

# Water Resources Research®

## RESEARCH ARTICLE

10.1029/2022WR033673

### Key Points:

- A novel method to utilize low-cost high-frequency in-situ sensor data in watershed water quality modeling was developed
- With proper error treatment, the new method can extract meaningful information from in-situ sensor data and improve modeling accuracy
- Assimilating 1 yr of subhourly EC sensor data reduced the modeling uncertainty and management cost by 70%

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

Y. Zheng,  
zhengy@sustech.edu.cn

### Citation:

Han, F., Hu, Z., Chen, N., Wang, Y., Jiang, J., & Zheng, Y. (2023). Assimilating low-cost high-frequency sensor data in watershed water quality modeling: A Bayesian approach. *Water Resources Research*, 59, e2022WR033673. <https://doi.org/10.1029/2022WR033673>

Received 13 SEP 2022

Accepted 2 APR 2023

### Author Contributions:

**Conceptualization:** Feng Han  
**Formal analysis:** Feng Han, Zhaoping Hu  
**Funding acquisition:** Yi Zheng  
**Investigation:** Zhaoping Hu, Yao Wang, Jiping Jiang  
**Methodology:** Feng Han, Zhaoping Hu, Yi Zheng  
**Resources:** Nengwang Chen  
**Supervision:** Nengwang Chen, Yi Zheng  
**Validation:** Nengwang Chen, Yao Wang, Jiping Jiang  
**Visualization:** Zhaoping Hu  
**Writing – original draft:** Feng Han  
**Writing – review & editing:** Feng Han, Yi Zheng

## Assimilating Low-Cost High-Frequency Sensor Data in Watershed Water Quality Modeling: A Bayesian Approach

Feng Han<sup>1,2</sup>, Zhaoping Hu<sup>1</sup>, Nengwang Chen<sup>3</sup> , Yao Wang<sup>3</sup>, Jiping Jiang<sup>1,2</sup> , and Yi Zheng<sup>1,2</sup> 

<sup>1</sup>School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China,

<sup>2</sup>Shenzhen Municipal Engineering Lab of Environmental IoT Technologies, Southern University of Science and Technology, Shenzhen, China, <sup>3</sup>Fujian Provincial Key Laboratory for Coastal Ecology and Environmental Studies, College of the Environment and Ecology, Xiamen University, Xiamen, China

**Abstract** Uncertainty reduction in watershed water quality (WWQ) modeling remains a major challenge. One important reason is the lack of sufficient available water quality observations because traditional laboratory analysis of water samples has high labor, financial and time costs. Low-cost high-frequency water quality data from in-situ sensors provide an opportunity to solve this problem. However, long-term sensing in complex natural environments usually suffers more significant errors. This study aimed to develop a novel method to utilize in-situ sensor data in WWQ modeling, namely, the Bayesian calibration using multisource observations (BCMSO), which can simultaneously assimilate laboratory-based observations and in-situ sensor data. Both synthetic and real-world cases of nitrate modeling were used to demonstrate the methodology, and the Soil and Water Assessment Tool was employed as the WWQ model. The results indicated that direct assimilation of sensor data using traditional Bayesian calibration generated obvious deviations in parameter inference and model simulation, which could consequently bias future predictions and affect management decision correctness. However, after proper treatment of errors in sensor data, the BCMSO method could extract meaningful information from sensor data and prevent negative impacts of errors. The modeling uncertainty was also greatly reduced. In the real-world case, with 1 yr of subhourly electrical conductivity sensor data incorporated, the modeling uncertainty of nitrate concentration and management cost of controlling nitrate pollution were reduced by 70%. The BCMSO method provides a flexible framework to accommodate nonconventional observations in environmental modeling and can be easily extended to other modeling fields.

## 1. Introduction

Watershed water quality (WWQ) models, such as Soil and Water Assessment Tool (SWAT) (Neitsch et al., 2011), Hydrological Simulation Program-FORTRAN (Bicknell et al., 1997) and Annualized Agricultural Non-Point Source Pollutant model (Bingner et al., 2018), have been widely used in environmental planning and management (Fu et al., 2019; Rode et al., 2010; Wellen et al., 2015). However, WWQ modeling usually exhibits a higher uncertainty than hydrologic modeling (Han & Zheng, 2018; Rode et al., 2010; Zheng & Keller, 2007a), which may significantly bias management decisions (Zheng & Keller, 2007a, 2007b). The scarcity of water quality observations signifies the problem of equifinality (Beven & Freer, 2001) in WWQ model calibration, representing one major cause of modeling uncertainty (Fu et al., 2019; Strokal et al., 2019). Traditional water quality observations are achieved via field sampling and laboratory chemical analysis, which involve high labor, financial and time costs (Das & Jain, 2017; Zulkifli et al., 2018). The data limitation is further exacerbated by policy, culture and technical barriers to free data sharing (Li et al., 2021).

Recently, in situ water quality monitoring techniques based on various sensors have rapidly developed (Kruse, 2018; Singh et al., 2022). Sensor-based in situ monitoring avoids tedious sampling procedures and complex analytical processes and produces continuous and high-frequency observations. This approach has been increasingly applied worldwide and has produced a large amount of water quality data (Meyer et al., 2019; Park et al., 2020), providing an opportunity to adopt environmental big data to improve WWQ modeling. However, compared to traditional data, long-term sensing in complex natural environments is subject to more significant measurement errors stemming from biofouling, lack of equipment calibration, background ion interference and other factors (Mahmud et al., 2020; Makarov et al., 2021). Typical error types in sensor data include outliers, noise, constant values, missing data, bias due to incorrect calibration, sensor drift from the calibration level and other errors (Horsburgh et al., 2015; Teh et al., 2020). In addition, as most existing sensors are installed to

measure water quality parameters at fixed locations (mobile measurements involving sensors onboard unmanned boats are technically feasible but currently still rare), the commensurability error (where the model-predicted variable differs from the observed variable) (Beven, 1989, 2006; Zheng & Keller, 2007a) could be notable in addressing the average water quality within a given spatial domain. The assimilation of sensor data with significant errors in WWQ modeling has not been systematically explored, which motivated this study.

Bayesian calibration (BC) refers to a category of model calibration and uncertainty analysis methods that drive posterior parameter distributions by adopting Bayes' rule in a formal way (Vrugt et al., 2008). In BC, a likelihood function of the model parameters is established by explicitly modeling the residual error between the model outputs and corresponding observations (Evin et al., 2013; Schoups & Vrugt, 2010), and the posterior distribution is often inferred using a Monte Carlo simulation algorithm such as the Markov chain Monte Carlo (MCMC) algorithm (Gelman et al., 2004). BC has been employed in both hydrological modeling (Smith & Marshall, 2008; Vrugt et al., 2008) and WWQ modeling (Zheng & Han, 2015). Advanced BC methods have also been developed to simultaneously address model structural errors and input uncertainty and consider multiple model outputs (Guzman et al., 2015; Han & Zheng, 2016, 2018). BC is a potential solution to the assimilation of both accurate but scarce laboratory observations and high-frequency but low-quality sensor data in WWQ modeling, but the feasibility, effectiveness and benefit of using the BC technique within this special context have yet to be explored.

In this study, a novel BC method, namely, the Bayesian calibration using multisource observations (BCMSO) method, was proposed to leverage both accurate but scarce laboratory observations and high-frequency but low-quality sensor data for WWQ model calibration. The BCMSO method involves unique strategies to treat the observational errors in sensor data and establish the likelihood function based on multisource data with different error characteristics. The proposed method was demonstrated in a synthetic case considering the Newport Bay Watershed (NBW) in the U.S. and a real-world case of the Fengpu Stream Watershed (FSW) in China, and the SWAT model was employed as the WWQ model in both cases. A series of numerical experiments were designed and implemented to address the following key questions: (a) can better model calibration be achieved by considering sensor data with significant errors through the BCMSO method? (b) What is the impact of sensor data assimilation on water quality simulation, uncertainty quantification, model prediction and associated water quality management? In this study, it is demonstrated that direct assimilation of sensor data using the traditional BC technique leads to obvious deviations in parameter inference and model simulation, biases future predictions and affects management decision correctness. However, meaningful information can be extracted from sensor data via the BCMSO method, thus preventing negative impacts of errors. The findings of this study could promote the rational use of existing sensor-based water quality observations and could help improve existing water quality monitoring programs.

## 2. Bayesian Calibration Using Multisource Observations (BCMSO)

The relationship between observation  $\mathbf{Z}$ , true watershed response  $\tilde{\mathbf{Y}}$  and model output  $\mathbf{Y}$  (given the model parameters  $\theta$ ) can be expressed as:

$$\mathbf{Z} = \tilde{\mathbf{Y}} + \mathbf{e} = \mathbf{Y}(\theta) + \delta + \mathbf{e} \quad (1)$$

where  $\mathbf{e}$  denotes the observational errors of  $\mathbf{Z}$ , and  $\delta$  denotes the total modeling errors of  $\mathbf{Y}$  (i.e.,  $\tilde{\mathbf{Y}} - \mathbf{Y}$ ) resulting from model structural, parametric and input errors. In the BC framework,  $\delta + \mathbf{e}$  is usually treated as a lumped residual error term  $\boldsymbol{\varepsilon}$ , which is commonly described by an independent, heteroscedastic and Gaussian error model (Schoups & Vrugt, 2010; Vrugt et al., 2008). Moreover,  $\boldsymbol{\varepsilon}$  is assumed to exhibit a zero mean and the following heteroscedasticity:

$$\sigma_{\varepsilon,t} = a_{\varepsilon} + b_{\varepsilon} \cdot y_t \quad (2)$$

where  $\sigma_{\varepsilon,t}$  is the standard deviation of  $\varepsilon_t$  (i.e., the value of  $\boldsymbol{\varepsilon}$  at the  $t$ th time step),  $a_{\varepsilon}$  and  $b_{\varepsilon}$  are two hyperparameters, and  $y_t$  denotes the value of  $\mathbf{Y}$  at the  $t$ th time step.

If  $\boldsymbol{\varepsilon}$  is self-independent and independent of  $\mathbf{Y}$ , the likelihood function (in logarithmic form) for this error model can be expressed as (Schoups & Vrugt, 2010; Vrugt et al., 2008):

$$\log L(\theta, \boldsymbol{\varphi}_{\varepsilon}, \mathbf{Z}) = -\frac{T}{2} \log(2\pi) - \sum_{t=1}^T \log(\sigma_{\varepsilon,t}) - \frac{1}{2} \sum_{t=1}^T \frac{(z_t - y_t)^2}{\sigma_{\varepsilon,t}^2} \quad (3)$$

where  $\varphi_e$  denotes the two hyperparameters for  $\epsilon$  (i.e.,  $a_e$  and  $b_e$ ),  $z_t$  is the value of  $\mathbf{Z}$  at the  $t$ th time step, and  $T$  is the number of time steps. The posterior distribution of the parameters can be inferred based on Bayes' rule, often via MCMC sampling (Gelman et al., 2004).

Now consider two distinct sources of observations within the context of WWQ modeling, accurate laboratory analysis-based observations (denoted as  $\mathbf{Z}$ ) and less accurate in situ sensor data (denoted as  $\mathbf{U}$ ). Their relationship can be conceptualized via regression analysis as follows:

$$\mathbf{Z} = f(\mathbf{U}) + \boldsymbol{\eta} \quad (4)$$

where  $f(\cdot)$  is a regression model,  $f(\mathbf{U})$  denotes the corrected sensor data based on  $\mathbf{Z}$  (denoted as  $\mathbf{Z}'$  hereafter), and  $\boldsymbol{\eta}$  denotes the regression errors. Notably, predictor  $\mathbf{Z}'$  is uncorrelated with  $\boldsymbol{\eta}$ .

By combining Equations 1 and 4, the following can be derived:

$$\mathbf{Z}' = \tilde{\mathbf{Y}} + \mathbf{e} - \boldsymbol{\eta} \quad (5)$$

Regard  $\mathbf{e} - \boldsymbol{\eta}$  as an independent error term  $\boldsymbol{\tau}$ . In addition,  $\boldsymbol{\tau}$  is independent of  $\mathbf{Z}'$  but correlated with  $\tilde{\mathbf{Y}}$  (Appendix A). To remove the dependence of  $\boldsymbol{\tau}$  on  $\tilde{\mathbf{Y}}$ , one solution is to divide  $\boldsymbol{\tau}$  into two parts, that is,  $\boldsymbol{\tau}_1$  and  $\boldsymbol{\tau}_2$ , in which  $\boldsymbol{\tau}_1$  depends on  $\tilde{\mathbf{Y}}$  and  $\boldsymbol{\tau}_2$  is independent of  $\tilde{\mathbf{Y}}$ . In this study,  $\boldsymbol{\tau}_1$  is described by a linear model as follows:

$$\boldsymbol{\tau}_1 = c \cdot [\tilde{\mathbf{Y}} - E(\tilde{\mathbf{Y}})] \quad (6)$$

where  $c$  is a linear adjustment factor,  $E(\cdot)$  denotes the mean, and  $E(\boldsymbol{\tau}_1)$  is 0. Therefore, Equation 5 can be rewritten as:

$$\mathbf{Z}' = \tilde{\mathbf{Y}}' + \boldsymbol{\tau}_2 \quad (7)$$

where  $\tilde{\mathbf{Y}}' = \tilde{\mathbf{Y}} + \boldsymbol{\tau}_1 = (1 + c) \cdot \tilde{\mathbf{Y}} - c \cdot E(\tilde{\mathbf{Y}})$ .  $E(\tilde{\mathbf{Y}}')$  is still equal to  $E(\tilde{\mathbf{Y}})$ . By changing the value of  $c$ , the correlation between  $\tilde{\mathbf{Y}}'$  and  $\boldsymbol{\tau}_2$  can be adjusted. To ensure independence of  $\boldsymbol{\tau}_2$  from  $\tilde{\mathbf{Y}}'$ ,  $c$  should take the following value (Appendix A):

$$c = R^2 - 1 \quad (8)$$

where  $R^2$  is the coefficient of determination of the regression model.

Substituting  $\tilde{\mathbf{Y}} = \mathbf{Y} + \boldsymbol{\delta}$  into Equation 7, the resulting equation can be modified as follows:

$$\mathbf{Z}' = \mathbf{Y}' + (1 + c) \cdot \boldsymbol{\delta} + \boldsymbol{\tau}_2 \quad (9)$$

where  $\mathbf{Y}' = (1 + c) \cdot \mathbf{Y} - c \cdot E(\mathbf{Y})$ . As in traditional BC,  $(1 + c) \cdot \boldsymbol{\delta} + \boldsymbol{\tau}_2$  is treated as a lumped residual error term  $\boldsymbol{\epsilon}'$ , which is independent of  $\mathbf{Y}'$ . Moreover,  $\boldsymbol{\epsilon}'$  can be described by an independent, heteroscedastic and Gaussian error model, and two hyperparameters,  $a_{\epsilon'}$  and  $b_{\epsilon'}$ , are used to describe the heteroscedasticity.

Therefore, the log-likelihood function for the sensor data can be obtained as:

$$\log L(\boldsymbol{\theta}, \boldsymbol{\varphi}_{\epsilon'}, \mathbf{U}) = -\frac{T}{2} \log(2\pi) - \sum_{t=1}^T \log(\sigma_{\epsilon',t}) - \frac{1}{2} \sum_{t=1}^T \frac{(z'_t - y'_t)^2}{\sigma_{\epsilon',t}^2} \quad (10)$$

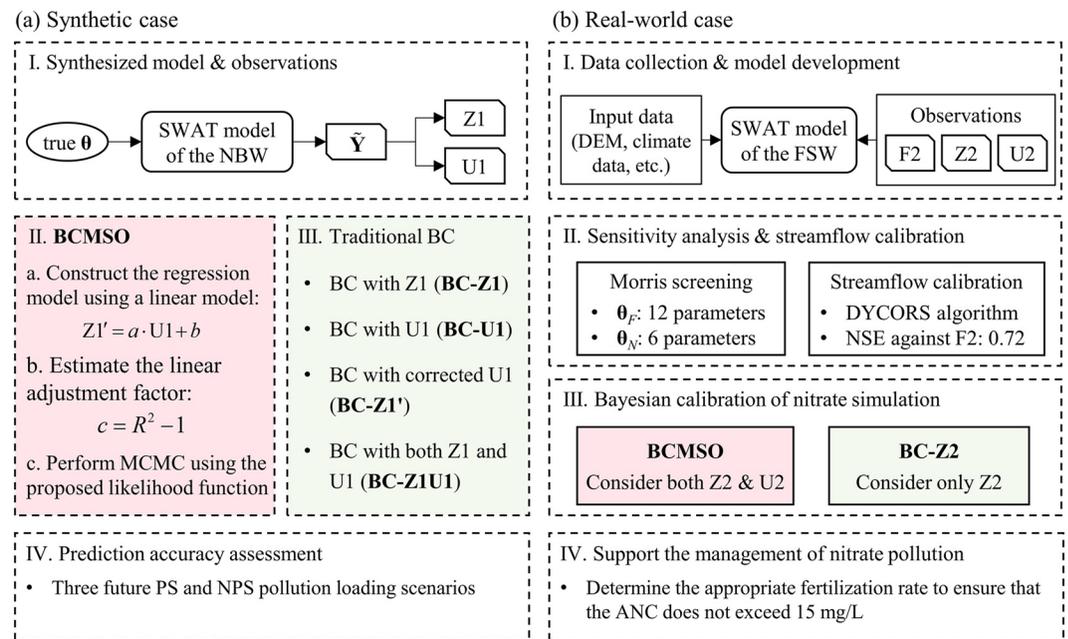
where  $\boldsymbol{\varphi}_{\epsilon'}$  denotes the two hyperparameters,  $\sigma_{\epsilon',t}$  is the standard deviation of  $\epsilon'_t$  (i.e., the value of  $\boldsymbol{\epsilon}'$  at the  $t$ th time step), and  $z'_t$  and  $y'_t$  are the values of  $\mathbf{Z}'$  and  $\mathbf{Y}'$ , respectively, at the  $t$ th time step. The MATLAB codes for Equations 3 and 10 are provided in Text S1 in the Supporting Information S1.

The likelihood function considering both laboratory observations and sensor data is simply the product of  $L(\boldsymbol{\theta}, \boldsymbol{\varphi}_{\epsilon}, \mathbf{Z})$  and  $L(\boldsymbol{\theta}, \boldsymbol{\varphi}_{\epsilon'}, \mathbf{U})$ , which is the sum of Equations 3 and 10 in logarithmic form. With this compound likelihood function, MCMC sampling can be performed to approximate the posterior parameter distribution and evaluate the simulation uncertainty.

### 3. Data and Methods

#### 3.1. Study Framework

Figure 1 shows a schematic overview of this study, which contains two nitrate modeling cases, a synthetic case based on the NBW in the U.S. and a real-world case of the FSW in China. In the synthetic case, a true SWAT



**Figure 1.** Schematic overview of the two nitrate modeling cases. (a) Synthetic case based on the Newport Bay Watershed (NBW), and (b) real-world case of the Fengpu Stream Watershed (FSW).

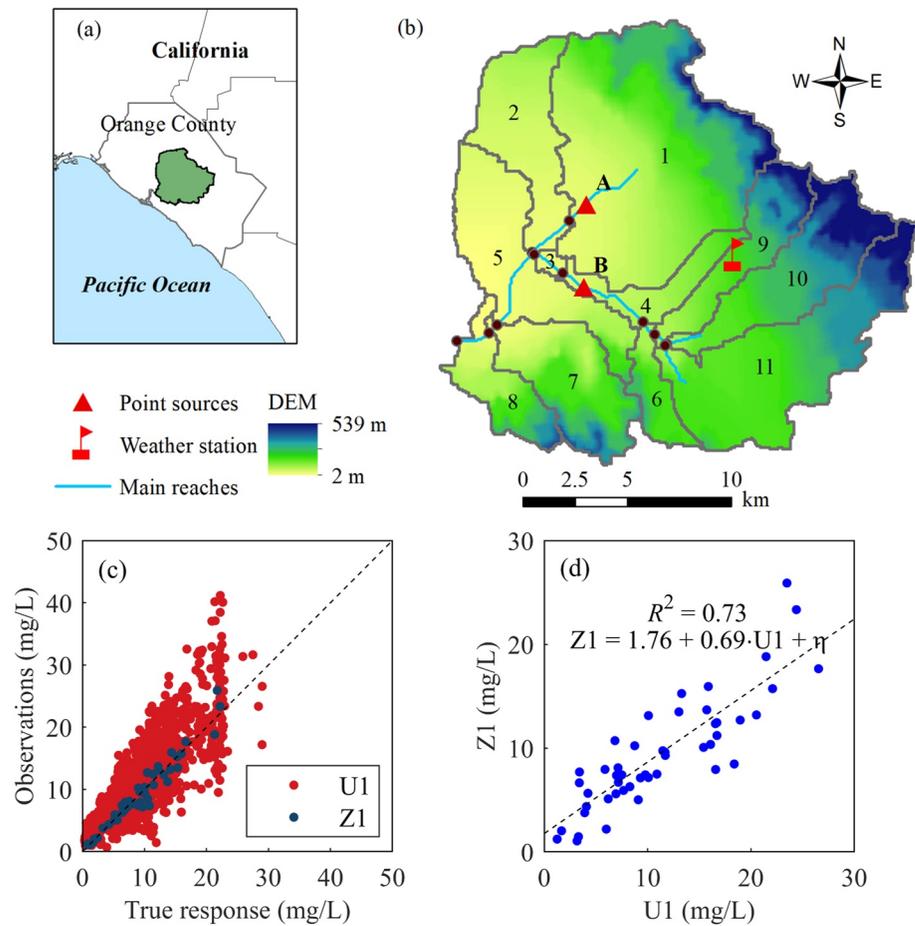
model of the NBW was synthesized, and its model outputs were regarded as the true response of the NBW. Two sets of artificial nitrate observations, Z1 and U1, were synthesized and used to represent laboratory observations and sensor data, respectively. The BCMSO method was applied to calibrate the above model (with unknown parameters). For comparison, the model was also calibrated with different observation data using the traditional BC technique (i.e., BC-Z1, BC-U1, BC-Z1', and BC-Z1U1). Notably, in BC-Z1U1 and BCMSO, if Z1 and U1 (or Z1') overlapped in time, only Z1 was used in model calibration. The calibrated models were employed to predict the nitrate concentration under three future scenarios to assess their prediction accuracy.

In the real-world case, a SWAT model of the FSW was first set up based on various input data. Multiple sets of observation data were collected, including streamflow observations (F2), laboratory nitrate observations (Z2) and sensor data of the electrical conductivity (EC) (U2). Sensitivity analysis was performed for the SWAT model to select sensitive parameters (i.e.,  $\theta_F$  and  $\theta_N$ ) in streamflow and nitrate simulation. Then, the model was calibrated against F2 to obtain one set of optimal values of  $\theta_F$ . With  $\theta_F$  fixed at the calibrated values, the nitrate response was adjusted using the BCMSO method. For comparison, the model was further calibrated using the traditional BC method only considering Z2 (i.e., BC-Z2). The calibrated models were used to support nitrate pollution management in the FSW. For more details on these two cases, please refer to Sections 3.2 and 3.3, respectively.

### 3.2. Synthetic Case

The NBW is located in southern California (Figure 2a), with an area of approximately 400 km<sup>2</sup>. The watershed is highly urbanized. It exhibits a typical Mediterranean climate, and the average annual precipitation is 330 mm. Nutrient pollution is a significant issue, where fertilization of urban lawns is the main type of nonpoint source (NPS) pollution, and effluents originating from commercial nurseries constitute the main type of point source (PS) pollution. A true SWAT model of the NBW was constructed based on real-world NBW data (Han & Zheng, 2016) (Figure 2b). In this model, the fertilization rate in urban lawns was 240 kg N/ha/a, and the average loads at two conceptualized PSs (A and B in Figure 2b) were set to 131 and 69 kg N/d, respectively. Except for eight parameters that are sensitive to nitrate modeling, the other parameters were fixed at their corresponding default values. The hypothesized true values of the eight parameters are listed in Table S1 in the Supporting Information S1.

Two sets of nitrate observations (i.e., Z1 and U1) were synthesized (please refer to Text S2 in the Supporting Information S1 for the generation method). Z1 exhibited a low frequency (monthly; 48 points) but small errors,

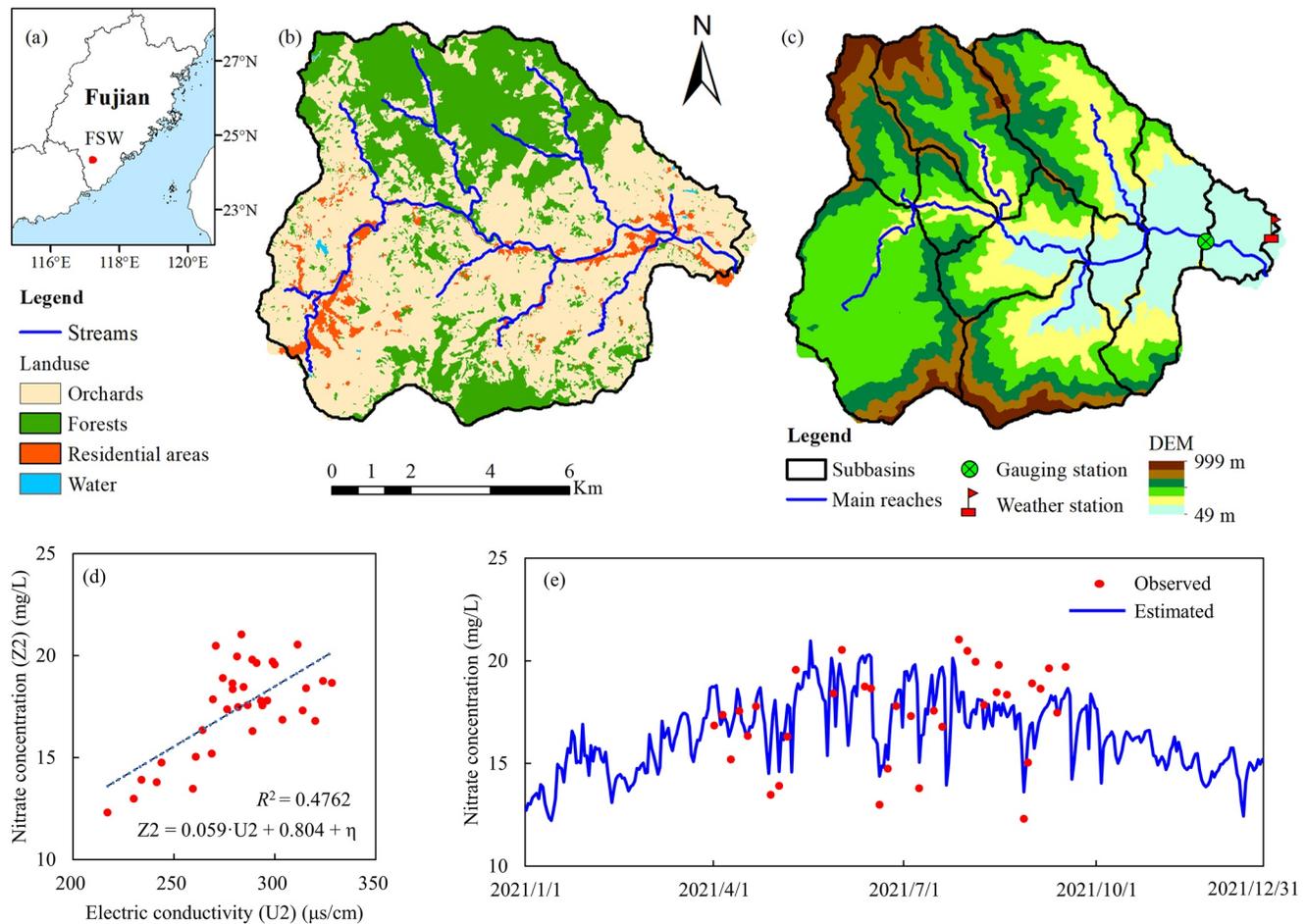


**Figure 2.** Synthetic case based on the Newport Bay Watershed (NBW). (a) Location of the NBW, (b) Soil and Water Assessment Tool model of the NBW, (c) comparisons between the artificial observations (Z1, laboratory observations; U1, sensor data) and true watershed response, and (d) linear relationship between Z1 and U1.

while U1 exhibited a high frequency (daily; 1,461 points) but large errors. Figure 2c shows a comparison of Z1 and U1 to the true watershed response. Notably, U1 not only attained a high variance but also attained a positive deviation (2 mg/L), which was used to represent sensor drift, a common type of error due to electronic drift from the instrument calibration level (Horsburgh et al., 2015). Based on 48 pairs of observations on the same day, a linear model was adopted to establish a relationship between Z1 and U1 (Figure 2d), and the  $R^2$  value was 0.73. Therefore, the linear adjustment factor  $c$  was estimated at  $-0.27$ .

As many parameters in the SWAT model are associated with spatially distributed values, the aggregate parameter method was employed to adjust these parameters, following Yang et al. (2007): varying the parameter value directly (Type I); adding a deviation to the prior parameter value and adjusting the deviation (Type II); and applying a multiplier to the prior parameter value and varying the multiplier (Type III). These parameters, deviations and multipliers were assumed to follow uniform distributions, and their ranges are provided in Table S1 in the Supporting Information S1. The Differential Evolution Adaptive Metropolis ( $DREAM_{(ZS)}$ ) algorithm (Laloy & Vrugt, 2012; Vrugt et al., 2009), a highly efficient MCMC algorithm, was employed for sampling from the posterior. To ensure the reliability of the calibration results,  $DREAM_{(ZS)}$  explored 200,000 samples in each calibration, and all the Markov chains achieved convergence according to the Gelman–Rubin diagnostic (Gelman & Rubin, 1992). The last 50% of all points of the Markov chains was used for subsequent analysis. Posterior justification of the Gaussian assumption of the residual errors is shown in Figure S1 in the Supporting Information S1.

To further compare the calibration results, three future loading scenarios (Table S2 in the Supporting Information S1) with relatively low, moderate and high loadings were proposed over the next 4 yr. The calibrated models were run to predict the nitrate concentration under the proposed future loadings, and the average nitrate



**Figure 3.** Real-world case of the Fengpu Stream Watershed (FSW). (a) Location of the FSW, (b) streams and land use, (c) Soil and Water Assessment Tool model of the FSW, (d) linear relationship between the laboratory nitrate observations ( $Z_2$ ) and electrical conductivity sensor data ( $U_2$ ), and (e) time series of the observed (i.e.,  $Z_2$ ) and estimated ( $Z_2'$ ) nitrate concentrations.

concentration (ANC) was selected to evaluate the model performance. In this synthetic case, the true ANC could be obtained by operating the SWAT model with the true parameters so that the predictions in the different calibration processes could be suitably evaluated.

### 3.3. Real-World Case

#### 3.3.1. Study Area, Model Development and Observation Data

The FSW is located in Zhangzhou city, Fujian Province, in southeastern China (Figure 3a), with an area of 86.2 km<sup>2</sup>. The FSW experiences a typical subtropical marine monsoon climate, with warm winters and humid summers. The annual rainfall reaches approximately 1,512 mm. The FSW is mainly characterized by hilly and mountainous landscapes. Most parts of the watershed exhibit a high slope and are unsuitable for cropland development. The predominant land use is orchards (Figure 3b), mainly honey pomelo orchards. The area of pomelo orchards is approximately 55.3 km<sup>2</sup>, accounting for 64.1% of the modeling area. Other land uses in the FSW include 31.5% evergreen forests, 4.3% residential areas and 0.1% water bodies (Figure 3b). Nitrate pollution is a serious water quality issue in the FSW because farmers apply large amounts of chemical fertilizers to promote honey pomelo yields. Excessive fertilization has led to considerable NPS pollution, which is exacerbated by the abundant rainfall in this area.

The SWAT model was adopted to simulate nitrate pollution in the FSW. Table 1 summarizes the data used for model development. In the SWAT model, the watershed was delineated into 10 subbasins (Figure 3c), which

**Table 1**  
Data for the Setup and Calibration of the Soil and Water Assessment Tool Model of the Fengpu Stream Watershed

Category	Data	Time range	Spatial resolution	Data sources
Basic data	Digital elevation model (DEM)	/	12.5 m × 12.5 m	Earth data ( <a href="https://www.earthdata.nasa.gov/">https://www.earthdata.nasa.gov/</a> )
	Soil data	/	1:1,000,000	Chinese Soil Database of the Institute of Soil Science, Chinese Academy of Sciences
	Land use	2019	0.6 m × 0.6 m	Obtained via satellite image interpretation
	River network	2019	0.6 m × 0.6 m	Obtained via satellite image interpretation
	Meteorological data	2015–2021	1 station	National Meteorological Science Data Center ( <a href="http://data.cma.cn/">http://data.cma.cn/</a> )
	Pomelo fertilization	2019	/	Interviews with local farmers
Observation data	Streamflow observations	2019–2021	Daily	Hydrology and Water Resources Survey Sub-center of Zhangzhou City
	Nitrate observations	2021	34 days	Analyzed via San++ analyzer, Germany
	EC observations	2021	Daily	Measured by In-Situ Aqua TROLL 600

were further divided into 101 hydrological response units. According to a survey among local farmers, the annual fertilizer application in orchards reaches approximately 1,118 kg N/ha/a, which was scheduled by date in eight operations in the SWAT model (Table S3 in the Supporting Information S1). There is no significant PS pollution in the FSW, and almost all human manure is applied in orchards as fertilizer. The simulation period ranged from 2015 to 2021, in which the first four years comprised the spin-up period (2015–2018) and the last three years comprised the calibration period (2019–2021).

Daily streamflow observations (i.e., F2) from 2019 to 2021 and 35 nitrate concentration observations (i.e., Z2) in 2021 (from April to September specifically) were collected for SWAT model calibration. No sensor data of the nitrate concentration were available, but high-frequency EC data were obtained via an in situ sensor (In-Situ Aqua TROLL 600) in 2021. The raw EC sensor data with a frequency of 20 min were averaged to a daily frequency. In the FSW, EC is highly correlated with the nitrate concentration because nitrate accounts for the majority of ions in surface streamflows, which is due to the large amount of chemical fertilizer applied in this area. Therefore, the EC data (i.e., U2) were used as alternative observations of the nitrate concentration in this study. Similar to the synthetic case, a linear regression model was employed to establish a relationship between Z2 and U2, and the  $R^2$  value was 0.48 (Figure 3d). Therefore, the linear adjustment factor  $c$  in this case was estimated at  $-0.52$ . Figure 3e shows the time series of the observed (i.e., Z2) and estimated (Z2') nitrate concentrations based on the EC data.

### 3.3.2. Sensitivity Analysis

To calibrate the SWAT model, 31 parameters affecting the streamflow and nitrate simulation performance were initially selected from hundreds of SWAT parameters (Table S4 in the Supporting Information S1). The aggregate parameter method was used to adjust the distributed parameters. All parameters were assumed to follow uniform distributions. The Morris screening method (Campolongo et al., 2007; Morris, 1991), an efficient global sensitivity analysis method, was employed to choose a subset of parameters sensitive to streamflow and nitrate concentration simulation. In Morris screening, the parameter space (mapped to the 0–1 space) is discretized into a  $p$ -level grid  $\Omega$ . Then, a number (e.g.,  $r$ ) of trajectories (of  $k + 1$  points) are constructed in  $\Omega$ , where  $k$  is the number of parameters. A trajectory is constructed by generating a random starting point and then completing the trajectory by moving one parameter at a time in a random order, and the movement of each parameter is  $\Delta$ .

In this study,  $k$  was 31,  $p$  was 10,  $\Delta$  was 5/9, and  $r$  was 200. In the literature (Sun et al., 2022),  $p$  has been reported to range from 4 to 10 in literatures. Here,  $p$  was set to 10 to better reflect the nonlinear effect of model parameters

**Table 2**  
Sensitive Parameters of the Soil and Water Assessment Tool Model of the Fengpu Stream Watershed Selected via the Morris Screening Method

Streamflow parameters $\theta_F$	Sensitivity index for streamflow (m <sup>3</sup> /s)	Calibrated values	Nitrate parameters $\theta_N$	Sensitivity index for nitrate (mg/L)
CN2	0.85	-14.96	FRT	18.48
SOL_AWC	0.51	0.83	NPERCO	15.10
CH_K2	0.50	50.00	HLIFE_NGW	7.97
ESCO	0.38	0.39	RCHRG_DP	5.57
TRNSRCH	0.35	0.14	SDNCO	4.25
CH_N2	0.31	0.06	GW_REVEP	1.63
GWQMN	0.30	3,174.65		
ALPHA_BNK	0.26	0.02		
GW_DELAY	0.20	56.75		
SHALLST	0.20	4,614.03		
CH_K1	0.17	42.17		
SURLAG	0.16	2.87		

on the simulation output. The elementary effect of the  $j$ th parameter on the output can be calculated as follows:

$$EE_j = \frac{\sum_{i=1}^T |y_i^{(i)} - y_i^{(i+1)}|/T}{\Delta} \quad (11)$$

where superscripts  $i$  and  $i + 1$  denote two consecutive points along a certain trajectory, and  $j$  denotes the number of parameters that change between these two points. The mean of the  $EE_j$  values obtained from  $r$  trajectories was adopted as the sensitivity index. Table 2 lists 12 parameters sensitive to streamflow simulation ( $\theta_F$ ), 6 parameters sensitive to nitrate simulation ( $\theta_N$ ), and their sensitivity index. Notably, fertilization adjustment factor (FRT) in  $\theta_N$  is not the original parameter in the SWAT model, but it denotes the adjustment factor of the fertilization rate. An FRT value of 0.1 indicates a 10% increase in fertilizer application. In this study, the FRT value ranged from -0.2 to 0.2. FRT reflects a certain degree of input uncertainty.

### 3.3.3. Model Calibration and Nitrate Management

A deterministic approach was applied in streamflow calibration, in which one set of optimal values of  $\theta_F$  was obtained in terms of the goodness of fit. The Nash–Sutcliffe efficiency coefficient (NSE) (Nash & Sutcliffe, 1970) was used to evaluate the streamflow simulation performance, and the DYNAMIC

Coordinate search using Response Surface models algorithm (Regis & Shoemaker, 2009, 2013), an advanced optimization algorithm, was employed to search the optimal values of  $\theta_F$  (with 2,000 model runs). After calibration, the NSE value for daily streamflow simulation was 0.72, indicating a satisfactory calibration performance (Krause et al., 2005). Table 2 lists the calibrated values of  $\theta_F$ . Figure 7a shows the observed and simulated streamflow.

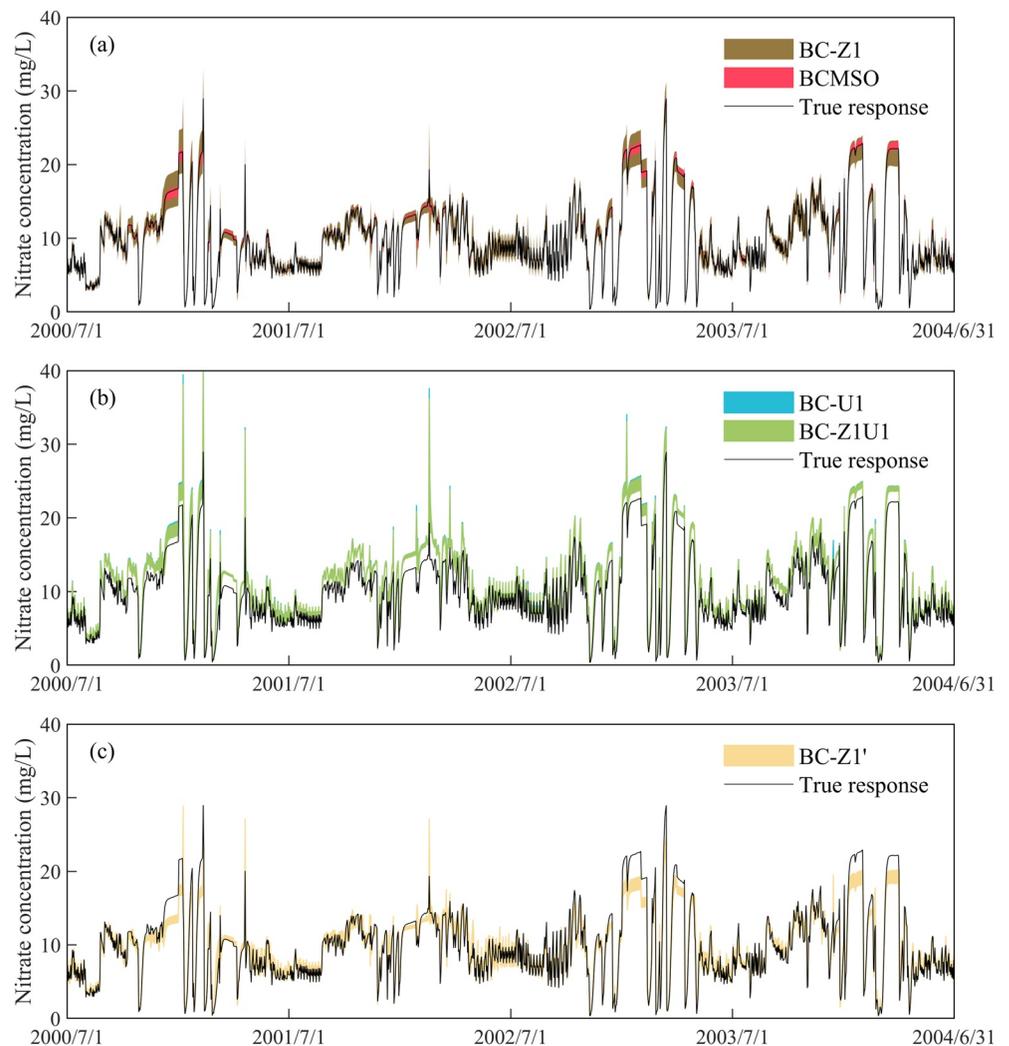
To evaluate the performance of the BCMSO method in real-world applications, the model was calibrated two times, that is, via the traditional BC technique with Z2 and the BCMSO method with both Z2 and U2. In nitrate calibration,  $\theta_F$  was fixed at the calibrated values. DREAM<sub>(ZS)</sub> was also used to sample from the posterior. A maximum of 200,000 samples were generated in each calibration, where Markov chain convergence was confirmed via the Gelman–Rubin diagnostic. The last 50% of all points of the Markov chains was used for subsequent analysis. Posterior justification of the Gaussian assumption of the residual errors is shown in Figure S2 in the Supporting Information S1.

Comparisons were performed from two perspectives: one considered the impact on model calibration, and the other considered the impact on nitrate pollution management. From the first perspective, the posterior distribution of the parameters and the uncertainty in nitrate simulation were compared. From the second perspective, we considered a specific management scenario, namely, the ANC at the watershed outlet should not exceed 15 mg/L, and this water quality improvement target should be achieved by managing the fertilization rate in honey pomelo orchards over the next three years (2022–2024). The decision-making process entailed the use of the SWAT model to predict the ANC under a series of planned fertilization rates, after which the appropriate rate could be determined via a comparison of the predictions to the improvement target (i.e., 15 mg/L).

## 4. Results

### 4.1. Synthetic Modeling Case

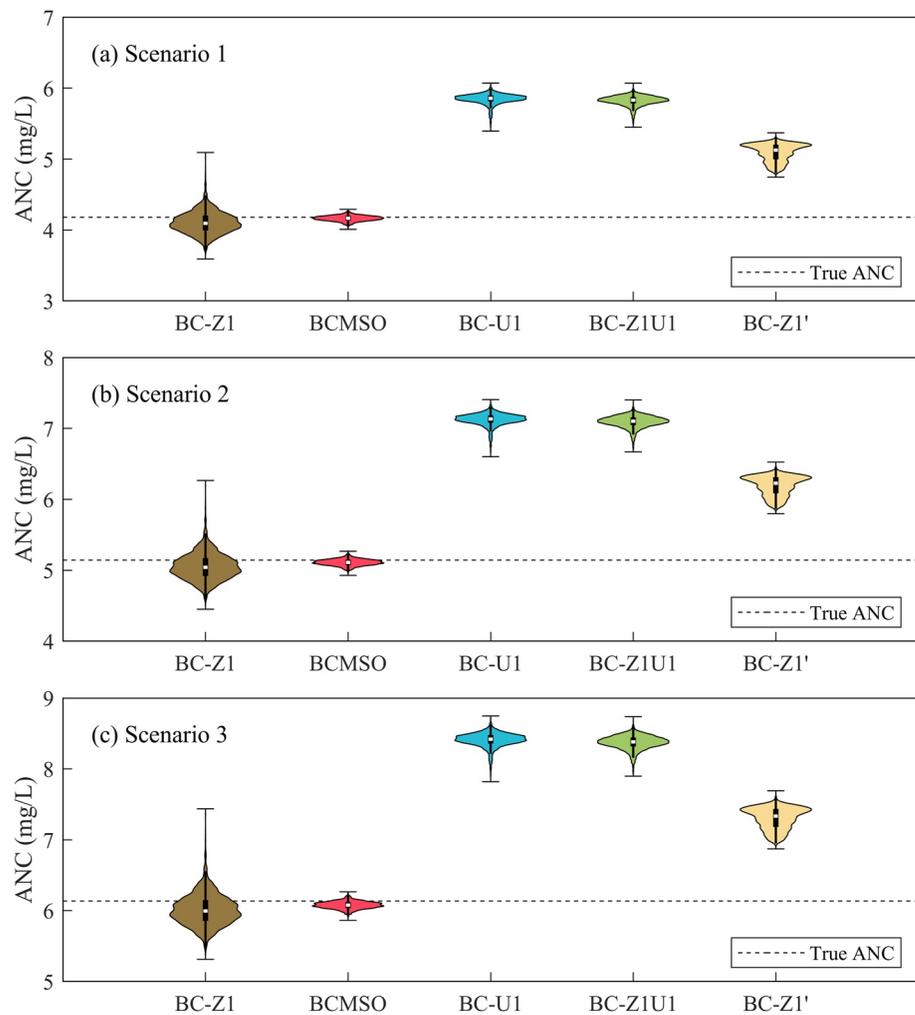
The posterior parameter distributions derived from the different calibration processes were summarized and compared (Figure S3 in the Supporting Information S1). The posterior distributions derived from BC-Z1U1 were very close to those derived from BC-U1 (Figure 4b). This occurred because in BC-Z1U1, Z1 is scarce, and the likelihood of U1 dominates the total likelihood in BC. The other distributions were quite different, which is due to the different observational data (Z1, U1 and Z1') and error models used in the different calibration processes. BC-Z1 achieved satisfactory results. The cumulative distribution function (CDF) curves generally included all



**Figure 4.** Parametric uncertainty bands derived from the five calibration processes in the synthetic case. Z1 and U1 denote the artificial laboratory observations and sensor data, respectively. Z1' denotes the corrected sensor data. BC-Z1, BC-U1, BC-Z1U1 and BC-Z1' represent four traditional Bayesian calibration (BC) processes based on different sets of observational data. BCMSO represents the new calibration process proposed in this study.

true values, although certain parameters, such as SHALLST and SDNCO, exhibited a relatively wide posterior range. When sensor data were directly used in BC (i.e., BC-U1 and BC-Z1U1), the corresponding posterior parameter distributions exhibited obvious deviations, especially ESCO, SHALLST and SDNCO. This could be attributed to the significant observational errors (especially positive deviations) in the sensor data. However, even if the sensor data were corrected (Z1'), deviations in the posterior parameter distributions would still exist. This result suggests that it may not be appropriate to treat Z1' like Z1 in the calibration process. The BCMSO method achieved an excellent performance. The CDF curves included the true values within a narrower range. This result indicates that the BCMSO method could properly and effectively extract meaningful information from the sensor data while preventing negative impacts of the associated significant observational errors. Posterior distributions of hyperparameters are shown in Figure S4 in the Supporting Information S1.

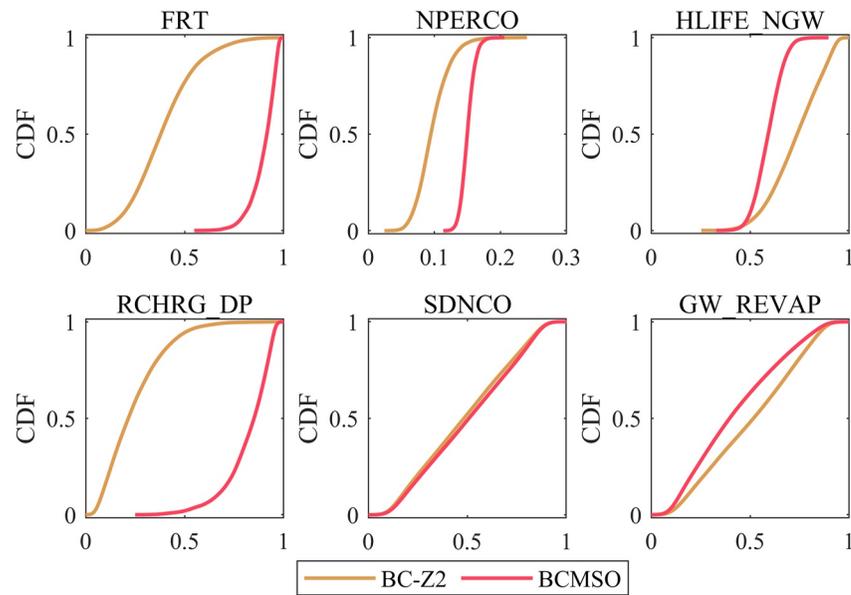
Figure 4 shows a comparison of the parametric uncertainty bands of the nitrate concentration to the true watershed response ( $\hat{Y}$ ). Here, the bands represent 95% uncertainty intervals (i.e., 2.5%–97.5%) of the model outputs corresponding to the last 50% of the samples of the Markov chains. The average band widths derived from BC-Z1, BC-U1, BC-Z1U1, BC-Z1' and BCMSO were 2.29, 0.73, 0.72, 1.25 and 0.62 mg/L, respectively. Therefore, after considering the sensor data, the parametric uncertainty was greatly reduced. In regard to the simulation accuracy, BC-Z1 successfully reproduced  $\hat{Y}$  (Figure 4a). This occurs because Z1 is accurate and unbiased. However, the



**Figure 5.** Predicted average nitrate concentration (ANC) under three future loading scenarios in the synthetic case. Z1 and U1 denote the artificial laboratory observations and sensor data, respectively. Z1' denotes the corrected sensor data.

bands derived from BC-U1 and BC-Z1U1 were obviously higher than  $\tilde{Y}$  (Figure 4b). This could be attributed to the positive deviation in the sensor data, indicating that the direct assimilation of sensor data in BC may cause biased simulation results. The band derived from BC-Z1' significantly deviated from  $\tilde{Y}$  (Figure 4c). This occurs because although Z1' is more accurate in the average sense, there exists a trend deviation between Z1' and  $\tilde{Y}$  in the time series, namely, Z1' is generally lower than  $\tilde{Y}$  at high values and higher than  $\tilde{Y}$  at low values. This trend deviation indicates that the observational errors of Z1' (i.e.,  $Z1' - \tilde{Y}$ ) are correlated to  $\tilde{Y}$  (and  $Y$ ). Therefore, the direct assimilation of Z1' could lead to a trend deviation in the simulation results in the time series. The BCMSO method performed the best, with a narrower band at approximately  $\tilde{Y}$  (Figure 4a). This result indicated that the method proposed in this study could properly and effectively consider sensor data in the BC process.

One important task of WWQ modeling involves future water quality prediction to support the decision-making process. Three future loading scenarios (Table S2 in the Supporting Information S1) were designed to evaluate the predictive ability of the calibrated models. Figure 5 shows the ANC distributions predicted with the calibrated models under the three scenarios. BC-Z1 and BCMSO performed well under all three scenarios, but BC-U1, BC-Z1U1 and BC-Z1' yielded overestimated ANC values. The overestimation of BC-U1 and BC-Z1U1 was attributable to the positive deviation in U1, and the inaccuracy of BC-Z1' occurred because the correlation between the observational errors of Z1' and  $\tilde{Y}$  is ignored in this process. The BCMSO method was hardly affected because it could greatly reduce the positive deviation in U1 through linear regression and eliminate the



**Figure 6.** Posterior parameter distributions derived from the two calibration processes in the real-world case. All parameters were mapped to the 0–1 domain via linear min-max normalization. Z2 denotes the laboratory observations.

correlation between the observational errors of  $Z1'$  and  $\tilde{Y}$  via a linear adjustment factor ( $c$  in Equation 8). Moreover, the prediction uncertainty of the BCMSO method was much lower than that of the BC-Z1 method. This result demonstrates that the consideration of sensor data could significantly reduce the simulation uncertainty of WWQ models through proper assimilation methods (such as the BCMSO method).

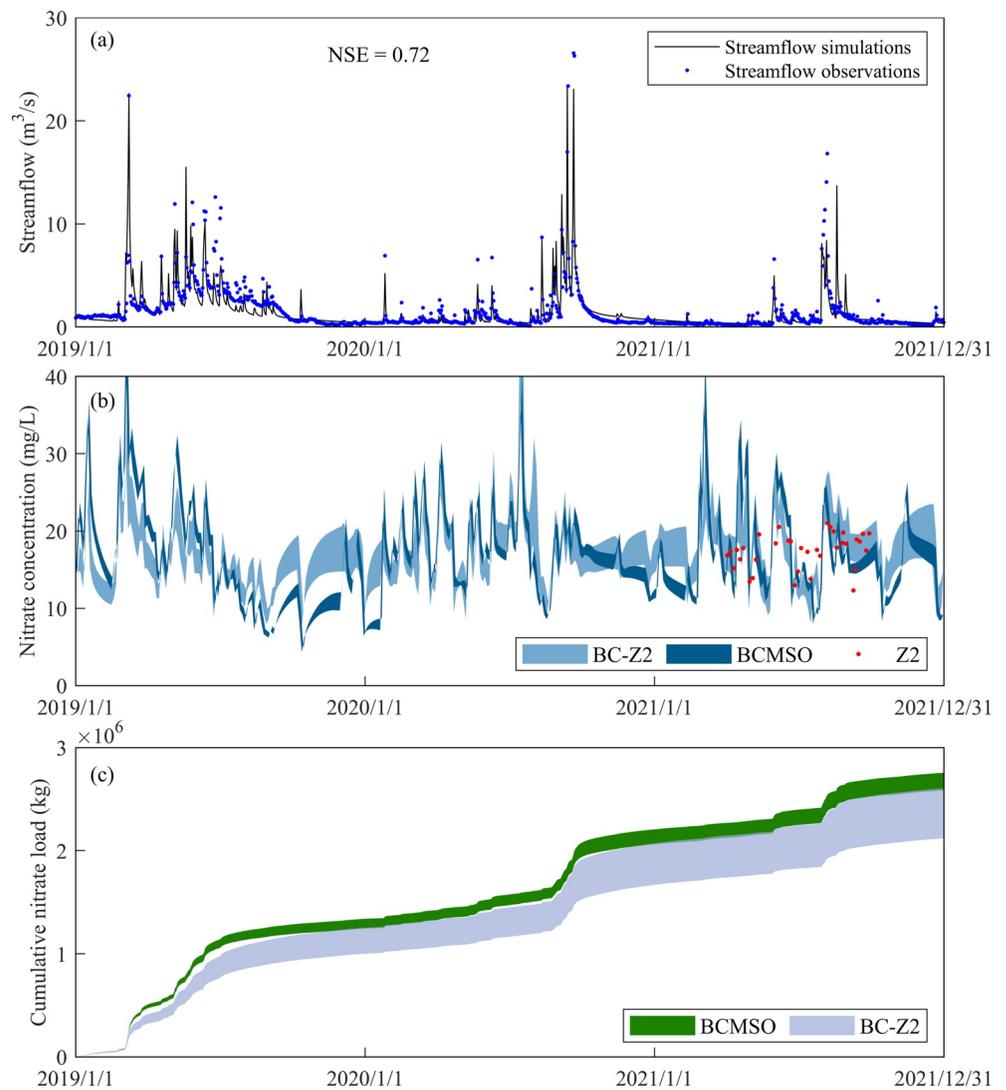
## 4.2. Real-Word Application

### 4.2.1. Impact on Model Calibration

It has been demonstrated thus far that the BCMSO method can reduce the simulation uncertainty and is less affected by the errors occurring in sensor data. Because WWQ modeling is susceptible to input and structural errors (Han & Zheng, 2016; Wellen et al., 2015), we introduced parameter FRT (Section 3.3.2) to address the uncertainty in fertilizer application in honey pomelo orchards, expecting calibration results less affected by the input uncertainty. Figure 6 shows the posterior parameter distributions obtained in the two calibration processes. The first four parameters were identified well, while SDNCO and GW\_REVAP, with the lowest sensitivity, could hardly be identified. The BCMSO method obtained narrower posterior distributions, especially for the more sensitive parameters. Regarding FRT, the posterior distributions also indicated that the fertilization rates were  $1,066 \pm 159$  kg N/ha/a and  $1,307 \pm 59$  kg N/ha/a (the margin of error was twice the standard deviation) inferred via BC-Z2 and BCMSO, respectively. In contrast, the fertilization rate inferred via BCMSO was higher, and the uncertainty was approximately 1/3 of that inferred via BC-Z2. Posterior distributions of hyperparameters are shown in Figure S5 in the Supporting Information S1.

Figure 7 shows a comparison of the obtained simulation results between the two calibration processes. Figure 7a shows the time series of the calibrated streamflow, which indicates that the model could successfully reproduce the streamflow in the FSW. Figure 7b shows parametric uncertainty bands of the nitrate concentration. Although the model inevitably exhibited significant input and structural errors, the two uncertainty bands basically matched those of the observations (Z2) and reproduced the temporal variation in the nitrate concentration. The ANCs derived from BC-Z2 and BCMSO were very close, at 16.97 and 16.82 mg/L, respectively. However, the temporal variation in the nitrate simulation derived from BCMSO was more significant (with higher peaks and lower bases). The average band widths determined by BC-Z2 and BCMSO were 4.75 and 1.61 mg/L, respectively. Therefore, the BCMSO method reduced the modeling uncertainty by approximately 70% by considering EC sensor data. This result underscores the constraining role of sensor data in Bayesian inference.

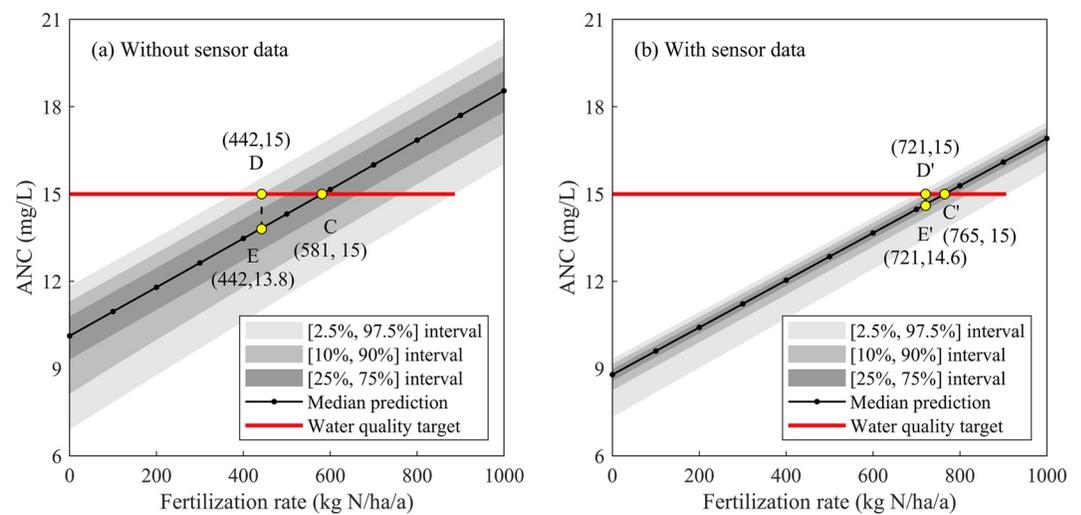
Figure 7c further shows a comparison of the cumulative nitrate loads derived from the two calibration processes. Although the ANCs derived from the BC and BCMSO methods were close, the cumulative loads differed.



**Figure 7.** Simulation results of the (a) streamflow, (b) nitrate concentration and (c) cumulative nitrate load in the real-world case with (Bayesian calibration using multisource observations (BCMSO)) and without assimilating sensor data (BC-Z2). The bands indicate 95% uncertainty intervals (i.e., 2.5%–97.5%). Z2 denotes the laboratory observations.

This result indicates that the difference in nitrate concentration in the time series affected the estimation of the total load. The cumulative nitrate load in the FSW over the 3-yr period derived via BC-Z2 was approximately  $2.34 \times 10^6$  kg. The estimate obtained with the BCMSO method was higher, at approximately  $2.71 \times 10^6$  kg, and the difference corresponded to an 16% difference in the total load. Furthermore, the cumulative uncertainty in the nitrate load over the 3-yr period derived by BC-Z2 was  $0.49 \times 10^6$  kg, accounting for 21% of the total load. However, the cumulative uncertainty in the nitrate load derived via the BCMSO method reached only  $0.17 \times 10^6$  kg, accounting for 6% of the total load.

The above comparisons (Figures 6 and 7) indicate that the BCMSO method could significantly reduce the uncertainty in both parameter inference and stochastic simulation. However, the difference in the results also raises the following question: which calibration result is more reliable? In this real-world case, the true values of the model parameters and true watershed responses are unknown, so we cannot directly determine which result is more accurate. Nevertheless, it is reasonable to believe that the results obtained with the BCMSO method could be more reliable because it introduced more observational information into the Bayesian inference process in an appropriate way. However, this reliability is only relative to the BC method with laboratory observations. Even if the sensor data were correctly included, the additional observational information would not effectively



**Figure 8.** Average nitrate concentration (ANC) under the different fertilization rates predicted with the calibrated models in the real-world case. Subplots (a) and (b) show the results without sensor data (BC-Z2) and with sensor data (Bayesian calibration using multisource observations (BCMSO)), respectively. The points along the median prediction lines correspond to the 11 fertilization rates used for model prediction.

compensate for the deviation due to input and structural errors. When significant input and structural uncertainties occur, it is recommended to directly address these uncertainties in BC (Ajami et al., 2007; Han & Zheng, 2018; Thyer et al., 2009; Xu et al., 2017; Yen et al., 2014), such as the introduction of FRT in this case.

#### 4.2.2. Impact on the Management of Nitrate Pollution

In water quality management practices (such as developing total maximum daily load standards), a high modeling uncertainty could lead to a large margin of safety (MOS), which could dramatically increase the management cost (Nunoo et al., 2020). Within this context, the calibrated models were used to support nitrate pollution management in the FSW. In this case, 11 future fertilization rates were considered, ranging from 0 to 1,000 kg N/ha/a, at intervals of 100 kg N/ha/a. Figures 8a and 8b show the predicted ANC values under the different fertilization rates derived from BC-Z2 and BCMSO, respectively. The uncertainty intervals were based on the parametric uncertainty. In general, there existed a strong linear relationship between the fertilization rate and ANC. The predictive uncertainty in the ANC derived from the BCMSO method was much lower.

In the management of nitrate pollution, if uncertainty were ignored, it would be reasonable to consider median predictions for decision-making purposes. Thus, to reduce the ANC to 15 mg/L, the fertilization rate should be reduced to 581 and 765 kg N/ha/a (interpolated values) according to BC-Z2 and BCMSO, respectively (i.e., C and C', respectively, in Figure 8). However, if uncertainty were considered and the probability of the ANC exceeding the target should be less than 10%, the fertilization rate should reach 442 and 721 kg N/ha/a according to BC-Z2 and BCMSO, respectively (i.e., D and D', respectively, in Figure 8). The median ANC predictions under these two rates were 13.8 and 14.6 mg/L, respectively (i.e., E and E', respectively, in Figure 8). Here, distances DE and D'E' (i.e., 1.2 and 0.4 mg/L, respectively) reflect the MOS accounting for the modeling uncertainty, and distances CD and C'D' (i.e., 139 and 44 kg N/ha/a, respectively) represent the MOS-associated cost. Obviously, the MOS and corresponding cost derived from BC-Z2 were much greater than those derived from BCMSO.

In the FSW, 1 kg of N fertilizer could yield approximately 42 kg of honey pomelo. Therefore, under BC-Z2, the cost of the MOS (i.e., the reduction in the fertilization rate) also indicated a reduction in honey pomelo production by approximately 32,284 tons in the FSW. According to a survey among local farmers, the prices of honey pomelo and fertilizer were approximately 4 Chinese yuan (CNY)/kg and 30 CNY/kg N, respectively, in 2020. Therefore, the economic cost of the MOS under BC-Z2 reached approximately 106.1 million CNY, accounting for 12.4% of the total economic income of honey pomelo cultivation in the FSW (i.e., 853.2 million CNY). In the BCMSO method, the reduction in honey pomelo production due to the MOS was only 10,219 tons, and the economic cost was only 33.6 million CNY, approximately 4% of the total economic income. Text S3 in the Supporting Information S1 provides details of the above estimates. This result underscores that sensor data introduction into the BC process could ensure a more accurate decision-making process and could greatly reduce the MOS-associated economic cost.

## 5. Discussion

The BCMSO method proposed in this study provides a very flexible framework and can be easily extended to other modeling fields (such as hydrology) involving multiple sets of observation data. In this study, one set of sensor data was considered in both the synthetic and real-world cases. The BCMSO method can simply be adopted to consider multiple sets of sensor data by constructing a likelihood function for each dataset and considering them in the total likelihood function. Multiple sets of sensor data are likely to occur because there may be multiple sensors of different individuals or organizations for water quality monitoring since sensors are relatively simple and low-cost instruments. Another advantage of the BCMSO method is that the sensor data used for calibration can comprise indirect observation data of the target water quality parameter, as long as the indirect observation data can capture the target parameter through a specific relationship, even if the estimations may suffer large errors. For example, EC data were used to calibrate the nitrate concentration in the real-world case in this study. This advantage potentially enables the application of more sensor data to assist in WWQ model calibration, which makes the BCMSO method very attractive against the background of environmental big data.

The BCMSO method also suffers limitations that must be mentioned. One limitation is the assumption of Gaussian residual errors ( $\epsilon$  and  $\epsilon'$ ). While independent Gaussian errors represent the most popular assumption, autocorrelated non-Gaussian errors may be relevant in certain cases (Ammann et al., 2019; Samadi et al., 2017; Schoups & Vrugt, 2010). Error models with more complex distributions or structures, such as the skew exponential power distribution (Schoups & Vrugt, 2010) and autoregressive models (Evin et al., 2013; Schoups & Vrugt, 2010), could also be considered to address these characteristics of residual errors, although they could lead to more complex likelihood functions. Notably, in the BC process, one should gradually increase the complexity of the error model to avoid introducing too much uncertainty. In addition, attention should be given to the reliability of the regression model of the considered laboratory observations and sensor data. For example, is the linear model established based on dozens of observations in this study adequate for all sensor data? Regarding this problem, we suggest using additional observations to evaluate the regression model reliability. If possible, a reasonable sampling plan should be designed, and representative observation data should be collected for regression model establishment.

## 6. Conclusions

In this study, a formal BC method was proposed for WWQ models that can simultaneously consider accurate but scarce laboratory observations and inaccurate but abundant sensor data. In the proposed method, namely, the BCMSO method, sensor data (i.e.,  $\mathbf{Z}'$ ) are corrected based on coupled laboratory observations through a regression model, and a linear adjustment factor (i.e.,  $c$ ) is employed to eliminate the correlation between the observational errors of  $\mathbf{Z}'$  and the model output. Based on the independent, heteroscedastic and Gaussian error model, likelihood functions for both the laboratory observations and sensor data were derived. Two modeling cases were designed to demonstrate the BCMSO method, namely, a synthetic case based on the NBW and a real-world case based on the FSW. In both cases, the SWAT model was used as the WWQ model to simulate nitrate pollution in the watershed. The major study findings are as follows:

1. Impact on parameter inference and model prediction. Due to the significant observational errors in the sensor data, direct sensor data application in WWQ model calibration could generate obvious parameter inference and model simulation deviations. These deviations could inevitably bias future predictions, thus affecting the correctness of model-based management decisions. However, after correcting the observational errors in the sensor data and eliminating their dependence on the model output, the BCMSO method could effectively extract meaningful information from the sensor data and prevent negative impacts of the inherent errors.
2. Effect on uncertainty reduction. The modeling uncertainty could be greatly reduced by considering sensor data in model calibration. In the FSW case, the modeling uncertainty was reduced by 70%. This could be particularly important for WWQ models because a higher uncertainty suggests a larger MOS, which could increase management costs. In the management case involving nitrate pollution control in the FSW, the economic cost of the MOS without considering sensor data reached 106.1 million CNY. After considering sensor data, the MOS-associated cost was only 33.6 million CNY, a reduction of approximately 70%.
3. Transferability of the method. The BCMSO method provides a very flexible framework allowing multiple sets of sensor data, even indirect observations of target water quality parameters (e.g., EC), to be used in model

calibration. This advantage makes the BCMSO method highly attractive against the background of environmental big data. In addition, the regression model and statistical assumptions of the errors ( $\epsilon$  and  $\epsilon'$ ) must be carefully examined in BCMSO application to ensure the validity of the results.

### Appendix A: Calculation of the Linear Adjustment Factor $c$

Let  $\tau$  denote  $\mathbf{e} - \eta$ . Because  $\mathbf{Z}'$  is independent of  $\mathbf{e}$  (unrelated observation errors) and  $\eta$  (regression errors), it is also independent of their difference  $\mathbf{e} - \eta$  (i.e.,  $\tau$ ). According to Equation 5,  $\tau = \mathbf{Z}' - \tilde{\mathbf{Y}}$ . Because  $\mathbf{e}$  is independent of  $\mathbf{Z}'$  and  $\tilde{\mathbf{Y}}$ , it is also independent of their difference  $\mathbf{Z}' - \tilde{\mathbf{Y}}$  (i.e.,  $\tau$ ). Therefore,  $\mathbf{Z}'$ ,  $\mathbf{e}$  and  $\tau$  are independent of each other. According to Equation 5,  $\tilde{\mathbf{Y}} = \mathbf{Z}' - \tau$ , so the covariance between  $\tau$  and  $\tilde{\mathbf{Y}}$  can be calculated as follows:

$$\text{cov}(\tau, \tilde{\mathbf{Y}}) = \text{cov}(\tau, \mathbf{Z}' - \tau) = \text{cov}(\tau, \mathbf{Z}') - D(\tau) \quad (\text{A1})$$

where  $D(\cdot)$  denotes the variance. Because  $\tau$  is independent of  $\mathbf{Z}'$ ,  $\text{cov}(\tau, \mathbf{Z}')$  is 0.  $\text{cov}(\tau, \tilde{\mathbf{Y}})$  equals  $-D(\tau)$ . Therefore,  $\tau$  is correlated with  $\tilde{\mathbf{Y}}$ .

According to Equation 7,  $\tau_2 = \mathbf{Z}' - \tilde{\mathbf{Y}}'$ . The covariance between  $\tilde{\mathbf{Y}}'$  and  $\tau_2$  can be calculated as:

$$\text{cov}(\tilde{\mathbf{Y}}', \tau_2) = \text{cov}(\tilde{\mathbf{Y}}', \mathbf{Z}' - \tilde{\mathbf{Y}}') = \text{cov}(\tilde{\mathbf{Y}}', \mathbf{Z}') - D(\tilde{\mathbf{Y}}') \quad (\text{A2})$$

If  $\tau_2$  is independent of  $\tilde{\mathbf{Y}}'$ , the above covariance is 0, which suggests  $\text{cov}(\tilde{\mathbf{Y}}', \mathbf{Z}') = D(\tilde{\mathbf{Y}}')$ . Since  $\tilde{\mathbf{Y}}' = (1 + c) \cdot \tilde{\mathbf{Y}} - c \cdot E(\tilde{\mathbf{Y}})$ , the following can be obtained:

$$\begin{aligned} \text{cov}(\tilde{\mathbf{Y}}', \mathbf{Z}') &= (1 + c)\text{cov}(\tilde{\mathbf{Y}}, \mathbf{Z}') \\ &= (1 + c)\text{cov}(\mathbf{Z}' - \tau, \mathbf{Z}') \\ &= (1 + c)D(\mathbf{Z}') \end{aligned} \quad (\text{A3})$$

and

$$D(\tilde{\mathbf{Y}}') = (1 + c)^2 D(\tilde{\mathbf{Y}}) \quad (\text{A4})$$

Combining Equations A3 and A4, we can obtain:

$$c = \frac{D(\mathbf{Z}')}{D(\tilde{\mathbf{Y}})} - 1 \quad (\text{A5})$$

Because  $\tilde{\mathbf{Y}}$  is unknown,  $c$  can be approximated as follows:

$$c \approx \frac{D(\mathbf{Z}'_s)}{D(\mathbf{Y}_s)} - 1 \approx \frac{D(\mathbf{Z}'_s)}{D(\mathbf{Z}_s)} - 1 \quad (\text{A6})$$

where subscript  $s$  denotes the subset of samples used to generate the regression model. Here,  $D(\mathbf{Y}_s)$  can be approximated by  $D(\mathbf{Z}_s)$  because  $\mathbf{Z}_s$  is considered to be relatively accurate.

When the regression model is unbiased, that is,  $E(\eta) = 0$ , we can obtain the following:

$$\frac{D(\mathbf{Z}'_s)}{D(\mathbf{Z}_s)} = \frac{D(\mathbf{Z}_s) - D(\eta)}{D(\mathbf{Z}_s)} = 1 - \frac{\sum_s (z_s - z'_s)^2}{\sum_s (z_s - \bar{z})^2} = R^2 \quad (\text{A7})$$

Therefore:

$$c \approx R^2 - 1 \quad (\text{A8})$$

### Data Availability Statement

For the Newport Bay Watershed (the synthetic case), the Digital elevation model (DEM), land use and soil data were obtained from U.S. Geological Survey (<https://www.usgs.gov/>), and the meteorological data (1997–2004) were obtained from U.S. National Climatic Data Center (<https://coast.noaa.gov/digitalcoast/contributing-part>).

ners/noaa-national-climatic-data-center.html). For the Fengpu Stream Watershed (the real-world case), the meteorological data, soil data and DEM were obtained from National Meteorological Science Data Center of China (<http://data.cma.cn/>), Institute of Soil Science (Chinese Academy of Sciences) (<http://english.issas.cas.cn/>) and EARTH DATA (<https://www.earthdata.nasa.gov/>), respectively.

#### Acknowledgments

This study is supported by the National Natural Science Foundation of China (NSFC) (No. 51961125203) and Shenzhen Science and Technology Innovation Commission (No. KCXFZ202002011006491).

#### References

- Ajami, N. K., Duan, Q., & Sorooshian, S. (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research*, *43*(1), W01403. <https://doi.org/10.1029/2005wr004745>
- Ammann, L., Fenicia, F., & Reichert, P. (2019). A likelihood framework for deterministic hydrological models and the importance of non-stationary autocorrelation. *Hydrology and Earth System Sciences*, *23*(4), 2147–2172. <https://doi.org/10.5194/hess-23-2147-2019>
- Beven, K. (1989). Changing ideas in hydrology — The case of physically-based models. *Journal of Hydrology*, *105*(1–2), 157–172. [https://doi.org/10.1016/0022-1694\(89\)90101-7](https://doi.org/10.1016/0022-1694(89)90101-7)
- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1–2), 18–36. <https://doi.org/10.1016/j.jhydrol.2005.07.007>
- Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, *249*(1–4), 11–29. [https://doi.org/10.1016/s0022-1694\(01\)00421-8](https://doi.org/10.1016/s0022-1694(01)00421-8)
- Bicknell, B. R., Imhoff, J. C., Kittle, J. J. L., Donigan, J. A. S., & Johanson, R. C. (1997). *Hydrological simulation program – FORTRAN, user's manual for version 11*. U.S. Environmental Protection Agency.
- Bingner, R. L., Theurer, F. D., & Yuan, Y. (2018). *AnnAGNPS technical processes: Version 5.5*. USDA-ARS National Sedimentation Laboratory.
- Campolongo, F., Carlini, J. B., Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, *22*(10), 1509–1518. <https://doi.org/10.1016/j.envsoft.2006.10.004>
- Das, B., & Jain, P. C. (2017). Real-time water quality monitoring system using Internet of Things. In *2017 International conference on computer, communications and electronics (comptelix)* (pp. 78–82).
- Evin, G., Kavetski, D., Thyer, M., & Kuczera, G. (2013). Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, *49*(7), 4518–4524. <https://doi.org/10.1002/wrcr.20284>
- Fu, B., Merritt, W. S., Croke, B. F. W., Weber, T. R., & Jakeman, A. J. (2019). A review of catchment-scale water quality and erosion models and a synthesis of future prospects. *Environmental Modelling & Software*, *114*, 75–97. <https://doi.org/10.1016/j.envsoft.2018.12.008>
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis*. Chapman & Hall/CRC.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. <https://doi.org/10.1214/ss/1177011136>
- Guzman, J. A., Shirmohammadi, A., Sadeghi, A. M., Wang, X., Chu, M. L., Jha, M. K., et al. (2015). Uncertainty considerations in calibration and validation of hydrologic and water quality models. *Transactions of the ASABE*, *58*(6), 1745–1762.
- Han, F., & Zheng, Y. (2016). Multiple-response Bayesian calibration of watershed water quality models with significant input and model structure errors. *Advances in Water Resources*, *88*, 109–123. <https://doi.org/10.1016/j.advwatres.2015.12.007>
- Han, F., & Zheng, Y. (2018). Joint analysis of input and parametric uncertainties in watershed water quality modeling: A formal Bayesian approach. *Advances in Water Resources*, *116*, 77–94. <https://doi.org/10.1016/j.advwatres.2018.04.006>
- Horsburgh, J. S., Reeder, S. L., Jones, A. S., & Meline, J. (2015). Open source software for visualization and quality control of continuous hydrologic and water quality sensor data. *Environmental Modelling & Software*, *70*, 32–44. <https://doi.org/10.1016/j.envsoft.2015.04.002>
- Krause, P., Boyle, D. P., & Båse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, *5*, 89–97. <https://doi.org/10.5194/adgeo-5-89-2005>
- Kruse, P. (2018). Review on water quality sensors. *Journal of Physics D: Applied Physics*, *51*(20), 203002. <https://doi.org/10.1088/1361-6463/aabb93>
- Laloy, E., & Vuigt, J. A. (2012). High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing. *Water Resources Research*, *48*(1), W01526. <https://doi.org/10.1029/2011wr010608>
- Li, X., Cheng, G., Wang, L., Wang, J., Ran, Y., Che, T., et al. (2021). Boosting geoscience data sharing in China. *Nature Geoscience*, *14*(8), 541–542. <https://doi.org/10.1038/s41561-021-00808-y>
- Mahmud, M. A. P., Ejeian, F., Azadi, S., Myers, M., Pejčić, B., Abbassi, R., et al. (2020). Recent progress in sensing nitrate, nitrite, phosphate, and ammonium in aquatic environment. *Chemosphere*, *259*, 127492. <https://doi.org/10.1016/j.chemosphere.2020.127492>
- Makarov, M., Aslamov, I., & Gnatovsky, R. (2021). Environmental monitoring of the littoral zone of Lake Baikal using a network of automatic hydro-meteorological stations: Development and trial run. *Sensors (Basel)*, *21*(22), 7659. <https://doi.org/10.3390/s21227659>
- Meyer, A. M., Klein, C., Funfrocken, E., Kautenburger, R., & Beck, H. P. (2019). Real-time monitoring of water quality to identify pollution pathways in small and middle scale rivers. *Science of the Total Environment*, *651*(Pt 2), 2323–2333. <https://doi.org/10.1016/j.scitotenv.2018.10.069>
- Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, *33*(2), 161–174. <https://doi.org/10.1080/00401706.1991.10484804>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Neitsch, S. L., Arnold, J. G., Kiniry, J. R., & Williams, J. R. (2011). *Soil and water assessment tool theoretical documentation version 2009*. Texas Water Resources Institute.
- Nunoo, R., Anderson, P., Kumar, S., & Zhu, J.-J. (2020). Margin of safety in TMDLs: Natural language processing-aided review of the state of practice. *Journal of Hydrologic Engineering*, *25*(4), 04020002. [https://doi.org/10.1061/\(asce\)he.1943-5584.0001889](https://doi.org/10.1061/(asce)he.1943-5584.0001889)
- Park, J., Kim, K. T., & Lee, W. H. (2020). Recent advances in information and communications technology (ICT) and sensor technology for monitoring water quality. *Water*, *12*(2), 510. <https://doi.org/10.3390/w12020510>
- Regis, R. G., & Shoemaker, C. A. (2009). Parallel stochastic global optimization using radial basis functions. *INFORMS Journal on Computing*, *21*(3), 411–426. <https://doi.org/10.1287/ijoc.1090.0325>
- Regis, R. G., & Shoemaker, C. A. (2013). Combining radial basis function surrogates and dynamic coordinate search in high-dimensional expensive black-box optimization. *Engineering Optimization*, *45*(5), 529–555. <https://doi.org/10.1080/0305215x.2012.687731>
- Rode, M., Arhonditsis, G., Balin, D., Kebede, T., Krysanova, V., van Griensven, A., & van der Zee, S. E. A. T. M. (2010). New challenges in integrated water quality modelling. *Hydrological Processes*, *24*(24), 3447–3461. <https://doi.org/10.1002/hyp.7766>

- Samadi, S., Tufford, D. L., & Carbone, G. J. (2017). Estimating hydrologic model uncertainty in the presence of complex residual error structures. *Stochastic Environmental Research and Risk Assessment*, 32(5), 1259–1281. <https://doi.org/10.1007/s00477-017-1489-6>
- Schoups, G., & Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46(10), W10531. <https://doi.org/10.1029/2009wr008933>
- Singh, S., Anil, A. G., Kumar, V., Kapoor, D., Subramanian, S., Singh, J., & Ramamurthy, P. C. (2022). Nitrates in the environment: A critical review of their distribution, sensing techniques, ecological effects and remediation. *Chemosphere*, 287(Pt 1), 131996. <https://doi.org/10.1016/j.chemosphere.2021.131996>
- Smith, T. J., & Marshall, L. A. (2008). Bayesian methods in hydrologic modeling: A study of recent advancements in Markov chain Monte Carlo techniques. *Water Resources Research*, 44(12), W00B05. <https://doi.org/10.1029/2007wr006705>
- Strokal, M., Spanier, J. E., Kroeze, C., Koelmans, A. A., Flörke, M., Franssen, W., et al. (2019). Global multi-pollutant modelling of water quality: Scientific challenges and future directions. *Current Opinion in Environmental Sustainability*, 36, 116–125. <https://doi.org/10.1016/j.cosust.2018.11.004>
- Sun, X., Croke, B., Jakeman, A., & Roberts, S. (2022). Benchmarking active subspace methods of global sensitivity analysis against variance-based Sobol' and Morris methods with established test functions. *Environmental Modelling & Software*, 149, 105310. <https://doi.org/10.1016/j.envsoft.2022.105310>
- Teh, H. Y., Kempa-Liehr, A. W., & Wang, K. I. K. (2020). Sensor data quality: A systematic review. *Journal of Big Data*, 7(1), 1–49. <https://doi.org/10.1186/s40537-020-0285-1>
- Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., & Srikanthan, S. (2009). Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resources Research*, 45(12), W00B14. <https://doi.org/10.1029/2008wr006825>
- Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., & Higdon, D. (2009). Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, 10(3), 271–288. <https://doi.org/10.1515/ijnsns.2009.10.3.273>
- Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., & Robinson, B. A. (2008). Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment*, 23(7), 1011–1026. <https://doi.org/10.1007/s00477-008-0274-y>
- Wellen, C., Kamran-Disfani, A. R., & Arhonditsis, G. B. (2015). Evaluation of the current state of distributed watershed nutrient water quality modeling. *Environmental Science & Technology*, 49(6), 3278–3290. <https://doi.org/10.1021/es5049557>
- Xu, T., Valocchi, A. J., Ye, M., Liang, F., & Lin, Y. F. (2017). Bayesian calibration of groundwater models with input data uncertainty. *Water Resources Research*, 53(4), 3224–3245. <https://doi.org/10.1002/2016wr019512>
- Yang, J., Reichert, P., Abbaspour, K. C., & Yang, H. (2007). Hydrological modelling of the Chaohe basin in China: Statistical model formulation and Bayesian inference. *Journal of Hydrology*, 340(3–4), 167–182. <https://doi.org/10.1016/j.jhydrol.2007.04.006>
- Yen, H., Wang, X., Fontane, D. G., Harmel, R. D., & Arabi, M. (2014). A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. *Environmental Modelling & Software*, 54, 211–221. <https://doi.org/10.1016/j.envsoft.2014.01.004>
- Zheng, Y., & Han, F. (2015). Markov Chain Monte Carlo (MCMC) uncertainty analysis for watershed water quality modeling and management. *Stochastic Environmental Research and Risk Assessment*, 30(1), 293–308. <https://doi.org/10.1007/s00477-015-1091-8>
- Zheng, Y., & Keller, A. A. (2007a). Uncertainty assessment in watershed-scale water quality modeling and management: 1. Framework and application of generalized likelihood uncertainty estimation (GLUE) approach. *Water Resources Research*, 43(8), W08407. <https://doi.org/10.1029/2006wr005345>
- Zheng, Y., & Keller, A. A. (2007b). Uncertainty assessment in watershed-scale water quality modeling and management: 2. Management objectives constrained analysis of uncertainty (MOCAU). *Water Resources Research*, 43(8), W08408. <https://doi.org/10.1029/2006wr005346>
- Zulkifli, S. N., Rahim, H. A., & Lau, W. J. (2018). Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications. *Sensors and Actuators B: Chemical*, 255, 2657–2689. <https://doi.org/10.1016/j.snb.2017.09.078>

## References From the Supporting Information

- Obreza, T. A. (1993). Program fertilization for establishment of orange trees. *Journal of Production Agriculture*, 6(4), 546–552. <https://doi.org/10.2134/jpa1993.0546>