International Journal of
Environmental Research and Public Health

MDPI

Article

Application of Parameter Optimization Methods Based on Kalman Formula to the Soil—Crop System Model

Qinghua Guo and Wenliang Wu *

Beijing Key Laboratory of Biodiversity and Organic Farming, College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193, China
* Correspondence: wuwenl@cau.edu.cn

Abstract: Soil–crop system models are effective tools for optimizing water and nitrogen application schemes, saving resources and protecting the environment. To guarantee model prediction accuracy, we must apply parameter optimization methods for model calibration. The performance of two different parameter optimization methods based on the Kalman formula are evaluated for a parameter identification of the soil Water Heat Carbon Nitrogen Simulator (WHCNS) model using mean bias error (ME), root-mean-square error (RMSE) and an index of agreement (IA). One is the iterative local updating ensemble smoother (ILUES), and the other is the DiffeRential Evolution Adaptive Metropolis with Kalman-inspired proposal distribution (DREAMkzs). Our main results are as follows: (1) Both ILUES and DREAMkzs algorithms performed well in model parameter calibration with the RMSE_Maximum a posteriori (RMSE_MAP) values were 0.0255 and 0.0253, respectively; (2) ILUES significantly accelerated the process to the reference values in the artificial case, while outperforming in the calibration of multimodal parameter distribution in the practical case; and (3) the DREAMkzs algorithm considerably accelerated the burn-in process compared with the original algorithm without Kalman-formula-based sampling for parameter optimization of the WHCNS model. In conclusion, ILUES and DREAMkzs can be applied to a parameter identification of the WHCNS model for more accurate prediction results and faster simulation efficiency, contributing to the popularization of the model.

Keywords: data assimilation; Bayesian calibration; WHCNS; parameter multimodal distribution; sampling efficiency

check for updates

Citation: Guo, Q.; Wu, W. Application of Parameter Optimization Methods Based on Kalman Formula to the Soil—Crop System Model. *Int. J. Environ. Res. Public Health* **2023**, *20*, 4567. https://doi.org/10.3390/ijerph20054567

Academic Editor: Lingxin Chen

Received: 28 January 2023
Revised: 26 February 2023
Accepted: 2 March 2023
Published: 4 March 2023

Copyright: © 2023 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 (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Unreasonable water and nitrogen management practices in intensive agricultural systems have caused serious resource waste and environmental pollution [1]. Field experiments and numerical simulations are the main approaches to optimize water and nitrogen application schemes for the construction of high-yield and low-pollution cropping systems [2–4]. Due to the long time consumption and high cost of field experiments, numerical models are crucial tools for analyzing how soil–crop systems respond to field management schemes. There are many relevant soil–crop system models, such as RZWQM [5], WNMM [6], SWAT [7] and SPWS [8]. DRAINMOD is often combined with other models, such as ADAPT, RZWQM2 and DSSAT [9–12]. There are also many other soil–crop models, such as HERMES [13], CHANI-EPIC [14], DAYCENT [15] and NLEAP-GIS [16]. DNDC can simulate the processes responsible for the production, consumption and transport of nitrous oxide [17]. The soil Water Heat Carbon Nitrogen Simulator (WHCNS) model can be applicable to studies of water and nitrogen management under the complex conditions of intensive cropping systems in North China [18].

However, parameters must be calibrated and validated before the application of soil–crop system models to ensure prediction accuracy [19–22]. There are several cases for exploring the calibration of system models through combining parameter optimization methods

with state variable observations [23,24]. Quantum-behaved Particle Swarm Optimization (QPSO) was used for the parameter calibration of the RZWQM2 model [21,25]. The Bayesian method was also useful as a method for quantifying uncertainty [26]. The RZWQM2 model was calibrated by the generalized likelihood uncertainty estimation (GLUE) method, suggesting the application of likelihood function with probability theory [27]. GLUE and Markov Chain Monte Carlo (MCMC) were compared for the parameter identification of the APSIM-Sugar model [28,29]. The DREAM algorithm was introduced for the parameter optimization of the STICS model and DayCent agroecosystem model [30,31]. However, the applicability of current popular parameter inversion methods to the data assimilation of soil–crop system models still needs further exploration.

In recent years, data assimilation methods have rapidly developed and are widely used in various fields. Ensemble Kalman filter (EnKF) can be considered for parameter optimization in nonlinear problems [32]. As its extension, ensemble smoother (ES) is a more efficient data assimilation method, which avoids updating model states and parameters simultaneously [33]. Surrogate-based iterative ES can also increase computation efficiency by constructing surrogates using the Gaussian process [34], polynomial chaos and so on. ES can also be coupled with other methods to solve data assimilation in non-Gaussian-distributed conductivity fields [35,36]. However, neither EnKF nor ES perform well when the model parameter distributions are multimodal. An iterative local updating ensemble smoother (ILUES) was proposed to explore the possible multimodal distributions of the hydrologic model's parameters [37]. Currently, MCMC algorithms are relatively popular for exploring multimodal parameter distributions, even though they are proposed based on the Gaussian assumption. However, the computational cost is prohibitive, especially for a high-dimensional model. Implementing the Kalman formula for the DREAM algorithm's proposal distribution improves sampling efficiency [38]. The introduction of the above-mentioned newly proposed parameter optimization methods for the data assimilation of soil–crop systems can improve the model's prediction accuracy and promote its wide application.

We hypothesize that the ILUES and DREAMkzs data assimilation methods will increase the WHCNS model's simulation accuracy and calibration efficiency. Our study objectives are to (i) compare the efficiency of exploring parameter space between ILUES and ESMDA using synthetic soil water content derived by the WHCNS model, (ii) evaluate the feasibility of exploring the multimodal distribution of the ILUES algorithm's model parameters, and (iii) compare the WHCNS model's sampling efficiency for DREAMkzs and DREAMzs using soil water content observations.

2. Materials and Methods

2.1. Introduction of Data Assimilation Methods

2.1.1. Basic Theory of Iterative Local Updating Ensemble Smoother

The general form of ES is shown as follows [39]:

$$\Delta \mathbf{m} = \mathrm{G}\left(\Delta \mathbf{y}\right), \tag{1}$$

where $\Delta\mathbf{m}$ is the update vector, $\Delta\mathbf{y}$ is the innovation vector, and G is a mapping from $\Delta\mathbf{y}$ to $\Delta\mathbf{m}$, which is defined by the Kalman gain matrix $\mathbf{K}$.

The essence of the ILUES method is to update local ensembles of each sample instead of the global ensembles. The local ensembles for a realization $\mathbf{m}_j^f$ can be derived through the following equation [40]:

$$J\left(\mathbf{m}\right) = \mathrm{a} * J_1\left(\mathbf{m}\right) / J_1^{max} + \mathrm{b} * J_2\left(\mathbf{m}\right) / J_2^{max}, \tag{2}$$

where $J_1\left(\mathbf{m}\right) = [\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_D}^{-1} [\mathbf{d} - f(\mathbf{m})]$ represents the distance between the simulation results $f(\boldsymbol{m})$ and the measurements $\mathbf{d}$; $J_2\left(\mathbf{m}\right) = [\mathbf{m} - \mathbf{m}_j^f]^T \mathbf{C}_{\mathbf{MM}}^{-1} [\mathbf{m} - \mathbf{m}_j^f]$ represents the distance between the realization $\mathbf{m}_j^f$ and parameters $\mathbf{m}$, and $\mathbf{C_{MM}}$ represents the model

parameter auto-covariance matrix; and the maximum values of J_1 (**m**) and J_1 (**m**) are defined as J_1^{max} and J_2 (**m**) is J_2^{max}, respectively. $\mathbf{M}_{j,L}^{f} = [\mathbf{m}_{j,1}^{f}, \ldots, \mathbf{m}_{j,N_L}^{f}]$ containing the realizations with $N_L = \alpha N_e$ and smallest J is defined as the local ensemble for $\mathbf{m}_j^f$. Then, ILUES updates $\mathbf{M}_{j,L}^{f}$ to $\mathbf{M}_{j,L}^{a}$ using the following equation:

$$\mathbf{m}_{j,l}^{a} = \mathbf{m}_{j,l}^{f} + \mathbf{C}_{\mathbf{MD}}^{L,f}(\mathbf{C}_{\mathbf{DD}}^{L,f} + \mathbf{C}_{\mathbf{D}})^{-1}[d_j - f(\mathbf{m}_{j,l}^{f})], \quad (3)$$

where $\mathbf{C}_{\mathbf{MD}}^{L,f}$ represents the cross-covariance matrix between the model parameters $\mathbf{M}_{j,L}^{f}$ and the corresponding simulation results $\mathbf{D}_{j,L}^{f}$, and $\mathbf{C}_{\mathbf{DD}}^{L,f}$ represents the $\mathbf{D}_{j,L}^{f}$ auto-covariance matrix, where $l = 1, \ldots, N_L$. Then, the possible multimodel distributions of model parameters may be achieved by choosing the updated realizations $\mathbf{m}_j^a$ randomly from the updated local ensemble $\mathbf{M}_{j,L}^{a}$.

The multiple data assimilation (MDA) scheme was applied as a supplementary for the model parameter update to achieve better optimization results for nonlinear problems [41,42]. The iterative form of an ES, which is defined as an ESMDA, assimilates the observations multiple times. The factor α_i should be applied to inflate the measurement error for obtaining the reasonable results in iteration i. The factor should obey $\sum_{i=1}^{N} 1/{\alpha_i}^2 = 1$, where N is the total assimilation time.

2.1.2. Theory of the DREAMkzs Algorithm

The formula of the Kalman-inspired proposal distribution, which accelerates the convergence of the chain to the posterior distribution, is shown as follows:

$$\boldsymbol{\theta}_p = \boldsymbol{\theta}_{(t-1)} + \boldsymbol{K}\mathbf{r}_{(t-1)} + \boldsymbol{K}\boldsymbol{\epsilon}_{(t-1)}, \quad (4)$$

where $\boldsymbol{\theta}$ is a vector of model parameters; $r_{(t-1)}$ represents the residual vector of parameters $\boldsymbol{\theta}_{(t-1)}$, and $\boldsymbol{\epsilon}_{(t-1)}$ denotes a vector of random draw from the distribution of measurement errors ($\boldsymbol{\epsilon}_{(t-1)} \sim \mathrm{N}(0, \mathbf{R})$, with $\mathbf{R}$ as the covariance matrix of measurement errors. The Kalman gain $\mathbf{K}$ is defined as:

$$\mathbf{K} = \mathbf{C}_{\boldsymbol{\theta}\mathbf{d}}(\mathbf{C}_{\mathbf{dd}} + \mathbf{R})^{-1}, \quad (5)$$

where $\mathbf{C}_{\boldsymbol{\theta}\mathbf{d}}$ signifies the cross-covariance matrix of model parameters and simulation results, and $\mathbf{C}_{\mathbf{dd}}$ denotes the covariance matrix of model simulation results.

The candidate can be derived from the samples in the archive, through the mixture of parallel direction, snooker and Kalman trial moves. However, the Kalman-inspired proposal distribution introduces asymmetry even though it can shorten the burn-in processes. The approaches that can remedy this defect at present will considerably deteriorate the sampling efficiency. To avoid this disadvantage, the Kalman-inspired proposal distribution is only applied to the first T_k steps of the Markov chain, after which the parallel direction and snooker proposal distributions are used for sampling.

2.2. Introduction of Forward Model Information

2.2.1. WHCNS Model

The WHCNS, which has a detailed introduction in the literature [43], was selected as the forward soil–crop system model for our simulation. The depth of the soil profile was 180 cm, assuming that nitrogen moved beyond this depth could not be utilized by the crop. In total, 32 uncertain hydraulic parameters are listed in Table 1. The soil profile was divided into eight layers with four parameters for each layer. Considering the 8 crop parameters and 5 nitrogen transformation parameters shown in Table 2, there were 45 parameters that needed to be calibrated in all. The prior distributions of the parameters were assumed to be uniform distributions, and the ranges are listed in Tables 1 and 2.

Table 1. Soil hydraulic parameters selected for calibration.

Name		Saturated Hydraulic Conductivity	Saturated Soil Water Content	Field Capacity	Wilting Point
Symbol		Ks	SAT	FC	PWP
Unit		cm/d	cm^3/cm^3	cm^3/cm^3	cm^3/cm^3
Interval	1	[56.59–69.17]	[0.30–0.36]	[0.20–0.25]	[0.11–0.13]
	2	[72.58–88.70]	[0.32–0.40]	[0.18–0.22]	[0.09–0.12]
	3	[46.22–56.50]	[0.32–0.40]	[0.18–0.22]	[0.08–0.10]
	4	[63.50–77.62]	[0.23–0.29]	[0.18–0.22]	[0.06–0.08]
	5	[29.81–36.43]	[0.26–0.31]	[0.16–0.19]	[0.06–0.08]
	6	[31.10–38.02]	[0.24–0.30]	[0.15–0.18]	[0.04–0.06]
	7	[37.37–45.67]	[0.23–0.28]	[0.14–0.18]	[0.04–0.06]
	8	[56.38–68.90]	[0.22–0.26]	[0.10–0.13]	[0.06–0.08]

Table 2. Nitrogen transformation and crop parameters selected for calibration.

Types of Parameters	Description	Symbol	Unit	Interval
Nitrogen transformation parameters	Maximum nitrification rate	Vn	$g/(m^3 \cdot d)$	[5–15]
	Nitrification semi saturation constant	Kn	g/m^3	[25–75]
	Denitrification ratio constant	Kd	-	[0.50–1.50]
	Empirical constant of denitrification	Ad	-	[0.05–0.15]
	First order kinetic constant of ammonia volatilization	Kv	d^{-1}	[0.015–0.045]
Crop parameters	Initial crop coefficient	K_{ini}	-	[0.36–0.54]
	Medium term crop coefficient	K_{mid}	-	[1–1.5]
	Late crop coefficient	K_{end}	-	[0.64–0.96]
	Accumulated temperature from emergence to maturity	T_{sum}	°C	[1480–2220]
	Maximum specific leaf area	SLA_{max}	m^2/kg	[24–36]
	Minimum specific leaf area	SLA_{min}	m^2/kg	[8–12]
	Maximum assimilation rate	AMAX	$kg/(hm^2 \cdot h)$	[36–54]
	Maximum root depth	R_{max}	cm	[96–144]

Mean bias error (ME), root mean square error (RMSE) and index of agreement (IA) are three different statistics applied for the evaluation of the model performance. The corresponding equations for them are as follows:

$$\mathrm{ME} = \sum_{i=1}^{n} \frac{\mathrm{Si} - \mathrm{Oi}}{n}, \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\mathrm{Si} - \mathrm{Oi})^2}{n}}, \tag{7}$$

$$\mathrm{IA} = 1 - \frac{\sum_{i=1}^{n} (\mathrm{Si} - \mathrm{Oi})^2}{\sum_{i=1}^{n} (|\mathrm{Si} - \overline{\mathrm{O}}| + |\mathrm{Oi} - \overline{\mathrm{O}}|)^2}, \tag{8}$$

where n is the number of observations, Si is the simulated value, Oi is the measured value, and $\overline{\mathrm{O}}$ is the mean of the measured values. The RMSE represents the average difference between the simulations and the observations. The range of IA is limited to 0–1, and a better model performance is indicated if the value is closer to 1.

2.2.2. Description of Field Experiment Conditions

The study area was located within Alxa Left Banner, Inner Mongolia, China. The total potential evaporation is 20 times more than the average annual precipitation, which reaches 116 mm/year. The oasis cropping system is single crop growth during middle April to early October, which is dominated by spring maize. Irrigation is typically applied

as flood irrigation and depends on groundwater, ranging from 40 to 70 m in depth. Typical N fertilizer application rates are about 280–350 kg N ha^{-1} [44]. The field experiment was undertaken at the experimental site over two spring maize growth periods. Four treatments were replicated three times in the 20 m × 20 m experimental plots each year. Maize was planted on April 12 and harvested on October 18. Volumetric soil water content was measured from 16 points of the soil profile by taking a value every 10 cm from a depth of 20 cm to 170 cm on 17 days (the 41st, 52nd, 54th, 63rd, 70th, 74th, 82nd, 91st, 95th, 104th, 109th, 115th, 124th, 135th, 146th, 154th and 160th day after sowing), using time domain reflectometry (TDR) probes. Detailed information on water and N management practices are presented in Table 3 [45]. Four irrigation-fertilizer treatments, namely $I_{std}N_{std}$ (W_1N_1), $I_{std}N_{csv}$ (W_1N_2), $I_{csv}N_{std}$ (W_2N_1) and $I_{csv}N_{csv}$ (W_2N_2), can be derived by combining two irrigation treatments with two N-fertilization applications.

Table 3. Irrigation and fertilizer schemes of spring maize.

Treatment	Date of Scheduled Irrigation/Fertilizer Application						Seasonal Total
	2008	3 June	21 June	13 July	4 August	29 August	
	2009	1 June	22 June	13 July	1 August	23 August	
				Irrigation (mm)			
Istd		150	150	150	150	150	750
Icsv		105	105	120	120	120	570
				N fertilization (kg Urea-N/ha)			
Nstd		138					138
Ncsv		92					92

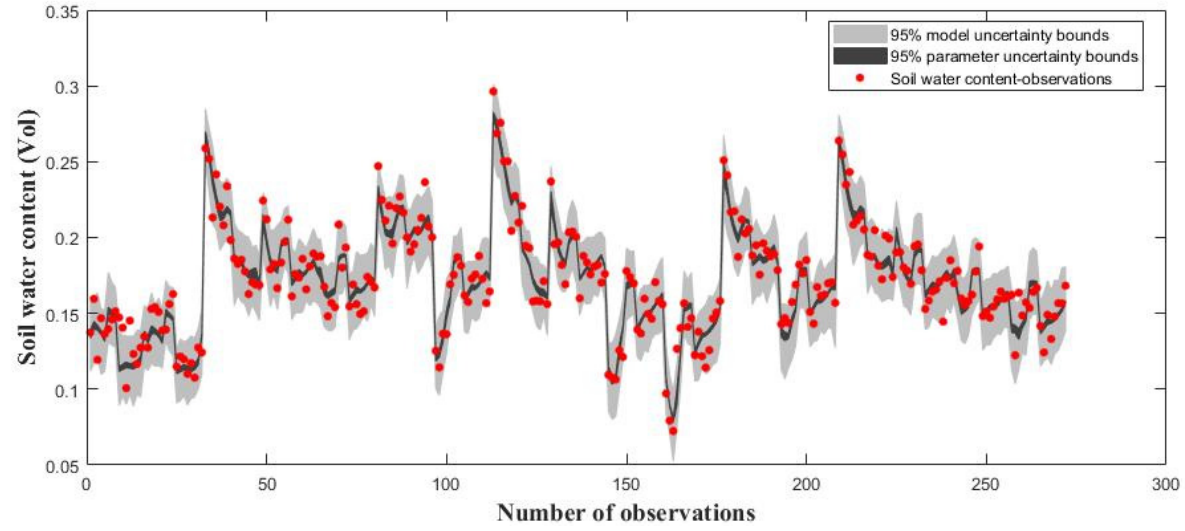
3. Results

We demonstrate the performance of the ILUES compared with the ESMDA through both synthetic and practical cases, which are mentioned in Sections 3.1.1 and 3.1.2, respectively. Then, the applicability of the DREAMkzs algorithm is evaluated in Section 3.2. The parameters that need to be optimized include 32 soil hydraulic parameters for 8 layers of the soil profile, 5 nitrogen transformation parameters and 8 crop parameters, which are contained in Tables 1 and 2. Uniform prior parameter distributions are assumed based on the ranges listed in Tables 1 and 2.

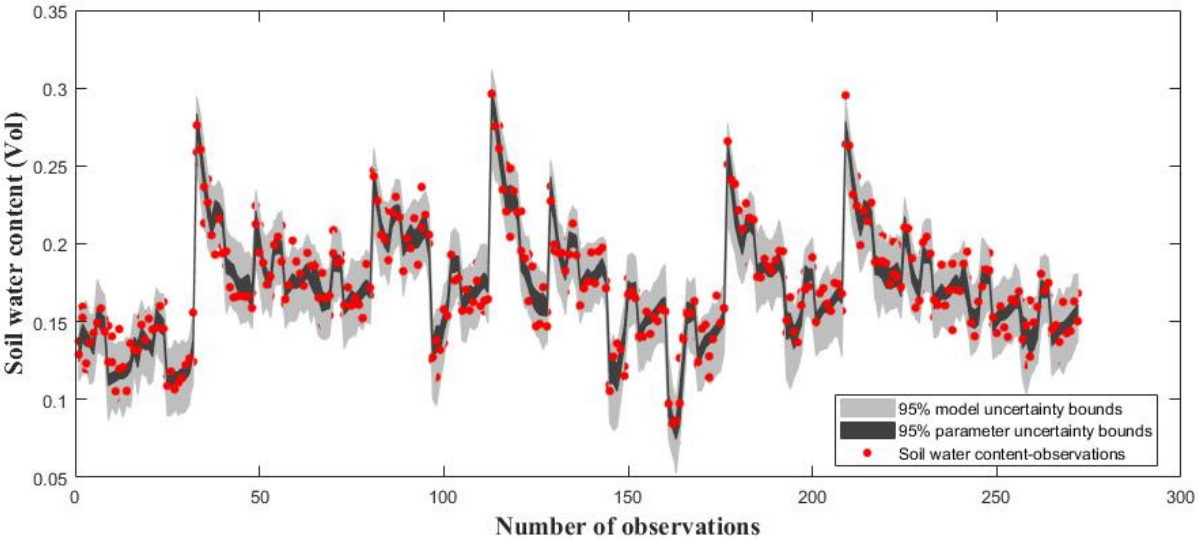
3.1. Comparison of ILUES and ESMDA

3.1.1. Synthetic Case

To demonstrate the performance of the ILUES algorithm, a synthetic case is discussed in this section. The reference soil water content values indicated with red points in Figure 1a,b are derived through the following steps: First, randomly sample the reference parameter values from the corresponding prior uniform distributions consistent with the ranges listed in Tables 1 and 2; second, generate the reference output values by collecting soil water content at 16 different points of the soil profile on the 17 days mentioned in Section 2.2.2, which are derived by running the forward WHCNS model; and third, disturb the simulated soil water content values using the assumed measurement error $\varepsilon \sim N\left(0,\ 0.005^2\right)$. The ILUES and MDAES algorithms are executed with the number of iterations as 3 and an ensemble size of 500. Figure 1a shows the fitting results of the simulated soil water content and the corresponding reference values derived by the ILUES algorithm with the RMSE_Maximum a posteriori (RMSE_MAP) is equal to 0.0104, while Figure 1b shows similar results derived by the MDAES algorithm with RMSE_MAP = 0.0093. Then, we come to a preliminary conclusion that the ILUES algorithm is able to obtain a satisfactory fitting effect equivalent to ESMDA.



(**a**)



(**b**)

Figure 1. Plots of simulated and reference soil water content. 95% simulation uncertainty intervals due to parameter (dark region) and total uncertainty (light gray). The reference soil water content values are indicated with red points. (**a**) iterative local updating ensemble smoother (ILUES); (**b**) ensemble smoother multiple data assimilation (ESMDA).

Figure 1 shows that acceptable simulation accuracy can be obtained by both the ILUES and ESMDA algorithms after three iterations. To further demonstrate the proposed algorithm's parameter estimation capabilities, the trace plots of model parameters calibrated by the two methods are compared in Figures 2 and 3. In Figure 2, the ILUES algorithm's sampling efficiency for all soil hydraulic parameters is higher than that of the ESMDA algorithm because its sampling trace derived converged to the corresponding reference values faster. Figure 3 shows a similar performance, except for T_{sum}. This result indicates that the ILUES method's ability to explore the parameter space is superior to the ESMDA method, as noted in the literature [40,41].

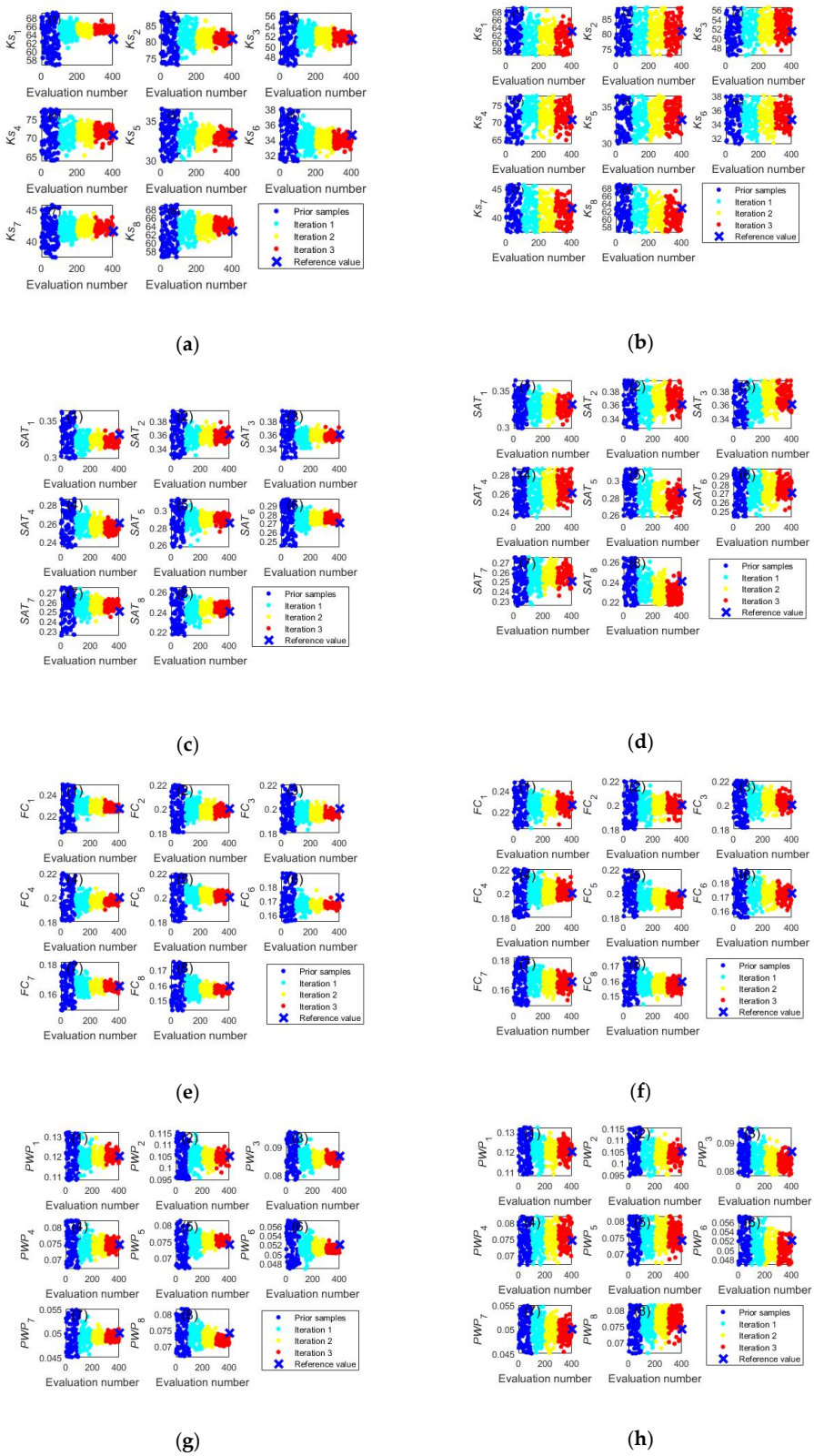


Figure 2. Trace plots of soil hydraulic parameters obtained by the ILUES (**left column**) and ESMDA (**right column**) algorithms. The blue cross symbol at the right side of each figure signifies the referred value of each parameter. (**a**) Saturated hydraulic conductivity derived by ILUES; (**b**) Saturated hydraulic conductivity derived by ESMDA; (**c**) Saturated soil water content derived by ILUES; (**d**) Saturated soil water content derived by ESMDA; (**e**) Field capacity derived by ILUES; (**f**) Field capacity derived by ESMDA; (**g**) Wilting point derived by ILUES; (**h**) Wilting point derived by ESMDA.

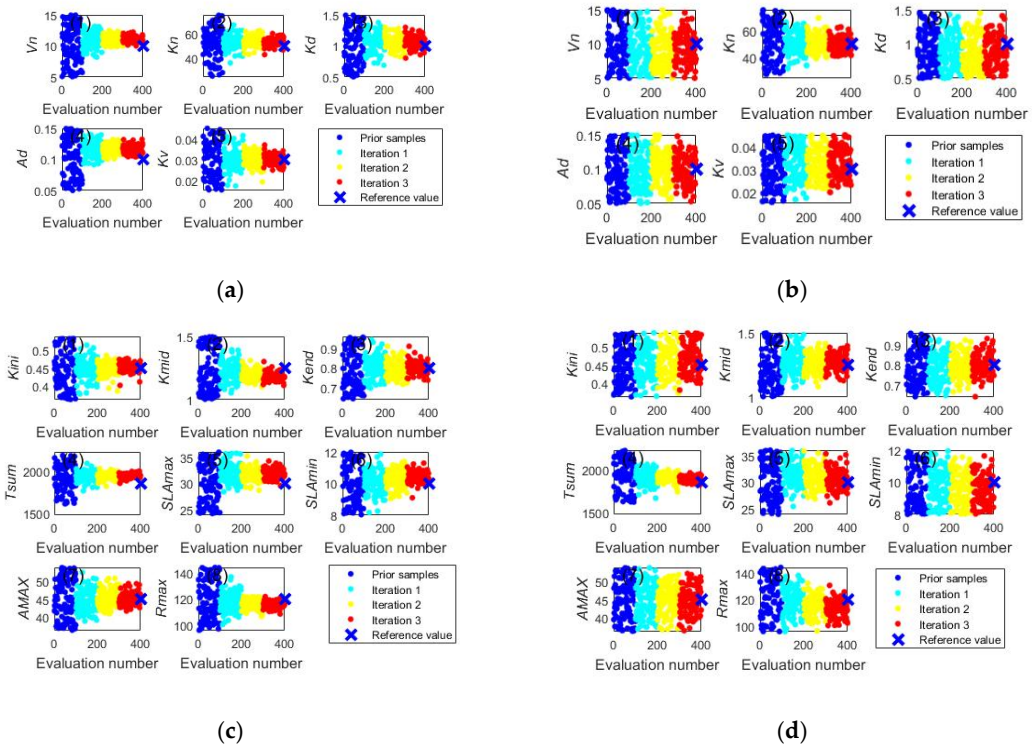
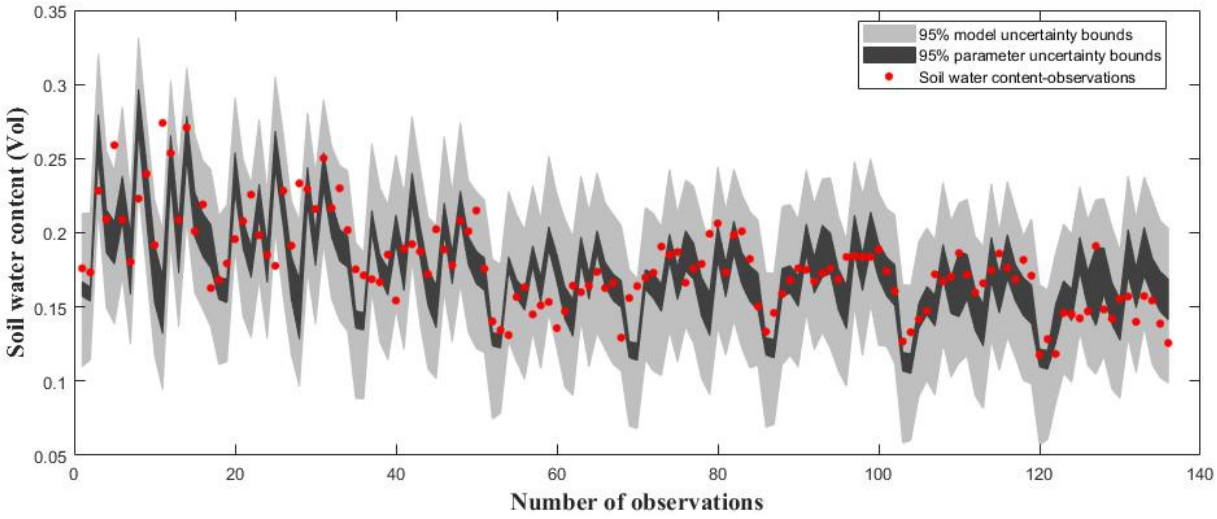


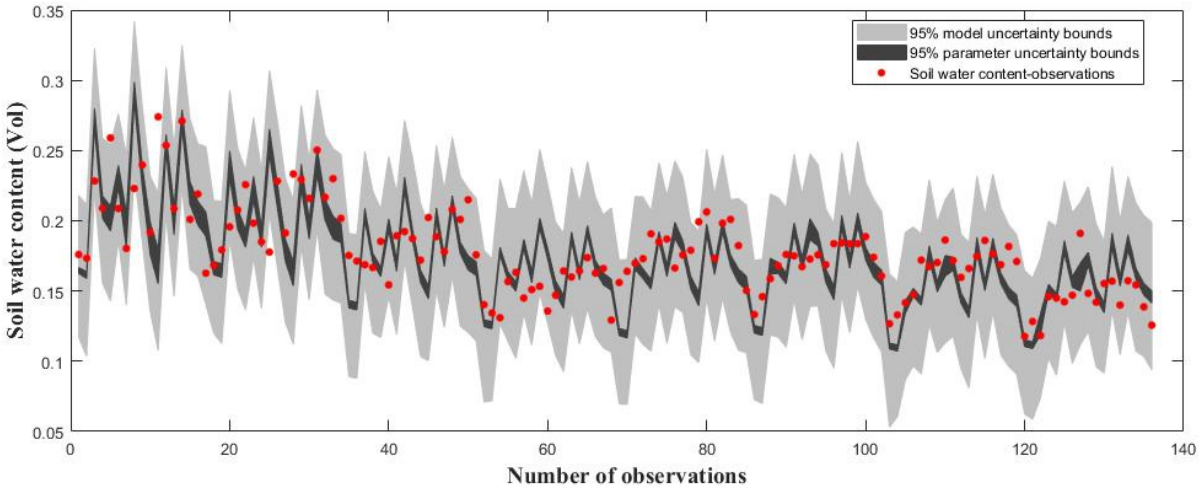
Figure 3. Trace plots of soil nitrogen transformation parameters and crop parameters obtained by the ILUES (**left column**) and ESMDA (**right column**) algorithms. The blue cross symbol at the right side of each figure signifies the referred value of each parameter. (**a**) Soil nitrogen transformation parameters derived by ILUES; (**b**) Soil nitrogen transformation parameters derived by ESMDA; (**c**) Crop parameters derived by ILUES; (**d**) Crop parameters derived by ESMDA.

3.1.2. Practical Case

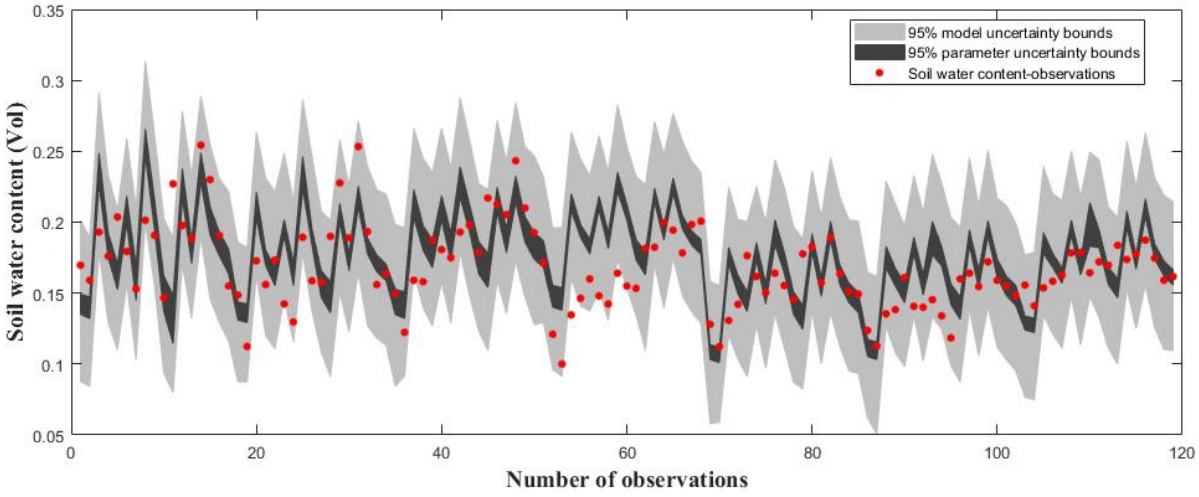
To further demonstrate the ILUES algorithm's performance compared with the ESMDA method, the soil hydraulic, crop and nitrogen parameters of the WHCNS model are simultaneously identified through the practical W_2N_1 case mentioned in Table 3. The red points shown in Figure 4a,b indicate volumetric soil water content measured from depths of 20 cm, 40 cm, 60 cm, 80 cm, 100 cm, 120 cm, 140 cm and 160 cm; furthermore, those displayed in Figure 4c,d indicate volumetric soil water content measured from depths of 30 cm, 50 cm, 70 cm, 90 cm, 110 cm, 130 cm and 170 cm. Volumetric soil water content at each selected depth of soil profile are measured at 17 days, which is mentioned in Section 2.2.2. That is to say, there are 136 observations at 8 layers for 17 days that need to be calibrated, and 119 measurement values at 7 layers for 17 days are used to validate the calibrated model's capacity for accurate prediction. Figure 4a,b are the calibration results of the ESMDA and ILUES algorithms, and the RMSE_MAP values for both are 0.0255, meaning both methods have a similar data assimilation capability. Figure 4c,d show validation results for the two different inverse methods with RMSE_MAPs, which are 0.0265 and 0.0269, respectively, indicating that the ILUES algorithm can derive reasonable results without overfitting. Therefore, both methods can reliably optimize the forward model system for accurate predictions.



(**a**)

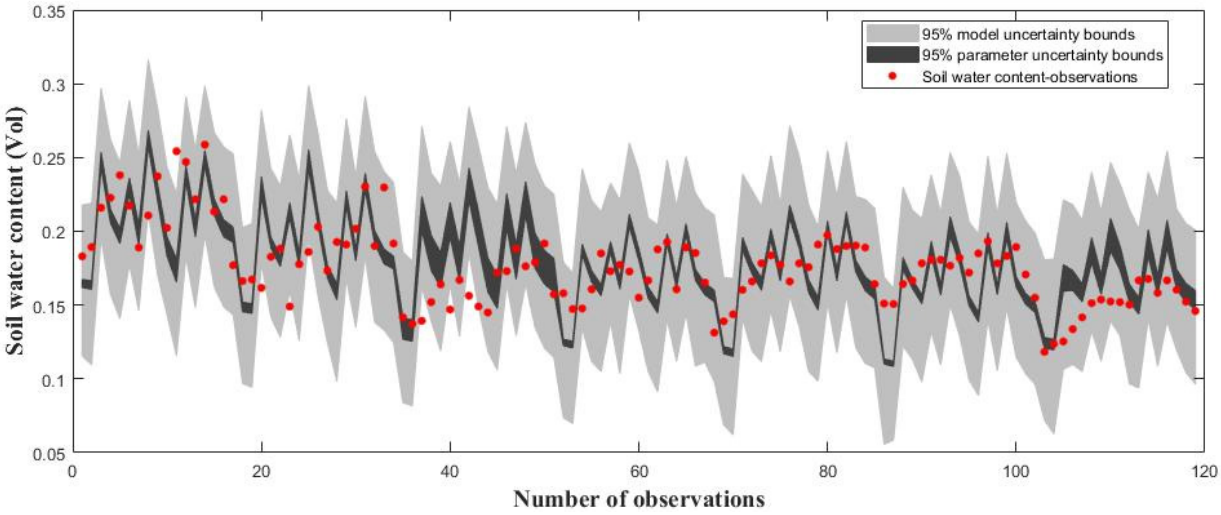


(**b**)



(**c**)

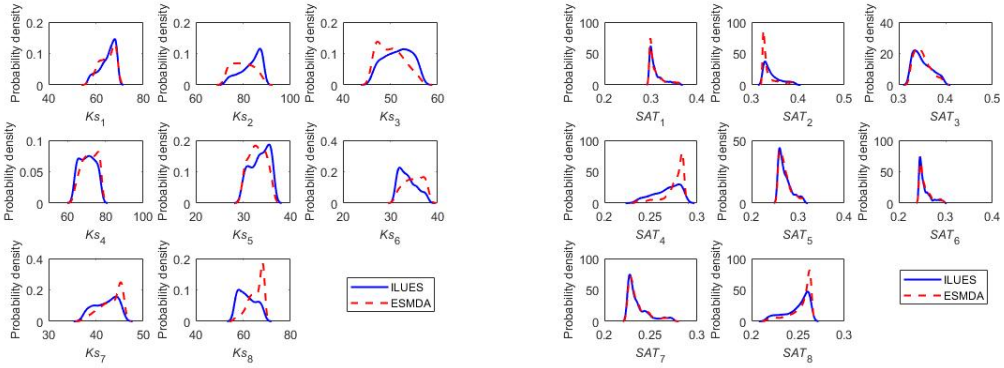
Figure 4. *Cont.*



(**d**)

Figure 4. Plots of simulated and reference soil water content. 95% simulation uncertainty intervals due to parameter (dark region) and total uncertainty (light gray). The measured soil water content values are indicated with red points. (**a**) ILUES-calibration; (**b**) ESMDA-calibration; (**c**) ILUES-validation; (**d**) ESMDA-validation.

Figure 5 presents the posterior distributions of model parameters obtained by the two inverse algorithms. Overall, reliable and robust estimations of the soil and crop parameters can be derived by the two ES-based methods. For Ks_1, Ks_4, SAT_1, SAT_3, SAT_5, SAT_6, SAT_7, FC_5, FC_8, PWP_8, T_{sum} and SLA_{max}, the posterior distributions derived by the ILUES method almost coincided with the corresponding results obtained by ESMDA. For Ks_5, Ks_7, Ks_8, FC_6, FC_7, PWP_3, PWP_4, PWP_6, K_{ini}, AMAX and five nitrogen transformation parameters, ILUES extracted the multimodal posterior distributions of parameters more obviously than ESMDA. Therefore, we conclude that, for a practical case with unknown posterior parameter distributions, ILUES is superior in exploring the multimodal distributions compared with ESMDA, which is consistent with the literature [38].



(**a**) (**b**)

Figure 5. *Cont.*

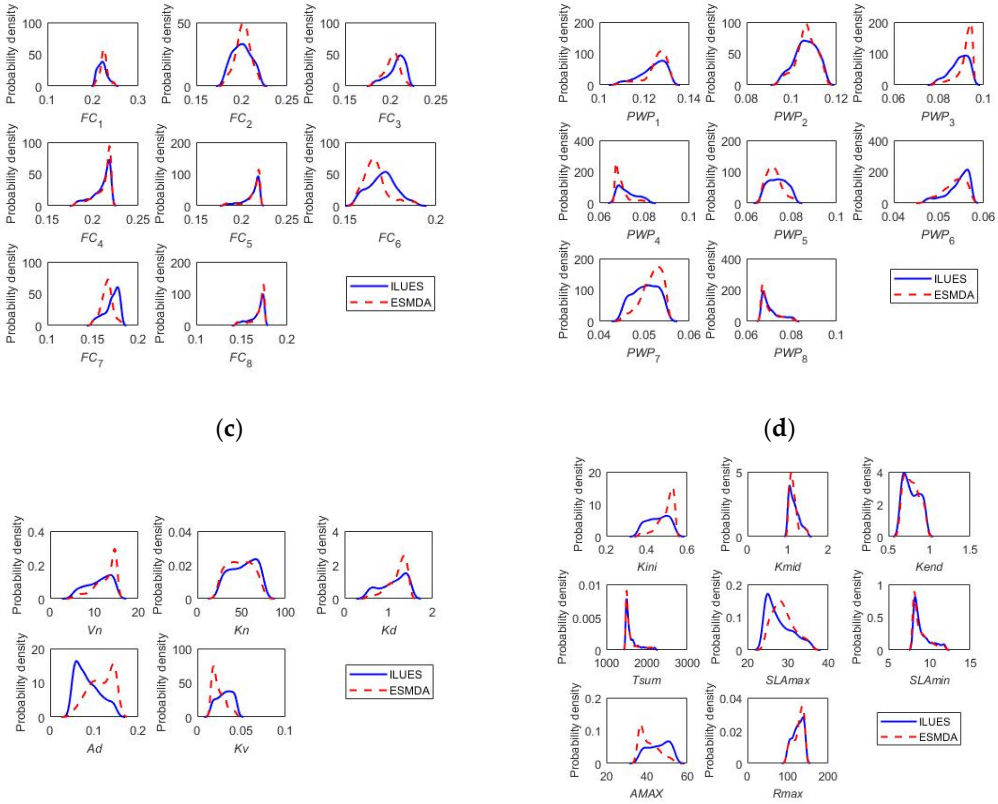
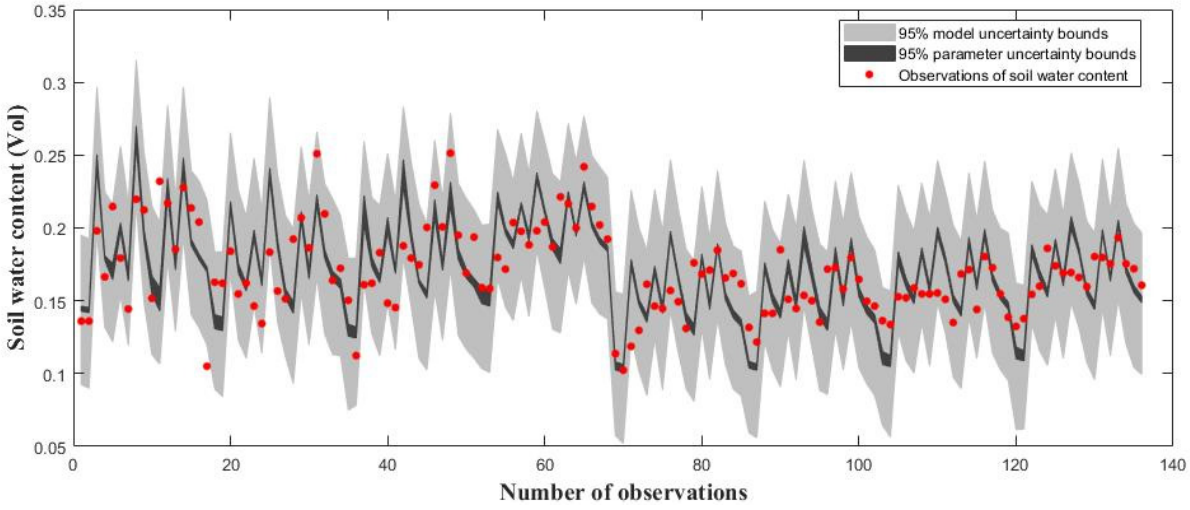


Figure 5. Posterior marginal distributions of model parameters estimated by ILUES (blue lines) and ESMDA (red dashed lines), respectively. (**a**) Ks; (**b**) Saturated soil water content; (**c**) Field capacity; (**d**) Wilting point; (**e**) Nitrogen transformation parameters; (**f**) Crop parameters, for W_2N_1 listed in Table 3.

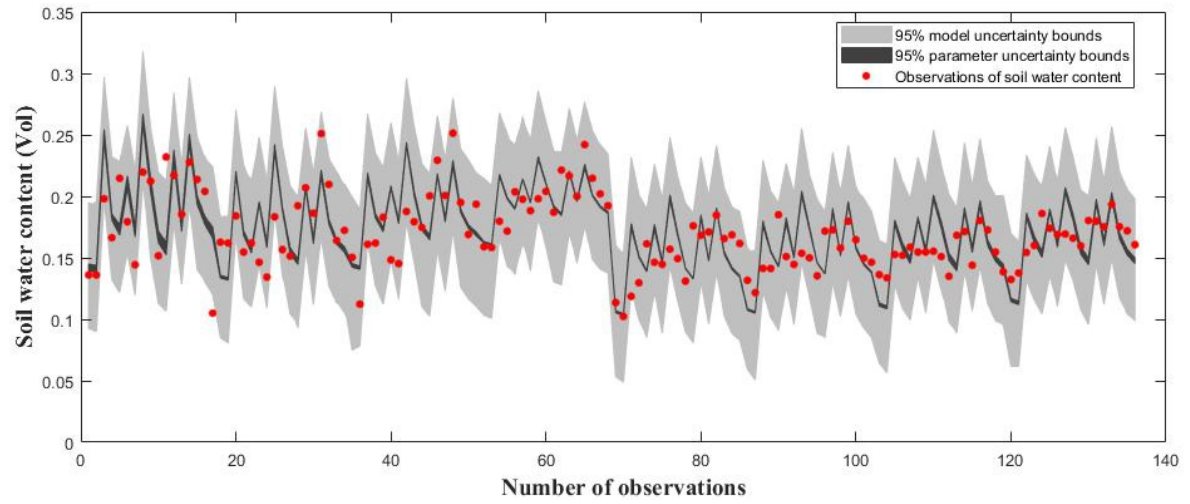
3.2. Comparison of DREAMkzs and DREAMzs

3.2.1. Model Performance Evaluation

Figure 6 shows evaluations of the DREAMzs and DREAMkzs algorithms' performances by applying the WHCNS model's parameter calibration based on the practical W_1N_1 case mentioned in Table 3. We take $P_p = 0.7$, $P_s = 0.1$, $P_k = 0.2$ and $T_k = 0.2T$, where P_p, P_s and P_k represent the selection probabilities of the parallel direction, snooker and Kalman-inspired proposal distributions, respectively, and T denotes the total number of model evaluations. Both MCMC algorithms were run with the number of chains N set to 3 and the number of generations in each chain T set to 1000. The red points shown in Figure 6a,b indicate the volumetric soil water content measured from depths of 20 cm, 40 cm, 60 cm, 80 cm, 100 cm, 120 cm, 140 cm and 160 cm. Volumetric soil water content at each selected soil profile depth is measured at 17 days, which is mentioned in Section 2.2.2. In total, 136 observations at 8 layers for 17 days need to be calibrated. Figure 6a,b show that the MCMC methods, including both DREAMkzs and DREAMzs, performed well in improving the level of fit between the observed states and simulated results, with RMSE_MAP values of 0.0253 and 0.0255, respectively. This preliminarily proved that the recommended DREAMkzs algorithm is feasible for accurately estimating practical observations.



(**a**)



(**b**)

Figure 6. Plots of simulated and reference soil water content. 95% simulation uncertainty intervals due to parameter (dark region) and total uncertainty (light gray). The measured soil water content values are indicated with red points. (**a**) DiffeRential Evolution Adaptive Metropolis with Kalman-inspired proposal distribution (DREAMkzs); (**b**) DiffeRential Evolution Adaptive Metropolis (DREAMzs).

3.2.2. Comparison of Simulation Statistics

Figure 7a,b display the evolution of the RMSE between the simulated and observed soil water content derived by DREAMkzs and DREAMzs, respectively. The horizontal red dashed line depicts the reference value of 0.0255. DREAMkzs significantly accelerates the optimization process compared with DREAMzs. Figure 7c,d show the distributions of ME and IA obtained by Equations (6) and (8), respectively. The results indicate that the simulated soil water content derived by both algorithms agreed well with the practical observations. Moreover, DREAMkzs improves the sampling efficiency considerably more than DREAMzs. However, the number of generations is small because the original parameter ranges are determined based on the LM algorithm's optimized results. More generations may be required if there are no preliminary optimization results.

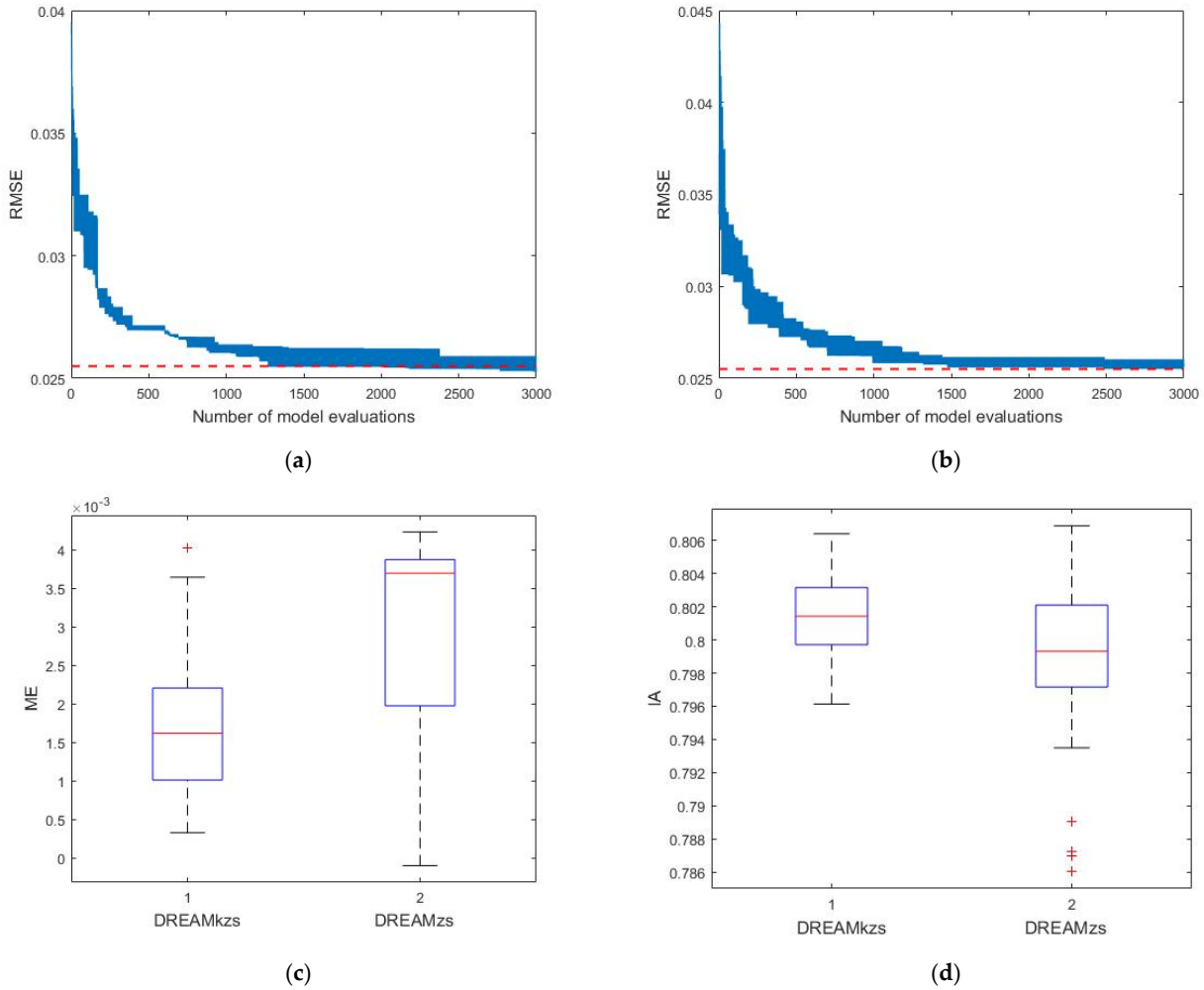


Figure 7. Statistics analysis of calibration results. (**a**) Evolution of RMSE derived by DREAMkzs. (**b**) Evolution of RMSE derived by DREAMzs. (**c**) Boxplots of ME. (**d**) Boxplots of IA.

4. Discussion

We found that the proposed parameter optimization methods can explore parameter sampling spaces and considerably improve sampling efficiency, increasing the WHCNS model's prediction accuracy. Reliable simulation results can be derived by the calibrated models, which popularize the WHCNS model. In addition, models with high prediction accuracy can be applied to analyze the effect of climate, types of plants and field management on soil water movement, crop growth and soil nitrogen transport, contributing to the development of field management practices that not only save resources but also decrease the environmental pollution induced by excessive water and fertilizer application.

Obvious differences exist between the two Kalman-formula-based parameter optimization methods mentioned above, even though both of them can identify the model with acceptable accuracy. First, the Kalman formula plays different roles in the two proposed parameter optimization algorithms. The Kalman formula is utilized to update the parameter ensemble, while the Kalman-inspired proposal distribution is just applied to sample a parameter candidate in the DREAMkzs algorithm. Second, the influence degree of the Kalman formula on inversion results is different. In ILUES, the Kalman gain exists throughout the whole parameter inversion process [40]. However, the Kalman-inspired proposal distribution is only used in the first several steps to avoid introducing asymmetry in the sampled candidate states [46]. In addition, parameter calibration results that enable a good fitness to the observed data can make a satisfying prediction under the same treatment.

However, it does not guarantee comparable estimations accuracy for other types of field management practices.

ILUES and DREAMkzs are newly proposed parameter inversion methods that have not been previously applied in the data assimilation of soil–crop system models. However, soil water content was the single-state variable used for the models' parameter calibrations, and improvements can be made by considering other state variables, such as soil nitrate concentration, crop N uptake and yield [30]. Meanwhile, we consider the soil hydraulic, soil nitrogen transformation and crop parameters simultaneously, which induced a high dimension. Global sensitivity analysis can be adopted to reduce parameter dimensions, as in [47]. In addition, the effects of applying the two proposed methods to high nonlinear models need to be further verified. Meanwhile, an increase in the parameter dimension can reduce the sampling efficiency. Therefore, more attention should be paid to further improving the parameter optimization performance of soil–crop system models.

5. Conclusions

Both ILUES and DREAMkzs are parameter optimization methods based on the Kalman formula. They can be successfully applied to a data assimilation of the WHCNS model, even though the Kalman formula produces a different effect for them. Our main conclusions are as follows: (1) The models calibrated using both ESMDA and ILUES algorithms under the condition of W_2N_1 are promisingly robust. The parameter distribution approaches the reference value faster under the artificial situation. Moreover, the performance of ILUES in modeling multimodal parameter distributions under a practical situation is superior to that of ESMDA under the same number of model evaluations condition. (2) DREAMkzs was successful in considerably accelerating the burn-in procedure compared with the DREAMzs with a comparative soil water content fitting accuracy under the same water and nitrate management practices.

Accurate model prediction is conducive to the formulation of field management measures for energy conservation and high yield. Furthermore, reducing input cost and increasing yield can raise farmers' incomes. The applicability of the latest parameter inversion methods in other fields to the WHCHS model can also be evaluated. At the same time, the two data assimilation methods mentioned in this paper can also be applied to other soil–crop system models. In addition, deep learning is currently popular and not limited by Gaussian hypothesis; therefore, it can be considered for introduction into the parameter inversion of a soil–crop system model, whether in whole or part.

Author Contributions: Conceptualization, W.W.; methodology, W.W.; software, Q.G.; validation, W.W. and Q.G.; formal analysis, Q.G.; writing—original draft preparation, Q.G.; writing—review and editing, W.W. and Q.G.; project administration, W.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Kelin Hu in China Agricultural University for providing original experimental data.

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

References

1. Lu, J.S.; Hu, T.T.; Zhang, B.C.; Wang, L.; Yang, S.H.; Fan, J.L.; Yan, S.C.; Zhang, F.C. Nitrogen fertilizer management effects on soil nitrate leaching, grain yield and economic benefit of summer maize in Northwest China. *Agric. Water Manag.* **2021**, *247*, 106739. [CrossRef]
2. Chen, S.M.; Wang, F.H.; Zhang, Y.M.; Qin, S.P.; Wei, S.C.; Wang, S.Q.; Hu, C.S.; Liu, B.B. Organic carbon availability limiting microbial denitrification in the deep vadose zone. *Environ. Microbiol.* **2018**, *20*, 980–992. [CrossRef]

3. Xiao, G.M.; Zhao, Z.C.; Liang, L.; Meng, F.Q.; Wu, W.L.; Guo, Y.B. Improving nitrogen and water use efficiency in a wheat-maize rotation system in the North China Plain using optimized farming practices. *Agric. Water Manag.* **2019**, *212*, 172–180. [CrossRef]
4. Karandish, F.; Šimůnek, J. A comparison of the HYDRUS (2D/3D) and SALTMED models to investigate the influence of various water-saving irrigation strategies on the maize water footprint. *Agric. Water Manag.* **2019**, *213*, 809–820. [CrossRef]
5. Wang, X.P.; Huang, G.H. Evaluation on the irrigation and fertilization management practices under the application of treated sewage water in Beijing, China. *Agric. Water Manag.* **2008**, *95*, 1011–1027. [CrossRef]
6. Hu, K.L.; Li, Y.; Chen, W.P.; Chen, D.L.; Wei, Y.P.; Edis, R.; Li, B.G.; Huang, Y.F.; Zhang, Y.P. Modeling Nitrate Leaching and Optimizing Water and Nitrogen Management under Irrigated Maize in Desert Oases in Northwestern China. *J. Environ. Qual.* **2010**, *39*, 667–677. [CrossRef]
7. Chen, Y.; Marek, G.W.; Marek, T.H.; Porter, D.O.; Brauer, D.K.; Srinivasan, R. Simulating the effects of agricultural production practices on water conservation and crop yields using an improved SWAT model in the Texas High Plains, USA. *Agric. Water Manag.* **2021**, *244*, 106574. [CrossRef]
8. Jin, L.; Hu, K.L.; Johannes Deelstra, J.; Li, B.G.; Wei, D.; Wang, H.Y. Evaluation of nitrogen fate, water and nitrogen use efficiencies of winter wheat in North China Plain based on model approach. *Acta Agric. Scand. Sect. B Soil Plant Sci.* **2014**, *63* (Suppl. 2), 127–138. [CrossRef]
9. Ale, S.; Gowda, P.H.; Mulla, D.J.; Moriasi, D.N.; Youssef, M.A. Comparison of the performances of DRAINMOD-NII and ADAPT models in simulating nitrate losses from subsurface drainage systems. *Agric. Water Manag.* **2013**, *129*, 21–30. [CrossRef]
10. Du, X.; Feng, H.; Helmers, M.J.; Qi, Z.M. Comparing Simulated Nitrate-Nitrogen Concentration In Subsurface Drainage Using Drainmod-N II and RZWQM2. *Irrig. Drain.* **2017**, *66*, 238–251. [CrossRef]
11. Liang, H.; Qi, Z.M.; Hu, K.L.; Li, B.B.; Prasher, S.O. Modelling subsurface drainage and nitrogen losses from artificially drained cropland using coupled DRAINMOD and WHCNS models. *Agric. Water Manag.* **2018**, *195*, 201–210. [CrossRef]
12. Singh, S.; Bhattarai, R.; Negm, L.M.; Mohamed, A.; Youssef, M.A.; Cameron, M.; Pittelkow, C.M. Evaluation of nitrogen loss reduction strategies using DRAINMOD-DSSAT in east-central Illinois. *Agric. Water Manag.* **2020**, *240*, 106322. [CrossRef]
13. Michalczyk, A.; Kersebaum, K.C.; Roelcke, M.; Hartmann, T.; Yue, S.C.; Chen, X.P.; Zhang, F.S. Model-based optimisation of nitrogen and water management for wheat–maize systems in the North China Plain. *Nutr. Cycl. Agroecosyst.* **2014**, *98*, 203–222. [CrossRef]
14. Wang, J.; Huang, G.H.; Zhan, H.B.; Mohanty, B.P.; Zheng, J.H.; Huang, Q.Z.; Xu, X. Evaluation of soil water dynamics and crop yield under furrow irrigation with a two-dimensional flow and crop growth coupled model. *Agric. Water Manag.* **2014**, *141*, 10–22. [CrossRef]
15. Qin, X.B.; Wang, H.; He, Y.; Li, Y.E.; Li, Z.G.; Gao, Q.Z.; Wan, Y.F.; Qian, B.D.; Brian McConkey, B.; DePauw, R.; et al. Simulated adaptation strategies for spring wheat to climate change in a northern high latitude environment by DAYCENT model. *Eur. J. Agron.* **2018**, *95*, 45–56. [CrossRef]
16. Li, Z.T.; Wen, X.M.; Hu, C.S.; Li, X.X.; Li, S.S.; Zhang, X.S.; Hu, B.Q. Regional simulation of nitrate leaching potential from winter wheat-summer maize rotation croplands on the North China Plain using the NLEAP-GIS model. *Agric. Ecosyst. Environ.* **2020**, *294*, 106861. [CrossRef]
17. Jiang, R.; Yang, J.Y.; Drury, C.F.; He, W.T.; Smith, W.; Grant, B.; He, P.; Zhou, W. Assessing the impacts of diversified crop rotation systems on yields and nitrous oxide emissions in Canada using the DNDC model. *Sci. Total Environ.* **2020**, *759*, 143433. [CrossRef]
18. Liang, H.; Lv, H.F.; Batchelor, W.D.; Lian, X.J.; Wang, Z.X.; Lin, S.; Hu, K.L. Simulating nitrate and DON leaching to optimize water and N management practices for greenhouse vegetable production systems. *Agric. Water Manag.* **2020**, *241*, 106377. [CrossRef]
19. Luo, Y.Q.; Edward, A.G.; Schuur, E.A.G. Model parameterization to represent processes at unresolved scales and changing properties of evolving systems. *Glob. Chang. Biol.* **2020**, *26*, 1109–1117. [CrossRef]
20. Peruta, R.D.; Keller, A.; Schulin, R. Sensitivity analysis, calibration and validation of EPIC for modelling soil phosphorus dynamics in Swiss agro-ecosystems. *Environ. Modell. Softw.* **2014**, *62*, 97–111. [CrossRef]
21. Xi, M.L.; Qi, Z.M.; Zou, Y.; Raghavan, G.S.V.; Sun, J. Calibrating RZWQM2 model using quantum-behaved particle swarm optimization algorithm. *Comput. Electron. Agric.* **2015**, *113*, 72–80. [CrossRef]
22. Confalonieri, R.; Bregaglio, S.; Acutis, M. Quantifying uncertainty in crop model predictions due to the uncertainty in the observations used for calibration. *Ecol. Model.* **2016**, *328*, 72–77. [CrossRef]
23. Falconnier, G.N.; Journet, E.P.; Bedoussac, L.; Vermue, A.; Chlébowski, F.; Beaudoin, N.; Justes, E. Calibration and evaluation of the STICS soil-crop model for faba bean to explain variability in yield and N2 fixation. *Eur. J. Agron.* **2019**, *104*, 63–77. [CrossRef]
24. Chen, Y.; Tao, F.L. Improving the practicability of remote sensing data-assimilation-based crop yield estimations over a large area using a spatial assimilation algorithm and ensemble assimilation strategies. *Agric. For. Meteorol.* **2020**, *291*, 108082. [CrossRef]
25. Xi, M.L.; Lu, D.; Cui, D.W.; Qi, Z.M.; Zhang, G.N. Calibration of an agricultural-hydrological model (RZWQM2) using surrogate global optimization. *J. Hydrol.* **2017**, *544*, 456–466. [CrossRef]
26. Gurung, R.B.; Ogle, S.M.; Breidt, F.J.; Williams, S.A.; Parton, W.J. Bayesian calibration of the DayCent ecosystem model to simulate soil organic carbon dynamics and reduce model uncertainty. *Geoderma* **2020**, *376*, 114529. [CrossRef]
27. Sun, M.; Zhang, X.L.; Huo, Z.L.; Feng, S.Y.; Huang, G.H.; Mao, X.M. Uncertainty and sensitivity assessments of an agricultural–hydrological model (RZWQM2) using the GLUE method. *J. Hydrol.* **2016**, *534*, 19–30. [CrossRef]
28. Sexton, J.; Everingham, Y.; Inman-Bamber, G. A theoretical and real world evaluation of two Bayesian techniques for the calibration of variety parameters in a sugarcane crop model. *Environ. Modell. Softw.* **2016**, *83*, 126–142. [CrossRef]

29. Gao, Y.J.; Wallach, D.; Liu, B.; Dingkuhn, M.; Boote, K.J.; Singh, U.; Asseng, S.; Kahveci, T.; He, J.Q.; Zhang, R.Y.; et al. Comparison of three calibration methods for modeling rice phenology. *Agric. For. Meteorol.* **2020**, *280*, 107785. [CrossRef]
30. Dumont, B.; Leemans, V.; Mansouri, M.; Bodson, B.; Destain, J.P.; Destain, M.F. Parameter identification of the STICS crop model, using an accelerated formal MCMC approach. *Environ. Modell. Softw.* **2014**, *52*, 121–135. [CrossRef]
31. Zhang, Y.; Arabi, M.; Paustian, K. Analysis of parameter uncertainty in model simulations of irrigated and rainfed agroecosystems. *Environ. Modell. Softw.* **2020**, *126*, 104642. [CrossRef]
32. Ines, A.V.M.; Das, N.N.; Hansen, J.W.; Njoku, E.G. Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. *Remote Sens. Environ.* **2013**, *138*, 149–164. [CrossRef]
33. Fossum, K.; Mannseth, T. Parameter sampling capabilities of sequential and simultaneous data assimilation: I. Analytical comparison. *Inverse Probl.* **2014**, *30*, 114002. [CrossRef]
34. Ju, L.; Zhang, J.J.; Meng, L.; Wu, L.S.; Zeng, L.Z. An adaptive Gaussian process-based iterative ensemble smoother for data assimilation. *Adv. Water Resour.* **2018**, *115*, 125–135. [CrossRef]
35. Li, L.P.; Stetler, L.; Cao, Z.D.; Arden Davis, A. An iterative normal-score ensemble smoother for dealing with non-Gaussianity in data assimilation. *J. Hydrol.* **2018**, *567*, 759–766. [CrossRef]
36. Kang, X.Y.; Shi, X.Q.; Revil, A.; Cao, Z.D.; Li, L.P.; Lan, T.; Wu, J.C. Coupled hydrogeophysical inversion to identify non-Gaussian hydraulic conductivity field by jointly assimilating geochemical and time-lapse geophysical data. *J. Hydrol.* **2019**, *578*, 124092. [CrossRef]
37. Mo, S.X.; Zabaras, N.; Shi, X.Q.; Wu, J.C. Deep Autoregressive Neural Networks for High-Dimensional Inverse Problems in Groundwater Contaminant Source Identification. *Water Resour. Res.* **2018**, *55*, 3856–3881. [CrossRef]
38. Huang, Y.; Zhang, J.J.; Zheng, F.F.; Jia, Y.Y.; Kapelan, Z.; Savic, D. Exploring the performance of ensemble smoothers to calibrate urban drainage models. *Water Resour. Res.* **2022**, *58*, e2022WR032440. [CrossRef]
39. Zhang, J.J.; Zheng, Q.; Wu, L.S.; Zeng, L.Z. Using Deep Learning to Improve Ensemble Smoother: Applications to Subsurface Characterization. *Water Resour. Res.* **2020**, *56*, e2020WR027399. [CrossRef]
40. Zhang, J.J.; Lin, G.; Li, W.X.; Wu, L.S.; Zeng, L.Z. An iterative local updating ensemble smoother for estimation and uncertainty assessment of hydrologic model parameters with multimodal distributions. *Water Resour. Res.* **2018**, *54*, 1716–1733. [CrossRef]
41. Emerick, A.A.; Reynolds, A.C. Ensemble smoother with multiple data assimilation. *Comput. Geosci.* **2013**, *55*, 3–15. [CrossRef]
42. Zhang, J.J.; Man, J.; Lin, G.; Wu, L.S.; Zeng, L.Z. Inverse modeling of hydrologic systems with adaptive multi-fidelity Markov chain Monte Carlo simulations. *Water Resour. Res.* **2018**, *54*, 4867–4886. [CrossRef]
43. Liang, H.; Hu, K.L.; Batchelor, W.D.; Qi, Z.M.; Li, B.G. An integrated soil-crop system model for water and nitrogen management in North China. *Sci. Rep.* **2016**, *6*, 25755. [CrossRef]
44. Wei, Y.P.; Chen, D.L.; Hu, K.L.; Willett, I.R.; Langford, J. Policy incentives for reducing nitrate leaching from intensive agriculture in desert oases of Alxa, Inner Mongolia, China. *Agric. Water Manag.* **2009**, *96*, 1114–1119. [CrossRef]
45. Liang, H.; Qi, Z.M.; Hu, K.L.; Prasher, S.O.; Zhang, Y.P. Can nitrate contaminated groundwater be remediated by optimizing flood irrigation rate with high nitrate water in a desert oasis using the WHCNS model? *J. Environ. Manag.* **2016**, *181*, 16–25. [CrossRef]
46. Zhang, J.J.; Vrugt, J.A.; Shi, X.Q.; Lin, G.; Wu, L.S.; Zeng, L.Z. Improving Simulation Efficiency of MCMC for Inverse Modeling of Hydrologic Systems with a Kalman-Inspired Proposal Distribution. *Water Resour. Res.* **2020**, *56*, e2019WR025474. [CrossRef]
47. Liang, H.; Xu, J.Z.; Chen, L.N.; Li, B.G.; Hu, K.L. Bayesian calibration and uncertainty analysis of an agroecosystem model under different N management practices. *Eur. J. Agron.* **2022**, *133*, 126429. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.