HYDROLOGICAL SCIENCES JOURNAL
https://doi.org/10.1080/02626667.2020.1810252

Comparison of single-site, multi-site and multi-variable SWAT calibration strategies

Ana Clara Lazzari Franco[a], Debora Yumi de Oliveira[a] and Nadia Bernardi Bonumá[b]

[a]Graduate Program in Environmental Engineering, Federal University of Santa Catarina - UFSC, Florianopolis, Brazil; [b]Department of Sanitary and Environmental Engineering, Federal University of Santa Catarina, Florianópolis, Brazil

ABSTRACT
This study compares single-site, multi-site and multi-variable SWAT calibration. The SWAT model was applied to a large basin (63 884 km^2) and calibrated at a monthly time step with the SUFI-2 algorithm, using the Kling-Gupta efficiency (KGE) as the objective function. Multi-variable calibration was performed by combining streamflow and remote sensing-derived actual evapotranspiration data. Parameter transferability was also investigated, by daily time step validation. The KGE for the outlet ranged from 0.73 to 0.86, and the average KGE of all streamflow gauge stations ranged from 0.73 to 0.80, reflecting a good overall simulation performance for the monthly time step. In daily time step validation, KGE ranged from 0.62 to 0.68, and the Nash-Sutcliffe efficiency ranged from 0.40 to 0.60, for the average of all gauge stations. Multi-site and multi-variable calibrations did not significantly improve inner sub-basin simulation performance but improved streamflow uncertainty when compared to single-site calibration.

ARTICLE HISTORY
Received 23 July 2019
Accepted 15 June 2020

EDITOR
A. Castellarin

ASSOCIATE EDITOR
M. Piniewski

KEYWORDS
uncertainty; SUFI-2; SWAT model; Kling-Gupta efficiency (KGE); Nash-Sutcliffe efficiency (NS); parameter transferability; remote sensed evapotranspiration; Iguaçu river basin; unsolved problems in hydrology (UPH); UPH #20

1 Introduction

Hydrological models are important and frequently used tools in the environmental field, with growing application for predicting climate change impacts (Daggupati *et al.* 2015b, Devia *et al.* 2015, Sudheer *et al.* 2007, Wi *et al.* 2015; Wu and Chen 2015). Hydrological model simulations are associated with many sources of uncertainty. How to disentangle and reduce model uncertainties in hydrological prediction was pointed out as one of the 23 unsolved problems in hydrology (UPH; *Blöschl et al.* 2019, UPH #20). Model uncertainties are intrinsic and arise from different sources, where model structure, initial conditions, observational data, and parameter values are the main ones (Wagener and Gupta 2005, Liu and Gupta 2007, Yen *et al.* 2016). Due to the large number of parameters and different sources of uncertainty, distributed hydrological model calibration and uncertainty analysis is inevitable (Zhang *et al.* 2016). When calibrating a model, different parameter sets can result in the same, or very similar, simulation output, resulting in parameter non-uniqueness (Abbaspour *et al.* 2004, Schuol and Abbaspour 2006, 2015) or equifinality (Beven and Binley 1992, Beven 2001) problems.

Good performing simulations for streamflow at the basin outlet can be produced at the expense of error compensation (Chiang *et al.* 2014), which is especially problematic in single-variable and single-site calibration (Cao *et al.* 2006). Multi-site and multi-variable calibration approaches are promising alternatives to work around this problem, since less parameter sets will be able to satisfy the calibration criteria for all variables at all considered locations at the same time (Bergstrom *et al.* 2002, White and Chaubey 2005, Daggupati *et al.* 2015a, 2015b).

With the development of (semi)distributed hydrological models, multi-site calibration approaches are being more frequently applied (Zhang *et al.* 2010, Wi *et al.* 2015, Leta *et al.* 2017, Nkiaka *et al.* 2018). Previous studies compared multi-site and single-site streamflow calibration with respect to model output performance. Moussa *et al.* (2007) calibrated a 543 km^2 French basin considering both nested and non-nested sub-basins and reported better performance for interior gauges when calibration was conducted with the multi-site approach, over the single-site outlet calibration. Zhang *et al.* (2008) calibrated the SWAT model on a 239 km^2 basin in the USA and reported better performance for both the outlet and interior points when using multi-site calibration, over single-site outlet calibration. Similarly, Daggupati *et al.* (2015b) also concluded that multi-site calibration outperforms outlet single-site calibration when calibrating a 28 330 km^2 basin in the USA using the SWAT model. In contrast, some authors have found similar results for calibration with interior flow data and single-site outlet flow data (Khakbaz *et al.* 2012, Lerat *et al.* 2012). Her and Chaubey (2015) investigated the impact of multi-site calibration on model output performance and uncertainty and suggested that it can reduce equifinality and

CONTACT Ana Clara Lazzari Franco ana.franco@ufsc.br Department of Sanitary and Environmental Engineering, Federal University of Santa Catarina, Florianópolis, Brazil
Supplemental data for this article can be accessed here.

increase model output uncertainty when compared to single-site calibration.

Remotely sensed evapotranspiration (ET) has high spatial and temporal resolution and can be used to infer soil moisture, an important component of the soil water balance (Allen *et al.* 2007, Githui *et al.* 2012). In areas where the measured data are inadequate or inexistent, remotely sensed data, such as ET and soil moisture, can be a promising alternative for model calibration and validation (Immerzeel and Droogers 2008).

Multi-variable calibration can increase model performance and model fidelity. Bergstrom *et al.* (2002) used multi-variable parameter estimation and internal control of the HBV model and concluded that the consideration of variables other than streamflow can increase confidence in model results. Rientjes *et al.* (2013) used streamflow and remotely sensed evapotranspiration to compare single and multi-variable calibration approaches on HBV model and reported better results when the model is calibrated against both variables. Wanders *et al.* (2014) used remotely sensed soil moisture and streamflow to calibrate the LISFLOOD model for the Upper Danube River basin and concluded that the simultaneous use of both variables can lead to a better simulation of discharge in upstream areas. The AWRA-L model was calibrated with evapotranspiration and soil moisture remotely sensed data and good streamflow simulation results were obtained, especially by using evapotranspiration data (Kunnath-poovakka *et al.* 2016, 2018).

More commonly, multi-variable calibration with the SWAT model uses streamflow and suspended sediment data (e.g. Brighenti *et al.* 2019). Remotely sensed data have been more frequently used in multi-variable calibration with SWAT in recent years. Githui *et al.* (2012) estimated groundwater recharge using remotely sensed evapotranspiration and streamflow measurements to calibrate the SWAT model in an irrigated basin in Australia. Rajib *et al.* (2016) reported reduced parameter uncertainty when calibrating the SWAT model with streamflow and remotely sensed soil moisture data simultaneously. SWAT model calibration was also carried out with satellite-based evapotranspiration only, in southwestern Nigeria (Odusanya *et al.* 2019). Finally, Briak *et al.* (2019) reinforce the importance of multi-variable calibration and compare different calibration approaches for the SWAT model.

Incorporating multiple objectives (multiple sites, multiple objective functions, or multiple variables) in model calibration is not straightforward. Many papers evaluate different strategies to integrate discharge data from multiple sites in model calibration (e.g. Feyen *et al.* 2008; Wi *et al.* 2015; Leta *et al.* 2017, Lavenne *et al.* 2019). For example, Feyen *et al.* (2008) and Wi *et al.* (2015) compared the use of single-site calibration, sequential calibration, and simultaneous multi-site calibration. When dealing with multiple objective functions, the most common approach is based on Pareto optimality. However, different approaches have also been investigated. For example, Fenicia *et al.* (2007) applied a stepwise calibration approach, in which the model was sequentially calibrated using different objective functions at each step. In their study, after calibration with one objective function, the parameters related to that objective function were fixed, and the next calibration (using the next objective function) was performed.

Transferring parameter sets between different temporal scales is a common practice on hydrological modeling (Sudheer *et al.* 2007, Melsen *et al.* 2016). Troy *et al.* (2008) investigated the effects of transferring calibrated parameters sets of the variable infiltration capacity (VIC) model from coarser to finer temporal and spatial scales. Their results indicated that transferring parameter sets between different time scales performed better than for different spatial resolutions. Chaney *et al.* (2015) used observed monthly and annual flows to constrain parameter sets for the VIC model and evaluate the reduction in the uncertainty of daily runoff estimates, which was strongly dependent on climate. Melsen *et al.* (2016) compared the VIC model calibration and validation for hourly, daily and monthly time steps, suggesting that the monthly time step calibration cannot ensure satisfactory daily or hourly model performance. Sudheer *et al.* (2007) calibrated SWAT model at monthly and daily scales and concluded that calibration at a coarser time scale (monthly) does not ensure satisfactory model performance at a finer time scale (daily). The daily calibrated model, however, significantly improved the monthly scale simulation. Conversely, Daggupati *et al.* (2015b) calibrated the SWAT model at monthly scale and obtained good results when validating the model at daily scale. The conflicting results suggest the need for more studies in this topic.

The main objective of this study is to evaluate the impact of different calibration strategies on model streamflow simulation performance and uncertainty. Model simulations were performed in the Iguaçu river basin, located in southern Brazil. We calibrated the SWAT model using SUFI-2 algorithm and KGE as the objective function. Different calibration strategies were compared: single-site (E1 – streamflow data), multi-site (E2 – streamflow data), and multi-variable (E3 – evapotranspiration and streamflow data). Results were assessed at the main basin outlet and interior streamflow gauges. We also investigated parameter transferability from monthly to daily time scale. For this purpose, monthly calibrated parameters were validated at daily time step. We compared the results in terms of output performance and uncertainty. The calibration dataset includes streamflow measurements for the main outlet and interior flow gauges and remotely sensed derived actual evapotranspiration data.

Franco and Bonumá (2017) calibrated the SWAT model with evapotranspiration and outlet streamflow data for the upper Negro River Basin (3453 km^2), a small sub-basin of the Iguaçu River Basin. This study follows the methodology applied by Franco and Bonumá (2017), using evapotranspiration data from METRIC and KGE as objective function to calibrate the SWAT model with SUFI-2. Major contributions of the present study, compared to Franco and Bonumá (2017), are the considerable larger drainage area (63 884 km^2) and the comparison of multi-site calibration strategy with more commonly applied single-site calibration. Franco and Bonumá (2017) reported better streamflow performance with multi-variable calibration (evapotranspiration and outlet streamflow). Also, that study reported poor streamflow simulation performance of SWAT model calibrated with evapotranspiration only. Therefore, calibration with evapotranspiration data only was not evaluated in this study.

2 Materials and methods

2.1 Study area

The Iguaçu River is a tributary of the Paraná River, with approximately 1320 km length, flowing from east to west. The Iguaçu River joins the Paraná River after the Itaipu Dam, one of the largest hydropower plants in the world. The Iguaçu River Basin drains about 67 455 km^2, of which approximately 65 720 km^2 is situated in Brazilian territory and 1735 km^2 in Argentina (Fig. 1). The average annual precipitation is 1842 mm/year. Due to its topography and natural conditions, the Iguaçu River Basin is favorable for the installation of hydropower plants. Currently, five hydropower plants located along the Iguaçu River are operating. Together, their present installed capacity totals 6556 MW (Machado 2012).

The basin is largely populated (approximately 4.5 million people) and the urban population, as well as the industrial, mining, agricultural, and livestock activities, are supplied mostly by surface water from rivers within the basin (SEMA 2010).

The Iguaçu River Basin is crucial for the ecological conservation of the Atlantic forest. Forest remnants from the original Subtropical Ombrophilous Forest (SOF) in the basin are important key areas for ecological conservation, such as the Iguaçu National Park. For the modeled portion of the basin, native forest covers about half of the area (49.8%), followed by agriculture (30.3%) and pasture (9.5%). Reforestation covers 6.8% of the area and urban areas occupy about 2.2% of the total basin area. A land cover map was generated by supervised image classification with Landsat 8 imagery of the year 2015.

2.2 METRIC model

The mapping evapotranspiration at high resolution with internalized calibration (METRIC) is a model to calculate evapotranspiration as a residual of surface energy balance using satellite images, producing ET “maps” (images) to quantify ET at field scale (Allen *et al.* 2007). The METRIC model is based on SEBAL (Bastiaanssen *et al.* 1998) and was modified to be applied to mountainous terrain (Allen *et al.* 2005). SEBAL and METRIC use “hot” and “cold” pixels, where two extreme conditions (dry and wet) are found, to internally calibrate the energy balance (Allen *et al.* 2002, 2007). Field measurements from meteorological stations are required to calculate reference ET, used to extrapolate instantaneous ET from the image to daily and monthly scales (Allen *et al.* 2007). SEBAL has been largely applied and validated under different climatic conditions worldwide, with a typical accuracy of 95% at field scale on seasonal basis and 96% for annual ET of large basins (Bastiaanssen *et al.* 2005).

For this study, we used an ET dataset produced by Uda (2016). The dataset was produced using the METRIC model applied to MODIS imagery for the Iguaçu river basin. The data comprises monthly actual ET for the years 2006, 2007, and 2009.

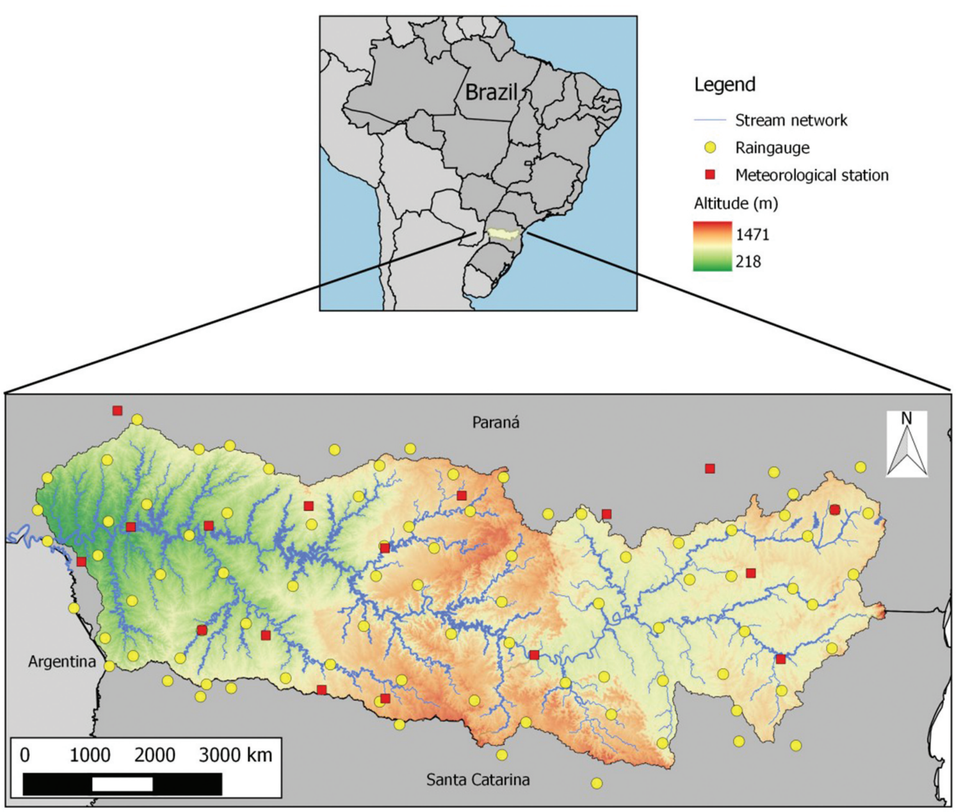

Figure 1. Location of the study area: Iguaçu River Basin.

2.3 SWAT model

The Soil and Water Assessment Tool (SWAT) model was selected because of its popularity among the scientific community, provided GIS interface (Dile *et al.* 2016) and open-source code. Research using SWAT has been largely reported in peer-reviewed papers, with relevant applications on streamflow calibration and calibration techniques (Gassman *et al.* 2007).

The SWAT model is a continuous-time, physically based, semi-distributed river basin model that operates commonly on a daily time step (Arnold *et al.* 2012). A brief description of SWAT methods is given below. The reader is referred to the manual (Neitsch *et al.* 2011) for further details.

Rainfall water can follow different paths in SWAT simulation. Rainfall can be intercepted and remain on canopy storage until it evaporates or falls to the soil surface. Water on the soil surface can infiltrate or flow overland as runoff, and reach a channel. The infiltrated water may be removed by evapotranspiration or flow to the surface water system via underground paths (Neitsch *et al.* 2011).

The land phase of the hydrological cycle is based on the following equation:

$$\mathrm{SW}_t = \mathrm{SW}_0 + \sum_{i=1}^{t} \left(R - Q_{\mathrm{surf}} - E_{\mathrm{a}} - w_{\mathrm{seep}} - Q_{\mathrm{gw}} \right) \quad (1)$$

where SW is the soil water content (mm), t is the time (days), R is the amount of precipitation (mm), Q_{surf} is the amount of surface runoff (mm), E_{a} is the amount of ET (mm), w_{seep} is the amount of water percolating to the shallow aquifer (mm), and Q_{gw} is the amount of return flow (mm) (Neitsch *et al.* 2011).

2.4 Model set-up and spatial discretization

The SWAT model for the studied basin was built within the QGIS interface (QSWAT 1.3) that uses a 2012 version of SWAT (Dile *et al.* 2016). In the spatial discretization of the modeled area in SWAT, the basin is first subdivided into sub-basins, which are divided into hydrological response units (HRUs). The sub-basins are generated according to the user-defined minimum contribution area for stream generation and topography and are divided into multiple HRUs consisting of the same land use, soil type and slope band each (Neitsch *et al.* 2011). Due to the large-modelled area, the minimum contribution area for stream formation and sub-basin division was set to 100 km^2, creating 365 sub-basins and 9741 HRUs. The modeled area comprises only the Brazilian territory of the basin and, therefore, the total modeled area is 63 884 km^2 (94.7% of the entire Iguaçu River Basin).

Table 1. Summary of data used for this study.

	Source	Resolution
Rainfall data	Hidroweb ANA[a]	Daily
Soil database	EMBRAPA[b]	
Digital Elevation Model (DEM)	EMBRAPA[b]	90 m
Meteorological data	EPAGRI[c] INMET[d] IAPAR[e] SIMEPAR[f]	Daily
Land cover map	Derived from Landsat 8 imagery	30 m
Streamflow data	Hidroweb ANAa Águas Paraná[g]	Daily
Evapotranspiration	METRIC from Uda (2016)	Monthly

[a]National Water Agency (ANA), Brazil.
[b]EMBRAPA (Miranda 2005).
[c]EPAGRI, Santa Catarina, Brazil.
[d]National Meteorological Institute (INMET), Brazil.
[e]Paraná Agronomic Institute (IAPAR), Paraná, Brazil.
[f]Paraná Meteorological System (SIMEPAR), Paraná, Brazil.
[g]Paraná Water Institute (Instituto das Águas do Paraná), Paraná, Brazil.

Despite being a large and complex basin, anthropic interventions were not included in the simulation due to the lack of reliable information and the higher complexity that would have been inserted in the model. The choice of a simpler model construction is based on the main intention of the study, which is to address changes in streamflow simulation uncertainty due to different calibration methods. The added complexity of considering these interventions would have had a very limited contribution to the main objective of the study.

2.5 Data

Table 1 summarizes the data and corresponding sources used in this study. Daily rainfall data from 90 raingauges (Fig. 1) were obtained for the period 2000–2009. Missing rainfall data were filled by inverse distance weighted interpolation with observations from the nearest raingauges. For the same time

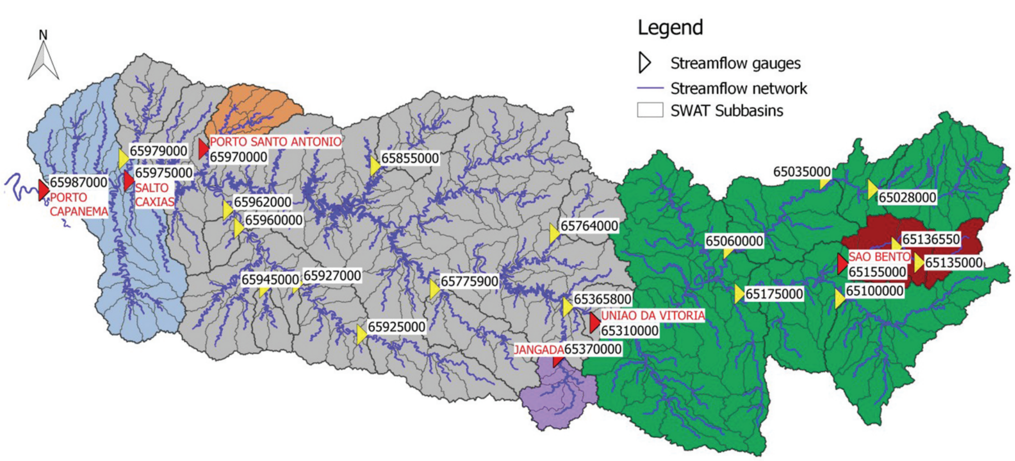

Figure 2. SWAT sub-basins and streamflow gauges. Sub-basins with red streamflow gauges were selected for posterior analysis.

interval, daily meteorological data (maximum and minimum daily temperature, solar radiation, average wind speed and average relative humidity) were obtained for 17 meteorological stations (Fig. 1). For model calibration and validation, daily measured streamflow data from 2000 to 2009 were collected from 23 streamflow gauges (Fig. 2). Evapotranspiration time series for calibration were generated from the average values of the pixels located inside each sub-basin delineated by SWAT.

The Brazilian Electric System National Operator maintains a regularly updated database with calculated naturalized streamflow for some hydropower plants in the Brazilian territory. Naturalized daily streamflow values are calculated by removing the effect of the upstream reservoir operation and the flow rate corresponding to the reservoir evaporation and consumptive water use, thereby reconstructing the natural streamflow regime (ONS 2007, 2016). For the Iguaçu River Basin, these data are available for the Salto Caxias hydropower plant reservoir, located 80 km upstream from the main outlet. The naturalized streamflow data were used only for daily validation of model results at the Salto Caxias location.

The basin soil type database was composed from Paraná (Bhering 2007) and Santa Catarina (Potter *et al.* 2004) soil type maps. The SWAT soil database was constructed with albedo from Post *et al.* (2000) and pedotransfer functions applied to each soil layer and soil type.

Table 2. Calibrated parameters and initial ranges.

Parameter name	Description	Adjustment	File	Min.	Max.
CN2	Initial soil conservation service runoff (Curve Number)	R	mgt	−0.30	0.30
SOL_AWC	Available soil water capacity (mm)	R	sol	−0.30	0.30
ALPHA_BF	Baseflow alpha factor	V	gw	0	1
GW_DELAY	Groundwater delay (days)	V	gw	0	500
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	V	gw	0	5000
LAT_TTIME	Lateral flow travel time (days)	V	hru	0	180
CH_K1	Effective hydraulic conductivity in the tributary channel (mm/h)	V	sub	0	300
CH_N1	Manning's *n* value for the secondary channel	V	sub	0.01	0.5
RCHRG_DP	Deep aquifer percolation fraction	V	gw	0	1
SLSOIL	Slope length for lateral subsurface flow (m)	V	hru	0	150
CANMX	Canopy storage capacity	V	hru	0	10
DEP_IMP	Depth to impervious layer in soil profile (mm)	V	hru	0	6000

V: existing parameter value is to be replaced by a given value,
R: existing parameter value is multiplied by (1 + a given value)

2.6 Model calibration and validation

Calibration and validation periods were established following the Klemeš (1986) split-sample approach, dividing the period 2002–2009 into two equal parts for calibration (2006–2009) and validation (2002–2005). Calibration was conducted only at a monthly time step. The calibration period comprises an extremely dry year, 2006 (average rainfall: 1318 mm/year), and a relatively rainy year, 2009 (average rainfall: 2004 mm/year). Validation was performed only for streamflow, since no ET data were available for the validation period. Daily streamflow validation was conducted separately for the calibration (2006–2009) and validation (2002–2005) periods, to assess parameter timescale transferability. All calibration and validation periods were preceded by 2 years of warm-up.

Sensitivity analysis, calibration, and validation were carried out within SWAT-CUP (calibration and uncertainty program), using the SUFI-2 (sequential uncertainty fitting) procedure. Global sensitivity analysis was performed with parameters chosen based on a review of the existing literature (Yang *et al.* 2007, Betrie *et al.* 2011, Githui *et al.* 2012, Zhang *et al.* 2015). The relative sensitivity of parameters is dependent on the variables included in the objective function and the time step considered. For the sensitivity analysis, streamflow and evapotranspiration were included in the objective function one at a time, using monthly data from the period 2006–2009. Both multi-site (average of all stations) and single-site (outlet only) conditions were investigated during the sensitivity analysis. The global sensitivity analysis is performed by regressing the Latin hypercube generated parameters values against objective function values (Abbaspour 2015). From this initial sensitivity analysis, parameters to which evapotranspiration and streamflow were more sensitive were included in model calibration, resulting in a total of 12 calibrated parameters. Evapotranspiration was sensitive to fewer parameters than streamflow (results not shown). Evapotranspiration was most sensitive to SOL_AWC, SLSOIL, and CN2, while changes in streamflow were more pronounced due to changes in ALPHA_BF, DEP_IMP, and CH_K1. Calibrated parameters and their initial range are listed in Table 2. Different processes of the hydrological cycle and channel routing phase are affected by the chosen parameters. A more detailed description of each parameter can be found in the SWAT manual. Sensitivity analysis was also conducted using daily discharge data. At the daily scale, streamflow was more sensitive to nine of the 12 parameters listed in Table 2.

2.6.1 SUFI-2 procedure

Calibration and validation were implemented on SWAT-CUP, using the sequential uncertainty fitting algorithm – version 2 (SUFI-2), an inverse modeling approach for combined optimization and uncertainty analysis (Abbaspour *et al.* 2004). The SUFI-2 algorithm was selected to calibrate the SWAT due to the large simulated area (63 884 km^2), which requires a large amount of processing time. Comparison studies reported that, from the algorithms available in SWAT-CUP, the SUFI-2 program is the least computationally expensive to run (Yang *et al.* 2008).

In the SUFI-2 algorithm, the parameter value ranges are reduced with successive iterations. To calibrate a model with SUFI-2, all iterations must be carried out with the same number of simulations. For each iteration, the parameters are homogenously sampled using Latin Hypercube according to the number of simulations to be carried out. Parameter ranges are reduced at each iteration, always centered on the parameter set that produced the best objective function value (Abbaspour *et al.* 2004, Abbaspour 2015). All calibration strategies were conducted with two iterations, and each iteration consisted of 500 simulations. Validation was conducted with one iteration consisting of 500 simulations.

As stated by Abbaspour *et al.* (2004), the objective of SUFI-2 is to find a "best range" of values for each calibrated parameter, and parameter set combinations inside this "best range" should ensure high-quality simulations. Although indicated in the calibration results, the SUFI-2 objective it is not to find only one set of parameter values. The final calibrated parameter ranges must be assessed for the validation periods with the same number of simulations as the calibration iterations. Therefore, 500 parameter sets sampled from the final calibrated ranges were assessed in terms of model goodness-of-fit statistics (KGE, NS and PBIAS). For each calibration strategy this resulted in 500 performance values of each statistic. Kundu *et al.* (2016) refers to this simulation set as the "best solution set". The "best simulation" is the simulation with one single set of parameters that yielded the best objective function value. The "best simulation" performance is also presented and discussed. In this study, the simulation achieving the highest average objective function value, considering all 23 streamflow gauges, is referred to as the "best simulation".

2.6.2 Objective function

The Kling-Gupta efficiency (KGE) (Gupta *et al.* 2009) objective function was used to calibrate the SWAT model. The KGE is based on the decomposition of the NS and the mean square error (MSE) into three components: correlation, variability error, and bias error. A three-dimensional criteria space is created, where the KGE values are calculated in terms of the Euclidian distance from the ideal point and the pareto three-dimensional surface:

$$\mathrm{KGE} = 1 - \mathrm{ED} \tag{2}$$

with

$$\mathrm{ED} = \sqrt{(r-1)^2 + (\beta-1)^2 + (\alpha-1)^2} \tag{3}$$

$$\alpha = {}^{\mu_s}/_{\mu_o} \tag{4}$$

$$\beta = {}^{\sigma_s}/_{\sigma_o} \tag{5}$$

where r is the linear correlation coefficient; α is the ratio between the mean simulated (μ_s) and mean observed flow (μ_0), and represents the bias; β is a measure of relative variability and is the ratio between simulated (σ_s) and observed (σ_0) standard deviation. KGE values range from $-\infty$ to 1, where 1 represents the ideal value. An interesting feature of the KGE is that it reflects the lower limit of its three components, i.e., the worst value between r, β and α is higher than or equal to the KGE value (Piniewski *et al.* 2017).

In case of multiple variables, the objective function is defined as

$$\mathrm{KGE}' = \sum_j w_j \mathrm{KGE}_j \tag{6}$$

where w_j is the weight of the jth variable. In this study, no distinction was made between the outlet and tributaries streamflow stations, therefore, the same weight was assigned to all 23 streamflow stations. Evapotranspiration calibration was carried out using sub-basin monthly evapotranspiration. As in the multi-site calibration, the objective function was calculated by attributing equal weights to the objective function computed for each sub-basin.

Streamflow is more sensitive than evapotranspiration to the parameters selected for calibration. For this reason, calibration iterations did not consider both variables at the same time to calculate the objective function value. Section 2.6.4 ("Calibration strategies") indicates the scheme adopted in the multi-variable calibration approach.

Since KGE is not a widely used indicator for assessing streamflow simulation performance, Nash-Sutcliffe (NS) efficiency coefficient and percent bias (PBIAS) are also reported. According to Moriasi *et al.* (2007), NS values above 0.50 are satisfactory for monthly flow, values from 0.65 to 0.75 are classified as "good" and values above 0.75 are "very good". Similarly, PBIAS ≤ ±25% are classified as acceptable, and PBIAS ≤ ±10% are "very good". These ratings, however, are not strict and the objective of the study and site conditions, as well as the modeling size area, must be taken into account when model performance is evaluated. For the KGE function, the value of 0.50, similar to the NS value suggested by Moriasi *et al.* (2007), was defined as acceptable and therefore simulations exhibiting KGE ≥ 0.50 were classified as behavioral. This was a simplified decision and we highlight that NS and KGE values are not directly comparable (Knoben *et al.* 2019). For example, the KGE mean flow benchmark (KGE = − 0.41) differs from the NS benchmark (NS = 0) (Knoben *et al.* 2019).

2.6.3 Uncertainty analysis

Uncertainty in SUFI-2 is expressed by the 95% probability distribution calculated from the 2.5% and 97.5% percentiles of the accumulated probability distribution of the output variable of interest. This result is presented as the 95% Probability Uncertainty (95PPU) "envelope" of solutions generated by the parameters value ranges. The *p* factor and *r* factor are statistical indicators used to quantify the output variable uncertainty and the agreement between the envelope of solutions and the measured data. The percentage of measured data bracket by the 95PPU envelope is quantified by the *p* factor. Uncertainty width is quantified by the *r* factor, which represents the thickness of the 95PPU envelope. The *r* factor is calculated as the ratio of the average thickness of the 95PPU band by the standard deviation of measured data (Abbaspour *et al.* 2004, , Abbaspour 2015). Abbaspour *et al.* (2015) recommended values of *p* factor ≥ 0.70 and *r* factor ≤ 1.50 for considering the calibration results as acceptable.

2.6.4 Calibration strategies

Three different calibration strategies were investigated (Fig. 3). In strategies E1 and E2, only streamflow data were used to calibrate the SWAT model. Strategy E3 included both

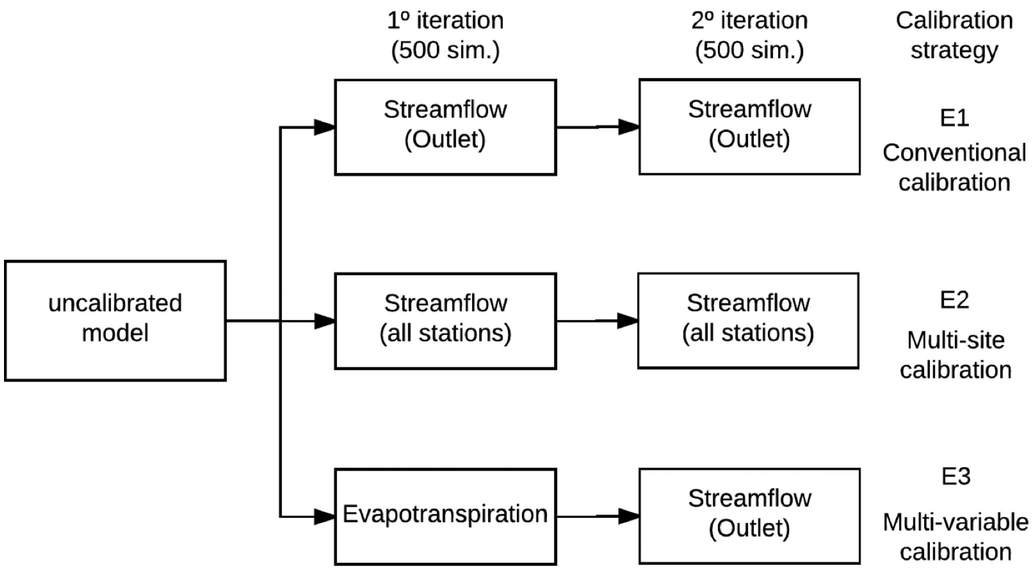

Figure 3. Scheme indicating the applied calibration strategies according to the variable included in the objective function of each iteration.

evapotranspiration and streamflow data. Strategy E1 considers only outlet streamflow in both iterations and is the benchmark for the other considered calibration approaches. In strategy E2, streamflow from all stations were used in both iterations, consisting of the multi-site calibration approach. Finally, strategy E3 was conducted using evapotranspiration in the first iteration and outlet streamflow in the second iteration, resulting in a multi-variable calibration approach. As stated in section 2.6.1 ("SUFI-2 procedure"), all calibration strategies were conducted with 2 iterations, each iteration consisting of 500 simulations. Therefore, in strategy E3 only one iteration is performed with each variable. This calibration strategy corresponds to two calibrations of one iteration each (each single-variable), with the exception that the initial parameter bounds changes from one calibration to the other. This was done to allow a fair comparison between all strategies. While the other two strategies have two iterations with the same variable, it would not be a fair comparison if we had used a higher number of total iterations in the third strategy. It would be desirable to get acceptable performance for evapotranspiration before switching the calibration to streamflow. However, we opted to perform only one iteration for each variable in order to keep this calibration strategy comparable with the other two (same total number of iterations). Since evapotranspiration was sensitive to fewer parameters, this variable was used in the first iteration, as an initial constraint of the parameter space, followed by calibration with outlet streamflow in the second iteration. Figure 3 indicates the calibration strategies according to the variable included in the objective function for each iteration.

3 Results

Model temporal and spatial validation were carried out with all the 23 streamflow gauges. Due to the high number of gauges, only the average value of all (including the outlet) and six different locations are presented hereafter (indicated in Fig. 2):

- Porto Capanema (code: 65,987,000), at the basin outlet (drainage area: 63 884 km^2);
- União da Vitória (code: 65,310,000), in the middle of the basin, which drains 38% of the total area (drainage area: 24 209 km^2);
- São Bento (code: 65,155,000), which constitutes a representative behavior of the headwater sub-basins, and exhibits the worst performance of all locations (drainage area: 2006 km^2);
- Jangada (code: 65,370,000), a small sub-basin with average acceptable performance (drainage area: 1014 km^2);
- Salto Caxias (code: 65,975,000), for daily validation only, streamflow gauge at the reservoir of the Salto Caxias hydropower plant where naturalized daily streamflow is available (drainage area: 57 060 km^2);

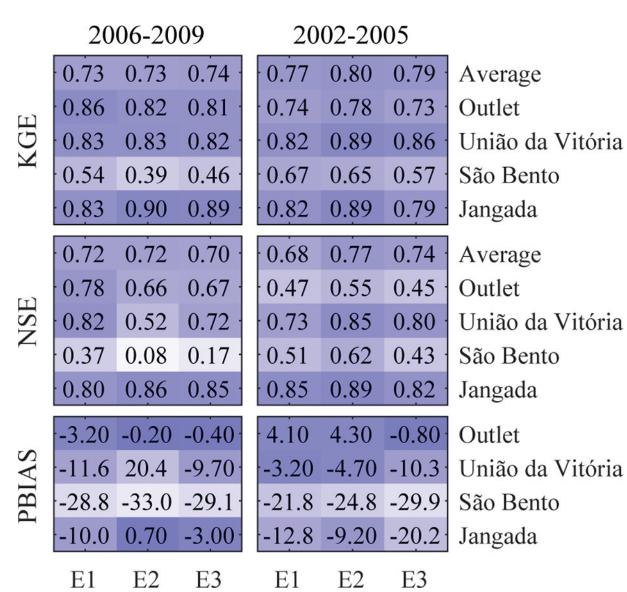

Figure 4. Goodness-of-fit statistics for the simulation leading to the best average KGE value, at monthly time-step, for calibration (2006–2009) and validation (2002–2005) periods.

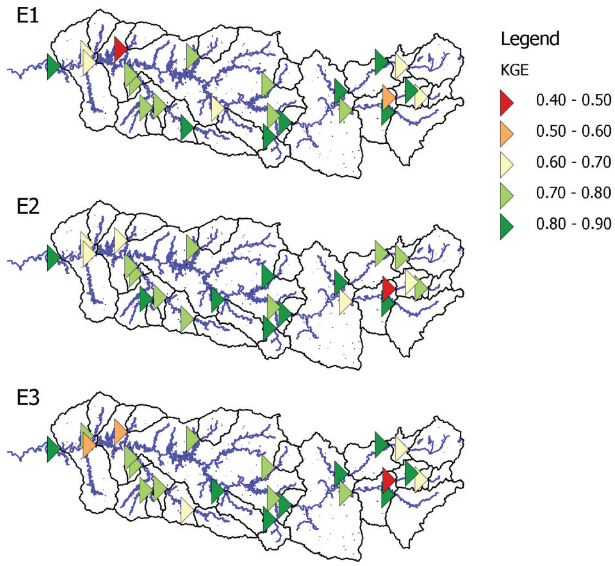

Figure 5. Best simulation performance for the monthly simulation of 2006–2009.

- Porto Santo Antônio (code: 65,970,000), for daily validation only, because it is a sub-basin exhibiting a distinct streamflow behaviour (drainage area: 1084 km^2).

3.1 Model performance at monthly time step

Figure 4 presents the goodness-of-fit statistics for the monthly flow simulations, calibrated using all strategies. Figure 5 indicates the "best simulation" results, which is the parameter set that resulted in the highest average KGE value of all locations. The average and main outlet KGE values are all above 0.73, reflecting a good overall simulation performance for all calibration strategies. At the main outlet, KGE values ranged from 0.73 to 0.86 among all strategies for both periods (2006–2009 and 2002–2005). Considering the average value of all gauging locations, KGE ranged from 0.73 to 0.80. According to the classification of Moriasi *et al.* (2007), NS values for the main outlet are "good" (0.65–0.75) for all three strategies on the calibration period (2006–2009). For the validation period (-2002–2005), the main outlet NS is acceptable (>0.50) only for the multi-site (E2) strategy. The main outlet PBIAS values are very good (≤±10%) for both calibration and validation periods. Considering the average value of all locations, NS values are "good" or "very good", ranging from 0.68 to 0.77.

The results suggest that conventional calibration (E1) can produce satisfactory performance for the inner sub-basins for monthly simulation (Fig. 5). The multi-site (E2) and multi-variable (E3) calibrations did not improve significantly the inner sub-basins simulation performance (Figs. 4 and 5). The average performance of the multi-variable calibration strategy (E3) was higher than the average performance of conventional calibration, but the main outlet performance was slightly depreciated. Considering the calibration period, the conventional strategy (E1) resulted in the highest values for NS and KGE at the main outlet. This result was expected since in strategy E1 both calibration iterations were performed using the main outlet data only. However, for the validation period, the best outlet performance was achieved by the multi-site strategy (E2), which also presents the best average performance within the three strategies.

Graphical analysis of the hydrograph (see the Supplementary Material, Fig. S1) and flow duration curves (FDC) (Fig. S2) reveals an overestimation of low flows for all strategies and locations, except for the outlet in E2. The KGE is influenced by the PBIAS, which are an average result of over and under prediction of flows from different magnitudes, and low absolute PBIAS can result from large compensations of under and overpredictions. The FDC, on the other hand, illustrates the over and underpredictions that occur at all flow magnitudes.

Monthly flow simulation for the year 2006, an exceptionally dry year in Iguaçu river basin, was satisfactory at the four locations and using all strategies, as it can be assessed by visual inspection of the hydrographs (Supplementary Material, Fig. S1). Satisfactory simulation for this extremely dry climate scenario suggests that the calibration method could be adequate for hydrological extreme climate scenarios.

3.2 Parameter temporal scale transferability: daily flow simulation performance

The goodness-of-fit statistics for the best simulation, at the daily time step, are presented in Fig. 6. A large dam is located at the Salto Caxias streamflow gauge where daily streamflow is controlled by the reservoir operation. The main outlet, Porto Capanema, is located 80 km downstream from Salto Caxias. As reservoir operations were not included in the simulation, daily streamflow simulated values at the main outlet are not expected to agree well with the available measurements, resulting in poor NS and KGE values. Therefore, the naturalized daily streamflow provided by the ONS for the Salto Caxias streamflow gauge was used to calculate the daily performance statistics. For the daily simulation analysis, the Salto Caxias location is indicated instead of the main outlet. Despite that,

		2006-2009			2002-2005		
		E1	E2	E3	E1	E2	E3
KGE	Average	0.62	0.67	0.64	0.67	0.68	0.63
	Salto Caxias	0.65	0.81	0.73	0.81	0.85	0.77
	União da Vitória	0.70	0.70	0.65	0.70	0.70	0.64
	São Bento	0.58	0.33	0.39	0.46	0.45	0.48
	Jangada	0.66	0.80	0.84	0.67	0.80	0.78
NSE	Average	0.60	0.49	0.43	0.47	0.47	0.40
	Salto Caxias	0.68	0.74	0.63	0.69	0.73	0.59
	União da Vitória	0.80	0.41	0.37	0.41	0.38	0.33
	São Bento	0.57	-0.24	-0.32	-0.12	-0.04	-0.15
	Jangada	0.60	0.76	0.70	0.73	0.80	0.68
PBIAS	Salto Caxias	-8.50	-5.70	-3.80	-11.9	-3.80	-3.40
	União da Vitória	-9.00	-10.8	-10.5	-10.5	-5.50	-6.30
	São Bento	-29.5	-33.9	-28.4	-28.5	-27.2	-24.9
	Jangada	-3.20	-2.20	-1.10	-29.7	-14.0	-15.7

Figure 6. Goodness-of-fit statistics for the best simulation, at the daily time-step, for calibration (2006–2009) and validation (2002–2005) periods.

the main outlet was included on the daily average performance statistics to allow comparison with the monthly values.

According to the objective function (KGE), daily simulation performance is reasonably good for the average, Salto Caxias, União da Vitória and Jangada locations (Fig. 6). São Bento daily simulation performance, however, is poor. Only Salto Caxias and Jangada locations achieved acceptable NS values (≥ 0.50) for all strategies in the validation period. Daily KGE (NS) at the Salto Caxias location ranged from 0.65 to 0.73 (0.63 to 0.68) for the calibration period and from 0.77 to 0.85 (0.59 to 0.73) for the validation period. The average KGE values are higher for the validation period than for the calibration period, indicating that there was no depreciation from the calibration to the validation period. Measured and simulated daily hydrographs for the period January 2005 to December 2006 are presented in the Supplementary material (Fig. S3).

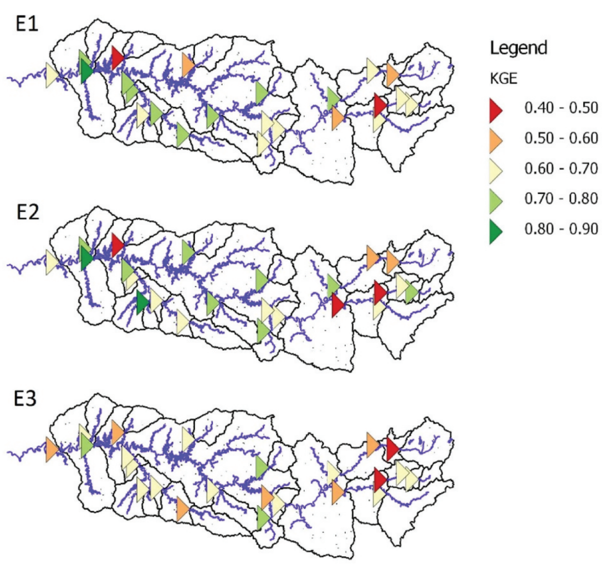

Figure 7. Best simulation performance for the daily simulation of 2002–2005.

Similar to the results obtained for monthly simulation, NS values are significantly smaller than KGE values, indicating that the flow peaks were poorly simulated despite the good average flow simulation performance. Is important to note that KGE was used as objective function to determine the "best simulation" for each location, having impacted NS values presented here. Better NS values could be obtained if this metric was chosen as the objective function.

All strategies and locations presented negative PBIAS values for the best simulation. Negative PBIAS values indicate streamflow overestimation, which occurs for low flows, as it can be visualized in the FDC (see the Supplementary Material, Fig. S4). The Salto Caxias daily FDCs indicate good agreement between the observed and simulated flows for all strategies. The negative PBIAS values for all strategies may suggest systematic modeling errors. One possible cause may be water withdrawal for agriculture or urban supply not accounted for in the present study.

The results for the best simulation at daily time step for the validation period are shown in Fig. 7. The overall good daily performance suggests that parameter scale transferability may

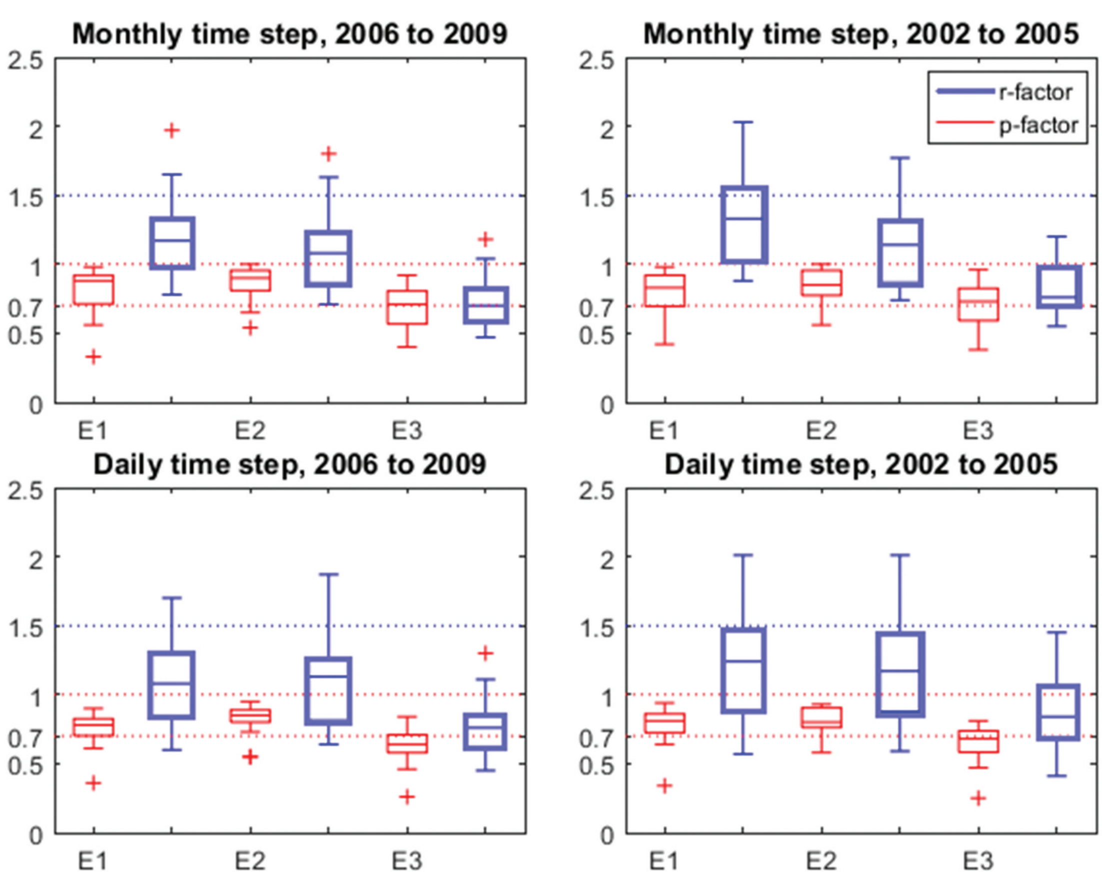

Figure 8. Comparison of p factor and r factor of all streamflow gauge locations by strategy and period.

be acceptable. Still, the poor performance for headwater sub-basins points out to the need of further studies. These headwater sub-basins are located upstream from União da Vitória, deteriorating its performance as well.

3.3 Predictive uncertainty

The quality of the predictive uncertainty was assessed by using two indicators: the *p* factor, which measures the percentage of observations bracket by the 95PPU envelope; and the *r* factor, which represents the thickness of the 95PPU envelope. Comparison between *p* and *r* factors of all 23 locations, for each strategy and time step, can be visualized in Fig. 8. A balance between *p* and *r* must be pursued, in order to capture most of the measured data within the 95PPU band with the smallest uncertainty possible.

The multi-site (E2) strategy resulted in the highest *p* factors, for both monthly and daily simulations. This result indicates that, while the multi-site calibration did not improve significantly the inner sub-basins simulation performance in terms of "best simulation" (as presented in Section 3.1), it provided a more reliable uncertainty.

The multi-variable calibration strategy (E3) reduced the uncertainty width when compared to the conventional calibration strategy (E1), for both monthly and daily time steps, which is indicated by the smaller *r* factors. However, this reduction in the uncertainty width was also followed by a reduction of *p* factor values. The reduced uncertainty for E3, compared to E1 and E2, is significant for monthly simulation, while still bracketing most of the observed data for, at least, half of the gauges.

3.4 Best solution set results

As indicated before, the objective of the SUFI-2 algorithm is not to define only one parameter set, but a "best range" for each calibrated parameter. Therefore, all parameter combinations inside the final ranges were assessed in terms of goodness-of-fit performance (KGE, NS and PBIAS) in order to assess the quality of the simulation for all possible calibrated parameter sets.

3.4.1 Kling-Gupta efficiency (KGE)

The range of KGE values obtained with the 500 simulations of the last iteration for each strategy is presented in Fig. 9. The high amplitude of KGE values reflects the large

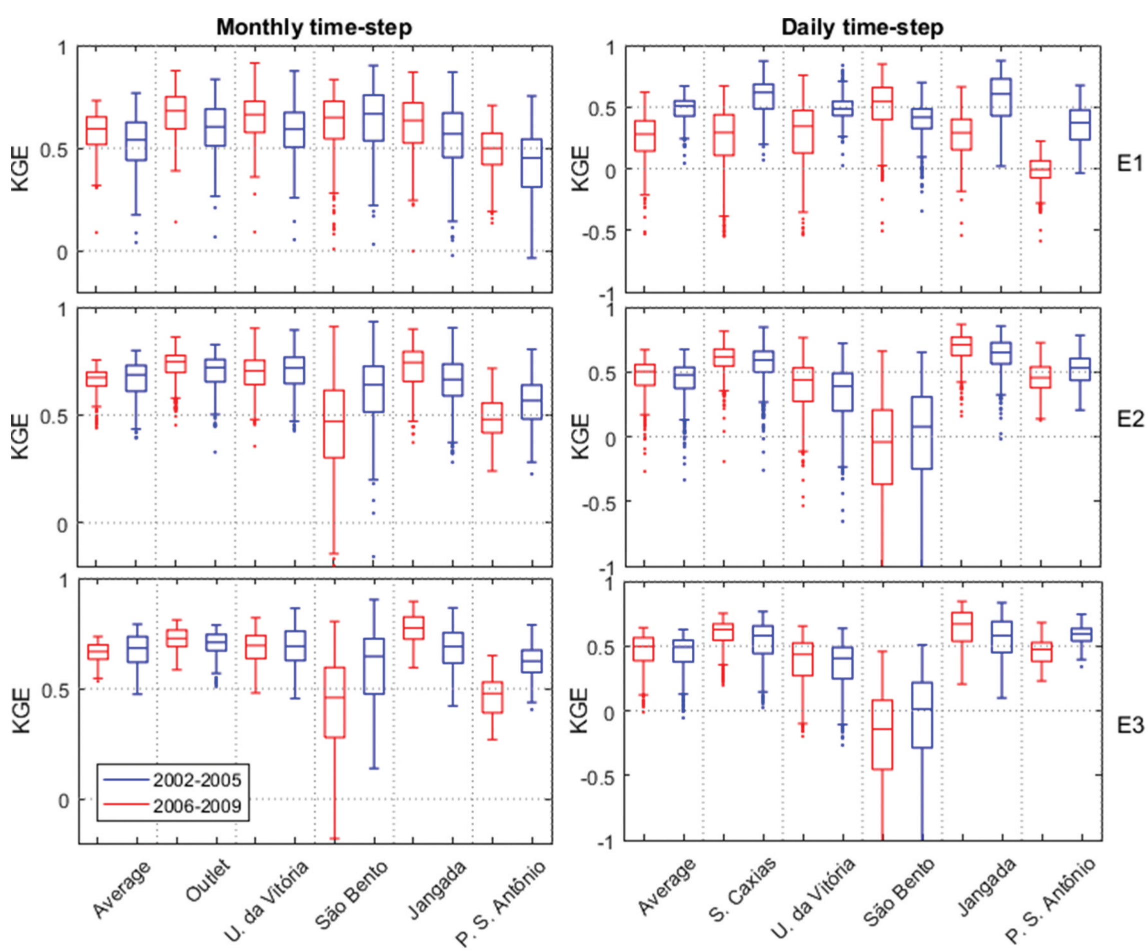

Figure 9. KGE values for the best solution set for daily and monthly simulations of calibration (2006–2009) and validation (2002–2005) periods.

uncertainty for the "best solution set" (Fig. 9). Monthly simulation shows positive KGE values for all strategies for all locations except for São Bento. São Bento has higher amplitude of KGE values and a few negative values. Gauges located on the headwater sub-basins, like São Bento, exhibit poor performance for daily simulation. The results shown in Fig. 9 suggest that daily simulation can be reasonably acceptable for the outlet and the average of all stations, but headwater sub-basins need further calibration to reach acceptable performance. Some of the calibrated parameters are related to the channel routing phase, which can be significantly different for the headwater sub-basins.

Simulations were considered as behavioral when KGE was equal to or greater than 0.50, for monthly and daily streamflow. Strategy E3 exhibits the highest number of behavioral simulations for the monthly simulation and for the daily simulation in the calibration period. For the daily simulation in the 2002–2005 period, the highest number of behavioral simulations is found for the conventional calibration (E1). The number of behavioral simulations for the average of all 23 stations is higher for monthly simulation than for the daily simulation. This result is expected, since the model was only calibrated at the monthly time step.

3.4.2 Nash-Sutcliffe efficiency (NS)

Monthly simulation in the calibration period shows good NS values for all stations, with a majority of positive and above 0.50 values, except for São Bento (Supplementary Material, Fig. S5). The NS values are significantly lower for the daily simulation, despite the acceptable NS values presented for the best simulation. This reveals insufficient performance of the monthly calibrated model to simulate daily flow peaks, especially for the headwater sub-basins.

4 Discussion

4.1 How well do multi-site and multi-variable calibration perform compared to single-site and single-variable calibration?

Single-site, single-variable calibration exhibited similar performance results than multi-site and multi-variable calibrations when only the performance of the "best simulation" is taken into account. However, the quality of the predictive uncertainty was higher in multi-site and multi-variable calibrations. Multi-site calibration provided a more reliable uncertainty for all 23 sub-basins than the single-site calibration. Multi-variable calibration resulted in a reduced uncertainty compared to single-site and multi-site calibrations, although also followed by reduction in the number of observations captured by the uncertainty bounds. Moreover, both multi-site and multi-variable calibrations resulted in a higher number of behavioral simulations from the best solution set (last 500 simulations).

The multi-site and multi-variable calibration strategies were especially useful for parameter transferability in time (i.e., daily time step simulation from monthly calibrated parameters). This approach can be used to work around the lack of data with the desirable time resolution when calibrating and validating hydrological models.

The performance and uncertainty results for the headwater sub-basins, which exhibited poorer performance and higher

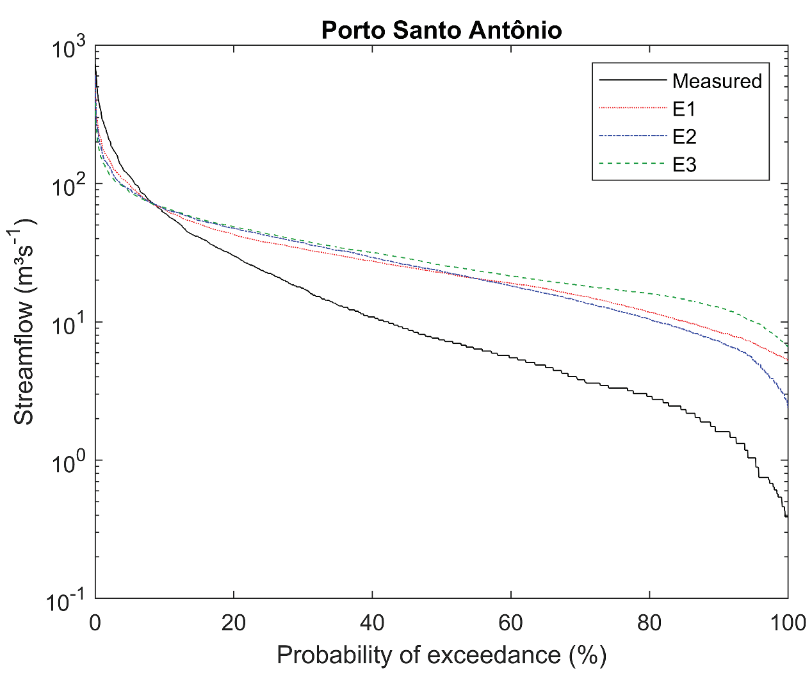

Figure 10. Flow duration curve for the Porto Santo Antônio daily flow.

uncertainty, is positively affected by the multi-site calibration strategy. This result reinforces that calibration and validation performance using the conventional single-site calibration approach can be strongly affected by basin characteristics.

4.2 Can we use a model calibrated at monthly time step to obtain simulations at daily time step?

Transferring parameters from monthly to daily time step resulted in streamflow simulation with satisfactory p and r factor values and acceptable performance regarding the objective function (KGE). However, daily simulation performances are inferior to the correspondent monthly simulation, especially for the headwater sub-basins. The reduced performance for the daily time step simulation may indicate that it can be related to the simulation of the peak flows. Daily simulation may be affected by the fact that peaks are more intimately related to hydraulic parameters that govern the routing phase, which were monthly calibrated and may not be adequate for daily simulation. This may also point out to the need for caution when evaluating results in different time steps, since the coarser time step (monthly) provided rather good results for the objective function despite the bad peak simulation performance in the headwater sub-basins.

4.3 On the use of KGE for model calibration

Porto Santo Antônio exhibits very good NS values, despite the substantial flow overestimation revealed by the PBIAS values and daily FDC (Fig. 10). The results for this location illustrate that a good NS value can be achieved while having significant overestimation of flows, which is reflected on high PBIAS values. This is explained by the fact that KGE tends to search for an overall performance, preserving the distribution of flows, and therefore, providing better agreement between measured and simulated FDC. Therefore, KGE value is negatively impacted by flow over or underestimations. On the other hand, NS criteria can lead to model simulations with large volume balance errors and underestimate flow variability (Gupta *et al.* 2009, Kling *et al.* 2012).

5 Conclusions

This study investigated the impact of multi-site and multi-variable SWAT model calibration on streamflow prediction and uncertainty for the Iguaçu River Basin and its inner sub-basins. Multi-variable calibration was carried out with remote sensed derived actual evapotranspiration and streamflow data. The results from multi-site and multi-variable calibration were compared with the conventional calibration, which uses only main outlet streamflow measurements. The impact of transferring parameters from monthly to daily simulation was also investigated.

The conventional calibration provided good streamflow performance for the main outlet and for the inner sub-basins. The multi-site and multi-variable calibrations did not improve significantly the performance for inner sub-basins. However, multi-site calibration provided a more reliable uncertainty for all 23 sub-basins than the single-site calibration and multi-variable calibration resulted in a reduced uncertainty compared to single-site and multi-site calibrations.

Transferring parameters from monthly to daily time step resulted in streamflow simulation with satisfactory p and r factor values and acceptable performance regarding the objective function (KGE). These results point out to the possibility of transferring parameters calibrated on a coarser time step to a finer time step. However, daily peak flow simulation (measured by the NS) was unsatisfactory, notably for the headwater sub-basins. These sub-basins exhibited poor performance and high uncertainty, even when a multi-site strategy was used. Calibration with daily streamflow measurements may help improve our understanding of the parameter transferability efficiency by comparing the results with those presented herein.

The results presented in this paper indicate that the incorporation of additional information in model calibration may be useful to improve the uncertainty estimate in streamflow simulation. This is true even with high uncertainties of the streamflow and evapotranspiration data.

The poor performance for some locations may be due to the presence of many hydropower plants reservoirs in the Iguaçu River Basin. Due to these reservoirs, some streamflow gauges data reflect the dam operation instead of the natural river behavior. We recommend further studies to implement these reservoirs operations in hydrological modeling of this basin.

We believe that the implementation of the multi-variable calibration strategy using all calibration variables simultaneously with different objective function weights could lead to new results, enriching the discussion about the use of multiple variables for model calibration. The proper selection of weights could be done by evaluating the use of different combinations of weights for the different variables and its impacts on model performance (e.g., Rientjes *et al.* 2013). Weights could also be assigned based on the sensitivity of the calibrated variables. Recommendations for future studies also include calibration with evapotranspiration data only and analysis of different model output variables, other than streamflow. The use of soil moisture remote-sensed estimations in multi-variable calibration are also encouraged for future studies.

Acknowledgements

The authors wish to thank CNPq and CAPES for providing scholarships for the first and second authors and the Federal University of Santa Catarina (UFSC) for the financial support of the research activities. The authors also would like to thank EPAGRI, COPEL, IAPAR, and INMET for kindly providing the data used in this study. We would also like to thank the Associate Editor and two anonymous reviewers for providing valuable insights and suggesting relevant and necessary improvements to our paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Ana Clara Lazzari Franco http://orcid.org/0000-0003-2489-2878
Debora Yumi de Oliveira http://orcid.org/0000-0003-3635-3249

References

Abbaspour, K.C., *et al.*, 2015. A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology*, 524, 733–752. doi:10.1016/j.jhydrol.2015.03.027

Abbaspour, K.C., 2015. SWAT-CUP: SWAT Calibration and Uncertainty Programs – A User Manual. Eawag: Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.

Abbaspour, K.C., Johnson, C.A., and van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone Journal*, 3 (4), 1340. doi:10.2136/vzj2004.1340

Allen, R.G., *et al.*, 2002. SEBAL (Surface energy balance algorithms for land) – advanced training and users manual - Idaho implementation. Version 1. Idaho: Waters Consulting, University of Idaho, WaterWatch, Inc.

Allen, R.G., *et al.*, 2005. Satellite-based evapotranspiration by energy balance for western states water management. In: EWRI 2005: Impacts of global climate change. Proc. World Water and Environ. Resour. Congr., Anchorage, AL. 15–19 May 2005. Am. Soc. Civ. Eng., New York. doi:10.1061/40792(173)556

Allen, R.G., Tasumi, M., and Trezza, R., 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC)—model. *Journal of Irrigation and Drainage Engineering*, 133 (4), 380–394. doi:10.1061/(ASCE)0733-9437-(2007)133:4(380)

Arnold, J.G., *et al.*, 2012. Swat: model use, calibration, and validation. *Transactions of the ASABE*, 55 (4), 1491–1508. doi:10.13031/2013.42256

Bastiaanssen, W.G.M., *et al.*, 1998. A remote sensing surface energy balance algorithm for land (SEBAL). *Journal of Hydrology*, 212 (213), 213–229. doi:10.1016/S0022-1694(98)00254-6

Bastiaanssen, W.G.M., *et al.*, 2005. SEBAL model with remotely sensed data to improve water-resources management under actual field conditions. *Journal of Irrigation and Drainage Engineering*, 131 (February), 2. doi:10.1061/(ASCE)0733-9437-(2005)131

Bergstrom, S., Lindstrm, G., and Pettersson, A., 2002. Multi-variable parameter estimation to increase confidence in hydrological modelling. *Hydrological Processes*, 16 (2), 413–421. doi:10.1002/hyp.332

Betrie, G.D., *et al.*, 2011. Sediment management modelling in the Blue Nile Basin using SWAT model. *Hydrology and Earth System Sciences*, 15 (3), 807–818. doi:10.5194/hess-15-807-2011

Beven, K. and Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6 (January), 279–298. doi:10.1002/hyp.3360060305

Beven, K.J., 2001. Calibration, validation and equifinality in hydrological modelling: a continuing discussion. *In*: *Model validation: perspectives in hydrological science*. Wiley, Chichester. p. 44–55.

Bhering, S. B., 2007. *Mapa de solos do estado do Paraná. EMBRAPA*.Rio de Janeiro: Embrapa Solos

Blöschl, G., *et al.*, 2019. Twenty-three unsolved problems in hydrology (UPH) – a community perspective. *Hydrological Sciences Journal*. doi:10.1080/02626667.2019.1620507

Briak, H., *et al.*, 2019. International soil and water conservation research use of a calibrated SWAT model to evaluate the effects of agricultural BMPs on sediments of the Kalaya river basin (North of Morocco). *International Soil and Water Conservation Research*. doi:10.1016/j.iswcr.2019.02.002

Brighenti, T.M., *et al.*, 2019. Two calibration methods for modeling stream fl ow and suspended sediment with the swat model. Ecological Engineering, 127, 103–113. doi:10.1016/j.ecoleng.2018.11.007

Cao, W., *et al.*, 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrological Processes*, 20 (5), 1057–1073. doi:10.1002/hyp.5933

Chaney, N.W., *et al.*, 2015. Flood and drought hydrologic monitoring: the role of model parameter uncertainty. *Hydrology and Earth System Sciences*, 19 (7), 3239–3251. doi:10.5194/hess-19-3239-2015

Chiang, L.C., *et al.*, 2014. Assessing SWAT's performance in the Kaskaskia River watershed as influenced by the number of calibration stations used. *Hydrological Processes*, 28 (3), 676–687. doi:10.1002/hyp.9589

Daggupati, P., *et al.*, 2015a. A recommended calibration and validation strategy for hydrologic and water quality models. *Transactions of the ASABE*, 58 (6), 1705–1719. doi:10.13031/trans.58.10712

Daggupati, P., *et al.*, 2015b. Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. *Hydrological Processes*, 29 (26), 5307–5320. doi:10.1002/hyp.10536

Devia, G.K., Ganasri, B.P., and Dwarakish, G.S., 2015. A review on hydrological models. *Aquatic Procedia*, 4, 1001–1007. doi:10.1016/j.aqpro.2015.02.126

Dile, Y.T., *et al.*, 2016. Introducing a new open source GIS user interface for the SWAT model. *Environmental Modelling and Software*, 85, 129–138. doi:10.1016/j.envsoft.2016.08.004

Fenicia, F., *et al.*, 2007. A comparison of alternative multiobjective calibration strategies for hydrological modeling. *Water Resources Research*, 43, W03434. doi:10.1029/2006WR005098

Feyen, L., Kalas, M., and Vrugt, J. A., 2008. Semi-distributed parameter optimization and uncertainty assessment for large-scale streamflow simulation using global optimization. *Hydrological Sciences Journal*, 53 (2), 293–308. doi:10.1623/hysj.53.2.293

Franco, A.C.L. and Bonumá, N.B., 2017. Multi-variable SWAT model calibration with remotely sensed evapotranspiration and observed flow. *Revista Brasileira de Recursos Hidricos*. 22. doi:10.1590/2318-0331.011716090

Gassman, P.W., *et al.*, 2007. The soil and water assessment tool: historical development, applications, and future research directions. *Transactions of the Asabe*, 50 (4), 1211–1250. doi:10.1.1.88.6554

Githui, F., Selle, B., and Thayalakumaran, T., 2012. Recharge estimation using remotely sensed evapotranspiration in an irrigated catchment in southeast Australia. *Hydrological Processes*, 26 (9), 1379–1389. doi:10.1002/hyp.8274

Gupta, H.V., *et al.*, 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *Journal of Hydrology*, 377 (1–2), 80–91. doi:10.1016/j.jhydrol.2009.08.003

Her, Y. and Chaubey, I., 2015. Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty. *Hydrological Processes*, 29 (19), 4220–4237. doi:10.1002/hyp.10487

Immerzeel, W.W. and Droogers, P., 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology*, 349 (3–4), 411–424. doi:10.1016/j.jhydrol.2007.11.017

Khakbaz, B., *et al.*, 2012. From lumped to distributed via semi-distributed: calibration strategies for semi-distributed hydrologic models. *Journal of Hydrology*, 418–419, 61–77. doi:10.1016/j.jhydrol.2009.02.021

Klemeš, V., 1986. Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31 (1), 13–24. doi:10.1080/02626668609491024

Kling, H., Fuchs, M., and Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, 424–425, 264–277. doi:10.1016/j.jhydrol.2012.01.011

Knoben, W.J.M., Freer, J.E., and Woods, R.A., 2019. Technical note: inherent benchmark or not? Comparing Nash–Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23:10, 4323–4331. doi:10.5194/hess-23-4323-2019

Kundu, D., van Ogtrop, F.F., and Vervoort, R.W., 2016. Identifying model consistency through stepwise calibration to capture streamflow variability. *Environmental Modelling & Software*, 84, 1–17. doi:10.1016/j.envsoft.2016.06.013

Kunnath-poovakka, A., *et al.*, 2016. The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction. *Journal of Hydrology*, 535, 509–524. doi:10.1016/j.jhydrol.2016.02.018

Kunnath-poovakka, A., *et al.*, 2018. Remotely sensed ET for streamflow modelling in catchments with contrasting flow characteristics: an attempt to improve efficiency. *Stochastic Environmental Research and Risk Assessment*, 32 (7), 1973–1992. doi:10.1007/s00477-018-1528-y

Lavenne, A., *et al.*, 2019. A regularization approach to improve the sequential calibration of a semidistributed hydrological model. *Water Resources Research*. 8. doi:10.1029/2018WR024266

Lerat, J., *et al.*, 2012. Do internal flow measurements improve the calibration of rainfall-runoff models? *Water Resources Research*, 48 (2), 1–18. doi:10.1029/2010WR010179

Leta, O.T., van Griensven, A., and Bauwens, W., 2017. Effect of single and multisite calibration techniques on the parameter estimation, performance, and output of a SWAT model of a spatially heterogeneous catchment. *Journal of Hydrologic Engineering*, 22 (3), 5016036. doi:10.1061/(ASCE)HE.1943-5584.0001471

Liu, Y. and Gupta, H.V., 2007. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework. *Water Resources Research*, 43 (7), 1–18. doi:10.1029/2006WR005756

Machado, B.P., 2012. Large hydropower plants of Brazil. *In*: *Comprehensive renewable energy*. Elsevier, 93–127. doi:10.1016/B978-0-08-087872-0.00607-7

Melsen, L., *et al.*, 2016. Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin. *Hydrology and Earth System Sciences*, 20 (6), 2207–2226. doi:10.5194/hess-20-2207-2016

Miranda, E.E.D. 2005. Brasil em relevo. [accessed 24 Aug 2015]. Available from: http://www.relevobr.cnpm.embrapa.br.

Moriasi, D.N., *et al.*, 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50 (3), 885–900. doi:10.13031/2013.23153

Moussa, R., Chahinian, N., and Bocquillon, C., 2007. Distributed hydrological modelling of a Mediterranean mountainous catchment - Model construction and multi-site validation. *Journal of Hydrology*, 337 (1–2), 35–51. doi:10.1016/j.jhydrol.2007.01.028

Neitsch, S.L., *et al.*, 2011. Soil and Water Assessment Tool: Theoretical Documentation, Version 2009. Texas Water Resources Institute Technical Report No. 406. Texas, USA: Texas A&M University.

Nkiaka, E., Nawaz, N.R., and Lovett, J.C., 2018. Effect of single and multi-site calibration techniques on hydrological model performance, parameter estimation and predictive uncertainty: a case study in the Logone catchment, Lake Chad basin. *Stochastic Environmental Research and Risk Assessment*, 32, 1665–1682. doi:10.1007/s00477-017-1466-0

Odusanya, A.E., *et al.*, 2019. *Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria*. Hydrology and Earth System Sciences, 23 (2), 1113–1144.

ONS, 2007. *Novo modelo de previsão de vazões com informação de precipitação para a bacia do rio Iguaçu*. [Accessed 13 Mar 2017]. Available from: https://www2.aneel.gov.br/aplicacoes/consulta_publica/documentos/NT_149_2007_Fuzzy_Igua%C3%A7u.pdf

ONS, 2016. *RE ONS 0181/2016: Atualização de séries históricas de vazões – período 1931 a 2015*, 1–38. [Accessed 13 Mar 2017]. Available from: https://ind01.safelinks.protection.outlook.com/?url=ftp%3A%2F%2Fftp.ecn.purdue.edu%2Fcklinkha%2FMMA&data=02%7C01%7Csathyan.dhanasekaran%40integra.co.in%7Ccd399e1576954753950908d84da7830f%7C70e2bc386b4b43a19821a49c0a744f3d%7C0%7C1%7C637344729144817243&sdata=ERzsAOXpeKy%2FHgr%2BGyWThSiSnfleCcJ3Zb9ZIJzfBMM%3D&reserved=0NetworksBrasil/Literature/PORTugueselitdocs/FlowRarehistoryforpowercomplexRJ.pdf

Piniewski, M., *et al.*, 2017. Hydrological modelling of the Vistula and Odra river basins using SWAT. *Hydrological Sciences Journal*, 62:8, 1266–1289. doi:10.1080/02626667.2017.1321842

Post, D.F., *et al.*, 2000. Predicting soil Albedo from soil color and spectral reflectance data. *Soil Science Society of America Journal*, 64 (3), 1027. doi:10.2136/sssaj2000.6431027x

Potter, R. O., *et al.*, 2004. *Levantamento de reconhecimento de solos do estado de Santa Catarina*. EMBRAPA. Rio de Janeiro: Embrapa Solos.

Rajib, M.A., Merwade, V., and Yu, Z., 2016. Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture. *Journal of Hydrology*, 536, 192–207. doi:10.1016/j.jhydrol.2016.02.037

Rientjes, T.H.M., *et al.*, 2013. Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *Journal of Hydrology*, 505, 276–290. doi:10.1016/j.jhydrol.2013.10.006

Schuol, J. and Abbaspour, K.C., 2006. Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. *Advances in Geosciences*, 9, 137–143. doi:10.5194/adgeo-9-137-2006

SEMA - Secretaria de Estado do Meio Ambiente e Recursos Hídricos. 2010. Bacias Hidrográficas do Paraná. [Accessed 14 Nov 2015]. Available from: http://www.meioambiente.pr.gov.br/arquivos/File/corh/Revista_Bacias_Hidrograficas_2015.pdf.

Sudheer, K.P., *et al.*, 2007. Impact of time-scale of the calibration objective function on the performance of watershed models. *Hydrological Processes*, 21 (25), 3409–3419. doi:10.1002/hyp.6555

Troy, T.J., Wood, E.F., and Sheffield, J., 2008. An efficient calibration method for continental-scale land surface modeling. *Water Resources Research*, 44 (9), 1–13. doi:10.1029/2007WR006513

Uda, P.K., 2016. *Evapotranspiração real da bacia do rio Iguaçu por meio do modelo METRIC*. Thesis. Universidade Federal de Santa Catarina.

Wagener, T. and Gupta, H.V., 2005. Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*, 19 (6), 378–387. doi:10.1007/s00477-005-0006-5

Wanders, N., *et al.*, 2014. The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models. *Water Resources Research*, 50 (8), 6874–6891. doi:10.1002/2013WR014639

White, K.L. and Chaubey, I., 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable Swat Model1. *JAWRA Journal of the American Water Resources Association*, 41 (5), 1077–1089. doi:10.1111/j.1752-1688.2005.tb03786.x

Wi, S., *et al.*, 2015. Calibration approaches for distributed hydrologic models in poorly gaged basins: implication for streamflow projections under climate change. *Hydrology and Earth System Sciences*, 19:2, 857–876. doi:10.5194/hess-19-857-2015

Wu, H., and Chen, B., 2015. Evaluating uncertainty estimates in distributed hydrological modeling for the Wenjing River watershed in China by GLUE, SUFI-2, and ParaSol methods. *Ecological Engineering*, 76 (110), 121. doi:10.1016/j.ecoleng.2014.05.014

Yang, J., *et al.*, 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology*, 358 (1–2), 1–23. doi:10.1016/j.jhydrol.2008.05.012

Yang, J., Reichert, P., and Abbaspour, K.C., 2007. Bayesian uncertainty analysis in distributed hydrologic modeling: A case study in the Thur River basin (Switzerland). *Water Resources Research*, 43 (10), 1–13. doi:10.1029/2006WR005497

Yen, H., *et al.*, 2016. Applications of explicitly incorporated/post-processing measurement uncertainty in watershed modeling. *Journal of the American Water Resources Association*, 52 (2), 523–540. doi:10.1111/1752-1688.12401

Zhang, D., *et al.*, 2015. Improved calibration scheme of SWAT by separating wet and dry seasons. *Ecological Modelling*, 301, 54–61. doi:10.1016/j.ecolmodel.2015.01.018

Zhang, D., *et al.*, 2016. Moving SWAT model calibration and uncertainty analysis to an enterprise Hadoop-based cloud. *Environmental Modelling and Software*, 84, 140–148. doi:10.1016/j.envsoft.2016.06.024

Zhang, X., Srinivasan, R., and Van, L.M., 2008. Multi-site calibration of the SWAT model for hydrologic modeling. *Transactions of the ASABE*, 51 (2007), 2039–2049. doi:10.13031/2013.25407

Zhang, X., Srinivasan, R., and Van Liew, M., 2010. On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model. *Hydrological Processes*, 24 (8), 955–969. doi:10.1002/hyp.7528