














Article

Additional Value of Using Satellite-Based Soil Moisture and Two Sources of Groundwater Data for Hydrological Model Calibration

Mehmet Cüneyd Demirel ^{1,*}, Alparslan Özen ¹, Selen Orta ^{1,2}, Emir Toker ³, Hatice Kübra Demir ¹, Ömer Ekmekcioğlu ¹, Hüsamettin Tayşi ¹, Sinan Eruçar ¹, Ahmet Bilal Sağ ¹, Ömer Sarı ¹, Ecem Tuncer ¹, Hayrettin Hancı ¹, Türkan İrem Özcan ¹, Hilal Erdem ¹, Mehmet Melih Koşucu ¹, Eyyup Ensar Başakın ¹, Kamal Ahmed ^{4,5}, Awat Anwar ⁶, Muhammet Bahattin Avcuoğlu ^{1,7}, Ömer Vanlı ⁸, Simon Stisen ^{9,*}, and Martijn J. Booij ^{10,*}

- ¹ Department of Civil Engineering, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey; ozenalp@itu.edu.tr (A.Ö.); selenorta@ayvansaray.edu.tr (S.O.); demirh17@itu.edu.tr (H.K.D.); ekmekcioglu17@itu.edu.tr (Ö.E.); tayysi18@itu.edu.tr (H.T.); erucar@itu.edu.tr (S.E.); saga@itu.edu.tr (A.B.S.); sari18@itu.edu.tr (Ö.S.); tuncere18@itu.edu.tr (E.T.); hanci18@itu.edu.tr (H.H.); ozcant18@itu.edu.tr (T.İ.Ö.); erdemhi@itu.edu.tr (H.E.); kosucu@itu.edu.tr (M.M.K.); basakin16@itu.edu.tr (E.E.B.); mbavcuoglu@dsi.gov.tr (M.B.A.)
 - ² Architecture and Urban Planning Department, Istanbul Ayvansaray University, 34087 Balat, Istanbul, Turkey
 - ³ Eurasia Institute of Earth Sciences Istanbul Technical University, 34469 Maslak, Istanbul, Turkey; tokerem@itu.edu.tr
 - ⁴ Faculty of Civil Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia; kamal_brc@hotmail.com
 - ⁵ Faculty of Water Resource Management, Lasbela University of Agriculture, Water and Marine Sciences, Balochistan 90150, Pakistan
 - ⁶ Department of Civil Engineering, Yıldız Technical University, 34210 Esenler, Istanbul, Turkey; awat.oanwar@gmail.com
 - ⁷ General Directorate of State Hydraulic Works (DSI), 06100 Ankara, Turkey
 - ⁸ Department of Geographical Information Technologies, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey; vanli@itu.edu.tr
 - ⁹ Geological Survey of Denmark and Greenland, Øster Voldgade 10, 1350 Copenhagen, Denmark
 - ¹⁰ Water Engineering and Management, Faculty of Engineering Technology, University of Twente, Enschede 7500 AE, The Netherlands
- * Correspondence: demirelmc@itu.edu.tr (M.C.D.); sst@geus.dk (S.S.); m.j.booij@utwente.nl (M.J.B.); Tel.: +90-212-285-7426 (M.C.D.)

Received: 4 September 2019; Accepted: 1 October 2019; Published: 6 October 2019



Abstract: Although the complexity of physically-based models continues to increase, they still need to be calibrated. In recent years, there has been an increasing interest in using new satellite technologies and products with high resolution in model evaluations and decision-making. The aim of this study is to investigate the value of different remote sensing products and groundwater level measurements in the temporal calibration of a well-known hydrologic model i.e., Hydrologiska Bryåns Vattenbalansavdelning (HBV). This has rarely been done for conceptual models, as satellite data are often used in the spatial calibration of the distributed models. Three different soil moisture products from the European Space Agency Climate Change Initiative Soil Measure (ESA CCI SM v04.4), The Advanced Microwave Scanning Radiometer on the Earth Observing System (EOS) Aqua satellite (AMSR-E), soil moisture active passive (SMAP), and total water storage anomalies from Gravity Recovery and Climate Experiment (GRACE) are collected and spatially averaged over the Moselle River Basin in Germany and France. Different combinations of objective functions and search algorithms, all targeting a good fit between observed and simulated streamflow, groundwater and

soil moisture, are used to analyze the contribution of each individual source of information. Firstly, the most important parameters are selected using sensitivity analysis, and then these parameters are included in a subsequent model calibration. The results of our multi-objective calibration reveal a substantial contribution of remote sensing products to the lumped model calibration, even if their spatially-distributed information is lost during the spatial aggregation. Inclusion of new observations, such as groundwater levels from wells and remotely sensed soil moisture to the calibration improves the model's physical behavior, while it keeps a reasonable water balance that is the key objective of every hydrologic model.

Keywords: HBV; GRACE; SMAP; ESA CCI SM v04.4; AMSR-E; Moselle River

1. Introduction

Hydrologic models are indispensable tools for predicting the rainfall-runoff response of catchments in order to anticipate hydrological droughts and (flash) flood events, causing loss of lives and economic damage. Hydrological processes transforming rainfall to runoff are represented by mathematical equations in different types of model structures. Based on their spatial discretization levels, the models are grouped as lumped, semi-distributed and fully distributed models. While the most sophisticated models fall in the latter group, lumped models are the most commonly used models, as they require less inputs and computational sources. Further, if lumped models are calibrated with appropriate observations and objective functions, they are skillful in obtaining reasonable discharge prediction performance. In this study, we are interested in examining the potential usage of remote sensing (RS) data to inform the parameter optimization, and evaluate different outputs of a bucket-type lumped hydrologic model.

Earth observations using satellite-based RS gained popularity with its easy-to-apply usage and large spatial coverage. RS products are recognized as an important source of data, and used directly as model input e.g., leaf area index, or used to assess model performance on simulated states and fluxes [1]. Both the availability and spatial resolution of remotely-sensed data continuously increase the provision of opportunities to constrain hydrological models [2,3]. Actual evapotranspiration (AET) [4,5] and soil moisture data [6] from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellites, and groundwater storage-related observations via the Gravity Recovery and Climate Experiment (GRACE; [7]) satellites are just some of the examples from the numerous satellites. To benefit from spatial observations, spatially-distributed models are utilized [8–11].

Demirel et al. [3] used MODIS-based AET to calibrate a spatially-distributed model, and revealed that there is little trade-off between either the streamflow performance or the spatial performance of the model. Appropriate objective functions were key to constrain the model and reduce the number of runs to search for an optimal solution, meeting the water balance and spatial AET patterns. Rakovec et al. [9] used GRACE-based total water storage data and evapotranspiration (ET) from FLUXNET data to identify model parameters and improve the streamflow simulation performance in more than 80 European catchments. Zink et al. [10] used remotely-sensed temperature data to constrain a spatially-distributed model, and could succeed in reducing the errors in predicting actual evapotranspiration by 8%, while the model streamflow performance slightly decreased by 6%, showing the trade-off between different objectives when applied in a multi-objective calibration framework. Xiong et al. [12] assessed the impact of the remotely-sensed soil moisture active passive level 3 product (SMAP L3) in a multi-objective calibration of a distributed hydrological model in the Qujiang and Ganjiang catchments in China. They documented minor improvements in the simulation performance of soil moisture and streamflow when adding the SMAP L3 product, which could be due to the short length of this high-resolution soil moisture product. Li et al. [6] used lumped and

semi-distributed hydrological models with different RS products and different calibration approaches in two Australian catchments.

Near-surface soil moisture retrievals from MODIS (500×500 m) and the Soil Moisture and Ocean Salinity level 3 product (SMOS L3, 25×25 km), together with remotely-sensed fractional vegetation cover data from MODIS (500×500 m), were incorporated with hydrologic models to test their contribution to the model simulation. They showed that the effect of RS soil moisture data for model calibration is usually substantial for the upstream sub-catchments.

There have been studies solely focused upon the effect of performance metrics (objective functions) on the model calibration. For instance, Kamamia et al. [13] calibrated a semi-distributed hydrologic model (Soil and Water Assessment Tool—SWAT) using a multi-metric framework including a Pareto front and five segment flow duration curve, and compared it with a standard calibration using Nash Sutcliffe Efficiency (NSE; [14]). Tobin et al. [15] calibrated the same model for a mountainous watershed in Idaho, US using ET estimates from MODIS, Simplified Surface Energy Balance and Global Land Evaporation: the Amsterdam Model (GLEAM). Their study showed that the MODIS ET product especially yielded the most to identify the model parameters, and improved streamflow simulations in the recession phase and summertime. They also used NSE as streamflow objective function that is criticized to be biased for high flows [16]. Krause et al. [17] studied nine different objective functions, including the index of agreement [18], the coefficient of determination (R^2), NSE and their revised versions, so as to adopt these metrics to high flows and low flows. They highlighted the importance of using multiple metrics depending on the modeling aim. All studies above and numerous other studies [19–22] have utilized mostly spatially-distributed hydrologic models with RS products and different calibration metrics [23,24].

However, to our knowledge, there have been few studies focusing on lumped models and RS products to improve model calibration, and none of them used groundwater level measurements from drilled wells in model evaluation and calibration. In one of the lumped model studies, Nijzink et al. [2] used soil moisture, evaporation, total water storage and snow state from relevant satellites to calibrate 27 catchments across Europe using the GLUE methodology [25]. They tested the applicability of satellite-based RS data in an evaluation of five different lumped models through a model calibration framework. Their study showed that using multiple information from satellites could be fruitful for estimating model parameters in data scarce regions without streamflow time series.

The objective of this study is to assess the value of remote sensing products, as well as groundwater level measurements in a multi-objective calibration of the Hydrologiska Bryåns Vattenbalansavdelning (HBV) lumped hydrologic model [26,27]. For that, we collected relevant satellite-based hydrologic data for the Moselle River Basin, and spatially averaged them first. Then different calibration scenarios are tested to quantify the individual effect of each RS data source on model calibration results. Two essential features of our study are: (1) The important parameters for the calibration are selected based on a multi-objective sensitivity analysis and, (2) three search algorithms i.e., one gradient descend and two global metaheuristic methods, are used to assess their effects on the calibration results. Further, the important parameters for the calibration are selected based on a multi-objective sensitivity analysis.

2. Study Area

The study area, the Moselle River, is a tributary of the Rhine River. The Rhine River Basin covers an area of $165,000 \text{ km}^2$ in eight countries, that is, the Netherlands, Germany, Switzerland, Luxemburg, Belgium, France, Austria and Liechtenstein [28,29]. The Moselle River Basin covers the mountains of the middle reaches in Germany, including the Black Forest (Schwarzwald) and the Vosges, reaching elevations over 1000 m (3280 feet) in the south, to around 600 m (1968 feet) towards the north. Land use in these areas include vineyards on sunny valley slopes, and agriculture and forest in the mountains and hillslopes. In the Moselle River Basin, winter is the season where the maximum discharge occurs. [28]. The mean discharge at Cochem is about $314 \text{ m}^3 \text{ s}^{-1}$ (11,088.8 cu. ft. per sec.), and its discharge varies

from $14 \text{ m}^3 \text{ s}^{-1}$ (494 cu. ft. per sec.), in dry summers to a maximum of $4000 \text{ m}^3 \text{ s}^{-1}$ (141,259 cu. ft. per sec.), during winter floods [30].

The Moselle River is 313 km (195 mi.) long and has a catchment area of $27,262 \text{ km}^2$, with an annual precipitation of $\sim 800 \text{ mm}$. The altitude ranges from 59 m to 1326 m (Figure 1), with a mean altitude of around 340 m. The Moselle Basin has 26 sub-basins, and the surface areas of these sub-basins vary between 102 and 3353 km^2 [31,32].

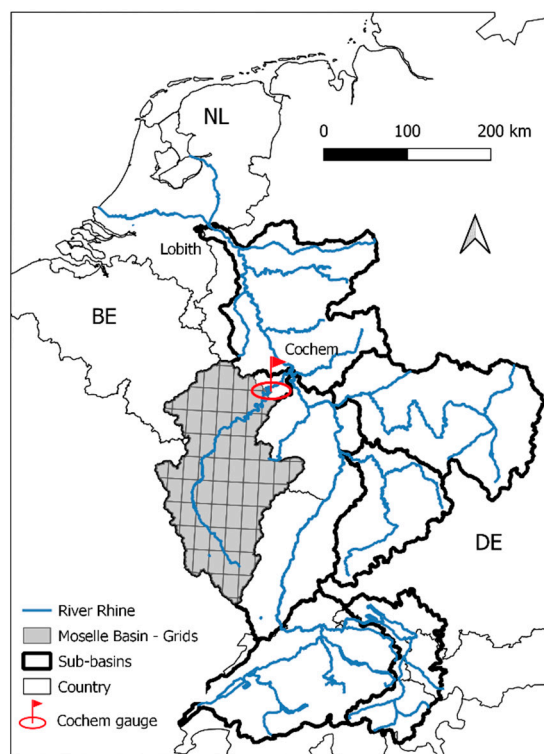


Figure 1. Moselle River Basin and Cochem Gauge.

3. Data

The basic inputs of a lumped hydrological model are daily precipitation (P) and potential evapotranspiration (PET). The P and PET time series for the Moselle Basin were obtained from the German Federal Institute of Hydrology (BfG) in Koblenz (Germany). Further, the outlet discharge (Q) measurements at Cochem (gauge #6336050) were provided by the Global Runoff Data Centre (GRDC) in Koblenz (Germany) [33]. Although Rhine has mostly alluvial aquifer, Moselle uplands comprised of Jurassic Limestone aquifers [2]. Also within the Moselle there is contrasting lithology and topography: granitic formations upstream in the Vosges mountains and carbonate platform downstream on the Lorraine plateau [2]. The groundwater well measurements (GWY) from 80 stations within the Moselle Basin were also obtained from the German Federal Institute of Hydrology (BfG) in Koblenz (Germany). GWY data from the stations falling on the same sub-catchment of Moselle are averaged using arithmetic mean. These sub-catchment averaged data are then aggregated to the Moselle Basin Level using areal weights of 26 sub-catchments. In addition to these ground-based data we also included remote sensing soil moisture data from different sources (Table 1). All remotely-sensed soil moisture (RS-SM) data are upscaled to catchment scale using an arithmetic average of all cells (Figure 2).

3.1. European Space Agency (ESA CCI SM V04.4)

Soil moisture (SM) is an important variable for many purposes, e.g., to anticipate possible floods using antecedent SM conditions, and to manage irrigation water. In 2010, the Global Climate Observing System (GCOS) panel recognized soil moisture as one of the 50 Essential Climate Variables (ECVs) [34].

The European Space Agency (ESA) started the CCI SM project, known as the Climate Change Initiative, in 2012, to contribute to the monitoring of the ECVs [35]. The purpose of the ESA CCI SM project is to merge active and passive microwave sensors and generate a new global SM data record. The latest-released ESA CCI SM v04.4 product is used in this study (<http://www.esa-soilmoisture-cci.org/>, last access: 4 September 2019; [36]). The last algorithm of this ESA CCI SM v4 project, which is also being used in the current Copernicus Climate Change Service (C3S) climate data records (CDRs) production system version v201812, uses an approach to better parameterize the least-squares merging [37]. The ESA CCI soil moisture v04.4 product with daily temporal and 0.25° spatial resolution, spans for the period from November 1978 and June 2018, and consists of three different datasets, the so-called active, passive and combined [38].

The active data set is generated by merging active microwave satellite data by using the TU Wien Water Retrieval Package (WARP) algorithm [39,40], estimates from the European Remote Sensing-Active Microwave Instrument Wind Scatterometer (ERS, AMI-WS, 5.3 GHz, July 1991–February 2007) [41,42] and the Meteorological Operational satellite program—Advanced Scatterometer (MetOp, ASCAT, 5.3 GHz, January 2007–June 2018) [43]. The passive data set is generated by merging active microwave satellite data by using the Land Parameter Retrieval Model (LPRM) algorithm [44], estimates from Nimbus 7—Scanning Multichannel Microwave Radiometer (Nimbus 7, SMMR, 6.6 GHz, January 1979–August 1987) [44], Defence Meteorological Satellite Program—The Special Sensor Microwave/Imager (DMSP, SSM/I, 19.3 GHz, September 1987–December 2007) [44], the Tropical Rainfall Measuring Mission’s Microwave Imager (TRMM, TMI, 10.7 GHz, January 1998–December 2013) [44], Aqua—The Advanced Microwave Scanning Radiometer for Earth Observing System (Aqua, AMSR-E, 6.9/10.7 GHz, July 2002–October 2011) [44], Coriolis—WindSat radiometer (6.8/10.7 GHz, October 2007–July 2012) [45], ESA—Soil Moisture and Ocean Salinity (ESA, SMOS, 1.4 GHz, January 2010–June 2018) [46,47]; and the Global Change Observation Mission 1st-Water—AMSR-2 (GCOM-W1, AMSR2, 6.9/10.6 GHz, July 2012–June 2018) [48]. The combined data set is obtained by merging the active and passive products. The method for merging based on the weighted averaging and merging scheme is parameterized using triple collocation analysis (TCA) [38]. In this paper, the ESA CCI SM v04.4 product combined data set is used. Studies of [36,38] are recommended to reader for more details about the evolution of the ESA CCI SM data record, details of sensor properties, and product merging methodology.

3.2. *The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E)*

Microwave remote sensing provides a solution to directly observe the soil moisture. Especially active sensors are generally not affected by cloud cover, although accurate soil moisture estimates are limited to regions that have either bare soil or low to moderate amounts of vegetation cover, since in the absence of significant vegetation cover, soil moisture is the dominant effect on the received signal [49]. Besides, one of the issues about using passive microwaves is that there are contrasting effects of soil moisture and vegetation water content on microwave emission. In other words, an increase of soil moisture and a decrease in vegetation water content have the same effect. Effects of vegetation and roughness become more evident as the frequency increases; that is why low frequencies in the L-band range (1–2 GHz) are better suited for soil moisture sensing [50]. Another issue is about the effect of the daytime temperature measurements affecting the top soil layers, making it difficult for soil moisture inversion. The Advanced Microwave Scanning Radiometer (AMSR-E) on the Earth Observing System (EOS) Aqua satellite was launched on 4 May 2002. The AMSR-E instrument was developed by the National Space Development Agency of Japan (NASDA), and provided to the U.S. National Aeronautics and Space Administration (NASA). AMSR-E records the brightness temperature at frequencies of 6.9, 10.7, 18.7, 23.8, 36.5 and 89 GHz, at horizontal (H) and vertical (V) polarizations. The mean spatial resolution at 6.9 GHz is about 56 km with a swath width of 1445 km. The AMSR-E soil moisture has a strong association to ground-based soil moisture data, with typical correlations of

greater than 0.8, and typical RMSD less than 0.03 vol/vol (for a normalized and filtered AMSR-E time series) [50,51].

For AMSR, various soil moisture algorithms that differ mainly in their approaches to the vegetation and surface temperature corrections have been considered. There are four groups of correction approaches, with the first one being sequential corrections using external ancillary data; another approach is the iterative parameter fitting to a forward multichannel brightness temperature model, an yet another approach consists of the use of brightness temperature indices and regression fits, and lastly combinations of the above approaches [50].

3.3. *The Soil Moisture Active Passive (SMAP)*

The Soil Moisture Active Passive (SMAP) mission has been developed by NASA as one of the first Earth observation satellites in return for the National Research Council's Decadal Survey. SMAP analyzes global measurements of soil moisture present at the Earth's surface, thus allowing us to make indirect observations of soil moisture and the thaw/freeze state from space to formulate considerably improved estimates of energy, water and carbon transfers between the atmosphere and the land. Correct characterization of these transfers is highly dependent on the accuracy of numerical atmosphere models that are used in weather prediction and climate projections. Measurement of soil moisture can be applied directly to flood assessment and drought monitoring. It is useful in estimating global water and energy fluxes at the land surface. Since April 2015, the Soil Moisture Active Passive (SMAP) mission of NASA has been successfully monitoring near-surface soil moisture, by mapping the globe between the latitude bands of 85.044° N/S in 2–3 days, depending upon the location [52–55]. In this research, daily data of soil moisture with a resolution of 36 × 36 km is downloaded and processed. However, soil moisture retrievals only from only ESA and AQUA are incorporated into the model calibration.

3.4. *Gravity Recovery and Climate Experiment (GRACE)*

The main purpose of GRACE (Gravity Recovery and Climate Experiment) is in the surveying of the Earth's gravity field anomalies of lands and oceans with a spatial resolution of 400 × 400 km [7]. Water storage in hydrologic reservoirs, the movements of oceans, atmospheric and land ice masses and mass swaps of Earth compartments cause gravity changes, and for GRACE monthly changes are considered. The gravity changes are measured in centimeters of equivalent water thickness every month. The GRACE mission consists of twin satellites, and the creators of the mission are the United States National Aeronautics and Space Administration (NASA) and the German Aerospace Centre (DLR). GRACE was launched on 17 March 2002, and its mission ended on 27 October 2017. As a continuation of the GRACE mission, on 22 May 2018, GRACE-FO (GRACE Follow On) was launched. The Science Data System (SDS), which consists of Centre for Space Research (CSR) (<http://www2.csr.utexas.edu/grace/>), the Geologic Research Centre in Potsdam (GFZ) (<https://www.gfz-potsdam.de/en/grace/>), and the Jet Propulsion Laboratory (JPL) (<https://gracefo.jpl.nasa.gov>), provides GRACE gravity field models. In this study, to track groundwater changes in the Moselle River Basin, monthly GRACE data for the period between April 2002 and June 2017 were used (Table 1 and Figure 2).

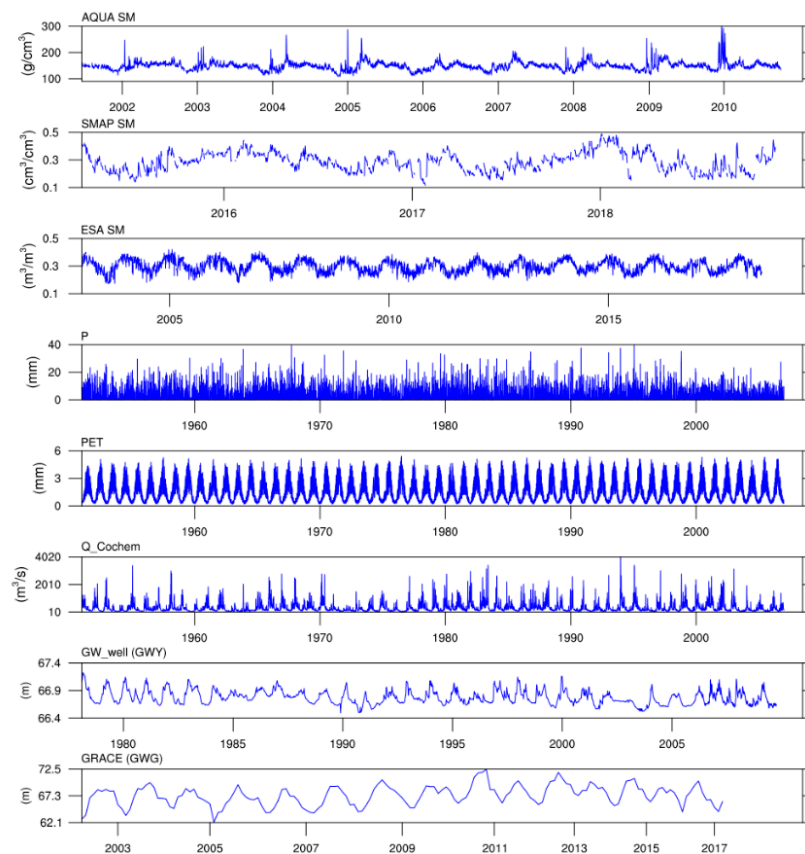


Figure 2. Catchment scale time series of remotely sensed soil moisture (RS-SM) products (AQUA, soil moisture active passive (SMAP) and ESA), hydrologic model inputs (P and PET), and discharge data (Q) from Cochem Gauge, GW well measurements (GWY) and Gravity Recovery and Climate Experiment (GRACE) groundwater storage data (GWG), respectively.

Table 1. Overview of remote sensing and hydro-meteorological data.

Abbreviation	Description	Number of Sub-Basins	Spatiotemporal Resolution	Spatial Averaging	Period	Source
Q	Discharge	1 gauge at Cochem	Catchment-scale (daily)	-	1.1.1951–31.12.2015	GRDC in Koblenz
P	Precipitation	26	10 km (daily)	Areal weighting (sub-basins)	1.1.1951–31.12.2015	BfG in Koblenz, van Osnabrugge et al. [29]
PET	Potential evapotranspiration	26	20 km (daily)	Areal weighting (sub-basins)	1.1.1951–31.12.2015	BfG in Koblenz, van Osnabrugge et al. [29]
GWG	Remotely sensed groundwater storage (GRACE)	1 (Rhine basin)	400 km (monthly)	-	2002–2017	GRACE
GWY	Groundwater well measurements	80	Point (daily)	Areal weighting (sub-basins)	8.2.1978–7.10.2009	BfG in Koblenz
ESA	Soil moisture (ESACCI_SM_V04.4 combined product)	-	0.25 degree (daily)	Grids—arithmetic averaging	01.11.1978–30.06.2018	ftp.geo.tuwien.ac.at
AQUA	Soil moisture (AMSR-E/Aqua)	-	25 km (daily)	Grids—arithmetic averaging	19.06.2002–03.10.2011	earthdata.nasa.gov
SMAP	Soil moisture	-	36 km (daily)	Grids—arithmetic averaging	31.03.2015–Present	earthdata.nasa.gov

4. Methods

Lumped hydrologic models are usually calibrated based solely on discharge data measured at the catchment outlet. In this study, we used well measurements (water levels), as well as temporal information from several remote sensing products, in order to evaluate outputs from different compartments (fast runoff and base flow) of the Hydrologiska Bryåns Vattenbalansavdelning (HBV) model. This is done to simultaneously improve the performance and physical meaning of the results. First, we applied a systematic sensitivity analysis using appropriate objective functions to identify the most important parameters. We then included the selected parameters in the calibration.

4.1. Hydrologic Model

In this study we used the Hydrologiska Bryåns Vattenbalansavdelning (HBV) lumped model that includes the basic rainfall-runoff processes, such as snow accumulation and melt, soil moisture, quick runoff and base flow (Figure 3). The model was developed by the Swedish Meteorological and Hydrological Institute [26,27], and applied in numerous projects [30,32,56–59]. The Moselle River is a rainfed river, therefore, we only used P and PET as model inputs. The parameter ranges for model calibration (Table 2) were taken from previous studies [32,58].

Table 2. Description of the Hydrologiska Bryåns Vattenbalansavdelning (HBV) model parameters.

Parameter	Unit	Range	Description
FC	mm	100–800	Maximum soil moisture capacity
LP		0.1–1	Soil moisture threshold for the reduction of evapotranspiration
BETA		1–6	Shape coefficient
CFLUX	mm day ⁻¹	0.1–1	Maximum capillary flow from upper response box to soil moisture zone
ALFA		0.1–2	Measure for nonlinearity of quick runoff
KF	day ⁻¹	0.005–0.5	Recession coefficient for quick runoff
KS	day ⁻¹	0.0005–0.2	Recession coefficient for base flow
PERC	mm day ⁻¹	0.01–6	Maximum flow from upper to lower response box

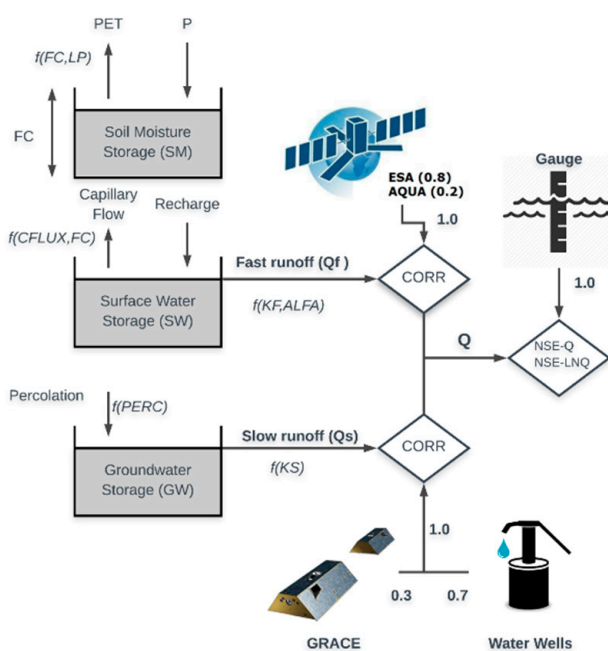


Figure 3. Conceptual diagram of the HBV model structure and evaluation procedure.

Figure 3 shows the three components (buckets) of the HBV and the flowchart of the model evaluation based on ESA, AQUA, Gravity Recovery and Climate Experiment (GRACE), water wells and gauge streamflow data. Three objective functions (CORR, NSE-Q and NSE-LNQ) and the weights (0.2, 0.8, 0.3, 0.7 and 1.0) of each data source used in the multi objective calibration framework are also depicted in this figure.

4.2. Objective Functions

There are various objective functions (i.e., performance metrics) to evaluate the hydrologic model. Some of them may focus on low flows [16], while others focus on high flows [14]. These objective functions (OFs) can use the crude difference between observed and simulated values, like the root mean square error, or the relative difference (i.e., normalized), such as the Nash Sutcliffe Efficiency (NSE) [14,16]. As highlighted in different studies [60–63], choosing an appropriate OF is very crucial for calibrating hydrologic models. In this paper, we carefully selected three OFs, namely NSE-Q, NSE-LNQ and the Pearson Correlation Coefficient (CORR) to assess the efficiency of the HBV model. The details of these OFs are summarized in Table 3.

The Pearson Correlation Coefficient (shown as CORR in Table 3) was used to determine the relationship between simulated variables from the HBV model (soil moisture and groundwater) and observed sensing data. Observed data contains groundwater well measurements (GWY) and remote sensing products (GRACE, ESA, AQUA and soil moisture active passive (SMAP)). Since soil moisture is highly relevant for fast runoff and high flows, NSE-Q was used as well. This metric uses squared differences which can be dominated by high values in the time series. To avoid this and to reduce the sensitivity of NSE to extreme high values, we used the logarithmic discharge values in NSE, i.e., NSE-LNQ, as well. Herein, peak values in the hydrograph are flattened, leading to a fair treatment of low flows in the model performance [17].

Table 3. Description of objective functions. In these formulas x , y , \bar{x} and \bar{y} indicate observation data, simulation data and their means, respectively.

Objective Functions	Formula	Range of Values
CORR	$\frac{\sum_{i=1}^n (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^n (x - \bar{x})^2} \sqrt{\sum_{i=1}^n (y - \bar{y})^2}}$	$[-1, 1]$
NSE-Q	$1 - \frac{\sum_{i=1}^n (x - y)^2}{\sum_{i=1}^n (x - \bar{x})^2}$	$[-\infty, 1]$
NSE-LNQ	$1 - \frac{\sum_{i=1}^n (\ln x - \ln y)^2}{\sum_{i=1}^n (\ln x - \ln \bar{x})^2}$	$[-\infty, 1]$

4.3. Latin Hypercube Sampling One-Factor-At-A-Time Sensitivity Analysis

Following the study of Demirel et al. [64], we used the Latin Hypercube Sampling One-at-a-Time (LHS-OAT) approach for sensitivity analysis with 20 different initial parameter sets. This number is assumed to be sufficient, as Demirel et al. [64] showed the stability of the average accumulated relative parameter sensitivities after nearly 20 initial parameter sets. Obviously the main advantage of using LHS is that it is a stratified sampling method improving the coverage of the parameter space, using a smaller number of random sets, as compared to other methods like crude Monte Carlo Sampling [65].

4.4. Model Calibration and Validation

In this study we tested three different optimization algorithms: (1) Gradient-based Levenberg Marquardt [66], (2) Shuffled Complex Evolutionary algorithm—University of Arizona (SCE-UA) [67] and (3) the Covariance Matrix Adaptation Evolution Strategy (CMAES) [68]. While the first method uses a local search algorithm, the other two use metaheuristic global search algorithms. SCE-UA is one of the most popular algorithms in hydrology, and CMAES emerges as a very efficient global algorithm that can converge quickly. The details of the model calibration and validation periods are presented in Table 4. The discharge data from the period of 2002–2006, GW well measurements and the satellite data (ESA and AQUA) from 2002–2004 are used to calibrate the model. The remaining part of the data is used to validate the HBV model (see Table 4).

It should be noted that the SMAP data have been available since 2015, and therefore they not included in the model calibration. A three year warm-up period is incorporated, and the evaluation of the discharge results started in 1954, instead of 1951. The objective functions are used individually or in different combinations based on our four cases: (1) Only discharge (only-Q), (2) only groundwater and soil moisture (GW-only and SM-only), (3) discharge and groundwater/soil moisture (Q + GW and Q + SM) and (4) all together (Q + GW + SM), i.e., discharge and remote sensing products (GRACE, ESA and AQUA), as well as groundwater measurements (GWY) are used together in the calibration with all objective functions. As the well observations (GWY) are a more reliable source of information about groundwater, as compared to the GRACE-based storage estimations, we gave a higher weight of 0.7 to GWY than GWG (weight of 0.3). Similarly, we gave higher weight to ESA (0.8) as compared to AQUA (0.2) since ESA is a blended product that contains AQUA (AMSR-E) as well. For the combined objective functions, we gave equal weights to Q, SM and GW.

In PEST toolbox, we set the max number of runs to 3000 in all three methods, and used two complexes for SCE-UA and default population values (i.e., lambda = 11, number of parents = 5, seed number = 1111) for CMAES. Also as a stopping rule, the maximum relative objective function change was set to 1% over five iterations in all three methods. We evaluated the differences between model-simulated and observed variables at their available temporal scales i.e., daily, except for the monthly GRACE data.

Table 4. The model calibration and validation periods.

Objective Function	Calibration		Validation	
	Start	End	Start	End
CORR_GWG	31.03.2002	31.12.2004	31.01.2005	31.12.2015
CORR_GWY	01.01.2002	31.12.2004	01.01.1979	31.12.2001
CORR_SM	19.06.2002	31.12.2004	01.01.2005	31.12.2015
NSE-Q	01.01.2002	31.12.2006	01.01.1954	31.12.2001
NSE-LNQ	01.01.2002	31.12.2006	01.01.1954	31.12.2001

5. Results

5.1. Sensitivity Analysis of the Model Parameters

Table 5 shows the LHS-OAT results based on the correlation coefficient and NSE, respectively. It should be noted that the sensitivity analysis is done separately for each objective function and RS product. The results are illustrated as descending in the third column i.e., SM for comparison and ease of following. The most influential parameters for SM are FC, LP, BETA and CFLUX, whereas the remaining four parameters (i.e., ALFA, KF, KS and PERC) have no effect upon the results. Interestingly for the two groundwater data sources (well and GRACE) the most influential parameters are different; i.e., KS for GWY and LP for GWG. As expected, the ALFA and KF parameters have no effect on the groundwater. Based on the results in the Table 5 columns Q and LNQ, ALFA is the most influential parameter for high flows (NSE), whereas BETA is the most important parameter for low flows (NSE-LNQ).

Table 5. Normalized (from 0 to 1) sensitivity indices of eight HBV parameters.

Parameter	Normalized Sensitivity				
	Correlation Coefficient			NSE	
	GWG (GRACE)	GWY (WELL)	SM (ESA and AQUA)	Q	LNQ
FC	0.733	0.970	1	0.075	0.695
LP	1	0.687	0.446	0.011	0.353
BETA	0.706	0.854	0.179	0.042	1
CFLUX	0.255	0.214	0.177	0.016	0.188
ALFA	0	0	0	1	0.112
KF	0	0	0	0.188	0.061
KS	0.926	1	0	0	0.450
PERC	0.352	0.376	0	0.086	0.106

Figure 4 shows the evolution of parameter rankings based on the average accumulated sensitivities. Interestingly, the rankings are unstable in the very first initial runs and stable after ten or twenty runs from different initial parameter sets.

This is consistent with the results of a recent study by Demirel et al. [64], stating that it is key to repeat local sensitivity analysis at different locations in the parameter space for fairly reaching global sensitivity results. In short, the most influential parameters from Table 5 can also be identified in Figure 4, and all parameters that have sensitivity above zero in this table are included in the model calibration.

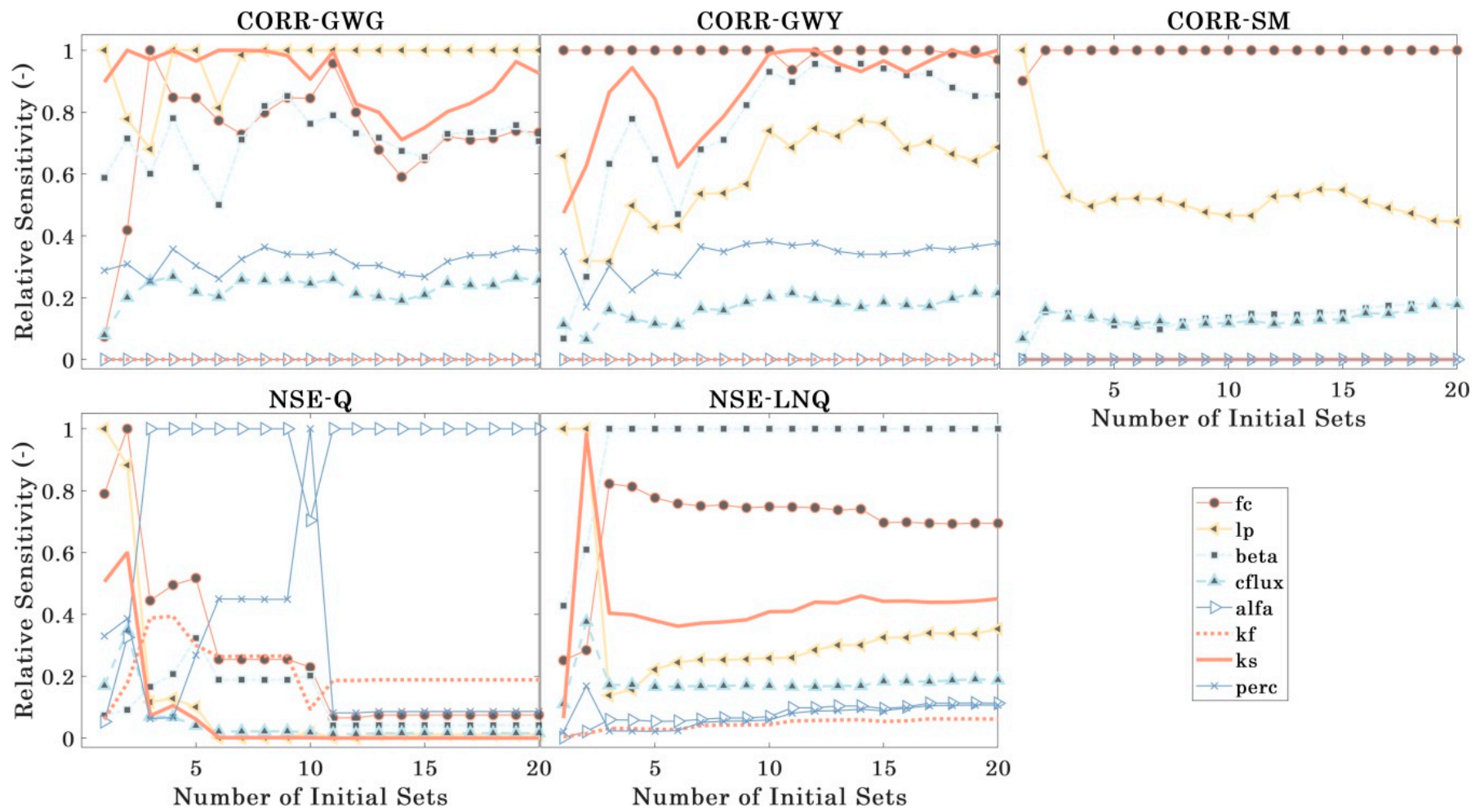


Figure 4. Relative sensitivity of objective functions to HBV model parameters as a function of the number of initial parameter sets.

5.2. Model Calibration and Validation

Table 6 shows different dimensions in our results (different cases, different objective functions, different optimization algorithms and calibration versus validation). Herein the case numbers from 1 to 4 indicate different hydrologic processes included in the model calibration. Appropriate objective functions are then included in the multi-objective calibration based on the considered processes (shown as bold in the Table 6). Three search algorithms, PEST-LM, SCE-UA and CMAES, are particularly selected, as they are widely used in hydrology. Model validation results are also presented next to the calibration in the same table.

Obviously, the validation results were mostly poorer than the calibration results. In addition, the two global optimization methods tested in this study (SCE-UA and CMAES) predominantly yielded better results than the gradient-based Levenberg Marquardt method. The best NSE-Q value obtained by the PEST-LM was 0.77 for case 1 (high flows), however with SCE-UA and CMAES, values of 0.90 and 0.89 could be obtained, respectively. The best NSE-LNQ value obtained by PEST-LM was 0.81 for case 1 (low flows), and SCE-UA and CMAES resulted in 0.81 and 0.82, respectively.

Table 6. Summary of the calibration and validation results for four cases.

Case	Processes	Objective Function	CAL			VAL		
			PEST_LM	SCE-UA	CMAES	PEST_LM	SCE-UA	CMAES
1	Q—High flows	CORR_GWG	−0.14	−0.36	−0.38	0.02	−0.05	−0.03
		CORR_GWY	0.51	0.71	0.73	0.61	0.67	0.68
		CORR_SM	0.82	0.83	0.78	0.68	0.67	0.67
		NSE-Q	0.77	0.90	0.89	0.77	0.90	0.89
		NSE-LNQ	0.53	0.86	0.85	0.44	0.84	0.84
		CORR_GWG	−0.38	−0.38	−0.36	−0.03	0.07	0.03
	Q—Low flows	CORR_GWY	0.73	0.62	0.72	0.66	0.75	0.73
		CORR_SM	0.82	0.79	0.78	0.69	0.64	0.67
		NSE-Q	0.84	0.84	0.81	0.84	0.84	0.81
		NSE-LNQ	0.81	0.80	0.82	0.81	0.79	0.76
		CORR_GWG	−0.37	−0.28	−0.27	−0.07	−0.10	−0.05
		CORR_GWY	0.73	0.75	0.75	0.63	0.70	0.75
2	GW—Only	CORR_SM	0.79	0.83	0.84	0.68	0.61	0.58
		NSE-Q	0.88	−286	0.72	0.88	−286	0.72
		NSE-LNQ	0.85	−0.68	−0.91	0.83	−1.18	−0.48
		CORR_GWG	−0.05	−0.38	0.18	−0.04	0.01	−0.08
		CORR_GWY	0.48	0.69	0.25	0.58	0.75	0.37
		CORR_SM	0.83	0.85	0.85	0.71	0.65	0.66
	SM—Only	NSE-Q	0.82	−26.7	0.20	0.82	−26.7	0.20
		NSE-LNQ	0.63	0.17	−1.39	0.54	−0.24	−1.39
		CORR_GWG	−0.11	−0.37	−0.31	−0.03	0.00	−0.04
		CORR_GWY	0.54	0.72	0.73	0.61	0.72	0.67
		CORR_SM	0.80	0.83	0.82	0.69	0.66	0.65
		NSE-Q	0.45	0.87	0.85	0.45	0.87	0.85
3	Q + GW	NSE-LNQ	0.64	0.84	0.83	0.57	0.78	0.82
		CORR_GWG	−0.36	−0.38	−0.36	0.10	0.00	0.01
		CORR_GWY	0.68	0.72	0.68	0.70	0.74	0.75
		CORR_SM	0.84	0.83	0.84	0.64	0.66	0.66
		NSE-Q	0.85	0.88	0.87	0.71	0.88	0.87
		NSE-LNQ	0.81	0.79	0.67	−0.69	0.73	0.62
	Q + SM	CORR_GWG	−0.15	−0.37	−0.34	−0.02	−0.06	−0.07
		CORR_GWY	0.55	0.73	0.73	0.64	0.67	0.65
		CORR_SM	0.83	0.84	0.83	0.69	0.68	0.66
		NSE-Q	0.80	0.88	0.87	0.80	0.88	0.87
		NSE-LNQ	0.58	0.86	0.83	0.50	0.82	0.81

It is worth pointing out that there is a trade-off between single and multi-objective calibrations. For instance, single objective calibration improved the SM-Only simulations from 0.83 to 0.85 using the SCE-UA method (case 2, CORR_SM). Multi-objective calibration (Q + SM) ensured a reasonable

hydrograph match (NSE-Q 0.88), whereas SM-Only calibration failed for both SCE-UA and CMAES, i.e., NSE-Q \sim -26.7 and 0.2, respectively.

Based on the validation results, the GW-Only calibration algorithm resulted in a value of 0.75 for CORR_GWY using the SCE-UA. Interestingly, the performance did not deteriorate, i.e., 0.72 for CORR_GWY, when we added NSE-LNQ to the set of objective functions. The case 3 (Q + GW) calibration using SCE-UA substantially improved the NSE-LNQ metric from -1.18 to 0.78 in the validation period; i.e., the main motivation of our multi-objective calibration framework proposed in this study (Table 6). With this in mind, the GRACE data from the Rhine basin apparently caused a conflict with the Moselle well measurements, like in the last case (Q + GW + SM); i.e., they showed negative CORR_GWG results. Out of curiosity, we tested GWG-Only and GWY-Only calibration scenarios using all three methods and obtained NSE of 0.18 and 0.76, respectively using the SCE-UA method. This confirms the low reliability of the coarse GRACE (GWG) data, as compared to the groundwater well measurements (GWY) in the Moselle River Basin.

Figure 5 shows the results of the model calibration framework comprising seven cases listed in the legend. For brevity, only the CMAES results are presented, as Table 6 gives the overall calibration and validation results. Here one can clearly see the trade-off between single and multiple objective functions in the calibration. Obviously, Only Q and Only Q-LN perform best in both calibration and validation periods, spring and fall seasons in particular. During low flows from June to September, the HBV model captures the hydrological behavior well. Only-GW calibration fails to capture the streamflow dynamics in the Moselle River, and Only-SM calibration fails to simulate low flow periods, since these constraints are not given to the model via the objective functions. In other words, if the objective function does not incorporate appropriate process information, the hydrologic model usually fails/ignores to simulate the underlying physics.

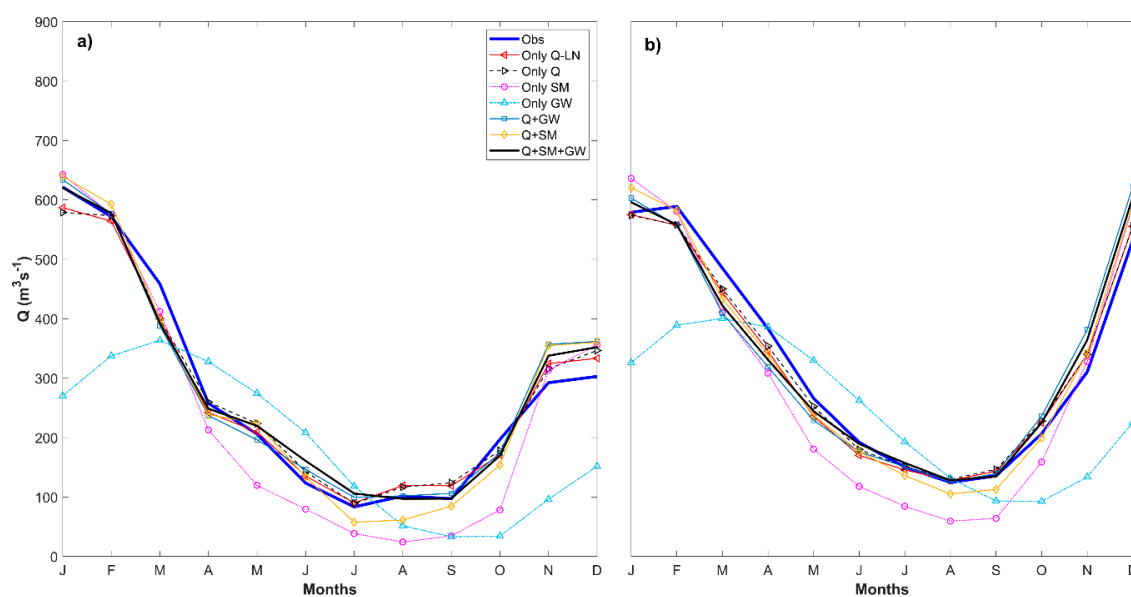


Figure 5. Observed average monthly discharge at the Cochem Gauge, and simulated average monthly discharge from HBV model for the (a) calibration (2002–2006) and (b) validation (1954–2001) periods for seven cases calibrated by the CMAES global search algorithm.

6. Discussion

Satellite-based observations have been incorporated into different model calibrations to constrain the spatially-distributed hydrologic model behavior towards reliable physics [3,6,10,12]. In this study, we sought an answer to the question whether satellite-based products can also be helpful for guiding lumped models throughout the calibration, an issue which has been rarely investigated [2]. Obviously spatial information is lost during the averaging (lumping) process, but the temporal information from

the satellite data are shown to be helpful for the model evaluations. SM-only and GW-only calibrations fail on discharge, and in combination with Q they add little to the performance.

Also the combination of Q, SM and G data gives good results for Q simulation for the right reasons (SM and G simulation is reasonable, in particular, compared to Q only simulations).

We received P and PET data from two sources: BfG-Koblenz (up to 2006) and van Osnabrugge et al. [29] (up to 2015). We noticed wetter characteristics in the latter data, and decided to stick to the first data from BfG-Koblenz for the model calibration, as the initial calibration results were promising. Instead, the more recent data were used only for the validation. We assume that three to five year daily data (2002–2004 and 2002–2006) are sufficient for model calibration. Besides, the long validation period is probably the main reason for good validation performances. In contrast to the conventional model calibration approaches, we used more recent data for calibration and older data for validation to benefit from the opportunities of new instrumentation.

We limited the PEST-LM calibration with 375 iterations corresponding to nearly 3,000 model runs, which is comparable with the maximum number of runs of SCE-UA and CMAES. However, PEST-LM usually ended much earlier than the maximum number of runs. This is from the fact that the initial point in the model calibration is key to determine a local minimum, or to coincidentally find a global optimum. Apparently, PEST-LM suffered from the local minimum issue. Comparing the NSE-LNQ values in the first column (PEST-LM) for low flows and GW-only, and also the NSE values in the first column of high flows and SM-only, clearly show the importance of the initial parameter set (start point in the solution space) and the local minimum problem of PEST-LM.

We only used the HBV model, which has GW and SM storages, whereas other similar models with three buckets can be also tested to assess the uncertainties from model selection. The effects of model structure and different remote sensing products on the calibration of five different rainfall-runoff models with different spatial discretization have been discussed in the previous studies [2,69].

Moreover, we only used NSE and CORR as objective functions, whereas there are many more objective functions, e.g., Kling and Gupta Efficiency [70], Spatial Efficiency (SPAEF) [3], Spearman and Kendall Rank Correlation Coefficients [71], than these two metrics tested in the spatial and temporal domain. Although SPAEF has been designed to evaluate spatial performance of distributed models, it can be applied to the time-series comparison in this study as well, since the three components in the SPAEF metric, i.e., the correlation coefficient, coefficient of variation and histogram match, are important characteristics of temporal data too. In a subsequent study, we plan to apply SPAEF in a temporal model calibration. Moreover, Spearman's rho and Kendall's tau coefficients can capture non-linear relations between two variables, while Pearson's metric (CORR) only shows the linear correlations.

We expected to see similar parameters showing a high sensitivity for fitting HBV model simulations to observed and remotely-sensed groundwater data (well and GRACE). However, the KS parameter showed the highest sensitivity for GWY, and the LP parameter showed the highest sensitivity for GWG. This can be an initial alert/indicator to the poor GRACE performance, since the LP parameter is relevant for the soil moisture reservoir of the model, and it hardly influences simulated groundwater behavior. We used GRACE-Rhine data directly instead of estimating or finding GRACE-Moselle data. This must have substantial effects on the results as the Rhine Basin size is large, and there can be different aquifer dynamics within the basin.

7. Conclusions

At the beginning of this study, we applied sensitivity analysis to determine the most important parameters for the calibration using the groundwater well measurements (GWY) and remotely-sensed products (GRACE, ESA, AQUA and SMAP) data. We selected three different objective functions, namely NSE-Q, NSE-LNQ and CORR. We can draw the following conclusions from the results of this study:

- Based on the sensitivity analysis results, we find that different hydrologic processes are sensitive for different parameters. The HBV model's groundwater performance (GWY) was most sensitive

to the KS parameter, whereas the model's soil moisture performance was most sensitive to the FC parameter. Also we confirm that 20 different initial parameter sets [64] using Latin Hypercube sampling are sufficient for globally encapsulating the most sensitive parameters.

- Based on the calibration and validation results, we show that the two global methods perform better than the local Levenberg Marquardt method. An upper limit of 3,000 model runs appeared to be plausible for both the local and global optimization of the HBV model. Also including groundwater and remotely-sensed soil moisture information slightly (up to ~10%) improved not only the GW and SM simulation performances of the model, but also the simulation of the observed discharge behavior of the HBV model. This is only exploiting the temporal information of the satellite-based data, since we spatially averaged the distributed data to get the time series of SM and GW. This was also confirmed by a recent study by Nijzink et al. [2].

Further work on the combination of lumped hydrological models with newly-available high resolution remote sensing products, e.g. SMAP, is necessary. Also it is recommended to test the value of spatially-distributed well measurements in addition to the GRACE on the performance of a fully distributed version of HBV (i.e., HYPE) or another hydrologic model like mHM [3]. Since ESA is a combined product including AQUA, only ESA can be used to avoid redundancy in the future studies. Spatial variation of droughts and high flows should also be assessed with different performance metrics to exploit both the temporal and spatial value of the remote sensing products. Promising deep learning methods such as Long Short-Term Memory Network (LSTM) [72,73] and Tensorflow can also be incorporated to simulate spatially-distributed runoff based on spatially-distributed layers of precipitation and potential evapotranspiration maps.

Author Contributions: Conceptualization, M.C.D., S.S., M.J.B. and S.O.; methodology, M.C.D., S.O., E.T. (Emir Toker), Ö.E., H.T., S.E., H.K.D., A.B.S., Ö.S., E.T. (Ecem Tuncer), H.H., T.İ.Ö., A.Ö. and M.B.A.; validation M.C.D., S.O., E.T. (Emir Toker), Ö.E., H.T., S.E., H.K.D., A.B.S., Ö.S., E.T. (Ecem Tuncer), H.H., T.İ.Ö. and M.J.B.; writing—original draft preparation, M.C.D., A.Ö., H.E., M.M.K., E.E.B., K.A. and S.O.; writing—review and editing, M.C.D., A.Ö., S.O., H.E., H.K.D., M.M.K., K.A., M.B., A.A., Ö.V., A.Ö., S.S. and M.J.B.; visualization, M.C.D., E.T. (Emir Toker), H.K.D.; supervision, M.C.D., S.S. and M.J.B.; project administration, M.C.D.

Funding: This research received funding from the TÜBİTAK grant 118C020.

Acknowledgments: The first author (M.C.D.) is supported by the Turkish Scientific and Technical Research Council (TÜBİTAK grant 118C020). The co-author K.A. is supported by the Professional Development Research University (PDRU) grant no. Q.J130000.21A2.04E10 of Universiti Teknologi Malaysia. Authors S.S. and M.C.D. acknowledge the financial support for the SPACE project by the Villum Foundation (<http://villumfonden.dk/>) through their Young Investigator Program (grant VKR023443). We also acknowledge the financial support of the Cornelis Lely Stichting (C.L.S.), Project No. 20957310. We thank Bart van Osnabrugge, Albrecht Weerts and Eric Sprokkereef for providing the meteorological data from the IMPREX project. Groundwater data for the French part of the Moselle sub-basin were obtained from Portail national d'Accès aux Données sur les Eaux Souterraines (<http://www.adeseaufrance.fr/>), and for the Saar sub-basin from the Wasser- und Schifffahrtsamt Saarbrücken. We thank Dennis Meissner from Referat M2-Mitarbeiter/innen group in BfG, Koblenz for the valuable discussions about Saar groundwater data during C.L.S. project meetings.

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

References

1. Jiang, D.; Wang, K. The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review. *Water* **2019**, *11*, 1615. [[CrossRef](#)]
2. Nijzink, R.C.; Almeida, S.; Pechlivanidis, I.G.; Capell, R.; Gustafssons, D.; Arheimer, B.; Parajka, J.; Freer, J.; Han, D.; Wagener, T.; et al. Constraining Conceptual Hydrological Models with Multiple Information Sources. *Water Resour. Res.* **2018**, *54*, 8332–8362. [[CrossRef](#)]
3. Demirel, M.C.; Mai, J.; Mendiguren, G.; Koch, J.; Samaniego, L.; Stisen, S. Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 1299–1315. [[CrossRef](#)]

4. Becker, R.; Koppa, A.; Schulz, S.; Usman, M.; aus der Beek, T.; Schüth, C. Spatially distributed model calibration of a highly managed hydrological system using remote sensing-derived ET data. *J. Hydrol.* **2019**, *577*, 123944. [[CrossRef](#)]
5. Rientjes, T.H.M.; Muthuwatta, L.P.; Bos, M.G.; Booij, M.J.; Bhatti, H.A. Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *J. Hydrol.* **2013**, *505*, 276–290. [[CrossRef](#)]
6. Li, Y.; Grimaldi, S.; Pauwels, V.R.N.; Walker, J.P. Hydrologic model calibration using remotely sensed soil moisture and discharge measurements: The impact on predictions at gauged and ungauged locations. *J. Hydrol.* **2018**, *557*, 897–909. [[CrossRef](#)]
7. Tapley, B.D.; Bettadpur, S.; Watkins, M.; Reigber, C. The gravity recovery and climate experiment: Mission overview and early results. *Geophys. Res. Lett.* **2004**, *31*. [[CrossRef](#)]
8. 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**, *17*, 287–307. [[CrossRef](#)]
9. 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](#)]
10. Zink, M.; Mai, J.; Cuntz, M.; Samaniego, L. Conditioning a Hydrologic Model Using Patterns of Remotely Sensed Land Surface Temperature. *Water Resour. Res.* **2018**, *54*, 2976–2998. [[CrossRef](#)]
11. Immerzeel, W.W.; Droogers, P. Calibration of a distributed hydrological model based on satellite evapotranspiration. *J. Hydrol.* **2008**, *349*, 411–424. [[CrossRef](#)]
12. Xiong, L.; Zeng, L. Impacts of Introducing Remote Sensing Soil Moisture in Calibrating a Distributed Hydrological Model for Streamflow Simulation. *Water* **2019**, *11*, 666. [[CrossRef](#)]
13. Kamamia, A.W.; Mwangi, H.M.; Feger, K.-H.; Julich, S. Assessing the impact of a multimetric calibration procedure on modelling performance in a headwater catchment in Mau Forest, Kenya. *J. Hydrol. Reg. Stud.* **2019**, *21*, 80–91. [[CrossRef](#)]
14. 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](#)]
15. Tobin, K.; Bennett, M. Improving Alpine Summertime Streamflow Simulations by the Incorporation of Evapotranspiration Data. *Water* **2019**, *11*, 112. [[CrossRef](#)]
16. Pushpalatha, R.; Perrin, C.; Le Moine, N.; Andréassian, V. A review of efficiency criteria suitable for evaluating low-flow simulations. *J. Hydrol.* **2012**, *420–421*, 171–182. [[CrossRef](#)]
17. Krause, P.; Boyle, D.P.; Bäse, F. Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* **2005**, *5*, 89–97. [[CrossRef](#)]
18. Willmott, C.J. On the validation of models. *Phys. Geogr.* **1981**, *2*, 184–194. [[CrossRef](#)]
19. Koch, J.; Siemann, A.; Stisen, S.; Sheffield, J. Spatial validation of large scale land surface models against monthly land surface temperature patterns using innovative performance metrics. *J. Geophys. Res. Atmos.* **2016**. [[CrossRef](#)]
20. Mendiguren, G.; Koch, J.; Stisen, S. Spatial pattern evaluation of a calibrated national hydrological model—A remote-sensing-based diagnostic approach. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5987–6005. [[CrossRef](#)]
21. Stisen, S.; Koch, J.; Sonnenborg, T.O.; Refsgaard, J.C.; Bircher, S.; Ringgaard, R.; Jensen, K.H. Moving beyond runoff calibration—Multi-variable optimization of a surface-subsurface-atmosphere model. *Hydrol. Process.* **2018**. [[CrossRef](#)]
22. Booij, M.J.; Krol, M.S. Balance between calibration objectives in a conceptual hydrological model. *Hydrol. Sci. J.* **2010**, *55*, 1017–1032. [[CrossRef](#)]
23. Koch, J.; Jensen, K.H.; Stisen, S. Toward a true spatial model evaluation in distributed hydrological modeling: Kappa statistics, Fuzzy theory, and EOF-analysis benchmarked by the human perception and evaluated against a modeling case study. *Water Resour. Res.* **2015**, *51*, 1225–1246. [[CrossRef](#)]
24. Koch, J.; Demirel, M.C.; Stisen, S. The SPATial Efficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models. *Geosci. Model Dev.* **2018**, *11*, 1873–1886. [[CrossRef](#)]
25. 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](#)]
26. Lindström, G.; Johansson, B.; Persson, M.; Gardelin, M.; Bergstrom, S. Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* **1997**, *201*, 272–288. [[CrossRef](#)]

27. Bergström, S. *Development and Application of a Conceptual Runoff Model for Scandinavian Catchments*; SMHI: Norrköping, Sweden, 1976.
28. Asselman, N.E.M.; Middelkoop, H.; Van Duijk, P.M. The impact of changes in climate and land use on soil erosion, transport and deposition of suspended sediment in the River Rhine. *Hydrol. Process.* **2003**, *17*, 3225–3244. [[CrossRef](#)]
29. Van Osnabrugge, B.; Weerts, A.H.; Uijlenhoet, R. genRE: A Method to Extend Gridded Precipitation Climatology Data Sets in Near Real-Time for Hydrological Forecasting Purposes. *Water Resour. Res.* **2017**, *53*, 9284–9303. [[CrossRef](#)]
30. Demirel, M.C.; Booij, M.J.; Hoekstra, A.Y. The skill of seasonal ensemble low-flow forecasts in the Moselle River for three different hydrological models. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 275–291. [[CrossRef](#)]
31. Descy, J.P. Ecology of the phytoplankton of the River Moselle: Effects of disturbances on community structure and diversity. *Hydrobiologia* **1993**, *249*, 111–116. [[CrossRef](#)]
32. Demirel, M.C.; Booij, M.J.; Hoekstra, A.Y. Effect of different uncertainty sources on the skill of 10 day ensemble low flow forecasts for two hydrological models. *Water Resour. Res.* **2013**, *49*, 4035–4053. [[CrossRef](#)]
33. Demirel, M.C.; Booij, M.J.; Hoekstra, A.Y. Identification of appropriate lags and temporal resolutions for low flow indicators in the River Rhine to forecast low flows with different lead times. *Hydrol. Process.* **2013**, *27*, 2742–2758. [[CrossRef](#)]
34. Secretariat, G. Implementation plan for the global observing system for climate in support of the UNFCCC (2010 Update). In Proceedings of the Conference of the Parties (COP), Copenhagen, Denmark, 7–18 December 2009; pp. 7–19.
35. Wagner, W.; Dorigo, W.; De Jeu, R.; Fernandez, D.; Benveniste, J.; Haas, E.; Ertl, M. Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *1*, 315–321. [[CrossRef](#)]
36. Dorigo, W.; Gruber, A.; Scanlon, T.; Hahn, S.; Kidd, R.; Paulik, C.; Reimer, C.; Van der Schalie, R.; Preimesberger, W.; De Jeu, R.W.W. ESA Soil Moisture Climate Change Initiative (Soil_Moisture_cci): Version 04.4 Data Collection. Available online: <https://catalogue.ceda.ac.uk/uuid/dce27a397eaf47e797050c220972ca0e?jump=related-docs-anchor> (accessed on 10 January 2019).
37. Gruber, A.; Scanlon, T.; Van Der Schalie, R.; Wagner, W.; Dorigo, W. Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. *Earth Syst. Sci. Data* **2019**, *11*, 717–739. [[CrossRef](#)]
38. Dorigo, W.; Wagner, W.; Albergel, C.; Albrecht, F.; Balsamo, G.; Brocca, L.; Chung, D.; Ertl, M.; Forkel, M.; Gruber, A.; et al. ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.* **2017**, *203*, 185–215. [[CrossRef](#)]
39. Wagner, W.; Lemoine, G.; Rott, H. A method for estimating soil moisture from ERS Scatterometer and soil data. *Remote Sens. Environ.* **1999**, *70*, 191–207. [[CrossRef](#)]
40. Naeimi, V.; Scipal, K.; Bartalis, Z.; Hasenauer, S.; Wagner, W. An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1999–2013. [[CrossRef](#)]
41. Scipal, K.; Wagner, W.; Trommler, M.; Naumann, K. The global soil moisture archive 1992–2000 from ERS scatterometer data: First results. *Int. Geosci. Remote Sens. Symp.* **2002**, *3*, 1399–1401. [[CrossRef](#)]
42. Wagner, W.; Blöschl, G.; Pampaloni, P.; Calvet, J.-C.; Bizzarri, B.; Wigneron, J.-P.; Kerr, Y. Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. *Hydrol. Res.* **2007**, *38*, 1–20. [[CrossRef](#)]
43. Wagner, W.; Brocca, L.; Naeimi, V.; Reichle, R.; Draper, C.; De Jeu, R.; Ryu, D.; Su, C.H.; Western, A.; Calvet, J.C.; et al. Clarifications on the “comparison between SMOS, VUA, ASCAT, and ECMWF Soil Moisture Products over Four Watersheds in U.S.”. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1901–1906. [[CrossRef](#)]
44. Owe, M.; de Jeu, R.; Holmes, T. Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res. Earth Surf.* **2008**, *113*, 1–17. [[CrossRef](#)]
45. Parinussa, R.M.; Holmes, T.R.H.; De Jeu, R.A.M. Soil moisture retrievals from the windSat spaceborne polarimetric microwave radiometer. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2683–2694. [[CrossRef](#)]

46. Li, L.; Gaiser, P.W.; Gao, B.C.; Bevilacqua, R.M.; Jackson, T.J.; Njoku, E.G.; Rüdiger, C.; Calvet, J.C.; Bindlish, R. WindSat global soil moisture retrieval and validation. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2224–2241. [[CrossRef](#)]
47. Kerr, Y.H.; Waldteufel, P.; Richaume, P.; Wigneron, J.P.; Ferrazzoli, P.; Mahmoodi, A.; Al Bitar, A.; Cabot, F.; Gruhier, C.; Juglea, S.E.; et al. The SMOS Soil Moisture Retrieval Algorithm. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1384–1403. [[CrossRef](#)]
48. Parinussa, R.M.; Holmes, T.R.H.; Wanders, N.; Dorigo, W.A.; De Jeu, R.A.M. A preliminary study toward consistent soil moisture from AMSR2. *J. Hydrometeorol.* **2015**, *16*, 932–947. [[CrossRef](#)]
49. Njoku, E.G.; Entekhabi, D. Passive Microwave Remote Sensing of Soil Moisture. *J. Hydrol.* **1996**, *184*, 101–129. [[CrossRef](#)]
50. Njoku, E.G.; Jackson, T.J.; Lakshmi, V.; Chan, T.K.; Nghiem, S.V. Soil moisture retrieval from AMSR-E. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 215–229. [[CrossRef](#)]
51. Gruhier, C.; De Rosnay, P.; Hasenauer, S.; Holmes, T.; De Jeu, R.; Kerr, Y.; Mougou, E.; Njoku, E.; Timouk, F. Soil moisture active and passive microwave products: Intercomparison and evaluation over a Sahelian site. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 141–156. [[CrossRef](#)]
52. Entekhabi, D.; Njoku, E.; O’Neill, P.; Spencer, M.; Jackson, T.; Entin, J.; Im, E.; Kellogg, K. The (Soil Moisture Active/Passive Mission (SMAP)). *IEEE Int. Geosci. Remote Sens. Symp.* **2008**, *3*, III-1–III-4. [[CrossRef](#)]
53. Entekhabi, D.; Njoku, E.G.; O’Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The soil moisture active passive (SMAP) mission. *Proc. IEEE* **2010**, *98*, 704–716. [[CrossRef](#)]
54. Sadri, S.; Wood, E.F.; Pan, M. A SMAP-Based Drought Monitoring Index for the United States. *Hydrol. Earth Syst. Sci. Discuss.* **2018**, 1–19. [[CrossRef](#)]
55. Brown, M.E.; Escobar, V.; Moran, S.; Entekhabi, D.; O’Neill, P.E.; Njoku, E.G.; Doorn, B.; Entin, J.K. NASA’s soil moisture active passive (SMAP) mission and opportunities for applications users. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1125–1128. [[CrossRef](#)]
56. Beck, H.E.; van Dijk, A.I.J.M.; de Roo, A.; Miralles, D.G.; McVicar, T.R.; Schellekens, J.; Bruijnzeel, L.A. Global-scale regionalization of hydrologic model parameters. *Water Resour. Res.* **2016**. [[CrossRef](#)]
57. Singh, S.K.; Bárdossy, A. Calibration of hydrological models on hydrologically unusual events. *Adv. Water Resour.* **2012**, *38*, 81–91. [[CrossRef](#)]
58. Tian, Y.; Booij, M.J.; Xu, Y.-P. Uncertainty in high and low flows due to model structure and parameter errors. *Stoch. Environ. Res. Risk Assess.* **2014**, *28*, 319–332. [[CrossRef](#)]
59. Uhlenbrook, S.; Leibundgut, C. Process-oriented catchment modelling and multiple-response validation. *Hydrol. Process.* **2002**, *16*, 423–440. [[CrossRef](#)]
60. ASCE. Criteria for Evaluation of Watershed Models. *J. Irrig. Drain. Eng.* **1993**. [[CrossRef](#)]
61. Dawson, C.W.; Abrahart, R.J.; See, L.M. HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environ. Model. Softw.* **2007**. [[CrossRef](#)]
62. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**. [[CrossRef](#)]
63. Reusser, D.E.; Blume, T.; Schaefli, B.; Zehe, E. Analysing the temporal dynamics of model performance for hydrological models. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 999–1018. [[CrossRef](#)]
64. Demirel, M.; Koch, J.; Mendiguren, G.; Stisen, S. Spatial Pattern Oriented Multicriteria Sensitivity Analysis of a Distributed Hydrologic Model. *Water* **2018**, *10*, 1188. [[CrossRef](#)]
65. Iman, R.L.; Consultants, S.T. Latin Hypercube Sampling. *Wiley StatsRef Stat. Ref. Online* **1998**. [[CrossRef](#)]
66. Doherty, J. *PEST: Model Independent Parameter Estimation. Fifth Edition of User Manual*; Watermark Numerical Computing: Brisbane, QLD, Australia, 2005.
67. Duan, Q.; Sorooshian, S.; Gupta, V. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **1992**. [[CrossRef](#)]
68. Hansen, N.; Ostermeier, A. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In Proceedings of the 2002 Congress on Evolutionary Computation, Honolulu, HI, USA, 12–17 May 2002.
69. Van Esse, W.R.; Perrin, C.; Booij, M.J.; Augustijn, D.C.M.; Fenicia, F.; Kavetski, D.; Lobligois, F. The influence of conceptual model structure on model performance: A comparative study for 237 French catchments. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 4227–4239. [[CrossRef](#)]

70. Kling, H.; Gupta, H. On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability. *J. Hydrol.* **2009**, *373*, 337–351. [[CrossRef](#)]
71. Spearman Rank Correlation Coefficient. In *The Concise Encyclopedia of Statistics*; Springer: New York, NY, USA, 2008; pp. 502–505.
72. Kratzert, F.; Klotz, D.; Shalev, G.; Klambauer, G.; Hochreiter, S.; Nearing, G. Benchmarking a Catchment-Aware Long Short-Term Memory Network (LSTM) for Large-Scale Hydrological Modeling. *Hydrol. Earth Syst. Sci. Discuss.* **2019**, 1–32. [[CrossRef](#)]
73. Bowes, B.D.; Sadler, J.M.; Morsy, M.M.; Behl, M.; Goodall, J.L. Forecasting Groundwater Table in a Flood Prone Coastal City with Long Short-term Memory and Recurrent Neural Networks. *Water* **2019**, *11*, 1098. [[CrossRef](#)]



© 2019 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/>).