An Efficient Estimation of Time-Varying Parameters of Dynamic Models by Combining Offline Batch Optimization and Online Data Assimilation

Yohei Sawada\textsuperscript{1} \( ^\ddagger \)

\textsuperscript{1}Institute of Engineering Innovation, The University of Tokyo, Tokyo, Japan

Abstract It is crucially important to estimate unknown parameters in process-based models by integrating observation and numerical simulation. For many applications in earth system sciences, a parameter estimation method which allows parameters to temporally change is required. In the present paper, an efficient and practical method to estimate time-varying parameters of relatively low dimensional models is presented. In the newly proposed method, called Hybrid Offline Online Parameter Estimation with Particle Filtering (HOOPE-PF), an inflation method to maintain the spread of ensemble members in a Sampling-Importance-Resampling Particle Filter (SIRPF) is improved using a non-parametric posterior probabilistic distribution of time-invariant parameters obtained by comparing simulated and observed climatology. HOOPE-PF outperforms the original SIRPF in synthetic experiments with toy models and a real-data experiment with a conceptual hydrological model when an ensemble size is small. The advantage of HOOPE-PF is that its performance is not greatly affected by the size of perturbation to be added to ensemble members to maintain their spread while it is important to get the optimal performance in the original particle filter. Since HOOPE-PF is the extension of the existing particle filter which has been extensively applied to many models in earth system sciences such as land, ecosystem, hydrology, and paleoclimate reconstruction, HOOPE-PF can be applied to improve the simulation of these process-based models by considering time-varying model parameters.

Plain Language Summary Computer simulation is widely used to understand earth systems such as atmosphere, ocean, land, ecosystem, and society. Many computer simulation models in earth system sciences inevitably have unknown coefficients of equations, which are called model parameters. It is important for accurate simulation of earth systems to tune these model parameters by comparing the results of computer simulation with observation. Although many previous works assumed that model parameters do not change over time, it is necessary to allow them to change over time in some applications, which makes it more difficult to estimate model parameters. In this study, a new method to estimate time-varying model parameters is proposed. The key idea of this paper is to combine two existing methods which integrate observations into computer simulation models. It is found that the proposed method works better than the existing method in three case studies. The proposed method can contribute to improving many simulations in earth system sciences by efficiently tuning time-varying model parameters.

1. Introduction

In Earth system sciences, numerical models are important tools to understand, monitor, and predict earth systems. Many numerical models in earth system sciences have many unknown parameters which cannot be directly determined by current theory and cannot be directly measured. This uncertainty in model parameters substantially affects the skill of numerical models to simulate real phenomena. It is a grand challenge to infer model parameters by integrating observations and models. Although model parameters are often assumed to be time-invariant, there are time-varying parameters in practical applications due to incomplete parameterizations and unconsidered dynamics that control parameters (e.g., Reichert et al., 2021). It is beneficial to develop an efficient and practical parameter estimation method which allows parameters to temporally change.

To estimate time-varying parameters, online parameter estimation by sequential data assimilation has been recognized as a useful method. Since ensemble data assimilation such as Ensemble Kalman filter (EnKF) and Particle Filter (PF) can sequentially adjust model parameters using real-time observations, it can be easily applied to the estimation of time-varying parameters. Ruiz et al. (2013) applied the Local Ensemble Transform Kalman Filter (LETKF; Hunt et al., 2007) to jointly estimate state variables and parameters in a simple low-resolution General...
Circulation Model (GCM). They successfully estimated time-varying parameters by sequentially adjusting them in an idealized experiment. Pathiraja et al. (2018) successfully detected a land use change in a river basin by applying the locally linear dual EnKF proposed by Pathiraja et al. (2016) to estimate a temporal change in parameters of hydrological models. Sawada and Hanazaki (2020) applied PF to a socio-hydrological model proposed by Di Baldassarre et al. (2013), in which flood-human interactions are simulated. They successfully estimated time-varying parameters driven by unknown social dynamics in an idealized experiment.

Whenever parameters are adjusted by ensemble data assimilation, the estimated ensemble variance of updated parameters is always smaller than that of background. If ensemble variances are too small compared to observation error variances, the information of observation cannot impact the state and parameter estimation. To avoid this filter degeneracy (Anderson, 2007), it is important to maintain the appropriate ensemble variance of updated parameters. Many parameter perturbation methods or inflation methods have been proposed to maintain ensemble variance of estimated parameters. Moradkhani, Hsu, et al. (2005) simply added Gaussian noise, whose variance is proportional to ensemble variance of background, to parameters of each ensemble member after adjusting them by the sampling-importance-resampling PF. Moradkhani, Sorooshian, et al. (2005) used the kernel smoother with location shrinkage (Liu & West, 2001) to perturb updated parameters and maintain the ensemble variance of parameters in their dual EnKF. Similar strategies to Moradkhani, Hsu, et al. (2005) and Moradkhani, Sorooshian, et al. (2005) have been used in many previous works (e.g., Ait-El-Fquih & Hoteit, 2017; Pathiraja et al., 2016, 2018; Sawada & Hanazaki, 2020; Yan & Moradkhani, 2016). It has been shown effective to evaluate perturbed parameters by running a numerical model again and iteratively sampling new perturbed parameters (e.g., Abbaszadeh et al., 2018; Moradkhani et al., 2012; Vrugt et al., 2013), although the computational cost is increased by this strategy. Ruiz et al. (2013) discussed how to objectively determine the appropriate variance of updated parameters using the ensemble variance of background and updated state variables. Kotsuki et al. (2018) used the relaxation to prior spread method (Whitaker & Hamill, 2012) to constrain the ensemble variance of parameters to initial prescribed variance in their LETKF framework.

As an alternative approach of parameter estimation, offline batch optimization, in which model parameters are optimized by minimizing a cost function based on the difference between simulation and observation in a long-term (e.g., longer than a year in hydrology) period, has been used. The advantage of the offline batch optimization is that non-parametric Bayesian inference can easily be applied in this framework, so that the reliable uncertainty estimation of model parameters can be obtained. Since it is necessary to iteratively perform the long-term integration of a numerical model with different parameters, the direct application of the offline batch optimization to models in earth system sciences is computationally expensive (e.g., Laloy & Vrugt, 2012). However, the computational cost of the offline batch optimization can be dramatically reduced by replacing the responses of numerical models with machine learning models called surrogate models. For example, Sawada (2020) used the Gaussian process regression (Rasmussen & Williams, 2006) to mimic the relationship between parameters of a land surface model and simulated microwave brightness temperature which is sensitive to soil moisture and vegetation water content. By evaluating parameters based on the computationally cheap statistical surrogate model, the Markov Chain Monte Carlo (MCMC) sampler (Hastings, 1970) could be efficiently applied to optimize soil and vegetation parameters of a land surface model. Zhang et al. (2020) proposed a method to reduce the approximation error of a surrogate model by adding samples based on the evaluation of the surrogate model. Their method was successfully applied to the parameter inference of a groundwater model. Teixeira Parente et al. (2019) applied the active subspace method (Constantine et al., 2014) to efficiently construct a surrogate model and calibrate parameters of a hydrological model. Dunbar et al. (2021) successfully performed the surrogate model based MCMC and realized the uncertainty quantification of convective parameters in an idealized GCM with only $O(10^5)$ GCM runs in parallel to construct a surrogate model. It is relatively cheap considering that the direct application of MCMC requires $O(10^6)$ GCM runs which cannot be efficiently parallelized and should be done iteratively.

Although several studies proposed combining data assimilation and offline batch optimization (e.g., Cleary et al., 2021; Lunderman et al., 2021; Tomizawa & Sawada, 2021), there is no contribution to the estimation of time-varying parameters by combining them. The aim of this study is to develop an efficient and practical model parameter optimization method which can allow parameters to temporally vary by combining online PF, which has been used in earth system sciences (e.g., Abolafia-Rosenzweig et al., 2019; Mairesse et al., 2013; Qin et al., 2009; Sawada et al., 2015), with offline batch optimization. The proposed method substantially reduces the sensitivity
of the PF’s performance to the choice of hyperparameters within a parameter perturbation method. Therefore, the proposed method contributes to solving the issues in the maintenance of ensemble variance discussed above. This feature will be demonstrated in idealized experiments with the Lorenz 63 model (Lorenz, 1963) and a real-data experiment with the HYMOD conceptual hydrological model (Moore, 2007).

2. Method
2.1. Online Data Assimilation: Particle Filtering

In this study, the Sampling-Importance-Resampling Particle Filter (SIRPF) algorithm was used as a data assimilation method. The implementation of Moradkhani, Hsu, et al. (2005) was mostly used in this study. While Moradkhani, Hsu, et al. (2005) performed resampling only when the effective sample size becomes below a prescribed threshold, this method was not adopted in this study. A discrete state-space dynamics system is defined as follows:

\[ x(t) = f(x(t-1), \theta(t-1), u(t-1)) + q(t-1) \]  

where \( x(t) \) is the state variables, \( \theta(t) \) is the model parameters, \( u(t) \) is the external forcing, and \( q(t) \) is the noise process which represents the model error at time \( t \). \( f(\cdot) \) is the dynamic model. Note that the model parameters are the function of \( t \), so that it is assumed that time-varying parameters exist. Since observable variables are not identical to the state variables of the dynamic model in general cases, data assimilation needs to formulate an observation equation:

\[ y^f(t) = h(x(t)) + r(t) \]  

where \( y^f(t) \) is simulated observations at time \( t \), \( r(t) \) is the noise process which represents the observation error at time \( t \). \( h(\cdot) \) is the observation operator. The purpose of PF (and any other sequential data assimilation methods) is to obtain the posterior distribution of the state variables and parameters based on the Bayesian update:

\[ p(x(t), \theta(t)|y^f(1:t)) \propto p(y^f(t)|x(t), \theta(t))p(x(t), \theta(t)|y^f(1:t-1)) \]  

where \( p(x(t), \theta(t)|y^f(1:t)) \) is the posterior probability of the state variables \( x(t) \) and parameters \( \theta(t) \) given all observation up to time \( t \), \( y^f(1:t) \). The prior knowledge, \( p(x(t), \theta(t)|y^f(1:t-1)) \), is updated using the likelihood, \( p(y^f(t)|x(t), \theta(t)) \) which includes the information of the new observation at time \( t \). PF directly approximates Equation 3 by Monte Carlo simulation.

The implementation of SIRPF in this study is as follows:

1. Forecasting model state variables and parameters from time \( t-1 \) to \( t \) using ensemble simulation.

\[ x_i^{f}(t) = f(x_i^{f}(t-1), \theta_i^{f}(t-1), u(t-1)) \]  
\[ \theta_i^{f}(t) = \theta_i^{f}(t-1) \]  

where superscripts \( f \) and \( a \) show forecasted and updated variables, respectively. Subscript \( i \) shows the ensemble number. Equation 5 implies that the model parameters do not change in the forecast step.

2. Calculating model simulation projected onto an observation space for all ensemble members:

\[ y_i^{f}(t) = h(x_i^{f}(t)) \]  

3. Calculating weights for all ensembles:

\[ w_i = \frac{p(y^f(t)|x_i^{f}(t), \theta_i^{f}(t))}{\sum_{i=1}^{N} p(y^f(t)|x_i^{f}(t), \theta_i^{f}(t))} \]  

where \( w_i \) is the normalized weight for the \( i \)th ensemble member, and \( N \) is the ensemble size. \( p(y^f(t)|x_i^{f}(t), \theta_i^{f}(t)) \) is the likelihood function. In this study, it is assumed that the observation error follows the Gaussian distribution, so that \( p(y^f(t)|x_i^{f}(t), \theta_i^{f}(t)) \) can be formulated as follows:
Resampling ensemble members based on normalized weights. The normalized weight $w_i$ is recognized as the probability that the ensemble $i$ is selected after resampling. Multinomial resampling (Gordon et al., 1993) was used as a resampling method. The resampled state variables and parameters are defined as $x_i^{\text{resamp}}$ and $\theta_i^{\text{resamp}}$, respectively.

5. Adding a perturbation to state variables and parameters to maintain appropriate ensemble variance:

$$x_i = x_i^{\text{resamp}} + \epsilon_{i, \text{state}}$$

$$\theta_i = \theta_i^{\text{resamp}} + \epsilon_{i, \text{para}}$$

where $N(.)$ is the Gaussian distribution, $\text{Var}^s$ and $\text{Var}^\theta$ are the ensemble variances of the state variables $x_i(t)$ and parameters $\theta_i(t)$ in the background field, respectively.

$S_{\text{state}}$ and $S_{\text{para}}$ are the “inflation” hyperparameters. The performance of the original PF is sensitive to them. Several approaches to determine them objectively and adaptively have been proposed (e.g., Leisenring & Moradkhani, 2012; Sawada & Hanazaki, 2020). In this study, these adaptive approaches were not used and these hyperparameters were fixed. Note that these adaptive approaches cannot eliminate hyperparameters in the filter since other hyperparameters appear in the adaptive method. The experiments of this study will show that the performance of a filter is less sensitive to these hyperparameters in the proposed method than the original SIRPF.

### 2.2. Offline Batch Optimization

The aim of the offline batch optimization used in this paper is to get a probabilistic distribution to check the quality of perturbed parameters. Note that the offline batch optimization of this study is not used to get the best set of parameters (see also Section 2.3). The MCMC sampling is used to sample model parameters which can accurately simulate observed climatology. In the offline batch optimization of this paper, parameters are assumed to be time-invariant. A discrete state-space dynamics system is re-written as:

$$x(t) = f \left(x(t-1), \hat{\theta}, u(t-1)\right) + q(t-1)$$

$$y^f(t) = h(x(t)) + r(t)$$

where $\hat{\theta}$ is the time-invariant model parameter. In this study, some climatological indices were calculated based on the long-term averages of the timeseries of the simulated observation $y^f(t)$ and the real observation $y(t)$. Since these climatological indices are designed to be independent to the initial condition, the problem of the offline batch optimization can be simplified as:

$$y^f = g \left(\hat{\theta}\right)$$

where $y^f$ is the simulated climatological index. $g(.)$ is the forward map. The purpose of the offline batch optimization in this study is to obtain the posterior probabilistic distribution of $\hat{\theta}$, which is defined as $Q \left(\hat{\theta}\right)$, based on the simulated climatological index $y^f$ and the observed climatological index $y^o$.

It is generally difficult to perform the Bayesian inference using Equation 15 since $g(.)$ is computationally expensive. In this study, the surrogate model of $g(.)$ was developed by the Gaussian process regression (Rasmussen & Williams, 2006). First, 500 ensemble members of model parameters were generated by the pseudo-Monte Carlo sampling which uniformly draws samples from the multi-dimensional parameter space. Second, the long-term integration of the dynamic models (i.e., the Lorenz 63 and HYMOD models) with the 500 parameter ensemble
members was performed in parallel. Third, the statistical surrogate model was constructed from the data of the 500 \( \hat{\theta} - \gamma' \) combinations by the Gaussian process regression. The surrogate model replaces Equation 15 with:

\[
\gamma' = g^{(i)}(\hat{\theta})
\]  

(16)

where \( g^{(i)}(.) \) is the computationally cheap surrogate model of \( g(.) \). While the long-term integration of a full dynamic model with input data such as initial and boundary conditions is necessary to obtain its climatology and evaluate \( g(\hat{\theta}) \), the trained surrogate model \( g^{(i)}(\hat{\theta}) \) requires only model parameters and directly estimates the climatology without performing the long-term integration. Therefore, the evaluation of Equation 16 is much cheaper and less complex than that of Equation 15 if the surrogate model can accurately mimic the responses of a dynamic model to parameters. The amount of the training data (i.e., ensemble size) affects the skill of the Gaussian process regression to mimic the responses of a dynamic model to its parameters. The skill of the Gaussian process regression to simulate \( \gamma' \) was evaluated by the independent test set of 1,000 simulation of dynamic models. In all applications of this paper (see also Section 3), the correlation coefficient between predicted and true \( \gamma' \) is more than 0.95, so that it is confirmed that the 500-ensemble size is enough in this study.

The Metropolis-Hastings algorithm (Hastings, 1970) was used as a MCMC sampler to draw samples from the probabilistic distribution \( Q(\hat{\theta}) \). The implementation of the MCMC sampler is the following:

1. For each iteration \( j \), generate a candidate parameter vector \( \tilde{\theta}_c \).
2. This candidate is sampled from the proposal distribution \( q(\hat{\theta},|\tilde{\theta}_c) \).
3. Calculating an acceptance probability of \( \alpha(\tilde{\theta}_c,\hat{\theta}_j) \):

\[
\alpha(\tilde{\theta}_c,\hat{\theta}_j) = \exp\left(\Phi(\tilde{\theta}_c,\hat{\theta}_j) - \Phi(\tilde{\theta}_j,\hat{\theta}_c)\right)
\]

(17)

\[
\Phi(\tilde{\theta}_c,\hat{\theta}_j) = \frac{1}{2}\|\gamma - g^{(i)}(\hat{\theta})\|^2_{R_g(\hat{\theta}),R_e}
\]

(18)

where \( R_g(\hat{\theta}) \) and \( R_e \) are the error variances of \( g^{(i)}(\hat{\theta}) \) and observation, respectively (see also Cleary et al., 2021). The acceptance probability is calculated based on the square difference between the simulated and observed climatological indices normalized by the total error variance of the Gaussian process and the observation measurement.

3. Determine if \( \tilde{\theta}_c \) is accepted as a new parameter or not.

A random number, \( b \), is generated from the uniform distribution of \([0,1]\). Then,

- If \( b \leq \alpha(\tilde{\theta}_c,\hat{\theta}_j) \), accept the candidate parameter and \( \hat{\theta}_{j+1} = \tilde{\theta}_c \).
- If \( b > \alpha(\tilde{\theta}_c,\hat{\theta}_j) \), reject the candidate parameter and \( \hat{\theta}_{j+1} = \hat{\theta}_j \).

These three steps were iterated 500,000 times in this study and the first 100,000 iterations were discarded as the spin-up period. From the remaining 400,000 samples, the probabilistic distribution of parameters \( Q(\hat{\theta}) \) was obtained. In this study, \( q(\hat{\theta},|\tilde{\theta}_c) \) is assumed to be Gaussian with zero mean. Since the Gaussian process can provide both mean and variance of their estimation in each point, it is straightforward to obtain \( R_g(\hat{\theta}) \) whenever \( g^{(i)}(\hat{\theta}) \) is evaluated. To calculate \( R_e \), the subset of the continuous timeseries of observations was randomly chosen and 1,000 observed climatological indices \( \gamma^o \) were generated. The variance of the observed climatological index \( R_e \) was calculated as the variance of these 1,000 \( \gamma^o \). In addition, \( \gamma^o \) was randomly chosen from these subsets of the observed climatological indices in every 100 iterations of MCMC. This procedure was found to be necessary to obtain the posterior of the time-invariant \( \theta \) when parameters actually change over time and make observed variables non-stationary. If this process is neglected and \( R_e \) is set to zero, the posterior distribution of \( \theta \) becomes underdispersive although the relatively large uncertainty of \( \theta \) needs to be obtained since the actual parameters temporally change.

### 2.3. Hybrid Offline Online Parameter Estimation With Particle Filtering (HOOPE-PF)

The schematic of the proposed method, Hybrid Offline Online Parameter Estimation with Particle Filtering (HOOPE-PF), is shown in Figure 1. The fundamental idea is that the posterior of PF, \( p(\theta(t),\theta(0)|y(1:n)) \) (see Section 2.1), is constrained by the probabilistic distribution of parameters from the offline batch optimization, \( Q(\hat{\theta}) \) (see Section 2.2) in HOOPE-PF. The parameter perturbation method was slightly changed to include the information of \( Q(\hat{\theta}) \) into the process of the parameter perturbation after resampling. The implementation of
5. Perturbing state variables and parameters considering

Steps 1–4 are identical to Steps 1–4 of the original SIRPF (see Section 2.1).

5a. Adding a perturbation to state variables and parameters to maintain appropriate ensemble variances:

\[
x'_i = x'_i^{\text{resamp}} + \epsilon_{i,\text{state}}
\]

\[
\epsilon_{i,\text{state}} \sim N(0, S_{\text{state}} \times \text{Var}^*)
\]

\[
\theta'_i = \theta'_i^{\text{resamp}} + \epsilon_{i,\text{para}}
\]

\[
\epsilon_{i,\text{para}} \sim N(0, S_{\text{para}} \times \text{Var}^p)
\]

where \( N(\cdot) \) is the Gaussian distribution, \( \text{Var}^* \) and \( \text{Var}^p \) are the ensemble variances of the state variables \( x'_i(t) \) and parameters \( \theta'_i(t) \) in the background field, respectively.

5b. Calculating an acceptance probability of \( \theta'_i \), \( a \left( \theta'_i^{\text{resamp}}, \theta'_i \right) \) by evaluating how \( \theta'_i \) corresponds to the results of the offline batch optimization, \( Q(\hat{\theta}) \):

\[
a \left( \theta'_i^{\text{resamp}}, \theta'_i \right) = \frac{Q(\hat{\theta})}{Q(\theta'_i^{\text{resamp}})}
\]

(19)

Here \( Q(\cdot) \) is used as a function to evaluate how a parameter vector fits the posterior distribution of \( \hat{\theta} \) obtained by the offline batch optimization. Equation 19 compares the fitness of the perturbed parameter \( \theta'_i \) to the posterior distribution of \( \hat{\theta} \) with that of the original parameter \( \theta'_i^{\text{resamp}} \).

5c. Determine if \( \theta'_i \) is accepted as a new parameter or not.

A random number, \( b \), is generated from the uniform distribution of [0,1]. Then,

If \( b \leq a \left( \theta'_i^{\text{resamp}}, \theta'_i \right) \), accept the updated parameter \( \theta'_i \).

If \( b > a \left( \theta'_i^{\text{resamp}}, \theta'_i \right) \), reject the current updated parameter and go back to the step 5a only for the parameters.

The new steps 5b and 5c can prevent from sampling parameters which cannot support the results of the offline batch optimization. If the fitness of the perturbed parameters to the results of the offline batch optimization, \( Q(\theta'_i) \), is substantially lower than that of the original resampled parameter \( Q(\theta'_i^{\text{resamp}}) \), this perturbed parameter is unlikely to be chosen. In the original SIRPF, large \( S_{\text{para}} \) moves particles far away from the resampled ones in any directions, which makes the filter unstable. Although large \( S_{\text{para}} \) of HOOPE-PF also moves particles far away from the resampled ones, it tends to move particles in the direction which increases \( Q(\cdot) \) since the perturbation which reduces \( Q(\cdot) \) tends to be rejected in the step 5c. The numerical experiments in the following sections will reveal that this new perturbation method substantially stabilizes PF.

Note that this new perturbation method can be used for any kinds of \( Q(\cdot) \), and \( Q(\cdot) \) is not necessarily obtained by the methods described in Section 2.2. It is possible to obtain \( Q(\cdot) \) by computationally cheaper methods such as assuming the Gaussian distribution. This
study proposes that $Q(t)$ is obtained by the fully nonparametric MCMC method to maximize the potential of PF to deal with non-Gaussian probabilistic distribution.

3. Experiment Design

3.1. Case Study I: The Lorenz 63 Model With an Abruptly Changed Parameter

HOOPE-PF is tested with the Lorenz 63 model (Lorenz, 1963):

$$\frac{dx}{dt} = -10(x - y)$$

(20)

$$\frac{dy}{dt} = -xz + \rho(t)x - y$$

(21)

$$\frac{dz}{dt} = xy - bz$$

(22)

$\rho(t)$ and $b$ are unknown parameters to be estimated. The true $b$ is assumed to be 8/3. Although the parameter in Equation 21 is time-invariant in the original Lorenz 63 model, here it is assumed that $\rho(t)$ was a time-varying parameter:

$$\rho(t) = \begin{cases} 
28 & (0 \leq t < 8000, 16000 \leq t < 24000) \\
24 & (8000 \leq t < 16000, 24000 \leq t < 32000) 
\end{cases}$$

(23)

In this case study, $\rho(t)$ abruptly changes every 8,000 timesteps. The size of the timestep was set to 0.01 and Equations 20–22 were numerically solved by the fourth-order Runge-Kutta scheme.

The two state variables, $y$ and $z$, are observed every 20 timesteps and the observation error was set to 1.0. The climatological index $\mathbf{y} = \begin{bmatrix} y^2, z^2 \end{bmatrix}$ was calculated by averaging simulated and observed variables during 4,000 timesteps. The 500 ensemble runs with different (time-invariant) parameters were performed for 4,000 timesteps to construct the surrogate model ($g^{(x)}(\hat{\theta})$). In the MCMC sampler, the observations up to $t = 32,000$ were used to generate the subsets of the observation timeseries (see Section 2.2). The period of each subset corresponded to 4,000 timesteps of the model. The initial conditions of state variables were generated by adding the Gaussian white noise whose mean and standard deviation are 0 and 1, respectively, to the true values. The initial conditions of $\rho(t)$ were drawn from the uniform distribution from 10 to 40. The initial conditions of $b$ were drawn from the uniform distribution from 0 to 15. The model tends to diverge when parameters outside this range are chosen, so that this prescribed range is reasonably wide. The performance of HOOPE-PF was compared with that of the original PF. The other settings of the hyperparameters and the ensemble size can be found in Table 1. The ensemble median was used as a representative value of the estimated probabilistic distribution to evaluate the filters because the median is often consistent to the mode of the probabilistic distribution in this experiment.

| Name          | Model                                      | Ensemble size | $S_{state}$                  | $S_{para}$                  |
|---------------|--------------------------------------------|---------------|------------------------------|------------------------------|
| Case Study I  | Lorenz63 with a time-varying parameter     | [30,100,250]  | [0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3, 0.325, 0.35, 0.375] | [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] |
| Case Study II | Lorenz63 with a time-varying parameter     | [30,100,250]  | [0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3, 0.325, 0.35, 0.375] | [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] |
| Case Study III| HYMOD                                      | [30,100,250]  | [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01] | [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] |

Table 1
Summary and Detailed Settings of the Case Studies in This Paper

| Name          | Model                                      | Ensemble size | $S_{state}$                  | $S_{para}$                  |
|---------------|--------------------------------------------|---------------|------------------------------|------------------------------|
| Case Study I  | Lorenz63 with a time-varying parameter     | [30,100,250]  | [0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3, 0.325, 0.35, 0.375] | [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] |
| Case Study II | Lorenz63 with a time-varying parameter     | [30,100,250]  | [0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3, 0.325, 0.35, 0.375] | [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] |
| Case Study III| HYMOD                                      | [30,100,250]  | [0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01] | [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0] |
3.2. Case Study II: The Lorenz 63 Model With Non-Periodic Forcing

In this second case study, the Lorenz 63 model described in Equations 20–22 was also used. However, the time-varying parameter was defined as:

$$\rho(t) = \rho_0 + \psi(t)$$  

$$\psi(t) = A \left( \frac{1}{3} \sin (2\pi f t) + \frac{1}{3} \sin \left( 2\pi \frac{\sqrt{3}}{f} t \right) + \frac{1}{3} \sin \left( 2\pi \frac{\sqrt{17}}{f} t \right) \right)$$

where $\rho_0 = 28$, $A = 5$, and $f = 1/20$. This version of the Lorenz 63 model has been intensively analyzed by Daron and Stainforth (2015). The settings of observation and a climatological index are the same as the case study I. The performance of HOOPE-PF was compared with that of the original PF. The other settings of the hyperparameters and the ensemble size can be found in Table 1.

3.3. Case Study III: Real-Data Experiment With the Conceptual Hydrological Model

While the previous two case studies are the synthetic experiments, here the parameters of the conceptual hydrological model, HYMOD (Moore, 2007), were estimated by assimilating real runoff observation. HYMOD has been widely used for studies on parameter optimization (e.g., Moradkhani, Hsu, et al., 2005; Pathiraja et al., 2018; Vrugt et al., 2013). HYMOD consists of a nonlinear rainfall excess model connected to four linear routing reservoirs (three quick-flow reservoirs and a slow-flow reservoir). There are five state variables and five parameters in HYMOD. Table 2 shows the description and ranges of these five parameters.

The study area is the Leaf River basin, Mississippi, the United States. The Leaf River basin has been used to test data assimilation and parameter optimization methods by previous works (e.g., Moradkhani, Hsu, et al., 2005; Sheikholeslami & Razavi, 2020; Vrugt et al., 2013). The data of precipitation, potential evapotranspiration, and runoff were extracted from the Catchment Attributes and Meteorology for Large Sample studies data set (CAMELS; Addor et al., 2017).

The climatological indices were set to the 5-year runoff ratio and the 5-year baseflow index. The runoff ratio is defined as the ratio of 5-year total runoff to 5-year total precipitation. The baseflow index is defined as the ratio of 5-year total baseflow to 5-year total runoff. Baseflow separation was performed by the one-parameter Lyne-Hollick digital filter (Lyne & Hollick, 1979). Following Gnann et al. (2019), the Lyne-Hollick digital filter was applied forward, backward, and forward again using a filter parameter of 0.925.

The 500 ensemble runs were performed for 5 years (from 1981 to 1985) to construct the surrogate model. This surrogate model directly estimates runoff ratio and baseflow index from the 5 parameters of HYMOD. In the MCMC sampler, observations from 1981 to 2000 were used to generate the subsets of the observation timeseries (see Section 2.2) and the period of each subset was set to 5-year. Although the HOOPE-PF was performed from 1981 to 2014, the results from 1981 to 2000 were not used for evaluation. The performance of HOOPE-PF was evaluated from 2001 to 2014, so that the observation data used for the offline batch optimization were not used for the evaluation of HOOPE-PF. The performance of the original PF was also evaluated from 2001 to 2014.

When the state variables are updated, the logarithm of them was taken to prevent from obtaining negative values of the state variables. The error variance of observed runoff was set to 10% of the observed value following Pathiraja et al. (2018). In addition, the minimum observation error variance was set to 0.1 since the filter degeneracy
was induced by too small observation errors. The state variables and parameters were updated by assimilating the daily runoff observation every day. The performance of HOOPE-PF was compared with the original PF. The other settings of the hyperparameters and the ensemble size can be found in Table 1.

The initial conditions of the state variables for each ensemble member were set to 0. The initial conditions of parameters for each ensemble were drawn from the uniform distribution whose upper and lower bounds were shown in Table 2.

The assimilated daily runoff observation was also used to evaluate the model simulation. The performance was evaluated by Kling-Gupta Efficiency (KGE, Gupta et al., 2009):

\[
KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{\text{sim}}}{\sigma_{\text{obs}}} - 1\right)^2 + \left(\frac{\mu_{\text{sim}}}{\mu_{\text{obs}}} - 1\right)^2}
\]  

where \( r \) is a correlation coefficient between simulated and observed runoff, \( \sigma_{\text{sim}} \) and \( \sigma_{\text{obs}} \) are standard deviations of simulated and observed runoff, respectively, \( \mu_{\text{sim}} \) and \( \mu_{\text{obs}} \) are means of simulated and observed runoff, respectively. KGE changes from \(-\infty\) to 1 and the larger KGE means the better skill to reproduce observed runoff. KGE \( = 1 \) indicates perfect correspondence between simulated and observed runoff. The ensemble median was used as a representative value of the estimated probabilistic distribution to evaluate the filters. Although the probabilistic distribution of the estimated runoff is highly skewed and does not follow the Gaussian distribution, the median is often consistent to the mode of the probabilistic distribution in this experiment.

4. Results

4.1. Case Study I: The Lorenz 63 Model With an Abruptly Changed Parameter

Figure 2 shows the probabilistic distribution of the parameters obtained by the offline batch optimization, \( Q(\tilde{\theta}) \). Both \( \rho \) and \( b \) are distributed around the synthetic truth. Although most of observation information is not used by averaging its long-term timeseries in the offline batch optimization, the posterior distribution which reasonably covers the synthetic truth can be obtained.

In Figure 3, the performances of HOOPE-PF and PF are evaluated by Root-Mean-Square-Error (RMSE) of \( \rho(t) \) as the function of the two inflation hyperparameters (Equations 10 and 12). Since this case study is the synthetic experiment, the estimated parameter can be directly compared with the synthetic truth of the parameter. First,
HOOPE-PF provides smaller RMSE of $\rho(t)$ than PF especially when the ensemble size is small ($N = 30$). The minimum RMSE of HOOPE-PF with the ensemble sizes of 30, 100, and 250 are 0.93, 0.60, and 0.56, respectively. On the other hand, the minimum RMSE of PF with the ensemble sizes of 30, 100, and 250 are 1.21, 0.71, and 0.58, respectively. Five different numerical experiments with different seeds for random numbers were performed to evaluate the statistical significance of the difference between minimum RMSEs of HOOPE-PF and PF. While the differences with the ensemble sizes of 30 and 250 are statistically significant, that with 100-member is insignificant (one-sided student’s t-test with a 5% significance level). Note that the mean of minimum RMSEs of these five experiments show that HOOPE-PF is better than PF for all ensemble sizes. Second, in HOOPE-PF, the sensitivity of RMSE to the hyperparameter $S_{para}$ is much smaller than that in PF. Except for the case of the smallest $S_{para}$, HOOPE-PF stably provides the accurate estimation of $\rho(t)$ especially when the ensemble size is large (Figures 3e and 3f). On the other hand, the performance of PF is substantially sensitive to $S_{para}$ (Figures 3b and 3c). In the case of the large $S_{para}$, PF collapses. This superiority of HOOPE-PF over PF can also be found in RMSE of the state variables (results not shown). The choice of $S_{state}$ does not greatly affect RMSE of $\rho(t)$ in both HOOPE-PF and the original PF compared with $S_{para}$ although smaller $S_{state}$ slightly reduces RMSE.

Figure 4 shows the timeseries of $\rho(t)$ in the PF and HOOPE-PF with the selected settings to demonstrate how differently these two data assimilation methods work. The purpose of Figure 4 is to demonstrate how the estimated probabilistic distribution of $\rho(t)$ is affected by $S_{para}$ in PF and HOOPE-PF. Note that the comparison of the performances between PF and HOOPE-PF is not the purpose of these figures. The superiority of HOOPE-PF over PF has already been shown in Figure 3. In the case of the small $S_{para}$, both PF and HOOPE-PF cannot accurately estimate the change in the parameter since the ensemble variance is too small to detect the abrupt regime change (Figures 4a and 4f). When $S_{para}$ increases to 0.3, both PF and HOOPE-PF maintain the appropriate ensemble variances and successfully estimate the time-varying parameter (Figures 4b and 4g). However, in PF, the filter becomes unstable when $S_{para}$ further increases, and PF completely collapses in $S_{para} = 0.9$ (Figures 4c, 4d, and 4e). On the other hand, HOOPE-PF stably works in the cases of the higher $S_{para}$ (Figures 4h, 4i, and 4j). This is because the perturbed
parameters are constrained by the result of the offline batch optimization in HOOPE-PF. The constraint of the offline batch optimization prevents ensemble members from going to the unrealistic space of the parameters in terms of model’s climate even if each ensemble member is strongly perturbed by the large $S_{para}$.

4.2. Case Study II: The Lorenz 63 Model With Non-Periodic Forcing

Figure 5 shows the probabilistic distribution of the parameters obtained by the offline batch optimization, $Q(\hat{\theta})$ in the case study II. Since true $\rho(t)$ changes from 23 to 33 (Equations 24 and 25), the posterior distribution of $\rho(t)$ obtained by the offline batch optimization captures the range of the synthetic true parameter.

Figure 6 indicates that HOOPE-PF outperforms PF in this case study as found in the case study I. In HOOPE-PF, the change in $\rho(t)$ can be accurately estimated when $S_{para}$ is relatively large even if the ensemble size is small. For the smallest ensemble size of 30, a large value of $S_{para}$ is needed to maintain the appropriate ensemble variance. In the original PF, the large value of $S_{para}$ substantially increases the risk to obtain unrealistic $\rho(t)$, which makes the filter unstable. On the other hand, HOOPE-PF can adopt the large value of $S_{para}$ avoiding such unrealistic $\rho(t)$.

Figure 4. (a–e) Timeseries of $\rho(t)$ by the original PF in the case study I (see Sections 3.1 and 4.1). Black and red lines show the synthetic truth and the median of the estimated parameter by the online data assimilation, respectively. Gray areas show the 5–95 percentile range. The ensemble size is set to 250. $S_{state}$ is set to 0.25. $S_{para}$ is set to (a) 0.1, (b) 0.3, (c) 0.5, (d) 0.7, and (e) 0.9. Panels (f–j) same as (a–e) but for Hybrid Offline Online Parameter Estimation with Particle Filtering.

Figure 5. Histograms of (a) $\rho(t)$ and (b) $b$ of the Lorenz 63 model estimated by the MCMC sampler in the case study II (see Sections 3.2 and 4.2). Red dashed line is the synthetic truth of the model parameter.
The minimum RMSE in HOOPE-PF with the ensemble sizes of 30, 100, and 250 are 0.94, 0.70, and 0.65, respectively. On the other hand, the minimum RMSE in PF with the ensemble sizes of 30, 100, and 250 are 1.06, 0.74, and 0.68, respectively. As the case study I (see Section 4.1), five different numerical experiments with different seeds for random numbers were performed to evaluate the statistical significance of the difference between minimum RMSEs of HOOPE-PF and PF. In this case study, the differences of minimum RMSEs with all ensemble sizes are statistically significant. The performance of PF is substantially sensitive to $S_{\text{para}}$ even in the cases of the large ensemble sizes. It implies that the intensive calibration of the inflation hyperparameter is necessary if one uses no objective estimation schemes of the hyperparameters. In HOOPE-PF, this intensive hyperparameter calibration is unnecessary. Note that some test runs to find proper hyperparameters are still necessary even in HOOPE-PF according to Figure 6.

Figure 7 shows the timeseries of $\rho(t)$ in PF and HOOPE-PF with the selected settings to demonstrate how differently these two data assimilation methods work. Analogous to Figure 4, the purpose of Figure 7 is to demonstrate how the estimated probabilistic distribution of $\rho(t)$ is affected by $S_{\text{para}}$ in PF and HOOPE-PF. Figure 7 demonstrates that $S_{\text{para}}$ plays the important role in controlling the ensemble variance and the skill to stably estimate the time-varying parameter in the original PF. On the other hand, the offline batch optimization largely determines the ensemble variance in HOOPE-PF and the sensitivity of the performance of the parameter estimation to $S_{\text{para}}$ in HOOPE-PF is much smaller than that in PF.

4.3. Case Study III: Real-Data Experiment With the Conceptual Hydrological Model

Figure 8 shows the histograms of the HYMOD’s 5 parameters (see Table 2) estimated by the offline batch optimization. All parameters are shown as normalized values by the prescribed maximum and minimum values (see Table 2). There is no known synthetic truth in this case study since the real runoff observation was used in this experiment. The 5-year runoff ratio and baseflow index effectively constrain $b_{\text{exp}}$ and $k_s$. Although the relatively
large variances remain in the other three parameters, all five parameters sampled by MCMC are mutually correlated with each other (not shown). Therefore, the offline batch optimization substantially reduces the total volume of the parameter space which should be searched by the particle filtering.

In Figure 9, the performances of HOOPE-PF and PF are evaluated by KGE as the function of the two inflation hyperparameters. Figure 9 clearly reveals that HOOPE-PF provides larger KGE than PF in most cases. The maximum KGE in HOOPE-PF with the ensemble sizes of 30, 100, and 250 are 0.79, 0.80, and 0.84, respectively. On the other hand, the maximum KGE in PF with the ensemble sizes of 30, 100, and 250 are 0.47, 0.73, and 0.79, respectively. In the supporting information, the timeseries of runoff and $k_q$ (see Table 2) provided by the optimal hyperparameters which maximize KGE are shown (Figures S1–S3 in Supporting Information). The effects of the two hyperparameters on the stability and accuracy of the simulations in this case study are different from those in the previous case studies. The performance of PF is not degraded when $S_{\text{para}}$ is set to the large values. This is mainly because the model parameters are bounded in the reasonable prescribed ranges shown in Table 2, which prevents the filter from collapsing when the ensemble variance of the parameters is strongly inflated. In this case study, ensemble members are already constrained by the prescribed ranges shown in Table 2, which stabilizes the original PF. The offline batch optimization further constrains ensemble members to the regions in which the observed runoff ratio and baseflow index can be reasonably reproduced. This additional constraint by the offline batch optimization significantly improves KGE.

Figure 9c shows the high $S_{\text{state}}$ is necessary for PF to obtain high KGE. This high KGE of PF with 250-member is maintained by further increasing $S_{\text{state}}$, and the KGE of PF with 100-member can be increased by increasing $S_{\text{state}}$ (not shown). However, the improvement of KGE by choosing extremely high $S_{\text{state}}$ could not achieve the goal of this study, which is the accurate estimation of the time-varying parameters of HYMOD. First, since the dimension of the observation space is small (i.e., one-dimension) and observations assimilated into HYMOD is directly used for verification, it is relatively easy to reproduce observed runoff by randomly perturbing state variables and finding appropriate particles without accurately estimating parameters when the ensemble size is large. Second, the perturbation of the state variables can represent model errors by lumping all possible sources of model errors together, so that the large $S_{\text{state}}$ can compensate the errors induced by the misspecification of the time-varying parameters. Therefore, the large $S_{\text{state}}$ makes both HOOPE-PF and PF to accurately reproduce observed runoff without accurately estimating model parameters. HOOPE-PF can provide high KGE with the small $S_{\text{state}}$, which shows that HOOPE-PF can accurately identify the errors induced by the misspecification of the time-varying parameters. It should be noted that in the case studies I and II, the performance tends to be slightly better when $S_{\text{state}}$ is small (see Figures 3 and 6). Since the skill to estimate the time-varying parameter is directly evaluated in these idealized experiments, this fact implies that the application of HOOPE-PF with small $S_{\text{state}}$ to HYMOD accurately estimates time-varying parameters in addition to runoff.
Figure 10 demonstrates, analogous to Figure 4, the difference between HOOPE-PF and PF by showing the time-series of runoff and $k_q$ (see Table 2). The ensemble size is 100. $S_{\text{state}}$ and $S_{\text{para}}$ are set to 0.008 and 0.7, respectively. These inflation hyperparameters provide the 50th highest KGE in HOOPE-PF. Note that $k_q$ shown in Figures 10c and 10d is a normalized value with respect to the prescribed minimum and maximum values (see Table 2). $k_q$ controls residence time for quick-flow tanks in HYMOD. Higher $k_q$ induces a quicker response of runoff to a heavy rainfall. Although PF accurately estimates flood peaks in many cases, PF overestimates baseflow (Figure 10b) due to the unstable behavior of the estimated parameters (Figure 10d). This behavior of PF can also be found even when the optimal hyperparameters were chosen to maximize KGE of PF (Figures S1 and S2 in...
Supporting Information S1). On the other hand, HOOPE-PF successfully simulates both high flow and base-flow (Figure 10a). The temporal change of \( k_q \) in HOOPE-PF is stabler than that in the original PF (Figures 10c and 10d). In HOOPE-PF, \( k_q \) rapidly increases mainly in the flood periods. It is generally difficult for hydrological models with a single set of parameters to accurately simulate runoff in both wet and dry periods (e.g., Deb & Kiem, 2020). The application of HOOPE-PF to HYMOD overcomes this limitation of hydrological modeling by allowing the model parameters to abruptly change based on the observed runoff. It should be noted that the original PF can also track the abrupt changes in the hydrological parameters although it needs the intensive hyper-parameter turning and/or the larger ensemble size, as shown in Figure 9. Figure 9 clearly shows that the parameter perturbation method in HOOPE-PF significantly improves the skill to successfully track the abrupt changes in the model parameters.

5. Conclusions and Discussion

In this paper, a new efficient and practical method to estimate model parameters allowing them to temporally change was proposed by combining offline batch optimization and online data assimilation. In the newly proposed method, HOOPE-PF, the estimated model parameters by the sequential data assimilation are constrained to the result of the offline batch optimization in which the posterior distribution of model parameters is obtained by comparing the simulated and observed climatological variables. HOOPE-PF outperforms the original SIRPF in the synthetic experiments with the Lorenz 63 model and the real-data experiment of the rainfall-runoff analysis especially when the ensemble size is small. The advantage of HOOPE-PF is that the performance of the online data assimilation is not greatly affected by the "inflation" hyperparameter for model parameters, so that the extensive tuning of this hyperparameter is not necessary. Considering the true dynamics of time-varying parameters is unknown and unobservable, it is difficult to determine which hyperparameters can realize the accurate estimation of time-varying model parameters in real-world applications. This makes the tuning of hyperparameters
complicated. In this case, the stability of HOOPE-PF contributes to improving the efficiency and practicality to estimate convincing time-varying parameters. The efficiency and stability of HOOPE-PF may contribute to solving computationally expensive problems related to time-varying model parameters such as large sample studies on conceptual hydrological models (e.g., Knoben et al., 2020). The inflation of model parameters has been intensively discussed in the previous works on data assimilation (e.g., Kotsuki et al., 2018; Moradkhani, Hsu, et al., 2005; Moradkhani, Sorooshian, et al., 2005; Moradkhani et al., 2012; Ruiz et al., 2013), and HOOPE-PF contributes to these published literature.

The proposed HOOPE-PF can be the practical approach to detect and quantify unmodeled processes such as land use change (e.g., Patheraja et al., 2018) from long-term observation. In addition, modeling with time-varying parameters realized stochastic process models whose potential was intensively discussed by Reichert et al. (2021). Reichert et al. (2021) clearly shows that the stochastic process models with time-varying parameters have the advantage to identify an intrinsic model uncertainty and improve model processes.

Since PF is robust to non-linear dynamics and non-Gaussian statistics, it has been widely applied to analyze and predict non-linear earth systems such as land (e.g., Abolafia-Rosenzweig et al., 2019; Qin et al., 2009), ecosystem

Figure 10. (a and b) Timeseries of daily discharge in the Leaf River basin. Black line shows the observed river discharge. Red lines and gray areas show median and the 10–90 percentile range of simulated river discharge by (a) Hybrid Offline Online Parameter Estimation with Particle Filtering (HOOPE-PF) and (b) PF. (c and d) Timeseries of the normalized quick flow routing coefficient (see Table 2) in the Leaf River basin. Red lines and gray areas show median and the 10–90 percentile range of the simulated normalized quick flow routing coefficient by (c) HOOPE-PF and (d) PF. The ensemble size is 100, $S_{state}$ and $S_{para}$ are set to 0.008 and 0.7, respectively. These inflation hyperparameters provide the 50th highest Kling-Gupta Efficiency in HOOPE-PF.
There are three limitations in HOOPE-PF. First, since HOOPE-PF requires the long-term observation before beginning the real-time online estimation, it cannot be applied to some types of the real-time estimation problem in which there is no observation at the initial time and the estimation of state variables and parameters needs to be gradually improved whenever new observations are obtained. This situation is often assumed in the previous works about data assimilation on groundwater modeling (e.g., Hendricks Franssen & Kinzelbach, 2008; Ramgraber et al., 2019) although the long-term observation timeseries can be obtained in many geoscientific applications such as numerical weather prediction and rainfall-runoff analysis. In addition, if long-term observations do not cover the information about “true” parameters for future dynamics, HOOPE-PF cannot accurately estimate them since the future “true” parameters are located outside the probability distribution estimated by the offline batch optimization. Climate change is one of the possible examples. It should be noted that even in this case HOOPE-PF can evaluate how far the online estimation of parameters are located from the mode of the probability distribution estimated by the offline batch optimization and objectively detect the unprecedented situation which cannot be explained in previous records.

Second, there is no generally applicable way to construct the appropriate “climatological index” for the offline batch optimization (Equation 15). Equation 15 needs to be the summary of the climatology of the important process for a user's purpose. It should be independent to the initial condition. The design of Equation 15 apparently depends on the problems. The design of the appropriate climatological index for the offline batch optimization is generally challenging (e.g., Springer et al., 2021). It is promising that the climatological index based only on the widely used hydrological indices, runoff ratio and baseflow index, appropriately works in the application of HOOPE-PF to the conceptual hydrological model. It implies that no complicated methods are needed to construct Equation 15.

Third, the current HOOPE-PF cannot be directly applied to high-dimensional problems because the number of required ensemble members increases exponentially with the size of problems (e.g., Snyder et al., 2008, 2015). The recent progress on the application of PF to the high-dimensional problem (i.e., numerical weather prediction) (e.g., Kawabata & Ueno, 2020; Poterjoy et al., 2019; Potthast et al., 2019, see also van Leeuwen et al. (2019) for the comprehensive review) needs to be included to directly apply HOOPE-PF to optimize parameters of high-dimensional models. In addition, the fundamental idea of HOOPE-PF, which is the combination of online data assimilation and offline batch optimization to stabilize the estimation of time-varying parameters, can be transferred to EnKF which has been successfully applied to the high-dimensional problems. The future work will be to develop the version of an EnKF (HOOPE-EnKF) based on the similar strategy which is shown in this paper.

Data Availability Statement
The CAMELS data set can be downloaded at https://ral.ucar.edu/solutions/products/camels. The author thank Ken Dixon for providing a useful Python source code of the Lorenz63 model which was used in this paper and can be downloaded at https://github.com/kendixon/Lorenz63. The author also thank Kel Markert for providing a useful Python source code of the HYMOD model which was used in this paper and can be downloaded at https://github.com/KMarkert/hymod.

References
Abbaszadeh, P., Moradkhani, H., & Yan, H. (2018). Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo. *Advances in Water Resources*, 111, 192–204. https://doi.org/10.1016/j.advwatres.2017.11.011
Abolafia-Rosenzweig, R., Livneh, B., Small, E. E., & Kumar, S. V. (2019). Soil moisture data assimilation to estimate irrigation water use. *Journal of Advances in Modeling Earth Systems*, 11(11), 3670–3690. https://doi.org/10.1029/2019MS001797
Addor, N., Newman, A. J., Muzukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, 21(10), 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Acknowledgments
This study was supported by the JST FOREST program (Grant No. JPMJ-FR205Q), the KAKENHI grant (Grant No. 21H01430), and the Foundation of River & basin Integrated Communication (FRICS). The author thank two anonymous reviewers for their helpful comments.
Qian, Y., Wan, H., Yang, B., Golaz, J. C., Harrop, B., Hou, Z., et al. (2018). Parametric sensitivity and uncertainty quantification in the version 1 of E3SM atmosphere model based on short perturbed parameter ensemble simulations. Journal of Geophysical Research: Atmospheres, 123(23), 13046–13073. https://doi.org/10.1002/2018JD028927

Qin, J., Liang, S., Yang, K., Kainohitsu, I., Liu, R., & Koike, T. (2009). Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal. Journal of Geophysical Research, 114(D15), 1–13. https://doi.org/10.1029/2008JD011358

Ramgraber, M., Albert, C., & Schirmer, M. (2019). Data assimilation and online parameter optimization in groundwater modeling using nested particle filters. Water Resources Research, 55(11), 9724–9747. https://doi.org/10.1029/2018WR024408

Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. The MIT Press.

Reichert, P., Ammann, L., & Fencic, F. (2021). Potential and challenges of investigating intrinsic uncertainty of hydrological models with stochastic, time-dependent parameters. Water Resources Research, 57(3), 1–28. https://doi.org/10.1029/2020WR028400

Ruiz, J. J., Pulido, M., & Miyoshi, T. (2013). Estimating model parameters with ensemble-based data assimilation: Parameter covariance treatment. Journal of the Meteorological Society of Japan, 91(4), 453–469. https://doi.org/10.2151/jmsj.2013-403

Sawada, Y. (2020). Machine learning accelerates parameter optimization and uncertainty assessment of a land surface model. Journal of Geophysical Research: Atmospheres, 125(20). https://doi.org/10.1029/2020JD032688

Sawada, Y., & Hanazaki, R. (2020). Socio-hydrological data assimilation: Analyzing human-flood interactions by model-data integration. Hydrology and Earth System Sciences, 24(10), 4777–4791. https://doi.org/10.5194/hess-24-4777-2020

Sawada, Y., Koike, T., & Walker, J. P. (2015). A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics. Journal of Geophysical Research, 120(12), 5910–5930. https://doi.org/10.1002/2015JF003285

Sheikholeslami, R., & Razavi, S. (2020). A fresh look at variography: Measuring dependence and possible sensitivities across geophysical systems from any given data. Geophysical Research Letters, 47(20), 1–11. https://doi.org/10.1029/2020GL089829

Snyder, C., Bengtsson, T., Bickel, P., & Anderson, J. (2008). Obstacles to high-dimensional particle filtering. Monthly Weather Review, 136(12), 4629–4640. https://doi.org/10.1175/2008MWR2529.1

Snyder, C., Bengtsson, T., & Morzfeld, M. (2015). Performance bounds for particle filters using the optimal proposal. Monthly Weather Review, 143(11), 4750–4761. https://doi.org/10.1175/MWR-D-15-0144.1

Springer, S., Haurio, H., Susiluoto, J., Bibow, A., Davis, A., & Marzouk, Y. (2021). Efficient Bayesian inference for large chaotic dynamical systems. Geoscientific Model Development, 14(7), 4319–4333. https://doi.org/10.5194/gmd-14-4319-2021

Teixeira Parente, M., Bittner, D., Mattis, S. A., Chiogna, G., & Wohlmuth, B. (2019). Bayesian calibration and sensitivity analysis for a karst aquifer model using active subspaces. Water Resources Research, 55(8), 7086–7107. https://doi.org/10.1029/2019WR024739

Tomizawa, F., & Sawada, Y. (2021). Combining ensemble Kalman filter and reservoir computing to predict spatiotemporal chaotic systems from imperfect observations and models. Geoscientific Model Development, 14(9), 5623–5635. https://doi.org/10.5194/gmd-14-5623-2021

van Leeuwen, P. J., Künsch, H. R., Nerger, L., Potthast, R., & Reich, S. (2019). Particle filters for high-dimensional geoscience applications: A review. Quarterly Journal of the Royal Meteorological Society, 145(723), 2335–2365. https://doi.org/10.1002/qj.3551

Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., & Schoups, G. (2013). Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. Advances in Water Resources, 51, 457–478. https://doi.org/10.1016/j.advwatres.2012.04.002

Whitaker, J. S., & Hamill, T. M. (2012). Evaluating methods to account for system errors in ensemble data assimilation. Monthly Weather Review, 140(9), 3078–3089. https://doi.org/10.1175/MWR-D-11-00276.1

Yan, H., & Moradkhani, H. (2016). Combined assimilation of streamflow and satellite soil moisture with the particle filter and geostatistical modeling. Advances in Water Resources, 94, 364–378. https://doi.org/10.1016/j.advwatres.2016.06.002

Zhang, J., Zheng, Q., Chen, D., Wu, L., & Zeng, L. (2020). Surrogate-based Bayesian inverse modeling of the hydrological system: An adaptive approach considering surrogate approximation error. Water Resources Research, 56(1), 1–25. https://doi.org/10.1029/2019WR025721