

Evaluating the Importance of Non-Unique Behavioural Parameter Sets on Surface Water Quality Variables under Climate Change Conditions in a Mesoscale Agricultural Watershed

B. Mehdi^{1,2,3}  · K. Schulz² · R. Ludwig⁴ · F. Ferber⁴ ·
B. Lehner¹

Received: 26 May 2016 / Accepted: 9 October 2017 /

Published online: 15 December 2017

© The Author(s) 2017. This article is an open access publication

Abstract The parameter uncertainty in the eco-hydrological model Soil and Water Assessment Tool (SWAT) was estimated using non-unique parameter sets for the Altmühl watershed (Bavaria, Germany). The Sequential Uncertainty Fitting Algorithm (SUFI-2) was used to calibrate SWAT. The non-unique parameter sets found were subsequently applied to SWAT concurrently with climate change simulations to determine the variables of streamflow, nitrate nitrogen (NO_3^- -N) and total phosphorus (TP). A suite of seven bias corrected climate change simulations provided reference (1970–2000) and future (2041–2070) climate data. The non-unique behavioural parameter sets that met an objective function of NSE >0.6 during calibration were applied to SWAT with the reference climate and with the future climate simulations. The best parameter set was also propagated through SWAT with each reference and future climate simulation in turn. Combining the non-unique behavioural parameter sets for estimating uncertainty bounds with an ensemble of climate change simulations led to a wider mean monthly spread (difference between maximum and minimum) of simulated NO_3^- -N and TP than using the best run with the future climate simulations. More monthly data was considered using the non-unique approach, resulting in statistical significances for more

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11269-017-1830-3>) contains supplementary material, which is available to authorized users.

✉ B. Mehdi
bano.mehdi@boku.ac.at

¹ Department of Geography, McGill University, 805 Sherbrooke Str. West, Montreal, QC H3A 0B9, Canada

² Present address: Institute for Water Management, Hydrology and Hydraulic Engineering, University of Natural Resources and Life Sciences, Vienna (BOKU), Muthgasse 18, 1190 Vienna, Austria

³ Division of Agronomy, University of Natural Resources and Life Sciences, Vienna (BOKU), Konrad-Lorenz Str. 24, 3034 Tulln, Austria

⁴ Department of Geography, Ludwig Maximilians University, Luisenstr. 37, 80333 Munich, Germany

months of the year and overall lower interquartile ranges. The study quantifies the non-unique behavioural parameter set contributions to the modelling prediction, which assists in making more informed decisions based on available knowledge, with its limitations, of the future simulations. We outline a simple approach that can easily be replicated for similar hydrological modelling studies.

Keywords Hydrological modelling · water quality · climate change · equifinality · uncertainty, nutrient transport

1 Introduction

The amount of (un)certainty that can be attributed to a hydrological prediction is necessary for decision-makers (Van der Keur et al. 2010; Warmink et al. 2017), particularly when adapting to climate change. Unfortunately, the uncertainty related to the simulated hydrological variables, especially water quality variables, when examining climate change impacts is not well known (Beven 2011), despite that some researchers (e.g. Sohrobi et al. 2003; Shirmohammadi et al. 2006) have advocated for quantifying this uncertainty, particularly for water quality simulations.

To date, in research that examines climate change impacts on agricultural nutrient loads, if uncertainties are considered, they are commonly limited to using an ensemble of future climate simulations, or several greenhouse gas emissions scenarios, and/or climate downscaling techniques. However, the uncertainties related to the hydrological parameters that are employed in such climate change impact studies are typically less rigorously investigated.

1.1 Uncertainties of Hydrological Modelling

The complexity of the hydrological system and the incomplete information available causes uncertainty to be intrinsic in modelling exercises. A hydrological model has three main sources of uncertainty: 1) input data (sampling and measurement); 2) conceptual (structural) uncertainty in the model where processes may not replicate the reality, or processes may be omitted (complexity); and 3) parameter uncertainty reflecting scale and/or inexact hydrological knowledge and understanding (Yang et al. 2008; Renard et al. 2010).

The uncertainty related to the parameters can be quantified in a sensitivity analysis (Norton 2015; Sarrazin et al. 2016) to better understand the model response to the influence of parameter values or model inputs. A calibration process of fitting the parameters to simulate a model output that best matches observed conditions can also provide information on parameter uncertainties (Abbaspour et al. 2007; Leta et al. 2015). Subsequent statistical methods are used to report on the change in parameter values and their impact on the model response, for an overview of these methods, see Yang et al. (2008), Duan et al. (2003), Draper and Smith (1998) and also to Razavi and Gupta (2015). It is beyond the scope of this paper to describe calibration methods in detail, however a main point is when a hydrological model calibration is undertaken, a performance metric (objective criterion) is specified to ensure a minimum level of performance.

During model calibration there can exist several parameter sets that provide equally suitable model responses that ensure a minimum level of model performance, and these are known as “non-unique parameter” sets which is a term coined as “equifinality” (Beven 1996). The

solution to a calibration therefore is several sets of parameters. A prediction with a calibrated model should propagate these parameter sets through the model to account for the uncertainty of the parameter values on the model outputs.

1.2 Uncertainties Related to Climate Change Simulations and Hydrological Modelling

The uncertainties associated with climate change impact simulations can be grouped into the following five types based on i) natural climate variability; ii) greenhouse gas emission scenario; iii) general circulation model (GCM) structure; iv) downscaling technique; and v) impact (hydrological) model (Wilby 2005; Poulin et al. 2011). In addition, the application of bias-correcting techniques adds to the uncertainty of the climate change signal (Muerth et al. 2013).

When examining hydrological and water quality changes in the future, some of the above (ideally all) uncertainties ought to be considered. Using an ensemble of climate models (Harvey et al. 1997) can cover one (or several) of the above uncertainty classes to force one (or several) hydrological models in order to determine a range of possible outcomes (i.e. Velázquez et al. 2013). A more limited approach is to only use one, or several, regional climate models forced by one GCM (e.g. Radermacher and Tomassini 2012). To increase the range of future projections, the GCM can also be run with different greenhouse gas concentrations (Nakicenovic et al. 2000; Moss et al. 2010).

1.3 Combined Uncertainties of Climate Change Impacts on Hydrology

Some statistical frameworks have been implemented to construct uncertainty bounds for hydrological quantity simulation estimates under climate change. Steinschneider et al. (2012) characterize the hydrological flow prediction with a likelihood function combined with prior distributions of parameters using Bayes Theorem, and then use MCMC sampling to evaluate the posterior distributions of hydrological and error model parameters. The uncertainties in the climate change projections were integrated with errors from the hydrological model.

Khan and Coulibaly (2010) used a Bayesian Neural Network approach to estimate the uncertainty (the mean ensemble flow and its 95% confidence intervals) of the hydrological prediction, and then generated the uncertainty of future streamflow and reservoir inflow from the mean of an ensemble of climate members.

Other studies have used confidence intervals such as the 95% prediction uncertainty to account for the uncertainties in discharge simulations under climate change (e.g. Faramarzi et al. 2009; Narsimlu et al. 2013).

Very few studies examine the uncertainty of modeled nutrient loads in a future climate. Ficklin et al. (2013) examined potential future temperature and precipitation ranges (0 to +6.4 °C, and -20% to +20%, respectively) to evaluate the sensitivity of these on sediment and nitrate outputs. They then determined the 95% confidence intervals under a range of temperature and precipitation conditions. However, they assumed the relationship between temperature and precipitation was independent.

For the first time, we report on the uncertainty of parameter sets that meet a specified objective function during the calibration of a hydrological model, and consider these “non-unique parameter” sets for simulating streamflow, nitrate nitrogen (NO_3^- -N) and total phosphorus (TP) variables in a reference period and using future climates simulations. The

uncertainty contribution of parameters is important information for decision-makers (Xu and Tung 2008), as the simulations capture a more complete range of plausible outcomes and therefore contribute pertinent knowledge for developing water management strategies to adhering to policies, such as the Water Framework Directive and Nitrate Directive in the long-term.

Our objective is to account for the uncertainty contribution of non-unique parameter sets to determine their role in climate change studies. We achieve this by quantifying the contribution of the non-unique behavioural parameter sets when assessing climate change impacts on future streamflow, NO_3^- -N and TP for the period 2041–2070. The uncertainty is quantified by the 95% confidence intervals, the interquantile ranges, the percentiles, and the total spread of the monthly variables. The evaluation of the future nutrient loads in the Altmühl per se has already been presented in Mehdi et al. (2015, 2016).

We use an approach that implements the Sequential Uncertainty Fitting Algorithm (SUFI-2; Abbaspour et al. 2004); a semi-automated calibration and uncertainty analysis tool available in SWAT CUP (Soil and Water Assessment Tool Calibration and Uncertainty Program). SUFI-2 was chosen because it is a global procedure and it applies to parameter sets, as opposed to one-at-a-time parameter analysis. Thus, an interaction between parameters in the model during each round of LHS sampling is preserved. Compared to other methods of uncertainty analysis, SUFI-2 requires relatively few runs to obtain satisfactory results (Yang et al. 2008). Although it may be desirable to apply an MCMC type of method, this requires much computing power and is suited to modeling systems that require less computational demands.

2 Materials and Methods

2.1 Description of the Study Watershed

The upper part of the Altmühl watershed (Bavaria, Germany) includes 130 km of river length from the source to the main outlet gauge in Treuchtlingen (10°54'48.91"E, 48°57'11.31"N) encompassing 980 km². The watershed is approximately 60% agricultural (summer and winter cereals, maize, oilseeds and permanent grassland), 30% forested and 5% urban.

Hydrological measurements (1961–1990) show the watershed to annually receive 700 mm of precipitation (46 mm is snowfall); evaporation comprises 475 mm, and runoff is 175 mm [BLfW, 1996]; the remaining 50 mm are presumably subsurface water flow, which is in the Karst landscape in the southern tip of the watershed. Water transfers for flood control of <4% of the annual water in the basin were not simulated.

The main water quality challenges are related to diffuse pollution of P levels. The EU environmental quality standards (2008/105/EG) established a TP threshold of 0.1 mg/L for reaching quality class II; but the TP concentrations are 0.3 and 0.2 mg/L in the Altmühl river and lake, respectively [BLfU, 2013 "Gewässerkundlicher Dienst Bayern" unpublished data available from www.gkd.bayern.de]. The NO_3^- -N concentrations in the watershed are <11 mg/L, the limit for the groundwater quality standard (91/676/EG).

2.2 The Hydrological Model SWAT

The hydrological model Soil and Water Assessment Tool (SWAT; Arnold et al. 1998) is a semi-distributed, process based hydrological model run on a daily time step. It was applied to

examine streamflow, as well as NO_3^- -N and TP loads for both a reference and a future period in the upper Altmühl watershed. One land use layer from 2008 with region-specific agricultural management practices and unchanging fertilizer inputs was used for the entire simulation period (both during the reference and future periods). To initialize soil processes, a 5-year spin up period was used prior to all SWAT simulations.

ArcSWAT version 510 was run on an ArcGIS 9.3.1 (ESRI, California, USA) platform. The setup was based on a 50 m Digital Elevation Model (Table 1S) that mapped the Altmühl watershed onto an area of 993.4 km² divided into 17 subbasins (based on an upstream drainage area > 200 ha). The watershed was further divided into hydrological response units (HRUs) which act as heterogeneous cells (grouping similar soil textures, land uses and slopes in each subbasin). An HRU threshold can be specified whereby areas below these thresholds are not considered in the subbasins, and the minority classes are reappointed so that 100% of the area is modeled. Thresholds of 0%, 10% and 0% were applied to land use, soil type and slope, respectively.

The potential evapotranspiration was estimated using the Penman-Monteith method. The baseflow filter program (Arnold and Allen 1999) was applied to streamflow records from three gauges in the watershed to determine the groundwater recharge and establish the baseflow recession constants for SWAT.

The SWAT model requires several types of data input relevant to climate, hydrological processes and plant growth (Table 1S). The observed climate data stemmed from measured sub-daily temperature, precipitation, relative humidity, cloud cover, and hours of sunshine for the period 1961–2005, provided by the German Meteorological Service (Deutscher Wetterdienst 2011). These were aggregated to a daily scale and interpolated to a 1 km grid using an elevation dependant inverse distance method (Mauser and Bach 2009). This climate data was input to calibrate SWAT.

Observed daily flows at the Thann (1981–2010), Aha (1975–2010) and Treuchtlingen (1948–2006) gauges were made available through the Water Management Authority in Ansbach. Measured monthly in-stream NO_3^- -N and TP concentrations were available at the Thann gauge (1982–2011) from the Bavarian State Office for the Environment.

More detailed information on the model set-up for the Altmühl watershed is described in Mehdi et al. (2015).

2.3 Climate Simulation Ensembles

In total, seven coherent sets of climate variables of temperature, precipitation, relative humidity, solar radiation and wind speed were available to drive the SWAT hydrological model. Each simulation from the RCMs was driven by a coupled GCM for the time periods 1970–2000 (“reference period”) and 2041–2070 (“future period”), with one of two SRES (Table 2S).

The global climate models available were based on projections using A2 or A1B greenhouse gas scenarios (Nakicenovic et al. 2000). In the A2 scenario, global CO₂ emissions reach 29 GtC by 2100; this is an increase of more than four times the 1990 levels (6 GtC). The A1B scenario has CO₂ emissions peaking around 2050, at 16 GtC; a level 2.7 times that of 1990, and fall to around 13 GtC by 2100. Both of the SRES represent pessimistic greenhouse gas emission futures and are comparable to the RCP scenario 8.5 (Rogelj et al. 2012).

Temperature for each member of the ensemble was bias-corrected using a monthly correction factor based on the difference between the ensemble-mean of the 30-year mean monthly minimum and maximum air temperature and the 30-year monthly means of the daily-observed

minimum and maximum air temperature. As well, a bias-correction method for precipitation was applied using the Local Intensity Scaling (Schmidli et al. 2006) which adjusts average monthly wet-day frequency and intensity (using a wet-day precipitation threshold of 1 mm). Using statistical algorithms, the RCM meteorological outputs (including the uncorrected variables of relative humidity, solar radiation and wind speed) were scaled to a finer resolution of a 1 km × 1 km grid with the scaling tool SCALMET (Marke 2008) using topography as the main predictor for small-scale patterns. SCALMET preserves energy and mass at the scale of the RCM grid. More detailed explanation of the climate change simulations and their post-processing is provided in Muerth et al. (2013).

2.4 SWAT Calibration and Quantification of Modelling Uncertainty

The Sequential Uncertainty Fitting algorithm (SUFI-2; Abbaspour et al. 2004) in SWAT-CUP version 4.3.2 (Abbaspour 2011) is a semi-automated inverse modelling procedure used for calibrating the SWAT simulated outputs to the available time series data of streamflow, NO_3^- -N and TP loads. It was used for finding the best run and the non-unique behavioural parameter sets. SUFI-2 is a stochastic procedure drawing independent parameter sets from a parameter hypercube using Latin Hypercube sampling (LHS). The flow parameters (19 in total) were calibrated first, followed by the NO_3^- (4 parameters) and then the P parameters (8 parameters). The bounds within each LHS sampling for the parameter values $[\theta_{\text{abs min}}, \theta_{\text{abs max}}]$ were based on the sensitivity analysis, the authors' expert knowledge of the research area and on values from the literature. The LHS leads to a uniform sampling in the hypercube with n parameter combinations; each parameter range is divided into equally distributed increments based on n , the number of desired model runs. Briefly, in SUFI-2 and according to Abbaspour et al. (2004), a global search algorithm examines the behaviour of the given objective function for each LHS parameter set for n runs. For each model run, a parameter sensitivity matrix (Jacobian matrix) is computed from the outputs. Using the Gauss-Newton method, and considering the first order derivatives, the Hessian matrix of the change in the objective function is calculated. Based on the Cramer-Rao theorem, an estimate of the lower bound of the parameter covariance matrix is calculated. The estimated standard deviation and 95% confidence interval of a parameter are calculated from the diagonal elements of the parameter covariance matrix. Parameter sensitivities (the average change in the objective function resulting from the change in each parameter while all other parameters are changing) are calculated with multiple regressions using the LHS parameters generated based on the simulated number of runs. The initial bounds for the parameter estimates are updated within the absolute ranges after each model run, whereby the new parameter ranges are determined centered on the best simulation.

The degree to which SUFI-2 algorithm accounts for the uncertainties (parameter, conceptual model and input data) in the calibrated model is defined by two measures (Abbaspour et al. 2004): the first is the p-factor and involves measuring the percentage of observed data that falls within the 95% prediction uncertainty (95PPU) of the simulated outputs. The second measure is the r-factor, of the average distance between the 2.5th percentile and the 97.5th percentile of the simulated output; it should be smaller than the standard deviation of the observed data (σ_{obs}).

$$r\text{-factor} = \frac{\frac{1}{n} \sum_{t=1}^n (Q_{t,u} - Q_{t,l})}{\sigma_{\text{obs}}} \quad (1)$$

Where n is the number of observed data points, $Q_{t,u}$ and $Q_{t,l}$ are the upper and lower bounds of the 95PPU, respectively. Ideally, for discharge the p-factor should be close to 100, and the r-factor should be less than 1.5 (Abbaspour et al. 2004; 2015).

The SUFI-2 algorithm focuses on one optimum area in the parameter space. It achieves this through an iterative process whereby it uses the outputs from the first model run (500–1000 runs) to narrow the parameter ranges to input into a second round of simulation (Abbaspour et al. 2004). The new (narrower) parameter ranges for the LHS reduce the 95PPU of the output variables. If the objective function, and the p-factor and r-factor are not satisfactory, SUFI-2 can be run 500–1000 times with the new parameter ranges until satisfactory results are obtained. Thus, the number of runs in each iteration is easily limited to a 500–1000 (Abbaspour 2011). In a study in which SWAT was applied to the European continent and where SUFI-2 was used as a calibration tool, satisfactory calibration results were achieved after three to five iterations (Abbaspour et al. 2015).

The number and nature of the input parameters, their values and the total number of runs conducted will determine how many behavioural parameter sets are obtained during the calibration. The number of runs to undertake depends on several factors, but is mainly dominated by the number of input parameters. After SWAT was calibrated for flow, NO_3^- -N and TP the final calibrated parameter ranges was the parameter space that was sampled to find the non-unique parameters. To test how many runs were necessary before convergence occurred, initial simulations were run 250, 400, 500, 800, 1000 and 1500 times (Table 1). For this study, we chose the non-unique behavioural parameter sets obtained from the 800 runs to conduct our analysis with. The 800 runs provided the best results with respect to the p-factor, the r-factor, and the objective function.

The screening of parameters used to calibrate SWAT were based on a literature review (Shen et al. 2008; Ullrich and Volk 2009; Sexton et al. 2011), combined with a sensitivity analysis carried out for parameters relevant for the streamflow, NO_3^- -N and TP variables. Also, if a parameter had values that were unknown, or less certain, it was included in the calibration. Table 2 lists the parameters with their final ranges after the calibration was completed.

To avoid over-parameterising the model (overfitting the noise), SWAT was calibrated sequentially for streamflow, NO_3^- -N, and TP (Arnold et al. 2012). SWAT was first calibrated (1964–1974) at the outlet gauge (Treuchtlingen) for surface flow at a daily time step (validated from 1975 to 1984). Because of data limitations, NO_3^- -N and TP were calibrated (1982–1983) at the monthly time step at the Thann gauge (and validated in 1984). After each time step, the

Table 1 Number of runs undertaken to determine the optimum TP calibration (1982–1983) with SUFI-2 (while the parameters for NO_3^- -N were held constant) and the number of behavioural parameter sets with NSE >0.6 found for each run

Number of runs	Overall p-factor		Overall r-factor		Behavioural parameter sets NSE > 0.6	Combined NSE
	Flow	TP	Flow	TP		
250	71	50	0.75	2.15	2	0.61
400	71	50	0.78	2.18	5	0.63
500	71	54	0.79	2.19	5	0.64
800	71	54	0.74	2.16	6	0.64
1000	71	54	0.76	2.19	11	0.64

Table 2 SWAT calibrated parameter ranges used for determining non-unique behavioural parameter sets

Parameter ^a	Description	Min	Max	Units
r_CN2	SCS Curve Number for soil moisture II	-12.8%	+6.3%	–
v_GWQMN	Threshold depth of water in the shallow aquifer	1.195	2.593	mm
v_ESCO	Soil evaporation compensation factor	0.809	0.923	–
v_CH_N2	Manning's "n" value for the main channel	0.072	0.216	–
v_CH_K2	Effective hydraulic conductivity in main channel alluvium	86.07	143.4	mm/h
v_ALPHA_BNK	Baseflow alpha factor for bank storage	0.284	0.852	days
r_SOL_AWC	Available water capacity of the soil layer	-4.9%	+26%	mm/mm
v_SURLAG	Surface runoff lag coefficient	7.952	13.69	–
v_CH_K1	Effective hydraulic conductivity in tributary channel alluvium	2.57	5.12	mm/h
v_GW_REVAP	Coefficient for groundwater transfer from the shallow aquifer to the root zone	0.151	0.200	–
v_CANMX	Maximum canopy storage	0	33.33	mm
v_TIMP	Snow pack temperature lag factor	0.632	0.877	–
v_SNOCOVMX	Minimum snow water content corresponding to 100% snow cover	0	319.8	mm
v_SNO50COV	Fraction of snow volume represented by SNOCOVMX that corresponds to 50% snow cover	0	0.590	–
v_SFTMP	Snowfall temperature	-1.686	4.946	°C
v_SMTMP	Snow melt base temperature	-5.336	1.556	°C
v_SMFMX	Melt factor for snow on June 21	3.344	10.04	mm/°C
v_SMFMN	Melt factor for snow on December 21	2.574	7.726	mm/°C
v_CH_N1	Manning's "n" value for the tributary channel	0.106	0.297	–
v_RCN	<i>Concentration of N in rainfall</i>	0.759	3.586	mg N/L
v_CMN	<i>Rate factor for humus mineralization of active organic N, P</i>	0.0001	0.0004	–
v_SDNCO	<i>Denitrification threshold water content</i>	0.547	1.641	–
v_SOL_NO3	<i>Initial NO₃⁻ concentration in the soil layer</i>	36.02	108.1	ppm
v_SOL_ORGP	Initial organic P concentration in the soil layer	66.09	288.7	ppm
v_SOL_SOLP	Initial soluble P concentration in the soil layer	36.57	109.7	ppm
v_ERORGP	P enrichment ratio for loading with sediment	0	2.724	–
v_PHOSKD	P soil partitioning coefficient	137.4	212.1	m³/Mg
v_K_P	Michaelis-Menton half-saturation constant for P	0.010	0.037	mg P/L
v_RS5	Organic P settling rate in the reach at 20 °C	0.008	0.069	day⁻¹
v_BC4	Rate constant for mineralization of organic P to dissolved P in the reach at 20 °C	0.282	0.825	day⁻¹
v_RSDIN	Initial residue cover	0	5268	kg/ha

^a Parameters in normal font were used to calibrate the flow, parameters in *italics* were used to calibrate nitrate nitrogen, parameters in **bold** were used to calibrate for total phosphorus

overall water balance and the simulated crop yields were verified against available data for the region (Table 1S).

In SUFI-2, during calibration, the user may specify the percentage error in the measured data, which is an independent error as it is a standard deviation added to the measured data. Here, we provided a 10% error for flow measured data and a 20% error for N and P related measurements (Harmel et al. 2006).

The Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe 1970) was chosen as the primary objective function for calibration because high flows are important for the transport of nutrient loads towards surface waters. The NSE is a statistical criterion that determines the relative magnitude of the variance of the residuals compared to the variance of the observed data. Due to the restrictions any single goodness-of-fit measure carries with it, several objective functions were performed post-validation (PBIAS, R², bR²) in Table 3a–b.

Table 3 a) SWAT calibration/validation results for the best run for streamflow (m^3/s) at Treuchtlingen, b) SWAT sequential calibration and validation (monthly time step) at Thann for the best run when NO_3^- -N and then TP were calibrated using final daily calibrated flow parameter ranges

a)	Calibration 1964–1974			Validation 1975–1984		
	Yearly	Monthly	Daily	Yearly	Monthly	Daily
NSE	0.96	0.77	0.57	0.81	0.75	0.68
PBIAS	1.4	13.5	13.8	0.67	13.3	3.34
R^2	0.96	0.79	0.59	0.83	0.78	0.69
bR ²	0.96	0.68	0.39	0.84	0.73	0.53
b)	Calibration 1982–1983			Validation 1984		
	NO_3^- -N + flow	TP + flow (stable NO_3^- -N parameters)		NO_3^- -N + flow	TP + flow (stable NO_3^- -N parameters)	
NSE	0.77	0.47		0.72		0.52
PBIAS	-11.8	33.5		16.6		27.7
R^2	0.77	0.71		0.71		0.70
bR ²	0.59	0.78		0.75		0.67
NSE for flow	0.83	0.80		0.87		0.59

The PBIAS measures the average tendency (%) of the simulated data to be larger or smaller than the observed data. A value of 0 represents a bias-free simulation. Negative values indicate the simulation overestimates the observed values, and positive values indicate an underestimation (Gupta et al. 1999). The coefficient of determination (R^2) describes the proportion of the observed variance that can be captured by the simulations as per Legates and McCabe (1999). Whereas the bR² multiplies the R^2 by the coefficient of the regression line to account for both the magnitude of the signal and their dynamics (Abbaspour 2011).

Once the SWAT model was calibrated and validated to at least meet the statistical criteria in Moriasi et al. (2007), whereby the NSE values are ≥ 0.5 for streamflow and nutrients at the monthly time-step, the hydrological model was run with i) the best parameter set that met the given objective function during calibration, and ii) all of the non-unique parameter sets that met the minimum objective function during calibration. Both i) and ii) were subsequently applied to the hydrological model, one at a time, using each of the reference climate as well as the future climate simulation data.

2.5 Quantifying the Uncertainties Related to Climate Change Simulations

SUFI-2 is one of several methods that can be used to calibrate SWAT that determines optimum sets of parameter values that minimize the difference between observed and simulated output variables. SUFI-2 can provide one “best estimate” parameter set (i.e. one optimum parameter set that achieved the highest fit with the objective criteria) found during the global calibration that best fits the defined objective function. Most hydrological studies use such a parameter set to validate and evaluate the model, and apply it to all subsequent model operations.

But, SUFI-2 implements a stochastic process; therefore, a number of parameter-set solutions can produce an equally satisfactory calibrated solution (Abbaspour et al. 1997). To capture some of the possible best parameter-set fits, a range of solutions that fall within the 95PPU for the variable(s) and that met the objective function was chosen.

The first part of our methodology consisted of running the SWAT model with the best estimate (optimum) parameter set found during calibration, and forcing it with each of the reference and the future climate simulations. Here, an objective function of $\text{NSE} > 0.6$ was used for all variables since it meets the satisfaction criteria for calibration purposes and it is a value

that modelers may realistically strive for. This approach will be referred to as the “*best run*” approach.

The second approach retained all of the calibrated parameter sets that allowed streamflow, NO_3^- -N and TP loads to meet the objective function of $\text{NSE} > 0.6$ (including the best run). The parameter sets which met a given objective function during the calibration period were subsequently propagated through the hydrological model. The same likelihood is attributed to each parameter set for which the objective function selected threshold is exceeded. Several sets of non-unique behavioural parameter sets were found using this approach; the number of sets was determined by the number of runs performed in SUFI-2.

The behavioral SWAT model parameter sets were propagated through SWAT with each of the future climate simulations, respectively to produce 42 simulations of streamflow, NO_3^- -N and TP (7 future climate simulations \times 6 sets of non-unique behavioural parameter sets). This approach will be referred to as the “*non-unique*” approach.

The future climate uncertainties were mainly captured by using a suite of climate models. And, to remove any bias in the climate model simulation, the reference climate simulations were compared to the future climate simulations.

3 Results

3.1 Future Climate Simulations with Temperature and Precipitation Changes

All climate simulations were bias-corrected. Since this correction intends to maintain the natural variability, and not force the model to match the observations, some differences between the observed data and the bias-corrected climate simulations from 1970 to 2000 may still occur (Fig. 1a–b), especially with precipitation (Teutschbein and Seibert 2012).

Compared to the respective reference climate (1970–2000), the future climate (2041–2070) projected mean monthly precipitation changes in the range of -20% to $+74\%$, and mean monthly temperatures increases of $0.75\text{ }^\circ\text{C}$ to $4.0\text{ }^\circ\text{C}$.

3.2 Evaluating the “Best Run” and “Non-Unique” Performance using Observed Data

The simulated annual water balance had an error of $< 1\%$. The simulated monthly outputs, resulting from the best set of calibrated parameters, demonstrate that SWAT reproduced the timing of dry spells and peak flows well (Fig. 2a and 1S). The magnitude of flows was also modelled satisfactory (Moriassi et al. 2007), although SWAT tended to underestimate the flow ($\text{PBIAS} = 13.8\%$; Table 3a). The NO_3^- -N and TP simulations reproduced the timing of the events (Fig. 2b–c and Table 3b), although modelled TP had overall lower values ($\text{PBIAS} 33.5\%$), whereas NO_3^- -N was overestimated ($\text{PBIAS} -11.8$). On the whole, based on the performance criteria, the sequential calibration of the three variables led to a satisfactory performing SWAT model for the Altmühl watershed.

Each non-unique behavioural parameter set found in SWAT caused streamflow, NO_3^- -N and TP loads to be simulated, so that combined, they met the specified objective criterion ($\text{NSE} > 0.6$) for the calibration period. From the 800 runs, 6 non-unique behavioural parameter sets met the objective function (Table 3S). The relatively few sets found were in part due to the purposefully high objective function chosen as a realistic goal for calibration, but it also

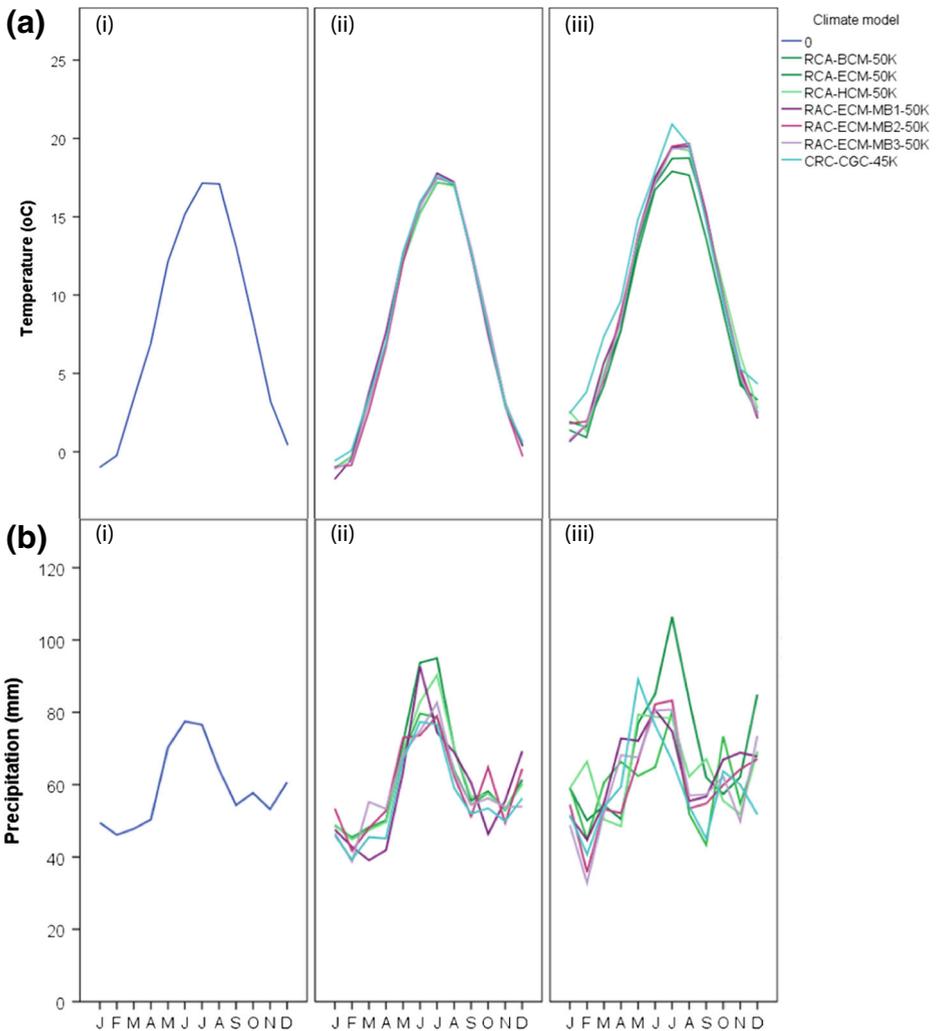
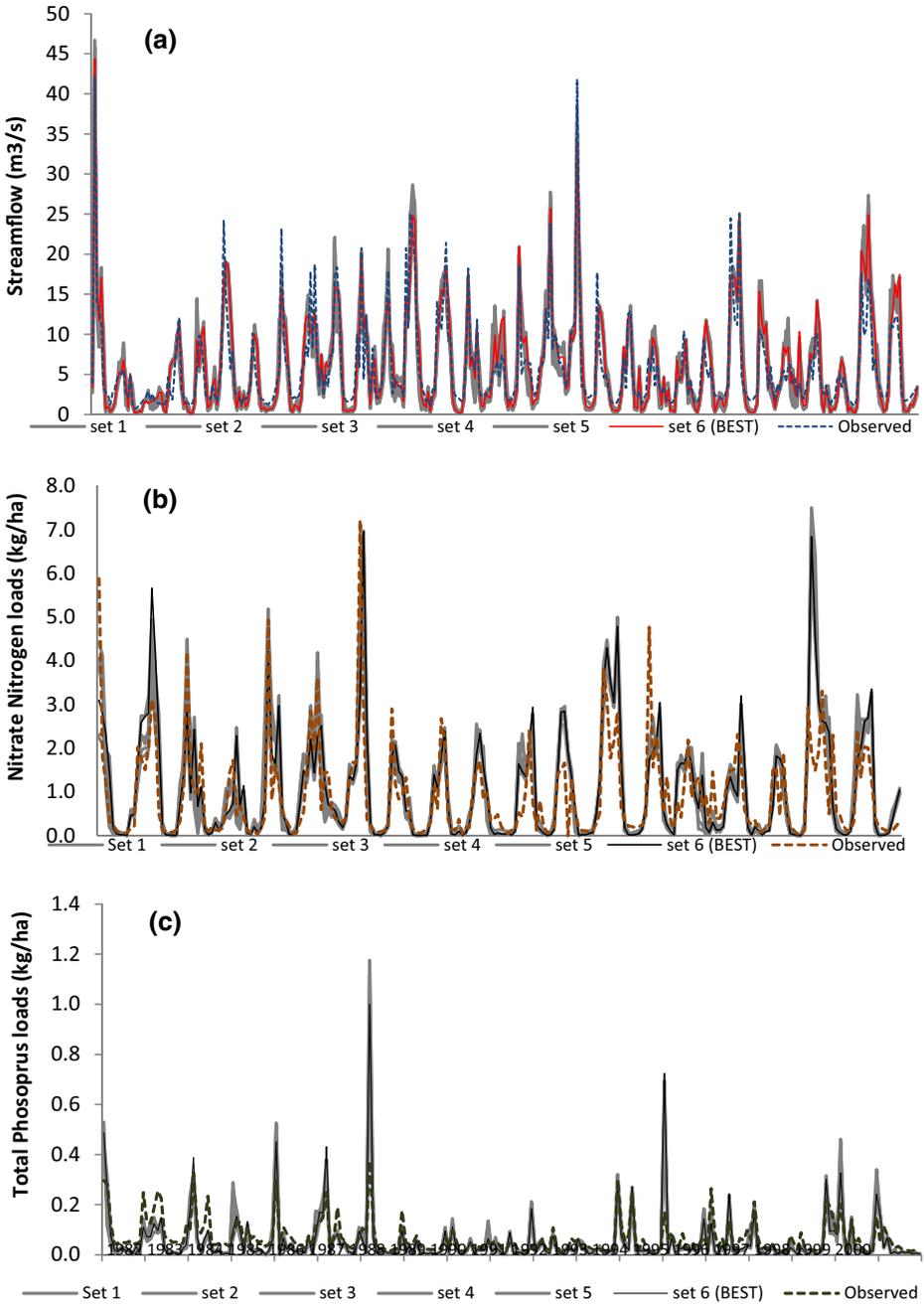


Fig. 1 **a** **i**) Observed mean monthly temperature (1970–2000) compared to the **ii**) climate simulated reference temperature data (1970–2000) and **iii**) future temperature data (2041–2070), **b** **i**) Observed mean monthly precipitation (1970–2000) compared to **ii**) climate simulated reference precipitation data (1970–2000) and **iii**) future precipitation data (2041–2070)

reduced the non-uniqueness. A stricter common objective criterion (i.e. a higher NSE) was tested (results not shown) and as expected, led to one or no behavioural parameter sets found, also because TP was the limiting variable to reach a higher NSE.

The non-unique behavioural parameter sets provided a first indication of the uncertainty information through the 95PPU for the simulated variables (Fig. 2a–c), where the p-factor indicated that 34% of the simulations encompassed the measured streamflow data, 28% of the simulated outputs bracketed the observed NO_3^- -N loads, and 14% of the TP simulations bracketed their observed values. These p-factors are smaller than when the overall uncertainty of the simulations is considered after the iterations are completed for the calibration (Table 1)



because the non-unique behavioural parameter sets are reporting p-factors that met our NSE objective function and thus examine a narrower range of outputs.

◀ **Fig. 2** **a)** Monthly streamflow at Treuchtlingen for the reference period (1971–2000) using SWAT simulated with observed climate and the best run simulated (red line) p-factor 34%; r-factor 0.38, and using observed climate and the non-unique behavioural parameter sets (grey lines), **b)** Monthly NO_3^- -N at Thann for the period of observed data 1982–2000 (orange line) and from SWAT simulated with the observed climate and the best parameter set (black line) and with the non-unique behavioural parameters (grey lines: NSE \geq 0.6; p-factor 28%; r-factor 0.42), **c)** Monthly TP at Thann for the period of observed data 1982–2000 (green line) and from SWAT simulated with the observed climate and the best parameter set (black line) and with the non-unique behavioural parameters (grey lines: NSE \geq 0.6; p-factor 14%; r-factor 0.37)

3.3 Simulated Streamflow and Water Quality Using the “Best Run” and “Non-Unique” Approach with Reference and Future Climate Simulations

The monthly streamflow, NO_3^- -N and TP loads simulated with SWAT using the best parameter set and the observed climate data (1970–2000) are depicted in Fig. 2a–c, respectively. The simulation based on the single best run provides little information about the simulated uncertainty for streamflow, NO_3^- -N and TP loads since it is a comparison between two signals (one modeled signal with one observed data set). Thus, the only source of uncertainty can be gleaned from the model error between observed and simulated data (Table 3a).

The best run was also applied with each of the seven reference climate simulations (period 1970–2000) in SWAT (grey lines in Fig. 2S); these depict the irreducible climate variability of modelled streamflow. The differences between the blue line and the grey lines depict discrepancies in the simulated results that are due to climate model disparities caused by the physical processes represented in the models, and/or in the different initial conditions with different members from the same model. The best run was finally applied in SWAT with each of the seven future (period 2041–2070) climate simulations (red lines in Fig. 2S).

When comparing the best run’s mean of the monthly streamflows for the reference period to the future period, statistically significantly higher streamflow was simulated in May and June. When the non-unique approach was used, the mean streamflow was significantly higher in May and significantly lower in September (Table 4S).

The non-unique approach applied to SWAT with the future climate simulations depict the parameter uncertainty (Fig. 3S). Using the non-unique approach, the complete spread between the minimum and the maximum of the points was greater (Table 4S) this is also depicted by the monthly extreme values in Fig. 3, even though the 10th and 90th monthly percentiles were mostly lower, or similar, compared to the best run.

The seasonal differences highlighted that the interquartile ranges using the two methods differed. For streamflow, the non-unique approach had smaller interquartile ranges (Table 5S), despite having a greater spread of data, indicating that the majority of the data points were centred within the quartile ranges.

Using the best run approach, the mean monthly future NO_3^- -N loads increased significantly in most months, except in March, April, August and September (Fig. 4 and Table 6S). In SWAT the NO_3^- -N loads in the stream are driven by surface flow, infiltration and throughflow. The NO_3^- -N loads are sensitive to precipitation and infiltration changes, this is reflected in the future climates by the higher monthly spread in the 25th and 75th percentiles compared to the reference period. For the best run approach with the future climate simulations, the 90th percentile NO_3^- -N loads were simulated to increase in all months (Table 6S) indicating a higher right hand tail of the load distribution and higher NO_3^- -N loadings. The non-unique approach showed statistical significances in all of the same months as the best run approach, but it additionally showed a statistically significant increase in loads in August and a decrease

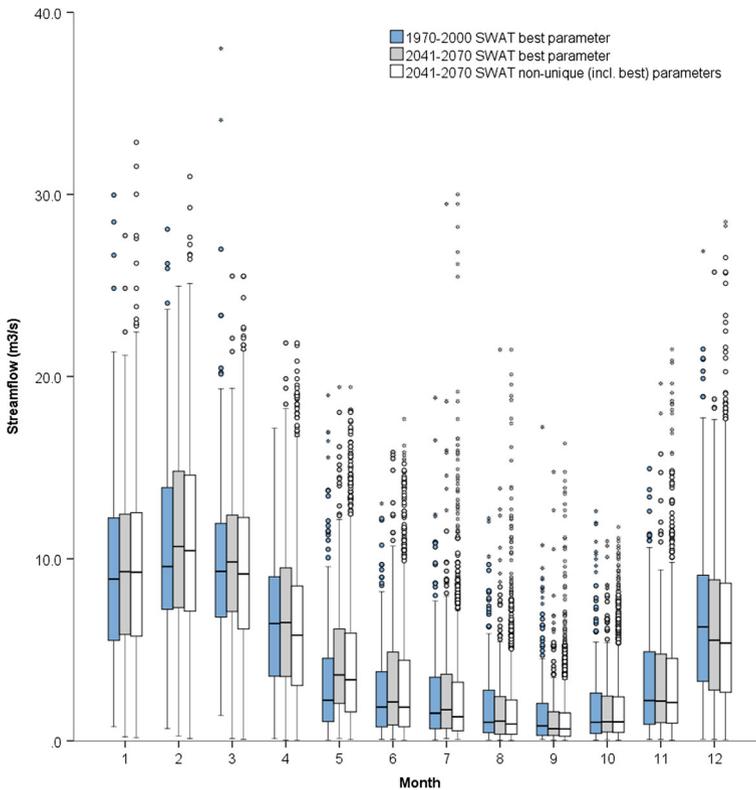


Fig. 3 SWAT simulated monthly flow at Treuchtlingen, using the best parameter set with reference climate simulations (blue boxes) and future climate simulations (grey boxes), and the non-unique approach with future climate simulations (white boxes). Boxplots show the central mark as being the median, the upper and lower edges of the box are the 75th and 25th percentile, respectively, and the whiskers extend to the values that lie inside one and half box lengths from the quartiles. The circles represent values which lie one and a half box lengths away from the quartile (considered outliers). Asterisks are values that lie more than three box lengths away from the quartile (considered extremes)

occurring in April. The total spread between the minimum and maximum of the monthly data points in the non-unique approach was greater than in the best run (Fig. 4), by up to 1 kg/ha (the difference in the maximum values in February) in spite of the seasonal interquartile ranges (Table 7S) being lower or the same as in the best run approach.

Using the best run approach, the mean TP loads in winter months (December to March) were simulated to be significantly lower in the future period. Decomposition and mineralization of fresh organic residues and of the humus add plant-available P to the soil. These processes are controlled by the decay rate constant, which is determined in part by the soil temperature, warmer temperatures increase the decay rate. The 90th percentiles were also lower in winter. However, in April, May and October the means were higher (Fig. 5 and Table 8S). The non-unique behavioural parameter sets also showed these significant changes and additionally indicated a significant increase in July. Again, the spread between the minimum and maximum was greater by 0.17 kg/ha (the difference in the maximum values in January) even though the seasonal interquartile ranges (Table 9S) were lower for the non-unique behavioural parameter sets compared to the best run approach.

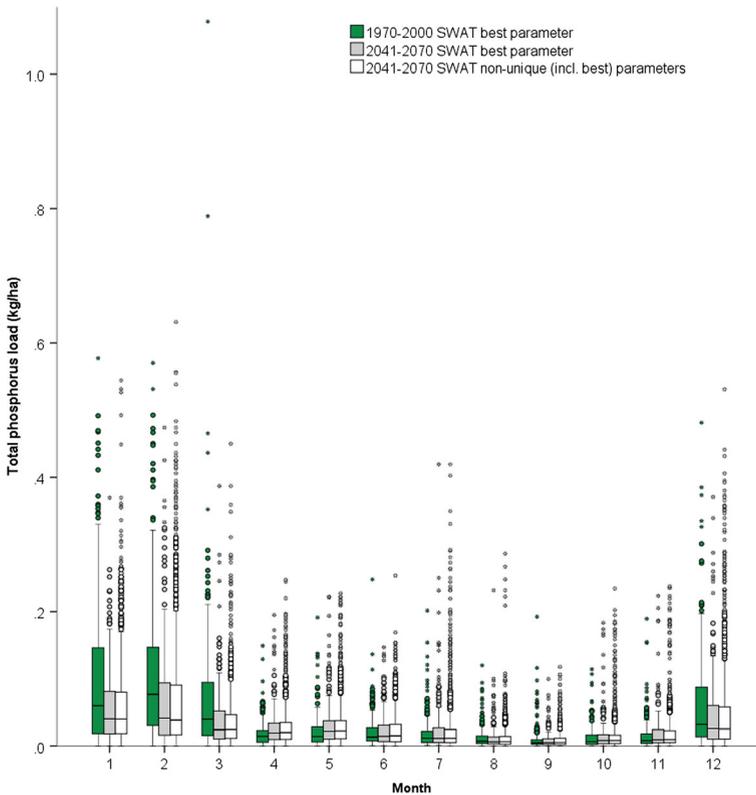


Fig. 4 SWAT simulated monthly TP loads (kg/ha) at Treuchtlingen, using the best parameter set with reference climate simulations (green boxes) and future climate simulations (grey boxes), and the non-unique approach with future climate simulations (white boxes)

The monthly differences between the mean of the future simulations using the best run and the mean of the simulations using the non-unique approach indicate changes ranging from 0 to $0.6 \text{ m}^3/\text{s}$ for streamflow; changes of -0.02 to 0.17 kg/ha in NO_3^- -N loads; and changes of -0.001 to 0.003 kg/ha in TP loads. Considering the basin size of $99,335 \text{ ha}$, the two approaches diverged by mean monthly NO_3^- -N loads of -2 to 17 Mg ; and mean TP loads of 0 to 0.3 Mg .

4 Discussion

Using the method outlined, more data is available for the non-unique approach; each month 1260 values ($30 \text{ years} \times 7 \text{ climate simulations} \times 6 \text{ non-unique behavioural parameter sets}$) are compared to the best run approach which has 210 values ($30 \text{ years} \times 7 \text{ climate simulations} \times 1 \text{ best run}$) which allows for more data to be analysed in the non-unique approach and statistical tests to be more robust, hence more months have statistical differences in nutrient loads with the non-unique approach.

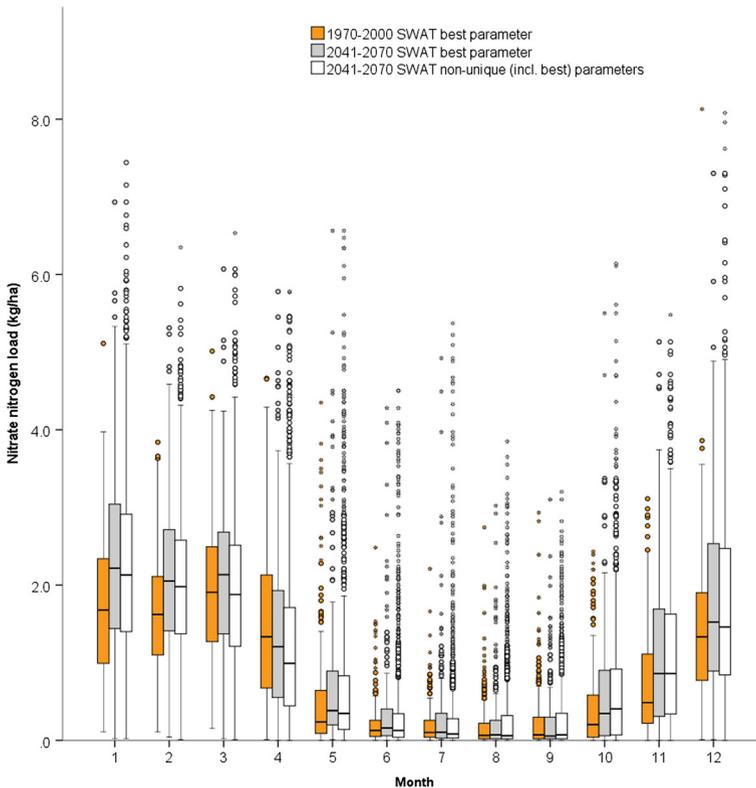


Fig. 5 SWAT simulated monthly NO_3^- -N loads (kg/ha) at Treuchtlingen, using the best parameter set with reference climate simulations (orange boxes) and future climate simulations (grey boxes), and the non-unique approach with future climate simulations (white boxes)

The climate change signal is provided by the best run approaches in SWAT; by taking the difference between the future climate and the reference climate simulations. The impacts of the non-unique parameter sets is shown by the difference between SWAT applied with the non-unique parameter sets with the future climate simulations and SWAT applied with the best runs using the future climate simulations.

Applying the non-unique behavioural parameter sets with future climate simulations to SWAT, simulated the outliers to have an overall greater spread between the minimum and maximum values, than with the best parameter approach. This indicates the number of simulated extreme events increased in the non-unique approach, so that the range of events was greater even though the interquartile range is narrower. As well, with the non-unique approach, the independent t-tests revealed statistically significant differences for more months during the year.

Using the non-unique behavioural parameter sets, the monthly nutrient loads were different than for the best run approach (Tables 6S and 8S). Using the non-unique approach, the seasonal interquartile ranges showed, with 50% confidence, that the simulated future NO_3^- -N loads transported from the fields would lie between 1.19 and 2.66 kg/ha in winter months. Using the best run approach, the range was 1.12 to 2.76 kg/ha. Although seemingly minimal, depending on the season, the differences between the approaches amount to a discrepancy at

the basin outlet of -3 to 18 Mg in the mean median NO_3^- -N loads and -20 to 100 kg in the mean median TP loads.

In climate change studies when a hydrological model is being applied to non-stationary conditions it is critical to use more than one suitable parameter set for the future predictions to cover a wider range of potential outcomes. Although the simulation cannot guarantee any expression of confidence in future predictions due to epistemic errors, using non-unique behavioural parameter sets provide a more honest range for the simulated variables by accounting for outcomes that are due to parameter uncertainties. A limitation of the SUFI-2 procedure is that the non-unique behavioural parameter sets are centered around one and the same local optimum in the parameter space.

4.1 Calibration of SWAT

A discussion on the calibration methodology of SWAT is warranted, as different approaches to calibration produce different results. Several different procedures were tried for calibration.

- i) The variables can be calibrated sequentially. After the flow is calibrated, the best flow ranges can be included with the NO_3^- -N parameter ranges, to calibrate for NO_3^- -N; and then, the same best flow ranges can be included with the TP parameters when calibrating for TP. The results of this calibration method are presented in this paper.
- ii) The variables can be calibrated sequentially only using the best parameter set each time, i.e. first streamflow, then NO_3^- -N then TP is calibrated; always using the best parameter set from the preceding variable to find the best fit of the new variable.
- iii) All three variable ranges can be calibrated together.

Calibrating on any one variable alone will yield a higher NSE for the best run. In our sequential calibration, for TP we obtained an NSE of 0.47 (Table 3b), yet the highest NSE achieved for TP during the multiple variable simulation (method iii listed) was 0.18 (results not shown); similarly, when NO_3^- -N loads were calibrated with flow alone, an NSE of 0.77 was achieved, but this dropped to 0.39 when TP parameters were added.

Arnold et al. (2012) provide a solid overview of best practices for model calibration and validation, but do not address multi-variable calibration techniques. Multi-variable objectives affected our solution by creating more limitations and constraints to be met, which is also observed in Ficklin et al. (2013) who simultaneously calibrated SWAT using streamflow, sediment, nitrate and pesticide variables at multiple gauging stations. The NSEs obtained during their multi-variable calibration varied from 0.11 (sediments) to 0.94 (streamflow). It may be possible to overcome this restriction by applying a constrained objective function to ensure that only the loads falling between specified value ranges are calibrated for. Using this approach, Abbaspour et al. (2007) calibrated SWAT simultaneously for discharge, sediments, NO_3^- and TP at the watershed outlet, and obtained p-factors and r-factors for discharge of 91 and 1.0, respectively; for sediment they were 80 and 1.5, respectively; TP had 78 and 1.35, respectively; and NO_3^- had 82 and 1.0, respectively.

Poor results in the calibration are due to a combination of influences, such as the correlation between streamflow and individual nutrients, such as TP, and the lack of correlation between the nutrients themselves. For example, the CN directly influences runoff and therefore all other aspects of the water balance. Soil erosion is greatly affected by the CN which directly impacts the amount of particulate P transported, but not so much NO_3^- -N transportation. Global

hydrological parameters (i.e. *bsn parameters) that pertain to the whole watershed and affect the hydrology of the entire calibration area are also responsible. If they are changed for one variable they will affect the hydrology of the other variables as well. The variation of any nutrient load parameter is affected by the uncertainty of the parameters associated with the streamflow process (Shen et al. 2008).

Finally, the choice of objective function is critical. For example, objective functions based on the NSE favour good agreements with peak flows and peak nutrient loads. Applying multi-objective functions may provide solutions which attenuate the partiality of any particular objective function statistic.

Wellen et al. (2015) highlighted how only approximately 10% of distributed, process-based, diffuse pollutant water quality modeling studies accounted for the uncertainty of their model predictions. Based on this information, methods and best practices to calibrate and to determine uncertainties should be made more widely available to the modelling community.

5 Conclusion

Applying the non-unique approach provides a greater range of the combined uncertainty when using the SWAT model for predicting future water quality. Here, using seven climate simulations, the future variability for NO_3^- -N and TP using the non-unique behavioural parameter sets was shown to have mean maximum monthly values that were up to 1.0 kg/ha and 0.17 kg/ha higher, respectively than those provided by the best run approach.

The uncertainty simulated using non-unique behavioural parameter sets as determined by SUFI-2 depends on several factors, such as the objective criteria threshold chosen; the number of objective criteria; the number of variables being calibrated for at any one time; and the number of gauges used during the calibration process.

Beyond the calibration process itself, predicting uncertainty bounds of the water quality variables also involves examining the hydrological model conceptual structure; the hydrological model input data uncertainties; and the period of simulation. As well, other changes occurring in the watershed, such as land use change, may also add to the uncertainty of the modelled outcome.

In this study, the seven climate change simulations dominated the uncertainty compared to the non-unique behavioural parameter sets, but reporting the parameter uncertainties associated with future streamflow and water quality variables portrays the range in existing ambiguities of the scientific tools available, and further provides added knowledge in potential extreme events compared to providing only the best run approach. Box plots, cumulative frequency distributions and probability distributions of non-unique simulated outcomes give a sense of the potential ranges of outcomes.

Further research on how uncertainty bounds are affected include testing the choice of the calibration method (i.e. SUFI-2, GLUE, Parasol, etc.); the calibration time period; the data available for calibration (sampling frequency and truthfulness); the number of simulation runs; the number of parameters; and the range of parameters used in the calibration.

Acknowledgements Open access funding provided by University of Natural Resources and Life Sciences Vienna. This work was partially supported by the Natural Sciences and Engineering Research Council of Canada, the Bourses Québec-Bavière, and the Klima- und Energiefonds KR13AC6KL11021.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abbaspour KC, Rouholahnejad E, Vaghefi S, Srinivasan R, Yang H, Kløve B (2015) A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *J Hydrol* 524:733–752
- Abbaspour KC (2011) SWAT-CUP4: SWAT Calibration and Uncertainty Programs - A User Manual. Swiss Federal Institute of Aquatic Science and Technology, Eawag, Dübendorf, p 103
- Abbaspour KC, Johnson CA, van Genuchten MT (2004) Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J* 3(4):1340–1352
- Abbaspour KC, van Genuchten MT, Schulin R, Schläppi E (1997) A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters. *Water Resour Res* 33(8):1879–1892
- Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R (2007) Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *J Hydrol* 333(2–4):413–430
- Arnold JG, Allen PM (1999) Automated methods for estimating baseflow and ground water recharge from streamflow records. *J Am Water Res Assoc* 35(2):411–424
- Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, van Griensven A, van Liew MW, Kannan N, Jha MK (2012) SWAT: Model use, calibration, and validation. *Trans ASABE* 55(4):1491–1508
- Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling and assessment part I: Model development. *J Am Water Res Assoc* 34(1):73–89
- Beven K (1996) The limits of splitting: Hydrology. *Sci Total Environ* 183(1–2):89–97
- Beven K (2011) I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrol Process* 25(9):1517–1520
- Deutscher Wetterdienst (2011) Daten der synoptischen Klimastationen. www.dwd.de/bvbw/appmanager/bvbw
- Draper NR, Smith H (1998) Applied Regression Analysis, 3rd edition. J. Wiley & Sons, Inc., New York
- Duan Q, Gupta HV, Sorooshian S, Rousseau AN, Turcotte R (2003) Calibration of Watershed Models, Water Science and Application 6. American Geophysical Union, Washington
- Faramarzi M, Abbaspour KC, Schulin R, Yang H (2009) Modelling blue and green water resources availability in Iran. *Hydrol Process* 23(3):486–501
- Ficklin DL, Luo Y, Zhang M (2013) Climate change sensitivity assessment of streamflow and agricultural pollutant transport in California's Central Valley using Latin hypercube sampling. *Hydrol Process* 27(18):2666–2675
- Gupta HV, Sorooshian S, Yapo PO (1999) Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J Hydrol Eng* 4(2):135–143
- Harmel RD, Cooper RJ, Slade RM, Haney RL, Arnold JG (2006) Cumulative Uncertainty in Measured Streamflow and Water Quality Data for Small Watersheds. *Trans Am Soc Agric Eng* 49(3):689–701
- Harvey D, Gregory J, Hoffert M, Jain A, Lal M, Leemans R, Raper S, Wigley T, de Wolde J (1997) An Introduction to Simple Climate Models Used in the IPCC Second Assessment Report. IPCC, Geneva
- Khan MS, Coulibaly P (2010) Assessing hydrologic impact of climate change with uncertainty estimates: Bayesian neural network approach. *J Hydrometeorol* 11(2):482–495
- Legates DR, McCabe GJ Jr (1999) Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resour Res* 35(1):233–241
- Leta OT, Nossent J, Velez C, Shrestha NK, van Griensven A, Bauwens W (2015) Assessment of the different sources of uncertainty in a SWAT model of the River Senne (Belgium). *Environ Model Softw* 68:129–146
- Marke T (2008) Development and Application of a Model Interface to couple Land Surface Models with Regional Climate Models for Climate Change Risk Assessment in the Upper Danube Watershed. Dissertation, Ludwig Maximilians Universität
- Mausser W, Bach H (2009) PROMET - Large scale distributed hydrological modelling to study the impact of climate change on the water flows of mountain watersheds. *J Hydrol* 376(3–4):362–377

- Mehdi BB, Ludwig R, Lehner B (2015) Evaluating the impacts of climate change and crop land use change on streamflow, nitrates and phosphorus: A modeling study in Bavaria. *J Hydrol Reg Stud* 4:60–90
- Mehdi B, Ludwig R, Lehner B (2016) Simulated future changes of extreme nutrient loads in a mesoscale agricultural watershed in Bavaria. *Die Bodenkultur J Land Manag Food Environ* 67(2): 77–90
- Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50(3):885–900
- Moss RH, Edmonds JA, Hibbard K et al (2010) The next generation of scenarios for climate change research and assessment. *Nature* 463(7282):747–756
- Muerth MB, Gauvin St-Denis B, Ricard S, Velázquez JA, Schmid J, Minville M, Caya D, Chaumont D, Ludwig R, Turcotte R (2013) On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrol Earth Syst Sci* 17:1189–1204
- Nakicenovic N, Alcamo J, Davis G et al (2000) Emission Scenarios. A special report of working group III of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge
- Narsimlu B, Gosain AK, Chahar BR (2013) Assessment of future climate change impacts on water resources of Upper Sind River Basin, India using SWAT model. *Water Resour Manag* 27:3647–3662. <https://doi.org/10.1007/s11269-013-0371-7>
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I - A discussion of principles. *J Hydrol* 10(3):282–290
- Norton J (2015) An introduction to sensitivity assessment of simulation models. *Environ Model Softw* 69:166–174
- Poulin A, Brissette F, Leconte R, Arsenault R, Malo JS (2011) Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. *J Hydrol* 409(3–4): 626–636
- Radermacher C, Tomassini L (2012) Thermodynamic causes for future trends in heavy precipitation over Europe based on an ensemble of regional climate model simulations. *J Clim* 25(21):7669–7689
- Razavi S, Gupta HV (2015) What do we mean by sensitivity analysis? The need for comprehensive characterization of “global” sensitivity in Earth and Environmental systems models. *Water Resour Res* 51(5):3070–3092
- Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW (2010) Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resour Res* 46(5): W05521
- Rogelj J, Meinhausen M, Knutti R (2012) Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nature Clim Change* 2:248–253
- Sarrazin F, Pianosi F, Wagener T (2016) Global Sensitivity Analysis of environmental models: Convergence and validation. *Environ Model Softw* 79:135–152
- Schmidli J, Frei C, Vidale PL (2006) Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. *Int J Climatol* 26(5):679–689
- Sexton AM, Shirmohammadi A, Sadeghi AM, Montas HJ (2011) Impact of parameter uncertainty on critical SWAT output simulations. *Trans ASABE* 54(2):461–471
- Shen Z, Hong Q, Yu H, Liu R (2008) Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. *Sci Total Environ* 405(1–3):195–205
- Shirmohammadi A, Chaubey I, Harmel RD, Bosch DD, Muñoz-Carpena R, Dharmasri C, Sexton A, Arabi M, Wolfe ML, Frankenberger J, Graff C, Sohrabi TM (2006) Uncertainty in TMDL models. *Trans ASABE* 49(4):1033–1049
- Sohrabi TM, Shirmohammadi A, Chu TW, Montas H, Nejadhashemi AP (2003) Uncertainty analysis of hydrologic and water quality predictions for a small watershed using SWAT2000. *Environ Forensic* 4(4): 229–238
- Steinschneider S, Polebitski A, Brown C, Letcher BH (2012) Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change. *Water Resour Res* 48(11):W11525
- Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J Hydrol* 456–457:12–29
- Ullrich A, Volk M (2009) Application of the Soil and Water Assessment Tool (SWAT) to predict the impact of alternative management practices on water quality and quantity. *Agr Water Manag* 96(8):1207–1217
- Velázquez JA, Schmid J, Ricard S, Muerth M, Gauvin St-Denis B, Minville M, Chaumont D, Caya D, Ludwig R, Turcotte R (2013) An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrol Earth Syst Sci* 17:565–578
- Van der Keur P, Brugnach M, Dewulf A, Refsgaard JC, Zorilla P, Poolman M, Isendahl N, Raadgever GT, Henriksen HJ, Warmink JJ, Lamers M, Mysiak J (2010) Identifying uncertainty guidelines for

- supporting policy making in water management illustrated for Upper Guadiana and Rhine Basins. *Water Resour Manag* 24:3901–3938
- Warmink JJ, Brugnach M, Vinke-de Kruijf J, Schielen RMJ, Augustijn DCM (2017) Coping with uncertainty in river management: Challenges and ways forward. *Water Resour Manag*. <https://doi.org/10.1007/s11269-017-1767-6>
- Wellen C, Kamran-Disfani A-R, Arhonditsis GB (2015) Evaluation of the current state of distributed watershed nutrient water quality modeling. *Environ Sci Technol* 49:3278–3290
- Wilby RL (2005) Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrol Process* 19(16):3201–3219
- Xu Y-P, Tung Y-K (2008) Decision-making in water management under uncertainty. *Water Resour Manag* 22: 535–550
- Yang J, Reichert P, Abbaspour KC, Xia J, Yang H (2008) Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *J Hydrol* 358(1–2):1–23