The name(s) of the author(s):

Ying Chen • Binbin Lu • Chong-Yu Xu • Xingwei Chen • Meibing Liu • Lu Gao • Haijun Deng

Title:

Uncertainty evaluation of best management practice effectiveness based on the AnnAGNPS model

The affiliation(s) and address(es) of the author(s):

Ying Chen[1,2]• Binbin Lu[1,2] • Chongyu Xu[3] • Xingwei Chen [1,2] • Meibing Liu[1,2] • Lu Gao[1,2] • Haijun Deng[1, 2]

[1] College of Geographic Sciences, Fujian Normal University, Fuzhou 350007, China

[2] Cultivation Base of State Key Laboratory of Humid Subtropical Mountain Ecology, Fuzhou 350007, China

[3] Department of Geosciences, University of Oslo, Oslo, Norway

The e-mail address, telephone and fax numbers of the corresponding author:

corresponding author: Ying Chen

e-mail: chenying_nju@163.com

Tel: +86-591-83465214

Fax: +86-591-83465397

Acknowledgements:

This work was supported by the National Key Research and Development Program of China (2018YFE0206400), the Scientific Project from Fujian Provincial Department of Science and Technology (2019R1002-3 and 2021R1002006), and the Scientific Project from Fujian Key Laboratory of Severe Weather (2020KFKT01).

Uncertainty evaluation of best management practice effectiveness based on the AnnAGNPS model

Ying Chen[1,2] • Binbin Lu[1,2] • Chong-Yu Xu[3] • Xingwei Chen[1,2] • Meibing Liu[1,2] • Lu Gao[1,2] • Haijun Deng[1,2]

Abstract

Uncertainty of best management practices (BMP) effectiveness is an important factor in the development of watershed management plans. This study explored the uncertainty of BMP effectiveness in reducing total nitrogen (TN) load owing to the uncertainty in hydrological parameters, thereby improving their reliability. A watershed model, annualized agricultural non-point source pollution (AnnAGNPS), was employed to evaluate the effectiveness of the four potentially feasible BMPs (i.e., riparian buffer, fertilization reduction, no-tillage, and parallel terraces) in the Shanmei Reservoir watershed, located in the southeastern coastal region of China. Annual and seasonal uncertainty variations in BMP effectiveness were evaluated based on ten parameter sets selected from 1000 parameter groups using Latin hypercube sampling. The results showed that the uncertainty of the BMP effectiveness in reducing the TN

✉ Ying Chen
chenying_nju@163.com

[1] College of Geographic Sciences, Fujian Normal University, Fuzhou 350007, China

[2] Cultivation Base of State Key Laboratory of Humid Subtropical Mountain Ecology, Fuzhou 350007, China

[3] Department of Geosciences, University of Oslo, Oslo, Norway

load was larger than the uncertainty of TN load simulation at annual and seasonal time scales. The BMP effectiveness tended to be higher during summer than during the other seasons. The uncertainty of BMP effectiveness varied seasonally, and it was always lower during summer for most BMPs. This indicated that the impact of BMPs on reducing TN load was more effective, with a higher reduction rate and lower uncertainty in summer. Among the BMPs, the parallel terrace was the most effective measure for reducing TN load since it had the highest reduction efficiency and relatively low uncertainty. Although this study is a case study, it can provide a scientific reference for decision-making in uncertain situations when AnnAGNPS is applied for water quality simulations.

Keywords Best management practice (BMP) · Watershed modeling · Non-point source pollution · Uncertainty · AnnAGNPS

1 Introduction

Non-point source pollutants, such as nitrogen (N) and phosphorus (P), sediments, and agricultural chemicals, have become a major cause of water quality degradation in rivers and streams worldwide. Large amounts of phosphorus and nitrogen loading loss may result in the eutrophication of freshwater, severely restricting water use (Schürz et al. 2019; Zhang et al., 2021). Thus, reduction or alleviation of nutrient load is a key issue for water quality improvement and aquatic ecosystem restoration (Schuwirth et al. 2019).

Environmental Conservation actions in agricultural districts, often referred to as Best management practices (BMPs), are recognized as effective measures to reduce the nutrient loads in reservoir watersheds (USEPA, 2008; Ni et al. 2020; Rodrigues e al. 2021). For example, these include the structural measures a) riparian buffer (RB) and b) parallel terraces (PT), and n-structural measures c) fertilization reduction (FR) and d) no-tillage (NT). Watershed modeling has commonly been used to evaluate the BMP effectiveness in watersheds because of the high cost of continuous water quality monitoring, complex physio-graphic nature of the watershed, and countless possible implementation scenarios (Schuwirth et al. 2019).

Conventionally, the modeling approach entails the calibration-validation-prediction process (Arabi et al. 2007; Bai et al. 2020; Casallas-Ojeda et al. 2021). In most cases, a single set of optimum parameter settings was obtained by model calibration and considered as the “true” solution. The calibrated model was then used to evaluate the BMP effectiveness by comparing the simulated results with and without BMPs. This BMP effectiveness evaluation was subject to the uncertainty problem of equifinality (i.e., different parameter sets reproduce acceptable model outputs; Beven, 2006). Equifinality makes it difficult to accurately recognize real parameters. The uncertainty problem of equifinality can propagate into simulation processes and affect BMP effectiveness evaluation, resulting in unrealistic and biased decisions. Therefore, the analysis of model parameter uncertainty in BMP effectiveness is of great significance (Engebretsen et al. 2019).

Previous studies have discussed the impact of parameter equifinality (i.e., parameter uncertainty) on the BMP effectiveness. However, there were no universal results for the impact of parameter uncertainty on BMP effectiveness. Arabi et al. (2007) and Karamouz et al. (2015) concluded that BMP effectiveness evaluation could be ascertained with high confidence using the soil and water assessment tool (SWAT) model. Lee et al. (2020) observed that the uncertainty associated with the estimated BMP effectiveness was relatively large, and the BMP effectiveness should be quantified using multiple parameter sets. Similar studies (Tzyy-woei et al. 2013; Woznicki et al. 2014; Taylor et al. 2016; Tasdighi et al. 2018) also identified that there was a relatively large degree of uncertainty associated with nutrient loss simulations or BMP effectiveness.

Furthermore, existing studies have mainly focused on the uncertainty in the annual total nutrient loads, rather than in their temporal distribution. However, temporal analysis of uncertainty on a monthly or seasonal bias can give insight into the risk and reliability of the BMP (Woznicki et al. 2003; Schuwirth et al. 2019). Meanwhile, SWAT was is the most common used model for the uncertainty evaluation of BMP effectiveness. Annualized agricultural non-point source pollution (AnnAGNPS; USDA ARS, 2006) is a semi-distributed and physically-based water quality model, and it has been commonly used to evaluate BMP effectiveness (e.g., Qi and Altinakar, 2011; Zhang et al. 2020). Therefore, this study aims to 1) evaluate the parameter uncertainty in BMP effectiveness, 2) examine the temporal variation of the parameter uncertainty in BMP effectiveness, and 3) test AnnAGNPS for the uncertainty evaluation of BMP effectiveness. The findings of this study can improve the reliability of BMP effectiveness and also can be used as a reference for alleviating non-point sources pollution in reservoir watersheds.

2 Study Area and Methodology

2.1 Study area and dataset

This study was conducted in the Shanmei Reservoir watershed, which was located in the upper region of the Jinjiang watershed in southeastern China (Fig. 1a). It covers an area of approximately 1,023 km^2. This study area is dominated by a subtropical monsoon climate, with an average annual temperature and precipitation of 20°C and 1600 mm, respectively. The dry and wet seasons are very distinct, with 80% of the annual precipitation occurring during the wet season (from March to September). The major land use types are forest (accounting for 68.34% of the area), followed by agricultural land (including orchards and farmlands, 22.64%), and urban areas (4.89%, Fig. 1b). The most common grown crops are ponkan, tea trees, paddy rice, sweet potatoes, and vegetables. The dominant hydrologic soil group (USDA, 1972) which refers to the classification of soils based on their runoff, producing characteristics and their infiltration rate, is D, followed by groups C and B (Fig. 1c).

The Shanmei Reservoir is the most important potable water resource for Quanzhou city, Fujian province, China. It supplies water for more than 6,000,000 residents in downstream areas. However, rapid population growth and agricultural development have resulted in high TN concentrations in the Shanmei Reservoir, thereby leading to eutrophication and deterioration of the reservoir water quality in recent years. Therefore, the implementation of BMPs in the upland watershed of the Shanmei Reservoir is urgently required to reduce or alleviate TN load and risk of eutrophication.

The input data for AnnAGNPS includes the digital elevation model (DEM), land use, soil, agricultural management schedule, and meteorological and hydrological data. Table 1 summarizes the input data and sources.

Table 1 Input data and sources of the AnnAGNPS model in Shanmei Reservoir watershed

Data	Description	Sources
DEM	A digital elevation model with 30 m × 30 m	International Scientific Data Platform of the Chinese Academy of Sciences (http://datamiffor.csdb.cn/admin/ datademMain/jsp)
Soil map	The digital soil map at 1:500,000 scale	Soil Fertilizer Laboratory of Fujian Province
Land use	Land use data in 1995 for the Taoxi sub-basin and the land use data in 2006 for the Shanmei Reservoir watershed	Landsat Thematic Mapper data by unsupervised classification and visual interpretation
Climate	Daily air temperature (maximum and minimum), relative humidity, wind speed, and cloud cover during 2000–2010 at Yongchun and Dehua stations	Meteorology Agency of Fujian Province
Precipitation	Daily precipitation from 16 precipitation stations during 2000–2010	Water Conservation Agency of Fujian Province
Runoff	Daily runoff data during 1995–1997 at Yongchun Station; Daily runoff during 2000-2010 at Shanmei Reservoir Station	Water Conservation Agency of Fujian Province
Sediment	Daily sediment form April 1995 to October 1995 and from April 1996 to October 1996 at Yongchun Station	Water Conservation Agency of Fujian Province
Water quality	Quarterly TN concentration data at the inlet of Shanmei Reservoir during 2002–2010	Environmental Protection Agency of Fujian Province
Point sources	Waste water discharge, TN concentrations	Environmental Protection Agency of Quanzhou Municipality
Agricultural management	Fertilizer application rates, livestock production (2000–2010)	Annual census published by local statistical bureau, and interviews with local farmers

2.2 AnnAGNPS model

AnnAGNPS represents the spatial heterogeneities of watershed properties by dividing the watershed into drainage areas (called cells) with homogenous land use, soil, etc. The cells are connected to each other by the network of channels and reaches, where runoff, sediment, and nutrients are transported. AnnAGNPS is built as a series of interconnected modules by integrating both empirical and quasi-physical models that simulate watershed processes (Villamizar and Brown, 2016). The Soil Conservation Service curve number (SCS-CN) is used to estimate the surface runoff (USDA, 1986). The Revised Universal Soil Loss Equation (RULSE) is used to estimate the daily sheet and rill erosion (Renard et al. 1997), whereas the Bagnold equation and a modified Einstein equation are used to express sediment transport into the stream system (Bingner et al. 2011). For the N simulation, bias mass conservation equilibrium is employed to determine nutrient generation for rainfall events. Nutrient dynamics include the fate of N, P, and organic carbons are determined by the Erosion Productivity Impact Calculator (Williams, 1995) and the Groundwater Loading Effects of Agricultural Management Systems (Leonard et al. 1995) algorithms.

2.3 Parameter sensitive analysis method

A sensitivity analysis (SA) method is required to select sensitive parameters for calibration. AnnAGNPS requires over 400 input parameters distributed across 34 modules (Bingner et al. 2011). In this study, the Morris screening method (Morris,

1991) and the Differential Sensitivity Analysis (DSA) method (Lenhart et al. 2002) were applied to identify the most important hydrological parameters that control the variability of runoff and sediment, and TN simulations, respectively.

Morris relies on a "one-factor-at-a-time" design of experiments. It perturbs one parameter at a time and measures the change in the model output. The Morris method evaluates a graphical representation of μ^* vs. σ to determine the most important parameters. The μ^* estimates the overall effect of each input on the output, while σ estimates the higher-order effects, such as nonlinearity and interactions between inputs.

DSA was selected for its simplicity and low computational time compared to other SA methods. DSA is calculated at one point in the parameter space by adjusting the parameter with a value or percentage (Δx) while the other parameters remain constant. In this study, the TN parameter sensitivity analyses were conducted by varying the initial parameter by a fixed percentage of $\Delta x = \pm 10\%$. The sensitivity index (I) (Lenhart et al. 2002) was computed to classify inputs parameters according to their sensitivity. The details of DSA can be found in Lenhart et al. (2002).

In this study, 16 parameters (Table 2) were selected to identify the sensitive parameters of AnnAGNPS in predicting runoff, sediment load and TN load. They have been identified as sensitive in previous studies (e.g., Chahor et al. 2014; Bisantino et al. 2015; Luo et al. 2015; Abdelwahab et al. 2016).

Table 2 Information of AnnAGNPS parameters

Category	Parameter Symbol (in this paper)	Description	Initial range		TYPE
			Min.	Max.	
Hydrology Parameters	CN*	Runoff curve number	−0.3	0.30	a
	RMN	Riverway Manning coefficient	0.01	0.05	b
	FC	Field moisture capacity	−0.3	0.30	a
	WP	Wilting coefficient	−0.3	0.30	a
	K	Soil erodibility factor K	−0.3	0.30	a
	RC	Remaining residue cover	−0.3	0.30	a
	RB	Root biomass	−0.3	0.30	a
	LS	LS factor	−0.3	0.30	a
	P	Water soil conservation factor	−0.3	0.30	a
	CC	Canopy cover	−0.3	0.15	a
	RCT	Remaining residue cover after tillage	−0.3	0.15	a
	RR	Random roughness	−0.3	0.15	a
Water quality parameters	FR	Fertilizing amount	-	-	a
	FD	Fertilizing depth	-	-	a
	NHL	Nitrogen half-life	-	-	a
	SBV	soil background value	-	-	a

Note: a indicates parameters whose baseline values are adjusted during the calibration by the multiplier sampled from the bound. b indicates parameters whose baseline values are replaced during the calibration by the value sampled from the bound.

Sediment observations were only available at the upstream Yongchun hydrological station, which was located at the outlet of the Taoxi sub-basin. The sensitive parameters of the runoff and sediment were identified based on daily runoff and sediment data at the Yongchun Station (Fig. 1). In addition, the sensitive parameters of TN were identified based on the observed TN data at the Shanmei Reservoir.

2.4 AnnAGNPS calibration and validation

In this study, AnnAGNPS was initiated with the calibration of runoff, followed by calibration of the sediment load, and finally, the TN load. Based on the identified sensitive parameters, calibration was performed using the dynamically dimensioned search (DDS) algorithm (Tolson and Shoemaker, 2007).

Multiple statistical values, including the Nash–Sutcliffe efficiency coefficient (*NSE*) (Nash and Sutcliffe, 1970), coefficient of determination (R^2), and percent bias (PBIAS) (Gupta et al. 1999), were selected to evaluate the monthly and daily model performances. Table 3 lists the calibration and validation periods. Because TN data were only available at the quarterly time scale from 2001 to 2010, the data were not split into calibration and validation periods.

2.5 BMP representation in the AnnAGNPS model

Based on the local socioeconomic conditions and pollution characteristics, four potential BMPs were considered to evaluate their effectiveness in reducing TN load under the current climate conditions including RB, FR, NT and PT. Under each scenario, BMP were applied to agricultural land in the Shanmei Reservoir watershed.

RB is a vegetated band set up between the agricultural land and the riverway to trap sediments and nutrients from the drained area to the channel segment. Based on the results of Haycock and Pinay (1993) and Lee et al. (2003), a grassland buffer width was set to 6m.

FR is mainly intended to reduce available N quantities on the ground for transporting to water to limit N loss. This BMP was directly implemented in

AnnAGNPS by reducing 30% of the total fertilization of N (Xing and Zhu, 2002).

Tillage practices influence the physical and biological properties of soil. The goal of **NT** is to reduce erosion caused by soil-disturbing activities. Under the NT scenario, the management operations for the agricultural land cells were modified to no till, and the value of RUSLE P factor was reduced by 10% (Abdelwahab et al. 2016).

PT is an earth embankment, or a combination ridge and channel constructed across the slope. It serves to decrease surface runoff volume, peak runoff rate, sheet and rill erosion, and erosive power of runoff, and also reduces the development of rills and gullies. Under the PT scenario, CN was reduced by six, and RUSLE_P and slope length were changed depending on the slope class according to Arabi et al. (2008).

2.6 Uncertainty analysis of BMPs effectiveness

Firstly, uncertainty analysis was accomplished based on the selection of behavioral parameter groups, which could cause equifinality in the model results. The calibrated parameter values were considered as the initial conditions. 1000 parameter groups within ±30 % of the parameter values were generated using the Latin hypercube sampling (LHS) method (Iman and Conover, 1980), and the related objective functions (i.e., *NSE*, R^2, and PBIAS) were calculated. Then, the behavioral parameter groups were selected with the following criteria: *NSE* > 0.70 and PBIAS < 10% for daily discharge, *NSE* > 0.4 and PBAIS < 30% for TN load. The threshold values for the objective functions were determined based on the results of the best set of calibrated parameter groups (Pfannerstill et al. 2014; Haas et al. 2016). In total, ten behavioral parameter groups were identified.

Secondly, for the ten parameter groups, AnnAGNPS were run for four BMPs and the baseline scenario using the current meteorological data (2000–2010), totaling up to 50 simulations. The TN reduction rate was quantified using Equation (1) to evaluate BMP effectiveness. Then, the degree of data discretization was evaluated among the ten parameter groups based on the coefficient of variation (CV) to evaluate the impact of parameter uncertainty on BMP effectiveness. Equation (1) can be expressed as follows:

$$r_i = \frac{y_{i,post-bmp} - y_{i,no-bmp}}{Y_{i,no-bmp}} \quad (1)$$

Where r_i is reduction rate for TN load for the parameter set i, $y_{i,-\mathrm{no}BMP}$ and $y_{i,post-BMP}$ are the simulated TN loads for parameter set i, before and after application of the BMP, respectively.

Additionally, the uncertainty of the seasonal BMP effectiveness. was completed using the following month groups: December–January–February (DJF), March–April–May (MAM), June–July–August (JJA), and September–October–November (SON). Seasonal CV was applied to understand the uncertainty of the BMP effectiveness varying at the seasonal scale.

3 Results and Discussion

3.1 SA and model performance

The SA results of the runoff , sediment and TN simulations are shown in Fig. 2. Based on the Morris results, the most sensitive flow parameter was CN with a μ* value of

0.0113, followed by RMN with a μ* value of 0.0034, while CN and RMN also had strong interactive effects with other parameters (Fig. 2a). In addition, RCT, LS, P, K, RMN, and K showed an important effect on the sediment simulations (Fig. 2b). Fig. 2c shows the results of DSA for the TN parameters. The sensitivity indices of FR, FD, NHL, and SBV parameters were 0.07, 0.001, 0.0014, and 0.45, respectively. Parameters with sensitivity index I less than 0.05 were classified as small to negligible sensitive parameters (Lenhart et al. 2002). Thus, the sensitive parameters of the Shanmei Reservoir watershed were CN, RMN, RCT, LS, P, K, FR, and SBV.

Table 3 Results of the AnnAGNPS model in calibration (Cal) and validation (Val) for runoff, sediment, and TN loads

Gauge	Variable	Cali period	Vali period	Time-step	*NSE*		R^2		PBIAS %	
					Cali	Vali	Cali	Vali	Cali	Vali
Yongcun	Streamflow	1995-1996	1997	Daily	0.73	0.77	0.79	0.77	-3.77	-4.31
				Monthly	0.94	0.93	0.95	0.94	-7.33	-4.31
	Sediment	1995.4.1-1995.10.31	1996.4.1-1996.10.31	Daily	0.73	0.68	0.74	0.69	-11.62	-9.65
				Monthly	0.9	0.96	0.95	0.97	-11.62	-9.65
Shanmei	Streamflow	2001-2005	2006-2010	Daily	0.74	0.83	0.83	0.87	-2.9	-2.7
				Monthly	0.91	0.97	-0.05	-0.02	-5.32	-9.37
	TN	2001-2010		Daily	0.4		0.47		23.4	

Table 3 shows the performances of AnnAGNPS with regard to runoff and sediment load in the Taoxi sub-basin, as well as runoff and TN load in the Shanmei Reservoir watershed. For the simulated runoff at the Shanmei Reservoir watershed and Taoxi sub-basin, *NSE* and R^2 were more than 0.70, and PBIAS was less than 5% at the daily scale, while *NSE* and R^2 were greater than 0.90, and PBIAS was less than 10% at the monthly scale for the calibration and validation periods. As for simulated sediment at the Taxi sub-basin, *NSE* and R^2 were higher than 0.60, and PBIAS was less than 15%

at the daily scale, while *NSE* and R^2 were greater than 0.90, and PBIAS was less than 15% at the monthly scale for both the calibration and validation periods. For TN in the Shanmei Reservoir watershed, *NSE*, R^2, and PBIAS were 0.42, 0.47, and 23.4%, respectively. According to Moriasi et al. (2007), model simulation evaluated at a monthly time step can be satisfactory if $NSE > 0.50$, and if PBIAS is within ±25% for streamflow, within ± 25% for sediment, and within ± 70% for N. AnnAGNPS produced satisfactory results for runoff and sediment simulations, while for TN simulations the result is less good partly because the number of observations for TN is limited (Table 3). However, the criteria set out by Moriasi et al. (2007) were on a monthly time scale. Generally, a relaxation of performance criteria is warranted when the evaluation time step decreases (Engebretsen et al. 2019). Moreover, a visual inspection of Fig. 3 reveals that the model could perform reasonably well in the simulation of temporal variation and average magnitude of TN, except for some pairs of simulated-measured TN loads (i.e., March 2, 2004; May 9, 2005), in which the AnnAGNPS simulated TN loads greatly deviated from the observed values. Additionally, the objective of this study was to evaluate the long-term average impact of different BMPs. Therefore, the model performed adequately in simulating TN load.

The weak performance of TN simulation may be attributed to the quality and relevance of its input factors, such as 1) inadequate frequency of TN monitoring data (i.e., the quarterly TN data) to calibrate and evaluate model performance, and 2) simplified management practices (i.e., the time of planting and fertilizer application rates). In the study, the time of planting and fertilizer application rates were regarded as spatially uniform in the watershed, and were determined using an average value based on the recorded data of actual practices during the study period. Witing and Volk (2013) identified that nutrient simulations were highly sensitive to input data on crop rotations based on SWAT modeling. Sahu and Gu (2009) and Motsinger et al. (2016) observed that the actual non-uniformity of fertilizer application rate and timing could have a greater impact than can be represented in the model. This may also be because of the model structure. The nutrient load is based on mass conservation in AnnAGNPS, and any missing input or output information of nutrients in the watershed will considerably affect the results (Shamshad et al. 2008). According to Bingner et al. (2011), the model assumed that there was no tracking of nutrients from one day to the next, implying that there would definitely be loss. That is, R^2=1 for nutrient loading is not possible (Shamshad et al. 2008).

3.2 Impact of parameter uncertainty on BMP effectiveness

Ten parameter groups with equifinality were identified based on the threshold of the objective functions. It was observed that the sensitive parameters were all evenly distributed within the sample range, except for CN (0 % – 10 %). It indicated that these sensitive parameters were highly uncertain. However, ten simulations were identified to have a similar model performance with the best parameter set identified in the model calibration. Compared to the best model performance in the model calibration, the differences in the values of NSE and R^2 of the ten simulations were both within 0.02, for runoff and TN load simulations, and the differences in the absolute values of *PBIAS* were within 3 % and 7 % for runoff and TN load simulation, respectively. In addition, ten parameter groups were run for the uncertainty analysis to evaluate the uncertainty associated with both model simulations and BMP effectiveness.

Table 4 Annual TN load and reduction efficiency for each BMP

BMPs	TN load（kg/ha/year）			Reduction rate of TN load（%）		
	Mean	range	CV	Mean	range	CV
Baseline	33.58	[25.16, 42.14]	18	-	-	-
RB	32.61	[24.59, 40.93]	17	2.82	[1.38, 4.20]	35
FR	30.11	[22.57, 36.59]	17	10.24	[6.17, 13.17]	20
NT	29.45	[22.36, 36.77]	16	11.96	[5.23, 18.63]	39
PT	26.23	[20.48, 32.56]	15	21.41	[14.53, 31.93]	25

3.2.1 Impact of parameter uncertainty on BMP effectiveness at annual scale

Table 4 lists the results of the average annual TN load and reduction rate under the baseline and the four BMP scenarios (RB, FR, NT, and PT). The results showed that simulated uncertainty bounds associated with the mean annual TN load were relatively large. For example, the mean annual TN loss ranged from 25.16 to 42.14 kg ha^{-1} yr^{-1} with an average of 33.58 kg ha^{-1} yr^{-1} under the baseline scenario. Similar results were observed for the four BMPs. It indicated that the uncertainty in the simulations of TN load cannot be ignored, even if the differences between the objective functions for each parameter group were small. This was similar to the results obtained by Arabi et al. (2007), Shen et al. (2008), and Woznicki and Nejadhashemi (2014). Among the four BMPs, the CVs of annual TN load varied from 15 % to 18 %. It implied that the impacts of parameter uncertainty on the TN load were similar among the four BMPs. Additionally, there were no obvious differences between the CV of annual TN load for each BMP scenario and that of the baseline scenario. However, according to Karamouz et al. (2015), BMP implementation effectively reduced the uncertainty in the prediction of watershed phosphorous load.

Relative to the baseline scenario, the RB, FR, NT, and PT scenarios can reduce the mean annual TN load by 2.82%, 10.24%, 11.96%, and 21.41%, respectively. The corresponding gaps in the reduction rate between the maximum and minimum values for the RB, FT, NT, and PT scenarios were 2.82%, 7%, 13.4%, and 17.4%, respectively. This suggested that the ranges of the reduction rates of the four BMPs derived from ten parameter groups were relatively wide. Moreover, the CV of the reduction rate of TN varied significantly (25 % to 35 %) among the four BMPs. The CV associated with the reduction rate of TN load was larger than that corresponding to the absolute prediction. For example, the CV of the mean annual TN loss was 16%, while the CV of the reduction rate increased to 39% under the NT scenario. This implied that BMP effectiveness did not translate into low variability. The uncertainty of the annual reduction rate of the TN load was larger than that of the TN load predictions. Lee et al. (2020) also identified that the effectiveness of riparian buffers in reducing annual total organic nitrogen differed by up to 30% based on the selection of acceptable parameter sets. However, Arabi et al. (2006) identified that the uncertainty associated with the estimated reduction rate of the TN load for BMPs was substantially smaller than the uncertainty associated with absolute predictions based on the SWAT model.

3.2.2 Impact of parameter uncertainty on BMP effectiveness at seasonal scale

The seasonal average TN load and reduction rate for each BMP scenario are shown in Fig. 5. The highest average TN load was observed in summer, while the lowest TN load was observed in winter. Summer was the period of frequent rainstorm events and continuous precipitation during agricultural management activities accounted for 49.5% of total rainfall, which was more likely to cause eutrophication in aquatic ecosystems. Similarly, the highest reduction rate of TN load was also observed in summer, while the lowest reduction rate was observed in winter, except for the FR scenario. This was because the reduction rate of the TN load was always high during the months when the TN load was large. This suggested that the implementation of BMP can improve the watershed's ability to reduce nutrient loads during summer, lowering the risk of eutrophic pollution in rivers. Similar results were obtained by Jiang et al. (2019), who observed that the implementation of riparian forest buffers achieved the highest reduction rate of TN during the wet and transition periods. However, for the FR scenario, the highest TN load was observed in summer, and the highest reduction rate occurred in spring in the current study. In spring, major crops (i.e., ponkan, tea trees, paddy rice, and vegetables) were in the sowing period in the study area, and relatively larger quantities of fertilizers are applied during this period. Under the FR scenario, a 30% reduction in fertilizer amount was adopted. Therefore, a greater reduction in fertilization causes a higher reduction rate of the TN load in spring.

Fig. 5b shows that the CVs of the reduction rate of TN load vary among seasons, and the variation in CVs differs among BMPs. Under the RB and NT scenarios, the CVs were the largest in spring and lowest in summer. Under the FR scenario, the CVs were the largest in summer and lowest in winter. Under the PT scenario, the CVs were the largest in winter and lowest in summer. However, for the majority of BMPs, the uncertainty of BMP effectiveness in reducing the TN load was the lowest in summer. This implied that the impact of BMPs on reducing TN load during summer would be considerably effective with a higher reduction rate and lower uncertainty. Moreover, similar to the results at the annual scale, the CVs associated with reduction rates of TN load were larger than those corresponding to the absolute predictions for each season.

In addition, there were large differences in the reducing efficiency of the TN load and the related uncertainty among the four BMPs. PT was identified to be the most effective management practice for the study area in terms of the highest reduction efficiency and relatively low uncertainty. This was consistent with the results at the annual scale. NT, which did not alter the landscape in a significant manner, had the highest uncertainty in the reduction efficiency of TN load. This suggested that the parameters pertaining to NT show higher sensitivity when quantifying the reduction efficiency. Under the NT scenario, the related parameters, including RCT, RR, and USLE_P, changed. The higher uncertainty of the BMP effectiveness in NT may be linked to the higher level of RCT sensitivity in sediment modeling (Fig. 2).

However, these BMP scenarios reflected the subjective assumptions of the modeler

(Schürz et al. 2019). Technical aspects, such as how to represent the BMPs in the model, involve highly subjective assumptions and thus introduce a new additional uncertainty (Tasdighi et al. 2018; Schürz et al. 2019). These assumptions can influence the simulation of BMP effectiveness. The potential uncertainties in the scenario development are essential for evaluating the BMP effectiveness.

4 Conclusions

This study evaluated the uncertainty of the BMP effectiveness in reducing TN load owing to hydrological parameter uncertainty at the annual and seasonal time scales by using AnnAGNPS. Four potentially feasible BMPs and ten behavioral parameter groups were adopted in uncertainty analysis. The following conclusions were drawn from this study.

(1) The uncertainty of the BMP effectiveness evaluation in reducing the TN load was larger than the uncertainty associated with absolute predictions for both the annual and seasonal time scales. This suggested that BMP effectiveness cannot be ascertained with high confidence using a modeling approach because of the model parameter uncertainty, and BMP effectiveness should be quantified based on the use of similar parameter sets to improve confidence in model predictions.

(2) The reduction efficiency was high in summer when frequent rainstorm events and continuous precipitation cause large quantities of surface runoff to carry considerable TN loads. The uncertainty of BMP effectiveness varied seasonally, and was always lower in summer for most BMPs. This indicates that the impact of BMPs

on reducing TN loads in summer would be considerably more effective with a higher reduction rate and lower uncertainty.

(3) There were significant differences in the reducing efficiency of the TN load and the related uncertainty among the four BMPs. PT was identified to be the most effective management practice for the study area in terms of the highest reduction efficiency and lower uncertainty.

This work is a new attempt on exploring the uncertainty of BMP effectiveness for reducing TN load at annual and seasonal scale using AnnAGNPS. The results of the BMP effectiveness for reducing TN load and the uncertainty bounds provided in the present study can fill the gap in BMP effectiveness with respect to the Shanmei Reservoir watershed. Extended studies should be undertaken to further evaluate the uncertainty of BMP effectiveness driven by different climatic conditions as well as other distributed hydrological models.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2018YFE0206400), the Scientific Project from Fujian Provincial Department of Science and Technology (2019R1002-3 and 2021R1002006), and the Scientific Project from Fujian Key Laboratory of Severe Weather (2020KFKT01).

Declarations

Ethics Approval Not applicable, because this article does not contain any studies with human or animal subjects.

Consent to Participate Not applicable, because this article does not contain any studies with human or animal subjects.

Competing Interests The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Abdelwahab O, Bingner RL, Milillo F, Gentile F (2016) Evaluation of alternative management practices with the AnnAGNPS model in the Carapelle Watershed. Soil Sci 181(7):293-305. https://doi.org/10.1097/ss.0000000000000162

Arabi M, Frankenberger JR, Engel BA, Arnold JG (2008) Representation of agricultural conservation practices with SWAT. Hydrol Process 22(16):3042-3055. https://doi.org/10.1002/hyp.6890

Arabi M, Rao SG, Hantush MM (2007) A probabilistic approach for analysis of uncertainty in the evaluation of watershed management practices. J Hydrol 333(2-4): 459-471. https://doi.org/10.1016/j.jhydrol.2006.09.012

Bai X, Shen W, Wang P, Chen X, He Y (2020) Response of Non-point Source Pollution Loads to Land Use Change under Different Precipitation Scenarios from a Future Perspective. Water Resour Manag 34:3987-4002. https://doi.org/10.1007/s11269-020-02626-0

Beven K (2006) A manifesto for the equifinality thesis. J Hydrol 320(1-2):18-36.

https://doi.org/10.1016/j.jhydrol.2005.07.007

Bingner RL, Theurer FD, Yuan Y (2011) AnnAGNPS technical processes documentation, version 5.2. USDA-ARS National Sedimentation Laboratory

Bisantino T, Bingner R, Chouaib W, Gentile F, Trisorio Liuzzi G (2015) Estimation of runoff, peak discharge and sediment load at the event scale in a medium-size Mediterranean watershed using the AnnAGNPS model. Land Degrad Dev 26(4): 340-355. https://doi.org/10.1002/ldr.2213

Casallas-Ojeda M, Soto-Paz J, Alfonso-Morales W, Oviedo-Ocaña1 ER, Komilis D (2021) Optimization of Operational Parameters during Anaerobic Co-digestion of Food and Garden Waste. Environ Process 8:769-791. https://doi.org/10.1007/s40710-021-00506-2

Chahor Y, Casalí J, Giménez R, Bingner RL, Campo MA, Goñi M (2014) Evaluation of the AnnAGNPS model for predicting runoff and sediment yield in a small Mediterranean agricultural watershed in Navarre (Spain). Agr Water Manage 134:24-37. https://doi.org/10.1016/j.agwat.2013.11.014

Engebretsen A, Vogt RD, Bechmann M (2019) SWAT model uncertainties and cumulative probability for decreased phosphorus loading by agricultural Best Management Practices. Catena 175:154-166. https://doi.org/10.1016/j.catena.2018.12.004

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. https://doi.org/10.1061/(asce)1084-0699(1999)4:2(135)

Haas MB, Guse B, Pfannerstill M, Fohrer N (2016) A joined multi-metric calibration of river discharge and nitrate loads with different performance measures. J Hydrol 536:534-545. https://doi.org/10.1016/j.jhydrol.2016.03.001

Haycock NE, Pinay G (1993) Groundwater Nitrate Dynamics in Grass and Poplar Vegetated Riparian Buffer Strips during the Winter. J Environ Qual 22(2):273-278. https://doi.org/10.2134/jeq1993.00472425002200020007x

Iman RL, Conover WJ (1980). Small sample sensitivity analysis techniques for computer models with an application to risk assessment. Commun Stat-Theor M 9(17):1749-1842. https://doi.org/10.1080/03610928008827996

Karamouz M, Taheriyoun M, Seyedabadi M, Nazif S (2015) Uncertainty based analysis of the impact of watershed phosphorus load on reservoir phosphorus concentration. J Hydrol 521:533-542. https://doi.org/10.1016/j.jhydrol.2014.12.028

Lee K H, Isenhart T M, Schultz RC (2003) Sediment and nutrient removal in an established multi-species riparian buffer. J Soil Water Conserv 58(1):1-7.

Lee S, McCarty GW, Moglen GE, Li X, Wallace CW (2020) Assessing the effectiveness of riparian buffers for reducing organic nitrogen loads in the Coastal Plain of the Chesapeake Bay watershed using a watershed model. J Hydrol 585:124779. https://doi.org/10.1016/j.jhydrol.2020.124779

Lenhart T, Eckhardt K, Fohrer N, Frede HG (2002) Comparison of two different approaches of sensitivity analysis. Phys Chem Earth 27(9):645-654. https://doi.org/10.1016/s1474-7065(02)00049-9

Leonard RA, Knisel WG, Davis FM (1995) Modelling pesticide fate with GLEAMS. Eur J Agron 4(4):485-490. https://doi.org/10.1016/s1161-0301(14)80100-7

Moriasi DN, Arnold JG, Liew M, Bingner RL, Harmel RD, Veith TL (2007) Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. T ASABE 50(3):885-900. https://doi.org/10.13031/2013.23153

Morris MD (1991) Factorial sampling plans for preliminary computational experiments. Technometrics 33(2):161-174. https://doi.org/10.1080/00401706.1991.10484804

Motsinger J, Kalita P, Bhattarai R (2016) Analysis of best management practices implementation on water quality using the soil and water assessment tool. Water 8(4):145. https://doi.org/10.3390/w8040145

Nash J E, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—A discussion of principles. J Hydrol 10(3):282-290. https://doi.org/10.1016/0022-1694(70)90255-6

Ni X, Parajuli PB, Ouyang Y (2020). Assessing Agriculture Conservation Practice Impacts on Groundwater Levels at Watershed Scale. Water Resour Manag 34:1553-1566. https://doi:10.1007/s11269-020-02526-3

Pfannerstill M, Guse B, Fohrer N (2014). Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. J Hydrol 510: 447-458. https://doi.org/10.1016/j.jhydrol.2013.12.044

Qi H, Altinakar MS (2011) Vegetation buffer strips design using an optimization approach for non-point source pollutant control of an agricultural watershed.

Water Resour Manag 25(2):565-578. https://doi.org/10.1007/s11269-010-9714-9

Renard KG, Foster GR, Weesies GA, McCool DK, Yoder DC (1997) Predicting soil erosion by water: A guide to conservation planning with the Revised Universal Soil Loss Equation (RUSLE). Agriculture handbook, 703:25-28.

Rodrigues A, Silv D, Filho F (2021). Methodology for allocation of best management practices integrated with the urban landscape. Water Resour Manag 35: 1353–1371. https://doi.org/10.1007/s11269-021-02791-w

Sahu M, Gu RR (2009) Modeling the effects of riparian buffer zone and contour strips on stream water quality. Ecol Eng 35(8):1167-1177. https://doi.org/10.1016/j.ecoleng.2009.03.015

Schürz C, Hollosi B, Matulla C, Pressl A, Ertl T, Schulz K, Mehdi B (2019) A comprehensive sensitivity and uncertainty analysis for discharge and nitrate-nitrogen loads involving multiple discrete model inputs under future changing conditions. Hydrol Earth Syst Sc 23(3):1211-1244. https://doi.org/10.5194/hess-23-1211-2019

Schuwirth N, Borgwardt F, Domisch S, Friedrichs M, Kattwinkel M, Kneis D, Kuemmerlen M, Langhans SD, Martínez-López J, Vermeiren P (2019) How to make ecological models useful for environmental management. Ecol Model 411: 108784. https://doi.org/10.1016/j.ecolmodel.2019.108784

Shamshad A, Leow CS, Ramlah A, Hussin W, Sanusi S (2008) Applications of AnnAGNPS model for soil loss estimation and nutrient loading for Malaysian conditions. Int J Appl Earth Obs 10(3):239-252.

https://doi.org/10.1016/j.jag.2007.10.006

Shen Z, Qian H, Hong Y, 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. https://doi.org/10.1016/j.scitotenv.2008.06.009

Taylor SD, Yi H, Hiscock KM (2016) Modelling the impacts of agricultural management practices on river water quality in Eastern England. J Environ Manage 180:147-163. https://doi.org/10.1016/j.jenvman.2016.05.002

Tasdighi A, Arabi M, Harmel D, Line D (2018) A Bayesian total uncertainty analysis framework for assessment of management practices using watershed models Environ Modell Softw 108:240-252. https://doi.org/10.1016/j.envsoft.2018.08.006

Tolson BA, Shoemaker CA. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. Water Resour Res 43(1): 208-214. https://doi.org/10.1029/2005wr004723

Tzyy-woei C, Adel S, Yu-cheng L, Yu-chi H (2013) BMP evaluation for nutrient control in a subtropical reservoir watershed using SWAT Model. In 2013 the International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE 2013) (pp. 912-915). Atlantis Press.

USDA ARS (2006) AnnAGNPS—Annualized Agricultural Non-point Source Pollution Model. USDA Agricultural Research Service, Washington DC.

USEPA (2008) Reducing stormwater costs through low impact development (LID) strategies and practices. U.S. Environmental Protection Agency (USEPA), Washington, D.C.

USDA (1972) National engineering handbook, section 4: Hydrology. Washington, DC.

USDA (1986) Technical Release 55: Urban Hydrology for Small Watershed. NRCS-USDA, Washington, DC.

Villamizar ML, Brown CD (2016). Modelling triazines in the valley of the River Cauca, Colombia, using the annualized agricultural non-point source pollution model. Agr Water Manage 177:24-36. https://doi.org/10.1016/j.agwat.2016.06.010

Williams JR (1995) The EPIC model. In Computer Models of Watershed Hydrology, Singh VP (ed). Louisiana State University: Baton Rouge, LA, 909–1000.

Witing F, Volk M (2013) Investigating SWAT model sensitivity to agricultural land cover and crop rotation parameterizations. 2013 SWAT Conference. Toulouse, France.A

Woznicki SA, Nejadhashemi AP (2014) Assessing uncertainty in best management practice effectiveness under future climate scenarios. Hydrol Process 28(4):2550-2566. https://doi.org/10.1002/hyp.9804

Xing GX, Zhu ZL. (2002) Regional nitrogen budgets for China and its major watersheds. Biogeochemistry 57(1):405-427.

Zhang M, Francis RA, Chadwick MA (2021). Nutrient Dynamics at the Sediment-Water Interface: Influence of Wastewater Effluents. Environ Proces

8: 1337–1357. https://doi.org/10.1007/s40710-021-00540-0

Zhang T, Yang Y, Ni J, Xie D (2020) Best management practices for agricultural non-point source pollution in a small watershed based on the AnnAGNPS model. Soil Use Manage 36(1):45-57. https://doi.org/10.1016/j.agwat.2012.10.018