

A new approach to identify the sensitivity and importance of physical parameters combination within numerical models using the Lund–Potsdam–Jena (LPJ) model as an example

Guodong Sun¹ · Mu Mu^{2,1}

Received: 13 July 2015 / Accepted: 29 November 2015 / Published online: 14 January 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract An important source of uncertainty, which causes further uncertainty in numerical simulations, is that residing in the parameters describing physical processes in numerical models. Therefore, finding a subset among numerous physical parameters in numerical models in the atmospheric and oceanic sciences, which are relatively more sensitive and important parameters, and reducing the errors in the physical parameters in this subset would be a far more efficient way to reduce the uncertainties involved in simulations. In this context, we present a new approach based on the conditional nonlinear optimal perturbation related to parameter (CNOP-P) method. The approach provides a framework to ascertain the subset of those relatively more sensitive and important parameters among the physical parameters. The Lund–Potsdam–Jena (LPJ) dynamical global vegetation model was utilized to test the validity of the new approach in China. The results imply that nonlinear interactions among parameters play a key role in the identification of sensitive parameters in arid and semi-arid regions of China compared to those in northern, north-eastern, and southern China. The uncertainties in the numerical simulations were reduced considerably by reducing the errors of the subset of relatively more sensitive and important parameters. The results demonstrate that our approach not only offers a new route to identify relatively more sensitive

and important physical parameters but also that it is viable to then apply “target observations” to reduce the uncertainties in model parameters.

1 Introduction

The predictability and uncertainty of weather and climate are hot topics in atmospheric and oceanic sciences. The uncertainties involved in numerical simulations and predictions are rooted not only in the uncertainty of initial error (Morris and Battisti 2004; Aberson 2011) but also in model error (Williams et al. 2001; Berthelot et al. 2005; Carrassi and Vannitsem 2011; Jarvinen et al. 2012; Hally et al. 2013; Wan et al. 2012) and belong to the “first kind” and “second kind” categories of weather and climate predictability problems (Mu et al. 2002). The uncertainties in model errors mainly arise not only from the mathematical descriptions of earth system processes (Cramer et al. 2001) but also from the uncertainties related to physical parameters in the model (Zaehle et al. 2005). Reducing the uncertainties in physical parameter errors in numerical models through observations, optimization methods, or data assimilation is thus a crucial area of research. Efforts in this area will help to improve the simulation ability and forecasting skill of such models in atmospheric and oceanic science studies (Lu and Hsieh 1997; Janiskova and Morcrette 2005; Pulido et al. 2012; Smith et al. 2013).

There are many physical parameters included in numerical models to describe the physical processes of atmosphere–ocean coupled general circulation models. It would be costly and impractical to reduce the uncertainties of all the physical parameter errors through observation and other methods. Instead, a more effective route is to choose certain key physical parameters, which are referred to be “sensitive and important” to improve the simulation abilities and forecasting

✉ Mu Mu
mumu@qdio.ac.cn

¹ State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

² Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

skill of models. There has been much discussion about how to identify the sensitivity and importance of parameters (White et al. 2000; Knorr and Heimann 2001; Li et al. 2012; Pappas et al. 2013; Wang et al. 2013). Pitman (1994) analyzed the percentage change of 18 physical parameters using the Biosphere Atmosphere Transfer Scheme (BATS), a land surface model, to reveal their relative importance. When the authors analyzed the percentage change of one parameter, the other 17 parameter values were fixed. The interaction among parameters was ignored during the process (Jackson et al. 2003; Bastidas et al. 2006). Other methods have also been employed to explore the sensitivity of certain physical parameters, such as the adjoint method, factorial experimentation, and the Multi-Objective Generalized Sensitivity Analysis (MOGSA) method (Henderson-Sellers 1992; Wang et al. 2001; Rayner et al. 2005; Bastidas et al. 2006; Rosero et al. 2010). The adjoint method is based on linear approximation, and while it may be valid for small parameter errors over short time periods, it is less applicable to large parameter errors and long integration times. The factorial experimentation, Monte Carlo, and MOGSA methods could be applied to consider the impact of interactions among physical parameters through producing a stratified sample under range of parameters values, as well as to identify key physical parameters (Zaehle et al. 2005; Bastidas et al. 2006). Henderson-Sellers (1992) employed factorial experimentation to show the rank of parameters using the BATS model. The top five most sensitive parameters are vegetation roughness length, vegetation albedo $< 0.7 \mu\text{m}$, maximum leaf-area index, vegetation albedo $\geq 0.7 \mu\text{m}$, and saturated soil hydrologic conductivity. Zaehle et al. (2005) applied a Monte Carlo-type stratified sampling approach to identify the sensitive and important physical parameters in a model and provide reasonable estimations of a model output variable. They used Latin hypercube sampling method to create random sample of parameter value. Based on the random sample, the parameters sensitivity and importance are determined by calculating ranked partial correlation coefficients (RPCC) to estimate the uncertainty contribution of a particular parameter to the total model output uncertainty. They reported that the top five most sensitive and important physical parameters were α_{C3} , α_a , θ , g_m , and r_{growth} for net primary production (NPP). Bastidas et al. (1999) identified which of the parameters within a model was sensitive using the MOGSA method. Their studies showed the number of sensitive parameters according to significance levels below 1 %, from 1 to 5 %, and above 5 %.

In this work, a new approach was established to ascertain which subset or combination of relatively more sensitive and important parameters causes maximum uncertainty in numerical simulation and forecast results because the parameters combination should be considered to reduce the uncertainty of numerical simulation, such as the data assimilation method and the optimal method (Vrugt et al. 2005; De Lannoy et al.

2006). After first providing an overview of the new approach in Section 2, we then describe its methodology in more detail in Section 3, as well as the experimental procedures used to test it. The results from the experiments are reported in Section 4, and a discussion and summary of the key findings are presented in Section 5.

2 Overview of the new approach

2.1 The new approach

To ascertain the combination of relatively more sensitive and important parameters that cause maximum uncertainty in numerical simulation and forecast results, we propose a new approach, which consists of the following three steps.

The first step is to choose the physical parameters. Generally, there are three types of parameters in a numerical model. Taking the Lund–Potsdam–Jena (LPJ) model as our example, the normalizing coefficient for the exponential distribution, which is randomly created and related to computational stability, and unrelated to observations, belongs to the first type of parameter. Wood density is obtained through direct observation data and falls into the second type of parameter. Finally, the parameters in the allometric equations are achieved using indirect observations of leaf area index and tree height and hence pertain to the third type of parameter. The last two types of parameters describe physical processes and can be obtained through observations. In this study, we only consider parameters that can be obtained through direct or indirect observations.

The second step is to examine the sensitivity of single physical parameter, of which there are many in numerical models. Supposing there are n physical parameters obtained by observation data, some of the less sensitive n parameters first need to be eliminated through implementing sensitivity tests of every physical parameter. There are two reasons to implement this step. Firstly, the cost in terms of human and material resources needed to identify the combination of relatively more sensitive and important physical parameters among all parameters would be enormous. Secondly, not all of the n parameters will cause large uncertainties in the numerical simulation. In the LPJ model, if $n = 24$, and the combination approach is applied, we are able to find the most significant combination of five relatively more sensitive and important parameters among the total of 24 (Table 1). Those five parameters are identified through C_{24}^5 optimization experiments, and we find that the computational cost is high. We also find, by implementing simple sensitivity experiments using the LPJ model, that not all of the 24 parameters bring large uncertainty to the numerical simulation. For example, the extent of uncertainty in simulated NPP is $72.55 \text{ g C m}^{-2} \text{ year}^{-1}$ when the g_m^* parameter error is 3.05, which is

Table 1 The chosen parameters within the LPJ model

| Number | Parameter | Standard | Minimum | Maximum | Description |
|--------|-----------------------------|----------|---------|---------|---|
| 1 | θ^* | 0.7 | 0.2 | 0.996 | Co-limitation shape parameter |
| 2 | α_a^* | 0.5 | 0.3 | 0.7 | fraction of PAR assimilated at ecosystem level relative to leaf level |
| 3 | $\lambda_{\max, C3}$ | 0.7 | 0.6 | 0.8 | Optimal $ci = ca$ for C3 plants (all PFTs except TrH) |
| 4 | α_{C3}^* | 0.08 | 0.02 | 0.125 | Intrinsic quantum efficiency of CO ₂ uptake in C3 plants |
| 5 | a_{C3}^* | 0.015 | 0.01 | 0.021 | Leaf respiration as a fraction of Rubisco capacity in C3 plants |
| 6 | $Q_{10, ko}$ | 1.2 | 1.1 | 1.3 | q10 for temperature-sensitive parameter ko |
| 7 | $Q_{10, kc}$ | 2.1 | 1.9 | 2.3 | q10 for temperature-sensitive parameter kc |
| 8 | $Q_{10, \tau}$ | 0.57 | 0.47 | 0.67 | q10 for temperature-sensitive parameter tau |
| 9 | r_{growth}^* | 0.25 | 0.15 | 0.4 | Growth respiration per unit <i>NPP</i> |
| 10 | g_m^* | 3.26 | 2.5 | 18.5 | Maximum canopy conductance analog [mm day ⁻¹] |
| 11 | α_m | 1.391 | 1.1 | 1.5 | Evapotranspiration parameter |
| 12 | k_{allom1} | 100 | 75 | 125 | Crown area = $k_{\text{allom1}} * \text{height}^{**} / k_{\text{rp}}$ |
| 13 | k_{allom2} | 40 | 30 | 50 | Height = $k_{\text{allom2}} * \text{diameter}^{**} / k_{\text{allom3}}$ |
| 14 | k_{allom3} | 0.67 | 0.5 | 0.8 | Height = $k_{\text{allom2}} * \text{diameter}^{**} / k_{\text{allom3}}$ |
| 15 | $k_{\text{la:sa}}^*$ | 6000 | 2000 | 8000 | Leaf-to-sapwood area ratio |
| 16 | k_{rp} | 1.5 | 1.37 | 1.6 | Crown area = $k_{\text{allom1}} * \text{height}^{**} / k_{\text{rp}}$ |
| 17 | k_{mort1}^* | 0.01 | 0.005 | 0.1 | Asymptotic maximum mortality rate [year ⁻¹] |
| 18 | k_{mort2} | 0.4 | 0.2 | 0.5 | Growth efficiency mortality scalar |
| 19 | $\text{est}_{\text{max}}^*$ | 0.24 | 0.05 | 0.48 | Maximum sapling establishment rate [m ² /year] |
| 20 | n_0 | 7.15 | 6.85 | 7.45 | Leaf N concentration (mg/g) not involved in photosynthesis |
| 21 | $\text{dens}_{\text{wood}}$ | 200 | 180 | 220 | Specific wood density [kg C/m ³] |
| 22 | τ_{litter}^* | 0.35 | 0.19 | 0.81 | Litter turnover time at 10 °C [year] |
| 23 | p_t | 1.32 | 1.12 | 1.52 | Priestley-Taylor coefficient |
| 24 | β | 0.17 | 0.15 | 0.19 | Global average short-wave albedo |

added to the standard value, using the LPJ model while the other parameters remain unchanged (Table 1). However, the extent of uncertainty in simulated NPP is just 2.48 g C m⁻² year⁻¹ when the $\text{est}_{\text{max}}^*$ parameter error is -0.04, which is added to the standard value. In a previous study, Zaehle et al. (2005) first identified 12 more sensitive and important parameters among a total of 36. Hence, it is necessary to implement this step to eliminate some of the less sensitive parameters. To accomplish the second step, the conditional nonlinear optimal perturbation related to parameter error approach (CNOP-P; Mu et al. 2010) is employed. A detailed explanation of the CNOP-P approach is provided in the following section. The n parameters are optimized individually using the CNOP-P approach. The CNOP-Ps and their cost function values for every parameter are obtained within the range of reasonable parameter error. The physical explanation of the cost function induced by the CNOP-P is the maximal extent of uncertainty in numerical simulations caused by the optimal value. The sensitivity of

every parameter can be identified according to the cost function value, and the more sensitive parameters can be chosen based on an appropriate threshold, i.e., we choose m physical parameters where $n > m$. The advantage of using the CNOP-P approach is to identify the sensitivity of every parameter, which in theory is the optimal way to ensure the ranking of parameters in terms of their sensitivity. However, it ignores the impact of interactions among parameters on the uncertainty of the numerical simulation. Therefore, this step alone does not identify the key combination of relatively more sensitive and important parameters.

The third step is to examine the sensitivity and importance of multiple parameters together, i.e., to reveal the most significant combination of parameters among all parameters previously identified in the second step. The aim of this step is to find the relatively more sensitive and important parameters (k) among the previously identified m parameters, where $m > k$. Continuing with the LPJ model as our example, we suppose

$n = 24$, $m = 10$, and $k = 5$ (i.e., 10 parameters were chosen following the second step). In the third step, we want to find the combination of five relatively more sensitive and important parameters among those 10 that cause the greatest uncertainty in the numerical simulation. To begin, $C_{10}^5 = 252$ groups of parameters combinations are built. For these groups, 252 CNOP-Ps and their cost function values are obtained using the CNOP-P approach within the range of reasonable parameter error. The parameters combination causing the maximal cost function value among the 252 cost function values is regarded as the relatively more sensitive and important subset of five parameters among the total of 10. The experiments in this step are different to those conducted in the second step (see Fig. 1), and in carrying out this step, the impact of nonlinear interactions among parameters on the uncertainty of the numerical simulation can be considered.

2.2 The CNOP-P approach

In the above new approach, the maximal uncertainty in numerical simulations is obtained within the range of reasonable parameter error in both single-parameter and multi-parameter experiments using the CNOP-P approach. The CNOP-P is a type of parameter perturbation that could cause the maximal cost function with a certain constraint and at an optimal time. This type of parameter perturbation, which could lead to the maximal uncertainty in numerical simulation and forecasting results, is a parameter error or parameter error combination. The advantages of the CNOP-P approach are not only its ability to obtain the parameter error combination causing the maximal uncertainty but also that it can be used to consider the impact of nonlinear interactions among parameters on the level of uncertainty.

In the work of Mu et al. (2010), the CNOP-P approach was proposed according to types of predictability, and it has been applied to study ENSO predictability, estimations of terrestrial ecosystems, and the Kuroshio large meander (KLM) (Mu et al. 2010; Duan and Zhang 2010; Sun and Mu 2012a, b; Wang et al. 2012). Yu et al. (2012) analyzed the roles of initial error and model error in generating a significant spring predictability barrier (SPB) for El Nino events using the CNOP-P approach and noted that initial errors play a more important role than parameter errors in causing a significant SPB for El Nino events. In addition, Wang et al. (2012) discussed the impact of model error on the KLM. They found that not only did the initial condition errors have greater effects on the prediction of the KLM than errors in model parameters but also that the latter cannot be ignored. We now review the derivation of the CNOP-P approach for the readers' convenience.

Let the nonlinear differential equations be as follows:

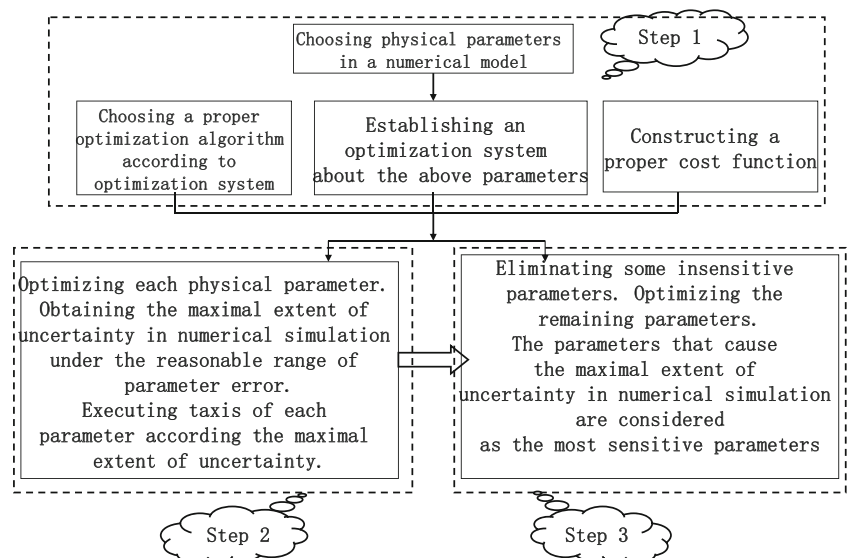
$$\begin{cases} \frac{\partial U}{\partial t} = F(U, P) & U \in R^n, t \in [0, T] \\ U|_{t=0} = U_0 \end{cases} \quad (1)$$

where F is a nonlinear operator, P is a parameter vector in Eq. (1), and U_0 is an initial value. Let M_τ be the propagator of the nonlinear differential equations from the initial time 0 to τ . u_τ is a solution of the nonlinear equations at time τ and satisfies $u(\tau) = M_\tau(u_0, p)$.

Let $U(T; U_0, P)$ and $U(T; U_0, P) + u(T; U_0, p)$ be the solutions of the nonlinear differential equations (1) with P and $P + p$, respectively, where P and p are parameter vectors. $u(T; U_0, p)$ describes the departure from the reference state $U(T; U_0, P)$ caused by p . The solutions satisfy:

$$\begin{cases} U(T; U_0, P) = M_T(U_0, P) \\ U(T; U_0, P) + u(T; U_0, p) = M_T(U_0, P + p) \end{cases} \quad (2)$$

Fig. 1 Flowchart depicting the steps involved in the new method



For a proper norm $\|\cdot\|$, a parameter perturbation p_δ is called a CNOP if and only if

$$J(p_\delta) = \max_{p \in \Omega} J(p), \tag{2}$$

where

$$J(p) = \left\| M_T(U_0, P + p) - M_T(U_0, P) \right\| \tag{3}$$

P is a reference state of the parameters in the Eq. (1), and p is the perturbation of the reference state. In the second and third steps, the dimension of the P is 1 and 5. $p \in \Omega$ is a constraint condition. The CNOP-P is the parameter perturbation whose nonlinear evolution attains the maximum value of the cost function J at time T .

3 Experimental procedures and model

3.1 Experimental design

Twenty-four parameters within the LPJ model were chosen for examination based on the study by Zaehle et al. (2005). Table 1 shows the physical meanings, standard values, and minimum and maximum values for all parameters. As we know, parameter values differ; therefore, for the convenience of data processing and implementing nonlinear optimization, the chosen parameters were normalized using linear transformations. The physical parameters were mapped into the range of -1 to 1 . The simple linear piecewise function was applied to normalize the physical parameters as follows:

$$\begin{cases} y = \frac{x - \text{Stavalue}}{\text{Maxvalue} - \text{Stavalue}} & \text{when } x \geq \text{Stavalue} \\ y = \frac{x - \text{Stavalue}}{\text{Stavalue} - \text{Minvalue}} & \text{when } x < \text{Stavalue} \end{cases}$$

where x and y are the values of front and rear transformation and Stavalue, Maxvalue, and Minvalue are the standard, maximum, and minimum values, respectively, of the physical parameters. When $x = \text{Minvalue}$, $y = -1$; when $x = \text{Maxvalue}$, $y = 1$. When $x = \text{Stavalue}$, $y = 0$. The constraint condition in Eq. (2), $p \in \Omega$, is a box constraint. The constraint condition parameter is 0.2 for parameters ($|p| \leq \delta$, $\delta = 0.2$). When $\delta = 1$, the parameter errors are also reasonable. However, the parameter errors will cause the terrestrial ecosystem to be unstable using the CNOP-P approach. Four study regions were chosen in China (northern, northeastern, southern, and arid/semi-arid); and 24 cases, whose longitude and latitudinal are shown in Table 2, were chosen as the study region in the below numerical results.

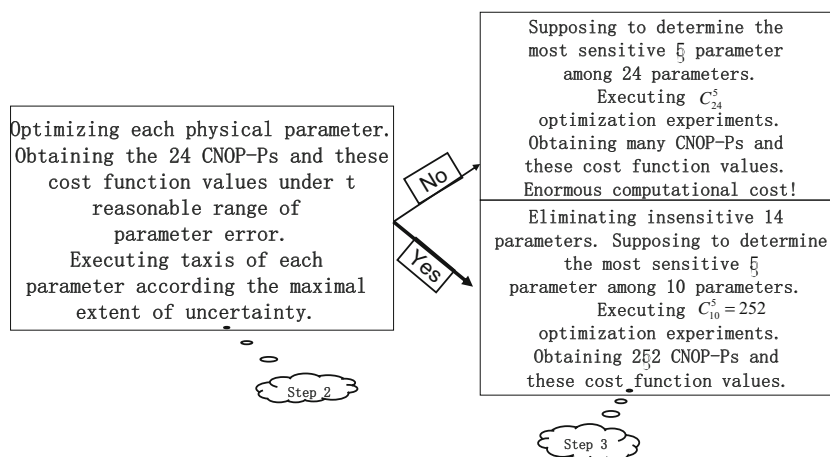
Figure 2 shows the detailed experimental design used to identify the relatively more sensitive and important parameters combination. Taking annual NPP as the variable in the cost function, we supposed $n = 24$, $m = 10$,

Table 2 The sensitivity of single parameter among 10 parameters using the CNOP-P approach (*numbers* represent the sequence number of single parameter as in Table 1, and 125.75 45.75 means 125.75°E 45.75°N. There are similar for other locations)

| Number | Location |
|--------|---|
| 1 | 125.75 45.75 10 4 2 9 11 1 8 17 23 5 |
| 2 | 125.75 46.25 10 4 2 9 1 11 17 8 23 19 |
| 3 | 125.75 46.75 10 4 2 9 1 11 8 17 5 24 |
| 4 | 126.25 45.75 4 10 2 9 1 11 8 5 23 17 |
| 5 | 126.25 46.25 4 10 2 9 1 11 15 8 5 17 |
| 6 | 126.25 46.75 4 10 2 9 1 17 8 11 5 23 |
| 7 | 126.75 45.75 4 10 2 9 1 8 11 17 5 15 |
| 8 | 126.75 46.25 4 10 2 9 1 11 8 17 5 15 |
| 9 | 126.75 46.75 4 10 2 9 1 17 8 15 11 5 |
| 10 | 115.75 26.25 4 2 1 9 10 17 5 8 24 19 |
| 11 | 115.75 26.75 4 2 1 17 9 10 5 24 8 19 |
| 12 | 115.75 27.25 4 2 1 17 9 16 15 10 11 5 |
| 13 | 115.75 32.25 4 2 10 1 9 5 24 8 19 15 |
| 14 | 115.75 32.75 4 2 1 9 11 10 5 19 23 8 |
| 15 | 115.75 33.25 4 2 1 9 10 12 16 17 15 23 |
| 16 | 115.75 36.25 11 10 4 9 2 24 16 19 13 14 |
| 17 | 115.75 36.75 11 10 4 2 5 8 1 24 9 16 |
| 18 | 115.75 37.25 11 10 4 2 1 5 24 16 12 9 |
| 19 | 116.25 36.25 4 2 10 11 17 9 1 5 16 14 |
| 20 | 116.25 36.75 11 10 4 2 9 1 24 17 16 19 |
| 21 | 116.25 37.25 11 10 2 9 4 1 16 5 8 14 |
| 22 | 116.75 36.25 4 2 1 10 11 16 5 14 24 19 |
| 23 | 116.75 36.75 16 2 4 5 10 24 9 1 8 11 |
| 24 | 116.75 37.25 10 11 4 9 14 2 8 24 16 19 |

$k = 5$ and wanted to find the five relatively more sensitive and important parameters among the 10 parameters. First, 10 physical parameters were chosen from the original 24 according to the second step introduced in Section 2. Next, we built $C_{10}^5 = 252$ groups of parameters combinations. For these groups, 252 CNOP-Ps and their cost function values were obtained using the CNOP-P approach within the range of reasonable parameter error. The five-parameter combination causing the maximal cost function value among the 252 cost function values was regarded as the most significant subset of relatively more sensitive and important parameters among the total of 10. In the second step, the sensitivity of single parameter was identified using the CNOP-P approach, which in theory is the optimal way to ensure the ranking of every parameter in terms of its sensitivity. For verification, we compared the CNOP-P approach with another method, the traditional one-at-a-time approach, to identify the sensitivity of every parameter (OAT; Pitman 1994; Saltelli 1999). The OAT approach supplied the variation due to the representative

Fig. 2 Detail of steps 2 and 3 of the new method



parameter perturbation value, such as $\pm 10\%$ or $\pm 20\%$. The other parameter values were fixed when the sensitivity of a certain parameter was identified. So, it is convenient to obtain the variation due to the perturbed parameter, and the computational cost is relatively small. The identification of the parameter sensitivity was performed according to the extent of variation in the numerical simulation. A finite difference method was employed to calculate the variation due to the perturbed parameter in the factor space to identify the sensitivity of the parameter. It was visible that the interaction among the parameters and the optimal settings of factors are ignored when the sensitivity of parameter is determined with the OAT approach. The parameter errors were ± 0.2 to run the LPJ model in our study, which were similar to those in the experiment using the CNOP-P approach.

In our studies, to obtain the maximum value of Eq. (2), an evolutionary algorithm (differential evolution (DE); Storn and Price 1997) was employed, because the cost function may be non-differentiable about the parameters as optimal variables. Some studies have applied evolutionary algorithms to investigate parameter estimations and uncertainties in land surface schemes and extreme events with the MM5 model and other models (Duan et al. 1992; Kruger, 1993; Zhang et al. 2000). The advantage of the evolutionary algorithms is that they obtain the optimal value of Eq. (2) without the gradient. The DE algorithm has been applied to explore terrestrial ecosystem responses to climate change, and details on the algorithm have been provided (Sun and Mu 2012a, b, 2013). In addition, the validity of the DE algorithm was checked before the it was applied to search for the optimal value with the LPJ model (Sun and Mu 2009). The CNOP-P was obtained by calculating Eq. (2), and an initial estimation value was given during the optimization process. In our study, to effectively obtain the CNOP-P, 12 random initial estimation

values were chosen. The final optimal value was repeatedly verified using the DE algorithm.

3.2 The LPJ model

We used the LPJ model in the present study as an example to validate the theoretical framework (Sitch et al. 2003). The LPJ model is process-based and can describe the dynamics of land processes in atmospheric and oceanic science studies, the carbon exchange between land and the atmosphere, and the hydrological cycle. The model, which originates from the biome model family (Prentice et al. 1992), can simulate the distribution of plant functional types (PFTs), with 10 PFTs used to distinguish different photosynthetic (C3, C4), phenological (deciduous, evergreen), and physiognomic (tree, grass) features. The parameters will be identified in all PFTs when the parameter sensitivity analysis is implemented. The LPJ model explicitly considers photosynthesis, mortality, fire disturbance, and soil heterotrophic respiration. Carbon is stored in seven PFT-associated pools in this model with leaves, sapwood, heartwood, fine roots, a fast and a slow decomposing above-ground litter pool, a below-ground litter pool, and two soil carbon pools for each grid cell. A detailed description and evaluation of the model can be found in Sitch et al. (2003). The LPJ model has been widely employed to discuss the variation in terrestrial ecosystems and the carbon cycle (Werner et al. 2007; Hickler et al. 2008), and its simulation of PFTs has been shown to be in agreement with the observations in China. However, owing to a lack of observational data regarding NPP in China, the NPP simulated using the LPJ model relies on this simulation using other models at similar spatial and temporal scales. Nevertheless, numerical results indicate that the LPJ model can be employed to examine variations in terrestrial ecosystems in response to climate change (Sun 2009).

The LPJ model was run with climate data comprising monthly precipitation, temperature, wet frequency, and cloud cover. Furthermore, a dataset of global atmospheric CO_2

concentrations obtained from a carbon cycle model, also including ice core measurements and atmospheric observations (Kicklighter et al. 1999). Soil texture data were based on the Food and Agriculture Organization (FAO) soil dataset (Zobler 1986).

To run the LPJ model over a period of 1000 model years, the equilibrium state was obtained by repeatedly using climate data from the Climatic Research Unit (CRU) 0.5° global climate dataset over the period 1901–1930 (Mitchell and Jones 2005). Generally, the equilibrium state will change, and the model will need to be spun-up over and over again when the parameters change. Therefore, in our study, the equilibrium state due to variational parameters was explored. The run time for the state to attain equilibrium as a result of changing parameters was not too long. The state attained equilibrium for 100 model years as the optimization time. Previous simulation results for 1000 model years using unvaried parameters were considered as the reference state.

4 Numerical results

4.1 The sensitivity and importance of single parameter

The CNOP-P and its cost function value using the CNOP-P approach were used to examine the sensitivity and importance of single parameter, and Table 2 shows the results among all parameters for 24 cases. Only the 10 chosen parameters are shown for convenience, as the cost functions of the remaining 14 parameters were negligible. We found that the most sensitive and important parameter controlling NPP was intrinsic quantum efficiency in C3 plants (α_{C3}), and this was the case in most of northern, northeastern, and southern China. In the three cases of northeastern China, the most sensitive and important parameter controlling NPP was the maximum canopy conductance analog (g_m); however, α_{C3} was still the second most sensitive and important parameter. In the other six cases of northeastern China, g_m was the second most sensitive and important parameter. For the six cases of northern and southern China, the second most sensitive and important parameter was the fraction of photosynthetically active radiation (PAR) assimilated at the ecosystem level relative to leaf level (α_a), while the co-limitation shape parameter (θ^*) was the third most sensitive and important parameter in these cases. The α_a and r_{growth} parameters, representing growth and respiration per unit NPP, and θ^* , were, respectively, the third, fourth, and fifth most sensitive and important parameters for northeastern China. The numerical results suggest that photosynthesis and canopy conductance in the soil hydrology are important physical processes for NPP in northern China.

Furthermore, in southern and in part of northern China, α_{C3} and α_a are the two important parameters, suggesting that photosynthesis is an important physical process for NPP in these

two regions and that precipitation is sufficient, while canopy conductance in soil hydrology may be secondary. However, in the arid and semi-arid regions, the most sensitive and important parameter was found to be different in different cases. In most cases, the evaporation parameter (α_m) was the most sensitive and important parameter. However, the α_{C3} , g_m , and α_a parameters were also important parameters in all cases of arid and semi-arid China. The results showed that the parameter related to evaporation was the most sensitive and important for water-limited regions. However, the most sensitive and important parameter was different for different water-limited regions, with other important parameters being those that describe photosynthesis and canopy conductance in soil hydrology, providing further indication that these are important physical processes overall. In summary, the numerical results showed that the most sensitive and important parameters might be different for different regions, especially in both water-limited and non-limited regions.

4.2 The sensitivity and importance of multiple parameters

In the results reported in the previous section, the sensitivity and importance of single parameter was identified using the CNOP-P approach. The dimension of p is 1 in Eq. (3). However, the results do not elucidate which combination or subset of parameters is the most sensitive and important. This is because the individual parameter is optimized only to determine its sensitivity, and the combination of top five ranked sensitive parameters for the single-parameter sensitive analysis method may not be equivalent to the combination of five parameters for the multiple parameter sensitive analysis method. Next, the sensitivity of parameters combination will be explored when the p is 5 in Eq. (3).

To address this, we first needed to discover, using the CNOP-P approach, which five parameters among the 10 were the most sensitive and important when the level of nonlinear interaction among the 10 parameters is considered. Table 2 shows the maximal cost function values and which five parameters among the 10 were the most sensitive and important using the CNOP-P approach for the different study regions. The results suggest that the five most sensitive and important parameters are similar in northern, northeastern, and southern China. For example, for the second and third cases, the five parameters were θ^* , α_a , α_{C3} , r_{growth} , and g_m . However, in the arid and semi-arid regions of China, the five most sensitive and important parameters were different. For example, for the 16th, 18th, 20th, and 24th cases, the five most sensitive and important parameters were α_a , α_{C3} , r_{growth} , α_m , and k_{TP} . The five parameters represent photosynthesis, respiration, hydrology, and allocation of annual carbon increment processes. The uncertainties of these physical process overestimate the GPP due to α_a and α_{C3} and underestimate the autotrophic respiration in virtue of r_{growth} , α_m and k_{TP} . So, the uncertainties of

these physical processes lead to the large uncertainty of the NPP, which is highly sensitive to these parameters. Meanwhile, for the 19th case, the five most important parameters were θ^* , α_a , α_{C3} , r_{growth} and k_{rp} . There are similar physical processes to influence the variation of the NPP.

While the sensitivity and importance of single parameter were identified in the single-parameter experiments, nonlinear interactions among all the parameters are neglected using the CNOP-P approach. Next, we try to answer whether the five most sensitive and important parameters using the CNOP-P approach were the same as the ranking of the foremost five parameters reported in Section 4.1. We found that in northern, northeastern, and southern China, the five most important parameters using the CNOP-P approach were similar to the foremost five parameters ranked in Section 4.1. For example, the important five parameters using the CNOP-P approach for the second case were α_a , α_{C3} , r_{growth} , g_m , and α_m , which was the same set of parameters as those determined using the single-parameter experiment. However, the five most sensitive and important parameters were different when using the CNOP-P approach for multiple parameters and using it for individual parameter in the arid and semi-arid regions. For example, for the 19th region, the five parameters were θ^* , α_a , α_{C3} , r_{growth} , and k_{rp} . However, based on the single-parameter experiments using the CNOP-P approach, the five parameters were α_a , α_{C3} , g_m , α_m , and k_{mort1} . The numerical results imply that the parameter r_{growth}^* , representing the respiration process, is very important when nonlinear interactions are considered during the plant growth process, although the single-parameter sensitivity regarding r_{growth} may not show it to be more sensitive than other parameters. The results also illustrate that nonlinear interactions among parameters and complex physical processes can be explored using the CNOP-P approach. The CNOP-P approach may be able to reasonably estimate the carbon cycle process; however, the single-parameter sensitivity method may overestimate the respiration process without considering the r_{growth} parameter (Table 3).

4.3 Sensitivity experiment using the OAT approach

In Section 4.1, we reported the results from single-parameter sensitivity experiments implemented using the CNOP-P approach. To compare these results with those based on other methods, the traditional OAT approach was employed to identify the sensitivity of the single parameter. The cost function was computed using Eq. (3) when $p = \pm 0.2$. The larger the cost function is, the more sensitive and important the parameter. As shown in Table 4, parameter sensitivity and importance were similar to those using the CNOP-P approach. Due to the design of the CNOP-P approach, the sensitivity and importance of every parameter are optimal. Zaehle et al. (2005) employed a Monte Carlo technique and reported a similar conclusion based

Table 3 The five most sensitive parameters using the CNOP-P approach (*numbers* represent the sequence number of single parameter as in Table 1, and 125.75 45.75 means 125.75°E 45.75°N. There are similar for other locations)

| Number | Location | Cost function | Parameters combination |
|--------|--------------|---------------|------------------------|
| 1 | 125.75 45.75 | 200.7070 | 1,2,4,9,10 |
| 2 | 125.75 46.25 | 199.3487 | 1,2,4,9,10 |
| 3 | 125.75 46.75 | 210.2074 | 1,2,4,9,10 |
| 4 | 126.25 45.75 | 189.8998 | 1,2,4,9,10 |
| 5 | 126.25 46.25 | 194.8346 | 1,2,4,9,10 |
| 6 | 126.25 46.75 | 198.5482 | 1,2,4,9,10 |
| 7 | 126.75 45.75 | 199.2951 | 1,2,4,9,10 |
| 8 | 126.75 46.25 | 199.9770 | 1,2,4,9,10 |
| 9 | 126.75 46.75 | 197.9136 | 1,2,4,9,10 |
| 10 | 115.75 26.25 | 168.7977 | 1,2,4,9,10 |
| 11 | 115.75 26.75 | 157.9411 | 1,2,4,9,10 |
| 12 | 115.75 27.25 | 149.2697 | 1,2,4,9,17 |
| 13 | 115.75 32.25 | 162.6690 | 1,2,4,9,10 |
| 14 | 115.75 32.75 | 150.9800 | 1,2,4,9,10 |
| 15 | 115.75 33.25 | 156.7094 | 1,2,4,9,10 |
| 16 | 115.75 36.25 | 276.1842 | 2,4,9,11,16 |
| 17 | 115.75 36.75 | 295.5089 | 2,4,8,9,10 |
| 18 | 115.75 37.25 | 314.7552 | 2,4,9,11,16 |
| 19 | 116.25 36.25 | 270.4804 | 1,2,4,9,16 |
| 20 | 116.25 36.75 | 297.9204 | 2,4,9,11,16 |
| 21 | 116.25 37.25 | 304.4899 | 1,2,4,9,11 |
| 22 | 116.75 36.25 | 244.2345 | 2,4,10,11,14 |
| 23 | 116.75 36.75 | 203.4058 | 4,5,9,16,24 |
| 24 | 116.75 37.25 | 297.2164 | 2,4,9,11,16 |

on 81 cases. These results reveal that sensitivity and importance are similar when parameters are considered individually.

4.4 Sensitivity experiment on the five most important parameters determined for different basic states

To discuss whether the sensitivity of the parameters is dependent on the choice of the parameters values, the parameters are identified when the reference values of the parameters change. The random number, which satisfies the normal distribution with zero average and a standard deviation of 0.1, is superimposed onto the reference state as the new reference state. The sensitivity of every parameter is determined with the CNOP-P approach for the 24 cases. The results showed that the sensitivity of every parameter given the new reference state is similar to that for the previous reference state (not shown). The five most important parameters are identified according to the sensitivity of every parameter. The numerical results show that the combinations of important parameters in 21 out of 24 cases given the new reference state are the same as those for the previous reference state (Table 5). However,

Table 4 The sensitivity of single parameter among 10 parameters using the OAT method (*numbers* represent the sequence number of single parameter as in Table 1, and 125.75 45.75 means 125.75°E 45.75°N. There are similar for other locations)

| Number | Location |
|--------|--|
| 1 | 125.75 45.75 10 4 2 9 11 1 8 17 23 15 |
| 2 | 125.75 46.25 10 4 2 9 1 11 17 8 23 12 |
| 3 | 125.75 46.75 10 4 2 9 1 11 17 8 5 24 |
| 4 | 126.25 45.75 4 10 2 9 1 11 8 5 23 15 |
| 5 | 126.25 46.25 4 10 2 9 1 11 15 8 5 17 |
| 6 | 126.25 46.75 4 10 2 9 1 17 8 11 5 23 |
| 7 | 126.75 45.75 4 10 2 9 1 11 8 17 5 15 |
| 8 | 126.75 46.25 4 10 2 9 1 17 8 11 5 23 |
| 9 | 126.75 46.75 4 10 2 9 1 8 17 15 11 5 |
| 10 | 115.75 26.25 4 2 1 9 10 17 5 8 24 19 |
| 11 | 115.75 26.75 4 2 1 17 9 10 5 24 8 19 |
| 12 | 115.75 27.25 4 2 1 17 9 16 15 10 11 5 |
| 13 | 115.75 32.25 4 2 1 9 17 5 10 24 8 19 |
| 14 | 115.75 32.75 4 2 10 1 9 5 24 8 19 15 |
| 15 | 115.75 33.25 4 2 10 1 9 11 5 19 23 24 |
| 16 | 115.75 36.25 10 4 9 2 24 11 14 13 17 6 |
| 17 | 115.75 36.75 11 10 4 2 1 17 8 24 13 14 |
| 18 | 115.75 37.25 10 11 4 2 1 8 9 5 14 24 |
| 19 | 116.25 36.25 4 2 17 1 10 11 14 9 5 19 |
| 20 | 116.25 36.75 10 4 2 11 9 19 8 16 3 23 |
| 21 | 116.25 37.25 10 4 11 6 9 2 5 1 16 24 |
| 22 | 116.75 36.25 4 2 1 11 16 5 24 14 12 19 |
| 23 | 116.75 36.75 10 4 9 2 11 24 17 1 8 15 |
| 24 | 116.75 37.25 10 11 4 9 2 8 15 14 5 19 |

the important parameters combinations for three of the cases are different than those for the previous reference state. In three of these cases, one parameter in combinations of important parameters for the new reference state is different than the parameters in combination for the precious reference state. The three cases are located in the arid and semi-arid region in China. The above numerical results imply that the combinations of identified important parameters are the same for different reference states in most cases. In the arid and semi-arid region, there are slight differences in the combinations of identified important parameters for different reference states.

4.5 Sensitivity experiment on the most important parameters determined for different sizes of the optimization set

In the above studies, the size of the optimization set is 5. In this section, the difference about the sensitivity of the parameters combination is explored when the size of the optimization set changes. The sensitive parameters combinations are determined for the three, four, and six parameters

Table 5 Same to Table 3, just for different reference state, and 125.75 45.75 means 125.75°E 45.75°N. There are similar for other locations

| Number | Location | Cost function | Parameters combination |
|--------|--------------|---------------|------------------------|
| 1 | 125.75 45.75 | 200.6966 | 1,2,4,9,10 |
| 2 | 125.75 46.25 | 198.6640 | 1,2,4,9,10 |
| 3 | 125.75 46.75 | 210.3192 | 1,2,4,9,10 |
| 4 | 126.25 45.75 | 189.2977 | 1,2,4,9,10 |
| 5 | 126.25 46.25 | 192.5827 | 1,2,4,9,10 |
| 6 | 126.25 46.75 | 199.4382 | 1,2,4,9,10 |
| 7 | 126.75 45.75 | 199.7737 | 1,2,4,9,10 |
| 8 | 126.75 46.25 | 198.6443 | 1,2,4,9,10 |
| 9 | 126.75 46.75 | 198.8266 | 1,2,4,9,10 |
| 10 | 115.75 26.25 | 170.9943 | 1,2,4,9,10 |
| 11 | 115.75 26.75 | 149.1599 | 1,2,4,9,10 |
| 12 | 115.75 27.25 | 149.0830 | 1,2,4,9,17 |
| 13 | 115.75 32.25 | 166.0491 | 1,2,4,9,10 |
| 14 | 115.75 32.75 | 147.4435 | 1,2,4,9,10 |
| 15 | 115.75 33.25 | 159.6410 | 1,2,4,9,10 |
| 16 | 115.75 36.25 | 230.7753 | 2,4,9,11,16 |
| 17 | 115.75 36.75 | 302.3711 | 2,4,8,9,10 |
| 18 | 115.75 37.25 | 311.1083 | 1,2,4,9,11 |
| 19 | 116.25 36.25 | 262.5241 | 1,2,4,9,16 |
| 20 | 116.25 36.75 | 274.9986 | 2,4,9,11,16 |
| 21 | 116.25 37.25 | 249.7583 | 1,2,4,9,11 |
| 22 | 116.75 36.25 | 303.4423 | 1,2,4,10,11 |
| 23 | 116.75 36.75 | 265.6096 | 4,5,8,16,24 |
| 24 | 116.75 37.25 | 243.1687 | 2,4,9,11,16 |

in the parameters combination (Table 6). It is found that the sensitive parameters combination with high size parameters contains that with low size parameters in northern, northeastern, and southern China. For example, for the case (115.75°E, 33.25°N), the sensitive parameters combinations with different dimensions are respectively $(\alpha_a^*, \alpha_{C3}^*, g_m^*)$, $(\theta^*, \alpha_a^*, \alpha_{C3}^*, g_m^*)$, $(\theta^*, \alpha_a^*, \alpha_{C3}^*, r_{growth}^*, g_m^*)$, and $(\theta^*, \alpha_a^*, \alpha_{C3}^*, r_{growth}^*, g_m^*, k_{mort1}^*)$. It is found that these regions are moisture or semi-moisture region in China. In the part of arid and semi-arid regions, there is similar character to the moisture and semi-moisture regions. For example, for the case (116.75°E, 37.25°N), the sensitive parameters combinations with different dimensions are respectively $(\alpha_a^*, \alpha_{C3}^*, \alpha_m)$, $(\alpha_a^*, \alpha_{C3}^*, r_{growth}^*, \alpha_m)$, $(\alpha_a^*, \alpha_{C3}^*, r_{growth}^*, \alpha_m, k_{tp})$, and $(\alpha_a^*, \alpha_{C3}^*, r_{growth}^*, \alpha_m, k_{allom3}, k_{tp})$. However, in the other part of arid and semi-arid regions, there is a different character. It is found that the sensitive parameters combination with high size parameters does not completely contain that with low size parameters in the other part of the arid and semi-arid region. For example, for the case (116.25°E, 36.75°N), the sensitive parameters combinations with different dimensions are respectively $(\alpha_a^*, \alpha_{C3}^*, \alpha_m)$, $(\alpha_a^*, \alpha_{C3}^*, \alpha_m, est_{max}^*)$, $(\alpha_a^*$

Table 6 The comparison for three, four, and six sensitive parameters combinations using the CNOP-P approach (*numbers* represent the sequence number of single parameter as in Table 1, and 125.75 45.75 means 125.75°E 45.75°N. There are similar for other locations)

| Number | Location | Parameters combination with three parameters | Parameters combination with four parameters | Parameters combination with six parameters |
|--------|--------------|--|---|--|
| 1 | 125.75 45.75 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 2 | 125.75 46.25 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 3 | 125.75 46.75 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 4 | 126.25 45.75 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 5 | 126.25 46.25 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 6 | 126.25 46.75 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 7 | 126.75 45.75 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 8 | 126.75 46.25 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 9 | 126.75 46.75 | 2,4,10 | 2,4,9,10 | 1,2,4,9,10,11 |
| 10 | 115.75 26.25 | 2,4,10 | 1,2,4,10 | 1,2,4,5,9,10 |
| 11 | 115.75 26.75 | 2,4,10 | 1,2,4,10 | 1,2,4,5,9,10 |
| 12 | 115.75 27.25 | 1,2,4 | 1,2,4,9 | 1,2,4,5,9,17 |
| 13 | 115.75 32.25 | 2,4,10 | 1,2,4,10 | 1,2,4,5,9,10 |
| 14 | 115.75 32.75 | 1,2,4 | 1,2,4,10 | 1,2,4,9,10,11 |
| 15 | 115.75 33.25 | 2,4,10 | 1,2,4,10 | 1,2,4,9,10,17 |
| 16 | 115.75 36.25 | 4,11,19 | 2,4,10,11 | 2,4,9,10,11,16 |
| 17 | 115.75 36.75 | 4,9,11 | 2,4,9,11 | 2,4,5,9,11,16 |
| 18 | 115.75 37.25 | 4,9,11 | 2,4,11,16 | 2,4,9,11,12,16 |
| 19 | 116.25 36.25 | 2,4,9 | 1,2,4,10 | 1,2,4,9,11,17 |
| 20 | 116.25 36.75 | 2,4,11 | 2,4,11,19 | 2,4,9,11,16,24 |
| 21 | 116.25 37.25 | 2,4,11 | 2,4,11,16 | 1,2,4,9,11,16 |
| 22 | 116.75 36.25 | 1,2,4 | 2,4,10,11 | 1,2,4,10,11,24 |
| 23 | 116.75 36.75 | 4,5,9 | 4,8,10,24 | 2,4,8,9,10,11 |
| 24 | 116.75 37.25 | 2,4,11 | 2,4,9,11 | 2,4,9,11,14,16 |

α_{C3}^* , r_{growth}^* , α_m , k_{TP}), and $(\alpha_a^*$, α_{C3}^* , r_{growth}^* , α_m , k_{TP} , β). The above numerical results imply that in the strong interaction region between land and atmosphere, such as arid and semi-arid region, the parameters combination may be dependent on the size of the parameters combination. However, in the weak interaction region between land and atmosphere, such as moisture region, the parameters combination may not be dependent on the size of the parameters combination.

4.6 Sensitivity experiment to identify the five most important parameters

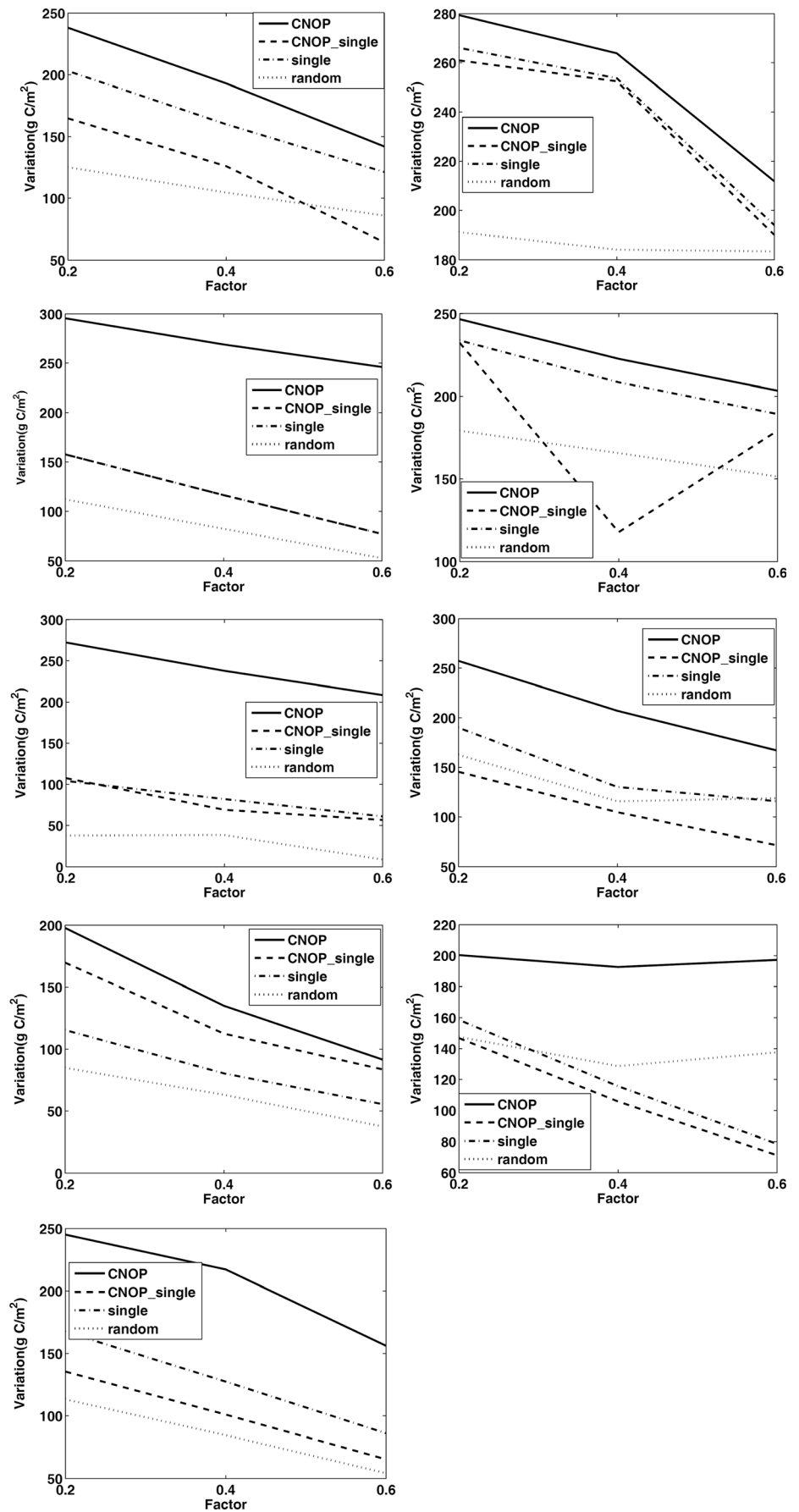
An important aim of this study was to improve the ability to estimate and predict the carbon cycle in terrestrial ecosystems using the identified sensitive and important parameters. Errors in these parameters are reduced through routine or additional observations. However, to show the extent to which the simulations were improved when the most sensitive and important parameters pattern was

destroyed, experiments were designed and carried out as follows. As reported in Section 4.2, 252 parameters combinations were computed to obtain the CNOP-Ps. Among those CNOP-Ps and their cost functions, the CNOP-P leading to the maximal cost function is called the CNOP. Excluding the CNOP-P-max, two CNOP-Ps were chosen to compare the extent of improvement. Among the 252 CNOP-Ps representing parameters combinations, a group of parameters, which was the same as the foremost five parameters identified using the CNOP-P approach for individual parameter, was named CNOP_single. Another group of parameters, which was same as the foremost five parameters identified using the sensitivity analysis method for single parameter, was simply called the “single” group. Finally, a random group of parameters was also chosen. Referred to Mu et al. (2009). who employed a formula to show the benefit of reducing the CNOP-type initial errors compared with another type of initial error, the extent of improvement in the uncertainty in the numerical simulation is measured as τ .

$$\tau = \frac{\|M_T(U_0, P + p) - M_T(U_0, P)\|}{\|M_T(U_0, P + \alpha p) - M_T(U_0, P)\|} \quad (4)$$

The parameters in Eq. (4) are similar to those in Eq. (3), and α represents the decreasing extent of the uncertainty of the parameters, which is a constant less than 1. In this study, the four types of parameters including the CNOP were compared with τ for each case. The larger τ was, the better the extent of improvement. The 252 parameters combinations must include the above four types of parameters. The numerical results in Section 4.2 showed that the most sensitive and important parameters were similar in most of the study regions, except for arid and semi-arid regions in China. Therefore, the sensitivity experiment was implemented for these regions. Figure 3 shows the variation using the four types of parameters combinations using three factors ($\alpha = 0.2, 0.4, \text{ and } 0.6$) by Eq. (4) in the arid and semi-arid regions in China. The average reduced extents of the uncertainties for nine cases in the simulated NPP due to the CNOP-P-type parameter error were 248.02, 215.35, and 180.43 g C/m² for $\alpha = 0.2, 0.4, \text{ and } 0.6$, respectively. Meanwhile, for the CNOP_single parameter error, the results were 168.98, 122.87, and 95.48 g C/m²; for the “single” parameter error they were 177.36, 141.60, and 108.74 g C/m²; and for the random errors they were 128.23, 107.48, and 92.30 g C/m², respectively. The results illustrate that the gain obtained by the CNOP-P-type parameters combinations was the best among all the parameters combinations for each case and each factor. The results also suggest that the uncertainties in the simulations of the carbon cycle due to parameter errors can be reduced through CNOP-P-type identification of the most sensitive and important parameters combination.

Fig. 3 Sensitivity experiments carried out to compute the gain from the five most sensitive parameters using the CNOP-P approach. *CNOP* represents the reduced extent due to the CNOP-P-type parameter error. *CNOP_single* represents the reduced extent due to the parameter error by optimizing single parameter. *Single* represents the reduced extent due to the parameter error by the OAT approach. *Random* represents the reduced extent due to the random parameter error



5 Discussion and summary

The uncertainties in physical parameters of numerical models are the main source of numerical simulation ability and forecast skill, such as numerical simulation for land process (Kuczera and Parent 1998; Vrugt et al. 2003). An approach of reducing uncertainty in numerical simulation is to calibrate the model parameters or parameters combination to closely match the input and output behavior to the observation data for model development. However, it is necessary to choose the model parameters or parameters combination to be calibrated because there are many parameters in the numerical model, such as the number of controlling the model parameters about from $O(10)$ to $O(100)$ (Li et al. 2013). Besides, to reduce the uncertainty of numerical modeling, the effects of the parameters combination should be considered, not to just calibrate parameter one by one. So, it is a key issue to determine which group of physical parameters should firstly be calibrated. Although the model parameters could be calibrated through the above methods, the calibrated model parameters values may not match with the true values. So, the calibrated model parameter values may not be applicable to other numerical simulations or forecast. It is feasible method that the physical parameter errors are reduced through observational data. Since the observation of the physical parameters is difficult and expensive, it is very important which parameter should be firstly observed to obtain the true values to improve the numerical models.

The physical parameters were ranked using the current parameter sensitivity analysis methods (Zaehle et al. 2005). For example, based on probability theory, such as Monte Carlo techniques, parameter sensitivity can be conducted according to model output through multiple model simulations. One method is to identify parameter sensitivity by analyzing the contribution of single parameter to the overall output uncertainty with variance (Verbeeck et al. 2006). Another is to determine parameter sensitivity by calculating the ranked partial correlation coefficient between model inputs and outputs (Zaehle et al. 2005). These methods can determine parameter sensitivity and show the ranking of the parameters. For example, Zaehle et al. (2005) determined the five top-most sensitive parameters for NPP to be α_{C3} , α_a , θ^* , g_m , and r_{growth} using a Monte Carlo-type stratified sampling approach for 81 cases from the class A data set, in which the regions in China maybe not included. In our study, the group of physical parameters among abundant parameters is identified using the new approach based on the CNOP-P approach. The results in Zaehle et al. (2005) were similar to those in the present study using the CNOP-P approach in the northern, northeastern, and southern regions. In the arid and semi-arid regions of China, the parameter sensitivity using the CNOP-P approach was different to that proposed by Zaehle et al. (2005). In the case (36.75°N, 116.25°E), the top five parameters for NPP using

the CNOP-P approach were α_m , g_m , α_{C3} , α_a , and r_{growth} . Furthermore, we found that the results emphasized water demand in water-limited regions. The sensitivity index was -0.11 for α_m in Zaehle et al. (2005). Their study did note that the evaporation parameter, α_m , is important, but the rank of α_m was more than 12, which implied that the importance and sensitivity of α_m were weaker than they were in 12 other parameters. In our study, α_m was the most important and sensitive parameter among the 24 parameters in the arid and semi-arid regions of China.

Owing to the nonlinearity among parameters, the most sensitive and important five parameters are not necessarily the top five parameters identified using the above methods. One advantage of the CNOP-P approach is that it can identify which group of parameters is the most sensitive and important compared with other groups of five parameters according to the cost function value. Using this approach, our results showed that for the northern, northeastern, and southern regions of China, the subsets of the five relatively more sensitive and important parameters were similar to the top five parameters proposed by Zaehle et al. (2005). However, for the arid and semi-arid regions, the combination of the five relatively more sensitive and important parameters using the CNOP-P approach was different to the combination proposed by Zaehle et al. (2005). The numerical results also show that the parameters in the arid and semi-arid region need to be carefully validated for the model validation.

In the arid and semi-arid region, soil water content is the key factor controlling plant growth and survival. The parameter α_m , representing evapotranspiration, plays a key role in vegetation growth in water-limited regions. The identified combination of parameters using the CNOP-P approach was able to exhibit the physical characteristics. The parameter g_m , representing the maximum canopy conductance analog, had a faint sensitivity in water-limited regions when the nonlinear interaction among parameters was considered. The numerical results suggest that not all parameters related to vegetation growth are key parameters and illustrate that the nonlinear interactions among parameters are crucial to the uncertainty of numerical simulations in complex regions associated with nonlinear physical processes. As demonstrated in the present study, the CNOP-P approach is able to consider nonlinear physical processes and is recommended for identifying the most significant subset or combination of relatively more sensitive and important parameters. It is important to determine the combination of relatively more sensitive and important parameters within a given number of parameters. The number of parameters combinations can be evaluated according to the practicalities of human and material resources, such as obtaining observations of the physical parameters and executing the optimization method. The relatively more sensitive and important parameters combination is determined based on the evaluated number of

parameters. Bastidas et al. (1999) employed the MOGSA method to identify which of the parameters were sensitive according to the significance level. They reported that there were nine sensitive parameters at a significance level below 1 % for sensible heat, latent heat, and ground temperature in the Tucson semi-arid region. However, our studies could answer which four or five parameters was the most sensitive among the abundant parameters.

The new approach established in the present study would be able to determine which subset of parameters was the most sensitive, as it is able to determine the most significant combination of relatively more sensitive and important parameters within a given total number of parameters. The core of the new approach is the CNOP-P method that considers the impact of nonlinear interaction among the parameters on the uncertainty of the numerical simulation. There are two key characteristics of the new approach. Firstly, it is able to explore the response of uncertainty in numerical simulations to nonlinear interactions among parameters. Secondly, in establishing the sensitivity of every parameter, it ensures that the sensitivity of every parameter is optimal, due to the CNOP-P approach optimizing every parameter. In addition, the most important combination of parameters is also identified through the new approach and is therefore optimized as well. According to current numerical results, for the different reference states, the top 10 sensitive parameters are similar using the CNOP-P approach to identify the sensitivity of every parameter. This will not lead to the determination of the five sensitive parameters combination using the CNOP-P approach, especially for the moisture regions. There is a little difference about the identification of the sensitive parameters in the arid and semi-arid regions if there is much difference among the reference states. It is possible and reasonable that these are different when the parameter values have large differences. The key parameters may be different for the different states due to climate condition or complex physical processes, or different initial conditions, and so on. However, the group of parameter, which could lead to the maximum uncertainty of numerical simulation, is identified using the new method based on the CNOP-P approach. The findings of the group of parameter could guide us firstly to calibrate the parameters to improve the ability of numerical simulation. On the other hand, because there are two group optimization processes during the two-stage sensitive parameter identification method in our study, the computational cost is more expensive than other methods, such as the OAT method, and the traditional Latin Hypercube method (Zaehle et al. 2005). However, the current dynamical vegetation models are single column models. The dynamical vegetation models could be run in parallel means. Based on the models' feature, the computational cost will be reduced using the CNOP-P approach to identify the sensitivity of parameter when the parallel skill is applied using many central processing units (CPUs) of computers.

The LPJ model was taken as an example to demonstrate how to implement the new approach. The numerical results showed that the most important combination of parameters in the arid and semi-arid regions of China was different than those in northern, northeastern, and southern China, which were similar to each other. The numerical results also indicated that the most important five-parameter combinations in northern, northeastern, and southern China were the same with or without consideration of the nonlinear interactions among the parameters. However, this was not the case in the arid and semi-arid regions, where the most important combinations were different with and without consideration of the nonlinear interactions among the parameters.

A specific example was shown in which we identified the most important combination of parameters. We wanted to determine the most important subset of five relatively more sensitive and important physical parameters among an initial total of 24. Firstly, 14 physical parameters were eliminated using the CNOP-P approach. Secondly, the most significant five-parameter combination from the remaining 10 more sensitive and important parameters was determined using the CNOP-P approach and a combination of methods. Whether the identified subset of five relatively more sensitive and important parameters is dependent on the number of initially eliminated parameters will be explored in the future. Moreover, combinations of four or seven parameters will also be identified in future work. In the strong interaction region between land and atmosphere, such as arid and semi-arid region, the parameters combination may be dependent on the number of the parameters combination. However, in the weak interaction region between land and atmosphere, such as moisture region, the parameters combination may not be dependent on the number of the parameters combination. These issues will be explored in the future. Moreover, the constraint condition is chosen as $\delta = 0.2$ in the current study. It is necessary to explore the sensitivity of parameters combination related to the constraint condition.

As we know, there are two types of predictability problems based on the different factors that lead to them. The first is related to initial errors (Palmer et al. 1998) and the second involves model uncertainties in numerical models performing simulations and forecasts of the Earth system (Williams et al. 2001; Sitch et al. 2003; Jackson et al. 2004; Berthelot et al. 2005; Lin et al. 2011). At present, targeted observations, also called adaptive observations, determine certain special areas (called sensitive areas) that cause large uncertainty in forecast results. These additional observations in these sensitive areas supply more reliable initial states for the model and thus a more accurate prediction is expected (Palmer et al. 1998; Bishop et al. 2001; Wu et al. 2005; Qin and Mu 2011). Mu (2013) proposed a new idea to address the second predictability problem by generalizing targeted observations from geographical space to phase space for model parameters. The

essential point is how to determine the more sensitive and important parameters among all model parameters. The present paper provides a concrete approach to realize this goal in a new way. Whether such types of targeted observations in the phase space can effectively reduce the uncertainty of physical processes and model parameters requires more work in the future.

Acknowledgments The authors thank Dr. Yanshan Yu and Dr. Xiaohao Qin for the valuable discussions held. Funding was provided by grants from the National Natural Science Foundation of China (Nos. 91437111, 41375111, 40830955), and by a grant from the State Key Development Program for Basic Research (No. 2012CB955202).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aberson SD (2011) The impact of drowindsonde data from the THORPEX Pacific area regional campaign and the NOAA hurricane field program on tropical cyclone forecasts in the global forecast system. *Mon Weather Rev* 139:2689–2703
- Bastidas LA, Gupta HV, Sorooshian S, Shuttleworth WJ, Yang ZL (1999) Sensitivity analysis of a land surface scheme using multicriteria methods. *J Geophys Res* 104(D16):19481–19490. doi:10.1029/1999JD900155
- Bastidas LA, Hogue TS, Sorooshian S, Gupta HV, Shuttleworth WJ (2006) Parameter sensitivity analysis for different complexity land surface models using multicriteria methods. *J Geophys Res* 111: D20101. doi:10.1029/2005JD006377
- Berthelot M, Friedlingstein P, Ciais P, Dufresne JL, Monfray P (2005) How uncertainties in future climate change predictions translate into future terrestrial carbon fluxes. *Glob Chang Biol* 11:959–970. doi:10.1111/j.1365-2486.2005.00957.x
- Bishop CH, Etherton BJ, Majumdar SJ (2001) Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. *Mon Weather Rev* 129:420–436
- Carrasi A, Vannitsem S (2011) State and parameter estimation with the extended Kalman filter: an alternative formulation of the model error dynamics. *Q J R Meteorol Soc* 137:435–451. doi:10.1002/qj.762
- Cramer W, Bondeau A, Woodward FI, Prentice IC, Betts R, Brovkin V, Cox PM, Fischer V, Foley JA, Friend AD, Kucharik C, Lomas MR, Ramankutty N, Sitch S, Smith B, White A, Young-Molling C (2001) Global responses of terrestrial ecosystem structure and function to CO₂ and climate change: results from six dynamic global vegetation models. *Glob Chang Biol* 7:357–373
- De Lannoy GJM, Houser PR, Pauwels VRN, Verhoest NEC (2006) Assessment of model uncertainty for soil moisture through ensemble verification. *J Geophys Res* 111:D10101. doi:10.1029/2005JD006367
- Duan Q, Sorooshian S, Gupta V (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour Res* 28:1015–1031. doi:10.1029/91WR02985
- Duan WS, Zhang R (2010) Is model parameter error related to a significant spring predictability barrier for El Niño events? Results from a theoretical model. *Adv Atmos Sci* 27(5):1003–1013
- Hally A, Richard E, Fresnay S, Lambert D (2013) Ensemble simulations with perturbed physical parametrizations: pre-HyMeX case studies. *Q J R Meteorol Soc*. doi:10.1002/qj.2257
- Henderson-Sellers A (1992) Assessing the sensitivity of a land-surface scheme to parameters used in tropical-deforestation experiments. *Q J R Meteorol Soc* 118:1101–1116. doi:10.1002/qj.49711850805
- Hickler T, Smith B, Prentice IC, Mjofors K, Miller P, Ameth A, Sykes MT (2008) CO₂ fertilization in temperate FACE experiments not representative of boreal and tropical forests. *Glob Chang Biol* 14: 1531–1542
- Jackson C, Sen MK, Stoffa PL (2004) An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions. *J Clim* 17:2828–2841. doi:10.1111/j.1365-2486.2008.01598.x
- Jackson C, Xia Y, Sen MK, Stoffa PL (2003) Optimal parameter and uncertainty estimation of a land surface model: a case study using data from Cabauw, Netherlands. *J Geophys Res* 108(D18):4583. doi:10.1029/2002JD002991
- Janiskova M, Morcrette J (2005) Investigation of the sensitivity of the ECMWF radiation scheme to input parameters using the adjoint technique. *Q J R Meteorol Soc* 131:1975–1995. doi:10.1256/qj.04.183
- Jarvinen H, Laine M, Solonen A, Haario H (2012) Ensemble prediction and parameter estimation system: the concept. *Q J R Meteorol Soc* 138:281–288. doi:10.1002/qj.923
- Kicklighter DW et al. (1999) A first-order analysis of the potential role of CO₂ fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models. *Tellus B* 51(2): 343–366. doi:10.1034/j.1600-0889.1999.00017.x
- Knorr W, Heimann M (2001) Uncertainties in global terrestrial biosphere modeling: I. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme. *Glob Biogeochem Cycles* 15(1):207–225
- Kruger J (1993) Simulated annealing—a tool for data assimilation into an almost steady model state. *J Phys Oceanogr* 23(4): 679–688
- Kuczera G, Parent E (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *J Hydrol* 211:69–85
- Li X, Su ZB, Wu BF (2012) Retrieval of key eco-hydrological parameters for cold and arid regions. *International Journal of Applied Earth Observation and Geoinformation* 17:1–2. doi:10.1016/j.jag.2012.03.001
- Li J, Duan QY, Gong W, Ye A, Dai Y, Miao C, Di Z, Tong C, Sun Y (2013) Assessing parameter importance of the common land model based on qualitative and quantitative sensitivity analysis. *Hydrol Earth Syst Sci* 17:3279–3293. doi:10.5194/hess-17-3279-2013
- Lin JC, Pejam MR, Chan E, Wofsy SC, Gottlieb EW, Margolis HA, McCaughey JH (2011) Attributing uncertainties in simulated biospheric carbon fluxes to different error sources. *Glob Biogeochem Cycles* 25:GB2018. doi:10.1029/2010GB003884
- Lu JX, Hsieh WW (1997) Adjoint data assimilation in coupled atmosphere-ocean models: determining model parameters in a simple equatorial model. *Q J R Meteorol Soc* 123:2115–2139. doi:10.1002/qj.49712354316
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol* 25(6):693–712
- Morss RE, Battisti DS (2004) Evaluating observing requirements for ENSO prediction: experiments with an intermediate coupled model. *J Clim* 17:3057–3073
- Mu M, Duan W, Wang J (2002) Nonlinear optimization problems in atmospheric and oceanic sciences. *East-West Journal of Mathematics, Thailand, Special Volume*, 155–164
- Mu M, Zhou FF, Wang HL (2009) A method for identifying the sensitive areas in targeted observations for tropical cyclone prediction:

- conditional nonlinear optimal perturbation. *Mon Weather Rev* 137: 1623–1639
- Mu M, Duan WS, Wang Q, Zhang R (2010) An extension of conditional nonlinear optimal perturbation approach and its applications. *Nonlinear Process Geophys* 17:211–220. doi:10.5194/npg-17-211-2010
- Mu M (2013) Methods, current status, and prospect of targeted observation. *Science China: Earth Sciences* 56(12):1997–2005. doi:10.1007/s11430-013-4727-x
- Palmer TN, Gelaro R, Barkmeijer J, et al. (1998) Singular vectors, metrics, and adaptive observations. *J Atmos Sci* 55:633–653
- Pappas C, Faticchi S, Leuzinger S, Wolf A, Burlando P (2013) Sensitivity analysis of a process-based ecosystem model: pinpointing parameterization and structural issues. *J Geophys Res Biogeosci* 118:505–528. doi:10.1002/jgrg.20035
- Pitman AJ (1994) Assessing the sensitivity of a land-surface scheme to the parameter values using a single column model. *J Clim* 7:1856–1869
- Prentice IC, Cramer W, Harrison SP, et al. (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *J Biogeogr* 19:117–134
- Pulido M, Polavarapu S, Shepherd TG, Thubum J (2012) Estimation of optimal gravity wave parameters for climate models using data assimilation. *Q J R Meteorol Soc* 138:298–309. doi:10.1002/qj.932
- Qin X, Mu M (2011) Influence of conditional nonlinear optimal perturbations sensitivity on typhoon track forecasts. *Quart J Roy Meteor Soc* 138:185–197
- Rayner PJ, Scholze M, Knorr W, Kaminski T, Giering R, Widmann H (2005) Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS). *Glob Biogeochem Cycles* 19:GB2026. doi:10.1029/2004GB002254
- Rosero E, Yang ZL, Wagener T, Gulden LE, Yatheendradas S, Niu GY (2010) Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season. *J Geophys Res* 115:D03106. doi:10.1029/2009JD012035
- Saltelli A (1999) Sensitivity analysis: could better methods be used? *J Geophys Res* 104:3789–3793
- Sitch S et al. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic vegetation model. *Glob Chang Biol* 9(2):161–185
- Smith PJ, Thornhill GD, Dance SL, Lawless AS, Mason DC, Nichols NK (2013) Data assimilation for state and parameter estimation: application to morphodynamic modelling. *Q J R Meteorol Soc* 139:314–327. doi:10.1002/qj.1944
- Storn R, Price K (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J Glob Optim* 11:341–359
- Sun GD (2009) Simulation of potential vegetation distribution and estimation of carbon flux in China from 1981 to 1998 with LPJ dynamic global vegetation model. *Climatic and Environmental Research (in Chinese)* 14(4):341–351
- Sun GD, Mu M (2009) A preliminary application of the differential evolution algorithm to calculate the CNOP. *Atmos Oceanic Sci Lett* 2(6):381–385
- Sun GD, Mu M (2012a) Inducing unstable grassland equilibrium states due to nonlinear optimal patterns of initial and parameter perturbations: theoretical models. *Adv Atmos Sci* 29(1):79–90. doi:10.1007/s00376-011-0226-1
- Sun GD, Mu M (2012b) Responses of soil carbon variation to climate variability in China using the LPJ model. *Theor Appl Climatol* 110(1–2):143–153
- Sun GD, Mu M (2013) Understanding variations and seasonal characteristics of net primary production under two types of climate change scenarios in China using the LPJ model. *Clim Chang* 120:755–769. doi:10.1007/s10584-013-0833-1
- Verbeeck H, Samson R, Verdonck F, Lemeur R (2006) Parameter sensitivity and uncertainty of the forest carbon flux model FORUG: a Monte Carlo analysis. *Tree Physiol* 26:807–817
- Vrugt JA, Gupta HV, Bouten W, Sorooshian S (2003) A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour Res* 39: 1201. doi:10.1029/2002WR001642
- Vrugt JA, Diks CGH, Gupta HV, Bouten W, Verstraten JM (2005) Improved treatment of uncertainty in hydrologic modeling: combining the strengths of global optimization and data assimilation. *Water Resour Res* 41:W01017. doi:10.1029/2004WR003059
- Wan SQ, He WP, Wang L, Jiang W, Zhang W (2012) Evolutionary modeling-based approach for model errors correction. *Nonlinear Process Geophys* 19:439–447. doi:10.5194/npg-19-439-2012
- Wang JD, Guo WD, Li HQ (2013) Application of extended Fourier amplitude sensitivity test (EFAST) method in land surface parameter sensitivity analysis. *Acta Phys Sin* 62(5):050202
- Wang Q, Mu M, Dijkstra HA (2012) Application of the conditional nonlinear optimal perturbation method to the predictability study of the Kuroshio large meander. *Adv Atmos Sci* 29(1):118–134. doi:10.1007/s00376-011-0199-0
- Wang YP, Leuning R, Cleugh HA, Coppin PA (2001) Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful? *Glob Chang Biol* 7:495–510. doi:10.1046/j.1365-2486.2001.00434.x
- Werner C, Butterbach-Bahl K, Haas E, Hickler T, Kiese R (2007) A global inventory of N₂O emissions from tropical rainforest soils using a detailed biogeochemical model. *Glob Biogeochem Cycles* 21:GB3010. doi:10.1029/2006GB002909
- White MA, Thornton PE, Running SW, Nemani RR (2000) Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: Net primary production controls. *Earth Interact* 4:1–85
- Williams KD, Senior CA, Mitchell JFB (2001) Transient climate change in the Hadley Centre models: the role of physical processes. *J Clim* 14:2659–2674
- Wu CC, Lin PH, Chou KH, Yeh TC (2005) Dropwindsonde observations for typhoon surveillance near the Taiwan region (DOSTAR): an overview. *Bull Am Meteorol Soc* 86:787–790
- Yu Y, Mu M, Duan W (2012) Does model parameter error cause a significant “spring predictability barrier” for El Niño events in the Zebiak–Cane model? *J Clim* 25:1263–1277. doi:10.1175/2011JCLI4022.1
- Zachle S, Sitch S, Smith B, Hatterman F (2005) Effects of parameter uncertainties on the modeling of terrestrial biosphere dynamics. *Glob Biogeochem Cycles* 19:GB3020. doi:10.1029/2004GB002395
- Zhang S, Zou X, Ahlquist J, Navon IM, Sela JG (2000) Use of differentiable and nondifferentiable optimization algorithms for variational data assimilation with discontinuous cost functions. *Mon Weather Rev* 128:4031–4044
- Zobler L (1986) A world soil file for global climate modeling, NASA Technical Memorandum, 87802. NASA, Washington, D.C., p. 32