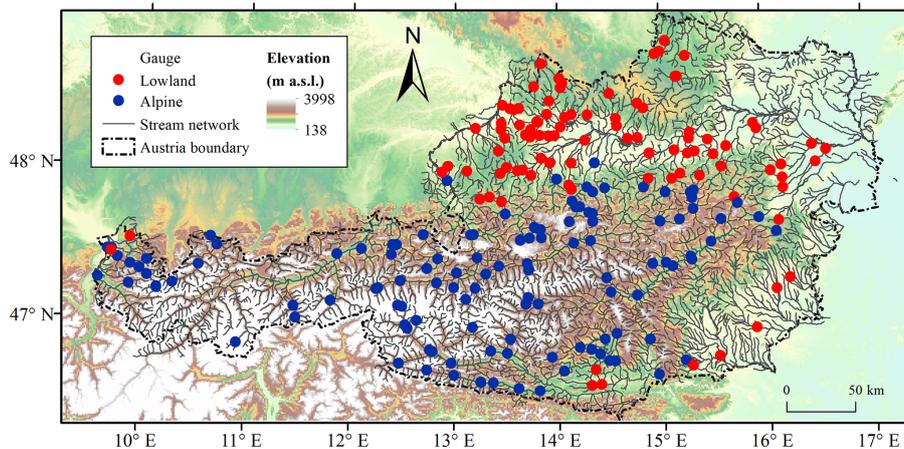
The study region is Austria, which represents a wide range of physiographic conditions. The topography varies from flat land in the east and north to alpine terrain in the west and south. Mean annual precipitation is less than 400 mm yr<sup>-1</sup> in the east and more than 2500 mm yr<sup>-1</sup> in the west. Land use is mainly agricultural in the lowlands and forest in the medium elevation ranges. Alpine vegetation and rocks prevail in the highest alpine regions.

The analysis is carried out for 213 catchments (Fig. 1). These catchments have been selected following previous studies (Vigliante et al., 2013; Slezziak et al., 2020; Tong et al., 2021) and represent catchments with no significant anthropogenic effects on the water balance. The size of the catchments ranges from 13.7 to 6214 km<sup>2</sup>, and the averaged slope varies from 1.74 % to 43.91 %. As previous studies have shown that the performance of regionalization methods differs between climatic zones (Parajka et al., 2013; Yang et al., 2020), separation of the effect of elevation and climate on the results was deemed important, and the catchments were split into two groups representing drier lowland and hilly regions (catchments with mean elevation below 900 m a.s.l. – above sea level) and wetter alpine conditions (catchments with mean elevation above 900 m a.s.l.). Out of the 213 catchments, 94 are classified as lowland catchments and 119 as alpine catchments (Fig. 1). The climatic statistics of the two groups are presented in Table 2. The threshold of 900 m is chosen as a compromise between balancing the number of catchments in the groups and representing different physiographic regions.

### 3.2 Hydrologic and climate data

The runoff data have been obtained from Central Hydrographical Bureau (HZB; <https://ehyd.gv.at/>, last access: 6 April 2022). The analysis period is from September 2000 to August 2014, which is split into the calibration (September 2000–August 2010) and validation periods (September 2010–August 2014).

Model inputs (i.e., the mean of daily climate characteristics for elevation zones) are derived from the SPARTACUS gridded dataset (Hiebl and Frei, 2016, 2018). This dataset includes gridded maps of the maximum and minimum daily air temperature and precipitation with a spatial resolution of 1 km. Daily mean air temperature is estimated as the mean between minimum and maximum air temperature. Potential evaporation model input is estimated by the Blaney–Criddle approach (Parajka et al., 2003). This approach estimates the potential evaporation from mean daily air temperature and a potential sunshine duration index, which is calculated from a 1 km digital elevation model of Austria.



**Figure 1.** Topography of the study region and location of 94 lowland (red symbols) and 119 alpine (blue symbols) catchments.

**Table 2.** Statistics of the climatic attributes of the 94 lowland catchments and 119 alpine catchments given with the abbreviation, unit, minimum, maximum, and median. The standard deviations refer to the spatial variability within each catchment.

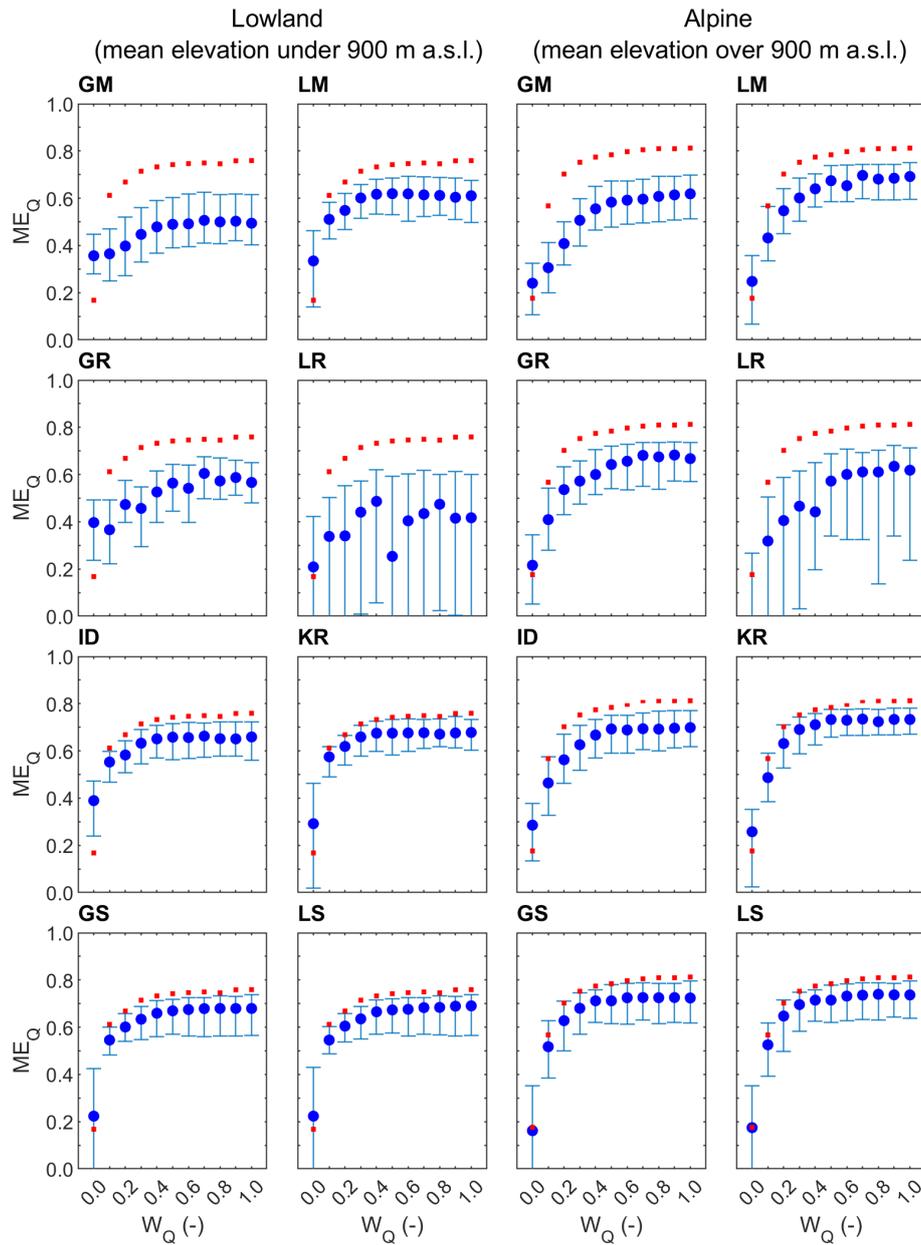
Attribute	Abbr.	Unit	Lowland (mean elevation under 900 m a.s.l.)			Alpine (mean elevation over 900 m a.s.l.)		
			Min.	Max.	Median	Min.	Max.	Median
Mean annual precipitation	MAP	mm	728.13	1828.40	999.46	913.66	2301.84	1476.64
Standard deviation of annual MAP	SDAP	mm	10.79	367.57	71.49	30.13	289.87	152.90
Mean air temperature	MAT	°C	7.26	10.30	8.98	-2.83	8.07	5.76
Standard deviation of MAT	SDAT	°C	0.06	1.71	0.57	0.40	3.55	1.64
Mean annual potential evaporation	MEPI	mm	618.36	740.45	690.08	233.49	657.01	563.00
Standard deviation of MEPI	SDEPI	mm	4.33	77.41	25.25	21.70	162.07	83.33
Catchment aridity index (MEPI/MAP)	CAI	–	0.36	0.98	0.66	0.18	0.69	0.37
Standard deviation of aridity index	SDAI	–	0.01	0.31	0.06	0.02	0.18	0.09

### 3.3 MODIS snow cover

The snow cover maps used in model calibration and regionalization validation are based on the combination of the daily, 500 m resolution, Terra (MOD10A1) and Aqua (MYD10A1) MODIS datasets (Hall and Riggs, 2016). We use the latest collection of six snow cover products, which includes the Normalized Difference Snow Index (NDSI). Snow cover mapping from MODIS products is performed in two steps. In the first step, NDSI pixels are classified into snow and land cover classes, based on seasonally varying NDSI thresholds (Tong et al., 2020). In the second step, the resulting snow cover maps from the Aqua and Terra products are combined to reduce the impact of clouds (Parajka and Blöschl, 2008). Finally, for each elevation zone of each catchment, the frequency of pixels classified as clouds, snow, and land is calculated. This allows for the estimation of the snow cover area percentage for each catchment that is needed for the calculation of the snow cover model objective function.

### 3.4 ASCAT soil moisture

The satellite soil moisture data used in this study is the soil water index (SWI) derived from an experimental version of the upcoming disaggregated Metop ASCAT surface Soil Moisture v2 product (H28) provided by the EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF). The original ASCAT surface soil moisture dataset at 12.5 km (before disaggregation) is based on a new parameterization for the vegetation correction (Hahn et al., 2021), which has shown improved performance over Austria (Pfeil et al., 2018). The disaggregation process consists of a directional resampling method utilizing the connection between regional- (12.5 km) and local-scale (0.5 km) Sentinel-1 backscatter observations describing temporally stable soil moisture patterns also reflected in the radar backscatter measurements (Wagner et al., 2008). Surface and root zone soil moisture are available, where the root zone soil moisture is represented by the SWI, which is determined by an exponential filter, introduced by Wagner et al. (1999a, b) and Albergel et al.



**Figure 2.** Median and regional scatter of the leave-one-out runoff model efficiency (Eq. 5) obtained by eight groups of parameter transfer methods and 11 calibration weights for lowland (94) and alpine (119) catchments in the calibration (2000–2010; blue symbols) period. Blue circles represent the median and whiskers the 25 % and 75 % quantiles of the leave-one-out efficiency. Red symbols show the median of the at-site runoff model efficiency.

(2008), with a characteristic time lag ( $T$ ). The  $T$  value represents the smoothing of soil moisture dynamics by infiltration, with higher  $T$  values corresponding to a higher degree of smoothing. In order not to lose information on the short-term soil moisture dynamics still present in deeper soil layers,  $T$  should be carefully chosen. Paulik et al. (2014) compared the ASCAT SWI dataset to in situ soil moisture and found that the SWI agrees better with in situ soil moisture from deeper layers than the original surface soil moisture dataset. More-

over, they related the  $T$  value with soil depth layers and found that the  $T$  values of 10 and 20 led to the highest correlations in the shallow subsurface (around 0–20 cm). To prevent the loss of short-term soil moisture dynamics,  $T$ -value = 10 d was selected in this study. Besides, to exclude invalid ASCAT measurements affected by snow and frozen ground, soil moisture is masked as no data when soil temperatures at a soil depth of 0–7 cm are below 1 °C or snow cover exceeds

30 % of the pixel with the information from the ECMWF Copernicus Climate Service (C3S) ERA5-Land.

## 4 Results

### 4.1 Efficiency of transfer methods for simulating runoff

Figures 2 and 3 show the median and scatter (i.e., 25 % and 75 % quantiles) of the leave-one-out runoff cross-validation efficiency for eight parameter transfer methods (panels) and 11 calibration variants (i.e., different runoff weight  $w_Q$  used in model calibration) in the calibration (Fig. 2) and validation (Fig. 3) periods. The red symbols indicate the at-site runoff efficiency, estimated in Tong et al. (2021), in the calibration and validation periods. Panels on the left and right show the results for the lowland and alpine group of catchments, respectively. The results for the runoff weight  $w_Q = 1.0$  represent the case when the model is calibrated to runoff only. The case  $w_Q = 0.0$  represents the case when the model is calibrated to satellite soil moisture and snow cover without using observed runoff.

The results show that the differences between the transfer methods are smaller than those between the different calibration variants, i.e., the different methods of calibrating the model in gauged catchments, for weights below 0.4. The impact of the choice of calibration variant (weight on runoff  $w_Q$ ) is smaller if the  $w_Q$  is larger than 0.4. For  $w_Q$  larger than 0.4, the differences between the transfer methods are larger, and the choice of transfer method is more important than that of the calibration variant. The worst parameter transfer (i.e., regionalization) methods are the global mean and the local regression approach. The median of runoff efficiency is particularly low, i.e., between 0.24 and 0.41, for calibration variants using  $w_Q < 0.3$  in the calibration period. If  $w_Q$  is larger than 0.7, the median of the runoff efficiency of global mean and local regression is between 0.42 and 0.5 for the lowland and between 0.61 and 0.63 for the alpine catchments. The best transfer methods are global and local similarity and kriging interpolation. If the  $w_Q$  is larger than 0.4, then the median efficiency is between 0.67 and 0.69 in the lowland and between 0.71 and 0.74 in the alpine region. The efficiency of the transfer of model parameters calibrated by multiple-objective approaches (for  $w_Q > 0.4$ ) for the similarity and kriging methods is the same as that for the transfer of model parameters obtained by calibration to runoff only ( $w_Q = 1$ ). The reason for the similar runoff efficiency of the regionalization methods between  $w_Q = 0.4$  to 1.0 is that the runoff model efficiencies of the donor catchments do not change obviously when the satellite soil moisture and snow cover were both included in the calibration with this  $w_Q$  range (Tong et al., 2021). In the validation period (Fig. 3), the median of multiple-objective calibrations ( $w_Q = 0.8$ ) of kriging in the lowland and similarity in the alpine catchments is even larger

than the runoff efficiency obtained by a transfer (kriging or similarity) based on parameters calibrated to runoff only (variant  $w_Q = 1$ ). Additionally, Table 3 also shows that the regional variability (runoff model efficiency between catchments) is small for kriging, while being large for local regression methods. A comparison of local and global variants of the transfer methods indicates that the local methods are only slightly better than the global methods in terms of runoff efficiency. The largest difference between the local and global methods occurs for the mean approach, but the runoff efficiency of the local mean is noticeably lower than for the spatial proximity or similarity approaches. An exception is the regression of model parameters, which has a larger runoff efficiency for the global than the local approach. The reason is a larger correlation between model parameters and catchment attributes estimated from all catchments. For example, for  $w_Q = 0.4$ , the median of the correlation between model parameters and catchment attributes for the local regression varies between 0.22 and 0.65. For the global regression approach, the median is larger and varies between 0.70 and 0.88. The results also show that the performance of transfer methods is better in the alpine than in the lowland catchments. In both groups, however, similarity and kriging are the best approaches for predicting daily runoff.

### 4.2 Efficiency of transfer methods in simulating soil moisture

The evaluation of eight parameter transfer methods in simulating root zone soil moisture is presented in Figs. 4 and 5. The results show the median and scatter (i.e., 25 % and 75 % quantiles) of the correlation between the satellite root zone soil moisture index and simulated relative root zone soil moisture in the lowland (left two panels) and alpine (right two panels) catchments. The comparison of the different transfer methods and calibration variants indicates that the difference between the transfer methods is similar to that found for the prediction of daily runoff. The best transfer methods are kriging and similarity (local and global) approaches. In the lowland catchments, the median of soil moisture correlation ranges between 0.62 and 0.70 for these methods. The impact of the calibration variants is, for each transfer method, smaller than what is found for runoff. Generally, the correlation increases with decreasing  $w_Q$ , and the soil moisture agreement tends to be larger if the  $w_Q$  is smaller than 0.4. A much smaller agreement (correlation) between soil moisture estimates is found in the alpine catchments; this may likely be due to the heterogeneity in temperature and snow cover in mountainous regions when the soil moisture is retrieved from the satellite (Tong et al., 2021). The best transfer methods in alpine catchments are local and global similarity and kriging. The median correlation between modeled and satellite soil moisture is, however, small and varies between 0.14 and 0.22. From Table 4, the regional variability of soil moisture correlation is simi-

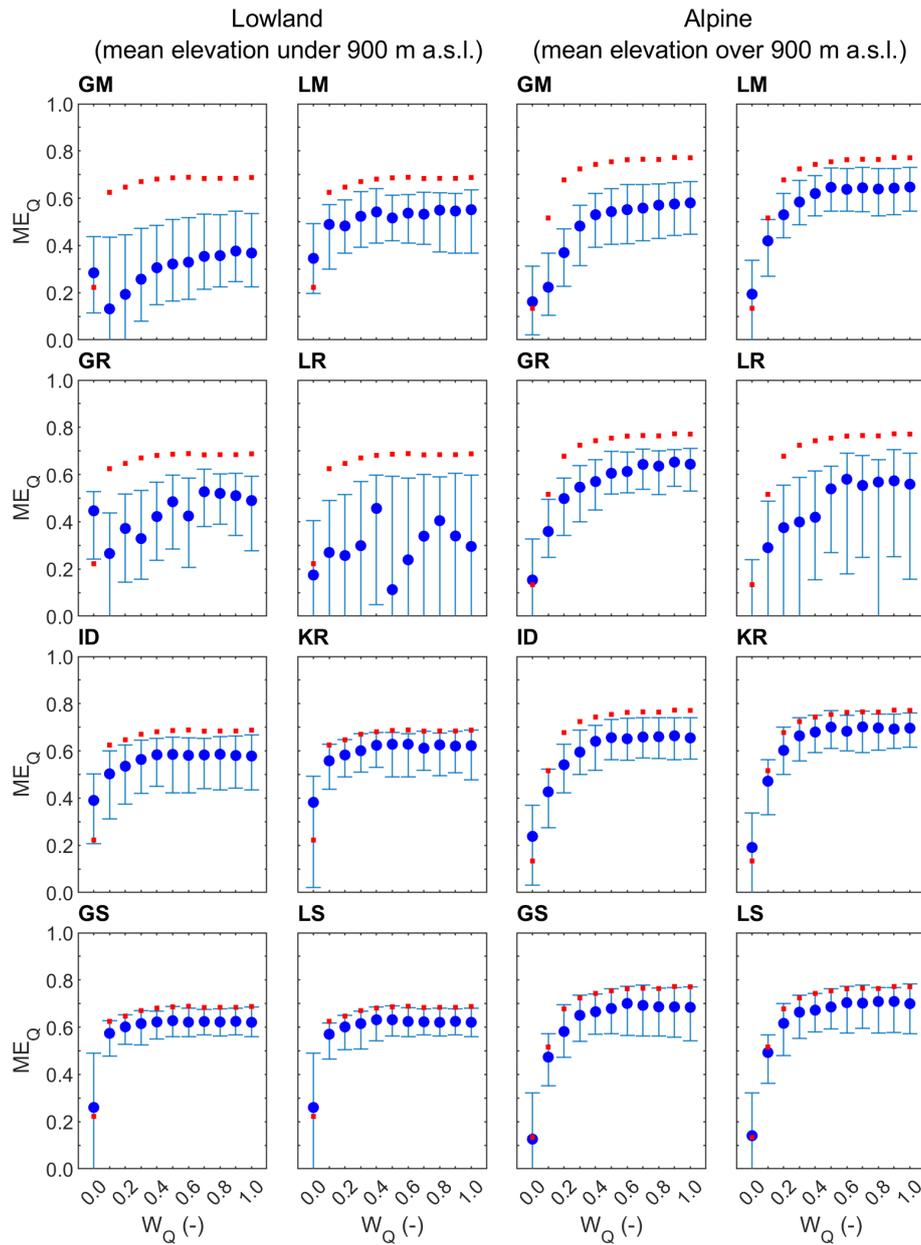


Figure 3. Same as Fig. 2 but for the validation period (2010–2014).

lar for each method and larger in the lowlands than that in the alpine regions. The comparison between the correlations of the calibration and validation periods shows a similar pattern. Interestingly, in the alpine catchments, the validation period correlations are slightly higher than those found for the transfer in the calibration period. This is likely related to the warmer validation period. Warming decreases the snow cover area, particularly in the alpine regions, and hence decreases the frequency of pixels which need to be masked in the soil moisture dataset.

### 4.3 Efficiency of transfer methods for simulating snow cover

The efficiency of eight transfer methods for simulating snow cover is evaluated in Figs. 6 and 7. The results indicate that the variability and differences between the regionalization approaches are the smallest for snow efficiency. A much larger difference and impact on snow efficiency is found for the runoff weight used in the model calibration. The difference in snow efficiency between transfer methods for the same  $w_Q$  is mostly between 1 % and 3 %, but the snow efficiency for different  $w_Q$  (and the same transfer method)

**Table 3.** Scatter (difference of 75 % and 25 % quantiles in model efficiency) of the leave-one-out runoff model efficiency (Eq. 5) obtained by eight groups of parameter transfer methods and 11 calibration weights for lowland (94) and alpine (119) catchments in the calibration (2000–2010; first value) and validation (2010–2014; second value) periods.

	$w_Q$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Lowlands	GM	0.17/0.32	0.22/0.56	0.25/0.47	0.23/0.39	0.22/0.34	0.21/0.34	0.22/0.35	0.21/0.32	0.21/0.31	0.20/0.30	0.21/0.31
	LM	0.32/0.29	0.15/0.27	0.15/0.22	0.14/0.24	0.15/0.23	0.15/0.19	0.19/0.21	0.17/0.22	0.16/0.25	0.17/0.25	0.18/0.27
	GR	0.26/0.28	0.27/0.66	0.18/0.37	0.25/0.37	0.22/0.33	0.20/0.31	0.24/0.38	0.18/0.24	0.17/0.21	0.15/0.26	0.17/0.32
	LR	0.94/1.82	0.51/0.85	0.88/1.45	0.56/0.72	0.56/0.55	1.10/1.61	0.63/1.05	0.68/1.15	0.58/0.86	0.61/0.84	0.61/0.84
	ID	0.23/0.30	0.13/0.29	0.13/0.25	0.14/0.22	0.14/0.20	0.15/0.23	0.15/0.23	0.15/0.21	0.15/0.23	0.15/0.22	0.16/0.23
	KR	0.45/0.47	0.13/0.19	0.12/0.16	0.13/0.16	0.13/0.15	0.15/0.19	0.14/0.18	0.12/0.17	0.12/0.18	0.13/0.18	0.13/0.21
	GS	0.50/0.57	0.12/0.15	0.12/0.13	0.14/0.15	0.15/0.12	0.15/0.13	0.16/0.12	0.17/0.11	0.17/0.11	0.17/0.11	0.17/0.13
	LS	0.53/0.57	0.12/0.15	0.12/0.14	0.14/0.16	0.15/0.14	0.15/0.13	0.16/0.12	0.16/0.11	0.17/0.12	0.17/0.11	0.17/0.12
Alpine	GM	0.22/0.29	0.21/0.26	0.18/0.24	0.20/0.26	0.18/0.23	0.19/0.24	0.18/0.25	0.19/0.24	0.19/0.23	0.19/0.22	0.18/0.22
	LM	0.29/0.35	0.23/0.24	0.19/0.19	0.18/0.19	0.14/0.17	0.15/0.18	0.15/0.18	0.15/0.19	0.15/0.20	0.15/0.19	0.15/0.19
	GR	0.29/0.35	0.26/0.25	0.20/0.24	0.18/0.24	0.19/0.21	0.18/0.18	0.19/0.17	0.19/0.17	0.20/0.19	0.16/0.16	0.17/0.18
	LR	0.96/0.99	0.51/0.59	0.61/0.61	0.58/0.61	0.45/0.46	0.35/0.37	0.38/0.51	0.37/0.43	0.57/0.71	0.38/0.45	0.48/0.53
	ID	0.24/0.34	0.25/0.25	0.21/0.21	0.19/0.19	0.16/0.19	0.16/0.17	0.16/0.18	0.15/0.17	0.17/0.17	0.16/0.18	0.15/0.17
	KR	0.33/0.46	0.21/0.23	0.18/0.20	0.15/0.18	0.13/0.16	0.12/0.16	0.11/0.15	0.11/0.17	0.11/0.15	0.11/0.15	0.11/0.14
	GS	0.48/0.59	0.24/0.22	0.21/0.22	0.17/0.20	0.14/0.17	0.16/0.19	0.17/0.21	0.16/0.22	0.17/0.20	0.17/0.21	0.18/0.23
	LS	0.50/0.53	0.23/0.20	0.22/0.22	0.17/0.18	0.13/0.16	0.16/0.17	0.15/0.20	0.16/0.18	0.16/0.19	0.14/0.19	0.16/0.21

**Table 4.** Scatter (difference of 75 % and 25 % quantiles of model efficiency) of the leave-one-out soil moisture correlation (Eq. 8) obtained by eight groups of parameter transfer methods and 11 calibration weights for lowland (94) and alpine (119) catchments in the calibration (2007–2010; first value) and validation (2010–2014; second value) periods.

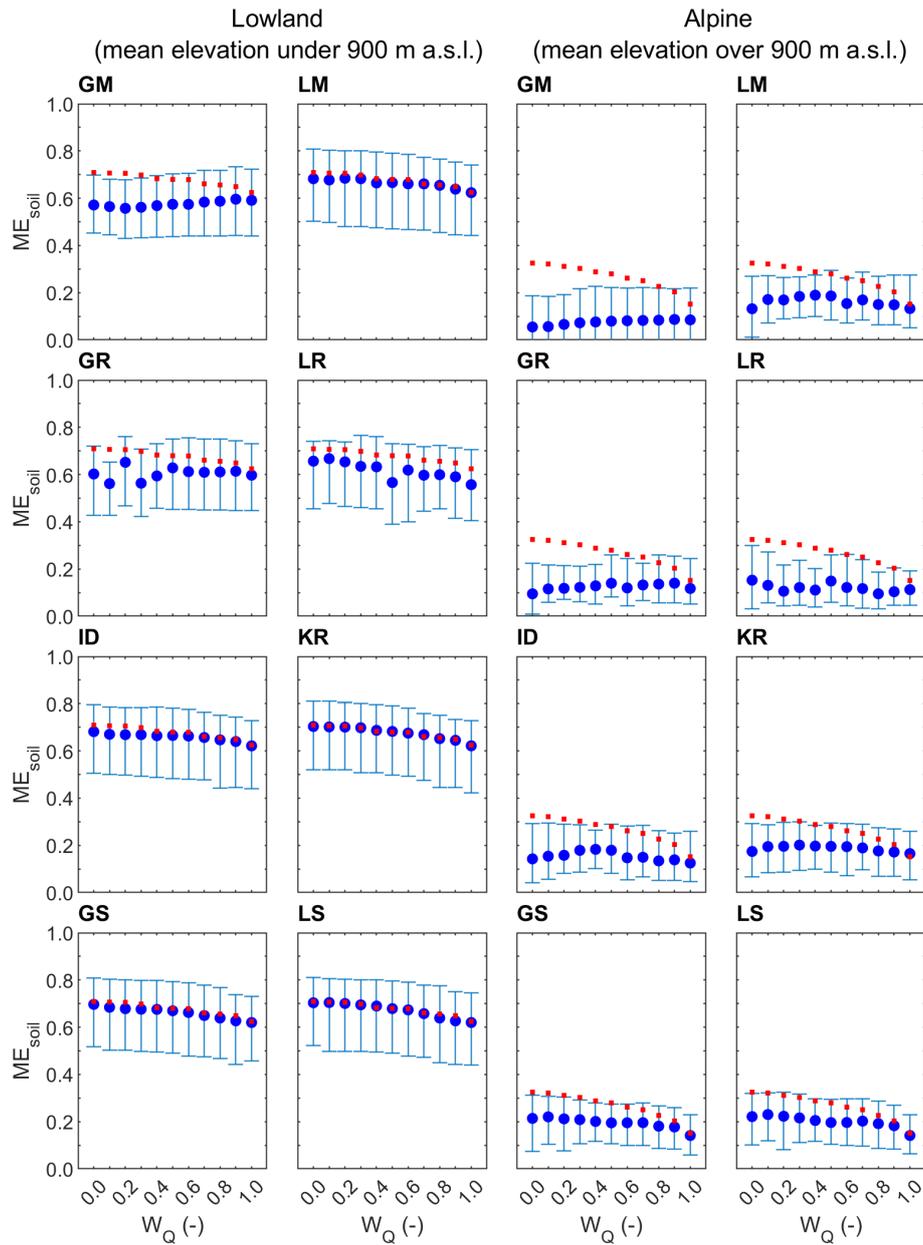
	$w_Q$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
Lowlands	GM	0.24/0.22	0.24/0.22	0.25/0.22	0.25/0.23	0.26/0.24	0.26/0.23	0.27/0.24	0.28/0.24	0.28/0.24	0.29/0.24	0.28/0.24	
	LM	0.30/0.31	0.31/0.28	0.32/0.28	0.32/0.29	0.32/0.29	0.32/0.28	0.32/0.28	0.31/0.26	0.31/0.26	0.31/0.25	0.30/0.25	
	GR	0.29/0.24	0.23/0.26	0.29/0.26	0.29/0.31	0.27/0.27	0.30/0.23	0.30/0.27	0.30/0.25	0.30/0.25	0.30/0.24	0.28/0.24	
	LR	0.28/0.26	0.27/0.24	0.27/0.22	0.30/0.25	0.31/0.29	0.34/0.34	0.33/0.31	0.27/0.25	0.27/0.25	0.30/0.29	0.30/0.29	
	ID	0.29/0.27	0.28/0.27	0.28/0.27	0.29/0.28	0.30/0.28	0.30/0.28	0.29/0.28	0.29/0.25	0.31/0.24	0.30/0.24	0.29/0.24	
	KR	0.29/0.26	0.29/0.26	0.29/0.27	0.29/0.26	0.29/0.27	0.29/0.28	0.29/0.29	0.28/0.25	0.31/0.25	0.29/0.25	0.30/0.26	
	GS	0.29/0.26	0.30/0.28	0.30/0.27	0.30/0.26	0.30/0.26	0.30/0.26	0.30/0.27	0.31/0.28	0.30/0.27	0.30/0.26	0.30/0.24	0.27/0.25
	LS	0.29/0.26	0.31/0.28	0.31/0.27	0.30/0.27	0.31/0.26	0.30/0.27	0.31/0.28	0.31/0.27	0.33/0.27	0.31/0.26	0.31/0.27	
Alpine	GM	0.23/0.19	0.23/0.18	0.23/0.19	0.24/0.19	0.24/0.19	0.23/0.20	0.23/0.20	0.23/0.19	0.23/0.19	0.23/0.19	0.22/0.19	
	LM	0.26/0.25	0.20/0.25	0.18/0.25	0.17/0.26	0.17/0.26	0.21/0.26	0.19/0.26	0.20/0.26	0.21/0.24	0.21/0.24	0.22/0.23	
	GR	0.22/0.21	0.16/0.25	0.14/0.39	0.15/0.37	0.17/0.26	0.18/0.28	0.20/0.24	0.16/0.25	0.20/0.28	0.20/0.22	0.19/0.24	
	LR	0.27/0.32	0.22/0.30	0.17/0.37	0.19/0.37	0.16/0.34	0.20/0.30	0.22/0.30	0.21/0.39	0.16/0.37	0.16/0.35	0.15/0.34	
	ID	0.25/0.24	0.24/0.25	0.21/0.26	0.20/0.25	0.17/0.24	0.21/0.27	0.23/0.28	0.22/0.28	0.21/0.25	0.20/0.23	0.21/0.22	
	KR	0.23/0.29	0.20/0.29	0.21/0.28	0.21/0.29	0.19/0.27	0.21/0.30	0.22/0.28	0.19/0.28	0.21/0.26	0.20/0.25	0.21/0.23	
	GS	0.24/0.29	0.20/0.28	0.23/0.25	0.18/0.25	0.16/0.24	0.17/0.23	0.18/0.23	0.18/0.24	0.18/0.21	0.17/0.20	0.17/0.19	
	LS	0.22/0.31	0.20/0.28	0.24/0.26	0.20/0.25	0.18/0.24	0.19/0.25	0.20/0.24	0.20/0.24	0.20/0.21	0.19/0.21	0.16/0.19	

ranges between 8 % and 17 %. The snow efficiency decreases with increasing  $w_Q$ , and it is generally larger in the lowland than in the alpine catchments. It is related to the different frequencies of snow cover conditions in these two regions, and generally the snow-free condition has with fewer errors for the simulation. Interestingly, in the alpine catchments, the similarity-based approaches and kriging have the smallest efficiency, and the most accurate results are obtained by the global mean and global regression methods. At the same time, the regional variability in the snow model efficiencies (Table 5) has a small difference between different  $w_Q$ . As also indicated in the Table 5, the local regression method performed relatively unstably for the snow simulations, and in the Alpine regions the regional variability in the snow model

efficiency is larger than that in the lowlands. The impact of the calibration variant is, however, much more important than the selection of transfer method. The comparison between calibration and validation periods indicates an overall larger snow efficiency in the validation period. This is likely linked with a warmer validation period and, generally, fewer days with snow cover associated with an increase in air temperature in the last few decades (Duethmann and Blöschl, 2018).

#### 4.4 Improvement of multi-objective regionalization vs. runoff-only regionalization

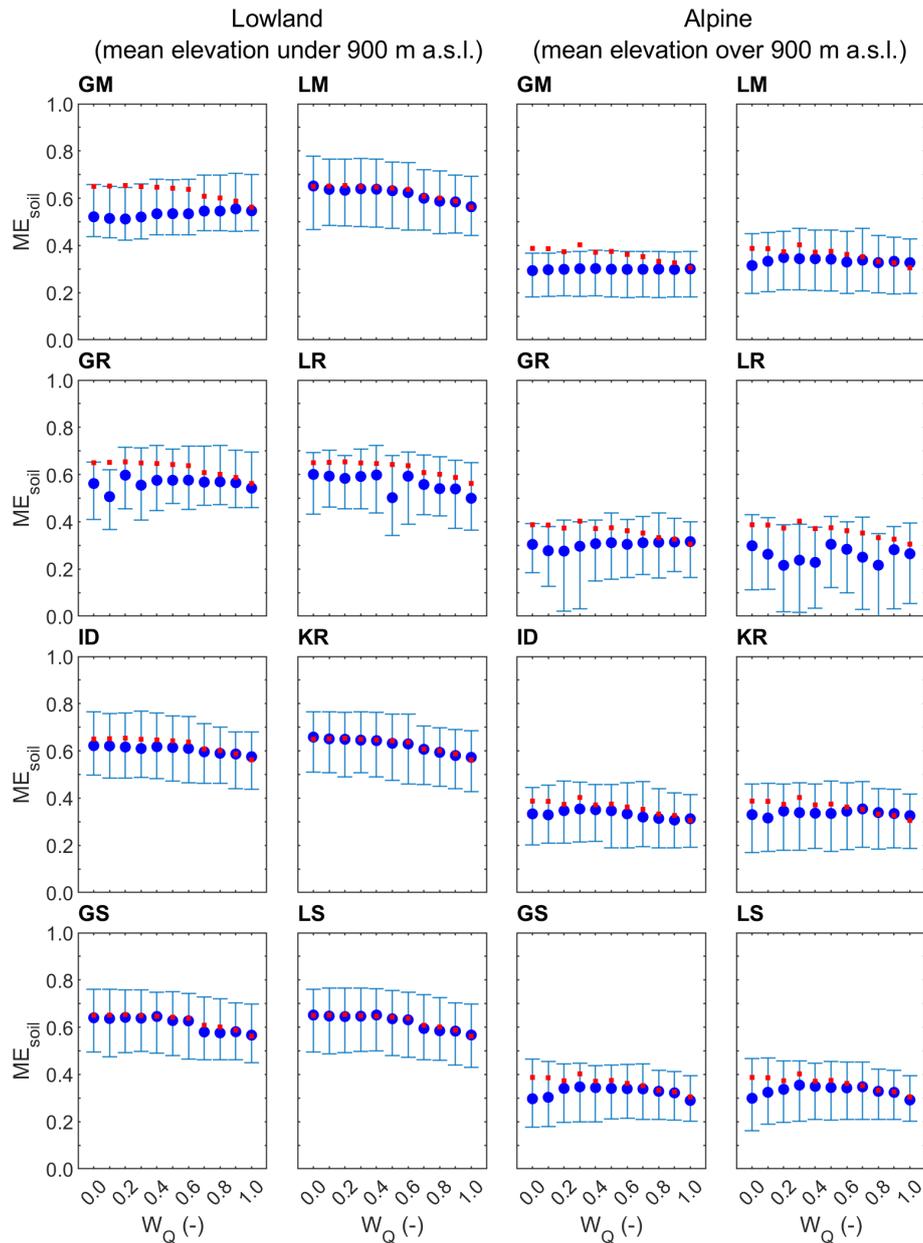
This section analyses the impact of calibration and runoff weight on the efficiency of different transfer methods for



**Figure 4.** Median and regional scatter of the leave-one-out soil moisture correlation (Eq. 8) obtained by eight groups of parameter transfer methods and 11 calibration weights for lowland (94) and alpine (119) catchments in the calibration (2007–2010; blue symbols) period. Blue circles represent the median and whiskers the 25 % and 75 % quantiles of the leave-one-out efficiency. Red symbols show the median of the at-site soil moisture correlations.

simulating runoff, soil moisture, and snow cover. We compared the efficiencies of the predictions obtained by transferring the model parameters from a multiple-objective calibration (i.e.,  $w_Q < 1$ ) with those obtained by parameters calibrated to runoff only ( $w_Q = 1$ ). The results are shown in Figs. 8 and 9 for the calibration and validation periods, respectively, in terms of the relative improvement in snow cover ( $x$  axis), soil moisture ( $y$  axis), and runoff (color of symbol) efficiencies for eight transfer methods (panels) and

10 calibration weights (symbol size) in the lowland (left) and alpine (right) catchments. Positive efficiencies indicate an improvement when using a multiple-objective calibration compared to a runoff-only calibration. The figures suggest that the runoff predictions are very similar (i.e., within a 1 % range) if the runoff weight is larger than 0.5. The largest impact of the multiple-objective calibration is found for the soil moisture prediction, where the improvement is larger than for snow cover. In the lowland catchments, the soil mois-



**Figure 5.** Same as Fig. 4 but for the validation period (2010–2014).

ture improvement is about 1 % to 7 % for  $w_Q$  larger than 0.4. In the case of snow cover, the improvement is within 5 % when  $w_Q$  larger than 0.4. The patterns of improvement for the best transfer methods (kriging and similarity approaches) are very similar. The improvement increases with decreasing runoff weight but at the expense of decreasing the runoff efficiency. The largest improvement in soil moisture (i.e., over 25 %) is found for the similarity approaches in the alpine catchments, but the overall soil moisture efficiency is rather low. The patterns of improvement are very consistent with those obtained for the transfer of model parameters in the validation period (Fig. 9). The improvement in soil moisture

and snow cover increases with the decreasing runoff weight, and the improvement is larger in the alpine than the lowland catchments. The patterns of improvement are very consistent with those obtained for the transfer of model parameters in the validation period. The most noticeable finding is that the improvement in soil moisture and snow cover increases with decreasing runoff weight, and the improvement is larger in the alpine than in the lowland catchments.

**Table 5.** Scatter (difference of 75 % and 25 % quantiles of model efficiency) of the leave-one-out snow model efficiency (Eq. 9) obtained by eight groups of parameter transfer methods and 11 calibration weights for lowland (94) and alpine (119) catchments in the calibration (2000–2010; first value) and validation (2010–2014; second value) periods.

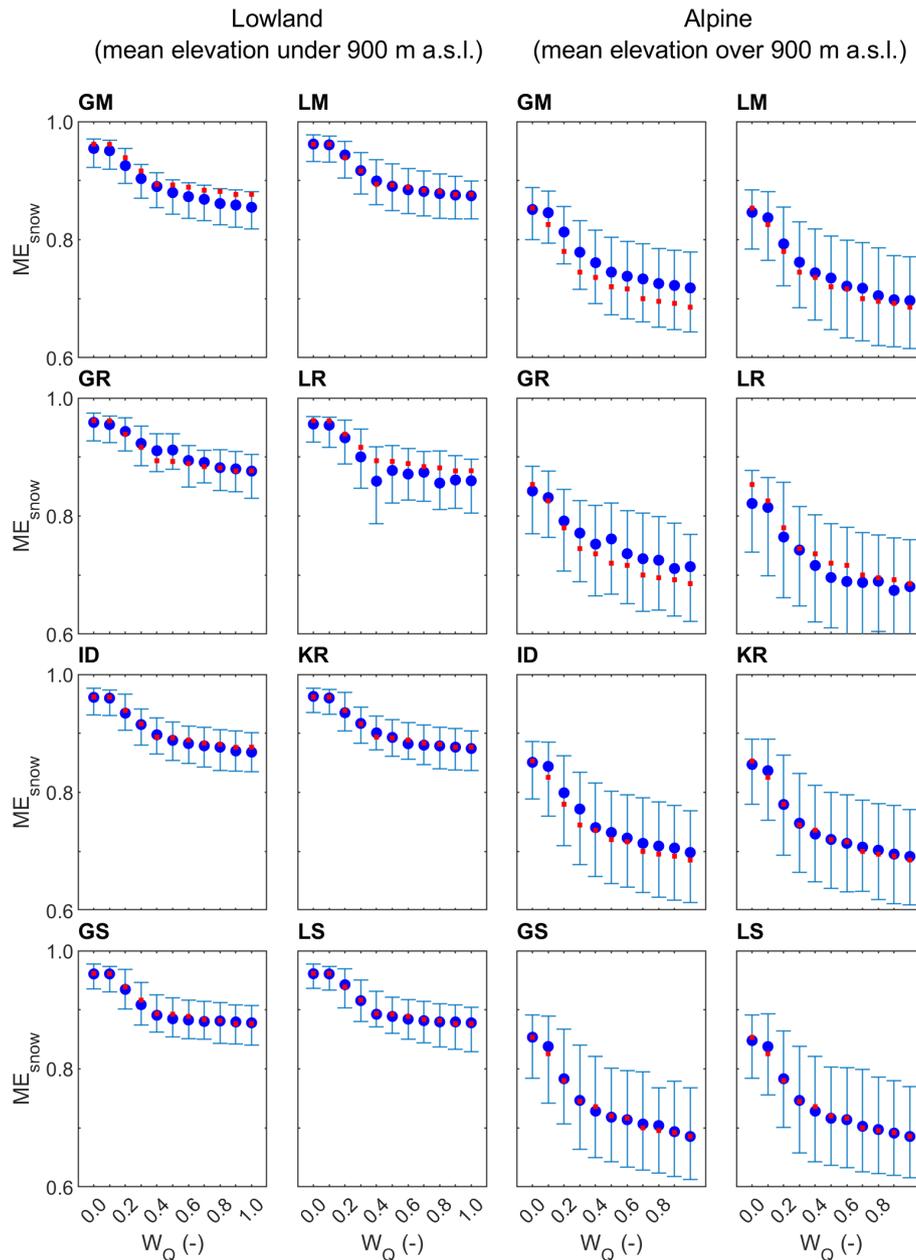
$w_Q$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	
Lowlands	GM	0.05/0.02	0.05/0.03	0.06/0.04	0.06/0.05	0.06/0.05	0.06/0.06	0.06/0.06	0.06/0.06	0.06/0.06	0.06/0.06	0.06/0.06
	LM	0.04/0.03	0.04/0.02	0.06/0.04	0.07/0.05	0.08/0.07	0.08/0.07	0.08/0.07	0.08/0.07	0.08/0.06	0.07/0.07	0.06/0.07
	GR	0.05/0.03	0.05/0.02	0.06/0.04	0.07/0.06	0.06/0.06	0.06/0.06	0.07/0.07	0.06/0.06	0.07/0.06	0.07/0.06	0.07/0.07
	LR	0.04/0.03	0.05/0.03	0.07/0.05	0.10/0.07	0.13/0.12	0.10/0.09	0.09/0.07	0.08/0.09	0.10/0.09	0.09/0.08	0.09/0.08
	ID	0.04/0.03	0.04/0.02	0.06/0.04	0.06/0.05	0.06/0.05	0.06/0.05	0.06/0.05	0.07/0.06	0.07/0.06	0.07/0.06	0.07/0.06
	KR	0.04/0.02	0.04/0.02	0.06/0.04	0.06/0.05	0.06/0.05	0.06/0.05	0.06/0.05	0.07/0.05	0.07/0.05	0.07/0.05	0.07/0.06
	GS	0.04/0.03	0.04/0.03	0.07/0.04	0.07/0.05	0.06/0.05	0.07/0.05	0.07/0.05	0.06/0.05	0.07/0.05	0.07/0.05	0.07/0.06
	LS	0.04/0.03	0.04/0.03	0.07/0.04	0.07/0.05	0.06/0.05	0.06/0.05	0.07/0.05	0.07/0.06	0.07/0.06	0.08/0.06	0.07/0.06
Alpine	GM	0.09/0.08	0.09/0.08	0.10/0.09	0.12/0.10	0.12/0.11	0.13/0.11	0.13/0.12	0.13/0.12	0.13/0.12	0.13/0.12	0.14/0.12
	LM	0.10/0.10	0.12/0.11	0.13/0.12	0.15/0.12	0.15/0.13	0.16/0.14	0.17/0.14	0.17/0.14	0.17/0.14	0.16/0.14	0.16/0.14
	GR	0.11/0.11	0.11/0.11	0.14/0.12	0.14/0.12	0.15/0.12	0.15/0.13	0.16/0.15	0.17/0.15	0.16/0.15	0.16/0.14	0.15/0.14
	LR	0.14/0.12	0.17/0.14	0.20/0.19	0.17/0.14	0.18/0.17	0.18/0.18	0.19/0.18	0.20/0.18	0.16/0.16	0.17/0.15	0.18/0.17
	ID	0.10/0.11	0.13/0.12	0.15/0.14	0.16/0.13	0.16/0.14	0.16/0.14	0.16/0.14	0.16/0.14	0.16/0.14	0.16/0.14	0.16/0.14
	KR	0.11/0.11	0.14/0.14	0.17/0.15	0.17/0.14	0.16/0.15	0.16/0.16	0.16/0.15	0.15/0.14	0.16/0.15	0.17/0.15	0.16/0.14
	GS	0.11/0.10	0.15/0.14	0.16/0.16	0.18/0.15	0.17/0.14	0.16/0.14	0.16/0.14	0.17/0.14	0.14/0.13	0.16/0.13	0.16/0.14
	LS	0.11/0.10	0.14/0.13	0.16/0.15	0.18/0.16	0.18/0.14	0.17/0.14	0.17/0.15	0.17/0.14	0.16/0.13	0.16/0.13	0.15/0.13

## 5 Discussion

The main aim of the study was to test to what extent satellite data can improve the prediction of daily runoff in ungauged catchments. Tong et al. (2021) showed that ASCAT and MODIS satellites in hydrological model calibration improve simulations of a conceptual hydrological model and that the improvements of runoff and soil moisture simulations were larger in low elevation and agricultural catchments. Here, we tested different methods for transferring model parameters from gauged to ungauged sites in the study region (Merz and Blöschl, 2004; Parajka et al., 2005). We examined which method and to what extent transferring model parameters calibrated to different objectives improves the prediction of runoff, soil moisture, and snow cover at ungauged sites. The results showed that the improvement is large in the simulation of soil moisture and snow cover, without a significant impact on runoff prediction accuracy. The assessment of the efficiencies between different transfer methods indicates that the similarity approach and kriging of model parameters are the best in the study region. This finding is in line with that of Yang et al. (2020), who concluded that the spatial proximity and similarity approaches are relatively better than the parameter average or regression method in Norwegian catchments. This finding is also entirely consistent with the results of Parajka et al. (2005), who calibrated the model parameters in a larger number of catchments but by using only runoff and interpolated snow depth (but not soil moisture). The results and efficiency of other similarity combinations tested in Parajka et al. (2005) are slightly lower or very similar in terms of runoff, soil moisture, and snow cover efficiency. The efficiency of global and local regression and arithmetic mean methods are very similar, even though here we use 107 catchments (33 %) fewer than what was tested in Parajka et

al. (2005). Overall, the lower prediction accuracy of global mean or local regression is also consistent with the global assessment of the accuracy of the transfer methods presented in Parajka et al. (2013).

The results of Tong et al. (2021) show that using only satellite soil moisture and snow cover data is not sufficient for calibrating a conceptual hydrological model in ungauged catchments. The lower runoff accuracy of the calibration variants with no or only a small weight to runoff is also reflected in the performance of the transfer methods. Satellite soil moisture and snow cover data are very useful for constraining model parameters related to the simulation of snow cover and soil moisture (Nijzink et al., 2018; Tong et al., 2021). However, the use of runoff is still essential for the accurate prediction of the complete runoff hydrographs; this is consistent with recent studies that used remotely sensed evaporation and/or total water storage anomalies for daily timescale hydrograph simulation (Zhang et al., 2020, 2021; Dembélé et al., 2020; Hulsman et al., 2021). This finding agrees with the previous assessment of using soil moisture estimates from ERS (European Remote Sensing) scatterometer observations to improve hydrological simulations in gauged and ungauged catchments (Parajka et al., 2006). The scatterometer data assimilation did not improve the prediction in ungauged sites but provided more consistent patterns of soil moisture estimates. The use of the experimental disaggregated ASCAT dataset showed that the detailed spatial and temporal resolution of satellite soil moisture improves the application and agreement of soil moisture estimates in smaller lowland catchments. Similarly, as found in Széles et al. (2020) for a small Austrian catchment, the use of soil moisture data has a larger impact on the overall consistency of the model simulations compared to snow cover observations.



**Figure 6.** Median and regional scatter of the leave-one-out snow model efficiency (Eq. 9) obtained by eight groups of parameter transfer methods and 11 calibration weights for lowland (94) and alpine (119) catchments in the calibration (2000–2010; blue symbols) period. Blue circles represent the median and whiskers the 25 % and 75 % quantiles of the leave-one-out efficiency. Red symbols show the median of the at-site snow model efficiency.

The transfer of model parameters to ungauged sites and the efficiency of different approaches for predicting runoff hydrographs are affected by different sources of uncertainty. During the conceptualization of the analysis, we considered the following potential sources of uncertainty: (a) model inputs, (b) model structure, (c) accuracy of the satellite data, (d) model calibration, and (e) model parameter regionalization. We considered the impact of sources (a) to (c) to be smaller than (d) and (e) for the following reasons. The uncertainty

of the model inputs (a) is generally mainly due to the spatial interpolation of point (precipitation and air temperature) observations, as catchment averages are needed for water balance reasons, and this is a topic that has traditionally attracted a lot of interest in hydrology (e.g., Faurès et al., 1995). In this study, the model inputs (mean daily precipitation and air temperature) are estimated from the gridded SPARTACUS dataset, with a grid resolution of 1 km that is small relative to the median catchment size 167 km<sup>2</sup>. Hiebl and Frei

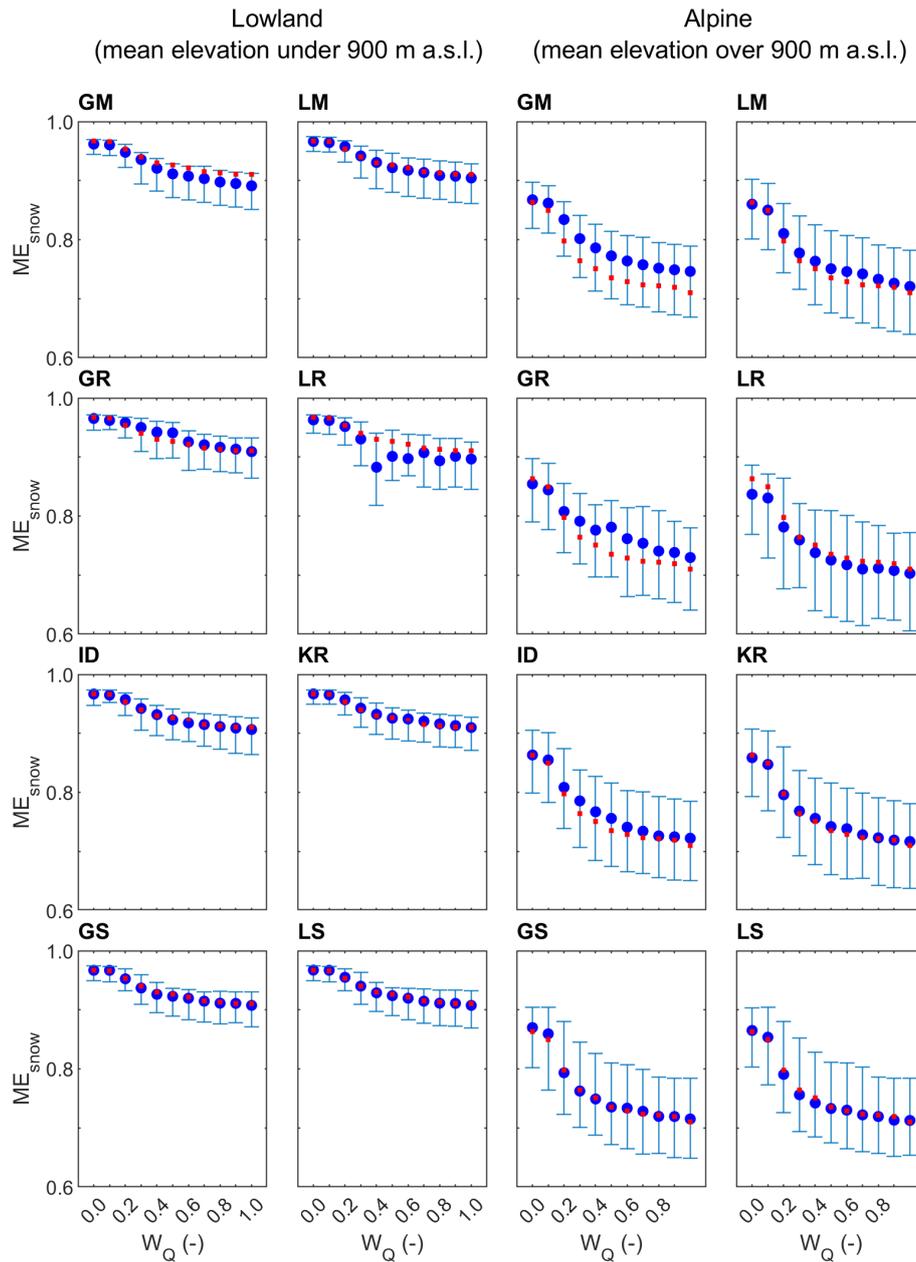
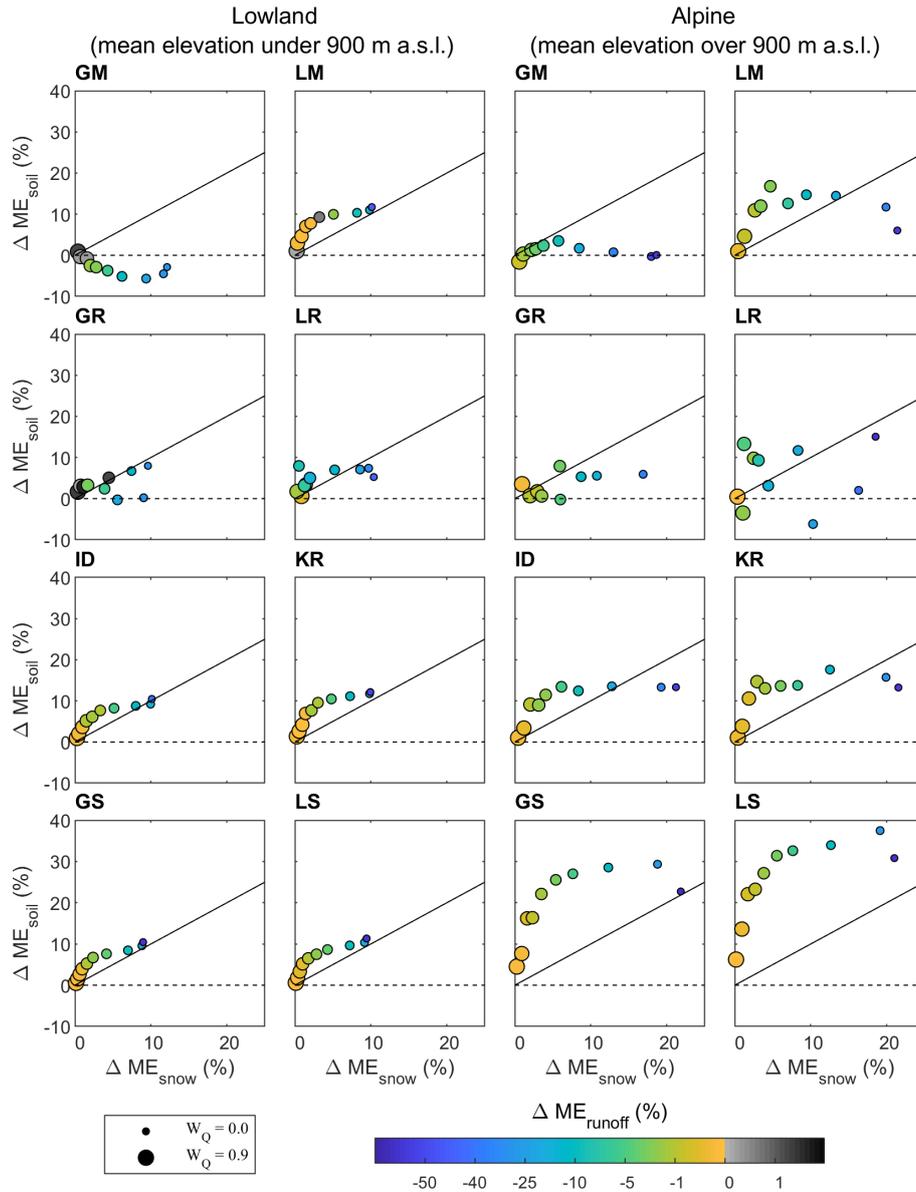


Figure 7. Same as Fig. 6 but for the validation period (2010–2014).

(2018) show the accuracy of the precipitation interpolation used in SPARTACUS to be high and the monthly biases to be very small (values are within  $\pm 2\%$ ). The cross-validation of the air temperature interpolation (Hiebl and Frei, 2016) indicates no systematic overestimation or underestimation, i.e., the compound mean error is  $0^\circ\text{C}$  and the root mean square error is  $1.4^\circ\text{C}$ .

Model structure (b) is, of course, more difficult to evaluate, and previous studies, in the context of regionalization performance (e.g., Petheram et al., 2012; Parajka et al., 2013; Yang et al., 2020), have shown that the simpler models are not superior to complex models (nor much worse) in pre-

dicting daily hydrographs in ungauged catchments, and more generally, the difference between hydrological models tends to be small (Petheram et al., 2012). Parajka et al. (2013) grouped models according to the number of model parameters and showed that the median of the regionalization performance (Nash–Sutcliffe efficiency) for each group of models is around 0.65. Yang et al. (2020) compared four daily rainfall–runoff models (GR4J, WASMOD, HBV, and XAJ, with 6, 8, 13, and 17 parameters) and reported that the difference in model structure has a smaller impact on the regionalization model performance than the difference in climate conditions. Yang et al. (2020) shows that the average Nash–



**Figure 8.** Relative difference (%) in the median of snow cover (horizontal axis), soil moisture (vertical axis), and runoff efficiency between model simulations obtained by transferring model parameters calibrated by multiple-objective calibration and calibration to runoff only. The relative difference is estimated for eight model transfer methods (panels) applied in the lowland (left panels) and alpine (right panels) catchments in the calibration period 2000–2010 (soil moisture from 2007 to 2010).

Sutcliffe runoff efficiency values are, for the best regionalization method (physical similarity methods with output averaging), larger than 0.6 for all tested model structures.

The evaluation of the MODIS snow cover (c) of Tong et al. (2021) indicates an overall classification accuracy of the most recent MODIS snow cover product of larger than 97 %, which implies much smaller uncertainties than most of the other sources. The accuracy assessment of the experimental S1ASCAT dataset at the regional scale is still work in progress. A preliminary assessment (Panic et al., 2020; [https://presentations.copernicus.org/EGU2020/EGU2020-16222\\_presentation.pdf](https://presentations.copernicus.org/EGU2020/EGU2020-16222_presentation.pdf), last access: 6 April 2022) demonstrates that S1ASCAT compares well with point-scale and area-representative in situ root zone measurements. The correlation between observed in situ (i.e., time-domain reflectometry (TDR) soil network and cosmic-ray neutron probe) and S1ASCAT soil moisture is 0.59 and 0.51, respectively. These correlations are higher than those obtained between in situ and existing COPERNICUS soil moisture products (surface soil moisture, SSM, 1 km and SWI 1 km), and it is to be expected that only a part of the differences between the data types is due to the satellite data,

org/EGU2020/EGU2020-16222\_presentation.pdf, last access: 6 April 2022) demonstrates that S1ASCAT compares well with point-scale and area-representative in situ root zone measurements. The correlation between observed in situ (i.e., time-domain reflectometry (TDR) soil network and cosmic-ray neutron probe) and S1ASCAT soil moisture is 0.59 and 0.51, respectively. These correlations are higher than those obtained between in situ and existing COPERNICUS soil moisture products (surface soil moisture, SSM, 1 km and SWI 1 km), and it is to be expected that only a part of the differences between the data types is due to the satellite data,

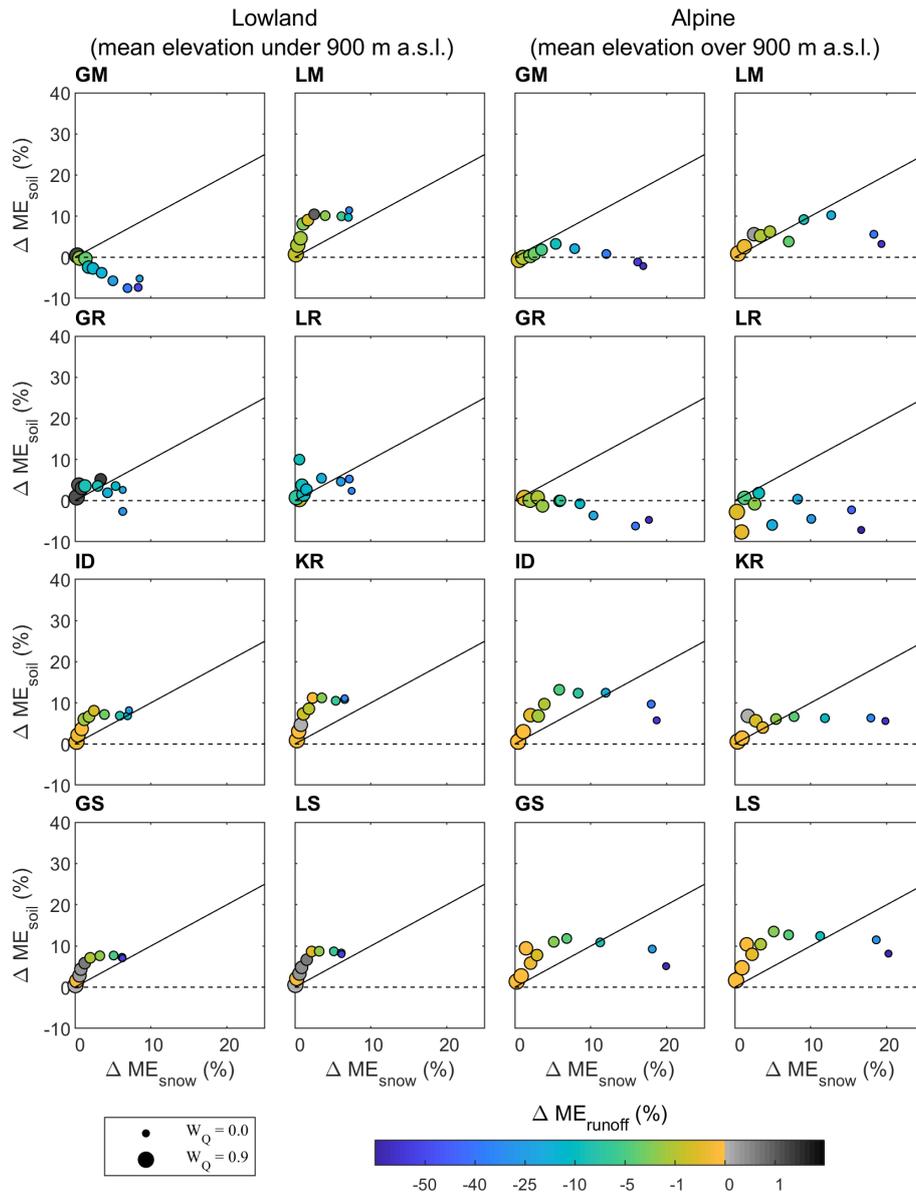


Figure 9. Same as Fig. 8 but for the validation period (2010–2014).

as TDR soil probes and cosmic-ray neutron probes also have some level of uncertainty.

We thus decided to focus on the uncertainties resulting from model calibration (d) and selection of regionalization method (e). The impact of using different time periods for the prediction of runoff hydrographs is evaluated by the split-sample uncertainty assessment proposed by Klemes (1985). Regionalization studies typically refer only to regionalization model efficiencies obtained for the same period as that used for model calibration. Our results indicate that regionalization efficiencies obtained in an independent validation period generally show a small decrease (loss) in runoff model performance. The median of the loss in the Nash–Sutcliffe efficiency varies between 0.02 and 0.07, depending on the re-

gionalization method and calibration weight. In the lowlands, the average median loss is 0.06, while it is 0.03 in the alpine basins. The results also show that the median loss of runoff efficiency tends to be smaller for multiple-objective variants (average median loss of 0.05) than for variants using parameters calibrated to runoff only (average median loss of 0.06). These results are consistent with Yang et al. (2020), who reported a small degradation of the regionalization runoff performance from the calibration to the validation period. The largest relative improvement of soil moisture efficiency is found in alpine catchments (more than 70 %), but the absolute value of the correlations (on average 0.31) is still lower than in lowland catchments (average correlation 0.59). These numbers suggest that the differences in performance (which

are an indicator of the uncertainties to be expected) are quite significant for the uncertainty source (d).

The evaluation of the uncertainty of runoff prediction, using different regionalization methods (e), shows that the variability in the medians of the runoff regionalization efficiency is smaller between regionalization methods than between different calibration variants (i.e., runoff weights). For example, the standard deviation of the medians obtained for 11 runoff weights for the local similarity regionalization method in alpine catchments is 0.17. The standard deviation of the medians between eight regionalization methods ranges (depending on the runoff weight) between 0.04 and 0.11. The differences are somewhat smaller in lowland catchments (i.e., the standard deviation of medians between runoff weights and regionalization methods is 0.14 and about 0.09, respectively).

## 6 Conclusions and outlook

This study shows that the recent advances in the remote sensing of water balance components contribute to improving the hydrological predictions in ungauged catchments. The main improvements are in estimating soil moisture and snow cover dynamics, mostly in alpine catchments. Future analyses may focus on assessing the value of satellite data for other types of regionalization approaches, such as regional calibration (Parajka et al., 2007) or multiscale parameter regionalization methods (Samaniego et al., 2010; Kumar et al., 2013). It will also be interesting to evaluate how much runoff information is needed in addition to existing satellite products to improve and constrain the model predictions in ungauged basins. Such an investigation can also include an analysis of the role of nested catchments in parameter transfer and the impact of stream gauge density on the regionalization model performance.

**Data availability.** The streamflow observation data can be obtained from HZB (<https://ehyd.gv.at/>; BMLRT, 2021). The meteorological data from the ZAMG are currently not publicly available and requests should be sent to [klima@zamg.ac.at](mailto:klima@zamg.ac.at). The ASCAT soil moisture data are available via the Copernicus Global Land Service (<https://land.copernicus.eu/global/products/swi/>; CGLS, 2018). MODIS C6 snow cover products are from NASA National Snow and Ice Data Center (<https://doi.org/10.5067/MODIS/MOD10A1>, Hall and Riggs, 2016a; <https://doi.org/10.5067/MODIS/MYD10A1.0>, Hall and Riggs, 2016b). Processed ASCAT SWI and MODIS snow cover maps used in this study are available upon request. Land use information is from the Copernicus Land Monitoring Service (<https://land.copernicus.eu/>; CLC2006, 2013). The R package of the TUW-model can be downloaded from CRAN (<https://CRAN.R-project.org/package=TUWmodel>; Viglione and Parajka, 2020).

**Supplement.** The supplement related to this article is available online at: <https://doi.org/10.5194/hess-26-1779-2022-supplement>.

**Author contributions.** RT and JP conceived and designed the study, wrote the codes, performed the analyses, and prepared the paper. BS, JK, and PV were responsible for the data management, including quality control, processing, and validating. IP and MV were responsible for developing, processing, and validating the ASCAT soil moisture data. GB supervised the study and contributed to the study design and interpretation of the results. All authors took part in the discussion of the results and revisions of the paper.

**Competing interests.** The contact author has declared that neither they nor their co-authors have any competing interests.

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**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 (grant no. DK W1219-N28), the Austrian Research Promotion Agency (FFG) through the BMon project (grant no. 866031), and the VEGA grant agency (grant no. VEGA 1/0632/19). The authors would also like to thank Sebastian Hahn, for his support in developing and processing the ASCAT soil moisture data. Rui Tong is grateful for the scholarship from the China Scholarship Council (CSC).

**Financial support.** This research has been supported by the Austrian Science Funds (FWF), as part of the Vienna Doctoral Programme on Water Resource Systems (grant no. DK W1219-N28), the Austrian Research Promotion Agency (FFG), through the BMon project (grant no. 866031), and the VEGA grant agency (grant no. VEGA 1/0632/19).

**Review statement.** This paper was edited by Jim Freer and reviewed by Luis Samaniego, Markus Hrachowitz, and one anonymous referee.

## References

- Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Pignatelli, B., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations, *Hydrol. Earth Syst. Sci.*, 12, 1323–1337, <https://doi.org/10.5194/hess-12-1323-2008>, 2008.
- Astagneau, P. C., Thirel, G., Delaigue, O., Guillaume, J. H. A., Parajka, J., Brauer, C. C., Viglione, A., Buytaert, W., and Beven, K. J.: Technical note: Hydrology modelling R packages – a unified analysis of models and practicalities from a user perspective, *Hydrol. Earth Syst. Sci.*, 25, 3937–3973, <https://doi.org/10.5194/hess-25-3937-2021>, 2021.

- Beck, H. E., Pan, M., Lin, P., Seibert, J., van Dijk, A. I. J. M., and Wood, E. F.: Global Fully Distributed Parameter Regionalization Based on Observed Streamflow From 4229 Headwater Catchments, *J. Geophys. Res.-Atmos.*, 125, e2019JD031485, <https://doi.org/10.1029/2019JD031485>, 2020.
- Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R. M., van Dijk, A. I. J. M., Du, J., Kimball, J. S., Vergopolan, N., and Wood, E. F.: Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors, *Hydrol. Earth Syst. Sci.*, 25, 17–40, <https://doi.org/10.5194/hess-25-17-2021>, 2021.
- Bergström, S.: The HBV – its structure and applications, Report No. 4, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, 1992.
- Blöschl, G. and Montanari, A.: Climate change impacts – throwing the dice?, *Hydrol. Process.*, 24, 374–381, <https://doi.org/10.1002/hyp.7574>, 2010.
- Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H. H. G. (Eds.): *Runoff Prediction in Ungauged Basins – Synthesis across Processes, Places and Scales*, Cambridge University Press, Cambridge, UK, 465 pp., ISBN 9781107028180, 2013.
- Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Allen, S. T., Amin, A., Andréassian, V., Arheimer, B., Aryal, S. K., Baker, V., Bardsley, E., Barendrecht, M. H., Bartosova, A., Batelaan, O., Berghuijs, W. R., Beven, K., Blume, T., Bogaard, T., Borges de Amorim, P., Böttcher, M. E., Boulet, G., Breinl, K., Brilly, M., Brocca, L., Buytaert, W., Castellarin, A., Castelletti, A., Chen, X., Chen, Y., Chen, Y., Chiffard, P., Claps, P., Clark, M. P., Collins, A. L., Croke, B., Dathe, A., David, P. C., de Barros, F. P. J., de Rooij, G., Di Baldassarre, G., Driscoll, J. M., Duethmann, D., Dwivedi, R., Eris, E., Farmer, W. H., Feiccabrino, J., Ferguson, G., Ferrari, E., Ferraris, S., Fersch, B., Finger, D., Foglia, L., Fowler, K., Gartsman, B., Gascoin, S., Gaume, E., Gelfan, A., Geris, J., Gharari, S., Gleeson, T., Glendell, M., Gonzalez Bevacqua, A., González-Dugo, M. P., Grimaldi, S., Gupta, A. B., Guse, B., Han, D., Hannah, D., Harpold, A., Haun, S., Heal, K., Helfricht, K., Hernegger, M., Hipsey, M., Hlaváčiková, H., Hohmann, C., Holko, L., Hopkinson, C., Hrachowitz, M., Illangasekare, T. H., Inam, A., Innocente, C., Istanbuluoglu, E., Jarihani, B., Kalantari, Z., Kalvans, A., Khanal, S., Khatami, S., Kiesel, J., Kirkby, M., Knoben, W., Kochanek, K., Kohnová, S., Kolechkina, A., Krause, S., Kremer, D., Kreibich, H., Kunstmann, H., Lange, H., Liberato, M. L. R., Lindquist, E., Link, T., Liu, J., Loucks, D. P., Luce, C., Mahé, G., Makarieva, O., Malard, J., Mashtayeva, S., Maskey, S., Mas-Pla, J., Mavrova-Guirguinova, M., Mazzoleni, M., Mernild, S., Misstear, B. D., Montanari, A., Müller-Thomy, H., Nabizadeh, A., Nardi, F., Neale, C., Nesterova, N., Nurtaev, B., Odongo, V. O., Panda, S., Pande, S., Pang, Z., Papacharalampous, G., Perrin, C., Pfister, L., Pimentel, R., Polo, M. J., Post, D., Prieto Sierra, C., Ramos, M.-H., Renner, M., Reynolds, J. E., Ridolfi, E., Rigon, R., Riva, M., Robertson, D. E., Rosso, R., Roy, T., Sá, J. H. M., Salvadori, G., Sandells, M., Schaeffli, B., Schumann, A., Scolobig, A., Seibert, J., Servat, E., Shafiei, M., Sharma, A., Sidibe, M., Sidle, R. C., Skaugen, T., Smith, H., Spiessl, S. M., Stein, L., Steinsland, I., Strasser, U., Su, B., Szolgay, J., Tarboton, D., Tauro, F., Thirel, G., Tian, F., Tong, R., Tus-supova, K., Tyralis, H., Uijlenhoet, R., van Beek, R., van der Ent, R. J., van der Ploeg, M., Van Loon, A. F., van Meerveld, I., van Nooijen, R., van Oel, P. R., Vidal, J.-P., von Freyberg, J., Vorogushyn, S., Wachniew, P., Wade, A. J., Ward, P., Westerberg, I. K., White, C., Wood, E. F., Woods, R., Xu, Z., Yilmaz, K. K., and Zhang, Y.: Twenty-three unsolved problems in hydrology (UPH) – a community perspective, *Hydrolog. Sci. J.*, 64, 1141–1158, <https://doi.org/10.1080/02626667.2019.1620507>, 2019.
- BMLRT: ehyd – Hydrographic data and analyses, <https://ehyd.gv.at/>, last access: 13 April 2021.
- Bouaziz, L. J. E., Fenicia, F., Thirel, G., de Boer-Euser, T., Buitink, J., Brauer, C. C., De Niel, J., Dewals, B. J., Drogue, G., Grelier, B., Melsen, L. A., Moustakas, S., Nossent, J., Pereira, F., Sprokkereef, E., Stam, J., Weerts, A. H., Willems, P., Savenije, H. H. G., and Hrachowitz, M.: Behind the scenes of streamflow model performance, *Hydrol. Earth Syst. Sci.*, 25, 1069–1095, <https://doi.org/10.5194/hess-25-1069-2021>, 2021.
- Burn, D. H. and Boorman, D. B.: Estimation of hydrological parameters at ungauged catchments, *J. Hydrol.*, 143, 429–454, [https://doi.org/10.1016/0022-1694\(93\)90203-L](https://doi.org/10.1016/0022-1694(93)90203-L), 1993.
- CGLS: Soil Water Index, TU Wien, CGLS [data set], <https://land.copernicus.eu/global/products/swi/> (last access: 6 April 2022), 2018.
- CLC2006: Copernicus Land Monitoring Service at a glance, <https://land.copernicus.eu/> (last access: 5 June 2020), 2013.
- Dembélé, M., Hrachowitz, M., Savenije, H. H. G., Mariéthoz, G., and Schaeffli, B.: Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, *Water Resour. Res.*, 56, e2019WR026085, <https://doi.org/10.1029/2019WR026085>, 2020.
- Duethmann, D. and Blöschl, G.: Why has catchment evaporation increased in the past 40 years? A data-based study in Austria, *Hydrol. Earth Syst. Sci.*, 22, 5143–5158, <https://doi.org/10.5194/hess-22-5143-2018>, 2018.
- Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review, *Hydrolog. Sci. J.*, 55, 58–78, <https://doi.org/10.1080/02626660903526292>, 2010.
- Faurès, J.-M., Goodrich, D. C., Woolhiser, D. A., and Sorooshian, S.: Impact of small-scale spatial rainfall variability on runoff modeling, *J. Hydrol.*, 173, 309–326, [https://doi.org/10.1016/0022-1694\(95\)02704-S](https://doi.org/10.1016/0022-1694(95)02704-S), 1995.
- Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J. C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., van der Schalie, R., Wigneron, J. P., and Wagner, W.: Validation practices for satellite soil moisture retrievals: What are (the) errors?, *Remote Sens. Environ.*, 244, 111806, <https://doi.org/10.1016/j.rse.2020.111806>, 2020.
- Guo, Y., Zhang, Y., Zhang, L., and Wang, Z.: Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review, *WIREs Water*, 8, e1487, <https://doi.org/10.1002/wat2.1487>, 2021.

- Hahn, S., Wagner, W., Steele-Dunne, S. C., Vreugdenhil, M., and Melzer, T.: Improving ASCAT Soil Moisture Retrievals With an Enhanced Spatially Variable Vegetation Parameterization, *IEEE T. Geosci. Remote*, 59, 8241–8256, <https://doi.org/10.1109/TGRS.2020.3041340>, 2021.
- Hall, D. K. and Riggs, G. A.: MODIS/Aqua Snow Cover Daily L3 Global 500 m Grid, Version 6, [September 2002 to August 2014], Boulder, Colorado, USA, NASA National Snow and Ice Data Center Distributed Active Archive Center [data set], <https://doi.org/10.5067/MODIS/MYD10A1.006>, 2016a.
- Hall, D. K. and Riggs, G. A.: MODIS/Terra Snow Cover Daily L3 Global 500 m Grid, Version 6, [September 2002 to August 2014], Boulder, Colorado, USA, NASA National Snow and Ice Data Center Distributed Active Archive Center [data set], <https://doi.org/10.5067/MODIS/MOD10A1.006>, 2016b.
- He, Y., Bárdossy, A., and Zehe, E.: A review of regionalisation for continuous streamflow simulation, *Hydrol. Earth Syst. Sci.*, 15, 3539–3553, <https://doi.org/10.5194/hess-15-3539-2011>, 2011.
- Hiebl, J. and Frei, C.: Daily temperature grids for Austria since 1961 – concept, creation and applicability, *Theor. Appl. Climatol.*, 124, 161–178, <https://doi.org/10.1007/s00704-015-1411-4>, 2016.
- Hiebl, J. and Frei, C.: Daily precipitation grids for Austria since 1961 – Development and evaluation of a spatial dataset for hydroclimatic monitoring and modelling, *Theor. Appl. Climatol.*, 132, 327–345, <https://doi.org/10.1007/s00704-017-2093-x>, 2018.
- Hirsch, R. M., Helsel, D. R., Cohn, T. A., and Gilroy, E. J.: Statistical analysis of hydrological data, in: *Handbook of Hydrology*, edited by: Maidment, R., McGraw-Hill, New York, 17.1–17.55, ISBN 9780070397323, 1992.
- 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., Fencic, 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., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) – a review, *Hydrolog. Sci. J.*, 58, 1198–1255, <https://doi.org/10.1080/02626667.2013.803183>, 2013.
- Huang, Q., Qin, G., Zhang, Y., Tang, Q., Liu, C., Xia, J., Chiew, F. H. S., and Post, D.: Using Remote Sensing Data-Based Hydrological Model Calibrations for Predicting Runoff in Ungauged or Poorly Gauged Catchments, *Water Resour. Res.*, 56, e2020WR028205, <https://doi.org/10.1029/2020WR028205>, 2020.
- Hulsman, P., Savenije, H. H. G., and Hrachowitz, M.: Learning from satellite observations: increased understanding of catchment processes through stepwise model improvement, *Hydrol. Earth Syst. Sci.*, 25, 957–982, <https://doi.org/10.5194/hess-25-957-2021>, 2021.
- Klemes, V.: Operational testing of hydrological simulation models, *Hydrolog. Sci. J.*, 31, 13–24, <https://doi.org/10.1080/02626668609491024>, 1985.
- Kovacs, A., Honti, M., Zessner, M., Eder, A., Clement, A., and Blöschl, G.: Identification of phosphorus emission hotspots in agricultural catchments, *Sci. Total Environ.*, 433, 74–88, <https://doi.org/10.1016/j.scitotenv.2012.06.024>, 2012.
- Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, *Water Resour. Res.*, 49, 360–379, <https://doi.org/10.1029/2012WR012195>, 2013.
- Lebecherel, L., Andréassian, V., and Perrin, C.: On evaluating the robustness of spatial-proximity-based regionalization methods, *J. Hydrol.*, 539, 196–203, <https://doi.org/10.1016/j.jhydrol.2016.05.031>, 2016.
- Li, H. and Zhang, Y.: Regionalising rainfall-runoff modelling for predicting daily runoff: Comparing gridded spatial proximity and gridded integrated similarity approaches against their lumped counterparts, *J. Hydrol.*, 550, 279–293, <https://doi.org/10.1016/j.jhydrol.2017.05.015>, 2017.
- Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, *J. Hydrol.*, 287, 95–123, <https://doi.org/10.1016/j.jhydrol.2003.09.028>, 2004.
- Narbondo, S., Gorgoglione, A., Crisci, M., and Chreties, C.: Enhancing Physical Similarity Approach to Predict Runoff in Ungauged Watersheds in Sub-Tropical Regions, *Water*, 12, 528, <https://doi.org/10.3390/w12020528>, 2020.
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6), 1970.
- Neri, M., Parajka, J., and Toth, E.: Importance of the informative content in the study area when regionalising rainfall-runoff model parameters: the role of nested catchments and gauging station density, *Hydrol. Earth Syst. Sci.*, 24, 5149–5171, <https://doi.org/10.5194/hess-24-5149-2020>, 2020.
- Nijzink, R. C., Almeida, S., Pechlivanidis, I. G., Capell, R., Gustafsons, D., Arheimer, B., Parajka, J., Freer, J., Han, D., Wagener, T., van Nooijen, R. R. P., Savenije, H. H. G., and Hrachowitz, M.: Constraining Conceptual Hydrological Models With Multiple Information Sources, *Water Resour. Res.*, 54, 8332–8362, <https://doi.org/10.1029/2017wr021895>, 2018.
- Panic, D., Pfeil, I., Salentinig, A., Vreugdenhil, M., Wagner, W., Wahbi, A., Fulajtar, E., Said, H., Franz, T., Heng, L., and Strauss, P.: Area-representative validation of remotely sensed high resolution soil moisture using a cosmic-ray neutron sensor, *EGU General Assembly 2020*, 4–8 May 2020, Online, EGU2020-16222, <https://doi.org/10.5194/egusphere-egu2020-16222>, 2020.
- Parajka, J. and Blöschl, G.: Validation of MODIS snow cover images over Austria, *Hydrol. Earth Syst. Sci.*, 10, 679–689, <https://doi.org/10.5194/hess-10-679-2006>, 2006.
- Parajka, J. and Blöschl, G.: The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models, *J. Hydrol.*, 358, 240–258, <https://doi.org/10.1016/j.jhydrol.2008.06.006>, 2008.
- Parajka, J., Merz, R., and Blöschl, G.: 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”, 20 November 2003, Slovak Academy of Sciences, Bratislava, 299–306, 2003.
- Parajka, J., Merz, R., and Blöschl, G.: A comparison of regionalisation methods for catchment model parameters, *Hydrol. Earth Syst. Sci.*, 9, 157–171, <https://doi.org/10.5194/hess-9-157-2005>, 2005.
- Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., and Scipal, K.: Assimilating scatterometer soil moisture data into

- conceptual hydrologic models at the regional scale, *Hydrol. Earth Syst. Sci.*, 10, 353–368, <https://doi.org/10.5194/hess-10-353-2006>, 2006.
- Parajka, J., Merz, R., and Blöschl, G.: Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments, *Hydrol. Process.*, 21, 435–446, <https://doi.org/10.1002/hyp.6253>, 2007.
- Parajka, J., Viglione, A., Rogger, M., Salinas, J. L., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies, *Hydrol. Earth Syst. Sci.*, 17, 1783–1795, <https://doi.org/10.5194/hess-17-1783-2013>, 2013.
- Parajka, J., Merz, R., Sköien, J. O., and Viglione, A.: The role of station density for predicting daily runoff by top-kriging interpolation in Austria, *J. Hydrol. Hydromech.*, 63, 228–234, <https://doi.org/10.1515/johh-2015-0024>, 2015.
- Paulik, C., Dorigo, W., Wagner, W., and Kidd, R.: Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network, *Int. J. Appl. Earth Obs.*, 30, 1–8, <https://doi.org/10.1016/j.jag.2014.01.007>, 2014.
- Petheram, C., Rustomji, P., Chiew, F. H. S., and Vleeshouwer, J.: Rainfall–runoff modelling in northern Australia: A guide to modelling strategies in the tropics, *J. Hydrol.*, 462–463, 28–41, <https://doi.org/10.1016/j.jhydrol.2011.12.046>, 2012.
- Pfeil, I., Vreugdenhil, M., Hahn, S., Wagner, W., Strauss, P., and Blöschl, G.: Improving the seasonal representation of ASCAT soil moisture and vegetation dynamics in a temperate climate, *Remote Sens.*, 10, 1788, <https://doi.org/10.3390/rs10111788>, 2018.
- Rakovec, O., Kumar, R., Mai, J., Cuntz, M., Thober, S., Zink, M., Attinger, S., Schäfer, D., Schrön, M., and Samaniego, L.: Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins, *J. Hydrometeorol.*, 17, 287–307, <https://doi.org/10.1175/jhm-d-15-0054.1>, 2016.
- Sachs, J. D. and McArthur, J. W.: The Millennium Project: a plan for meeting the Millennium Development Goals, *Lancet*, 365, 347–353, [https://doi.org/10.1016/S0140-6736\(05\)17791-5](https://doi.org/10.1016/S0140-6736(05)17791-5), 2005.
- Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, *Water Resour. Res.*, 46, W05523, <https://doi.org/10.1029/2008WR007327>, 2010.
- Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., and Boni, G.: Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data, *Hydrol. Earth Syst. Sci.*, 19, 1727–1751, <https://doi.org/10.5194/hess-19-1727-2015>, 2015.
- Sleziak, P., Szolgay, J., Hlavčová, K., Danko, M., and Parajka, J.: The effect of the snow weighting on the temporal stability of hydrologic model efficiency and parameters, *J. Hydrol.*, 583, 124639, <https://doi.org/10.1016/j.jhydrol.2020.124639>, 2020.
- Széles, B., Parajka, J., Hogan, P., Silasari, R., Pavlin, L., Strauss, P., and Blöschl, G.: The Added Value of Different Data Types for Calibrating and Testing a Hydrologic Model in a Small Catchment, *Water Resour. Res.*, 56, e2019WR026153, <https://doi.org/10.1029/2019WR026153>, 2020.
- Tegegne, G. and Kim, Y.-O.: Modelling ungauged catchments using the catchment runoff response similarity, *J. Hydrol.*, 564, 452–466, <https://doi.org/10.1016/j.jhydrol.2018.07.042>, 2018.
- Tong, R., Parajka, J., Komma, J., and Blöschl, G.: Mapping snow cover from daily Collection 6 MODIS products over Austria, *J. Hydrol.*, 590, 125548, <https://doi.org/10.1016/j.jhydrol.2020.125548>, 2020.
- Tong, R., Parajka, J., Salentinig, A., Pfeil, I., Komma, J., Széles, B., Kubáň, M., Valent, P., Vreugdenhil, M., Wagner, W., and Blöschl, G.: The value of ASCAT soil moisture and MODIS snow cover data for calibrating a conceptual hydrologic model, *Hydrol. Earth Syst. Sci.*, 25, 1389–1410, <https://doi.org/10.5194/hess-25-1389-2021>, 2021.
- Viglione, A. and Parajka, J.: TUVmodel: Lumped/Semi-Distributed Hydrological Model for Education Purposes, R package version 1.1-1, CRAN [code], <https://CRAN.R-project.org/package=TUVmodel> (last access: 6 April 2022), 2020.
- Viglione, A., Parajka, J., Rogger, M., Salinas, J. L., Laaha, G., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins – Part 3: Runoff signatures in Austria, *Hydrol. Earth Syst. Sci.*, 17, 2263–2279, <https://doi.org/10.5194/hess-17-2263-2013>, 2013.
- Wagner, W., Lemoine, G., Borgeaud, M., and Rott, H.: A study of vegetation cover effects on ERS scatterometer data, *IEEE T. Geosci. Remote.*, 37, 938–948, <https://doi.org/10.1109/36.752212>, 1999a.
- Wagner, W., Lemoine, G., and Rott, H.: A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, *Remote Sens. Environ.*, 70, 191–207, [https://doi.org/10.1016/S0034-4257\(99\)00036-X](https://doi.org/10.1016/S0034-4257(99)00036-X), 1999b.
- Wagner, W., Pathe, C., Doubkova, M., Sabel, D., Bartsch, A., Hasenauer, S., Blöschl, G., Scipal, K., Martínez-Fernández, J., and Löw, A.: Temporal Stability of Soil Moisture and Radar Backscatter Observed by the Advanced Synthetic Aperture Radar (ASAR), *Sensors*, 8, 1174–1197, <https://doi.org/10.3390/s80201174>, 2008.
- Yang, X., Magnusson, J., Rizzi, J., and Xu, C.-Y.: Runoff prediction in ungauged catchments in Norway: comparison of regionalization approaches, *Hydrol. Res.*, 49, 487–505, <https://doi.org/10.2166/nh.2017.071>, 2017.
- Yang, X., Magnusson, J., Huang, S., Beldring, S., and Xu, C.-Y.: Dependence of regionalization methods on the complexity of hydrological models in multiple climatic regions, *J. Hydrol.*, 582, 124357, <https://doi.org/10.1016/j.jhydrol.2019.124357>, 2020.
- Zhang, L., Zhao, Y., Ma, Q., Wang, P., Ge, Y., and Yu, W.: A parallel computing-based and spatially stepwise strategy for constraining a semi-distributed hydrological model with streamflow observations and satellite-based evapotranspiration, *J. Hydrol.*, 599, 126359, <https://doi.org/10.1016/j.jhydrol.2021.126359>, 2021.
- Zhang, Y., Chiew, F. H. S., Liu, C., Tang, Q., Xia, J., Tian, J., Kong, D., and Li, C.: Can Remotely Sensed Actual Evapotranspiration Facilitate Hydrological Prediction in Ungauged Regions Without Runoff Calibration?, *Water Resour. Res.*, 56, e2019WR026236, <https://doi.org/10.1029/2019WR026236>, 2020.