Article

Modeling Nitrogen Dynamics in a Waste Stabilization Pond System Using Flexible Modeling Environment with MCMC

Hussnain Mukhtar 1, Yu-Pin Lin 1,*, Oleg V. Shipin 2 and Joy R. Petway 1

1 Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei 10617, Taiwan; d04622005@ntu.edu.tw (H.M.); d05622007@ntu.edu.tw (J.R.P.)
2 Environmental Engineering and Management Program, School of Environment, Resources and Development, Asian Institute of Technology, Pathum Thani 12120, Thailand; oshipin@ait.asia
* Correspondence: yplin@ntu.edu.tw; Tel.: +886-2-3366-3467

Academic Editor: Jason K. Levy
Received: 19 May 2017; Accepted: 9 July 2017; Published: 12 July 2017

Abstract: This study presents an approach for obtaining realization sets of parameters for nitrogen removal in a pilot-scale waste stabilization pond (WSP) system. The proposed approach was designed for optimal parameterization, local sensitivity analysis, and global uncertainty analysis of a dynamic simulation model for the WSP by using the R software package Flexible Modeling Environment (R-FME) with the Markov chain Monte Carlo (MCMC) method. Additionally, generalized likelihood uncertainty estimation (GLUE) was integrated into the FME to evaluate the major parameters that affect the simulation outputs in the study WSP. Comprehensive modeling analysis was used to simulate and assess nine parameters and concentrations of ON-N, NH3-N and NO3-N. Results indicate that the integrated FME-GLUE-based model, with good Nash–Sutcliffe coefficients (0.53–0.69) and correlation coefficients (0.76–0.83), successfully simulates the concentrations of ON-N, NH3-N and NO3-N. Moreover, the Arrhenius constant was the only parameter sensitive to model performances of ON-N and NH3-N simulations. However, Nitrosomonas growth rate, the denitrification constant, and the maximum growth rate at 20 °C were sensitive to ON-N and NO3-N simulation, which was measured using global sensitivity.

Keywords: flexible modeling environment; waste stabilization pond; nitrogen dynamic; parameterization; sensitivity; MCMC; GLUE; global uncertainty

1. Introduction

A significant number of processes can be used for wastewater treatment prior to its disposal into water bodies [1,2]. A waste stabilization pond (WSP) is an economical, effective and common approach in biological wastewater treatment [3–5], especially in tropical climates due to its simple and manageable operation, low capital and operational cost requirements, and favorable climatic factors [6–9]. WSPs are influenced by stochastically varying environmental factors such as temperature, rainfall, and evaporation [4]. In WSP systems, increased oxygenation attributed to aquatic micro- and macrophytes present during the photosynthesis process has been observed to affect the interfacial oxygen mass transfer across the air–water interface [4]. Numerous studies have indicated that oxygen transfer in WSPs displays high temporal variation, and that mass transfer of oxygen across the air–water interface is a major design parameter [4,10] when modeling WSP systems. Therefore, accurately understanding and modeling the nitrogen dynamics and mechanisms in WSP systems is fundamentally important in assessing performance of, and in designing, WSP systems. WSP systems include complex nitrogen removal mechanisms (i.e., bioassimilation,
volatilization, adsorption and chemical precipitation), and long-term adverse effects of nitrogen-rich effluents on water bodies (i.e., eutrophication) [2,11–18].

Many studies have applied mathematical models to simulate nitrogen dynamics and mechanisms in WSP systems [4]. These include simulations of the rate of oxygen absorption incorporating interfacial oxygen by using a lumped parameter model [4], and mass balance equations using system dynamic modeling [11,19]. Recently, various types of models, such as mathematical, deterministic, stochastic, empirical, and dynamic models, have been developed to simulate the process of nitrogen removal and parameter calibration in waste stabilization and high rate algal ponds [13,16,17,20–22]. Process-based models have been widely used to examine the dynamics of nitrogen and represent the physical–chemical systems in WSP systems, which usually consist of major parameters that must be calibrated to account for environmental heterogeneity and variations [23,24]. Furthermore, model parameter complexity and the process-based model structure are each sources of uncertainty in model predictions. However, uncertainty estimations can be made when considering how the variability of model parameters affects model simulations. A dynamic model with parameter sensitivity and uncertainty analyses for nitrogen simulation in WSP is a potential model for uncertainty estimations, since the model can deal with dynamic mechanisms of simulated processes both deterministically and empirically [23,24]. The Flexible Modeling Environment developed in R package (R-FME) by Soetaert and Petzoldt [25], provides a flexible and complex modeling environment with modeling algorithms, identifiability analysis [24,26], sensitivity, uncertainty and Monte Carlo analysis [24,26] functionalities. Wu and Liu [27] applied FME with the Soil and Water Assessment Tool (SWAT) model, using an R-SWAT-FME framework to perform parameter identifiability, optimization, sensitivity, and uncertainty analyses. Wu et al. [24] applied FME to simulate the spatial pattern of grain yield production for crops and net primary production (NPP) in their studies. Vigiak et al. [28] used FME in simulating sediment and changes in sediment sources during drought periods. To date, FME-based modeling has not been used for wetland modeling.

In the last decade, a number of sensitivity analysis approaches have been developed and widely applied in hydrological and wetland modeling, such as the Metropolis–Hastings algorithm (MH) [29,30], the Hornberger–Spear–Young (HSY) method of global sensitivity analysis [31,32], the Markov chain Monte Carlo (MCMC) method [17,24,26,33,34], and generalized likelihood uncertainty estimation (GLUE) [24,31,35–37]. The above approaches have their advantages, and have been successfully used in parameter sensitivity analysis, optimization, and uncertainty analysis of water quantity and quality models [24,36,38]. GLUE analysis has also been successfully applied in water quantity and quality modeling of wetlands [36,39,40]. Given a specific distribution of inputs, the GLUE analysis procedure estimates the likelihood of all possible outcomes for determinations of behavioral parameter sets of a model [24,36]. The results from any GLUE analysis are an ensemble of behavioral models based on likelihood values [31]. Integrating GLUE analysis into R-FME can therefore provide additional uncertainty analysis for modeling uncertainty in FME [25,41].

The objective of this study is to develop a dynamic simulation model for WSP by using R-FME with MCMC in order to evaluate the nitrogen dynamics and perform a comprehensive modeling analysis, with a focus on parameter optimization, local sensitivity, parameter global sensitivity, and uncertainty analysis. GLUE analysis was integrated into the R-FME model to optimize model parameters and simulate the concentrations of ammonia nitrogen (NH3-N), organic nitrogen (ON-N), and Nitrate (NO3-N) in the study WSP. Global sensitivity and parameter sensitivity analyses for the simulations were performed using MCMC and GLUE.

2. Materials and Methods

Figure 1 illustrates the WSP model development process including calibration, validation, local and global sensitivity analysis, and GLUE analysis using R-FME.
2.1. Site and Experiment Design

Observed data from a pilot-scale WSP system for treating septic tank effluent was used for model calibration and validation. The experiment measured the nitrogen removal capacity of ponds designed for flood control on a university campus in Thailand (at coordinates: 14°04′46.00″ N, 100°36′50.51″ E). A hydraulic retention time of 10 days was maintained in the pond, and the average influent nitrogen concentration was 10.24 mg/L, 16.8 mg/L, and 1.03 mg/L for ON-N, NH₃-N, NO₃-N, respectively. The pond length, width, and depth measured 20.5 m, 1.2 m, and 0.80 m, respectively (Figure 2).

Water samples were collected and monitored weekly for one year during 10:00–11:00 am of the sampling day from June 2014 to July 2015. Influent and effluent water samples from each of the systems (400 mL each) were analyzed for total Kjeldahl nitrogen (TKN), ammonia nitrogen (NH₃-N and NH₄-N), nitrite, and nitrate nitrogen (NOₓ-N). Temperature, dissolved oxygen, and pH were measured at each sampling location. All measurements were conducted according to methodology endorsed by the Standard Methods for the Examination of Water and Wastewater [42]. Statistical procedures and one-way analysis of variance (ANOVA) to evaluate the significance of the differences in influent and effluent were carried out using Statistical Packages for Social Scientists (SPSS version 12.0).

2.2. R-FME Functions Adopted in This Study

R package FME [25,41] was developed to perform comprehensive analysis for inverse modeling with optimization techniques (“Nelder-Mead”, “BFGS”, “CG”, “LBFGS-B”, “SANN”, “Newton”, “Pseudo”), parameter identification, sensitivity analyses, and Monte Carlo analyses. Detailed descriptions of FME functions were introduced by Soetaert and Petzoldt (2010) [25] and are available.
on the R Archive Network at http://CRAN.R-project.org/package5FME. This study developed WSP models with R-FME, as described in Section 2.4. We considered three target variables, including NH3-N, ON-N, and NO3-N, and multiple parameters with allowable ranges for simulating nitrogen dynamics in a WSP. Two functions within R-FME, “sensFun” and “Collin”, were used to estimate local sensitivities and parameter identifiability of nine parameters. The sensFun function used initial parameter values and an allowable parameter range based on the sum of squared residuals and standard deviation of the observed data in order to calculate the local sensitivity of parameters for model output. The Collin function provided collinearity index values to identify the corresponding parameter subsets. Table S1 lists the descriptions of the R-FME functions used in this study.

2.3. GLUE Analysis

Developed by Beven and Binley [35], GLUE analysis executes numerous model runs with generated random parameter values from a priori probability distributions [35,37,43]. Each run is defined as “non-behavioral” when its likelihood value is less than a defined threshold. The GLUE analysis samples from an assumed probability model of uncertain parameters, and estimates the uncertainty in model outputs as a function of likelihood [44]. Detailed discussion of the GLUE method can be found in Beven and Binley (1992, 2014) [31,35]. In this study, several steps of the general GLUE method were adapted for our investigation. The simulated variables of the nitrogen removal model (ON-N, NH3-N, and NO3-N) were compared with observed data using the likelihood functions of the Nash–Sutcliffe coefficient of efficiency (NSE) [38,45]. A likelihood value threshold was used to evaluate non-behavioral and discarded parameters sets. The NSE is used as a goodness-of-fit criterion and as the likelihood measure, defined as follows:

\[
L(\theta|Y) = (1 - \frac{\sigma_{\theta}^2}{\sigma_{\theta|Y}^2}),
\]

where 
\(L(\theta|Y)\) denotes model simulation likelihood measures for a set of observations, \(\sigma_{\theta}^2\) denotes the associated error variance, and \(\sigma_{\theta|Y}^2\) denotes the observed variance. In this study, MCMC was used to generate 5000 parameter sets (within the range of -10% to +10%), whereby each parameter set consisted of nine parameters (Tables 1 and 2). Uniform distribution was assumed, since an unknown a priori distribution is also assumed. Parameter ranges are based on published values from established studies (Table 2).

2.4. Model Construction

A WSP system dynamic model was developed using the R-FME package. R-FME functions were used to write the set of differential equations, and were solved numerically by using the “deSolve” package. It was assumed that the study WSP is completely mixed and acts as a single continuous stirred tank reactor [46]. Mass balanced equations for organic nitrogen (ON-N), ammonia nitrogen (NH3-N), and nitrate nitrogen (NO3-N) concentrations are as follows [21,46]:

\[
\frac{d(\text{Org}-\text{N})}{dt} = \frac{Q_i}{V} (\text{Org} - \text{N})_i - \frac{Q_e}{V} (\text{Org} - \text{N}) + r_m - r_v + r_1 + r_2, \quad (2)
\]

\[
\frac{d(\text{NH}_3-N)}{dt} = \frac{Q_i}{V} (\text{NH}_3 - \text{N})_i - \frac{Q_e}{V} (\text{NH}_3 - \text{N}) + r_1 + r_m - r_v - r_n, \quad (3)
\]

\[
\frac{d(\text{NO}_3-N)}{dt} = \frac{Q_i}{V} (\text{NO}_3 - \text{N})_i - \frac{Q_e}{V} (\text{NO}_3 - \text{N}) + r_n - r_2 + r_d, \quad (4)
\]

where \(Q_i\) and \(Q_e\) are influent and effluent flow rate (m³/d), respectively; \(V\) is volume of the pond (m³); \(r_m\) is nitrification rate (mg/L.d); \(r_v\) is denitrification rate (mg/L.d); \(r_m\) is mineralization rate (mg/L.d), \(r_v\) is net loss of Org-N (mg/L.d); \(r_v\) is volatilization rate (mg/L.d); \(r_n\) is uptake rate of NH3-N by microorganisms (mg/L.d); and \(r_d\) is uptake rate of NO3-N by microorganisms (mg/L.d).

Table 1 lists mass balance reaction rate models used in this study and each model’s process descriptions. This study considers biomass as proportional to organic nitrogen, represented as Org-N [11,13]. Organic nitrogen converts to ammonia through a mineralization process \(r_m\) that can be simulated by first-order kinetics as a function of ON-N concentration [47]. Nitrification \(r_n\) of
NH3-N, where ammonia nitrogen is converted to nitrate through biological oxidation, is modeled as a single step process that is a function of temperature, dissolved oxygen, and pH [13]. The half saturation constant for Nitrosomonas (Kθ), the temperature dependence factor (C_T), and Nitrosomonas growth-inhibiting factor for pH (C_pH) are computed from Downing’s model (Table 1) [13]. In the nitrification rate function, Y_n represents the mass of Nitrosomonas as per unit mass of substrate (mg/mg) and µ_n represents the Nitrosomonas growth rate. A rate-limiting factor for Nitrosomonas growth is pH, which is corrected by C_pH. When pH is below 7.2, the presence of free ammonia reduces the growth of nitrifying bacteria that react, according to equations shown in Table 1. However, when pH is greater than 7.2, no significant inhibition occurs for nitrifying bacteria and C_pH = 1 [13]. Nitrification is also influenced by temperature. An exponential function is used to describe the temperature correction factor (C_T) when T_o equals the reference temperature (15 °C).

Volatilization of ammonia through ammonia stripping is dependent on pH, pond depth (d), temperature (T), and ammonia concentration in the WSP, and is modeled using equations proposed in recent studies [47,48]. Ammonia and nitrate convert into organic compound mostly by autotrophs [11,49]. Photosynthesis is the main means by which autotrophs produce organic compounds and oxygen, which is done by consuming carbon dioxide, water, and nutrients from their surroundings. Regarding microorganism growth (r_1, r_2), microorganisms consume ammonia nitrogen as well as nitrate nitrogen, but prefer ammonia to nitrate for cell synthesis [11]. Preference factors p_1 and p_2 are therefore introduced to represent the ammonia and nitrate uptake by microorganisms, respectively.

In the WSP system, non-biodegradable organic nitrogen settles into sediments and is considered to be the permanent removal of organic nitrogen from the influent. Previous studies showed that 10 days of hydraulic retention time was sufficient for proper sedimentation in ponds and wetlands [50,51] where the sedimentation coefficient value (Rs) varies between 0.001 to 0.1 d^{-1} [46,47]. Net loss of nitrogen to sediment (r_s) in WSP is calculated as a function of ON-N concentration where the sedimentation coefficient (Rs) was obtained from model calibration. Denitrification is performed primarily by heterotrophic bacteria. The rate-limiting factors for denitrification are temperature, nitrite, and nitrate concentration [11]. However, nitrite is not modeled in this study due to very low concentrations in the influent (less than 0.1 mg/L). Nitrification and denitrification within aquatic systems can occur in sediments, and have been studied extensively [52,53]. In treatment ponds and wetlands, however, sediments play a negligible role in the system’s nitrification and denitrification processes [11]. Moreover, if the WSP is working properly, denitrification (r_d) may occur near the bottom layer of the system and is modeled using first-order Arrhenius kinetics [13].

**Table 1. Mass balance reaction rate models.**

<table>
<thead>
<tr>
<th>Process</th>
<th>Model</th>
<th>Source Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mineralization</td>
<td>r_m = 0.002T × (Org−N)</td>
<td>[47,54]</td>
</tr>
<tr>
<td></td>
<td>r_n = μ_n \frac{NH_3−N}{K_θ + NH_4−N} (DO \times K_T × C_pH)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>K_θ = 10^{(0.053(T−15)−5.58)}, C_T = \exp (0.098 \times (T−15)) IF pH &lt; 7.2 THEN C_pH = 1 ELSE C_pH</td>
<td>[54]</td>
</tr>
<tr>
<td></td>
<td>= 1 − 0.833(7.2−pH) ELSE C_pH</td>
<td></td>
</tr>
<tr>
<td>Nitrification</td>
<td>r_s = R_20 \theta^{(T−20)} NO_3−N</td>
<td>[46]</td>
</tr>
<tr>
<td>Sedimentation</td>
<td>r_s = R_s \theta^{(T−20)} NO_3−N</td>
<td>[46]</td>
</tr>
<tr>
<td>Denitrification</td>
<td>r_d = R_20 \theta^{(T−20)} NO_3−N</td>
<td>[46]</td>
</tr>
<tr>
<td>Volatilization</td>
<td>r_v = NH_3−N × 0.0566 × \exp(0.13(T−20)) \times d × (1 + 10^{(0.5−0.037pH)})</td>
<td>[46]</td>
</tr>
<tr>
<td>Microorganisms growth 1</td>
<td>r_1 = \mu_{max} \theta^{(T−20)}NH_3−N \times (Org−N) × P1</td>
<td>[13,55]</td>
</tr>
<tr>
<td>Microorganisms growth 2</td>
<td>r_2 = \mu_{max} \theta^{(T−20)}NO_3−N \times (Org−N) × P2</td>
<td>[55]</td>
</tr>
</tbody>
</table>

Note: r_m = mineralization rate (mg/L.d), r_d = denitrification rate (mg/L.d), r_n = nitrification rate (mg/L.d), r_s = net loss of Org-N (mg/L.d), r_v = volatilization rate (mg/L.d), r_1 = uptake rate of NH3-N by microorganisms (mg/L.d), r_2 = uptake rate of NO3-N by microorganisms (mg/L.d), d = depth of pond (m).
The performance of the proposed model was evaluated by comparing its simulated output with measured outflow concentrations of NH$_3$-N, ON-N, and NO$_3$-N using NSE values. The “Pseudo” optimization technique with 500 iterations was used to obtain the optimal parameter values, and was based on the first six months of weekly observations of the three target variables between June 2014 and December 2014. Pseudo optimization was employed to bring the algorithm near vicinity, after which the MCMC algorithm was used to maximize the NSE value. Iterations were stopped upon criteria satisfaction (NSE > 0.5). The model was then validated using the observed data from the subsequent six months, January–July 2015.

3. Results

3.1. Local Sensitivity and Parameter Analyses

Table 2 summarizes the parameter ranking results of the dimensionless parameter sensitivity values. A value with high absolute sensitivity index (L1) represents high importance and ranking. Compared to other narrow range parameters, the L1 results show that $\theta$, $K_1$, and $K_2$ are ranked 1st, 2nd, and 3rd, respectively, followed by $R_s$, $\mu_n$, and $\mu_{max}$. Parameter identifiability is important for parameter estimation to examine the degree of near-linear dependence of the sensitivity function of all possible parameter subsets. Sensitivity measurements estimate the identifiability of a parameter on an individual basis, while collinearity estimates a parameter subset’s compensation in joint estimation. A collinearity index value below 15 indicates that the corresponding parameter subsets are highly identifiable. In this study, the selected parameters were highly identifiable as the collinearity index value for various parameters was always less than four. The collinearity index for the selected nine parameters are shown in Figure S1.

Table 2. Model calibration parameter values with summary statistics of parameters’ sensitivities.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Literature Value</th>
<th>Calibrated Value</th>
<th>References</th>
<th>L1 *</th>
<th>L2</th>
<th>Rank *</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$</td>
<td>Arrhenius constant</td>
<td>1.01 to 1.09</td>
<td>1.01</td>
<td>[56]</td>
<td>1341.9</td>
<td>171.0</td>
<td>1</td>
</tr>
<tr>
<td>$R_s$</td>
<td>ON. Sedimentation constant (d$^{-1}$)</td>
<td>0.015</td>
<td>0.050</td>
<td>[46,47]</td>
<td>48.1</td>
<td>6.4</td>
<td>7</td>
</tr>
<tr>
<td>$K_1$</td>
<td>Nitrosom. half sat. const. (mg/L)</td>
<td>0.13 to 0.15</td>
<td>1.26</td>
<td>[57]</td>
<td>1157.9</td>
<td>151.9</td>
<td>2</td>
</tr>
<tr>
<td>$K_2$</td>
<td>Ammonium. half sat. const. (mg/L)</td>
<td>18.0</td>
<td>16.20</td>
<td>[55]</td>
<td>870.6</td>
<td>115.5</td>
<td>3</td>
</tr>
<tr>
<td>$R_20$</td>
<td>Nitrate. half sat. const. (mg/L)</td>
<td>2.0</td>
<td>2.0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>$Y_n$</td>
<td>Yield coeff.</td>
<td>0.0 to 1.0</td>
<td>0.26</td>
<td>[56]</td>
<td>274.6</td>
<td>37.3</td>
<td>4</td>
</tr>
<tr>
<td>$\mu_{max}$</td>
<td>Nitrosom. (g microbes/g N)</td>
<td>0.13</td>
<td>0.163</td>
<td>[58]</td>
<td>177.4</td>
<td>23.1</td>
<td>5</td>
</tr>
<tr>
<td>$\mu_n$</td>
<td>Nitrosom. growth rate (d$^{-1}$)</td>
<td>0.008</td>
<td>0.009</td>
<td>[21,55]</td>
<td>144.2</td>
<td>18.1</td>
<td>6</td>
</tr>
</tbody>
</table>

* L1 = $\sum_{i,j} |S_{ij}| / n$ and L2 = $\sqrt{\sum_{j} S_{i,j}^2 / n}$ are the L1 and L2 norm, respectively, and $S_{ij}$ is sensitivity of parameter $i$ for model variable $j$. Note: L1 is high absolute sensitivity index; L2 is used for weighted least squares estimation of parameter subset. Large differences between L1 and L2 indicate a high variability, or outliers in $S_i$.

3.2. Parameter Optimization

Figure 3 shows the weekly time series of three observed and simulated variables (i.e., ON-N, NH$_3$-N, and NO$_3$-N) during both the calibration and validation periods. The calibrated parameter values are listed in Table 2. It was observed that many of the parameters were either identical to or within the range of those values obtained by other researchers. Results show that the optimized $\theta$ (Arrhenius constant) and $K_2$ (“Ammonium. half sat. const.”) are very close to their lower bound (1.01).
This value is in line with most other applications. Some literature reported a θ value of 1.02 for waste stabilization ponds [14,47] while Wang et al. [43] reported a higher θ value (1.04) when simulating pollution removal in a constructed wetland, than those found in the current study (1.01).

Figure 3. Weekly time series of observed and simulated ammonia nitrogen (NH3-N), organic nitrogen (ON-N), and Nitrate (NO3-N) during a one-year calibration and validation period from 2014–2015. Sensitivity ranges of weekly ON, NH3, and NO3 were generated with the Markov chain Monte Carlo (MCMC) application. Gray shading indicates minimum and maximum model responses at each time step.

3.3. Model Performance Evaluation

A significant difference in the nitrogen concentration of influent and effluent (p < 0.001) was observed. Table S2 shows ANOVA test results for ON-N concentrations in influent and effluent. Similar results for NH3-N and NO3-N are not shown. Mean removal efficiency for ON-N, NH3-N, and NO3-N was 46.3%, 62.8%, and 61.4%, respectively. However, there was no significant difference (p > 0.1) in removal efficiency during calibration and validation periods, since there is minimal seasonal variation in tropical climates [2,9].

Table 3 summarizes the statistical analysis for the three variables used in this study for both calibration and validation. The results show that the simulations of the three targeted variables (ON-N, NH3-N, and NO3-N) were quite close and fit well with their corresponding observations for both calibration and validation periods. The NSE for NH3-N and ON-N were 0.69–0.70 and 0.53–0.66, respectively, while the corresponding correlation constant (r) values for both were 0.83–0.87 and 0.76–0.84, respectively. The NSE value for NO3-N, however, was lower when compared to NH3-N and ON-N.

Table 3. Calibration and validation results for simulated concentrations with measured concentrations of ON, NH3, and NO3.

<table>
<thead>
<tr>
<th>Water Quality</th>
<th>Nash-Sutcliffe Coefficient (NSE)</th>
<th>Correlation Constant (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calibration</td>
<td>Validation</td>
</tr>
<tr>
<td>ON</td>
<td>0.66</td>
<td>0.53</td>
</tr>
<tr>
<td>NH3</td>
<td>0.70</td>
<td>0.69</td>
</tr>
<tr>
<td>NO3</td>
<td>0.58</td>
<td>0.64</td>
</tr>
</tbody>
</table>

3.4. Uncertainty and Global Sensitivity Analyses

Model uncertainty for the WSP model, with an allowable parameter range, was estimated with R-FME-based uncertainty analysis and GLUE analysis. The Bayesian–MCMC method has been widely used in uncertainty analysis [59]. Parameter uncertainty, with an allowable parameter range, was quantified with the R-FME “modMCMC” function.

Figure 4 shows the MCMC sample pair plot based on 5000 runs, with a scatterplot matrix of coefficients and marginal frequency distributions for each parameter in top, bottom, and diagonal
plots [23]. The results show that the parameters (Arrhenius constant (θ) and sedimentation constant (R_s)) are related, because both are closely linked to nitrogen simulation in ponds [43]. However, there is no positive correlation between other parameters. Regarding parameter identifiability, the correlation coefficient values among parameters are small (<0.25), so the parameter is identifiable [33].

The “modCRL” (global sensitivity analysis) function was used to estimate the effect of a uniform parameter distribution on the mean model output in their allowable parameter range. Figure 5 shows the results of the modMCMC global sensitivity analysis of ammonia nitrogen (NH_3-N), organic nitrogen (ON-N), and nitrate (NO_3-N) for the one-year calibration period (2014–2015), as a function of uniform distribution of the parameter values. The results indicate that the MCMC parameter values for each iteration are well mixed within upper and lower ranges of allowable parameter values. There is a clear linear and negative relationship between NH_3-N and θ (Arrhenius constant) \( r = -0.95, p < 0.01 \), but a positive relationship between ON-N and θ. However, there is a near-linear and negative relationship between output variable NO_3-N and the parameters Y_n (yield coeff. Nitrosomonas) and R (denitrification constant).

In GLUE analysis, a range of model responses was estimated to simulate the proposed model output using all behavioral parameter sets, which are either accepted as behavioral (likelihood value ≥0) or rejected as non-behavioral (Likelihood value <0), and are a function of its goodness of fit. Figure 6 shows the results of the GLUE analysis, where dots represent the Nash–Sutcliffe coefficient value for each iteration of MCMC within the parameter range, and the red horizontal line indicates the NSE minimum criteria. The scatter plots also revealed the feasible values of all parameter ranges within the calibration values [31]. The parameter range and sensitivity of nine parameters individually sensitive to ammonia nitrogen (NH_3-N), organic nitrogen (ON-N), and Nitrate (NO_3-N), are associated with likelihood values. GLUE sensitivity analysis results indicate that sensitivity parameters included θ (Arrhenius constant) in simulating NH_3-N and ON-N, and that slight or no sensitivity is observed for the remaining parameters set [17]. Figure 7 shows cumulative density functions of behavior and non-behavior solutions based on the NSE likelihood functions for the sensitive parameter (Arrhenius constant). The cumulative distributions represent the behavioral and non-behavioral distributions, which show the NSE ≥0 and NSE <0, respectively. The results show that the behavioral and non-behavioral distributions for NO_3-N are close to each other when compared to ON-N and NH_3-N.

![Figure 4](image.png)

**Figure 4.** Pairs plot of the Markov chain Monte Carlo (MCMC) samples for the nine parameters. Note: Top is the pairwise relationship; bottom is the correlation coefficient; diagonal is marginal distribution.
Figure 5. Global sensitivity analysis results of ammonia nitrogen (NH3-N), organic nitrogen (ON-N), and nitrate (NO3-N) for the one-year calibration and validation period (2014–2015), as a function of uniform distribution of the parameter values.

Figure 6. Scatter plot of Nash–Sutcliffe coefficient of efficiency (NSE)-sensitive parameters based on generalized likelihood uncertainty estimation (GLUE) method for ammonia nitrogen (NH3-N), organic nitrogen (ON-N), and Nitrate (NO3-N). Note: dots represent the Nash–Sutcliffe coefficient value for each iteration; red horizontal line indicates the NSE minimum criteria.
Figure 7. Cumulative density functions of behavior and non-behavior solutions based on the NSE likelihood functions for (a) ammonia (NH$_3$-N), (b) organic nitrogen (ON), and (c) nitrate (NO$_3$-N). Black indicates behavioral likelihood (NSE $\geq$0); gray indicates non-behavioral likelihood (NSE <0).

4. Discussion

4.1. Parameter Optimization and Model Performance

The accuracy of predicted organic nitrogen concentration in effluent was improved by increasing the calibrated value of the sedimentation constant ($R_s$) to 0.050, instead of using the value reported in literature (0.015) [47]. $K_1$, $K_2$, and $K_3$ represent the Nitrosomonas half-saturation constant, ammonium half-saturation constant, and nitrate half-saturation constant, respectively [21]. $K_1$ was increased from 0.15 to 1.26 to improve the simulation of ammonia and nitrate. The calibrated value for $K_3$ was the same as the initial value (2.0), and likely caused by high proportions of NH$_3$-N in influent, since microorganisms tend to prefer ammonia over nitrate, and nitrate consumption does not occur until the ammonia is almost completely consumed [60]. In this study, the calibrated $R_20$ value was within the “literature range” obtained by other researchers (0.0–1.0) [47,56]. Yet, the simulation of pollution removal in a constructed wetland by Wang et al. [43] reported an even higher value (3.74) than those found in the current study (0.26). Model performance can be improved by further reducing the $\theta$ value (Arrhenius constant); however, in this study, real-world application was prioritized over obtaining a better parameter value.

Similar to results from reviewed literature, this study observed that the major pathway of nitrogen removal in WSP could be ammonia volatilization [61,62] influenced by environmental factors, macro-microphyte abundance, and influent NH$_3$-N concentrations. In the current study, volatilization played only a negligible role in nitrogen removal. Similar to results reported by Valero and Mara [63], the ammonia volatilization process accounted for only 3% of the total nitrogen removal in the WSP system.

Some studies showed that nitrification-denitrification is not likely to be a major pathway for nitrogen removal in pond systems. However, mineralization, nitrogen uptake by microorganism growth, and ammonia stripping are the most important mechanisms in WSP systems operating at various hydraulic retention times [64]. Moreover, microorganisms are assumed to consume NO$_3$-N only when NH$_3$-N is not available. This may explain the low NSE value for NO$_3$-N. Nutrient removal in algal-based wastewater treatment systems could be affected by the interaction of algae with other microorganisms, such as bacteria [17,65,66]. Increased pH and inhibitory metabolites attributed to algae presence has been observed to affect bacterial activity [66]. On the other hand, algae can be damaged by the soluble cellulose enzymes released by some bacterial species. The interactions between algae and bacteria are complex and will be influenced by environmental conditions [67]. Further studies are therefore needed to understand the interactions between algae and bacteria to improve our understanding about nitrogen removal dynamics in WSP systems.
4.2. Uncertainty and Global Sensitivity Analyses

The parameters with allowable ranges were analyzed with MCMC and GLUE procedures to estimate uncertainty in nitrogen simulations. The results indicate that both methods perform similarly in terms of determining best parameter values. The NSE index scatter plots provide feasible values within the parameter range over the most calibration values. However, the parameter sensitivity results estimated by MCMC and GLUE were significantly different. MCMC sensitivity analysis results indicated that sensitivity parameters included $\theta$, $R_20$, $Y_n$ and $\mu_n$. However, GLUE sensitivity analysis indicated that $\theta$ was the only sensitive parameter in model simulations, while no sensitivity was observed for the remaining parameters set. In GLUE uncertainty analyses, parameter sensitivity is subject to certain threshold values that separate the total parameter set into behavioral and non-behavioral parameters as a function of goodness of fit and measured by comparing the model simulation results and observations [39]. In this study, all parameter sets that resulted in NSE <0 were discarded. Due to this, MCMC sensitive parameters ($R_20$, $Y_n$ and $\mu_n$) were not variants of GLUE behavioral parameters (NSE >0), and cumulative distribution for behavioral solutions always equaled one for a given parameter range. Likelihood functions with threshold values have a significant effect on the parameter uncertainty estimation. It is recommended that modelers choose the best option based on the factors involved in their decision-making approach. The *Nitrosomonas* species accomplishes an important function of aerobic ammonia oxidation [17,18]. Ochoa et al. [17] indicated that the yield factor for heterotrophic bacteria contributed to the simulation of NO$_3$. In this study, $R_20$, $Y_n$, and $\mu_n$ are sensitive in MCMC analysis of NO$_3$ concentration simulations. Though sedimentation of organic nitrogen is a major nitrogen removal mechanism reported by Møller et al. [16], in this study, Rs (sedimentation constant) was only slightly sensitive in the MCMC analysis of ON concentration simulations.

Nash and Stucliffe’s efficiency measure with a threshold value has been widely used to identify a set of behavioral models [31]. Parameter sensitivity can also be described by orientation to and gradation of the likelihood distribution of functions. A parameter will be highly sensitive to a likelihood function if cumulative distributions are far apart from each other. On the other hand, a parameter will be less, or not sensitive, to a given set of likelihood functions if cumulative distributions of likelihood functions are close to each other [39]. Differing from the results of Møller et al. [16] but similar to that of Ochoa et al. [17], this study’s sensitivity analysis results show that the Arrhenius constant is highly sensitive to ammonia (NH$_3$-N) and organic nitrogen (ON) simulation, as behavioral and non-behavioral distribution tends to be far apart. Alternatively, the closely distributed behavioral and non-behavioral solutions show that the Arrhenius constant is not sensitive to nitrate (NO$_3$-N) simulation. These results reiterate the need for the careful selection of likelihood functions, due to their strong influence on parameter sensitivity [31,68].

5. Conclusions

This study presented an approach that integrates R-FME with MCMC and GLUE in simulations of concentrations of ON-N, NH$_3$-N, and NO$_3$-N at a pilot WSP system in Thailand. Using the proposed approach, WSP models were developed in R-FME to successfully derive: an optimal parameter set; global sensitive parameters for ON-N, NH$_3$-N, and NO$_3$-N concentrations; and delineated sensitive sets of parameters from GLUE behavior sets based on predefined thresholds of model performance. The major advantages of using an FME-GLUE approach include its modeling comprehensiveness, which incorporates parameter sensitivity analysis and global sensitivity analysis for application to other physical-chemical processes. The results indicate that the Arrhenius constant is the parameter sensitive to simulated ON-N concentrations, the denitrification constant, and the yield coefficient. Further, the Arrhenius constant was the only parameter sensitive to model performances of ON-N and NH$_3$-N simulations. The *Nitrosomonas* growth rate, denitrification constant, and the maximum growth rate at 20 °C were sensitive to simulated concentrations of ON-N and NO$_3$-N, which was measured using global sensitivity. The proposed approach is a potential tool to indicate and quantify the model parameter uncertainty of various target variables for WSP
systems. Additionally, our proposed approach, which uses R functions, is adaptable for users to develop their own models or additional functions based on model structures.

**Supplementary Materials:** The following are available online at www.mdpi.com/1660-4601/14/7/765/s1, Figure S1: Collinearity index for selected 9 parameters, Table S1: List of R-FME functions used, Table S2: One-way analysis of variance (ANOVA) for ON-N. Similar results for NH3-N and NO3-N are not shown.

**Acknowledgments:** The authors would like to thank National Taiwan University for partially finical support.

**Author Contributions:** The scope of this study was developed by Yu-Pin Lin and Hussnain Mukhtar. The first manuscript draft was written by Yu-Pin Lin and Hussnain Mukhtar, and was substantially revised by Yu-Pin Lin, Hussnain Mukhtar and Joy Petway.

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

**References**


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).