

## Article

# Application of Different Weighting Schemes and Stochastic Simulations to Parameterization Processes Considering Observation Error: Implications for Climate Change Impact Analysis of Integrated Watershed Models

Eunhee Lee <sup>1</sup>, Hyeonju Lee <sup>1</sup> , Dongkyu Park <sup>1</sup>, Hyoun-Tae Hwang <sup>2,3,\*</sup> and Changhui Park <sup>4</sup>

<sup>1</sup> Korea Institute of Geoscience and Mineral Resources, 124 Gwahak-ro, Yuseong-gu, Daejeon 34132, Republic of Korea; eunheelee@kigam.re.kr (E.L.)

<sup>2</sup> Aquanty Inc., 600 Weber Street North, Unit B, Waterloo, ON N2V 1K4, Canada

<sup>3</sup> Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, ON N2L 3G1, Canada

<sup>4</sup> GeoGreen 21 Inc., 55 Digital-ro, Guro-gu, Seoul 08376, Republic of Korea

\* Correspondence: hthwang@aquanty.com

**Abstract:** We investigated the potential impact of observation error on the calibration performance of an integrated watershed model. A three-dimensional integrated model was constructed using HydroGeoSphere and applied to the Sabgyo watershed in South Korea to assess the groundwater–surface water interaction process. During the model calibration, three different weighting schemes that consider observation error variances were applied to the parameter estimation tool (PEST). The applied weighting schemes were compared with the results from stochastic models, in which observation errors from surface discharges were considered a random variable. Based on the calibrated model, the interactions between groundwater and surface water were predicted under different climate change scenarios (RCP). Comparisons of calibration performance between the different models showed that the observation-error-based weighting schemes contributed to an improvement in the model parameterization. Analysis of the exchange flux between groundwater and surface water highlighted the significance of groundwater in delaying the hydrological response of integrated water systems. Predictions based on different RCP scenarios suggested the increasing role of groundwater in watershed dynamics. We concluded that the comparison of different weighting schemes for the determination of error covariance could contribute to an improved characterization of watershed processes and reduce the model uncertainty arising from observation errors.

**Keywords:** integrated model; surface water; groundwater; observation error; climate change



**Citation:** Lee, E.; Lee, H.; Park, D.; Hwang, H.-T.; Park, C. Application of Different Weighting Schemes and Stochastic Simulations to Parameterization Processes Considering Observation Error: Implications for Climate Change Impact Analysis of Integrated Watershed Models. *Water* **2023**, *15*, 1880. <https://doi.org/10.3390/w15101880>

Academic Editor: Renato Morbidelli

Received: 6 April 2023

Revised: 7 May 2023

Accepted: 11 May 2023

Published: 16 May 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

Climate change impacts on hydrological processes have received increasing attention in recent decades [1]. Climate change is expected to have a significant impact on the global water environment, affecting both the quantity and quality of water resources. The impacts of climate change on the global water environment are complex and varied, and they have the potential to affect many aspects of human and ecosystem health. In order to study the potential impact of future climate change, different scenarios were developed to provide a consistent framework for climate model simulations and to facilitate the comparison of results across different research groups. The RCP (Representative Concentration Pathway) scenario is a set of projections of future greenhouse gas (GHG) emissions and atmospheric concentrations, as well as associated climate changes, used in climate research [2,3]. Therefore, predicting the integrated water systems' response under different climate change scenario is necessary for taking action to better adapt to the changes and reduce the water-related risks. Numerous studies have shown that projected climate change is likely to alter regional water cycles and, subsequently, impact

the spatiotemporal behavior of hydrological systems [4–8]. While climate change can affect the surface water system directly with a relatively short time lag (~days), the hydrological response of the subsurface system to the changing climate is slower and more complex. The role of groundwater as a buffer to environmental changes has, thus, been widely acknowledged and investigated [9,10].

Surface water and groundwater bodies comprise integral parts of the hydrologic cycle, with a strong influence on their water budgets [11,12]. This indicates that a more realistic evaluation of climate change impacts on surface water resources should consider the changes in groundwater bodies, and vice versa. Several studies have focused on this interlinkage and, in particular, examined the role of groundwater in surface water systems [13–17]. Despite the increasing attention focused on the feedbacks between groundwater–surface water systems, only a limited number of studies have been conducted to evaluate the contribution of groundwater in altering hydrological responses with different climate change scenarios [9,18–20].

Integrated models provide a valuable tool for examining complex watershed processes and predicting future hydrological changes. Physically-based three-dimensional (3D) numerical models can be used to represent the complexity of watershed characteristics by fully encompassing various hydrological feedbacks between groundwater and surface water [21]. A successful representation of a complex hydrological process into an integrated model mainly depends on the proper characterizations of the model parameters. Inverse models, such as PEST (Parameter Estimation Tool) [22], have provided an efficient tool for identifying the set of model parameters by minimizing the weighted difference between measured and simulated observations.

Although the PEST model has provided an efficient method to obtain the parameters that can best represent a watershed system [23–25], uncertainties associated with the model prediction still exist. Most hydrological/hydrogeological models contain two types of error: observation and structural errors [22]. Observation errors are caused by the accompanying noise during the measurements of the system state, whereas structural errors mainly occur when the simplified model fails to capture the complexity of the real-world system. Although the second term is known as the dominant contributor to prediction errors [26,27], measurement errors still play a significant role in determining the uncertainty of the integrated model. For example, a considerable level of observation errors in the surface water discharge—an observation type controlling the water balance of the integrated model—stems from the stage–discharge rating curve approximation and the accompanying errors during measurements of stream velocity and stage.

To consider observation errors, parameter inversion models such as PEST use a weight matrix, an element of which is calculated as the inverse of the observation error (co)variance [28–30]. The approach has been proven to provide the smallest possible variances of the estimated parameters [31]. The key assumption underlying this approach is that the relation between the model inputs and outputs could be approximated as having linear behavior. Another approach used to represent the variability of the observation error is stochastic characterization [32,33]. This method generates multiple sets of observations using a probability distribution function of measurement error and calculates the corresponding model outputs for each random realization. The latter is conceptually simple and can better reflect complex (nonlinear) watershed processes. The practical application of the stochastic approach, however, has been limited to rather simple hydrological models due to its intensive computational requirement; thus, few studies have adopted the approach for complex 3D integrated models.

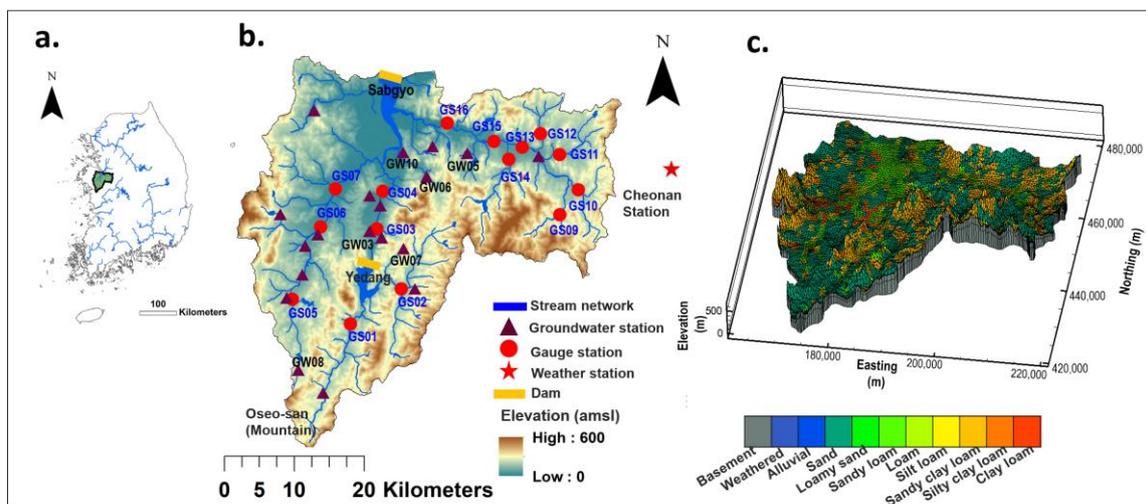
The objective of this study was to predict the future hydrological response of the watershed system and evaluate the interactions between groundwater and surface water under different climate change scenarios using a 3D integrated groundwater–surface water flow model. The constructed model was applied to the Sabgyo watershed and calibrated using the parameter inversion tool PEST. In this study, we focused on exploring the potential impact of observation errors on parameter inversion processes. Three different weighting

schemes that consider observation error variances were suggested and applied to PEST. The applied weighting schemes were compared with the results from stochastic models, in which observation errors from the principal observation points—a set of observation points that mainly controls inversion process—were considered as a random variable. Based on the calibrated model setup and error analysis, the interactions between groundwater and surface water were analyzed, and the future variations in the integrated flow system under the impact of climate change were investigated. The results of this study can provide new insights on addressing observation errors within complex watershed models and reduce the predictive uncertainties arising from observation errors.

## 2. Materials and Methods

### 2.1. Study Site

The Sabgyo watershed is located in the northwestern part of South Korea (Figure 1a). The Sabgyo-cheon (stream) originates from the Oseo-san (Mountain) in the southern part of the watershed and then flows approximately 65 km to the north until it meets the West Sea. The watershed drains an area of approximately 1650 km<sup>2</sup>. The surface topography relief ranges from zero in the northern part to 650 m in the southeastern part (amsl). The measured rainfall and temperature at Cheonan station (Figure 1b) from 2000 to 2019 were  $1203.1 \pm 126.3$  mm/year and  $12.1 \pm 9.8$  °C, respectively [34]. Seasonal variations in rainfall and temperature range from  $560.1 \pm 45.8$  mm/year (dry season, October to May) to  $2528.4 \pm 158.7$  mm/year (wet season, June to September) and from  $6.5 \pm 7.0$  (dry season) to  $23.1 \pm 2.3$  °C (wet season), respectively.



**Figure 1.** (a) Location of the study site; (b) monitoring stations and stream network within the watershed; and (c) 3D subsurface model and dominant hydrostratigraphy/soil types of the study site.

The surficial geology of the study site is mostly covered by Precambrian granite-biotite gneiss, Mesozoic granite, and Quaternary alluvial deposits [35]. The surface soil of the site is categorized into eight different types, including sand (50.9%), sandy clay loam (26.3%), sandy loam (12.8%), silt loam (4.8%), etc. [36]. The dominant land-use types within the watershed were forest (44.5%) and agricultural land (42.9%), followed by urban area (4.7%) and grassland (2.9%) as of 2010 [37].

The estimated total groundwater extraction rate of the study area from 2000 to 2019 was approximately  $255,000 \pm 21,000$  m<sup>3</sup>/day [38]. Two dams (Yedang and Sabgyo) within the watershed have been operated for the purpose of agricultural water supply (Figure 1b). In addition to the water supply, the Sabgyo Dam, located at the outlet of the watershed, controls water flow and downstream flooding.

## 2.2. Numerical Model

A 3D integrated HydroGeoSphere (HGS) model was developed to simulate integrated groundwater and surface water flows [39]. HGS calculates surface and subsurface water flow based on the two-dimensional (2D) diffusion-wave approximation of the Saint Venant equation [40] and the Richards' equation for variably saturated flow, respectively.

The surface domain was generated into 2D triangular meshes using AlgoMesh [41]. The grid sizes are uniformly distributed along the entire domain with a mean length of 420 m. Based on the geological logging database [38], the subsurface domain was vertically discretized into four divisions (surface soil, alluvial deposits, weathered rock zone, and basement rock), the depths of which were determined using inverse distance weighting interpolation [42]. These vertical geological units were then divided into 10 sublayers to increase the vertical resolution of the model. The total numbers of 3D nodes and elements are 89,530 and 156,420, respectively (Figure 1c).

In the surface domain, a rainfall boundary was assigned using the observed monthly means at the Cheonan weather station (Figure 1b). A critical depth boundary condition was given along the perimeter of the surface domain, while a no flow boundary condition was applied to those of the subsurface domain. The hydraulic structure of the Yedang dam was represented using a mesh cutoff and boundary condition linking scheme [43]. The water discharge rate from the dam was expressed by varying the flux boundary in which the discharge rate was determined depending on the dam stage and the reservoir capacity. At the subsurface domain, constant pumping boundaries were assigned. Since the location of each pumping well cannot be properly represented in the model domain due to the limited pumping well database and grid size, we assumed that groundwater pumping wells were uniformly distributed within the watershed, thus assigning constant pumping boundaries to all nodes (except the nodes beneath the streams) that lie between alluvial deposits and weathered rock layers (major aquifers). The watershed includes parts of six different counties that have different groundwater utilization rates. Accordingly, different groundwater pumping rates were given for each administrative unit by dividing the total groundwater usage rate by the total number of 2D nodes for each district.

Evapotranspiration (ET) processes were described using the relation suggested by Kristensen and Jensen [44] and Wigmosta et al. [45]. The potential evapotranspiration (PET) was calculated using the simplified FAO Penman–Monteith equation [46]. Meteorological data (minimum and maximum temperature, humidity, wind speed, and solar radiation) observed at the Cheonan station from 2000 to 2019 were used to obtain monthly varying PET. The monthly extraterrestrial radiation values were retrieved from [47]. The leaf area index (LAI) was obtained from the MODIS Data [48]. Different monthly averaged LAI curves were assigned (2000–2019) for each vegetation type by averaging cell-by-cell LAIs (500 m in resolution) for each zone.

Tables 1–3 display the model parameters used for simulation. Among these parameters, hydraulic conductivities, transpiration limiting saturation index (oxic limit, anoxic limit), and minimum and maximum evaporation-limiting saturation index were chosen as the model calibration parameters. Whereas the hydraulic conductivities of alluvial deposits and weathered rock were represented as lumped values (homogeneous conditions), the top soil was subdivided into eight zones based on the soil types; thus, different hydraulic conductivities were assigned for each zone. Other model parameters were referred from the literature. A total of 19 national groundwater monitoring wells and 15 gauge stations were used as observation targets [38,49,50]. For each observation station, available monthly data since 2000 to 2019 were used for model calibration.

**Table 1.** Evapotranspiration (ET) parameters used for simulation.

Parameter		Values	Sources/Notes
Evaporation depth		1(Urban)–3 m (Mixed trees)	Cubic decay with depth [51]
Root depth		0.1 (Urban)–3.5 (Mixed trees)	Cubic decay with depth [51,52]
LAI		From 0.26–2.9 (Vegetation) to 0.45–4.1 (Deciduous)	Monthly averages used for simulation [48]
Transpiration limiting saturation	Wilting point	0.19	[53]
	Field capacity	0.3	[53]
	Oxic limit	Calibration target (tox101) <sup>1</sup>	-
	Anoxic limit	Calibration target (tax101) <sup>1</sup>	-
Evaporation limiting saturation	Minimum	Calibration target (emin01) <sup>1</sup>	-
	Maximum data	Calibration target (emax01) <sup>1</sup>	-

Note: <sup>1</sup> Symbols indicate parameter names used for model calibration.

**Table 2.** Surface flow parameters used for simulation.

Parameter	Values	Sources/Notes
Manning’s Roughness Coefficients	0.0016 (Urban)–0.03 (Forest)	[54]
Rill Storage Height	$2.0 \times 10^{-5}$ (urban)– $5.0 \times 10^{-3}$ (Wetland)	[38,51]
Obstruction Storage Height	$1.0 \times 10^{-5}$ (Urban)– $5.0 \times 10^{-3}$ (Wetland)	[38,51]
Coupling Length	0.01 m	[55]

**Table 3.** Subsurface flow parameters used for simulation.

Parameter	Values	Sources/Notes	
Hydraulic conductivity (m/s)	Basement rock	$1.0 \times 10^{-10}$	
	Weathered rock	Calibration target (kx001) <sup>1</sup>	[56]
	Alluvial layer	Calibration target (kx002) <sup>1</sup>	
	Surface soil	Calibration target (kx003-kx010) <sup>2</sup>	
Anisotropy (Kv/Kh)	Basement rock	1	
	Weathered rock	1	[56]
	Alluvial layer	1	
	Surface soil	0.1	
Specific Storage	Ss	$1 \times 10^{-4}$ – $5.0 \times 10^{-4}$	[56]
Porosity	n	0.05 (basement rock)–0.35 (Soil)	[56]
Van Genuchten parameters	$\alpha$	2.25	
	$\beta$	1.89	[57]
	Residual saturation	0.18	

Notes: <sup>1</sup> Symbols indicate parameter names used for model calibration. <sup>2</sup> kx0003 to kx010 represent sand to clay loam in Figure 1.

### 2.3. Model Calibration Process Considering Observation Errors

In this study, we calibrated the constructed model using the monthly normal concept, which can represent the monthly mean hydrological behavior of the water system under varying meteorological and water-use conditions. The monthly normal inputs of the integrated model was developed by averaging the weather forcing data (precipitation, potential evapotranspiration, and temperature) of each month over the target period. These monthly normal models are considered to represent the long-term average seasonal cycle. A similar approach was adopted by [17] for the effective employment of large-scale integrated models in climate change impact studies. The applied target period for the monthly normal calibration was from 2000 to 2019. Since the goal of this work was to predict a long-term hydrological response (until 2100), we chose the calibrated model setup to properly reflect a long period of meteorological and hydrological characteristics of the

study site. Simulation with the meteorological inputs at daily scale was not considered in this study because the daily scale model requires intensive computational resources. The calibration targets of the constructed model consisted of the groundwater level (GWL) and surface discharges. Similar to input datasets, the calibration target datasets from 2000 to 2019 were processed for the monthly normal format. Model calibrations were performed by combining HGS with PEST. The effect of observation errors on the PEST model was tested with two different approaches: (1) adjustment of the weight matrix considering observation errors and (2) stochastic realization of observation sets using observation error functions. Figure 2 shows the conceptual flow chart of this study for incorporating observation errors in PEST simulation.

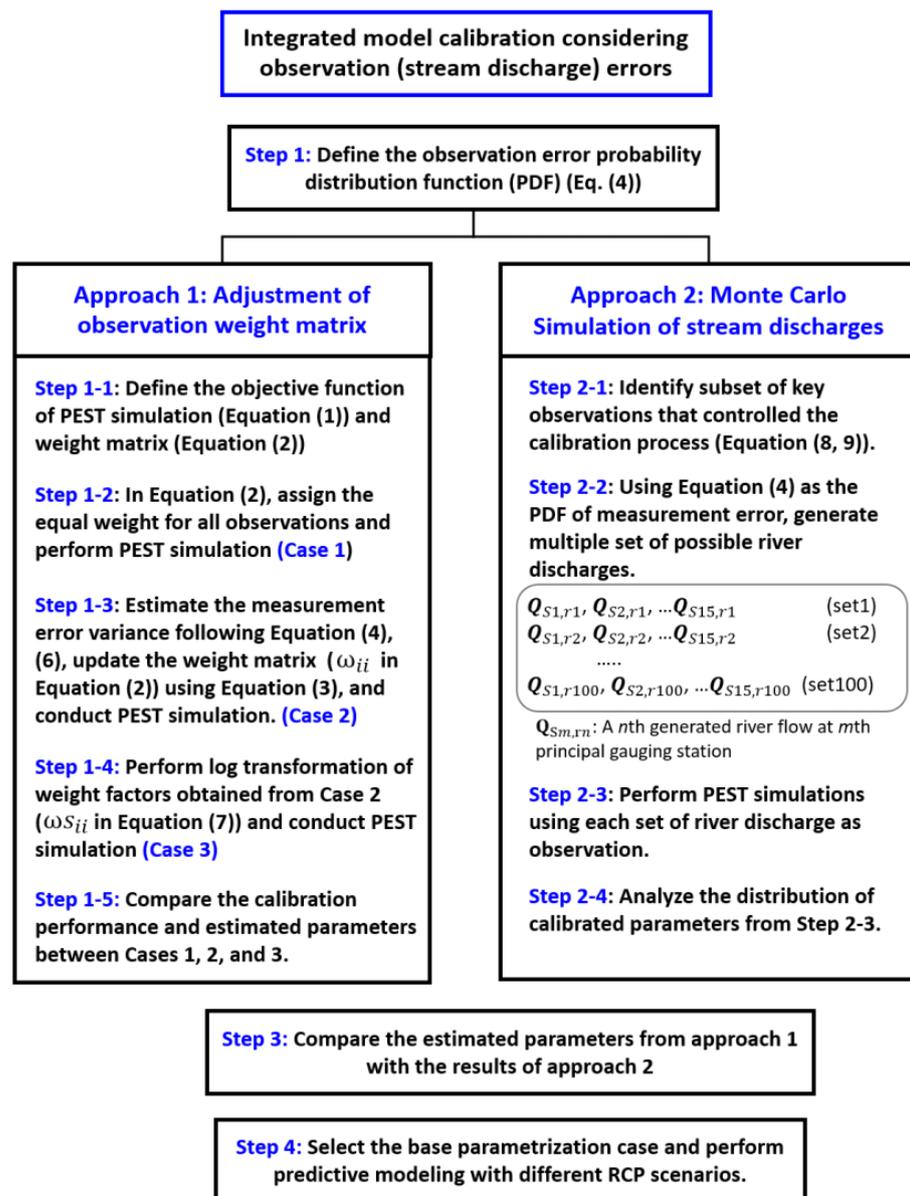


Figure 2. Conceptual flow chart of this study for incorporating observation errors in PEST simulations.

2.3.1. Estimation of Observation Error Weight Matrix for PEST

In PEST, a set of target parameters was obtained by minimizing the objective function:

$$S = (y - y'(\mathbf{b}))^T \omega (y - y'(\mathbf{b})) \tag{1}$$

where  $\mathbf{S}$  is the objective function,  $\mathbf{y}$  is the observation vector,  $\mathbf{y}'(\mathbf{b})$  is the simulated equivalent vector,  $\mathbf{b}$  is the parameter vector, and  $\omega$  is the observation weight matrix [30]. Assuming that the observation errors are uncorrelated,  $\omega$  can be approximated as a diagonal matrix as follows:

$$\omega_{ii} = \frac{\alpha_i}{\sigma_{y_i}^2} \quad i = 1, \dots, n, \dots, nd \tag{2}$$

where  $nd$  is the number of total observations (including  $n$  observations from GWL and  $nd-n$  observations from surface discharges),  $\alpha_i$  is the weight adjustment factor to ensure that no subobjective functions dominate the inversion process [28], and  $\sigma_{y_i}^2$  is the observation error variance [30]. In this study, we assumed that the observation error variance stemmed from two sources:

$$\sigma_{y_i}^2 = \sigma_{ym_i}^2 + \sigma_{ys_i}^2 \quad i = 1, \dots, n, \dots, nd \tag{3}$$

where  $\sigma_{ym_i}^2$  is the measurement error variance and  $\sigma_{ys_i}^2$  is the sampling error variance of the  $i$ th observation. The measurement error variance indicates the accompanying errors during the observation data acquisition. Previous studies that measured stream discharges at the study site reported that the error ranges of the stream discharges varied from 8 to 29% of their average depending on the location [58,59]. Following the suggestion by [33], we assumed that the measurement error probability function for stream discharge followed the truncated Gaussian distribution:

$$\begin{aligned} & Q_{i,measured} \\ & \propto N(Q_{i,True}, \sigma_{ym_i}^2) \text{ where } |Q_{i,measured} - Q_{i,True}| < 1.96\sigma_{ym_i} \\ & = 0 \text{ where } |Q_{i,measured} - Q_{i,True}| \geq 1.96\sigma_{ym_i} \quad i = n + 1, \dots, nd \end{aligned} \tag{4}$$

where  $Q_{i,measured}$  is the measured discharge taking into account both the true river discharge ( $Q_{i,True}$ ) and measurement error. The variance of each discharge observation was chosen to give a 95% confidence interval at the error ranges (8–29%) of the measured discharges (i.e.,  $\sigma_{ym_i} = (0.04 \sim 0.15) \times Q_{i,measured}$ ,  $i = n + 1, \dots, nd$ ). The measurement error variance of GWL was relatively small compared to stream discharge and, thus, was not considered in this work  $\sigma_{ym_i}^2 \approx 0$ ,  $i = 1, \dots, n$ ).

The sampling error was defined as the potential variance of the observation caused by the limited number of data acquisitions. Because the constructed model was calibrated using the monthly normal state, all model inputs (rainfall, temperature, etc.) and observations were processed as the monthly normal means for the period from 2000 to 2019. Some observations, however, did not possess full data during the target period for different reasons (some data were lost due to device malfunction, while some observation stations were installed recently). In such cases, the true monthly normal observations covering the whole duration of the target period could be statistically inferred from the sampled observations. Assuming a normal approximation of the distribution of the sample mean, the difference between the true mean and the sample mean of the  $i$ th observation lies within  $[-\sigma_{ys_i}, \sigma_{ys_i}]$  at a 95% confidence level:

$$Q_{i,measured} \propto N(Q_{i,True}, \sigma_{ys_i}^2) \tag{5}$$

$$\sigma_{ys_i} = 1.96 \frac{\sigma_{yo_i}}{\sqrt{m_i}} \tag{6}$$

where  $\sigma_{ys_i}$  is the 95% standard error taking into account the uncertainty of observation errors,  $\sigma_{yo_i}$  is the standard deviation of sampled values corresponding to the  $i$ th observation, and  $m_i$  is the size of samples (number of available data) for the  $i$ th observation during the target period. Next, the weight adjustment factor,  $\alpha_i$ , was estimated from the preliminary calibration process to ensure that the contributions from multiple subobjective functions were almost equal. Finally, the estimated  $\omega_{ii}$  ranged from 0.002 to 265.9 for surface discharges and from 0.001 to 12.77 for GWLs. In terms of the weight for surface discharge, we observed that smaller values of weights were assigned for surface peak flows because the

$\sigma_{y_{m_i}}^2$  in Equation (4) was proportional to the magnitude of surface discharges. With smaller values of weight for the peak flow, the calibrated model might tend to fit the hydrograph preferably to the low-flow conditions [60]. Accordingly, we introduced a log-transformed weight factor  $\omega_{s_{ii}}$  to reduce the difference between the maximum and minimum weight factors using the following equation:

$$\omega_{s_{ii}} = \ln \left( 1 + \frac{\alpha_i}{\sigma_{y_i}^2} \right) \quad (7)$$

The estimated  $\omega_{s_{ii}}$  ranged from 0.61 to 18.26 for surface discharges and from 0.005 to 2.0 for GWLs. Next, PEST was operated with three different weighting schemes: (1) weights assigned equally for all observations (Case 1); (2) weights assigned using Equation (2) (Case 2); and (3) weights assigned using Equation (7) (Case 3).

### 2.3.2. Stochastic Realization of Observation Sets Considering Measurement Uncertainty

We applied a stochastic random generation approach to reflect potential uncertainty arising from the measurement errors. Among the two observation types used for the model calibration, stream discharges were chosen as the random variable considering the larger uncertainty of its measurements compared to GWL. Since the stochastic model combined with the 3D integrated model requires intensive computational resources and simulation time, we introduced principal component analysis (PCA) of the sensitivity matrix to identify a subset of key observations that controlled the parameter inversion process. The adopted method can effectively reduce the number of target variables while minimizing information loss stemming from limited simulation cases.

A lumped sensitivity matrix (**S**) was first estimated, the element of which was defined as follows:

$$s_{ij} = \frac{\partial \sum_{o=1}^{u(i)} y(b)_{i,o}}{\partial b_j} \quad i = 1, \dots, k, \quad j = 1, \dots, l \quad (8)$$

where  $b_j$  is the  $j$ th calibration parameter;  $y(b)_{i,o}$  is the simulated stream discharge at the  $i$ th observation point and  $o$ th time;  $u(i)$  is the number of observations at the  $i$ th location during the calibration period ( $\sum_i u(i) = nd - n$  in this case);  $k$  is the number of observation points (gauge stations); and  $l$  is the number of calibration parameters. The covariance matrix of **S**, **V**(**y**(**b**)), can be described as follows:

$$\mathbf{V}(\mathbf{y}(\mathbf{b})) = E \left[ (\mathbf{S} - \boldsymbol{\mu})^T (\mathbf{S} - \boldsymbol{\mu}) \right] \quad (9)$$

Using PCA of the covariance matrix, eigenvectors explaining more than 90% of its variance were identified, and the observations comprising those eigenvectors were chosen as the principal observations. Among the 15 gauge stations, 7 stations (stations 3, 4, 6, 7, 13, 15, and 16 in Figure 1b) were identified as a subset of observation points that mainly controlled the inversion process. Accepting the assumption that the measurement probability distribution function followed a multidimensional Gaussian distribution, the variances of which were defined as in Equation (4), random samples may be drawn from the distribution to give river discharge at each gauge station. Using a Monte Carlo approach, multiple sample sets of river discharge (100 possible cases) were taken to approximate the true joint distribution of the gauging points, and each sample was considered as a possible observation of river discharge [33]. The PEST algorithm was then applied to the integrated model with each random scenario. During the stochastic simulations, the initial parameters were assumed to be equal for all simulation cases, and  $\boldsymbol{\omega}$  was given as an identification matrix. Based on the results from the stochastic experiments, the distributions of the inversion parameter sets were approximated using a Monte Carlo approach.

2.4. Prediction of Future Variability of Watershed Processes with Different RCP Scenarios

Based on the calibrated model setup, the watershed response to changing climate forcing was examined with different RCP scenarios [2,3]. Among four different RCP scenarios (RCPs 2.6, 4.5, 6.0, and 8.5), this study used RCPs 2.6 and 8.5 to predict future water system with the best (RCP 2.6: assuming the strong decrease in greenhouse gas (GHG) emissions) and the most extreme (RCPs 8.5: assuming high GHG emissions and little mitigation efforts) scenarios. The downscaled climate models from HadGEM3-RA were then chosen for analysis [34]. The projected climate data over the study site were processed to monthly means and applied to prediction models. The simulation period was from 2011 to 2100. A statistical summary of projected climate information over the study site is provided in Supplementary Table S1 and Supplementary Figure S1.

3. Results

3.1. Comparisons of PEST Performance between Different Weighting Schemes

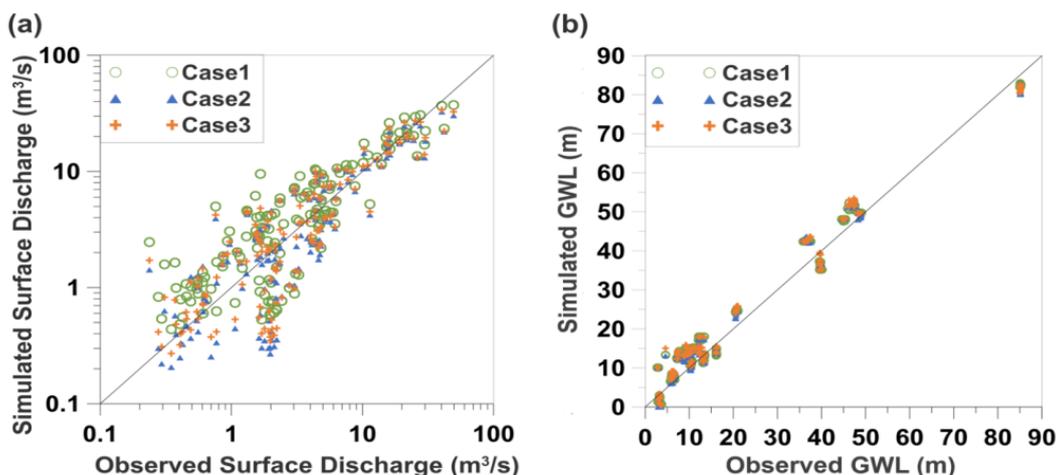
Table 4 summarizes the PEST performance results under different weighting schemes (Cases 1, 2, and 3). Four efficiency criteria were selected to estimate the model calibration performance, including the root mean square error (RMSE), the weighted residual sum of squares (RSS) normalized by the sum of weights, the coefficient of determination ( $R^2$ ), and the Nash–Sutcliffe efficiency (NSE). The RMSE and weighted RSS measure the discrepancy between the data and an estimation model, and a smaller value of these criteria indicates a better fit of the model to the data. The  $R^2$  and NSE, which are widely used to evaluate the goodness-of-fit of modeled surface discharges, range from 0.0 to 1.0 and  $-\infty$  to 1.0, respectively. For flow evaluation,  $R^2$  and NSE values closer to 1.0 indicate a better fit between the simulation and the observations [61].

Table 4. Comparisons of calibration performance between different weighting schemes.

Model Performance Criteria	Components	Case 1	Case 2	Case 3
RMSE *	Surface flow [m <sup>3</sup> /s]	3.5	3.6	3.1
	GWL [m]	3.9	3.8	4.0
Weighted RSS *	Surface flow [m <sup>3</sup> /s]	12.2	0.5	1.4
	GWL [m]	15.3	9.7	4.5
$R^2$ *	Surface flow [m <sup>3</sup> /s]	0.83	0.84	0.83
NSE *	Surface flow [m <sup>3</sup> /s]	0.83	0.82	0.83

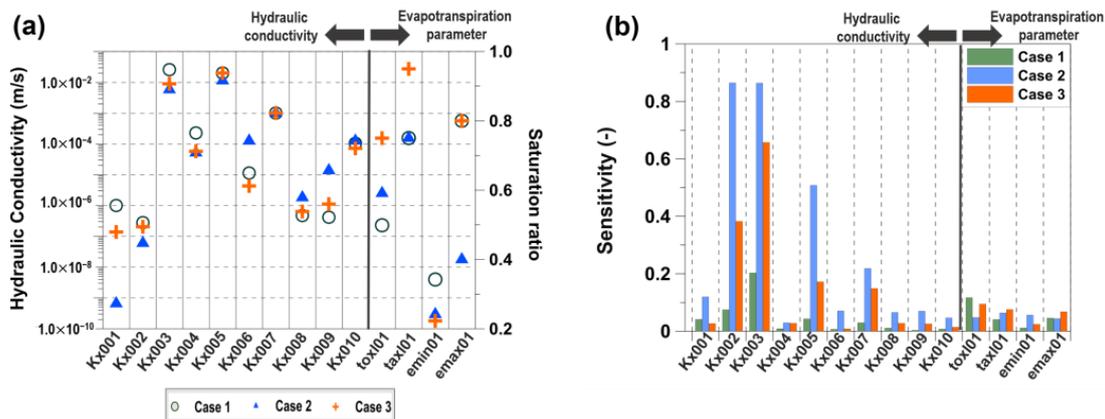
Notes: \*  $RMSE = \left[ \frac{1}{N} \sum_i (Ob_i - Sim_i)^2 \right]^{0.5}$ ;  $Weighted\ RSS = \frac{\sum_i (Ob_i - Sim_i)^2 w_i}{\sum_i w_i}$ ;  $R^2 = \left\{ \frac{\sum (Q_{ob,i} - \bar{Q}_{ob})(Q_{sim,i} - \bar{Q}_{sim})}{\left[ \sum (Q_{ob,i} - \bar{Q}_{ob})^2 \right]^{0.5} \left[ \sum (Q_{sim,i} - \bar{Q}_{sim})^2 \right]^{0.5}} \right\}^2$ ;  $NSE = 1 - \frac{\sum (Q_{ob,i} - Q_{sim,i})^2}{\sum (Q_{ob,i} - \bar{Q}_{ob})^2}$  where  $ob_i$ : *i*th observation;  $sim_i$ : simulation at *i*th observation;  $N$ : number of total observations;  $w_i$ : weights assigned for *i*th observation;  $Q_{ob,i}$ : observed surface flow at *i*th observation;  $\bar{Q}_{ob}$ : mean of observed surface flow inclusive of all surface observations;  $\bar{Q}_{sim}$ : mean of simulated surface flow inclusive of all surface observations, and  $Q_{sim,i}$ : simulated surface flow at *i*th observation.

Comparisons of the three different weighting schemes revealed that the total error of the modeled system was the smallest in Case 3 (case using log-transformed weight factors) for both the RMSE and the weighted RSS (Table 4). Model performance efficiency was analyzed for each calibration type (surface flow and GWL) and showed that Case 2 and Case 3 slightly outperformed Case 1, but the differences between the models were minor. The  $R^2$  and NSE did not show significant variation among the three test cases. These results indicate that the observation error-based weighting schemes used in this study could slightly increase the inversion model performance in terms of error analysis, although their effect was small. The one-to-one plots between observed and simulated variables (Figure 3) also showed that the differences in the simulation results among the three weighting schemes were not significant.



**Figure 3.** Comparisons of observed and simulated (a) surface flow ( $m^3/s$ ) and (b) GWL (m; amsl) under different simulation cases.

Compared to the error evaluation (Table 4), the estimated model parameters showed a more scattered pattern depending on the applied weighting schemes (Figure 4a). Discrepancies in estimated hydraulic conductivity were noticeable in weathered rock (kx001), loam (kx006), and silty clay loam (kx009), whereas those from the remaining subsurface zones were similar for all cases. Figure 3 also reveals that the effect of the applied weighting scheme on the parameterization was more prominent for ET-related properties. In particular, the estimated *tox101*, *tax101*, and *emax01* exhibited largely scattered distributions for each test case. Despite similar error values obtained from different simulation cases (Table 4), the estimated parameters could be considerably different depending on the applied weighting schemes.



**Figure 4.** Comparisons of (a) estimated model parameters and (b) parameter sensitivities between different weighting schemes.

A sensitivity analysis of observations (including both GWL and surface discharge from all observation points in the watershed) to changes in the calibration parameters indicated that *kx002*, *kx003*, *kx005*, and *kx007* were major properties dominating the observation sensitivity (Figure 4b). In general, observations were more sensitive to the model properties in Case 2, with the exception of the ET parameters. The high sensitivities of Case 2 and Case 3 compared to Case 1 could be associated with the larger weights applied to those models.

Figure 5 compares the measured monthly averaged surface discharges and GWL with the simulation results (Case 3). The figure shows that the constructed model reasonably matched the measured surface discharges. The time-series plot revealed that the surface

water system of the study site was significantly affected by seasonal forcing, whereas the temporal variation in GWL was comparatively small. The simulation results from Cases 1 and 2 also exhibited similar distributions to that of Case 3 with minor variations.

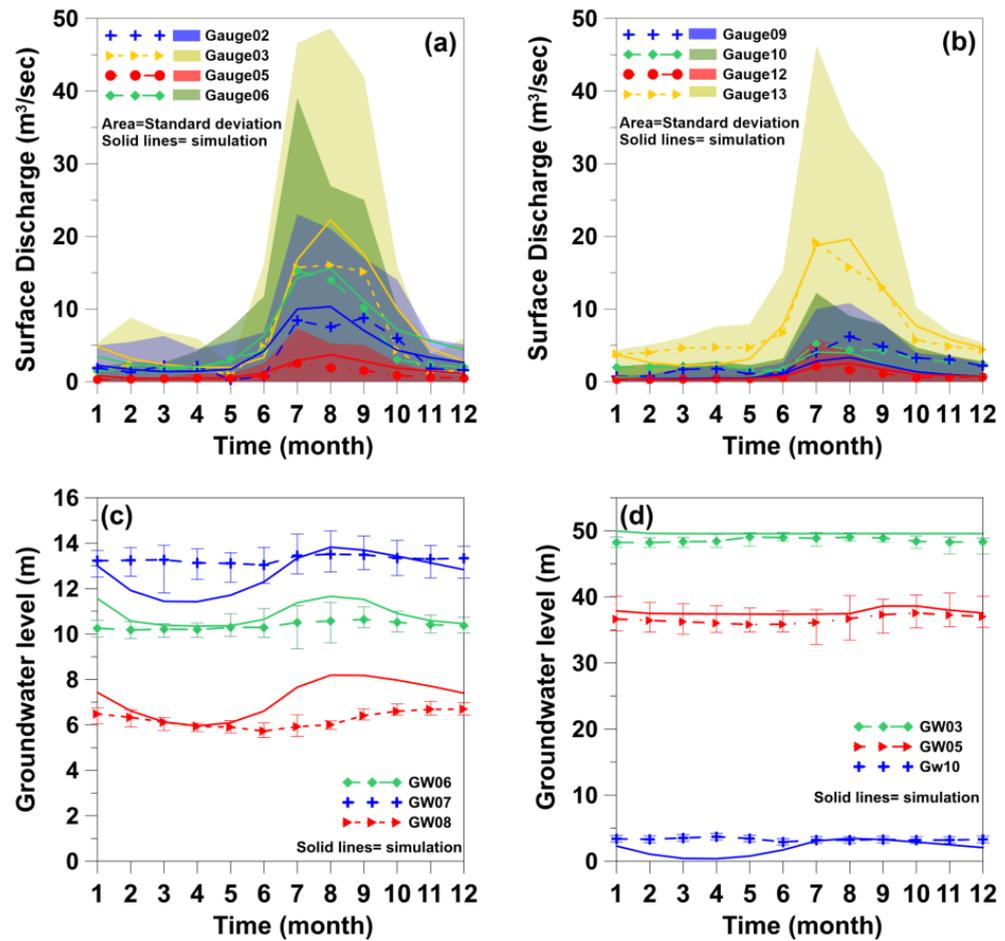


Figure 5. (a,b) Monthly variations in surface discharges ( $m^3/s$ ) from (a) Gauge02, 03, 05, 06; (b) Gauge09, 10, 12, 13, and GWL (m; amsl) observed from (c) GW06, 07, 08; (d) GW03,05,10 of the study site.

### 3.2. Results of Stochastic Simulations

Figure 6 shows the histograms of the log-transformed hydraulic conductivities obtained from the stochastic models and PEST inversion. We observed that the histogram of each calibration target had different distribution shapes. The distributions for  $kx002$ ,  $kx004$ ,  $kx005$ , and  $kx006$ , for example, resembled a skewed pattern, whereas those from other calibration targets had normal ( $kx001$ ), uniform ( $kx003$ ), or bimodal ( $kx007$  and  $kx008$ ) shapes. Some calibration targets ( $kx002$  and  $kx005$ ) showed relatively high peak values ( $>0.8$ ) in their modes (the most frequently predicted value), with minor variations in their predicted values. Other calibration targets exhibited a more dispersed distribution with reduced probability at the peak. For example, the models predicted that log-transformed  $kx001$  could be present between  $[-8.5, -6]$  (over two orders of magnitude difference) with a peak value of 0.6 at the interval of  $[-7.5, -7]$ .

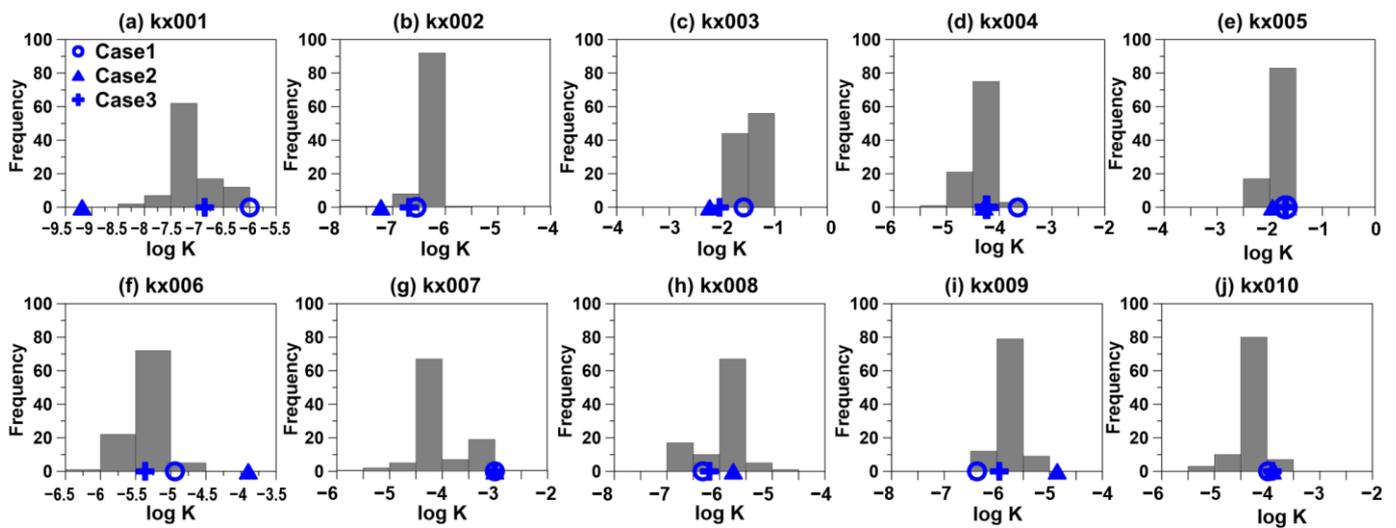


Figure 6. Histograms of estimated hydraulic conductivities from the stochastic models and comparisons with the results from the weighting tests (Cases 1, 2, and 3).

The histograms for ET-related parameters (Figure 7) could be interpreted as having bimodal (tox1001 and tax1001) or skewed (emin001 and emax001) distributions. The ET parameters exhibited a more dispersed distribution than did the hydraulic conductivities. In the cases of tox1001 and emax001, the peak values were lower than 0.5 and, even in the case with a peak over 0.5 (tax1001), its histogram showed a distinct bimodal shape.

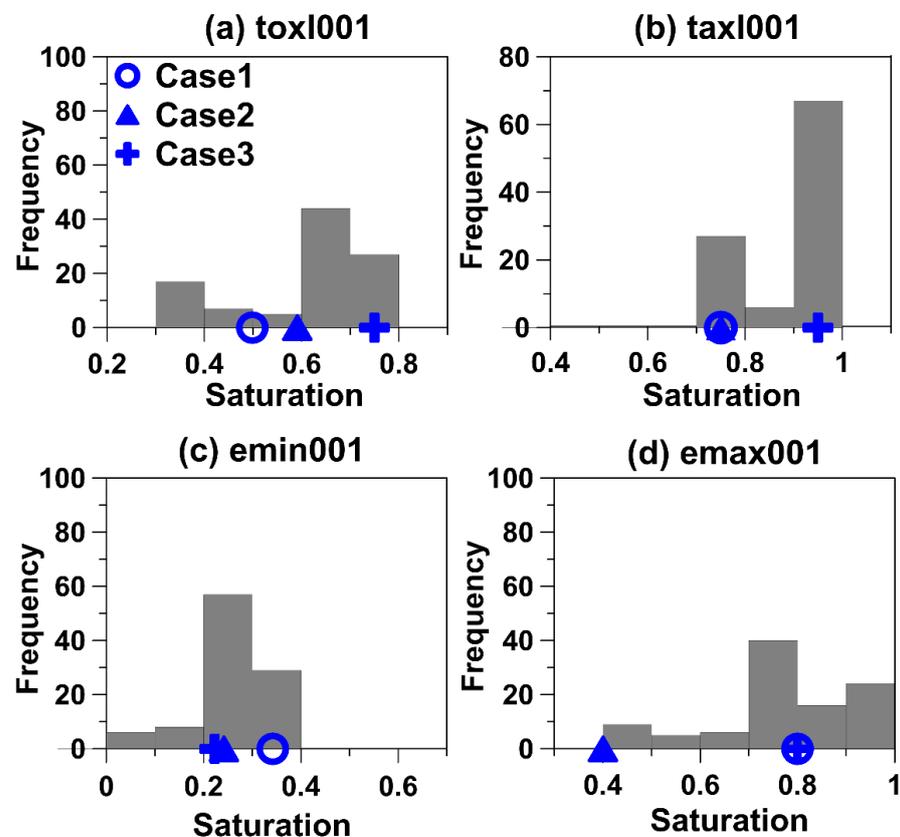


Figure 7. Histograms of estimated ET parameters from the stochastic model comparisons with the results from the weighting tests (Cases 1, 2, and 3).

The symbols in Figures 6 and 7 indicate the values of the calibration parameters predicted from the different weighting scenarios. Graphical comparison indicates that, among the three tested weighting schemes, predictions from Case 3 were closest to the mode of stochastic simulations for most of the calibration targets. For some parameters (kx007), however, all tested weighting schemes failed to capture the mode value. Comparisons with the mean value from the stochastic simulations (Table 5) revealed that predictions from Case 3 reasonably matched the averages of calibration targets without significant discrepancy to each other (except for kx003, kx007, kx008, toxl001, and emin001).

**Table 5.** Estimated values of calibration targets under different weighting schemes and comparison with the result from the stochastic model.

Parameter	PEST with Different Weights			Stochastic Model		
	Case 1 (Difference Ratio) *	Case 2 (Difference Ratio)	Case 3 (Difference Ratio)	Mean	Standard Deviation	
log(K) (m/s)	kx001	$1.00 \times 10^{-6}$ (4.43)	$6.60 \times 10^{-10}$ (1.00)	$1.39 \times 10^{-7}$ (0.24)	$1.84 \times 10^{-7}$	$2.86 \times 10^{-7}$
	kx002	$2.80 \times 10^{-7}$ (0.33)	$6.04 \times 10^{-8}$ (0.85)	$2.04 \times 10^{-7}$ (0.51)	$4.16 \times 10^{-7}$	$8.61 \times 10^{-8}$
	kx003	$2.59 \times 10^{-2}$ (0.20)	$5.86 \times 10^{-3}$ (0.82)	$8.97 \times 10^{-3}$ (0.72)	$3.23 \times 10^{-2}$	$3.42 \times 10^{-3}$
	kx004	$2.26 \times 10^{-4}$ (4.50)	$5.17 \times 10^{-5}$ (0.26)	$5.76 \times 10^{-5}$ (0.40)	$4.11 \times 10^{-5}$	$2.74 \times 10^{-5}$
	kx005	$2.00 \times 10^{-2}$ (0.34)	$1.14 \times 10^{-2}$ (0.23)	$2.00 \times 10^{-2}$ (0.34)	$1.49 \times 10^{-2}$	$4.93 \times 10^{-3}$
	kx006	$1.14 \times 10^{-5}$ (1.79)	$1.27 \times 10^{-4}$ (30.13)	$4.32 \times 10^{-6}$ (0.06)	$4.08 \times 10^{-6}$	$2.88 \times 10^{-6}$
	kx007	$1.00 \times 10^{-3}$ (3.41)	$1.00 \times 10^{-3}$ (3.41)	$1.00 \times 10^{-3}$ (3.41)	$2.27 \times 10^{-4}$	$3.42 \times 10^{-4}$
	kx008	$4.74 \times 10^{-7}$ (0.72)	$1.82 \times 10^{-6}$ (0.08)	$6.39 \times 10^{-7}$ (0.62)	$1.69 \times 10^{-6}$	$1.79 \times 10^{-6}$
	kx009	$4.19 \times 10^{-7}$ (0.78)	$1.39 \times 10^{-5}$ (6.20)	$1.11 \times 10^{-6}$ (0.42)	$1.93 \times 10^{-6}$	$9.80 \times 10^{-7}$
	kx010	$1.05 \times 10^{-4}$ (0.77)	$1.26 \times 10^{-4}$ (1.13)	$7.16 \times 10^{-5}$ (0.21)	$5.92 \times 10^{-5}$	$4.02 \times 10^{-5}$
ET Properties	toxl001	0.50 (0.18)	0.59 (0.03)	0.75 (0.23)	0.61	0.14
	taxl001	0.75 (0.16)	0.75 (0.16)	0.95 (0.07)	0.89	0.08
	emin001	0.34 (0.21)	0.24 (0.14)	0.22 (0.21)	0.28	0.09
	emax001	0.80 (0.05)	0.40 (0.47)	0.80 (0.05)	0.76	0.16

Note: \* Difference ratio was defined as  $\frac{|\text{predicted value from each Case} - \text{mean of stochastic models}|}{\text{mean of stochastic models}}$ .

### 3.3. Analysis of Seasonal Hydrologic Variation in Groundwater-Surface Water Integrated System

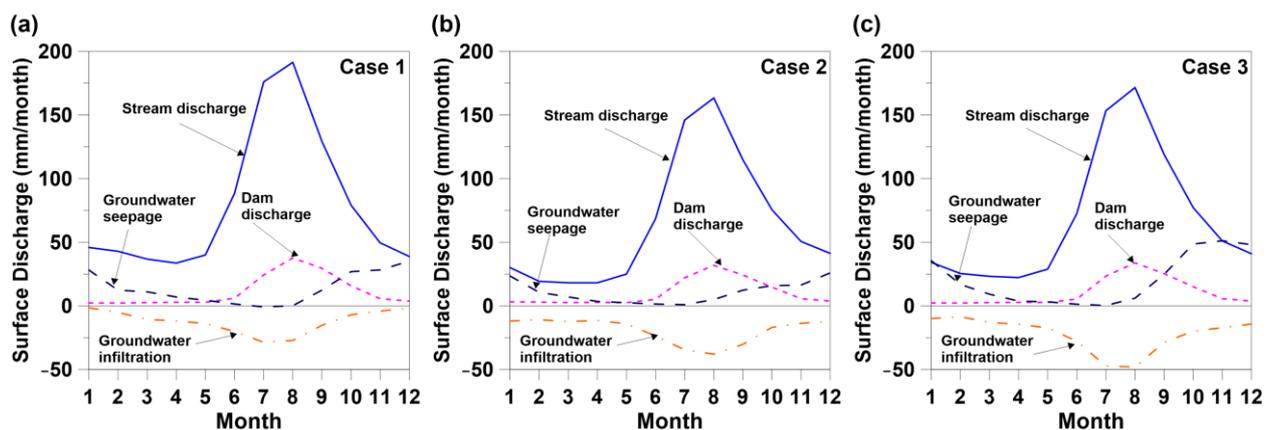
Based on the calibrated model setup, we analyzed and compared seasonal hydrological variations in the integrated flow system for three tested simulation cases. Table 6 summarizes the hydrological water balance of the study site. The surface discharge showed significant seasonal variations, characterized by high discharge rates (129 to 146 mm/month) during the rainy season (June to September) and low flow rates (35 to 46 mm/month) during the dry season (December to May). During the rainy season, surface discharge comprises approximately 61 to 72% of the total rainfall, whereas actual evapotranspiration (AET) accounts for approximately 12 to 15% of the rainfall. With a substantial decline in rainfall during the dry season, however, the contribution of AET to the water balance increases (32 to 40% of the total rainfall) while the amount of surface discharge significantly decreases.

**Table 6.** Simulated hydrological balance of the study site with groundwater and dam contributions to stream discharge (2000~2019).

Hydrologic Component		Amount (mm/Month)		
		Case 1	Case 2	Case 3
Rainy Season	Rain	203.4	203.4	203.4
	AET	25.3	31.5	31.1
	Surface Discharge	146.3	123.2	129.1
	Groundwater Seepage <sup>1</sup>	3.3	4.9	8.1
	Stream Water Infiltration <sup>1</sup>	22.8	34.0	38.0
	Dam Discharge	24.4	20.8	22.0
Dry Season	Rain	49.1	49.1	49.1
	AET	15.9	19.6	18.7
	Surface Discharge	45.8	34.7	37.9
	Groundwater Seepage <sup>1</sup>	19.1	17.1	27.0
	Stream Water Infiltration <sup>1</sup>	7.0	11.9	14.3
	Dam Discharge	4.7	4.8	4.6

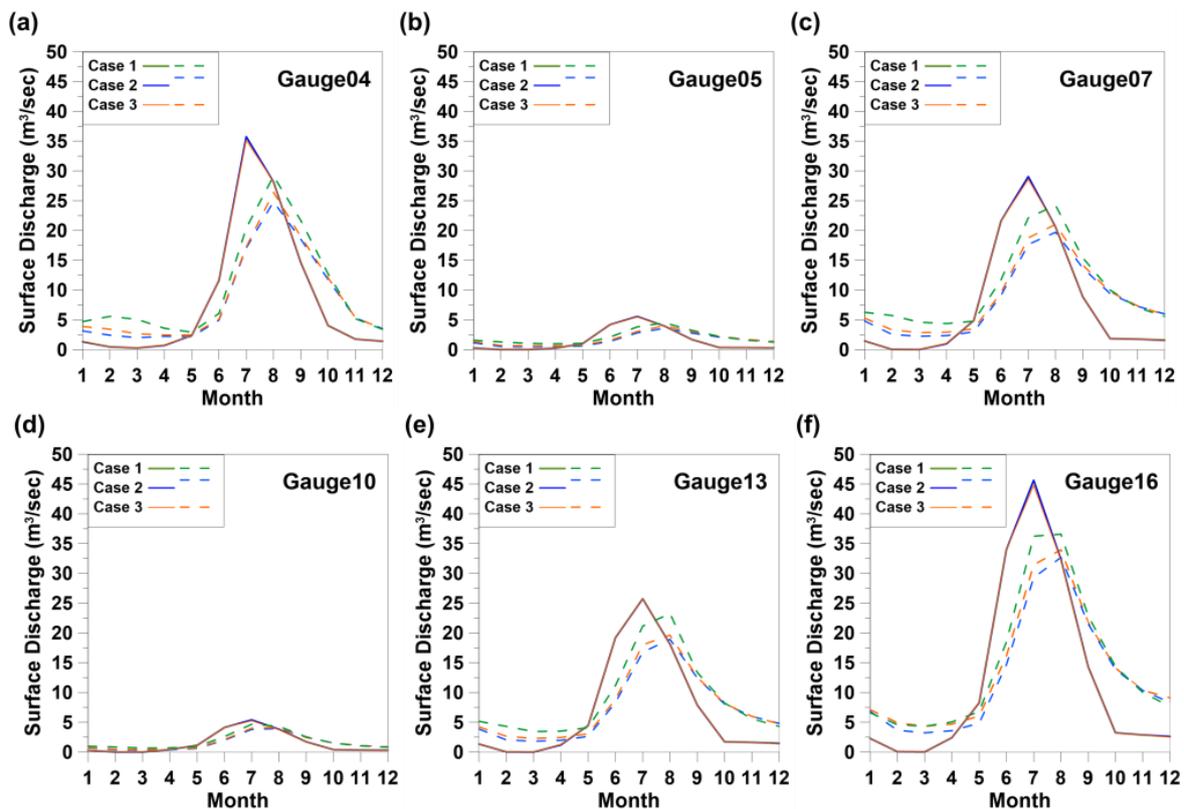
Note: <sup>1</sup> Components indicate groundwater seepage/infiltration to/from the streambed.

The exchange flux between aquifers and streams also showed strong seasonal variations. The direction of the water exchange flux during the rainy season could be characterized by net surface water infiltration into the aquifer, whereas, during the dry season, groundwater seepage mainly feeds the surface discharge (Table 6). During the rainy season, for example, the total amount of infiltrated stream water accounted for 13 to 24% of the total surface discharge. The contribution of groundwater to surface discharge was greater in the dry season, as the net groundwater discharge comprised 15 to 34% of the total surface discharge. Details on the seasonal variations in groundwater-stream dynamics can also be found in Figure 8. The seepage-dominated exchange flux during the dry season and infiltration-dominated water exchange during the rainy season demonstrated the direct contribution of groundwater to surface discharges.



**Figure 8.** Hydrograph separation results for stream discharges, groundwater infiltration, groundwater seepage, and dam discharge for (a) Case 1, (b) Case 2, and (c) Case 3.

The dynamic interaction between groundwater and streams was also demonstrated by comparing hydrographs between the integrated model and the surface flow model that excluded the groundwater flow system. In Figure 9, solid lines indicate the predicted hydrographs at different gauge stations without groundwater components and dashed lines indicate those from the integrated model. The comparison of hydrographs revealed that groundwater storage, which was filled during the rainy season and then slowly released during the dry season, supported the stream system and contributed to the buffering of temporal variations in stream discharge.



**Figure 9.** Monthly varying hydrographs at the selected gauge stations (a) Gauge04, (b) Gauge05, (c) Gauge06, (d) Gauge10, (e) Gauge13, and (f) Gauge16. Solid lines indicate hydrographs obtained from the surface flow model, and dashed lines show the hydrograph from the surface water-groundwater integrated models.

In summary, the comparisons of seasonal hydrological variations among the three different weighting schemes showed that differences existed for both water balance and hydrographs. Among the three tested cases, Case 3 predicted a more dynamic interaction between groundwater and streams. In Case 2, the groundwater seepage contribution to stream discharge was smaller than that in Cases 1 and 3, resulting in a decrease in estimated surface discharge. Despite the differences observed from the model predictions, however, all tested models exhibited similar seasonal variations for both water budget and groundwater-surface water interaction processes.

### 3.4. Predictions of Hydrological Responses and Groundwater-Surface Water Interactions under Different Climate Change Forcings

Tables 7 and 8 show the predicted hydrologic balance of the study site under the RCP2.6 and RCP8.5 scenarios, respectively. Climate change impacts were mainly analyzed using the Case 3 integrated model because the parameters obtained from Case 3 were closest to the mode of stochastic simulations for most calibration targets (Figures 6 and 7). The results from Case 1 and Case 2 also showed similar trends, with minor variations in their estimated values. The prediction model showed a general increasing trend of AET for both the RCP2.6 and RCP8.5 scenarios, which was likely to be associated with increasing temperature. Surface discharges also exhibited decadal variations, but their varying magnitudes were relatively small compared to the temporal variations in rainfall.

**Table 7.** Predicted hydrological balance of the study site with groundwater and predicted dam contributions to stream discharge under the RCP2.6 scenario (Case 3).

Hydrologic Component		Amount (mm/Month)		
		2020s (2011–2040)	2050s (2041–2070)	2080s (2071–2100)
Rainy Season	Rain	210.3	187.5	206.5
	AET	36.5	37.2	36.8
	Surface Discharge	118.4	106.3	114.8
	Groundwater Seepage	34.3	38.3	38.3
	Stream Water Infiltration	62.3	57.8	59.7
	Dam Discharge	20.4	18.5	19.6
Dry Season	Rain	70.4	73.5	68.0
	AET	21.3	21.6	21.7
	Surface Discharge	42.9	42.9	39.5
	Groundwater Seepage	45.2	44.5	44.7
	Stream Water Infiltration	25.5	27.2	26.6
	Dam Discharge	6.0	5.8	5.2

**Table 8.** Predicted hydrological balance of the study site with groundwater and predicted dam contributions to stream discharge under the RCP8.5 scenario (Case 3).

Hydrologic Component		Amount (mm/Month)		
		2020s (2011–2040)	2050s (2041–2070)	2080s (2071–2100)
Rainy Season	Rain	211.3	206.8	249.9
	AET	36.7	38.2	39.9
	Surface Discharge	106.2	106.3	126.3
	Groundwater Seepage	17.6	17.8	17.4
	Stream Water Infiltration	45.7	43.9	50.3
	Dam Discharge	18.3	18.5	22.0
Dry Season	Rain	68.7	61.3	62.4
	AET	20.7	22.1	23.1
	Surface Discharge	41.9	42.9	44.3
	Groundwater Seepage	36.5	39.8	43.1
	Stream Water Infiltration	16.7	18.4	19.3
	Dam Discharge	5.6	5.8	6.7

The models predicted that groundwater infiltration and seepage would play a significant role in reducing stream discharge during the rainy season and maintaining it during the dry season. The predicted net infiltration of stream water during the rainy season accounted for 18 to 24% and 25 to 26% of the total surface discharge under RCP2.6 and RCP8.5, respectively. Additionally, the net groundwater seepage during the dry season contributed 40 to 46% and 47 to 54% of the surface discharge under the RCP2.6 and RCP8.5 scenarios, respectively. It is notable that, compared to the current state model (Table 6), the total amount of both groundwater seepage and stream infiltration tended to increase under the projected climate conditions, suggesting a more dynamic interaction between groundwater and streams. Increasing feedbacks between groundwater and streams could also be observed in Figure 10, as shown by the higher infiltration and seepage flux, especially during the dry season. Similarly, the predicted hydrographs at several gauge stations under the RCP8.5 scenario suggested that groundwater would attenuate the hydrological response of the integrated system under changing climate forcing (Figure 11).

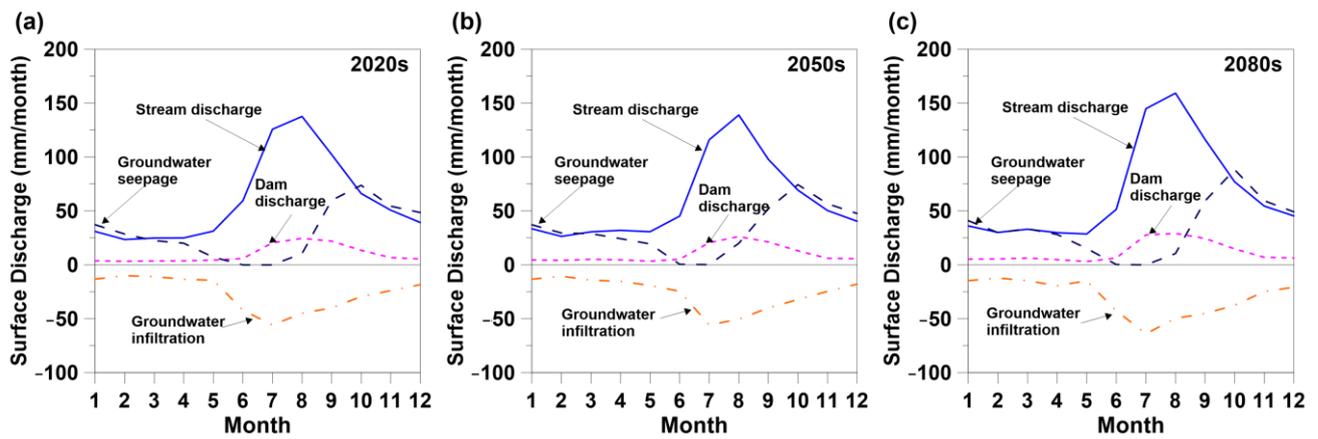


Figure 10. Predicted hydrograph separation results under RCP8.5 projection with the calibrated hydrogeologic condition from Case 3. Monthly averages at (a) 2020s, (b) 2050s, and (c) 2080s.

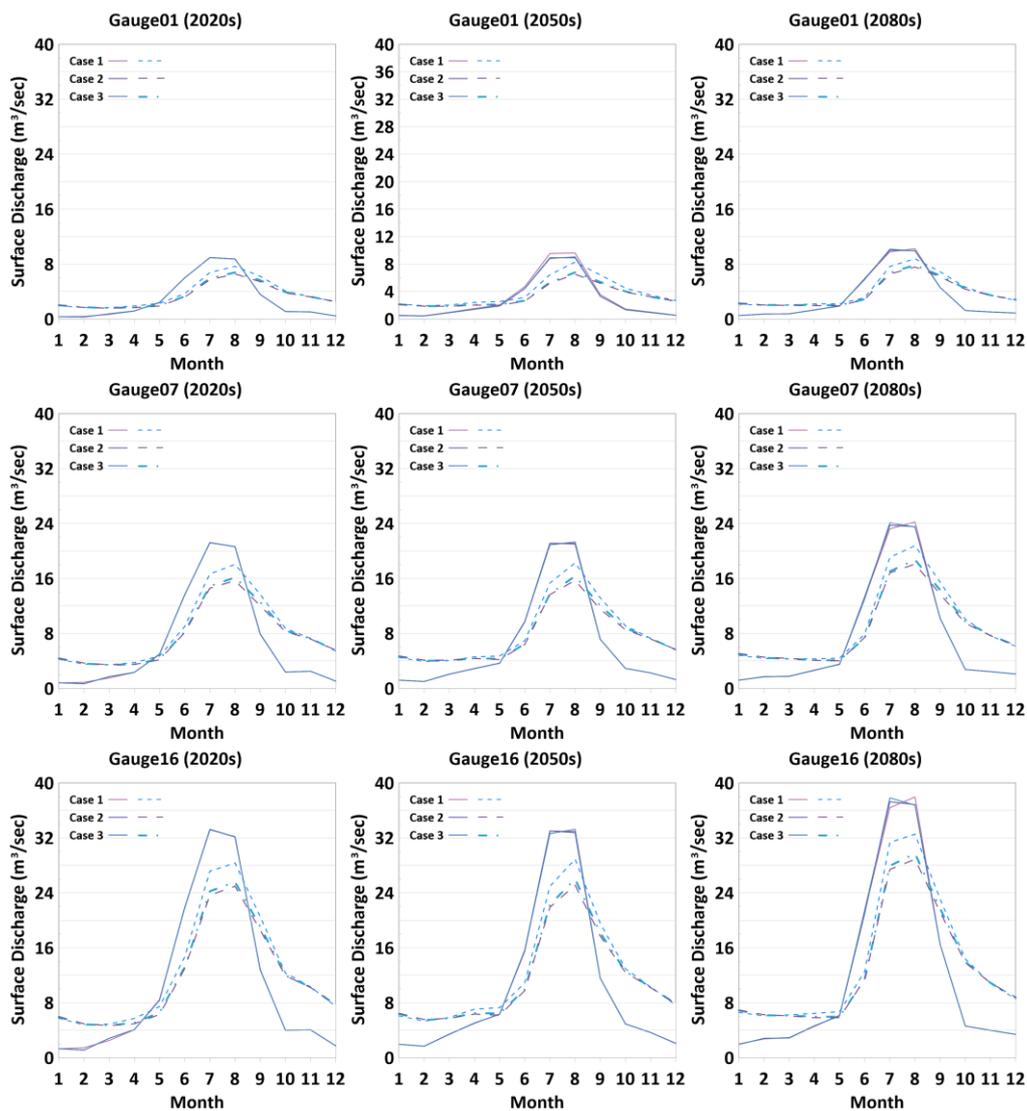


Figure 11. Monthly varying hydrographs at the selected gauge stations of the study site. Solid lines indicate hydrographs obtained from the surface flow model and dashed lines show the hydrograph from the surface water-groundwater integrated models.

## 4. Discussions

### 4.1. Effect of the Different Observation-Error Weighting Schemes on the Parametrization of the Integrated Model

In this study, we applied three different weighting schemes to analyze the potential impact of observation error in the parameterization processes of the 3D integrated model using PEST. The observation error considered in this work includes measurement error and sampling error. The measurement error was used to describe the difference between a measured quantity and its true value. The sampling error was defined as the difference between the true monthly means (i.e., population means) for the whole duration of the simulation period and the monthly averages estimated from the available sample data (i.e., sample means) from each observation point. Although the statistics that defined sampling error in this study were associated only with the temporal mean, the error could be used to represent the difference in areal mean between the observations and a true value. For example, Linsley and Kholer [62] reported that areal rainfall estimates could have standard errors exceeding 30% of their average if the gauge network is sparse [63].

The observation weight matrix was formulated to consider the potential variances in the observation error. Assuming that the variance in the observation error followed the truncated Gaussian distribution and that both measurement and sampling errors were uncorrelated and independent, we estimated the error covariance and the corresponding weight for each observation (Case 2). Furthermore, we introduced a log-transformed weight matrix to reduce the difference between the maximum and minimum weights to avoid larger weights dominating the parameterization process (Case 3). Comparisons of calibration performance between different weighting schemes showed that, in terms of error analysis, Case 2 and Case 3 slightly outperformed the case with equal weights (Case 1). This result demonstrates that the observation weight matrix applied in this work performed properly and contributed to a minor improvement in the model parameterization.

In terms of estimated parameter values, the estimated parameter values could be different depending on the applied weighting scheme. A comparison of the estimated parameter value with its sensitivity revealed that the calibration targets that had lower sensitivities to the observations showed larger variations in their predicted values. This result could be associated with the relatively small contribution of these parameters to the model calibration. PEST seeks the solutions of Equation (1) by minimizing the weighted least square between the observation and the simulation. Accordingly, for the parameters whose variation results in significant changes in the objective function, the numerical model can readily determine the steepest gradient; therefore, the solution of the matrix is likely to converge to the same optimized values if the initial condition is equal. However, for the other set of parameters that have a minor effect on the observation, the numerical model could have difficulty searching for the unique solution, thus leading to considerable variations in its solutions depending on the model condition.

### 4.2. Potential Implication of Observation Error to the Model Parameterization and Performance

The results of the stochastic model showed that the potential uncertainty of observation error could affect the parameterization process of the integrated model. We observed that the histogram of each calibration target showed a different pattern, implying that the effect of observation error on model calibration varied depending on the calibration targets. The comparison of the histogram with sensitivity analysis showed that, similar to the results from the weighting scheme test (Section 4.1), the calibration targets that had lower sensitivities to the observations tended to have large intervals in their predicted values.

In terms of ET parameters, the histograms of these parameters showed large variations; thus, it was difficult to estimate the 'representative values' of these parameters. This result indicates that the model constructed for use in this study may not fully capture varying dynamics of the ET process of the study site. This could arise from the lack of structural representation of the watershed system, particularly due to the limited discretization of the surface zone associated with ET dynamics. Additionally, the observation type used in this

study (GWL and surface discharge) may not be enough to represent the spatiotemporal ET processes of the study site. This result demonstrates that the inclusion of different types of observations, such as soil moisture and measured ET rate, could improve the model calibration performance and site characterization.

The observation error-based weighting schemes used in this study were shown to be suitable for obtaining the optimized values of the calibration targets. Among the three test cases, predictions from Case 3 were closest to the “representative value” from the stochastic simulations for most calibration targets. The total error between the simulation and the observations was also the smallest in Case 3. For Case 2, some of the calibrated parameters were out of the ranges of those estimated from the stochastic model. This result implies that the error function applied in this study (Gaussian distribution with variances obtained from the measurement error) may not be the best option to derive the error covariance. Because our chosen error covariance was proportional to the magnitude of surface discharge, it assigned a smaller weight to the peak flow and a larger weight to the low-flow condition. Under metrological conditions with distinct seasonal variations, such as those in our study site, this could result in large variations in the observation weights with seasons, creating a subset of the observation period that dominates the calibration process. The improved calibration performance of Case 3 also demonstrates that reducing the weight difference between observations would be more suitable for our study site. This result implies that the proper choice of weighting scheme could be important in the model calibration process, particularly when the observation errors show large variations. Therefore, additional investigation of the error covariance evaluation for surface discharge would be necessary for the proper selection of the observation error function and the improved parametrization of the integrated model of the study site.

We used the PCA of the sensitivity matrix to identify a subset of key observations that controlled the parameter inversion process. The stochastic models were then generated using the chosen observations as random variables. A similar approach was adopted by previous studies to identify super-parameters corresponding to pilot point values of a highly parameterized inversion problem [24] or to reduce computational efficiency in highly parameterized inversion problems [64]. Although the variable chosen for PCA in this study was the observation sensitivity rather than the calibration parameters, it was shown that the adopted approach could reasonably generate different random outputs with a limited number of simulations while saving the computational requirements of the models. The approach, in particular, could be useful for large-scale integrated models that require intensive computational resources. Additionally, a comparison with predictive error analysis and uncertainty analysis in PEST could provide new information to understand the model behavior associated with observation errors and the parameterization process.

We demonstrated that, despite the abovementioned limitations, the observation error weighting scheme applied in this work reasonably generated the hydrological dynamics of the study site and provided some insights to better understand the parametrization process of the integrated models. Although the adopted weighting schemes may possess local characteristics and might not properly work for other integrated water systems, our study showed that the application and comparison of different weighting schemes for the determination of error covariance could contribute to assessing the stability of the constructed models and to reducing the model uncertainty arising from the observation error.

#### *4.3. Effect of Groundwater–Surface Water Feedbacks on the Integrated Water System and Implications for Integrated Water System Management*

In this study, we analyzed feedbacks between groundwater and surface water considering the potential uncertainty of observation errors. We observed that all three tested models with different error-weighting schemes predicted dynamic interactions between groundwater and surface water systems. The models also showed that groundwater seepage that was filled in the high-flow season and then released in the dry season served as a major water source for the stream network of the study site. Such a direct groundwater

contribution to stream flow highlights the significance of groundwater for preserving riparian ecology, especially during the low-flow season, implying that the spatial and temporal dynamics of groundwater–surface water interactions need to be considered during the assessment and management of the ecohydrological system of the study site.

Because this study adopted a 3D integrated model in a basin scale, some important aspects in groundwater–surface-water interaction study, such as local-scale variability or consideration of vertical forces at the groundwater–surface-water interface interaction, were not analyzed in this work. Furthermore, in order to gain a better understanding of the integrated water system and improve integrated water system management, it would be necessary to conduct an analysis on the potential effects of riverbed geography, seasonal inputs, and geological variability on the calibration performance of the integrated model.

The predictions based on different RCP scenarios all suggested the increasing role of groundwater in watershed dynamics, as demonstrated by increasing water infiltration and seepage fluxes. With an increasing frequency of extreme weather events such as droughts and heavy rainfall, the contribution of groundwater as a buffer for attenuating the hydrological response of integrated water systems would become more significant. It is notable that, although hydrological structures such as the Yedang dam also play a significant role in regulating stream discharges, the influence of groundwater could be more important for integrated water management because the latter has a more positive contribution to reducing seasonal and interannual variations in both the stream stages and the stream discharges. The comparison of the predicted hydrological balance between RCPs 2.6 and 8.5 scenarios highlights the larger temporal variations in rainfall and AET in RCP 8.5. This indicates that extreme weather events such as heavy rainfall or drought, which are expected to occur more frequently under the extreme RCP scenarios, would pose a greater threat to water environment and sustainability. Given these circumstances, groundwater is expected to play an increasingly important role in attenuating the hydrological response of the integrated system under changing climate forcing. Therefore, a sustainable water management strategy that successfully adapts to climate change should incorporate the integrated water management strategy that considers the connection between groundwater and surface water systems.

Among the three tested model cases using different weighting schemes, we observed that Case 2 predicted the least groundwater–surface water interaction and smallest groundwater seepage rate. Accordingly, the predicted surface discharges were the lowest in Case 2. The relatively low peak flow predicted in Case 2 could be associated with small weights assigned to high surface discharge data during the model calibration process. Due to the larger measurement errors related to the observed peak flow, model calibration was performed assuming greater uncertainty for the high-flow season and, consequently, the model suggested that the measured stream discharge during the peak flow might preferably overestimate the actual surface discharges. Despite these differences, however, the simulated seasonal hydrologic behaviors of the watershed were similar to all tested models, suggesting that groundwater of the study site is of critical importance for preservation of the surface water and hydro-ecosystem under the varying projected climate conditions.

## 5. Summary and Conclusions

In this study, we investigated the potential impact of observation error on the model calibration performance of an integrated watershed-scale model. Three different weighting schemes that consider observation error variances were suggested and applied to the PEST model. The applied weighting schemes were compared with the results from stochastic models, in which observation errors from surface discharges were considered a random variable. Comparisons of the calibration performances between different models showed that the applied weighting schemes in this work performed properly and contributed to a minor improvement in the model parameterization. Among the three test cases, predictions from Case 3 were closest to the “representative value” from stochastic simulations for most calibration targets. This result implies that the proper choice of weighting scheme

is important in the model calibration process, particularly when the observation error variances show large seasonal variations. Additional investigation of the error estimation would, thus, be necessary to improve the calibration performance of the integrated model. The inclusion of different observation types and a detailed characterization of the study site through a highly parameterized problem approach or regularized inversion could be useful to improve the calibration performance and better characterize the spatially varying dynamics of integrated water systems [19,65]. Additionally, incorporation of better performing metaheuristic evolutionary algorithms would be necessary for improved performance of stochastic models.

The analysis of the feedbacks between groundwater and surface water revealed that groundwater plays a key role in attenuating the hydrological response of integrated water systems and sustaining surface flows under seasonally varying metrological conditions. Predictions using different RCP scenarios suggested the increasing role of groundwater in watershed dynamics. This result clearly highlights the significance of groundwater for preserving riparian ecology, indicating that the varying dynamics of groundwater–surface water interactions need to be considered for the management of the ecohydrological system of the study site.

The predicted hydrological behaviors of the study site showed variations depending on the applied weighting scheme. Among the three tested cases, Case 3 predicted a more dynamic interaction between groundwater and streams, whereas Case 2 suggested a decreased contribution of groundwater to the watershed system. Despite the differences observed from the model predictions, however, all tested models exhibited similar seasonal variations in both water budget and groundwater–surface water interaction processes. Finally, we concluded that the application and comparison of different weighting schemes for the determination of error covariance could contribute to a better understanding of integrated water dynamics and to reducing the model uncertainty arising from the observation error.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/w15101880/s1>. Figure S1: Projected monthly variation of monthly rainfall under RCP2.6 and 8.5 scenarios during (a) the 2020s (b) the 2050s, and (c) the 2080s over the study area; Table S1: Summary of predicted climate change scenarios over the study site.

**Author Contributions:** Conceptualization, E.L. and H.-T.H.; methodology, E.L.; software, E.L., H.L. and D.P.; formal analysis, E.L. and H.L.; investigation, E.L., H.L. and D.P.; resources, E.L. and C.P.; data curation, E.L. and C.P.; writing—original draft preparation, E.L.; writing—review and editing, E.L., D.P. and H.-T.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Basic Research Project of Korea Institute of Geoscience and Mineral Resources (grant number: 23-3411).

**Data Availability Statement:** The data that support the findings of this study are available from <https://data.mendeley.com/datasets/8xyspzfwvp/1> (accessed on 13 May 2023).

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

## References

1. Taylor, R.G.; Scanlon, B.; Döll, P.; Rodell, M.; van Beek, R.; Wada, Y.; Longuevergne, L.; Leblanc, M.; Famiglietti, J.S.; Edmunds, M.; et al. Ground water and climate change. *Nat. Clim. Chang.* **2013**, *3*, 322–329. [[CrossRef](#)]
2. IPCC. *Climate Change 2013: The Physical Science Basis*; Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change; Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013.
3. Van Vuuren, D.P.; Edmonds, J.; Kainuma, M.; Riahi, K.; Thomson, A.; Hibbard, K.; Hurtt, G.C.; Kram, T.; Krey, V.; Lamarque, J.-F.; et al. The representative concentration pathways: An overview. *Clim. Chang.* **2011**, *109*, 5–31. [[CrossRef](#)]
4. Kløve, B.; Ala-Aho, P.; Bertrand, G.; Gurdak, J.J.; Kupfersberger, H.; Kværner, J.; Muotka, T.; Mykrä, H.; Preda, E.; Rossi, P.; et al. Climate change impacts on groundwater and dependent ecosystems. *J. Hydrol.* **2014**, *518*, 250–266. [[CrossRef](#)]
5. Scibek, J.; Allen, D.M. Modeled impacts of predicted climate change on recharge and groundwater levels. *Water Resour. Res.* **2006**, *42*, W11405. [[CrossRef](#)]

6. Ayugi, B.; Dike, V.; Ngoma, H.; Babaousmail, H.; Mumo, R.; Ongoma, V. Future Changes in Precipitation Extremes over East Africa Based on CMIP6 Models. *Water* **2021**, *13*, 2358. [[CrossRef](#)]
7. Peters, D.L.; Dibike, Y.B.; Shudian, J.; Monk, W.A.; Baird, D.J. Effects of Climate Change on Navigability Indicators of the Lower Athabasca River, Canada. *Water* **2023**, *15*, 1373. [[CrossRef](#)]
8. Majone, B.; Avesani, D.; Zulian, P.; Fiori, A.; Bellin, A. Analysis of high streamflow extremes in climate change studies: How do we calibrate hydrological models? *Hydrol. Earth Syst. Sci.* **2022**, *26*, 3863–3883. [[CrossRef](#)]
9. Miguez-Macho, G.; Fan, Y. The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *J. Geophys. Res. Atmos.* **2012**, *117*, D15113. [[CrossRef](#)]
10. Pokhrel, Y.N.; Fan, Y.; Miguez-Macho, G. Potential hydrologic changes in the Amazon by the end of the 21st century and the groundwater buffer. *Environ. Res. Lett.* **2014**, *9*, 084004. [[CrossRef](#)]
11. Stergiadi, M.; Di Marco, N.; Avesani, D.; Righetti, M.; Borga, M. Impact of Geology on Seasonal Hydrological Predictability in Alpine Regions by a Sensitivity Analysis Framework. *Water* **2020**, *12*, 2255. [[CrossRef](#)]
12. Tian, Y.; Zheng, Y.; Wu, B.; Wu, X.; Liu, J.; Zheng, C. Modeling surface water-groundwater interaction in arid and semi-arid regions with intensive agriculture. *Environ. Model. Softw.* **2015**, *63*, 170–184. [[CrossRef](#)]
13. Ala-Aho, P.; Rossi, P.M.; Isokangas, E.; Kløve, B. Fully integrated surface–subsurface flow modelling of groundwater–lake interaction in an esker aquifer: Model verification with stable isotopes and airborne thermal imaging. *J. Hydrol.* **2015**, *522*, 391–406. [[CrossRef](#)]
14. Brannen, R.; Spence, C.; Ireson, A. Influence of shallow groundwater–surface water interactions on the hydrological connectivity and water budget of a wetland complex. *Hydrol. Process.* **2015**, *29*, 3862–3877. [[CrossRef](#)]
15. Kalbus, E.; Reinstorf, F.; Schirmer, M. Measuring methods for groundwater–Surface water interactions: A review. *Hydrol. Earth Syst. Sci.* **2006**, *10*, 873–887. [[CrossRef](#)]
16. Kiel, B.A.; Cardenas, M.B. Lateral hyporheic exchange throughout the Mississippi River network. *Nat. Geosci.* **2014**, *7*, 413–417. [[CrossRef](#)]
17. Lautz, L.K.; Siegel, D.I. Modeling surface and ground water mixing in the hyporheic zone using MODFLOW and MT3D. *Adv. Water Resour.* **2006**, *29*, 1618–1633. [[CrossRef](#)]
18. Erler, A.R.; Frey, S.K.; Khader, O.; d’Orgeville, M.; Park, Y.J.; Hwang, H.T.; Lapen, D.R.; Peltier, W.R.; Sudicky, E.A. Simulating climate change impacts on surface water resources within a lake-affected region using regional climate projections. *Water Resour. Res.* **2019**, *55*, 130–155. [[CrossRef](#)]
19. Havril, T.; Tóth, Á.; Molson, J.W.; Galsa, A.; Mádl-Szőnyi, J. Impacts of predicted climate change on groundwater flow systems: Can wetlands disappear due to recharge reduction? *J. Hydrol.* **2018**, *563*, 1169–1180. [[CrossRef](#)]
20. Persaud, E.; Levison, J.; MacRitchie, S.; Berg, S.J.; Erler, A.R.; Parker, B.; Sudicky, E. Integrated modelling to assess climate change impacts on groundwater and surface water in the Great Lakes Basin using diverse climate forcing. *J. Hydrol.* **2020**, *584*, 124682. [[CrossRef](#)]
21. Schilling, O.S.; Cook, P.G.; Brunner, P. Beyond classical observations in hydrogeology: The advantages of including exchange flux, temperature, tracer concentration, residence time, and soil moisture observations in groundwater model calibration. *Rev. Geophys.* **2019**, *57*, 146–182. [[CrossRef](#)]
22. Doherty, J. *Calibration and Uncertainty Analysis for Complex Environmental Models*, 1st ed.; Watermark Numerical Computing: Brisbane, QLD, Australia, 2015.
23. Albers, B.M.C.; Molson, J.W.; Bense, V.F. Parameter sensitivity analysis of a two-dimensional cryo-hydrogeological numerical model of degrading permafrost near Umiujaq (Nunavik, Canada). *Hydrogeol. J.* **2020**, *28*, 905–919. [[CrossRef](#)]
24. Christensen, S.; Doherty, J. Predictive error dependencies when using pilot points and singular value decomposition in groundwater model calibration. *Adv. Water Resour.* **2008**, *31*, 674–700. [[CrossRef](#)]
25. Doherty, J.; Johnston, J.M. Methodologies for calibration and predictive analysis of a watershed model. *J. Am. Water Resour. Assoc.* **2003**, *39*, 251–265. [[CrossRef](#)]
26. Moore, C.; Doherty, J. The role of the calibration process in reducing model predictive error. *Water Resour. Res.* **2005**, *41*, W05020. [[CrossRef](#)]
27. Refsgaard, J.C.; van der Sluijs, J.P.; Brown, J.; van der Keur, P. A framework for dealing with uncertainty due to model structure error. *Adv. Water Resour.* **2006**, *29*, 1586–1597. [[CrossRef](#)]
28. James, S.C.; Doherty, J.E.; Eddebarh, A.A. Practical postcalibration uncertainty analysis: Yucca Mountain, Nevada. *Ground Water* **2009**, *47*, 851–869. [[CrossRef](#)]
29. Liu, X.; Kitanidis, P.K. Large-scale inverse modeling with an application in hydraulic tomography. *Water Resour. Res.* **2011**, *47*, W02501. [[CrossRef](#)]
30. Tiedeman, C.R.; Green, C.T. Effect of correlated observation error on parameters, predictions, and uncertainty. *Water Resour. Res.* **2013**, *49*, 6339–6355. [[CrossRef](#)]
31. Hill, M.C.; Tiedeman, C.R. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*; Wiley: Hoboken, NJ, USA, 2007.
32. Gelhar, L.W. *Stochastic Subsurface Hydrology*; Prentice Hall: Englewood Cliffs, NJ, USA, 1993.
33. McMillan, H.; Freer, J.; Pappenberger, F.; Krueger, T.; Clark, M. Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. *Hydrol. Process.* **2010**, *24*, 1270–1284. [[CrossRef](#)]

34. Korea Meteorological Administration. Available online: <http://www.weather.go.kr/> (accessed on 3 April 2023).
35. Geological Map of Korea. Available online: <http://mgeo.kigam.re.kr> (accessed on 3 April 2023).
36. Korean Soil Information System. Available online: <http://www.soil.rda.go.kr/> (accessed on 3 April 2023).
37. Environmental Geographic Information Service. Available online: <https://egis.me.go.kr/> (accessed on 3 April 2023).
38. Groundwater Information Monitoring System. Available online: <http://www.gims.go.kr/> (accessed on 3 April 2023).
39. Aquanty. *HydroGeoSphere User Manual*; Aquanty Inc.: Waterloo, ON, Canada, 2015.
40. Viessman, W.J.; Lewis, G.L. *Introduction to Hydrology*, 4th ed.; Harper Collins College Publisher: New York, NY, USA, 1996.
41. HydroAlgorithmics Pty Ltd. *AlgoMesh 2 User Guide*; HydroAlgorithmics Inc.: Melbourne, VIC, Australia, 2020.
42. Environmental Systems Research Institute. *ArcGIS Desktop: Release 10.6.1*; ESRI, Inc.: Redlands, CA, USA, 2018.
43. Hwang, H.T.; Park, Y.J.; Frey, S.K.; Callaghan, M.V.; Berg, S.J.; Lapen, D.R.; Sudicky, E.A. Efficient numerical incorporation of water management operations in integrated hydrosystem models: Application to tile drainage and reservoir operating systems. *J. Hydrol.* **2019**, *575*, 1253–1266. [[CrossRef](#)]
44. Kristensen, K.J.; Jensen, S.E. A model for estimating actual evapotranspiration from potential evapotranspiration. *Hydrol. Res.* **1975**, *6*, 170–188. [[CrossRef](#)]
45. Wigmosta, M.S.; Vail, L.W.; Lettenmaier, D.P. A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.* **1994**, *30*, 1665–1679. [[CrossRef](#)]
46. Valiantzas, J.D. Simplified versions for the Penman evaporation equation using routine weather data. *J. Hydrol.* **2006**, *331*, 690–702. [[CrossRef](#)]
47. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*; FAO Irrigation and Drainage Paper 56; FAO: Rome, Italy, 1998.
48. MCD15A3H MODIS/terra+aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006 [Data Set]. NASA EOSDIS Land Processes DAAC. Available online: <https://doi.org/10.5067/MODIS/MCD15A3H.006> (accessed on 3 April 2023).
49. Water Resources Management Information System (WAMIS). Available online: <http://www.wamis.go.kr> (accessed on 3 April 2023).
50. Rural Groundwater Net. Available online: <http://www.groundwater.or.kr> (accessed on 3 April 2023).
51. Panday, S.; Huyakorn, P.S. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Adv. Water Resour.* **2004**, *27*, 361–382. [[CrossRef](#)]
52. Canadell, J.; Jackson, R.B.; Ehleringer, J.B.; Mooney, H.A.; Sala, O.E.; Schulze, E.D. Maximum rooting depth of vegetation types at the global scale. *Oecologia* **1996**, *108*, 583–595. [[CrossRef](#)]
53. Ok, J.; Kim, D.J.; Han, K.; Jung, K.H.; Lee, K.D.; Zhang, Y.; Cho, H.-R.; Hwang, S.-A. Relationship between measured and predicted soil water content using soil moisture monitoring network (in Korean). *Korean J. Agric. For. Meteorol.* **2019**, *21*, 297–306.
54. Chow, V.T. *Open-Channel Hydraulics*; McGraw-Hill: New York, NY, USA, 1959.
55. Ebel, B.A.; Mirus, B.B.; Heppner, C.S.; VanderKwaak, J.E.; Loague, K. First-order exchange coefficient coupling for simulating surface water-groundwater interactions: Parameter sensitivity and consistency with a physics-based approach. *Hydrol. Process.* **2009**, *23*, 1949–1959. [[CrossRef](#)]
56. Freeze, R.A.; Cherry, J.A. *Groundwater*; Prentice Hall: Englewood Cliffs, NJ, USA, 1979.
57. Fetter, C.W.; Boving, T.; Kremer, D. *Contaminant Hydrogeology*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 1999.
58. Namown Eng. Inc. *Measurements and Computation of Streamflow of Sabgyo Watershed—Gangcheong, Wonypyeong, Hannaedari*; Geum River Flood Control Office: Gongju, Republic of Korea, 2005. (In Korean)
59. Namown Eng. Inc.; Korea Institute of Construction Technology. *Measurements and Computation of Streamflow of Sabgyo Watershed*; Geum River Flood Control Office: Gongju, Republic of Korea, 2006. (In Korean)
60. Perin, R.; Trigatti, M.; Nicolini, M.; Campolo, M.; Goi, D. Automated calibration of the EPA-SWMM model for a small suburban catchment using PEST: A case study. *Environ. Model. Assess.* **2020**, *192*, 374. [[CrossRef](#)]
61. Kim, S.M.; Benham, B.L.; Brannan, K.M.; Zeckoski, R.W.; Doherty, J. Comparison of hydrologic calibration of HSPF using automatic and manual methods. *Water Resour. Res.* **2007**, *43*, W01402. [[CrossRef](#)]
62. Linsley, R.K.; Kohler, M.A. *Hydrology for Engineers*; McGraw-Hill: London, UK, 1988.
63. Thyer, M.; Renard, B.; Kavetski, D.; Kuczera, G.; Franks, S.W.; Srikanthan, S. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resour. Res.* **2009**, *45*, W00B14. [[CrossRef](#)]
64. Park, E. A geostatistical evolution strategy for subsurface characterization: Theory and validation through hypothetical two-dimensional hydraulic conductivity fields. *Water Resour. Res.* **2020**, *56*, W026922. [[CrossRef](#)]
65. Doherty, J.E.; Hunt, R.J. *Approaches to Highly Parameterized Inversion: A Guide to Using PEST for Groundwater-Model Calibration*; US Geological Survey Scientific Investigations Report 2010-5169; US Geological Survey: Reston, VA, USA, 2010.

**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.