

Article

# Benefits of Combining Satellite-Derived Snow Cover Data and Discharge Data to Calibrate a Glaciated Catchment in Sub-Arctic Iceland

Julia de Niet <sup>1</sup>, David Christian Finger <sup>2,3</sup> , Arvid Bring <sup>1,4</sup>, David Egilson <sup>5</sup>, David Gustafsson <sup>6</sup> and Zahra Kalantari <sup>1,4,\*</sup> 

<sup>1</sup> Department of Physical Geography, Stockholm University, 106 91 Stockholm, Sweden; juliarobin.deniet@gmail.com (J.d.N.); arvid.bring@natgeo.su.se (A.B.)

<sup>2</sup> School of Science and Engineering, Reykjavik University, 101 Reykjavik, Iceland; davidf@ru.is

<sup>3</sup> Sustainability Institute and Forum (SIF), Reykjavik University, 101 Reykjavik, Iceland

<sup>4</sup> Bolin Centre for Climate Research, Stockholm University, 106 91 Stockholm, Sweden

<sup>5</sup> Icelandic Meteorological Office, 105 Reykjavik, Iceland; davide@vedur.is

<sup>6</sup> Swedish Meteorological and Hydrological Institute, 601 76 Norrköping, Sweden; David.Gustafsson@smhi.se

\* Correspondence: zahra.kalantari@natgeo.su.se

Received: 9 January 2020; Accepted: 26 March 2020; Published: 30 March 2020



**Abstract:** The benefits of fractional snow cover area, as an additional dataset for calibration, were evaluated for an Icelandic catchment with a low degree of glaciation and limited data. For this purpose, a Hydrological Projections for the Environment (HYPE) model was calibrated for the Geithellnaá catchment in south-east Iceland using daily discharge (Q) data and satellite-retrieved MODIS snow cover (SC) images, in a multi-dataset calibration (MDC) approach. By comparing model results using only daily discharge data with results obtained using both datasets, the value of SC data for model calibration was identified. Including SC data improved the performance of daily discharge simulations by 7% and fractional snow cover area simulations by 11%, compared with using only the daily discharge dataset (SDC). These results indicate that MDC improves the overall performance of the HYPE model, confirming previous findings. Therefore, MDC could improve discharge simulations in areas with extra sources of uncertainty, such as glaciers and snow cover. Since the change in fractional snow cover area was more accurate when MDC was applied, it can be concluded that MDC would also provide more realistic projections when calibrated parameter sets are extrapolated to different situations.

**Keywords:** glaciated-catchment modeling; conceptual hydrological model; multi-dataset calibration; Hydrological Predictions for the Environment; Geithellnaá; Iceland

## 1. Introduction

### 1.1. Background

Hydrological models are essential for a range of important functions for society, including freshwater resource management, flood prediction, and hydropower development. Limited availability and poor hydrological data quality pose a challenge in achieving satisfactory performance of hydrological models [1,2]. Calibrating hydrological models for glaciated areas is particularly challenging, as ice and snowmelt represents an additional source of uncertainty [3–5]. Models are continually being updated and improved to overcome these and other obstacles, through various approaches to calibration and validation [6]. This study assessed the benefits of including satellite-derived data as an additional dataset for calibration, a process known as multi-dataset

calibration (MDC). To this end, a case study was conducted for which a hydrological model was calibrated using both daily discharge data and daily discharge data in combination with fractional snow cover area data. Compared to other studies examining MDC approaches, this study focused on a remote region with limited hydrological and meteorological data, with a small fraction of glaciation and high fluctuation of snow cover area in the catchment. Potential advantages of MDC in discharge simulation were examined and the implications for hydrological projections under changing conditions were assessed.

Methods to improve calibration and validation approaches have been studied for decades. Early attention focused on the role of objective functions [7,8]. Some studies focused on identifying the optimal temporal resolution and the minimum requirements of data samples [1,9], while others examined the role of parameter uncertainty in the calibration process [10,11]. Recent studies have tested various approaches to calibration [12–14] and have developed and improved methods to calibrate models based on minimal data [9,15].

A key challenge encountered in many studies is that even if a conceptual model can reproduce discharge or rainfall-runoff correctly, it may still be “right for the wrong reasons”, as other components of the water cycle could be simulated incorrectly [16–18]. If input data are limited, this can compromise the accuracy of the model setup, due to the issue of equifinality [19]. Model simulations that are right for the wrong reason become especially problematic when the results are extrapolated to different settings, for example to ungauged catchments or to make hydrological projections under changing conditions, such as climate change [20]. A study within the hydrological research community recently identified the 23 most pressing research questions within hydrology, which include the issue of extrapolating hydrological models under changing conditions (question 19) [21]. To avoid the issue of equifinality, recent research has focused on the correct representation of all relevant components of the water cycle during the calibration and validation phases [1,22,23]. One method to ensure correct calibration of all relevant water cycle components is to use multi-objective calibration strategies. In various studies, parameters have been constrained based on alternative or multiple objectives, such as proper simulation of the hydrograph shape or timing of peak flows, in addition to calibration using daily discharge [1,12,24]. During the 1990s, the idea of MDC developed. It involves calibrating hydrological models using (an) additional dataset(s), besides observed daily discharge, to improve model performance. Calibration based on continuous daily discharge data has long been common practice in hydrological research and continues to be applied in current research (e.g., [1,25]). However, there is a widespread consensus in the hydrological community on the need for additional datasets in the calibration phase, complementing discharge observations, to achieve the correct representation of all relevant variables [26]. One study has shown that calibrating against daily discharge observations is necessary [27], but not sufficient, to correctly simulate all internal variables and physical processes in a catchment. As MDC in hydrological modeling has only recently become more common, there is a clear need for more research on various aspects of the technique, applying it under different conditions.

Parameter regionalization is another strategy used to make hydrological simulations more realistic. With this technique, parameters are constrained by relating them to the physical characteristics of a (sub-)basin, such as land use, elevation, or soil type. Parameter regionalization is suggested to solve the issues of over-parameterization in conceptual models and non-transferability to other basins of calibrated parameter sets [28]. The method can also be used to assess water resources in ungauged catchments, by transferring parameter sets from gauged areas with similar physical characteristics [29]. However, regionalization is challenging in regions with heterogeneous and complex physical characteristics, such as catchments with diverse vegetation, porous underground layers, and glaciated or snow-covered areas [30,31]. Hence, MDC applied in conjunction with parameter regionalization could be valuable in such areas. Various studies have assessed the benefits and drawbacks of parameter regionalization. For example, in one study, the HBV model was calibrated for 95 sub-catchments of the Rhine river basin, by associating parameters to four basin characteristics [32]. The authors found reasonable results but concluded that calibration against observed discharge alone

does not account for the performance of the various components in the runoff generation process. Hence, further research should be conducted on integrating parameter regionalization and MDC to calibrate a conceptual model.

In high alpine areas, fractional snow cover area has proven to be a valuable complementary dataset to daily discharge data to improve the calibration of models, because snow and ice melt contributes to river discharge [22,23,28]. Previous studies have shown that the degree of glaciation in a catchment influences the importance of accurate representation of the glacial melt process in hydrological models [33]. Therefore, the value of adding fractional snow cover area as an additional calibration dataset also depends on the amount of glaciation in a catchment. For example, Finger et al. [23] conducted a study using the MDC approach with snow cover images for three medium glaciated catchments and found that the improvement in performance was greatest for the catchment with the lowest degree of glaciation. The catchment examined in the present study, the Geithellnaá catchment in south-east Iceland, also has a low degree of glaciation, with 4.83% of the catchment covered by the Prándarjökull glacier. However, snow cover area is highly variable throughout the seasons, with the Geithellnaá catchment fully covered in snow during winter and almost zero snow cover during summer. Hence, we tested the hypothesis that the fractional snow cover area can be valuable as an additional calibration dataset.

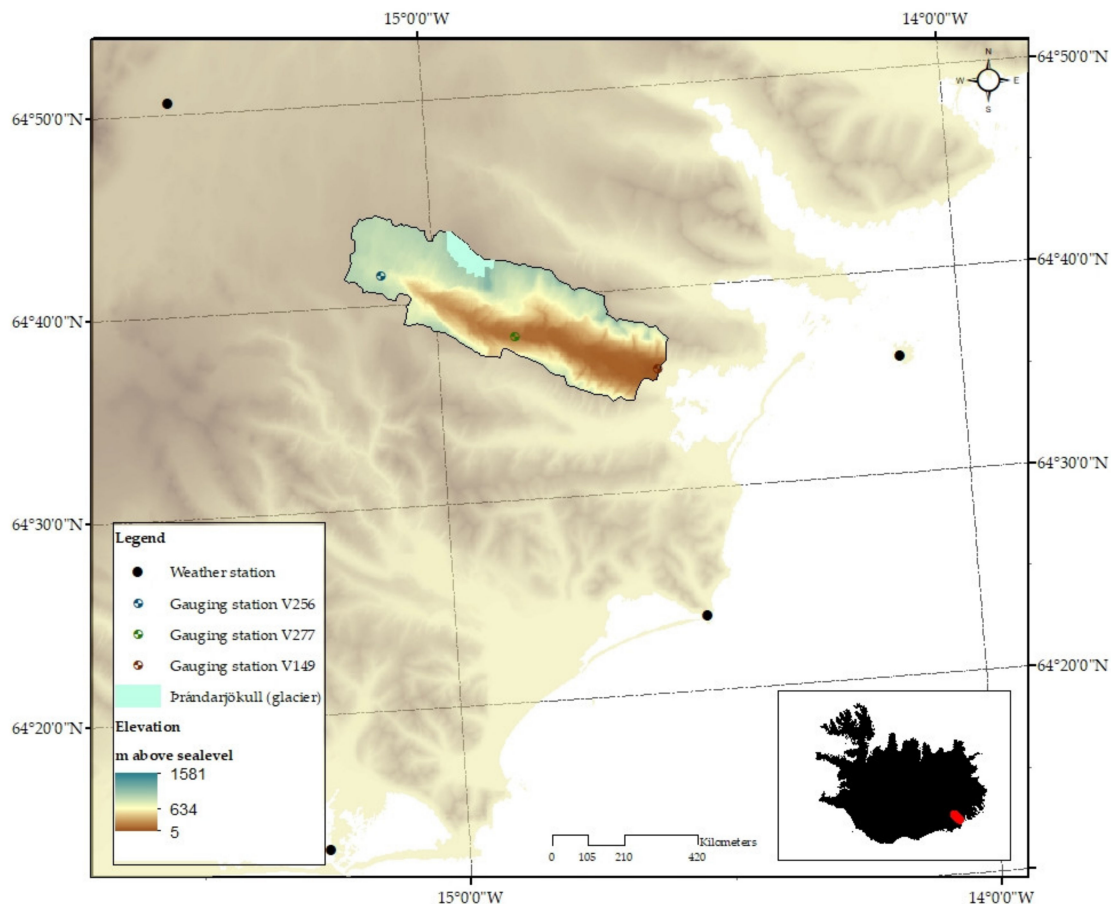
As many catchments around the world lack sufficient ground observations or data of satisfactory quality, earth observations or re-analyzed products from satellites are increasingly being integrated into hydrological models [2]. In this study, Moderate Resolution Imaging Spectroradiometer (MODIS) snow cover data were used to generate fractional snow cover area data for the study area. Previous studies have found earth observations to be suitable as an additional dataset when applying MDC [34]. Examples of additional datasets used for MDC include snow cover [20], glacier mass balance [35–37], a combination of those two [22,23,28], groundwater level [11], and soil saturation [38]. In the present study, the potential of fractional snow cover area as an additional dataset for calibration was evaluated in a low glaciated catchment with limited data (the Geithellnaá valley), using the Hydrological Predictions for the Environment (HYPE) model [39], calibrated using discharge data and MODIS snow cover images. The objective was to assess the benefits of calibrating the model against the two datasets, to improve the internal consistency and overall performance of the model. Compared to previous studies, this study focused on a region with limited data and a catchment with only a small fraction covered by a glacier but a great fluctuation in the amount of snow cover throughout the seasons. The fluctuation in snow-covered area is important, as the objective of the study was to assess the benefits of including satellite-derived snow cover data as an additional calibration dataset. Additionally, due to the remote nature of the catchment, data availability is limited in various ways. The study area has one operational discharge gauging station, while there are four meteorological stations in the vicinity of the catchment. Surface composition data are available from the Corinne database [40], but data about the underlying soil-layers is nonexistent. Furthermore, the abundance of cloud-cover and long winter nights in sub-arctic Iceland limit the use of satellite images, compared to other regions. One of the major benefits of MDC is that it can be valuable in helping to ensure that several relevant variables of a modeled catchment are simulated realistically. Seibert and McDonnell [11,18] claim that MDC can help ensure that a model provides the “right answers for the right reasons”, a claim confirmed by several other studies [22,41] and for most areas in the world, as reviewed by Helmert et al. [42]. Based on these previous studies and the modeling structure in HYPE, we expected discharge and snow cover observations to improve HYPE model performance in the study catchment.

HYPE is a semi-distributed conceptual model developed by the Swedish Meteorological and Hydrological Institute (SMHI) between 2005 and 2007. It is designed to simulate discharge and model flow paths of nutrients in water [39]. HYPE has since been applied in some countries outside Sweden [43]. It is similar to HBV and several other conceptual models in the sense that it uses lumped sub-catchments and a concept similar to hydrological response units (HRUs) to characterize these sub-catchments. HYPE was chosen for this study due to its versatility in performance criteria

and calibration methods [39], and the option to include additional calibration datasets. As HYPE is similar in setup to several conceptual models like HBV, we believe that this paper is of interest to the wider hydrological modelling community. The remainder of this paper is structured as follows: In Section 1.2, the study area is described. Section 2 provides an overview of materials and methods. Results regarding the overall model performance, discharge simulations, and fractional snow cover area simulations are discussed, in that order, in Section 3. The implications of the results obtained are discussed in Section 4, together with the limitations of the study and recommendations for future research. Finally, some conclusions based on the results and discussion are presented in Section 5.

## 1.2. Study Area

The Geithellnaá catchment is located in south-east Iceland, east of the Vatnajökull icecap (Figure 1). The valley was created due to erosion by retreating glaciers after the last Ice Age and has a long, narrow shape, with steep slopes in upper regions and flat lower plains. The National Energy Authority of Iceland (NEA) has classified Geithellnaá as “direct snowmelt runoff” with little to no groundwater contribution. The main river, also named Geithellnaá, collects water from the highlands and a small icecap, Prándarjökull, which is partly located in the valley [44]. The highest peak in the catchment is also the highest point of Prándarjökull and reaches an elevation of 1231 m a.s.l. The glacier accounts for 4.8% of the valley area and approximately half the catchment is covered by vegetation, mainly on the flat lower plains. Accordingly, it can be assumed that glacier melt and snowmelt contribute significantly to the runoff in Geithellnaá. The River Geithellnaá drains a 190 km<sup>2</sup> large watershed and is host to only three farms, near the outlet [44]. The mean annual precipitation is 2680 mm, while mean the annual temperature is 1.1 °C in the upper reaches and 3.9 °C at sea level [45]. Daily temperature and precipitation data for the present analysis were interpolated from four meteorological stations in the vicinity of the study catchment, using the Harmonie model operated by the Icelandic Meteorological Office [45–47]. Monthly mean precipitation and temperature reveal a typical sub-arctic seasonal pattern, with monthly mean precipitation in October of over 10 mm day<sup>-1</sup> in May–June less than 5 mm day<sup>-1</sup> and mean monthly air temperatures in winter of −5°C and around 5°C in summer [29]. The regional weather is characterized by wet ocean air coming from the east [29]. There are three gauging stations in the catchment, V256, and V277. This study relied only on discharge data from gauging station V149, which is located at the estuary of River Geithellnaá. More detailed meteorological data are presented in Finger, 2018 [29].



**Figure 1.** Overview of the Geithellnaá catchment in South East Iceland. Glacial, river, and boundary data from National Geosurvey of Iceland [47].

## 2. Materials and Methods

This section presents the methods and data used in this study. First, an overview of the input data is given, followed by a description of the model, current model configuration, calibration-parameter selection, performance criteria and, finally, calibration methods.

### 2.1. Input Data and Calibration Data

Based on data availability, it was decided to use 1 September 2005–31 December 2010 for calibration and 1 January 2011–26 December 2015 for validation. HYPE uses daily time series of precipitation (Pobs) and temperature (Tobs). For the MDC simulations, daily fractional snow cover data (Fsc) were included as an additional dataset. Full data coverage for all datasets was available for the period 1 September 2005–26 December 2015.

Temperature and precipitation time series were created by averaging interpolated gridded Harmonie datasets, which comprise re-analyzed weather predictions [45], for the period 2000–2017. The Harmonie datasets were selected over other datasets as their temporal coverage overlaps most with the other observed datasets (the data can be accessed through a linked repository in MyGeohub, DOI:10.13019/G5BM-SD63). Table 1 shows mean observed daily precipitation and temperature during the whole calibration period and per year of the validation period.

**Table 1.** Daily meteorological conditions (mean and standard deviation (Std)) for the calibration period and each year in the validation period.

Period	Precipitation (mm/day)		Temperature (°C)	
	Mean	Std	Mean	Std
Calibration	7.2	14.1	−0.3	5.1
2011	8.3	13.5	−0.3	5.5
2012	6.5	12.4	−0.7	5.6
2013	6.3	11.9	−0.6	6.0
2014	9.8	17.1	0.5	4.4
2015	8.9	17.6	−0.9	4.8

Continuous discharge data for calibration and validation purposes were only available from gauging station V149, operated by the Icelandic Meteorological Office (IMO). It was decided to use data from this station due to its downstream location (Figure 1) and sufficient data availability (continuously available for the periods 1972–2000 and 2005–2015). The two other gauging stations in the catchment, V256 and V277, are manual stations with only a few sporadic discharge measurements.

Fractional snow cover area data were computed from gridded data, using the MOD10A.1 dataset, version V006 [48]. These snow data consist of satellite images obtained with MODIS using two satellites, Terra and Aqua [48]. Both the Terra and Aqua datasets, which are available from 2000 and 2002 onwards, respectively, were used. Gridded data with 500-m resolution were averaged for the whole catchment. Fractional snow cover area was calculated by dividing the snow-covered area by the whole area of the catchment. Due to frequent cloud-covered days, it was decided that only days with less than 40% cloud cover data would be used for calibration and validation purposes. This resulted in an average of 65 days per year with useable data, or 18% of days.

## 2.2. Model Description

HYPE uses lumped sub-catchments, all of which are defined by area and divided into soil type and land use combinations (SLCs), a concept similar to HRUs. A more detailed description of the HYPE model can be found in Lindström et al. [39] and on SMHI.net [43]. For the creation of sub-catchments in the Geithellnaá catchment, we used a digital elevation model (DEM) with resolution 10 m obtained from the National Land Survey of Iceland [47]. With these raster data, sub-catchments were created in ArcMap (v.10.2). After tests with varying numbers of sub-catchments, it was found that a minimal number of sub-catchments led to the highest performance of the model. Hence, the area was divided into three sub-catchments, with the outlets located at the three gauging stations. Information provided to the model about the sub-catchments included area, elevation (mean and standard deviation), soil type and land use characteristics, and downstream sub-catchments. For each sub-catchment, a water balance was calculated for each time step, including precipitation, evaporation based on temperature, storage in snow, water, and soil, inflow from upstream catchments, and outflow. Each of these processes was simulated based on various parameters and aspects of the model setup. Various sub-models can be added to the structure to determine how specific processes are generated, but if a sub-model is included, separate parameters incorporated in this process have to be defined (see SMHI.net [43]).

For the creation of SLCs, soil data were obtained from the Icelandic Agricultural Research Institute [49], and five different soil classes were identified within the catchment, including “glacier”. For land cover data, the EU Corine Land Cover dataset (Version 18.5.1) was used [40]. Nine different land cover types were initially identified within the catchment and were reclassified to six, including “water” (lake) and “glacier”. SLCs were created by overlaying the soil type and land cover layers. Based on the five soil classes and six land cover types identified, this led to a potential total of 30 different SLCs. However, both the land and soil type classes included “glacier”, and “water” was one of the land cover types. Glaciers and lakes are simulated as separate classes in HYPE, existing as only one layer without sub-surface water flow. Hence, only 18 combinations of SLCs were possible,

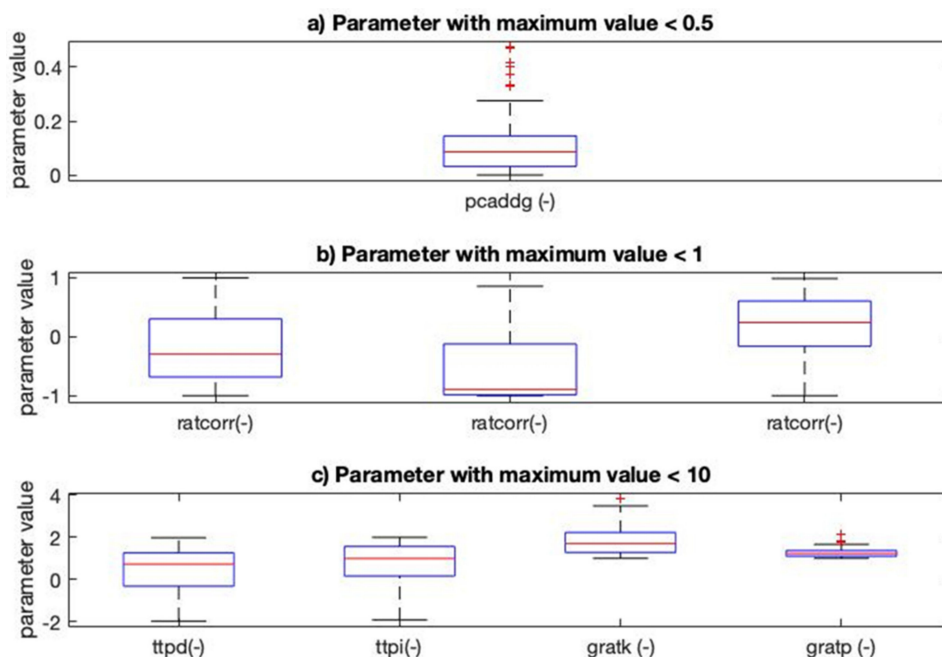
16 of which were found in the catchment. For each sub-catchment, the coverage of each SLC was expressed as a fraction. SLCs do not require information about elevation or geographical location in the sub-catchment, but, for each SLC, soil depth has to be defined in the model setup. Glacier mass is calculated per time step, using a coefficient and exponent parameter for the area–volume relation and a factor for the density of the ice (see SMHI.net [43]).

Fractional snow cover area was calculated per land use type for every sub-catchment. There are two different functions in HYPE, for the snow accumulation and ablation phases. During the snow accumulation phase, fractional snow cover is a function of snow water equivalent (snow) and the maximum snow cover equivalent during winter (fscmax) [43,50]. Snow water equivalent (snow) is calculated per time step and is dependent on snowmelt and snowfall. Snowmelt occurs when the observed temperature is above the threshold temperature (ttmp) and the amount is determined by the snowmelt parameter (cmlt). Similarly, snowfall occurs on days with precipitation when air temperature is below the threshold temperature (ttmp). Both cmlt and ttmp are land-use-dependent, meaning they vary with land use type. As soon as a threshold value is reached for fractional snow cover (fscmax-fsclim), the model switches to the ablation phase and uses a function that represents snow cover during winter. Fractional snow cover is then calculated based on snow water equivalent, maximum snowpack during winter (snowmax), and a distribution factor. The distribution factor is governed by the parameters fscdistexp, fscdist0, and fscdist1, and is dependent on the standard deviation of elevation in the sub-catchment. In this study the maximum value, fscmax, was set at 1, following observed data, while minimum snow cover, fscmin, was set at 0.048, as this is the fractional area of the glacier.

### 2.3. Parameter Selection and Calibration

In total, 74 parameters were included in the HYPE setup for this study. Initially, 40 parameters were assessed for sensitivity, to test which parameter values had the most influence on the performance of the model (Figure A1 in Appendix A). These 40 parameters were tested because they are incorporated in the general flow simulation and not related to any sub-models included in the model structure. Those sub-models and the other 34 parameters were defined separately, based on different studies or default values suggested by SMHI [43]. The sensitivity of the 40 general parameters was assessed using the Monte Carlo approach with 50,000 runs. Based on the 100 parameter sets with the highest Nash–Sutcliffe efficiency (NSE) value, six parameters were selected for calibration, as model performance proved most sensitive to variation in the value of these parameters (Figure A1). Figure 2 shows the selected parameters and the distribution of the values of the 100 best parameter sets for the calibration. The selected parameters showed more convergence within their given range than other parameters (Figure A1), indicating that variation in the value of these six parameters influenced model performance. These six parameters were further calibrated, applying MDC and single dataset calibration (SDC), again using the Monte Carlo approach with 50,000 runs. The values of the other parameters were based on other studies and values suggested as common or default by SHMI in the model description [43].

None of the parameters selected for final calibration have a direct influence on HYPE simulation of snow cover area (Table 2). However, parameters ttpd and ttpi are correcting parameters for the threshold temperature for snowmelt and snowfall, and thus influence the snowpack. Parameter pcaddg influences precipitation distribution based on elevation, which in turn influences fractional snow cover area. One of the parameters included (ratcorr) is region-dependent and was hence calibrated per sub-catchment, also known as parameter regionalization.



**Figure 2.** Distribution of parameter values for the 100 best parameter sets obtained during Monte Carlo analysis for the calibration period (1 September 2005–31 December 2010) when parameters were given an upper range limit of: (a) 0.3 (pcaddg), (b) 1 (ratcorr), and (c) 2 (ttpd, ttpi) and 4 (gratk, gratp). Whiskers indicate the maximum and minimum values found during the Monte Carlo analysis among the 100 best parameter sets, while the red line is the mean and the boxes represent the first and third quartile of the value distribution. The parameters were tested in various setups, with different combinations of input datasets. Some parameters showed little convergence but were important for model performance in different setups.

**Table 2.** Parameters used for calibration and fractional snow cover area simulations.

Abbreviation <sup>1</sup>	Units	Type <sup>2</sup>	Description
ttpd	°C	General	Deviation from threshold temperature for snow-/rainfall
ttpi	°C	General	Half of temperature interval with mixed snow and rainfall (temperature interval = (threshold temperature + ttpd) +/- ttpi)
ratcorr	-	Parameter region	Correction factor for discharge (calculated per sub-catchment).
gratp	-	General	Corrects parameter gratk to different regions
gratk	-	General	Parameter for rating curve for lake outflow
pcaddg	-	General	Parameter for rating curve for lake outflow
fscmax	-	General	Spatial correction parameter for precipitation depending on elevation
fscmin	-	General	Maximum fractional snow cover area
fsclim	-	General	Minimum fractional snow cover area
fscdistmax	-	Land-use	Limit of fractional snow cover area for onset of snowmax
fscdist0	-	Land-use	Maximum snow distribution factor
fscdist1	m <sup>-1</sup>	Land-use	Minimum snow distribution factor
fsck1	-	General	Std coefficient for snow distribution factor
fsckexp	s <sup>-1</sup>	General	Parameter for snowmax
			Parameter for snowmax

<sup>1</sup> Code used by HYPE. <sup>2</sup> Whether the parameter is general or specified for a certain land use, soil type, or parameter region (parreg, i.e., a sub-catchment).

#### 2.4. Performance Criteria

Model performance was evaluated using an overall performance criterion based on the average of two separate criteria, one for discharge simulations and one for fractional snow cover area. HYPE offers a range of performance criteria to cater for various variables that can be included (a full list can be found at SMHI.net [43]). Here, Nash-Sutcliffe Efficiency (NSE) was used to evaluate discharge



simulations and normalized root mean square error (NE) was chosen for the fractional snow cover area. Average NSE was calculated for the whole Geithellnaá-catchment as [51]:

$$NSE = 1 - \frac{\sum_{i=1}^{mi} (Q_{obs} - Q_{sim})^2}{\sum_{i=1}^{mi} (Q_{obs} - \overline{Q_{sim}})^2} \quad (1)$$

where  $Q_{obs}$  is observed discharge,  $Q_{sim}$  is modeled discharge, and  $mi$  is the number of values in a time series of data from a meteorological station.

To evaluate simulations of fractional snow cover area, the mean NE value was calculated for all sub-catchments [43], as:

$$NE = \frac{\sqrt{\frac{1}{mi} \sum_{i=1}^{mi} (F_{scobs} - F_{scsim})^2}}{\max(F_{scsim})} \quad (2)$$

where  $F_{scobs}$  is observed fractional snow cover,  $F_{scsim}$  is simulated fractional snow cover, and  $mi$  is the number of values in a time series of a sub-catchment.

To determine overall model performance, NSE and NE values were averaged. NE should be minimized, meaning that better performance of the simulation leads to a lower NE value. For NSE the best score is 1, meaning that better performance leads to a higher value [43,51]. To combine the two performance criteria with opposing maximum scores, the following formula was used:

$$\text{Overall performance} = xNSE - yNE \quad (3)$$

where  $x$  and  $y$  refer to the weights given to NSE and NE, respectively. HYPE allows the user to assign weights to each evaluation criterion when using multiple criteria, with the sum equaling 1. For this study, both performance criteria were given a score of 0.5. We decided to weigh both the NSE and NE scores equally, as we deemed correct simulation of discharge and fractional snow cover area to be equally important when applying the MDC approach.

### 2.5. Calibration Using SDC and MDC

In order to evaluate SDC simulations, only the NSE values were considered, as only daily discharge was used as calibration dataset. To evaluate MDC simulations, NE and NSE values were averaged, as both daily discharge and fractional snow cover area data were used as calibration sets. Only the parameters listed in Table 2 were calibrated, while all other parameters were kept the same in both MDC and SDC. For the validation phase, the optimal parameter sets obtained during the calibration phase were used to simulate discharge and fractional snow cover area. The overall performance and the performance of discharge and fractional snow cover area simulations were compared.

## 3. Results

In this section, overall model performance is presented first, followed by the discharge simulations and fractional snow cover area simulations, respectively.

### 3.1. Overall Performance

The results showed that MDC using discharge data and fractional snow cover led to 17% higher overall performance of the model than SDC during the validation phase (January 2011–December 2015). The largest relative improvement derived from the higher performance of the fractional snow cover area simulations with MDC, which increased model performance by 12% from 0.25 to 0.22 (Table 3, lowest column). The largest absolute improvement with the inclusion of fractional snow cover area for calibration was found in the discharge simulations, with an increase from 0.73 to 0.78.

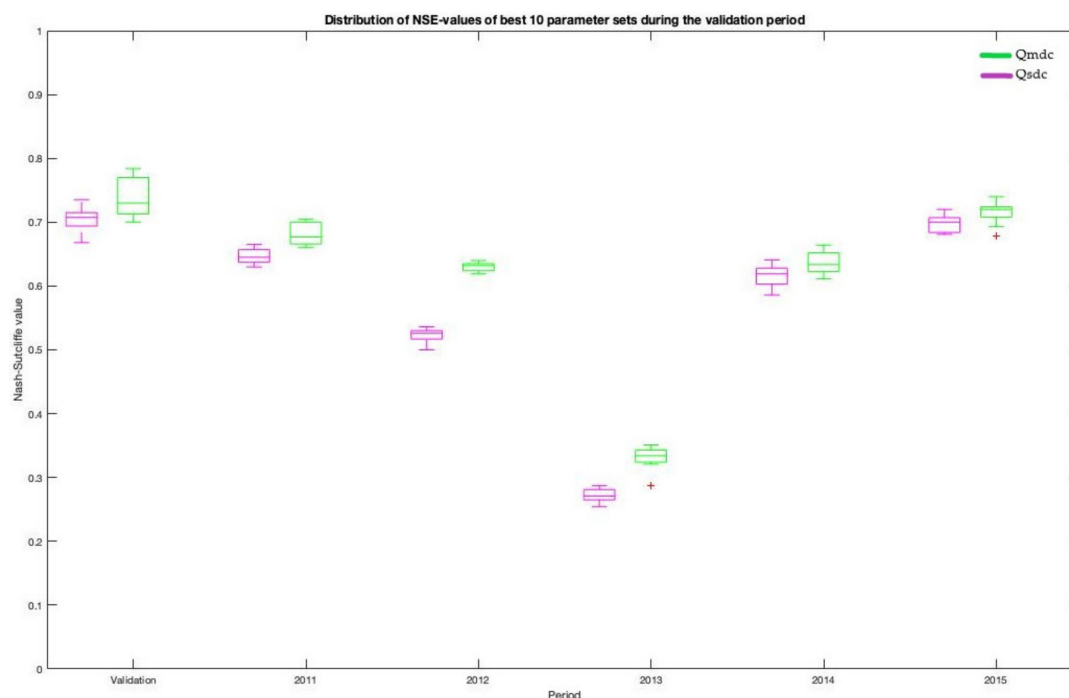
**Table 3.** Performance of single-dataset and multi-dataset calibration (SDC and MDC) in the validation period, based on different criteria.

Criterion <sup>1</sup>	SDC	MDC	Relative Change (%)
Overall performance	0.24	0.28	17
Nash-Sutcliffe efficiency (NSE) (for discharge data)	0.73	0.78	7
Normalized root mean square error (NE) (for fractional snow cover data)	0.25	0.22	12

<sup>1</sup> Computed using equations 3, 1, and 2, respectively. The values shown are for the best datasets only.

### 3.2. Discharge Simulations

During the validation period, the MDC method resulted in more realistic simulations of daily discharge, with an average NSE value of 0.78, compared with 0.73 with SDC, an improvement of 7%. Moreover, the parameter sets created using MDC led to higher NSE values during each individual year of the validation period (Table 3). Figure 3 shows the NSE values for the 10 highest performing parameter sets when considering the overall performance criterion.

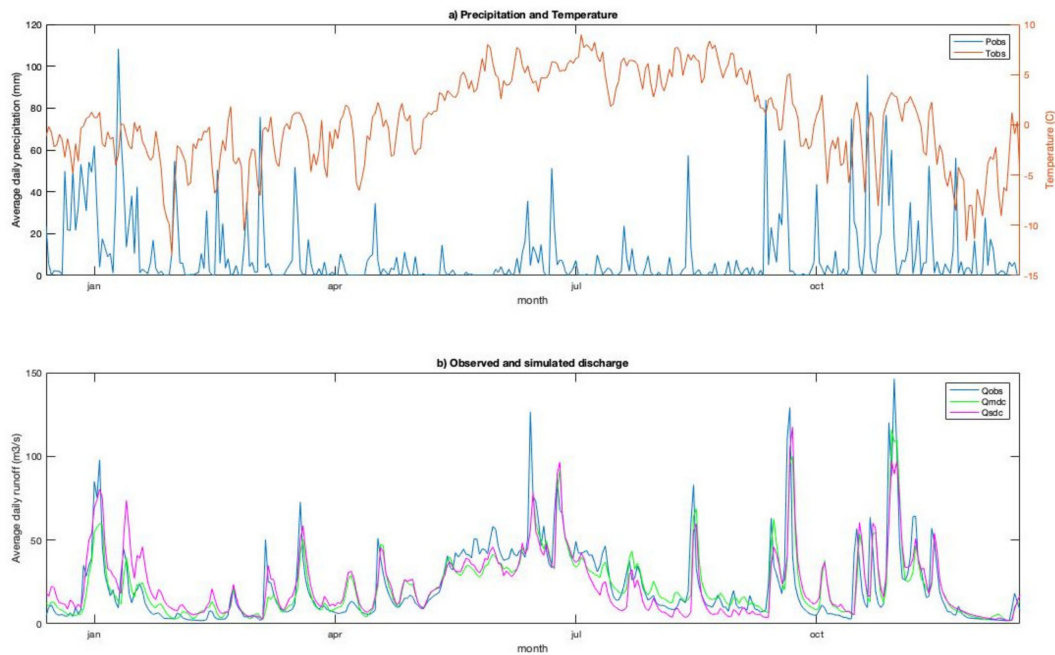


**Figure 3.** Distribution of Nash–Sutcliffe efficiency (NSE) values for the 10 best-performing parameter sets (when considering the overall performance criterion) for SDC and MDC simulations during the entire validation period (1 January 2011–26 December 2015) and in each year.

Meteorological circumstances and changes in these seemingly influence the ability of the model to simulate discharge correctly. During years when temperature and precipitation were similar to the calibration period, such as 2011 and 2015, model performance was high; SDC led to (almost) equally realistic discharge simulations (Figure 3). Conversely, for those years in the validation period when meteorological conditions deviated most from the calibration period, model performance was lowest, and the difference in performance between simulations created with MDC and SDC was greatest (Figure 3, Table 1). During the years with the lowest performance, 2012 and 2013, both temperature and precipitation were low on average compared to the calibration period, except for several peaks in temperature in summer (Table 1). Yet, since model performance was higher for the

MDC simulations, this suggests that including fractional snow cover data leads to model calibrations that are less susceptible to changing conditions. It should be noted that, despite the increase in performance when applying MDC, the discharge simulations during 2013 were not of sufficient quality, with a mean NSE value of 0.34 for the best 10 parameter sets (Figure 3).

Figure 4 shows a timeseries of meteorological circumstances and discharge simulations in 2014, which was taken as a representative year for the results during the validation period.



**Figure 4.** Observed precipitation and temperature, and observed and simulated discharge, during the validation period (2014). (a) Observed temperature and precipitation data from the Harmonie dataset [46]. (b) Observed and simulated discharge using single-dataset and multi-dataset calibration (MDC and SDC).

The results show that discharge is high when either precipitation or temperature is high: the graph displays discharge peaks during January and October/November when precipitation is high and in June/July when observed temperatures are relatively high (Figure 4).

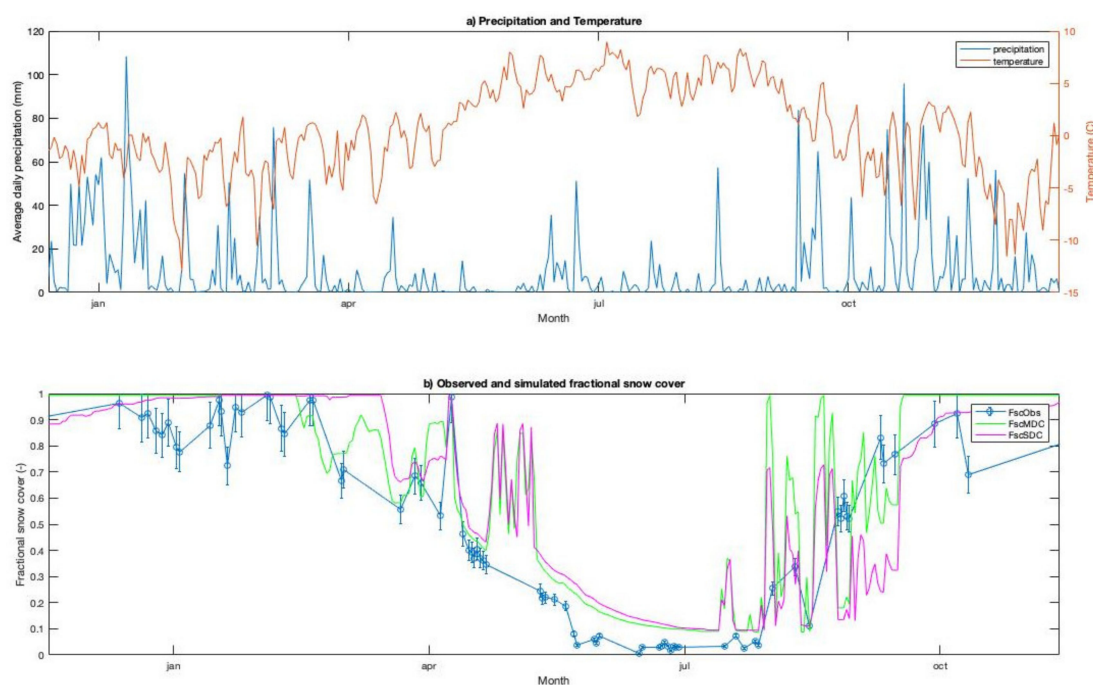
### 3.3. Fractional Snow Cover Area Simulations

Fractional snow cover area simulations created with HYPE were compared to MODIS snow cover data. Fractional snow cover area was modeled more realistically when MDC was applied during the overall validation period (Table 3) and for each year during this period (Table 4). During 2012, 2014 and 2015 fractional snow cover area simulations were most realistic. The performance of fractional snow cover was lower during 2011 and 2013, and, corresponding to the discharge simulations, the difference in performance between SDC and MDC is largest for these years. Interesting in this case is the year 2012, for which discharge simulations are relatively low and the difference in performance is highest between MDC and SDC. Fractional snow cover area simulations are relatively realistic for the MDC simulations but less so in the SDC simulations, compared to other years of the validation period. A similar observation can be made for 2011, when the difference in the performance of fractional snow cover simulation is the same, and there is a difference in performance for discharge simulations despite overall model performance being high that year. These results imply, together with the observation that MDC is less susceptible to changes in meteorological circumstances, that MDC simulations are more likely to be right for the right reasons.

**Table 4.** Normalized root-mean-square error (NE) values for fractional snow cover area simulations when applying single-dataset and multi-dataset calibration (MDC and SDC) for each year during the validation period.

Year	2011	2012	2013	2014	2015
SDC	0.28	0.23	0.31	0.22	0.21
MDC	0.22	0.19	0.28	0.21	0.20

Figure 5 shows the meteorological circumstances and fractional snow cover area timeseries for 2014. The simulations with MDC and SDC follow largely the same trend. There are two periods with noticeable differences: during the snowmelt period (February–April) when the MDC simulations show an earlier onset of snow-melt trends, and during the snow-build period (October–December). This is reflected in the discharge simulations (Figure 4), where there is a delay in discharge, leading to an overestimation during the period of February to April.



**Figure 5.** Observed and simulated precipitation, temperature, and fractional snow cover area during the validation period (2014). (a) Observed temperature and precipitation data from the Harmonie dataset [46]. (b) Observed and simulated fractional snow cover using single-dataset and multi-dataset calibration (MDC and SDC).

#### 4. Discussion

The results obtained indicated that including satellite-derived fractional snow cover data as an additional dataset for calibration helps to improve hydrological models in glaciated areas with limited data. Calibration with both datasets led to more realistic simulations of fractional snow cover and discharge for almost every year of the validation period, compared with SDC (Tables 3 and 4, Figure 3). The difference in accuracy between SDC and MDC was particularly apparent during years when model performance was low, such as 2012, 2013, and 2014 in the validation period. Data requirements for the additional dataset are low, as shown in this study, where only 18% of days could be used for fractional snow cover observation data. The improved overall model performance with MDC supports earlier findings (e.g., [1,20,22,23,31]).

Besides improving discharge simulations, MDC could be particularly valuable when extrapolating parameter sets, for example to ungauged catchments or to make discharge projections. Parameter

regionalization has been proven to be useful when extrapolating constrained parameter sets to ungauged areas [28]. However, the technique becomes challenging when applied to catchments with high uncertainty, for example with diverse vegetation, porous underground layers, and glaciated or snow-covered areas [30,35]. The results of the present study show that MDC could be useful in overcoming this challenge, since fractional snow cover area simulations created with MDC were consistently more accurate during the validation period. This indicates that some internal processes related to snow cover were simulated less accurately when calibrating towards daily discharge only. In other words, the SDC simulations were right for the wrong reasons. Model performance when simulating discharge could be influenced by changing circumstances, such as different meteorological conditions or physical characteristics when applied in another catchment. Indeed, as concluded in the results section, during years in the validation period when meteorological conditions deviated most from the calibration period, the difference in performance between simulations created with MDC and SDC was greatest (Figures 4 and 5, Table 1). It can be concluded that applying MDC leads to model calibrations that are less susceptible to changing conditions. Hence MDC could be valuable when extrapolating model setups, particularly in combination with parameter regionalization as in this study, for example for ungauged catchments or climate change projections. In this study, it was not possible to carry out actual climate projections, due to limited temporal data coverage.

The degree of glaciation in a catchment influences the quality of discharge simulations, mainly as the difference in performance between SDC and MDC. Finger et al. [23] conducted a study using MDC with snow cover images for three separate catchments and found that the improvement in performance was highest for the catchment with the lowest degree of glaciation. When a smaller fraction of the catchment is glaciated, the fractional snow cover area in the catchment is more variable throughout the year. Geithellnaá has a relatively low degree of glaciation (Figure 1). Therefore, based on our results and those in previous studies, it can be concluded that applying MDC with fractional snow cover area is more beneficial for the overall performance in low to medium glaciated catchments. In highly glaciated catchments, glacier mass balance would be more appropriate as an additional dataset for calibration purposes, as glacier mass balance has been shown to contain important information that could improve the reliability of simulations [52]. MDC can also improve discharge simulations in non-glaciated catchments, as shown in various studies (e.g., [1,12,20]). For regions with seasonal snowfall, MDC could be applied using fractional snow cover area as an additional dataset, while in areas in warmer climates other variables, such as groundwater level or soil saturation, could be used. It could be argued that MDC has more benefits in glaciated areas or regions with seasonal snow coverage, as snow and ice are an additional source of water, and hence add uncertainty [3–5].

While calibrating MDC can improve model performance, several issues might arise that can compromise the quality of the simulations, as noted by Freer et al. [53]. First, as the input data may include uncertainty, researchers risk degrading the quality of their simulations and the ability of the model to represent a system when including additional datasets [53]. In this study, the frequent cloud cover was a major source of uncertainty regarding the observed fractional snow cover data. Despite this uncertainty, the overall model performance improved with MDC and the model also simulated discharge better, indicating that the snow cover data were of sufficient quality. Second, the chosen additional dataset should be selected carefully, to be commensurate with the model setup [53]. The additional variable used for discharge calibration should have a significant influence on runoff generation throughout the year. The results obtained in the present and previous studies (e.g., [54]) indicate that accurate modeling of snow cover is important when attempting to model discharge in glaciated catchments. In this study, when the performance of fractional snow cover simulations was low, as in 2013, discharge simulations were also modeled less accurately. Given the importance of accurate modeling of snow cover in our study setting in Iceland, fractional snow cover can be considered an appropriate choice of an additional dataset for calibration. Hence, our results confirm that the additional dataset should be chosen with care.

The results obtained in this study indicate that HYPE is a suitable model for use with the MDC method and that adding an additional dataset is advisable for users of the open-access version of HYPE. HYPE is versatile and allows for easy inclusion of additional datasets and variables. It is possible to average performance criteria to evaluate the overall performance, making it easy to calibrate towards multiple datasets at once. One drawback of HYPE is the limited choice of performance criteria available for use in the calibration process. For example, in this study it was initially decided to use mean absolute error (MSE) to evaluate the performance of fractional snow cover area simulations, but this option was not available in the version of HYPE used [55,56].

#### *Uncertainty, Limitations, and Future Research*

Simulation results are only as reliable as the inputs, parameter estimates, and underlying assumptions used in the model [7]. Here, MODIS snow cover data, which were frequently obscured and included 40% cloud cover, were used for calibration purposes. An additional limitation is that snow cover fractions were computed for the whole catchment, while HYPE requires input per sub-catchment for additional observation data. Hence for every sub-catchment, the whole catchment average was given, which deviates from reality. Furthermore, the precipitation and temperature data were based on modeled interpolated data, as there were no meteorological stations within the catchment. Both the interpolation method and the averaging of spatial data add some uncertainty.

Based on the results and conclusions presented here, some recommendations can be made for future research. This study had limited time coverage due to the availability of observed data. Ultimately, the benefits of MDC for hydrological projections should be studied over a longer period, to test whether MDC does indeed lead to improved hydrological projections under climate change. Furthermore, a study using multiple catchments could lead to an enhanced in-depth analysis of the benefits of MDC. As MDC was found to lead to satisfactory results, further studies can employ other options for additional calibration datasets. The present study could be repeated for non-glaciated catchments with temporary snow cover during winters, to assess the usefulness of fractional snow cover data in such cases. For glaciated catchments, the application of MDC for discharge projections under future climate change should be examined for various types of models, as in Etter et al. [22].

## **5. Conclusions**

This study examined the possible benefits of using satellite-derived snow cover area data as an additional calibration dataset (MDC) in a low-glaciated catchment with limited data. A case study was carried out using the Swedish HYPE model to simulate daily discharge in the Geithellnaá catchment in south-east Iceland. The results revealed that simulations using MDC with snow cover area as an additional calibration dataset led to 17% better overall model performance than simulations using only discharge observations, or Single Dataset Calibration (SDC). Both discharge and fractional snow cover area were modeled more realistically when applying MDC, with the performance of discharge simulations created with MDC improving by 7% and that of fractional snow cover area simulations by 12%. This improvement was found despite limited days with fractional snow cover area observations due to the abundant cloud cover in Iceland. We concluded that, for this case study, the MDC simulations were likely to be right for the right reasons. The performance of the model with SDC decreased most during years in the validation period when meteorological conditions differed most from the calibration period. The fractional snow cover area simulations created with MDC were more realistic for every year of the validation period than those created with SDC. This indicates that internal hydrological processes in the study area are likely to be more realistic when modeled using MDC. Since the MDC results are more likely to be right for the right reasons, it can be concluded that including satellite-derived data as an additional dataset for calibration has potential when working in low-glaciated catchments with low data. For example, MDC with fractional snow cover area could be valuable when extrapolating constrained parameter sets to ungauged catchments or when making hydrological projections in remote, glaciated areas with limited data. Finally, we concluded that the

data requirements for additional calibration datasets are low, as only 18% of the days had usable snow cover area observations.

**Author Contributions:** Conceptualization: D.C.F., J.d.N.; Methodology: D.C.F., J.d.N.; Software: J.d.N., D.E.; Validation: J.d.N., D.G.; Formal Analysis: J.d.N., D.C.F.; Investigation: J.d.N.; Resources: D.E.; Data Curation: D.G.; Writing—Original Draft preparation: J.d.N.; Writing—Review & Editing: Z.K., D.C.F.; Visualization: J.d.N., Z.K.; Supervision: D.C.F., Z.K., A.B.; Project Administration: Z.K.; Funding Acquisition: N/A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding. During the period in which this study was conducted, the authors received funding from their respective institution, namely the Department of Physical Geography (Stockholm University), School of Science and Engineering (Reykjavik University), Sustainability Institute and Forum (Reykjavik University), Bolin Centre for Climate Research (Stockholm University), Icelandic Meteorological Office and the Swedish Meteorological and Hydrological Institute.

**Acknowledgments:** We would like to extend our gratitude to the Icelandic Meteorological Office for supplying the data and much useful advice and feedback. We would also like to thank the Swedish Meteorological and Hydrological Institute for making HYPE freely available.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Data Availability:** The data used in this study can be accessed through a linked repository in MyGeohub, (doi:10.13019/G5BM-SD63).

## Appendix A

Figure A1 shows the distribution of values for all 40 parameters that were tested for sensitivity. They were included as they are related to the main flow generation process in the model, and not to any sub-models such as snowmelt or glacier melt. All 40 parameters were tested in a large Monte Carlo analysis, with 50,000 runs. Figure A1 shows the distribution of values for the 100 highest performing parameter sets for each individual parameter. For many parameters, the distribution showed a “normal” shape, with the mean roughly in the middle of the range of possible values for that parameter, the first and third quartile at around one quarter and three-quarters of the range, and the minimum and maximum at the ends of the range. Other parameters showed clear divergence within their set range of possible values in the Monte Carlo analysis (e.g., *ssrrate* (panel D) and *preccorr* (panel F)). Convergence showed that the model was sensitive to variation in the value of a parameter, as the best performing 100 parameter sets showed some consistency as regards the value of this parameter. For the six parameters used in the final calibration of the model, the convergence in model values was strong (see Figure 2 in main text).

It should be noted that for the final calibration, only six parameters were included in a Monte Carlo analysis with 50,000 runs. For all parameters presented in Figure A1, values suggested by the creators of HYPE or in other literature were used. We realize that 50,000 runs is low for Monte Carlo analysis with 40 parameters, but this initial test was intended solely to determine which model parameters were most important for model performance, not for model calibration purposes.

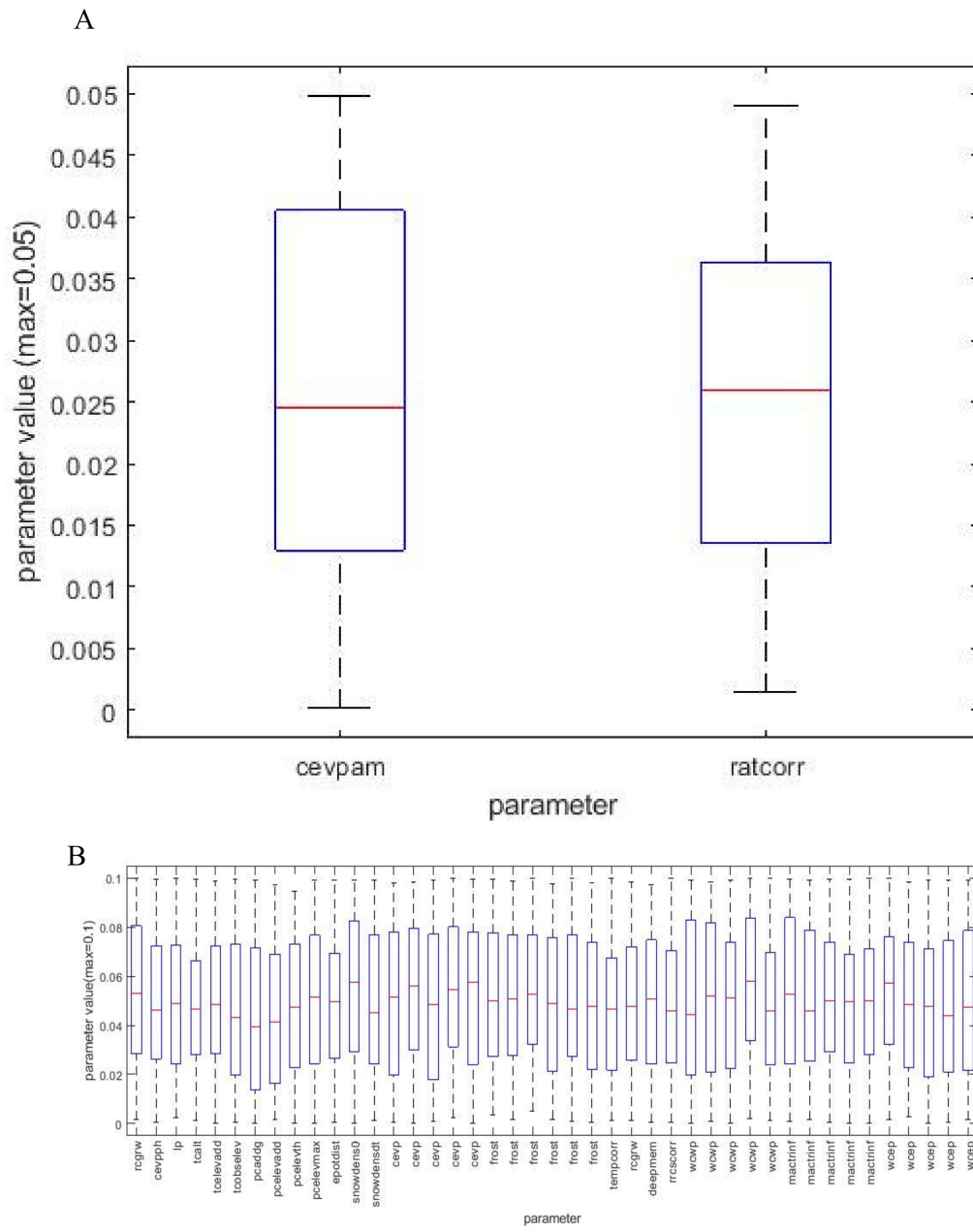


Figure A1. Cont.



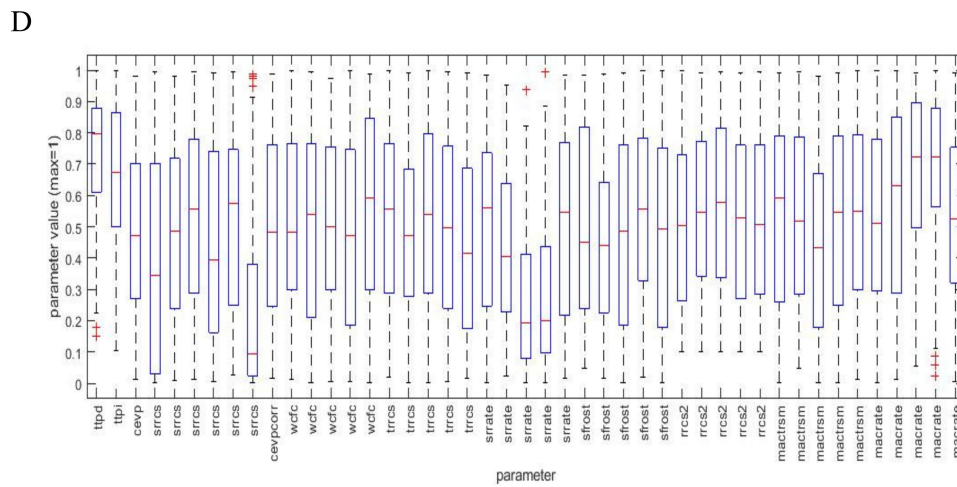
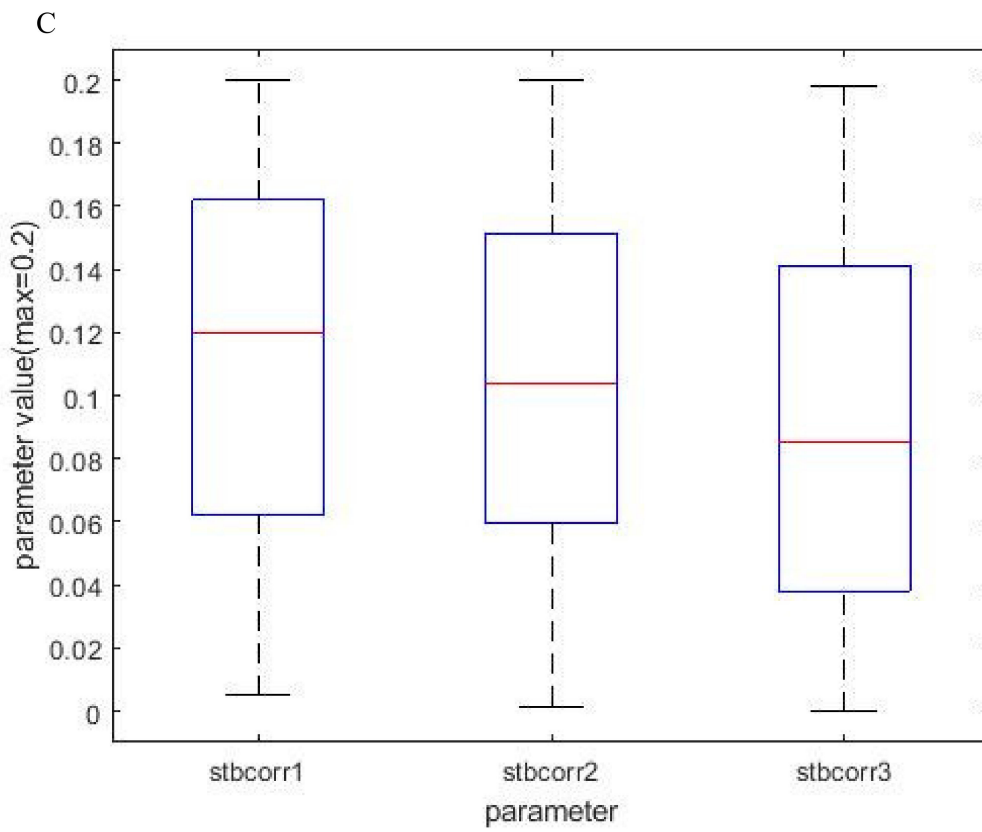


Figure A1. Cont.

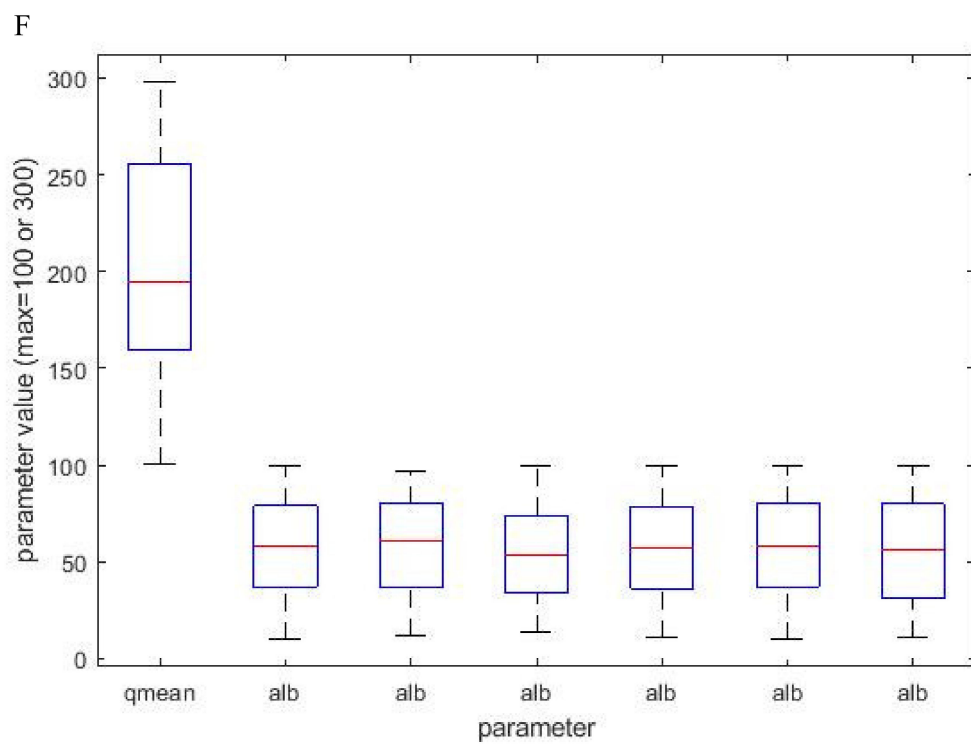
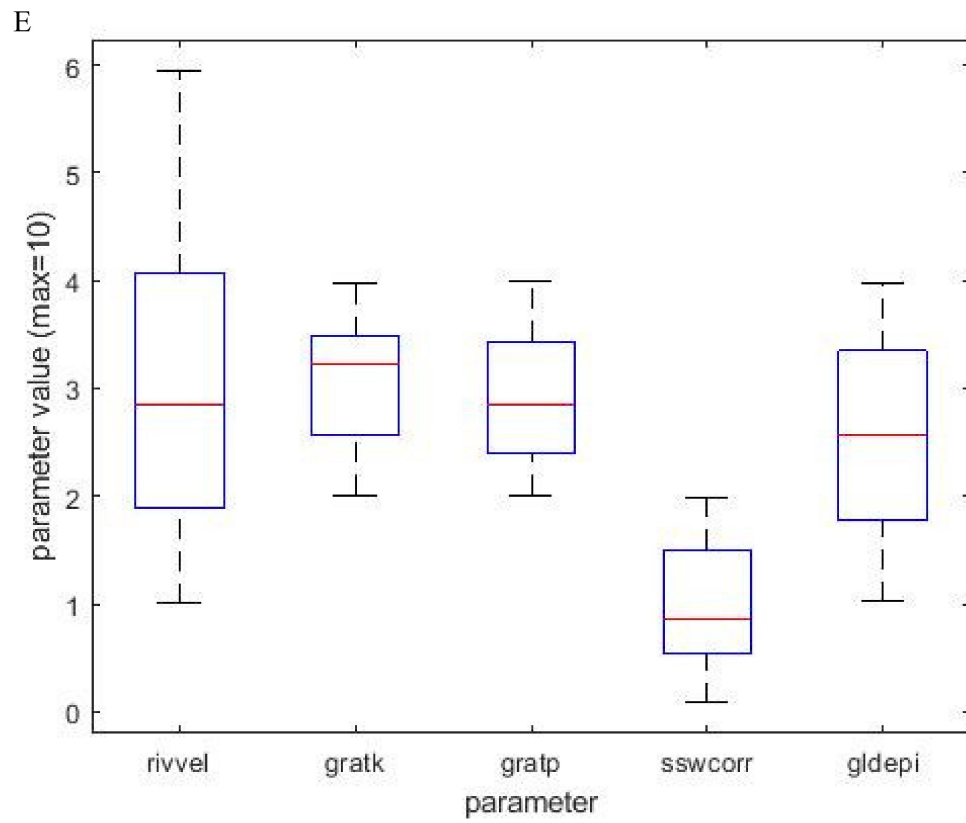
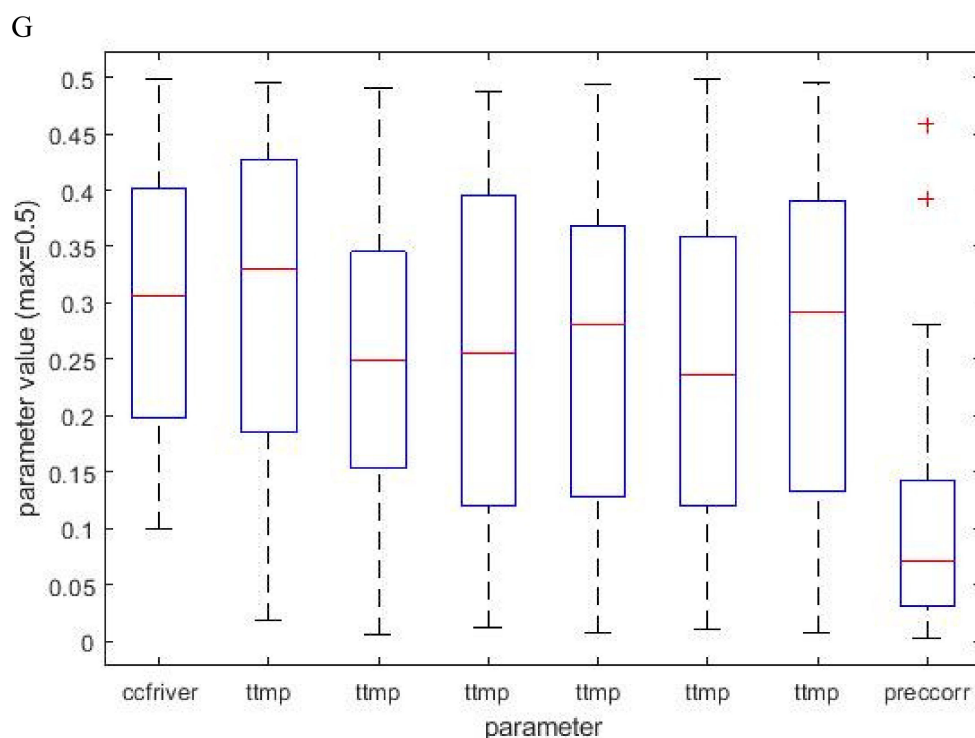


Figure A1. Cont.



**Figure A1.** Value distribution in the 100 best parameter sets for all 40 parameters used in initial Monte Carlo analysis. For clarity, the parameters are split between seven different diagrams based on the maximum value in the Monte Carlo analysis. Calibration was carried out using all four possible combinations in precipitation and temperature datasets. These diagrams were created using Harmonie data [46] for both precipitation and temperature, during the calibration period (1 September 2005–31 December, 2010). (A) Parameters with a maximum range in the MC-analysis of 0.05. (B) Parameters with a maximum range in the MC-analysis of 0.1. (C) Parameters with a maximum range in the MC-analysis of 0.2. (D) Parameters with a maximum range in the MC-analysis of 1. (E) Parameters with a maximum range in the MC-analysis of 2 (sswcorr), 4 (gratk, gratp and gldepi) or 6 (rivvel). (F) Parameters with a maximum range in the MC-analysis of 100 (albedo) or 300 (qmean). (G) Parameters with a maximum range in the MC-analysis of 0.5.

## References

1. Juston, J.; Seibert, J.; Johansson, P.O. Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrol. Process.* **2009**, *23*, 3093–3109. [\[CrossRef\]](#)
2. Kalantari, Z.; Briel, A.; Lyon, S.W.; Olofsson, B.; Folkesson, L. On the utilization of hydrological modelling for road drainage design under climate and land use change. *Sci. Total. Environ.* **2014**, *475*, 97–103. [\[CrossRef\]](#)
3. Hannah, D.M.; Kansakar, S.R.; Gerrard, A.J.; Rees, G. Flow regimes of Himalayan rivers of Nepal: Nature and spatial patterns. *J. Hydrol.* **2005**, *308*, 18–32. [\[CrossRef\]](#)
4. Nijssen, B.; O'donnell, G.M.; Hamlet, A.F.; Lettenmaier, D.P. Hydrologic sensitivity of global rivers to climate change. *Clim. Chang.* **2001**, *50*, 143–175. [\[CrossRef\]](#)
5. Reedyk, S.; Woo, M.K.; Prowse, T.D. Contribution of icing ablation to streamflow in a discontinuous permafrost area. *Can. J. Earth Sci.* **1995**, *32*, 13–20. [\[CrossRef\]](#)
6. Refsgaard, J.C.; Knudsen, J. Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **1996**, *32*, 2189–2202. [\[CrossRef\]](#)
7. Sorooshian, S.; Gupta, V.K.; Fulton, J.L. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: Influence of calibration data variability and length on model credibility. *Water Resour. Res.* **1983**, *19*, 251–259. [\[CrossRef\]](#)
8. Yapo, P.O.; Gupta, H.V.; Sorooshian, S. Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data. *J. Hydrol.* **1996**, *181*, 23–48. [\[CrossRef\]](#)

9. Anctil, F.; Perrin, C.; Andréassian, V. Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models. *Environ. Model. Softw.* **2004**, *19*, 357–368. [[CrossRef](#)]
10. Harlin, J.; Kung, C.S. Parameter uncertainty and simulation of design floods in Sweden. *J. Hydrol.* **1992**, *137*, 209–230. [[CrossRef](#)]
11. Seibert, J. Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrol. Earth Syst. Sci. Discuss.* **2000**, *4*, 215–224. [[CrossRef](#)]
12. Madsen, H. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J. Hydrol.* **2000**, *235*, 276–288. [[CrossRef](#)]
13. Brath, A.; Montanari, A.; Toth, E. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. *J. Hydrol.* **2004**, *291*, 232–253. [[CrossRef](#)]
14. Vrugt, J.A.; Gupta, H.V.; Dekker, S.C.; Sorooshian, S.; Wagener, T.; Bouten, W. Application of stochastic parameter optimization to the Sacramento soil moisture accounting model. *J. Hydrol.* **2006**, *325*, 288–307. [[CrossRef](#)]
15. Stehr, A.; Debels, P.; Romero, F.; Alcayaga, H. Hydrological modelling with SWAT under conditions of limited data availability: Evaluation of results from a Chilean case study. *Hydrol. Sci. J.* **2008**, *53*, 588–601. [[CrossRef](#)]
16. Kalantari, Z.; Lyon, S.W.; Jansson, P.E.; Stolte, J.; French, H.K.; Folkesson, L.; Sassner, M. Modeller subjectivity and calibration impacts on hydrological model applications: An event-based comparison for a road-adjacent catchment in south-east Norway. *Sci. Total Environ.* **2015**, *502*, 315–329. [[CrossRef](#)]
17. Kirchner, J.W. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.* **2006**, *42*, W03S04. [[CrossRef](#)]
18. Seibert, J.; McDonnell, J.J. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resour. Res.* **2002**, *38*, 1241–1252.
19. Beven, K.; Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **2001**, *249*, 11–29. [[CrossRef](#)]
20. Borja, S.; Kalantari, Z.; Destouni, G. Global Wetting by Seasonal Surface Water over the Last Decades. *Earth's Future* **2020**, *8*, e2019EF001449. [[CrossRef](#)]
21. Blöschl, G.; Bierkens, M.F.; Chambel, A.; Cudennec, C.; Destouni, G.; Fiori, A.; Kirchner, J.W.; McDonnell, J.J.; Savenije, H.H.; Sivapalan, M.; et al. Twenty-three unsolved problems in hydrology (UPH)—a community perspective. *Hydrol. Sci. J.* **2019**, *64*, 1141–1158. [[CrossRef](#)]
22. Etter, S.; Addor, N.; Huss, M.; Finger, D. Climate change impacts on future snow, ice and rain runoff in a Swiss mountain catchment using multi-dataset calibration. *J. Hydrol. Reg. Stud.* **2017**, *13*, 222–239. [[CrossRef](#)]
23. Finger, D.; Vis, M.; Huss, M.; Seibert, J. The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments. *Water Resour. Res.* **2015**, *51*, 1939–1958. [[CrossRef](#)]
24. Schoups, G.; Addams, C.L.; Gorelick, S.M. Multi-objective calibration of a surface water-groundwater flow model in an irrigated agricultural region: Yaqui Valley, Sonora, Mexico. *Hydrol. Earth Syst. Sci. Discuss.* **2005**, *2*, 2061–2109. [[CrossRef](#)]
25. Kalantari, Z.; Lyon, S.W.; Folkesson, L.; French, H.K.; Stolte, J.; Jansson, P.E.; Sassner, M. Quantifying the hydrological impact of simulated changes in land use on peak discharge in a small catchment. *Sci. Total Environ.* **2014**, *466–467*, 741–754. [[CrossRef](#)]
26. Rakovec, O.; Kumar, R.; Attinger, S.; Samaniego, L. Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resour. Res.* **2016**, *52*, 7779–7792. [[CrossRef](#)]
27. Rakovec, O.; Kumar, R.; Mai, J.; Cuntz, M.; Thober, S.; Zink, M.; Attinger, S.; Schäfer, D.; Schrön, M.; Samaniego, L. Multiscale and multivariate evaluation of water fluxes and states over European river basins. *J. Hydrometeorol.* **2016**, *19*, 287–307. [[CrossRef](#)]
28. Samaniego, L.; Kumar, R.; Attinger, S. Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.* **2010**, *46*. [[CrossRef](#)]
29. Finger, D. The value of satellite retrieved snow cover images to assess water resources and the theoretical hydropower potential in ungauged mountain catchments. *Jökull* **2018**, *68*, 47–66.
30. El Maayar, M.; Chen, J.M. Spatial scaling of evapotranspiration as affected by heterogeneities in vegetation, topography, and soil texture. *Remote. Sens. Environ.* **2006**, *102*, 33–51. [[CrossRef](#)]

31. Finger, D.; Hugentobler, A.; Huss, M.; Voinesco, A.; Wernli, H.R.; Fischer, D.; Weber, E.; Jeannin, P.Y.; Kauzlaric, M.C.; Wirz, A.C.; et al. Identification of glacial meltwater runoff in a karstic environment and its implication for present and future water availability. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 3261–3277. [[CrossRef](#)]
32. Hundecha, Y.; Bárdossy, A. Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *J. Hydrol.* **2004**, *292*, 281–295. [[CrossRef](#)]
33. Klok, E.J.; Jasper, K.; Roelofsma, K.P.; Gurtz, J.; Badoux, A. Distributed hydrological modelling of a heavily glaciated Alpine river basin. *Hydrol. Sci. J.* **2001**, *46*, 553–570. [[CrossRef](#)]
34. Parajka, J.; Blöschl, G. The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models. *J. Hydrol.* **2008**, *358*, 240–258. [[CrossRef](#)]
35. Pokhrel, P.; Yilmaz, K.K.; Gupta, H.V. Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J. Hydrol.* **2012**, *418*, 49–60. [[CrossRef](#)]
36. Stahl, K.; Moore, R.D.; Shea, J.M.; Hutchinson, D.; Cannon, A.J. Coupled modelling of glacier and streamflow response to future climate scenarios. *Water Resour. Res.* **2008**, *44*. [[CrossRef](#)]
37. Huss, M.; Farinotti, D.; Bauder, A.; Funk, M. Modelling runoff from highly glacierized alpine drainage basins in a changing climate. *Hydrol. Process.* **2008**, *22*, 3888–3902. [[CrossRef](#)]
38. Franks, S.W.; Gineste, P.; Beven, K.J.; Merot, P. On constraining the predictions of a distributed model: The incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resour. Res.* **1998**, *34*, 787–797. [[CrossRef](#)]
39. Lindström, G.; Pers, C.; Rosberg, J.; Strömqvist, J.; Arheimer, B. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrol. Res.* **2010**, *41*, 295–319. [[CrossRef](#)]
40. European Union. *Copernicus Land Monitoring Service (2012)*; European Environment Agency (EEA): Copenhagen, Denmark, 2018.
41. Finger, D.; Heinrich, G.; Gobiet, A.; Bauder, B. Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century. *Water Resour. Res.* **2012**, *48*, W02521. [[CrossRef](#)]
42. Helmert, J.; Şensoy Şorman, A.; Alvarado Montero, R.; De Michele, C.; De Rosnay, P.; Dumont, M.; Pullen, S. Review of snow data assimilation methods for hydrological, land surface, meteorological and climate models: Results from a cost harmonsnow survey. *Geosciences* **2018**, *8*, 489. [[CrossRef](#)]
43. Smhi.net (info.txt: HYPE Model Documentation). Available online: [http://www.smhi.net/hype/wiki/doku.php?id=start:hype\\_file\\_reference:info.txt](http://www.smhi.net/hype/wiki/doku.php?id=start:hype_file_reference:info.txt) (accessed on 13 February 2018).
44. Tockner, K.; Uehlinger, U.; Robinson, C. (Eds.) Arctic rivers. In *Rivers of Europe*, 1st ed.; Elsevier Academic Press: Cambridge, MA, USA, 2009; pp. 366–367. Available online: [https://notendur.hi.is/gmg/Serprent/Brittain\\_et\\_al\\_2009\\_Arctic\\_Rivers\\_Rivers\\_of\\_Europe%20ch%209.pdf](https://notendur.hi.is/gmg/Serprent/Brittain_et_al_2009_Arctic_Rivers_Rivers_of_Europe%20ch%209.pdf) (accessed on 27 February 2019).
45. Bengtsson, L.; Andrae, U.; Aspelien, T.; Batrak, Y.; Calvo, J.; de Rooy, W. The HARMONIE-AROME Model Configuration in the ALADIN-HIRLAM NWP System. *Mon. Weather Rev.* **2017**, *145*, 1919–1935. [[CrossRef](#)]
46. Icelandic Meteorological Office. HARMONIE-Numerical Weather Prediction Model. Available online: <https://en.vedur.is/weather/articles/nr/3232> (accessed on 2 May 2018).
47. National Land Survey of Iceland: Download page. Available online: <http://atlas.lmi.is/LmiData/index.php?id=963540976962> (accessed on 16 March 2018).
48. Hall, D.K.; Riggs, G.A.; Salomonson, V.V.; DiGirolamo, N.E.; Bayr, K.J. MODIS snow-cover products. *Remote Sens. Environ.* **2002**, *83*, 181–194. [[CrossRef](#)]
49. Arnalds, O.; Grétarsson, E. *Soil map of Iceland*; Agricultural Research Institute (RALA): Reykjavík, Iceland, 2010.
50. Samuelsson, P.; Gollvik, S.; Ullerstig, A. *The Land-Surface Scheme of the Rossby Centre Regional Atmospheric Climate Model (RCA3)*; SMHI: Norrköping, Sweden, 2006.
51. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [[CrossRef](#)]
52. Konz, M.; Seibert, J. On the value of glacier mass balances for hydrological model calibration. *J. Hydrol.* **2010**, *385*, 238–246. [[CrossRef](#)]
53. Freer, J.E.; McMillan, H.; McDonnell, J.J.; Beven, K.J. Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures. *J. Hydrol.* **2004**, *291*, 254–277. [[CrossRef](#)]

54. Gurtz, J.; Zappa, M.; Jasper, K.; Lang, H.; Verbunt, M.; Badoux, A.; Vitvar, T.A. comparative study in modelling runoff and its components in two mountainous catchments. *Hydrol. Process.* **2003**, *17*, 297–311. [[CrossRef](#)]
55. Kalantari, Z.; Ferreira, C.S.S.; Page, J.; Goldenberg, R.; Olsson, J.; Destouni, G. Meeting sustainable development challenges in growing cities: Coupled social-ecological systems modeling of land use and water changes. *J. Environ. Manag.* **2019**, *245*, 471–480. [[CrossRef](#)]
56. Kalantari, Z.; Ferreira, C.S.S.; Koutsouris, A.J.; Ahmer, A.-K.; Cerdà, A.; Destouni, G. Assessing flood probability for transportation infrastructure based on catchment characteristics, sediment connectivity and remotely sensed soil moisture. *Sci. Total Environ.* **2019**, *661*, 393–406. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).