Replies to Referee #1, GMD-2023-164

Xianwei Wu, Liang Hu, Lanning Wang, Haitian Lu, and Juepeng Zheng

November 30, 2023

Thank you very much for your patient and detailed comments on our work [1]. These valuable comments are very helpful for us to improve this paper. After carefully reading all the questions, we have answered each of them and will make appropriate corrections in the revised version of our manuscript.

In this attachment, the blue paragraphs represent your comments, and the black paragraphs below are our corresponding replies.

1 Replies to major comments

Replies to major comments are as follows:

1. The manuscript structure, particularly the method section, needs to be reorganized to improve the compactness. There are several areas that require clarification. For instance, Algorithm 1 calculates the RMSE, but its definition is found in section 3.2.2. It would be more appropriate to move the definition to section 2. Additionally, in Line 4 of Algorithm 2, it is unclear whether the new parameters are obtained using CAND. Furthermore, it is not explained why the local-level surrogate utilizes Gaussian Process. In addition, it could describes the difference between the algorithm used in this work and the ASMO. Typically, optimization algorithms require hundreds of steps to achieve convergence, but in this work, only around 20 steps of local optimization are performed. It is hard to say the algorithms get convergence. It appears that the ASMO method can achieve local optimization more quickly. The conclusion is not convinced. The description of CAND is difficult to follow, particularly the calculation vs and vd, which is lack of calculation details. The cross validation describe can move from result section to the method section.

Thanks for your comment. This comment contains multiple questions, we will reply these questions separately.

1) We will improve the structure of the study in revised manuscript according to the comments.

2) The new parameters are not obtained by CAND, CAND is just used for global-level surrogate model. x_t is the optimal solution within the trust region of the surrogate model. $f(x_t)$ represents the objective function (in this problem is RMSE) of the simulation results of this parameter set in CAM5.

3) In the process of constructing the global surrogate, because of the relatively large amount of samples, we chose some relatively complex learning-based models and selected the optimal method based on cross-validation results. The results in this paper indicates that the GBRT is the best method to construct global surrogate model.

In the process of building local surrogate models, we take into account the insufficient number of samples. Using some relatively complex learning-based methods may lead to under-fitting. Therefore, we consider selecting some relatively simple construction methods. Among them, Polynomial Response Surface (PRS), Gaussian Process (GP), and Radial Basis Function (RBF) are the most commonly used regression-based or statistics-based methods, widely used for various complex parameter optimization

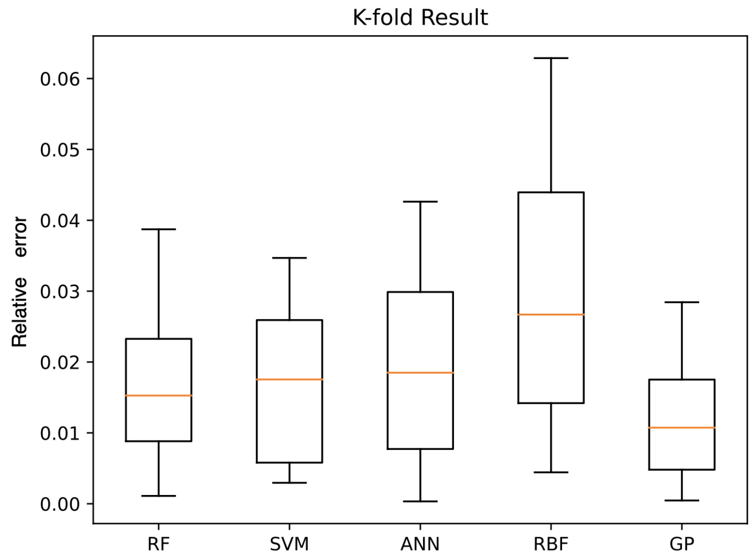

Figure 1: Local surrogate model cross-validation results

problems in the industry [2]. Therefore, we consider choosing one of these three as the method for constructing local surrogate models. Among them, we first consider PRS. Although it is the simplest method, the fitting performance of the PRS model to complex curves is relatively poor. PRS model is a simple model based on polynomials, and it may not perform well for complex, nonlinear, or highly interactive systems. Its expressive capacity is limited and may not accurately capture certain complex relationships. According to [3], the polynomial surrogate model does not perform well in terms of fitting accuracy for multivariate and nonlinear problems. Therefore, we choose either GP or RBF to construct the local surrogate model.

In order to choose a model that is more suitable for our study, we conducted cross-validation experiments based on the selection method for the global model. We selected three learning-based methods: Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN) for comparison. The results are shown in the figure 1. The results indicate that compared with RBF method, GP has a smaller error in cross-validation, providing more accurate predictions. Moreover, the cross-validation results are better than the three learning-based methods. In contrast, the RBF method not only has a larger error but also a wider range of upper and lower relative error bounds. The prediction results are unstable, and the predictive performance is lower than the three learning-based surrogate model construction methods.

Please note that in these paragraphs, "global" and "local" represent the different surrogate model, rather than "global/local optimal" in optmization process or simulation results in "global/region".

We will add the relevant experimental results to the revised manuscript.

4) The entire optimization process consists of over 20 steps. After obtaining the current optimal solution, there are several validation steps. Once these validation steps are completed, and no new optimal solution is found, we consider the current optimal solution as the final result of the optimization. In the figure, we illustrate the reduction in RMSE during the optimization process and the associated errors.

In terms of errors, it's possible that our method may indeed have slightly higher errors compared to ASMO in the end. However, our method demonstrates greater stability throughout the entire optimization process, with errors consistently maintained at a lower level. In contrast, AMSO exhibits initial oscillations in errors, indicating that our surrogate model remains stable. While our final error may be slightly higher than that of ASMO, we believe that in cases where the errors are relatively close, the reduction in the number of optimization iterations is a highlight of our method, resulting in resource savings.

5) CAND (candidate points) strategy is proposed in [4], has been widely used in surrogate model-based optimization method [5, 6]. It is used for balance the global exploration and local exploitation. **Please note that the terms "global and local" represent characters of the optimal solution in optimization methods rather than "best parameter for global precipitation or region precipitation".**The inputs of the strategy are a set of generated points Ω, a set of initial sampling points A, and the output of the strategy is the best candidate point selected from the generated points Ω.

In each iteration step, there are two criteria in CAND to select to select the best candidate point:

i). Estimated function value obtained from the surrogate model.

ii). Minimum distance from previously evaluated points.

The first criterion represents the exploitation, which means that search a better solution based on known regions. The second criterion represents the exploration, which means that search a better solution in an unknown region. **Please note that in this sentence "region" represents a part of parameter space rather than precipitation simulation results over each "region" like East Asia.** In order to find the next candidate for evaluation, we do not minimize that fitness function over a continuous set. Instead, we select the best among a finite set of randomly generated points. Fitness functions are made up from these two criteria: the value of fitness function obtained from the surrogate model at each point and its minimum distance to existing data points.

The description of CAND is described in Algorithm 1:

Algorithm 1 Candidate point strategy

1: Compute $s^{max} \leftarrow \max_{x\in\Omega} s(x)$ and $s^{min} \leftarrow \min_{x\in\Omega} s(s)$
2: **for** each $x \in \Omega$ **do**
3: $\quad V^S(x) = \begin{cases} \frac{s(x)-s^{min}}{s^{max}-s^{min}} & if\ s^{max} > s^{min} \\ 1 & else \end{cases}$
4: $\quad$ Calculate corresponding value of objective function for each sample.
5: **end for**
6: **for** each $x \in \Omega$ **do**
7: $\quad \Delta(x) = min_{y\in A} d(x, y);$
8: **end for**
9: Compute $\Delta^{max} \leftarrow \max_{x\in\Omega} s(x)$ and $\Delta^{min} \leftarrow \min_{x\in\Omega} s(x)$
10: **for** each $x \in \Omega$ **do**
11: $\quad V^D(x) = \begin{cases} \frac{\Delta(x)-\Delta^{min}}{\Delta^{max}-\Delta^{min}} & if\ \Delta^{max} > \Delta^{min} \\ 1 & else \end{cases}$
12: **end for**
13: **return** $argmin_{x\in\Omega} wV^S(x) + (1-w)V^D(x)$

Where, the Ω represents the random samples generated in this iteration process. $S(x)$ represents the predict value of point x generated by surrogate model. $\Delta(x)$ is the minimum distance from point x to the current sampling point set A and y represents each point in set A. In line 1, fitness function values of generated point sets are calculated according to the surrogate model and their maximum and minimum are marked. In line 2-4, the value of V^S is calculated and the for loops represents the criterion 1, a smaller value of V^S means that the current point is an effective exploitation. In line 6-8, the minimum distance $\Delta(x)$ between the point sample Ω and current sampling point set is calculated in the for loops. The maximum and minimum of $\Delta(x)$ are marked in line 9. The for loops in line 10-12 represent the criterion 2, a smaller value of V^D means that the current point is an effective exploration. The last line means the weighted sum to balance the exploration and exploitation.

We will add these to the revised manuscript in section 3.2.3.

6) We will move the cross validation to section method according to your comment.

2. The manuscript lacks a thorough mechanism analysis of how parameters affect precipitation on a global and regional scale. While section 4 presents optimization results, it lacks organization and falls short in providing a detailed understanding of the underlying mechanisms. To enhance the manuscript, it is recommended to delve deeper into the analysis. By investigating the cause-effect relationships between parameters and precipitation patterns, physics insights can be gained to improve the parameterization scheme.

Thanks for your comment. We try to reply to this comment from the following points.

1).The purpose of this work is to improve the CAM5 precipitation simulation result accord to parameter tuning method, rather than analyzing the mechanism of physics process. We believe that the goal has been achieved and it is a complete work. In general, the calibration of parameters in parameterization schemes relies on statistical models and expert knowledge, leading to significant uncertainty. Small variations in parameter values can result in substantial changes in simulation results. In this paper, we propose a surrogate model based method which can quickly calibrate parameters, and improve CAM5 precipitation using the proposed multi-level surrogate model method. During the optimization process, we found that the same parameter value has different effects on simulation results in different regions, called “rocker effect”. This suggests that it is challenging to achieve precipitation optimization for all regions using the same set of parameters. Therefore, we design a non-uniform parameterization scheme, employing different parameter values for distinct regions. We find a more suitable set of parameters for each region according to the multi-level surrogate model-based method and integrate these different sets of parameter values into one case.

2).We do not change the physical process in the parameterization schemes, we only changed the values of the parameters, or different values in different regions. In this paper, we use CAM5 and the selected parameters belong to three different parameterization schemes: The cloud microphysics parameterization scheme is proposed by [7, 8, 9]. The deep convective parameterization scheme is proposed by [10] and modified by [11]. The cloud fraction parameterization scheme is proposed by [12]. We believe that these studies have analyzed the physical meanings of each parameter and their effects on simulation results. While maintaining these physical processes, we improve the parameterization scheme based on the optimal values obtained for each region through the proposed tuning methods. In the improved parameterization scheme, the parameter values for each selected region are set to the tuned values, while the values for other regions remain at default parameters. In the parameterization scheme file, there have been no changes to the descriptions of physical processes. The modifications we made involve selecting different numerical values based on the judgment of region latitudes and longitudes. So that we believe that these positive improvements achieved by changing the values of the parameters. In [13], there is more introduction to the mechanism of physics process, including the impact of each parameter on precipitation results in different regions. We also refer to this work when selecting parameters and determining the range of these parameters. We will try to explain these effects from the parameter value changes according to this work and any other related works.

3. In equations 10-11, it could be possible for the numerator to be very large, and the denominator

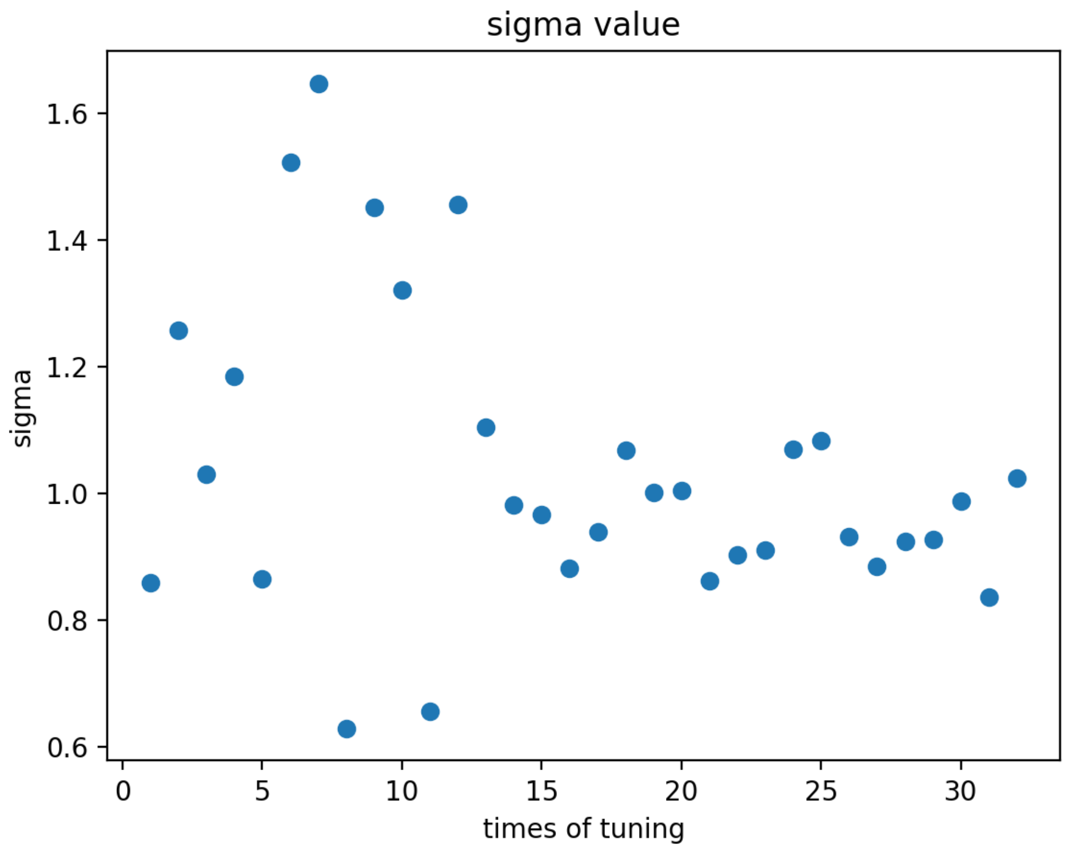

Figure 2: Sigma value

can be very small. This implies that the value of sigma could exceed 0.75, but the fitness is bad. If the fitness is good, the value of sigma could be close to 1 rather than just being greater than 0.75.

In order to confirm the radius of the trust region, we some works about trust region [14, 15], the update parameter $\eta 1, \eta 2$ are both less than 1 and they satisfy $0 < \eta 1 < \eta 2 < 1$. In this paper we set $\eta 1 = 0.25$ and $\eta 2 = 0.75$. If the $\sigma > 0.75$, we consider "increase the radius if the change is very successful , $\sigma \geq \eta 2$ "[16]. In our method, the surrogate model ensures a certain level of accuracy, preventing scenarios where the numerator significantly outweighs the denominator.

To validate whether the value of σ is reasonable, we extract and plot the sigma values during the optimization and validation convergence processes, as shown in Figure 2. It can be observed that all values are distributed between 0.6 and 1.6, indicating that the error of the surrogate model is generally controlled within a certain range, without exceptionally small or large outliers.

We believe that the setting of σ is free from anomalies and can effectively optimize the precipitation parameters of CAM5.

4. Improving the clarity of motivation for the nonuniform parameter parameterization scheme.

Thanks for your comment. We try to reply to this comment from the following points.

1) In this paper, the motivation for the nonuniform parameter parameterization scheme is as follows:

- It is well known that CAM5 is a well tuned model, however the holistic optimal parameters do not necessarily mean they are the best solutions for every region. The simulation results over these regions are challenging to improve through global tuning experiments. **Please note**

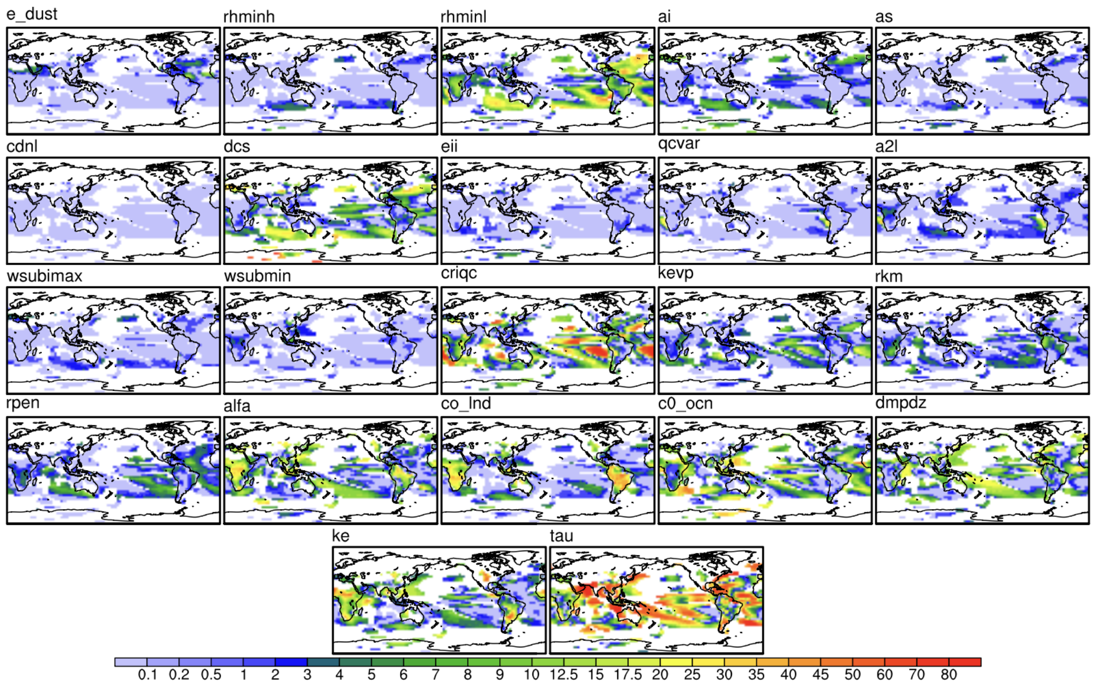

Figure 3: Global spatial distributions of relative contributions (%) of parameters to total variance of annual mean precipitation in [13].

that in this sentence "global" tuning represents the CAM5 tuning experiments to improve the global simulation results of precipitation rather than "global optimal" in optimization process.

- Regional optimization experiments demonstrate that some regions have optimal parameters, leading to better results than default parameters. However, applying these parameters to global simulations may not obtain optimal results; they are effective only within the selected regions for achieving the best simulation outcomes.

- Our experiments show that there is a "rocker effect" in the influence of parameters on precipitation. The same parameter values have different effects on different regions. When the simulation results in one region improve due to changes in parameter values, the results in other regions may decline. This implies that optimizing precipitation for all regions using a single set of parameters is challenging.

2) In [13], the authors discuss the contributions of different parameters to precipitation in different regions. The research results indicate that the contribution of different parameters to precipitation varies across regions. As shown in the figure 3, it can be observed that the contribution of parameters to precipitation cannot be simply judged based on the relationship with ocean or land. Even in adjacent regions, there can be some degree of differences. When using globally uniform parameter values, in order to pursue a holistic optimal solution, approximate mean value is employed to achieve a better overall simulation performance. If there is a significant difference between the local optimum and mean value in certain regions, the simulation results for that region will have a large error. **Please note that "local optimum" in this sentence means that the best parameters over this region, rather than "Local optimum solution" in the optimization process.**

3) We must also consider the diversity of the oceans regions. In CAM5, the physical processes

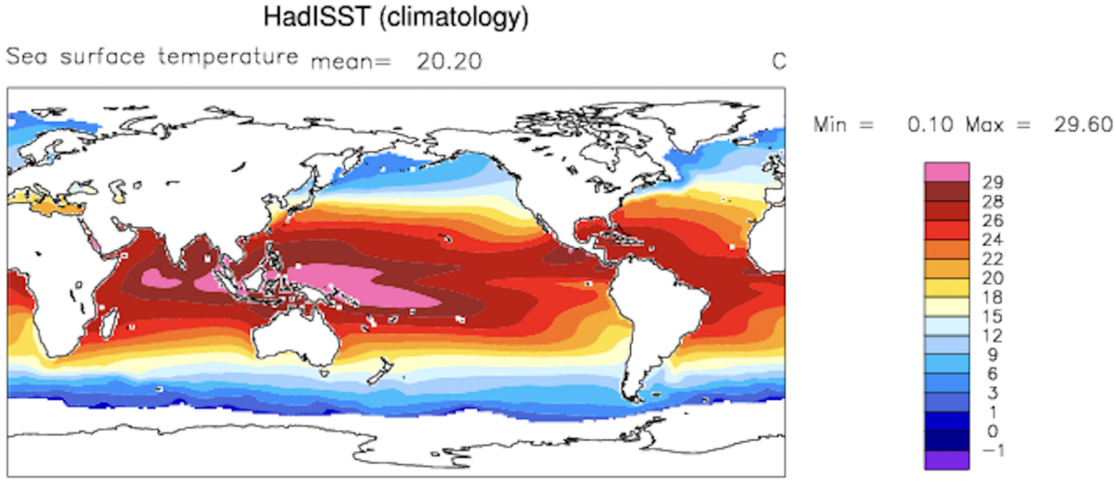

Figure 4: climatology SST

related to the ocean include optical reflection and some complex thermodynamic processes. The variations in sea surface temperatures have a significant impact on these physical processes, leading to substantial differences in simulation results across different regions. As can be seen in the Figure 4, the sea surface temperatures vary significantly across different oceanic regions, such as the Pacific region and warmpool. In the presence of such differences, dividing parameters based solely on ocean/land distinctions is not precise enough. Therefore, we try to select multiple regions and utilize faster parameter tuning method to find better parameters. These papameters are then integrated into the same case through a non-uniform parameterization scheme.

In summary, we proposed the nonuniform parameter parameterization scheme. We search different parameter combinations for different regions by surrogate model based tuning method and integrate them into a single case in a non-uniform parameterization scheme.

5. Line 55, while previous methods involved running the climate model, it is important to note that this work also requires running the climate model in each iteration. However, the manuscript does not provide a direct comparison of the efficiency of this method with other approaches. To enhance the evaluation of the proposed method, it would be beneficial to include an assessment of the computational cost compared to existing methods. This evaluation can provide valuable insights into the efficiency and computational advantages of the proposed approach, strengthening the manuscript's contribution in terms of computational performance.

Thanks four your comment. We are very willing to conduct some performance-related comparisons. However, for some commonly used parameter optimization algorithms, such as DE, PSO, GA, and so on, using these methods for parameter tuning in CAM can yield relatively good results. Nevertheless, these algorithms require more computational resources and time during execution, which makes it challenging to evaluate these methods based on performance.

We attempt to use algorithms such as PSO and GA for parameter tuning for the CWRF model (Climate-Weather Research and Forecasting model) [17]. We know that CWRF, as a weather forecasting model, has much shorter runtime and resource consumption compared to CAM5. We refer to these commonly used parameter settings, setting the population size to 50 and the number of iterations to 100. The results show that even for a model with significantly shorter execution time and computational resource consumption than CAM5, these optimization methods still struggle to obtain optimal parameters within a reasonable time frame. The results may even be insufficient to meet the timeliness requirements for CWRF predictions. While comparing the performance advantages of our method with other methods can demonstrate the superiority of our approach, the performance of these methods is challenging to quantify in a short time frame due to limitations such as time constraints and allocated computational resources.

Considering computational resources and time costs, we compare our method with the AMSO algorithm. The ultimate advantage in performance is the advantage in the number of iterations.

2 Replies to minor comments

Replies to minor issues are as follows:

1. The title uses CAM5, but the contexts use CAM. They could be consistent.

Thanks for your comment. We will use consistent definition of CAM5 and other technical terms over the whole manuscript.

2. Line 11: "selected points.." to "selected points."

Thanks for your comment. We will correct it in revised manuscript.

1. Line 29: traditional tuning methods in climate modeling have certain limitations. However, they remain highly useful. The majority of climate models employ traditional tuning approaches due to their reliance on well-established physics knowledge. In fact, automatic tuning methods require a solid understanding of physics to enhance their efficiency.

Thanks for your comment. We agree with your comment that manual parameter tuning remains necessary. This is because optimizing the parameters of atmospheric models requires a solid understanding of the underlying physics. Our proposed method is not intended to completely replace manual tuning but to enhance the efficiency of optimization. It is built on a foundation of substantial knowledge about the model. Using automated optimization methods, we aim to improve the tuning efficiency and reduce the consumption of computational resources. Perhaps the term "less useful" is not quite accurate. We will reconsider and use a more appropriate word to express our viewpoint.

2.Line 35, The statement that "WRF physics process is simple" is not accurate. In fact, it is known to be complex and intricate.

Thanks for your comment. We agree with your comment. WRF is indeed a complex model that involves many intricate physical processes. The confusion may have arisen from our choice of words. What we are trying to emphasize is not the complexity of the model but rather that WRF is geared towards local execution and short-term forecasting, which generally incurs lower resource costs for repeated runs. In contrast, CAM5 primarily focuses on global, long-term simulations, which result in longer execution times. Therefore, when it comes to optimizing parameters for CAM, it's challenging to apply methods involving many iterations. We will replace the ambiguous terms in line with your comment.

3. Line 37, The statement that "MVFSA may become infeasible for CAM tuning" may require further consideration. Fast simulated annealing, which is utilized in MVFSA, actually requires only one population to search for the next optimal parameters. The MVFSA requires thousands of steps to get a stable solution. But CAM requires a lot of computational cost for each optimization iteration. The authors should thoroughly discuss the challenges associated with MVFSA to provide a comprehensive understanding of its feasibility for CAM tuning.

Thanks for your comment. The term "infeasible" does not imply that these methods cannot be used for parameter tuning of CAM, as mentioned in the comments, MVSFA requires thousands of iterations. After these iterations, a better set of parameters can be obtained. However, from an efficiency perspective, even though this method can yield improved parameters, the computational cost and time required for thousands of iterations are deemed unacceptable. Therefore, in this paper, "infeasible" not only refers to the capability for optimization but also encompasses whether better parameters can be obtained through optimization within acceptable resource costs.

4. Line 51, When the optimization process reaches convergence, further iterations do not lead to any improvement. Similarly, once the optimization algorithm has obtained a local solution, additional iterations do not result in further enhancements. The effectiveness of the algorithm is also a determining factor in this regard.

Thanks for your comment. We agree this comment. Typically, in the normal operation of an algorithm, the optimal solution improves as the number of iterations increases. However, in some cases, increasing the number of iterations may not yield any better results. This can happen when the algorithm gets stuck in a local optimum, as mentioned in the comments, or when it has already converged. We will modify this sentence to make it less absolute in revised manuscript.

5. Line 58, It is confusing that 'the mathematical expression is complex and time-consuming'. Could you explain it?

Thanks for your comment. This sentence contains a punctuation error. We will rewrite it in reviesd manuscript.

The sentence can be rewrite as:

The objective function of these problems is difficult to describe as a mathematical expression or the mathematical expression is complex and time-consuming.

6. Line 59. Revise the sentence "Wang et al. ... ; a SCM-SMA hydrologic model"

Thanks for your comment. We will rewrite this sentence in revised manuscript.

The sentence can be rewrite as:

Wang et al. (2014) established a connection between the optimization of mathematical benchmarks and complex geoscientific models, and proposed the adaptive surrogate model-based optimization (ASMO) method. In order to demonstrate the performance of ASMO method for parameter tuning ofcomplex geoscientific models, parameters of a SCA-SMA hydrologic model was tuned based on the ASMO method.

7. Line 85, the authors could carefully analyze the challenge of ASMO used in atmospheric model. The method has been successfully used in WRF, CLM. what's the real challenge for atmospheric model?

Thanks for your comment. WRF is a regional model that focuses more on simulating specific regions, with a smaller spatial scale. The differences and variations in parameter values within a region are not particularly large. Therefore, the computational resource cost of parameter tuning is not exceptionally high. CLM is a land model designed specifically for simulating processes occurring on Earth's land. Considering that a significant portion of the Earth's surface is covered by oceans, which involve complex physical and chemical changes, as well as direct energy and substance exchanges between land and ocean, these aspects are not incorporated into CLM. In contrast, CAM includes physical processes related to the ocean, resulting in notable differences between the two models.

So that ASMO has been successfully applied to models like WRF and CLM, the application of surrogate model-based parameter tuning to CAM remains a challenge. In this paper, we designe a multi-level surrogate model-based method and demonstrate that the surrogate-based methods can enhance precipitation simulation results in CAM5.

8. Line 91, the above sentences discuss the tuning algorithms. The sentence "The precipitation process ..." talk about the metrics. It would be beneficial to separate these statements into individual

paragraphs.

Thanks for your comment. Yes, These sentences talk about challenge of surrogate model for precipitation parameter tuning, we will separate these statements into individual paragraphs in revised manuscript.

9. Line 110, it is hard to say the nonlinearity and complexity of CAM5 are much higher than WRF.

Thanks for your comment. We agree with your point. Perhaps it's not straightforward to conclude that the nonlinearity and complexity of CAM are necessarily higher than those of WRF, as both involve a significant amount of computation and complex physical processes. We will rephrase this sentence accordingly.

10. Section 2.1, describe more details of CAM5, such as horizontal resolution, vertical level, how long does CAM5 run, the sst and sea ice are used prescribed seasonal climatology.

Thanks for your comment. We will add more description about CAM5 and the compset used in this study. They include the modes used in the compset, the description of the grid and their specific meanings. We will supplement to the revised manuscript based on [18]

11. Line 138: define the six main regions, giving a table including the range of latitude and longitude.

Thanks for your comment. In the manuscript, the Table 1 we introduce the region selected in the study and we will add the cite of the table in revised manuscript.

12. Line 143, why not use GPCP to estimate precipitation but use ERA5.

Thanks for your comment. Both ERA5 and GPCP can be used for precipitation analysis. They both provide global precipitation data. However ERA5 provides higher-resolution data. Data of GPCP are provided on a 2.5 degree grid and ERA5 precipitation data are provided on a 0.25 degree grid. We believe that choosing data with higher resolution can significantly contrast the tuning results, thereby demonstrating the effectiveness of the proposed method. So that we select ERA5 instead of GPCP as the metric for precipitation parameter tuning. The RMSE is calculated between the CAM5 simulation results and ERA5 reanalysis data.

13. Line 147, "Makes" to "makes"

Thanks for your comment. We will correct it in revised manuscript.

14. Line 163, is the "sampling method" is the latin hypercube sampling? How many samples do you conduct?

Thanks for your comment. Yes, the "sampling method" is latin hypercube sampling. There are 60 samples we conduct. They are described in section 3.2.1.

15. Line 280, use the correct ref for GP.

Thanks for your comment. We will add the correct ref for GP in revised manuscript.

16. Line 308, It is confusing that the surrogate model is built as the quadratic function. Does it use GP?

Thanks for your comment. We use the GP to construct the local-level surrogate model, the "quadratic function" means that in the initial mathematical theory of trust region , a quadratic function is used to fit the real function. These sentence are used to describe the trust region theory rather than introduce our proposed method.

17. For fig2, what is the y-axis? Is it the relative error? How calculate it?

Thanks for your comment. the y-axis represents the relative error between the predict value and real value, it is calculated based on Eq.1:

$$relative\ error = \frac{|precdictvalue - realvalue|}{realvalue} \tag{1}$$

Where, the *precdictvalue* represents the objective function value (in this paper is RMSE) predicted by surrogate model and *realvalue* represents the objective function value (in this paper is RMSE) obtained by CAM5 simulation. We will revise this paragraphs according to this comment.

18. Line 343, 'lower' to 'lowest'.

Thanks for your comment. We will correct it in revised manuscript.

19. Section 4.1.3 should be merged into section 4.1.2.

Thanks for your comment. We will reorganize the structure according to the the comment and other comments about the structure.

20.In figure 4, it should include the obs pattern, or the difference between opt/default and observation.

Thanks for your comment. We will add new results figures in revised manuscript according to this comment.

21.Line 366, it is confusing for this sentence "Therefore, we need to further ..."

Thanks for your comment. In previous sentences. We illustrate that the result of global-based tuning is not significant, some regions still need to be tuned. So that we try to find the best way to use the proposed method to improve the simulation result. Perhaps our choice of words was not precise. We will use more accurate term in this sentence.

References

[1] X. Wu, L. Hu, L. Wang, H. Lu, and J. Zheng. Surrogate model-based precipitation tuning for cam5. *Geoscientific Model Development Discussions*, 2023:1–30, 2023.

[2] Jakub Kudela and Radomil Matousek. Recent advances and applications of surrogate models for finite element method computations: a review. *Soft Computing*, 26(24):13709–13733, December 2022.

[3] Timothy W Simpson, Andrew J Booker, Dipankar Ghosh, Anthony A Giunta, Patrick N Koch, and R-J Yang. Approximation methods in multidisciplinary analysis and optimization: a panel discussion. *Structural and multidisciplinary optimization*, 27:302–313, 2004.

[4] Rommel G. Regis and Christine A. Shoemaker. A Stochastic Radial Basis Function Method for the Global Optimization of Expensive Functions. *INFORMS Journal on Computing*, 19(4):497–509, November 2007.

[5] Ky Khac Vu, Claudia d'Ambrosio, Youssef Hamadi, and Leo Liberti. Surrogate-based methods for black-box optimization. *International Transactions in Operational Research*, 24(3):393–424, 2017.

[6] Yong Wang, Da-Qing Yin, Shengxiang Yang, and Guangyong Sun. Global and local surrogate-assisted differential evolution for expensive constrained optimization problems with inequality constraints. *IEEE transactions on cybernetics*, 49(5):1642–1656, 2018.

[7] Hugh Morrison and Andrew Gettelman. A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (cam3). part i: Description and numerical tests. *Journal of Climate*, 21(15):3642–3659, 2008.

[8] Andrew Gettelman, Hugh Morrison, and Steven J Ghan. A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (cam3). part ii: Single-column and global results. *Journal of Climate*, 21(15):3660–3679, 2008.

[9] Andrew Gettelman, Xiaohong Liu, Steven J Ghan, Hugh Morrison, Sungsu Park, AJ Conley, Stephen A Klein, James Boyle, DL Mitchell, and J-LF Li. Global simulations of ice nucleation and ice supersaturation with an improved cloud scheme in the community atmosphere model. *Journal of Geophysical Research: Atmospheres*, 115(D18), 2010.

[10] GJ Zhang and NA McFarlane. Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate centre general circulation model. *Atmosphere Ocean*, 33:407–446, 1995.

[11] Richard B Neale, Jadwiga H Richter, and Markus Jochum. The impact of convection on enso: From a delayed oscillator to a series of events. *Journal of climate*, 21(22):5904–5924, 2008.

[12] S Park, CS Bretherton, and PJ Rasch. The revised cloud macrophysics in the community atmosphere model. *Journal of Climate*, 2010.

[13] Yun Qian, Huiping Yan, Zhangshuan Hou, Gardar Johannesson, Stephen Klein, Donald Lucas, Richard Neale, Philip Rasch, Laura Swiler, John Tannahill, Hailong Wang, Minghuai Wang, and Chun Zhao. Parametric sensitivity analysis of precipitation at global and local scales in the Community Atmosphere Model CAM5. *Journal of Advances in Modeling Earth Systems*, 7(2):382–411, June 2015.

[14] Stefan M Wild, Rommel G Regis, and Christine A Shoemaker. Orbit: Optimization by radial basis function interpolation in trust-regions. *SIAM Journal on Scientific Computing*, 30(6):3197–3219, 2008.

[15] Ruobing Chen, Matt Menickelly, and Katya Scheinberg. Stochastic optimization using a trust-region method and random models. *Mathematical Programming*, 169:447–487, 2018.

[16] Kely DV Villacorta, Paulo R Oliveira, and Antoine Soubeyran. A trust-region method for unconstrained multiobjective problems with applications in satisficing processes. *Journal of Optimization Theory and Applications*, 160:865–889, 2014.

[17] Xin-Zhong Liang, Hyan I Choi, Kenneth E Kunkel, Yongjiu Dai, Everette Joseph, Julian XL Wang, and Praveen Kumar. Development of the regional climate-weather research and forecasting (cwrf) model: surface boundary conditions. *Scientific report*, 1, 2005.

[18] Richard B Neale, Chih-Chieh Chen, Andrew Gettelman, Peter H Lauritzen, Sungsu Park, David L Williamson, Andrew J Conley, Rolando Garcia, Doug Kinnison, Jean-Francois Lamarque, et al. Description of the ncar community atmosphere model (cam 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, 1(1):1–12, 2010.