Parameter identification and uncertainty quantification of a non-linear pump-turbine governing system based on the differential evolution adaptive Metropolis algorithm

Chu Zhang$^{1,2}$ | Tian Peng$^1$ | Jianzhong Zhou$^2$ | Muhammad Shahzad Nazir$^1$

1 College of Automation, Huaiyin Institute of Technology, Huaian, China
2 School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan, China

Abstract
Pumped storage units (PSUs) are now widely used for energy storage. However, the uncertainty of the identification results of the pump-turbine governing system (PTGS) caused by the random observation noises and the lack of prior knowledge remains an unaddressed issue. In recent years, the differential evolution adaptive Metropolis algorithm (DREAM) based on the Bayesian theory has been extensively used for parameter estimation and uncertainty analysis, but its application to the uncertainty analysis of PTGS has been rare. This study systematically evaluates the applicability of DREAM in the parameter identification and uncertainty quantification of PTGS. A real PSU in China has been employed as a case study for numerical experiments. Four groups of control experiments with different proportions of observation noises and different prior search spaces have been constructed in this study. It can be concluded from this study that: (a) accurate point identification results and effective uncertainty quantification can be obtained simultaneously using DREAM, (b) a lower proportion of observation noises can enhance the efficiency and effectiveness of the DREAM algorithm when applying to PTGS simulation. The DREAM method can narrow the prior search space of PTGS parameters effectively and thus help the hydropower engineers to make decisions.

1 | INTRODUCTION

Pumped storage units (PSUs) are important energy storage technologies that can effectively alleviate the impact of intermittent fluctuating energy sources such as wind and solar on the power system [1, 2]. They are widely used to improve the power grid’s ability to absorb clean energy and thus solve the safety and stability problem brought by the integration of renewable energy generation systems [3]. With the growing market of clean and renewable energy and the increasing requirements for safe and stable operation of the power grid [4], the construction of PSUs is increasingly prominent [5]. In order to model the behaviours of different modules of PSU, the pump-turbine governing system (PTGS) has a set of parameters. However, it is difficult to obtain the actual value of the PSU parameters directly from on-site measurements, so it is becoming more and more important to identify the PTGS parameters to model the PSU accurately [6].

In the past 20 years, great progress has been made in parameter identification theory for PTGS [7, 8]. The identification methods of PTGS is becoming more and more mature. However, it should be pointed out that due to the limitations of the on-site test conditions and researchers’ cognition of the inherent complexity and non-linearity of PSUs [9], the uncertainty of the parameter identification results will inevitably occur. A typical flowchart of the parameter identification process is shown in Figure 1. As shown in Figure 1, the parameter identification results of PTGS are affected by the fuzziness of prior knowledge, the randomness of measurement signal noise, the incompleteness of the identification model, and the influences of different identification algorithms, thus the parameter identification of PTGS is likely to be inaccurate or even wrong [2, 10–12].

Traditional parameter identification methods and strategies based on optimisation theory belong to single-point estimation from the statistical point of view [13–15]. Especially,
Various intelligent evolutionary algorithms have been employed for the parameter identification of PTGS or hydraulic turbine-governing system (HTGS), such as gravitational search algorithm and its improved versions [10,12,16,17], improved backtracking search algorithm (IBSA) [11], improved ant lion optimisation algorithm [18] and so forth. Only one theoretical optimal solution can be obtained, and the uncertainty of the identification results cannot be evaluated [19]. With the development of the control and stability analysis theory, point parameter identification results can no longer meet the requirements of theoretical research and industrial application. Parametric uncertainty analysis methods that can obtain the probability distribution and confidence interval of the model parameters have become more and more important [20].

A number of approaches have been developed for parametric uncertainty analysis [20–23], among which the Bayesian theory has been extensively used for parameter estimation and uncertainty analysis in recent years [24, 25]. The probability inversion method based on Bayesian theory sets the unknown parameters as random variables, and the probability density distribution function is used to describe the uncertainty of the results. The Markov chain Monte Carlo (MCMC) is widely used to sample the search space of the model parameters randomly, and the probability description of the posterior distribution of the model parameters is obtained according to the likelihood values and the prior information of the parameter values [19].

The traversal efficiency and estimated performance of the MCMC method mainly depend on the way the prior distribution is constructed and the parameter updated strategy of the Markov chain. The differential evolution adaptive metropolis (DREAM) is a fast and efficient MCMC algorithm that combines the multi-chain parallelism idea and the differential evolution (DE) algorithm. It has been widely used in the research of parameter inversion and quantitative uncertainty evaluation in hydrology [26–28], geophysics [29, 30], agronomy [19] and other fields. Zheng and Han [26] applied the DREAM algorithm to access the uncertainty of the parameters of watershed-scale water quality models. Sheng et al. [19] first used the DREAM method to estimate the parameters of the maize module of the Agricultural Productions Systems sIMulator model (APSIM-maize). Results indicate that the DREAM method performed well in theoretical and real-world evaluation. Zahmatkesh et al. [27] used the DREAM method to investigate the model parameters and input uncertainty of the rainfall-runoff process in urban areas. Results turn out that the developed method can provide reliable tools for rainfall and runoff modelling.

Many studies have applied the DREAM method to the parametric uncertainty analysis of complex non-linear models, but very few of the DREAM analyses addressed the PTGS parameter identification problem. In order to quantitatively evaluate the uncertainty of the parameter identification results caused by the random observation noises and the lack of prior knowledge, a Bayesian uncertainty analysis framework for parameter identification of PTGS is established in this study. To the best of the authors’ knowledge, it is the first time to investigate the application of the DREAM algorithm in the parameter identification and uncertainty analysis of PTGS. The DREAM method can not only obtain the point identification results of PTGS parameters but also quantify the uncertainty of the identification results. The posterior probability density function of the parameters to be identified is sampled and traversed using the DREAM method. The posterior probability distribution including all the information of the parameter set is obtained, which can intuitively express the uncertainty of the identification results and thus provide a new idea for the improvement and development of the existing parameter identification strategies for PTGS.

The rest of this study is arranged as follows: Section 2 introduces the theoretical background involved in this study, Section 3 presents the detailed parametric uncertainty analysis process of PTGS based on DREAM, Section 4 gives a real-world case study with experiments and results analysis, and Section 5 presents the conclusions of this study.

2 | BACKGROUND

2.1 | Model of the PTGS

PTGS is designed as a modular system with governor, electro-hydraulic servomechanism, penstock, pump-turbine and generator modules [11, 31]. The governor module outputs the control signals with certain regulation characteristics according to the
unit frequency feedback, signal setting and operation instructions. The electrohydraulic servomechanism module is used to amplify the control signal of the governor [32]. The penstock module describes the dynamic process of the penstock pressure and the water head of the pump-turbine. The pump-turbine module is the core module of PSU, which is the key to realise the efficient conversion of hydropower energy. The generator module operates as a generator under the condition of a water turbine and as a motor under the condition of the water pump. The framework of the non-linear simulation model of PTGS is given in Figure 2.

PTGS is the terminal actuator and core control system for automatic generation control of PSU. As shown in Figure 2, the generator of PTGS is designed with the given speed \( x_c \) and the unit speed \( x \) as input signals. After the non-linear transformation of the governor module and the amplification of the electrohydraulic servomechanism module, the guide vane opening \( y \) is obtained. The guide vane opening \( y \) is then sent into the penstock and pump-turbine modules to obtain the mathematical relationship between the unit flow (or moment) and the guide vane opening, rotating speed and working water head. The unit speed is obtained finally after the generator module. The definition of the PTGS parameters regarding different modules are as follows: (a) Governor module: The proportional gain \( K_p \), the integral gain \( K_i \), the differential gain \( K_d \); (b) electrohydraulic servomechanism module: The main relay \( T_y \) and the inertial time constant of the main pressure valve \( T_{y,1} \); (c) penstock module: The flow inertia time constant \( T_w \), the head-loss coefficient \( f \), the water-hammer pressure wave time constant \( T_{w,1} \); (d) generator module: The inertial time constant of the generator \( T_a \) and the adjusting coefficient \( \epsilon_a \). The PTGS parameters to be identified include \( K_p, K_i, K_d, T_y, T_{y,1}, T_w, T_a \) and \( \epsilon_a \) [11].

2.2 | DREAM

DREAM proposed by Vrugt et al. [33] is an MCMC sampling method with high-solving efficiency. Based on an adaptive random sampling of subspaces, useless chain removal technology, cross and mutation strategies, the DREAM algorithm can quickly converge to a stable distribution state. The DREAM algorithm has good sampling efficiency and estimation accuracy for high-dimensional decision spaces with complex multi-peaks. Therefore, it is widely used for uncertainty quantitative assessment in various fields, such as hydrology, geology, environment and so forth. For a better understanding of the DREAM algorithm, the Bayesian theory and the MCMC method are first introduced in this study, and then the implementation process of DREAM algorithm is introduced.

2.2.1 | The Bayesian theory

The Bayesian theory is proposed by British mathematician Thomas Bayes as a statistical inference method and is an important branch of mathematical-statistical theory. The Bayesian theory treats the unknown parameters \( \theta \) as random variables and describes their uncertainties in the form of probability distributions. The core of the Bayesian theory is to first calculate the posterior distribution of unknown parameters based on the prior distribution of the parameters when the sample is unknown and the likelihood function of the parameters when the sample is known, and then calculate or infer the unknown parameters through the posterior distribution. Compared with the classical statistical theory methods, the Bayesian inference can make full use of the information of sample sampling and
prior distribution to obtain an estimate with a smaller error. Assuming that \( x_1, x_2, \ldots, x_n \) represents the sample data, the theoretical framework of Bayesian inference is

\[
p(\theta | x_1, x_2, \ldots, x_n) = \frac{p(\theta) / (x_1, x_2, \ldots, x_n | \theta)}{\int p(\theta) / (x_1, x_2, \ldots, x_n | \theta) d\theta} \tag{1}
\]

where \( p(\theta | x_1, x_2, \ldots, x_n) \) represents the posterior distribution of \( \theta \); \( f(x_1, x_2, \ldots, x_n | \theta) \) represents the likelihood function of \( \theta \) when the sample \( x_1, x_2, \ldots, x_n \) is known; \( p(\theta) \) represents the prior distribution of \( \theta \) when the sample is unknown, \( \int p(\theta)p(y | \theta)d\theta \) represents the normalisation constant.

### 2.2.2 The MCMC sampling

In most cases, the integral of the denominator of Equation (1) cannot be calculated in a closed form, such that the posterior distribution cannot be solved analytically. Therefore, the Monte Carlo algorithm based on stochastic simulation, such as the MCMC sampling, is usually used to calculate the posterior distribution of model parameters.

Suppose that \( X_0, X_1, X_2, X_n, \ldots \) is a Markov chain in which the variables are randomly generated with Markov property. Given any \( X_t \), the variable of the next moment \( X_{t+1} \) is only related to \( X_t \) and has nothing to do with \( X_0, X_1, X_2, \ldots, X_{t-1} \). The mathematical expression of the Markov chain can be expressed as follows:

\[
P \left( X_{t+1} = \theta' | X_0, X_1, \ldots, X_t \right) = P \left( X_{t+1} = \theta' | X_t = \theta \right) \quad \forall \theta, \theta' \in S \tag{2}
\]

where \( \theta \) and \( \theta' \) are the states at time \( t \) and \( t + 1 \), respectively; \( S \) represents the state-space of the Markov chain. The Markov chains can be represented by the state-space \( S \) and the transition kernel \( \rho(\theta, \theta') \) between the states.

Assuming that the Markov chain is in a state \( \theta \) at the current moment, and the transfer kernel \( \rho(\theta, \theta') \) represents the probability of it being in state \( \theta' \) at the next moment, after a finite number of iterations, the Markov chain finally converges to the stationary distribution \( \pi(\theta) \) that satisfies the following equation:

\[
\int \rho(\theta, \theta') \pi(\theta') d\theta = \pi(\theta') \tag{3}
\]

where \( \pi(\theta) \) represents the stationary distribution of the transfer kernel \( \rho(\theta, \theta') \).

The MCMC method first generates several parallel Markov chains with a stationary distribution \( \pi(\theta) \) and then samples the stationary distribution \( \pi(\theta) \). By updating the sample information iteratively, the Markov chain can effectively search the feasible region of the model parameters and finally converge to the high-probability density region, that is, the maximum posterior estimation. The key to the MCMC method is to construct an appropriate transfer kernel.

#### 2.2.3 Process of the DREAM algorithm

The DREAM algorithm employs the adaptive Metropolis (AM) sampling to construct the transfer kernel of the Markov chain. It combines the Markov chain based on AM sampling with the DE algorithm to form an algorithm with strong global convergence ability [33]. The specific steps of the DREAM algorithm are as follows:

**Step 1:** The posterior probability density of each parameter is initialised. Assuming that the number of unknown parameters to be identified is \( n \). \( N \) initial individuals \( \{\theta^i, i = 1, 2, \ldots, N\} \) are first generated according to the prior distribution of the unknown parameters randomly. And then the posterior probability density of the \( N \) initial individuals is calculated, and the iteration process of the \( N \) parallel Markov chains is carried out.

**Step 2:** Mutation. For the \( j \)th Markov chain, mutation operation is performed on each individual of the \( j \)th iteration according to Equation (4):

\[
r_{j,t+1}^i = \theta_{j,t} + (I_q + \varepsilon) \gamma(\delta, n)
\]

where \( \theta_{j,t} \), \( j = 1, 2, \ldots, N \) represents the individual of the \( j \)th generation of the \( j \)th Markov chain; \( r_{j,t+1}^i \) denotes the mutated individual; \( \varepsilon \) and \( \delta \) are two random numbers taken from the \( n \)-dimensional uniform distribution \( U(-b, b) \) and the \( n \)-dimensional normal distribution \( N_0(0, b^2) \), respectively; \( b \) represents the width less than the target distribution; \( \delta \) denotes the number of the Markov chains employed to generate the mutated individuals; \( r_1(p), r_2(q) \in \{1, 2, \ldots, N\} \) and \( r_1(p) \neq r_2(q) \neq j \); \( p, q = 1, 2, \ldots, \delta \); \( n \) is the number of parameters to be identified; \( \gamma(\delta, n) \) denotes the jump rate and generally taken as \( 2.38 / \sqrt{2 \delta n} \).

**Step 3:** Cross. The cross operation of the DREAM algorithm is used to determine whether to update the \( i \)th component of the current individual. The process of the cross operation is as follows:

\[
r_{j,t+1}^i = \begin{cases} \theta_{j,t}^i & u \leq p_i \\ r_{j,t+1}^i & \text{else} \end{cases} \tag{5}
\]

where \( p_i \) denotes the crossover probability, \( u \) denotes the random number in \([0,1] \). The \( i \)th component of the current individual is retained if \( u \leq 1 - p_i \), otherwise the \( i \)th component of the current individual is updated.
Step 4: The posterior probability of the parameter $\theta_{i,t+1}$ and the reception probability are calculated according to Equation (6):

$$\alpha(\theta_{i,t}, \theta_{i,t+1}) = \begin{cases} \min \left( \frac{p(\theta_{i,t+1}|y)}{p(\theta_{i,t}|y)}, 1 \right) & \text{if } p(\theta_{i,t}|y) > 0 \\ 1 & \text{if } p(\theta_{i,t}|y) < 0 \end{cases}$$

where $\alpha(\theta_{i,t}, \theta_{i,t+1})$ represents the receiving probability; $p(\theta_{i,t+1}|y)$ represents the posterior probability of $\theta_{i,t+1}$; $p(\theta_{i,t}|y)$ represents the posterior probability of $\theta_{i,t}$.

Step 5: Determine whether to accept the cross-mutated individual according to the reception probability. If $\alpha(\theta_{i,t}, \theta_{i,t+1}) \geq R$, accept the new individual, otherwise let $\theta_{i,t+1} = \theta_{i,t}$ and proceed to the next iteration.

Step 6: The useless chains according to the inter-quartile-range (IQR) statistical test is removed [33].

Step 7: If $S_R < 1.2$, the Markov chain converges to the posterior probability distribution and stops iteration; otherwise, steps 2–6 are to be repeated. $S_R$ represents the scale-reduction-scoring factor of which the mathematical expression is as follows [34]:

$$S_R = \sqrt{\frac{k-1}{k} + \frac{s+1}{s-k} B/W}$$

where $k$ denotes the number of the iterations, $s$ represents the number of parallel chains used for evaluation, $B/k$ denotes the variance of the average value of all Markov chains, $\bar{n}$ denotes the average value of $n_i$, and $W$ denotes the average value of the variance $s_i$ of the $s$ parallel chains.

The flowchart of the DREAM method is given in Figure 3.

2.2.4 Performance evaluation indices

After the posterior parameter distributions for each parameter are obtained using the DREAM method, the posterior mean and variance are calculated to evaluate the parameter identification results of PTGS using DREAM. The posterior mean and variance are calculated as follows:

$$\hat{\mu}(\theta)_{post} = \frac{1}{(j*n) - 1} \sum_{j=1}^{N} \sum_{i=a/2}^{a} \theta_{ij}$$

$$\hat{\sigma}^2(\theta)_{post} = \frac{1}{(j*n) - 1} \sum_{j=1}^{N} \sum_{i=a/2}^{a} (\theta_{ij} - \hat{\mu}(\theta)_{post})^2$$

where $N$ denotes the number of chains and $n$ denotes the number of maximum iterations, that is, the length of each chain.

3 PARAMETRIC UNCERTAINTY ANALYSIS PROCESS OF PTGS BASED ON DREAM

The PTGS is a high-dimensional and strong non-linear system. It cannot be expressed analytically and the posterior probability density distribution of the system parameters cannot be directly derived. For Bayesian parametric uncertainty analysis, the parameters to be identified are considered as a set of random variables, and the MCMC method is used to sample the search space of the model parameters randomly. This study introduces the DREAM algorithm that combines multi-chain parallelism and differential evolution. It has high-sampling efficiency and estimation accuracy. The parametric uncertainty analysis steps based on DREAM for PTGS are as follows:
TABLE 1  The prior distribution of the eight PTGS parameters

| Parameter | Definition                                    |  System value | [50%, Min] | 150% Max | [10%, Min] | 300% Max |
|-----------|-----------------------------------------------|---------------|------------|----------|------------|----------|
| $K_p$     | Proportional gain                             | 8.7           | 4.35       | 13.05    | 0.87       | 26.1     |
| $K_i$     | Integral gain                                 | 0.35          | 0.175      | 0.525    | 0.035      | 1.05     |
| $K_d$     | Differential gain                             | 1.9           | 0.95       | 2.85     | 0.19       | 5.7      |
| $T_y$     | Open-loop time constant of the main relay    | 0.2           | 0.1        | 0.3      | 0.02       | 0.6      |
| $T_{yl}$  | Inertial time constant of the main pressure valve | 0.02     | 0.01       | 0.03     | 0.002      | 0.06     |
| $T_w$     | Flow inertia time constant                   | 1.3           | 0.65       | 1.95     | 0.13       | 3.9      |
| $T_a$     | Inertial time constant of the generator      | 10.8          | 5.4        | 16.2     | 1.08       | 32.4     |
| $e_n$     | Adjusting coefficient of the generator       | 1             | 0.5        | 1.5      | 0.1        | 3        |

Step 1: The observation data points of the PTGS system are collected. Because of the fuzziness of the prior information of the parameters to be identified, this study selects the uniform distribution on search space of the parameters as the prior distribution.

Step 2: For the $n$-dimensional parameters set to be identified $\mathbf{\hat{\theta}}$, $N$ initialised Markov chains are generated and the parameter decision space is randomly tranversed.

Step 3: The $N$ candidate samples are sent to the non-linear model of PTGS and the model output is recorded, and then the posterior probability of the $N$ candidate samples is calculated.

Step 4: Cross and mutate steps are performed according to Equations (4) to (6), and then whether to accept the sample according to the acceptance probability $a$ and remove the useless chains according to the IQR value are determined.

Step 5: If the convergence criteria are reached, then we proceed to the next step, otherwise we skip to step 3.

Step 6: The calculated Markov chain is saved and the posterior probability density distribution of each parameter is calculated.

4 | EXPERIMENTS AND RESULTS ANALYSIS

To verify the effectiveness of the parametric uncertainty analysis method of PTGS based on DREAM, a real pumped storage unit (PSU) located in the Bailian River pumped storage in China [5] is simulated. The type of the pump-turbine model is HLN-LJ-550 [7]. A group of parameters $\mathbf{\hat{\theta}} = [K_p, K_i, K_d, T_y, T_{yl}, T_w, T_a, e_n]$ is set as the real value of PTGS. The PTGS simulates the real PSU at a 50 s time step. The sampling frequency is set to 100 Hz.

4.1 | Parameters setting

The main parameters of the DREAM algorithm are set as follows: The number of the parameters to be identified is eight; the number of the Markov chains is 16; the number of maximum iterations is 1000. The prior distribution of the eight PTGS parameters is set as the uniform distribution on the search space. This study sets the search space of the eight PTGS parameters as [50%, 150%] of the system value. The eight PTGS parameters and their pre-defined prior uniform distributions are given in Table 1. The search space of the eight PTGS parameters is chosen according to the published literatures [11, 16].

In order to test the influences of the environmental interference, such as the operation of water machine and water flow, the electromagnetic interference of the generator, and the measurement noise on the parameter identification results of PTGS, different proportions of noises are added to the outputs of the governor [11]. The parameter identification experiments are designed under different signal-to-noise ratio (SNR). The SNR is set as 90 and 20 dB for the parameter identification experiments, respectively. This study chooses 90 and 20 dB as the noise scenarios because they represent two typical noise scenarios, namely, high definition and low definition of noises, respectively.

4.2 | Results and discussions

4.2.1 | Analysis of scaled-reduction score curves

The influence of the observation data uncertainty on the parameter identification results is analysed in this section. The scaled-reduction score curves of each parameter of the PTGS system using the DREAM method under 90 and 20 dB noise scenarios are illustrated in Figures 4(a) and (b), respectively. $T_y$ stands for $T_y$ in the following figures of this study for better exhibition, and so do the other seven symbols.

From Figure 4, it is known that the DREAM method tends to converge to a stationary posterior distribution finally for all the eight PTGS parameters. As shown in Figure 4, it is noted that both the scaled-reduction score curves for the two scenarios have a sharp decline in the first 4000 model evaluations. The scaled-reduction score turns to remain unchanged when it comes to the 6000th model evaluation. The convergence speed of scenario 20 dB is a little faster than that of scenario 90 dB.
4.2.2 | Analysis of posterior probability distribution curves

The posterior probability distribution curves of the eight PTGS parameters estimated using DREAM under scenarios 90 and 20 dB are illustrated in Figures 5 and 6, respectively. The red vertical line in the posterior probability distribution curve represents the system value of the parameters. The shapes of the posterior probability distribution curves indicate the degree of uncertainty for the parameter identification results. Flat and wide distributions indicate more uncertainty for the parameters, while sharp and narrow ones indicate less uncertainty [19].

From Figure 5, it can be found that: (a) The marginal distribution intervals are relatively small compared with the predefined search range with uniform distribution, which indicates that the eight PTGS parameters are reasonably well identifiable; (b) the posterior probability distribution curves of the eight parameters are different from each other, which indicates different degrees of parameter uncertainty for different parameters; (c) the posterior probability density distribution of the eight parameters is relatively concentrated. The maximum posterior estimations of the eight parameters are relatively close to the system value, which demonstrates the effectiveness of the DREAM method in the parameter identification of PTGS; (d) the marginal distribution intervals of $K_p$, $K_i$, $T_a$ and $e_n$ are very close to the system

FIGURE 4  Scaled-reduction score of each parameter of PTGS using the DREAM method under 90 and 20 dB noise scenarios (a) 90 dB, (b) 20 dB

FIGURE 5  Posterior probability distribution curves of the eight PTGS parameters estimated using DREAM under scenario 90 dB
value and generally fluctuates around the system value, which indicates that the uncertainty of the parameter estimation for these parameters using DREAM is relatively small.

From the comparison analysis of Figures 5 and 6, it can be found that the uncertainty degree of the parameter estimation results under scenario 20 dB is much larger than that under 90 dB, which demonstrates that the uncertainty degree will become larger along with the increase in the observation noises, and the width of the parameter marginal distribution interval generally increases accordingly. It is worth noting that the $K_d$, $T_y$, and $T_{yl}$ are irregularly distributed, and the posterior probability density distribution is relatively flat under scenario 20 dB, which indicates that different values of the parameters may obtain the same fitness value, that is, the uncertainty of these parameters increases significantly with the increasing of noises. The uncertainty of the identification results of $K_p$ and $T_a$ is relatively small under scenario 20 dB. The increasing trend of the uncertainty is not obvious along with the increasing of the noises. It should be emphasized that the maximum posterior estimation of $K_d$, $T_y$, and $T_{yl}$ gradually deviate from the system value, indicating that the non-linear characteristics of the separate modules of PTGS can limit the parameter identification accuracy of PTGS.

Through the above analysis, it is found that the observation data noise is an important factor that causes the uncertainty of the parameters. It should be pointed out that the above experiments set the search space of the parameters to be identified as [50%, 150%] of the system value. However, it may not be possible to accurately obtain the search space in practical engineering applications. In addition, with the increase of the unit operation life and the degradation of the equipment performance, the parameters may also shift. Therefore, the uncertainty of the prior search space of the parameters is also worth further research and discussion. In order to simulate the influence of the prior knowledge on the parameter uncertainty, this study adjusts the search space of the decision variable to [10%, 300%] of the system value. The noise scenarios are still set to 90 and 20 dB, and the posterior probability distribution curves of PTGS obtained by the DREAM method are shown in Figures 7 and 8 for scenarios 90 and 20 dB, respectively.

Compare Figures 5 with 7, it can be found that with the expansion of the prior search space, the width of the posterior probability density distribution interval of the parameters increases accordingly, and the maximum posterior estimates of $K_d$, $T_y$, and $T_{yl}$ gradually deviate from the system value. It can also be found from the comprehensive analysis of Figures 5–8 that the parameter uncertainty degree increases significantly under the dual effects of the expansion of the prior search space and the noise interference of the observed signal. Although the parameter uncertainty increases with the increase of the prior search space and the noise interference of the observed signal, the posterior probability density distribution of the parameters obtained by the DREAM method is still relatively accurate and compact. That is to say, the uncertainty analysis results can narrow the search space of the decision variables. Based on the contracted search space, refined parameter identification of PTGS can be further researched using intelligent algorithms.

### 4.2.3 Statistical performance evaluation

The posterior mean (mean) and the maximum posterior estimation (max) can be regarded as the identified value of the PTGS parameters. Therefore, five indices that include the max value, the ARE between max and the system value ($\text{ARE}_{\text{map}}$), the mean value, the ARE between mean and the system value ($\text{ARE}_{\text{mean}}$), and the standard deviation (SD) of the posterior
probability density distributions are calculated to evaluate the parameter identification results of PTGS using DREAM. The parameter identification results for PTGS with [50%, 150%] and [10%, 300%] search spaces are given in Tables 2 and 3, respectively. To evaluate the performance of the point identification results of the DREAM algorithm, the point identification results of the IBSA algorithm [11] have also been provided in Table 4.

By comparing the results of Tables 2 with 4, it can be found that the mean and max of $K_p$, $K_i$, $T_w$, $T_y$ and $e_a$ are similar to the point identification results based on the IBSA algorithm. However, the identification results of $T_p$, $T_d$ and $T_w$ are worse than those of the IBSA algorithm. From Figure 5, it is obvious that the posterior probability distribution curves of $T_p$, $T_d$ and $T_w$ are relatively wide and contain rich information, so the results are not as accurate and direct as the identification results based on IBSA.

It can be concluded from Tables 2 to 4 that: (a) The mean and max of most PTGS parameters obtained by DREAM are similar to the point identification results obtained by the intelligent
A real PSU in China is simulated and the PTGS parameters are estimated using the DREAM method. The main results and conclusions of this study are summarized as follows: (a) The marginal distribution intervals are relatively small compared with the evolutionary algorithm, which demonstrates the effectiveness of the DREAM method to a certain extent; (b) the SD values of the posterior distributions for scenario 90 dB are much smaller than those of scenario 20 dB, which demonstrates that the parameter identification accuracy of PTGS is greatly affected by the observation noises; (c) the SD values of the posterior distributions for [10%, 300%] search space are much larger than those for [50%, 150%] search space, which demonstrates that narrower and more accurate prior search space is beneficial to enhance the identification accuracy of PTGS using DREAM.

TABLE 2 Parameter identification results for [50%, 150%] search space

| Parameter | System value | Max ARE\(_{\text{max}}\) (%) | Mean ARE\(_{\text{mean}}\) (%) | SD | Max ARE\(_{\text{max}}\) (%) | Mean ARE\(_{\text{mean}}\) (%) | SD |
|-----------|--------------|----------------------------|----------------------------|----|----------------------------|----------------------------|----|
| K\(_p\)    | 8.7          | 8.703                     | 0.037                      |     | 8.704                     | 0.047                      | 0.019 |
| K\(_f\)    | 0.35         | 0.350                     | 0.005                      |     | 0.350                     | 0.001                      | 0.000 |
| K\(_d\)    | 1.9          | 2.011                     | 5.842                      |     | 1.991                     | 4.790                      | 0.427 |
| T\(_s\)    | 0.2          | 0.207                     | 3.256                      |     | 0.205                     | 2.644                      | 0.026 |
| T\(_c\)    | 0.02         | 0.020                     | 1.364                      |     | 0.020                     | 0.131                      | 0.002 |
| T\(_c\)    | 1.3          | 1.300                     | 0.023                      |     | 1.300                     | 0.017                      | 0.001 |
| T\(_c\)    | 10.8         | 10.801                    | 0.009                      |     | 10.798                    | 0.015                      | 0.011 |
| \(\epsilon\) | 1            | 1.000                     | 0.001                      |     | 1.000                     | 0.017                      | 0.001 |

Note: ARE is the absolute relative error.

TABLE 3 Parameter identification results for [10%, 300%] search space

| Parameter | System value | Max ARE\(_{\text{max}}\) (%) | Mean ARE\(_{\text{mean}}\) (%) | SD | Max ARE\(_{\text{max}}\) (%) | Mean ARE\(_{\text{mean}}\) (%) | SD |
|-----------|--------------|----------------------------|----------------------------|----|----------------------------|----------------------------|----|
| K\(_p\)    | 8.7          | 8.757                     | 0.652                      |     | 8.810                     | 1.262                      | 0.317 |
| K\(_f\)    | 0.35         | 0.350                     | 0.122                      |     | 0.344                     | 0.172                      | 0.034 |
| K\(_d\)    | 1.9          | 3.576                     | 88.209                     |     | 5.178                     | 172.500                    | 1.673 |
| T\(_s\)    | 0.2          | 0.300                     | 50.232                     |     | 0.353                     | 76.716                     | 0.227 |
| T\(_c\)    | 0.02         | 0.029                     | 45.822                     |     | 0.033                     | 66.935                     | 0.036 |
| T\(_c\)    | 1.3          | 1.300                     | 0.004                      |     | 1.414                     | 8.799                      | 1.336 |
| T\(_c\)    | 10.8         | 10.786                    | 0.131                      |     | 9.804                     | 9.220                      | 3.894 |
| \(\epsilon\) | 1            | 1.001                     | 0.122                      |     | 0.948                     | 5.213                      | 0.940 |

TABLE 4 Parameter identification results based on improved backtracking search algorithm for [50%, 150%] search space [11]

| Parameter | System value | Scenario 90 dB | Scenario 20 dB |
|-----------|--------------|---------------|---------------|
| K\(_p\)    | 8.7          | 8.703         | 8.542         |
| K\(_f\)    | 0.35         | 0.356         | 0.341         |
| K\(_d\)    | 1.9          | 1.917         | 2.234         |
| T\(_s\)    | 0.2          | 0.198         | 0.172         |
| T\(_c\)    | 0.02         | 0.02          | 0.026         |
| T\(_c\)    | 1.3          | 1.293         | 1.387         |
| T\(_c\)    | 10.8         | 10.883        | 10.583        |
| \(\epsilon\) | 1            | 0.996         | 0.979         |

5 CONCLUSIONS

Parameter identification of PTGS has received a lot of attention in recent years to model the PSU accurately. However, recent parameter identification methods are mostly based on optimisation theory, which can only obtain a set of optimal parameters and cannot quantify the uncertainty of the parameters. This study employs an MCMC algorithm based on Bayesian theory, that is, the DREAM method, for efficient parameter identification and accurate uncertainty quantification for PTGS. A real PSU in China is simulated and the PTGS parameters are estimated using the DREAM method. The main results and conclusions of this study are summarized as follows: (a) The marginal distribution intervals are relatively small compared with the...
prior search range with uniform distribution and the posterior probability density distribution of the eight parameters is relatively concentrated, which indicates that the eight PTGS parameters are reasonably well identifiable; (b) the maximum posterior estimations of the eight parameters are relatively close to the system value. The posterior mean and the maximum posterior estimation of most PTGS parameters obtained by DREAM are similar to the point identification results obtained by intelligent evolutionary algorithm [11], which demonstrates the effectiveness of the DREAM method; (c) the uncertainty degree will become larger along with the increase of the observation noises, and the width of the parameter marginal distribution interval generally increases accordingly; (d) the uncertainty of the prior search space of the parameters has a great influence on the parameter identification accuracy of PTGS. The DREAM algorithm can narrow the prior search space of PTGS parameters effectively and thus help the hydropower engineers to make decisions.

After all, it can be concluded from the results and conclusions of this study that the DREAM method can be employed as an effective tool for parameter identification and accurate uncertainty quantification of PTGS. On the one hand, the DREAM algorithm can identify the parameters of PTGS with unknown prior search space and noise; on the other hand, it can also reduce the search space of PTGS parameters. The above research framework of parameter identification and uncertainty analysis not only brings new ideas and methods for hydropower research but also has certain enlightenment significance for the parameter estimation and uncertainty analysis of complex non-linear systems in other industries. Further research will be focused on applying more parameter uncertainty methods on PTGS and compare the performances of different uncertainty analysis methods for parameter identification of PTGS on real-world simulations.

ACKNOWLEDGEMENTS
This work is supported by the Natural Science Foundation of Jiangsu Province (No. BK20191052), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 19KJB470012), the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 19KJB480007), the Construction system of Jiangsu Province of China (No.20200037). Special thanks are given to the Huai’an Innovative and Entrepreneurial Talents Project, and the Hua’ian Innovative Talents Project.

REFERENCES
1. Lai, X., et al.: A multi-objective optimization strategy for the optimal control scheme of pumped hydropower systems under successive load rejections. Appl. Energy 261, 114474 (2020)
2. Zheng, Y., et al.: Hierarchical MPC scheme for the speed governing of psu with complex conduit system. IET Gener. Transm. Distrib. 14(2), 316–329 (2020)
3. Wu, X., et al.: Characteristics analysis and fuzzy fractional-order pid parameter optimization for primary frequency modulation of a pumped storage unit based on a multi-objective gravitational search algorithm. Energies 13, (1), 137 (2020)
4. Guo, P., Zhang, H., Gou, D: Dynamic characteristics of a hydro-turbine governing system considering draft tube pressure pulsation. IET Renewable Power Gener. 14(7), 1210–1218 (2020)
5. Xu, Y., et al.: Adaptive condition predictive-fuzzy pid optimal control of start-up process for pumped storage unit at low head area. Energy Convers. Manage. 177, 592–604 (2018)
6. Zhao, Z., et al.: A coordinated optimization framework for flexible operation of pumped storage hydropower system: Nonlinear modeling, strategy optimization and decision making. Energy Convers. Manage. 194, 75–93 (2019)
7. Li, C., et al.: Design of a fractional-order pid controller for a pumped storage unit using a gravitational search algorithm based on the cauchy and gaussian mutation. Inform. Sci. 396, 162–181 (2017)
8. Zhang, C., et al.: Modeling and synchronization optimization of pump turbine governing system using sparse robust least squares support vector machine and hybrid backtracking search algorithm. Energies 11(11), 3108 (2018)
9. Zhang, C., et al.: An improved autoencoder and partial least squares regression-based extreme learning machine model for pump turbine characteristics. Appl. Sci. 9(19), 3987 (2019)
10. Chen, Z., et al.: Improved gravitational search algorithm for parameter identification of water turbine regulation system. Energy Convers. Manage. 78(30), 306–315 (2014)
11. Zhou, J., et al.: Parameter identification of pump turbine governing system using an improved backtracking search algorithm. Energies 11(7), 1668 (2018)
12. Xu, Y., et al.: A parameter adaptive identification method for a pumped storage hydro unit regulation system model using an improved gravitational search algorithm. Simulation 93(8), 679–694 (2017)
13. Tian, Y., et al.: Takagi–Sugeno fuzzy generalised predictive control of a time-delay non-linear hydro-turbine governing system. IET Renewable Power Gener. 13(13), 2338–2345 (2019)
14. Liu, J., et al.: Grid connection analysis of hydro-turbine generator unit with stochastic disturbance. IET Renewable Power Gener. 13(3), 500–509 (2019)
15. Chen, Q., et al.: Multi-objective heuristic guide vane closure scheme optimisation of hydroturbine generating unit. IET Renewable Power Gener. 14, 1087–1094, (2020)
16. Li, C., Zhou, J.: Parameters identification of hydraulic turbine governing system using improved gravitational search algorithm. Energy Convers. Manage. 52(1), 374–381 (2011)
17. Zhang, N., et al.: A mixed-strategy based gravitational search algorithm for parameter identification of hydraulic turbine governing system. Knowledge-Based Syst. 109, 218–237 (2016)
18. Tian, T., et al.: An improved ant lion optimization algorithm and its application in hydraulic turbine governing system parameter identification. Energies 11(1), 95 (2018)
19. Sheng, M., et al.: Comparison of glue and dream for the estimation of cultivar parameters in the apsim-maize model. Agric. For. Meteorol. 278, 107569 (2019).
20. Xu, B., et al.: Mathematical model and parametric uncertainty analysis of a hydraulic generating system. Renewable Energy 136, 1217–1230 (2019)
21. Xu, B., et al.: Parametric uncertainty in affecting transient characteristics of multi-parallel hydropower systems in the successive load rejection. Int. J. Electr. Power Energy Syst. 106, 444–454 (2019)
22. Jia, H., et al.: Bayesian framework of parameter sensitivity, uncertainty, and identifiability analysis in complex water quality models. Environ. Model. Software 104, 13–26 (2018)
23. Guo, J., et al.: Uncertainty assessment and optimization of hydrological model with the shufflet complex evolution metropolis algorithm: An application to artificial neural network rainfall-runoff model. Stochastic Environ. Res. Risk Assess. 27(4), 985–1004 (2013)
24. Ma, J., Huang, B., Ding, F.: Parameter estimation of markov-switching hammerstein systems using the variational bayesian approach. IET Control Theory Appl. 13(11), 1646–1655 (2019)
25. Xu, Y., et al.: A bayesian approach to real-time dynamic parameter estimation using PMU measurement. IEEE Trans. Power Syst. 35(2), 1109–1119 (2019)
26. Zheng, Y., Han, F.: Markov chain Monte Carlo (MCMC) uncertainty analysis for watershed water quality modeling and management. Stochastic Environ. Risk Assess. 30(1), 293–308 (2016)

27. Zahmatkesh, Z., Karamouz, M., Nazif, S.: Uncertainty based modeling of rainfall-runoff: Combined differential evolution adaptive metropolis (DREAM) and K-means clustering. Adv. Water Resour. 83, 405–420 (2015)

28. Liu, Y., et al.: A bayesian-based multilevel factorial analysis method for analyzing parameter uncertainty of hydrological model. J. Hydrol. 553, 750–762 (2017)

29. Balkaya, Ç.: An implementation of differential evolution algorithm for inversion of geoelectrical data. J. Appl. Geophys. 98, 160–175 (2013)

30. Ren, H., et al.: Bayesian inversion of seismic and electromagnetic data for marine gas reservoir characterization using multi-chain markov chain monte carlo sampling. J. Appl. Geophys. 147, 68–80 (2017)

31. Zhang, C., et al.: Multiobjective optimization of a fractional-order pid controller for pumped turbine governing system using an improved nageli algorithm under multiworking conditions. Complexity 2019(1) 1–18 (2019)

32. Lai, X., et al.: Stability and dynamic characteristics of the nonlinear coupling system of hydropower station and power grid. Commun. Nonlinear Sci. Numer. Simul. 79, 104919 (2019)

33. Vrugt, J.A., et al.: Accelerating markov chain monte carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Int. J. Nonlinear Sci. Numer. Simul. 10(3), 273–290 (2011)

34. Gelman, A., Rubin, D.B.: Inference from iterative simulation using multiple sequences. Statist. Sci. 7(4), 457–472 (1992)

How to cite this article: Zhang C, Peng T, Zhou J, Nazir MS. Parameter identification and uncertainty quantification of a non-linear pump-turbine governing system based on the differential evolution adaptive Metropolis algorithm. IET Renew Power Gener. 2021;15:342–353. https://doi.org/10.1049/rpg2.12027