Manuscript prepared for Geosci. Model Dev.
with version 2014/07/29 7.12 Copernicus papers of the LaTeX class copernicus.cls.
Date: 14 January 2015

An automatic and effective parameter optimization method for model tuning

Tao Zhang[1]

[1]Tsinghua

Correspondence to: Tao Zhang (t-zhang11@mails.tsinghua.edu.cn)

Abstract.

Physical parameterizations in GCMs, having various uncertain parameters, greatly impact model performance and model climate sensitivity. Traditional manual and empirical tuning of these parameters is time consuming and ineffective. In this study, a "triple-step" methodology is proposed to automatically and effectively obtain the best/optimum combination of some key parameters in cloud and convective parameterizations based on a comprehensive objective evaluation metrics. It is found that the optimum combination of these parameters determined using this method is able to improve the model overall performance by 9% in a GCM. The method can be easily applied to other GCMs to speed up model development process, especially regarding unavoidable comprehensive parameter tuning in model development.

1 Introduction

Sub-grid scale physical processes are presented as empirical or statistic parameters in climate system models (Hack et al., 1994). The parameterization physical processes approximate the unresolvable scale dynamic and thermodynamics (Williams, 2005), consequently, introducing to simulations uncertainties for investigating the climate change using climate system models (Warren and Schneider, 1979). The uncertain parameters are required to calibrate when new or improved parameterized schemes are integrated into models (Li et al., 2013).

Traditionally, the uncertain parameters are manually tuned by analysis the relationship between simulations and observations. This calibration is somewhat subjective and hard to manipulate due to the tedious labor intensive work (Hakkarainen et al., 2012; Allen et al., 2000). Currently, the automatic calibration technique is a hot topic in the climate system model uncertainty quantification. The previous work focuses on the method of posterior range and probability, optimization algorithms, and data assimilation technique. The first class method, the optimization parameters confidence range is evaluated based on likelihood and bayesian estimation. Cameron et al. (1999) improves the forecast by the generalized likelihood uncertainty estimation (GLUE), a method of obtaining parameters uncertain range of a specified confidence level. The Bayesian Markov chain Monte Carlo (MCMC) is widely used to obtain posterior probability distribution from prior knowledge. Hararuk et al. (2014) calibrates soil C data in CLM-CASA model, a global land model consisting of biogeophysics and biogeochemistry processes using MCMC approach. Sun et al. (2013) demonstrates the possibility of calibration of hydrologic parameters with MCMC in CLM4. Jackson et al. (2008) obtains 6 parameters posterior probability from clouds and convection physical process in CAM3.1 by Multiple Very Fast Simulated Annealing (MVFSA) to optimize a comprehensive metrics, including cloud, radiation, temperature, and precipitation, wind, as well as humidity variables.The second class method, the optimization algorithms search the maximum or minimum metrics value in the given parameters space.Severijns and Hazeleger (2005) calibrates parameters of radiation, clouds, and convection in Speedy with down-hill simplex, to improve the radiation budget at the top of the atmosphere and at the surface, and the large scale circulation. Down-hill simplex is a fast convergence algorithm when the parameters space is not high. The changed geometry represents the optimal direction in the method, instead of the gradient information, such as Newton and quasi-Newton. Compared with down-hill simplex, the evolution algorithm, such as MVFSA, (Jackson et al., 2004; Yang et al., 2014), simulated stochastic approximation annealing (SSRR) (Yang et al., 2013), multi-objective genetic algorithm (MOGA) (Swiler et al.) are the global optimization algorithms, also used to automatically tune the uncertain parameters. Gill et al. (2006) Another class method, data assimilation method become another research direction of parameters calibration, such as Ensemble Kalman filter (ENKF) (Aksoy et al., 2006; DelSole and Yang, 2010), Extended Kalman filter (EKF) (Carrassi and Vannitsem, 2011), as well as Particle filtering (PF) (Snyder et al., 2008).

However, posterior parameter distribution methods based MCMC and the global optimization evolutionary algorithms require at least ten thousand steps to obtain the stability solution. The latter and filter method, such as ENKF and PF, require multi-individuals in each iterations, leading to high computation cost. Taken into account high dimension uncertain parameters space, strong nonlinear relationship between parameters and simulations, the optimization algorithms can not search the optimal parameters inefficiently. The above work mostly use the single step to calibrate uncertain parameters. Zhang et al. presents a “two-steps” strategy, conducting a step of pre-processing initial values of optimization algorithms before tuning with down-hill simplex. Nevertheless, the tuned uncertain parameters are also selected by experts, suffering to subjective experience, hardly quantifying the parameters sensitivity.

In this paper, we propose a “three-steps” strategy based on the “two-steps”. In the first step, a global sensitivity analysis method, Morris (Morris, 1991; Campolongo et al., 2007), eliminates the insensitivity parameters by analyzing the main and interaction effects among parameters. Another global method, Sobol (Sobol, 2001), is used to verify the Morris results. Taken into account the complex configuration and manipulation of the "three-steps", invoking parameters sampling, namelist reset, models running simultaneously, optimization iteration, sensitivity analysis, as well as metrics diagnostics, a automatic workflow is provided to make the calibration process more efficient. The "three-step" calibration strategy is applied to GAMIL2, a Grid-point Atmospheric Model developed by State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, China, to improve the comprehensive simulation performance of the climate mean state.

2 Model and observations

2.1 GAMIL2 atmospheric model

In this paper, the Grid-point Atmospheric Model of IAP LASG version 2 (GAMIL2) is used. It takes part in the Atmospheric Model Inter-comparison Project (AMIP) of IPCC AR5, Cloud Feedback Model Inter-comparison Project (CFMIP) and Coupled Model Intercomparison Project Phase 5 (CMIP5) as an atmospheric component of Flexible Global-Ocean-Atmosphere-Land System Model grid version 2 (FGOALS-g2). The horizontal resolution is 2.8 x 2.8, with 26 vertical levels. The dynamical core of GAMIL2 is a finite difference scheme, and conserves mass and effective energy (Wang et al., 2004). The moisture equation adopts the two- step shape-preserving advection scheme (Rucong, 1994). Compared with GAMIL1, the previous version, GAMIL2 upgrade cloud- related process (Li et al., 2013), such as the deep convection parameterization (Zhang and Mu, 2005), the convective cloud fraction (Xu and Krueger, 1991), and the cloud microphysical (Morrison and Gettelman, 2008). The initial calibrated parameters are selected from deep convection, shallow convection, cloud fraction, cloud microphysical processes and boundary layer scheme, as table 1 shown. The default parameters values are the configuration of the standard version, which takes part in IPCC AR5 and is called CNTL.

2.2 Observational data

Observed wind, humidity, and geopotential height derive from the European Center for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (EEA) - Interim reanalysis, 1989 to 2004 and 1.5x1.5 horizontal resolution (Simmons et al., 2007). The precipitation comes from the Global Precipitation Climatology Project (GPCP), 1989 to 2004 and 2.5x2.5 horizontal resolution (Adler et al., 2003) . The radiation variables use the Earth Radiation Budget Experiment (ERBE), 1985 to 1989 and 1.875x1.875 horizontal resolution (Barkstrom, 1984). The climate mean state of observational data are require to remap to the grid of GAMIL2. However, the duration of simulation (2000-2004) is inconsistent with the observational data. We conduct a long simulation (1989-2004) with the optimal parameters to insure the tuning parameters validity. The results show the consistent with the short simulation.

3 Methods

3.1 Metrics

A comprehensive metrics, including wind, temperature, humidity, geopotential height, precipitation and radiation flux is used to quantitatively evaluate the simulation performance, to improve overall simulation skills (Murphy et al., 2004; Gleckler et al., 2008; Reichler and Kim, 2008) . These variables are shown as table 2. The model starts up in the 2000th year, and simulates 5 years. Climate mean state of the last three years is used to diagnostic metrics. The calibration RMSE is defined as the spatial standard deviation (SD) of model simulation against observations, as equation (1) (Taylor, 2001; Yang et al., 2013). In the propose of integrating the 16 variables into a unique metrics, the SD of default GAMIL2 simulations against observations is used to scale each variable calibration RMSE, as equation (2). Finally, the metrics is computed as the average value of each scaling variable factor, as equation (3). Therefore, if the metrics is lower than one, the uncertain parameters have been improved.

$$(\sigma_m^F)^2 = \sum_{i=1}^{l} w(i)(x_m^F(i) - x_o^F(i))^2 \tag{1}$$

$$(\sigma_r^F)^2 = \sum_{i=1}^{l} w(i)(x_r^F(i) - x_o^F(i))^2 \tag{2}$$

$$\chi^2 = \frac{1}{N^F} \sum_{F=1}^{N^F} (\frac{\sigma_m^F}{\sigma_r^F})^2 \tag{3}$$

$x_m^F(i)$ is the model outputs according to selected shown in the Table 2. $x_o^F(i)$ is the corresponding observation or reanalysis data. $x_r^F(i)$ is the reference results from CMIP5. w is the weight due to the different grid area. I is the total grid number in model. N^F is the number of the chosen variables.

3.2 "Three-steps" calibration strategy

With contrast to the manual and optimal algorithms calibration, the "three-steps" provides an effective and efficient automatic calibration strategy, and aims at reducing the dimension of tuning space by eliminating the insensitivity parameters, and reducing the iteration steps by pre-processing initial values of optimal algorithms. Finally, an inexpensively computational algorithm, down-hill simplex, is used to search the optimal solution.

3.2.1 Reducing Dimension

Due to the high complex of physical parameterization process, there are a large number of uncertain parameters in climate system model. Moreover, the prior parameters values are usually set with a relatively large range. Most of optimization algorithms, such as genetic algorithm, down-hill simplex, and simulated annealing are ineffective in the high dimension space. Additionally, the atmospheric models require a long time to spin up, leading to the extremely long calibration computational cost.

Instead of the correlation coefficient, only presenting the linear relationship, Morris, is a qualitative global sensitivity method based on one-step-at-a-time (OAT) experiment design using relatively few samples. Not only the single parameter sensitivity can present, but also the interaction sensitivity among parameters can describe. It introduces MOAT sampling technology, requiring only $(n+1) \times M$ points, where n is the number of calibration parameters and M is number of trajectories, usually 10-20. Consider the n parameters $x_i(i=1,...,n)$, normalized to [0,1]. The influence of each variables is defined as a elementary effect, show as equation (4), where Δ is the value of $1/p-1,...,p-2/p-1$, and p is the sampling level. The starting point of a trajectory is selected randomly and completed it by adjusting one unchanged parameter value at a time in random order, consisting $n+1$ samplings. The mean of $|d_j|$ can stand for the main effect of a single parameter, and the standard deviation presents the interactive effect of multi parameters. Therefore, those with low mean and low standard deviation will be eliminated. In this paper, parameters in table 2 are required to analyze sensitivity. We sample 80 points, and the results are showed in Figure 1. The insensitivity parameters, ke, capelmt and c0 of shallow convection, are removed.

$$d_{ij} = \frac{y(X_1,...,X_j+\Delta,...,X_N) - y(X_1,...,X_j,...,X_N)}{\Delta} \tag{4}$$

$$\mu_j = avg(|d_{i,j}|), \sigma_j = stddev(d_{i,j}) \tag{5}$$

In order to verify the Morris results, we adopt another sensitivity method, Sobol. It is a quantitative method based on variance decomposition and requires more samples than the Morris, with a higher computation cost. The variance of model output can be decomposed as equation (7), where n is the number of parameters, and V_i is the variance of the i_{th} parameter, and V_{ij} is variance of iterative effect between the i_{th} and j_{th} parameters, and so on. The total sensitivity effect can be presented as equation (8), where V_{-i} is the total variance except for the x_i parameter. The Sobol results are showed as figure 2. The screened out parameters are the same as the Morris.

$$V = \sum_{i=1}^{n} V_i + \sum_{1 \le i < j \le n} V_{ij} + ... + V_{1,2,...,n} \tag{6}$$

$$S_{T_i} = 1 - \frac{V_{-i}}{V} \tag{7}$$

V_{-i} is the total variance except for X_i

3.2.2 The initial values preprocessed Downhill Simplex

Parameters tuning in climate system model is a global optimal problem. The popular evolutionary algorithms, requires a population of individuals in each iterations, bringing in extremely computation cost. Nevertheless, the downhill simplex searches the optimal solution by changing the shape of a simplex , which represents the optimal direction and step, similar to the gradient information of Newton optimal method. A simplex is the geometry, consisting $N + 1$ vertexes and their interconnecting edges, where N is the number of calibration parameters screened by the 3.2.1 section. The vertexes stand for the pair of a set of parameters and its metrics. The new vertex is determined by expanding and shrinking the vertex with the highest metrics value, restructuring the new simplex. The detail of the downhill simplex is described in Press et al. (1992) and Nelder and Mead (1965)

This optimal algorithm can be fast convergence when the dimension of calibration parameters is not high. If the simplex is a large span, it can jump out the local searching area by reflecting geometry and retractable transforming. Otherwise, it oozes down the valley area and traps into the local optimal space. Therefore, the searching speed rely heavily on the initial values of this method. It is necessary to sample in the full range to determine the likely area of the optimal solution. The full factor method is an equilibrium distance sampling strategy, and is suitable to analyze the parameters values sensitivity in the low dimension space. We use it to sample in the given range as shown table 1, and refine sampled in a sensitivity range. The parameter sets with the smaller metrics will be selected. It is likely that one of the initial value is near to the optimal solution. The down-hill simplex algorithm can converge toward this point.

Besides, the inappropriate initial values may lead to the ill-conditioned simplex geometry. It means that some parameters keep the same value in the initial values. These parameters can not change in the optimal algorithm. Consequently, as much as different sampled parameters values are required to select to improve the parameters freedom of initial values. The preprocessing initial values of downhill simplex is presented as the Algorithm 1.

4 Design of an end-to-end calibration workflow

Taken into account the complex configuration and operation of the “three-steps”, resulting in the inefficient manipulation, we design and implement the automatic parameters calibration workflow, as figure 1 shown. It consists of four components, dimension reducing, calibration algorithms, post -processing, and tasks schedule. Meanwhile, it integrates some open source tools, such as PSUADE for sensitivity analysis, DAKOTA for calibration algorithms, NCO and NCL for metrics diagnostic. The input of the framework is calibration parameters of interest and their initial value range. The output is the optimal parameters and its corresponding diagnostic results after calibrating with the “three-steps” or other strategy.

Other than Morris and Sobol, the dimension reducing module also provide multi sampling methods, such as full factor, Latin Hypercube (LH), Morris one-at-a-time (MOAT) and Central Composite Designs (CCD), used to preprocessing the initial values of calibration algorithms and others sensitivity analysis. It could produce the duplicate sampling point in some sampling method, such as MOAT and CCD. The preprocessing module supports automatically eliminate duplicate points, reducing the unnecessary computing loads. As well, Markov Chain Monte Carlo (MCMC) method based on adaptive Metropolis-Hastings algorithms is also provided to get the posterior distribution of uncertain parameters. The calibration algorithm module offers the local and global optimization algorithms as figure 2 shown. Due to most of optimization algorithms and samplings require multi cases running synchronization, the schedule module supports flexible schedule to take fall use of computation resource. The post-processing module is response for metrics diagnostic, reanalysis and observation data management. Additional, all the intermediate metrics and their corresponding parameters are stored in DATABASE, used for posterior knowledge analysis.

5 The optimization results and mechanism analysis

5.1 Comparison the effective and efficient with different strategies

We compare the effective and efficient performance among the “three-steps”, “two-steps” of initial values pre-processing and down-hill simplex, and “one-step” directly using optimal algorithms with downhill simplex, differential evolution (DE) and particle swarm optimization (PSO) , as showed in table 3. The “two-steps” and “three-steps” are required extra 25 samples for the initial values pre-processing and 80 samples for parameters space reducing. The 3rd column in table 3 is the number of iterations when each method gets the optimal solution in the 2nd column. The real iteration steps are greater than the number in table 3. The size of population of DE and PSO is set as 12. The effective of each strategy is evaluated by the optimal solution, and the efficiency is evaluated by the core hours, computed as $N_{step} \times N_{size} \times 30 \times 6$, where N_{step} is the number of iterations, and N_{size} is the size of population. The model runs as 30 processes, each assigned one core, and simulates 5 model years, about 6 hours. In the “one- steps” strategy, though the PSO gets a good optimal solution, it spends a expensive computational cost. On the contrary, the “three-steps” gets the best solution using half core hours than the PSO. The “two-steps” has the best efficiency, while the key factor, optimal solution is worse than the “three-steps”.

5.2 The optimal mechanism analysis

Figure 4 shows the metrics of each variables with the global, tropic, and northern / southern middle and high latitudes (NMHL / SMHL). Most variables of global are improved compared with CNL. Specific Humidity at 400hPa (Q400) is the best improved. Two variables, Meridional wind at 200hPa (V200) and clearsky short wave net flux at TOA (FSNTOAC), keep the same magnitude. 850hPa and 200hPa temperature are worse than CNL. From the spatial distribution perspective, the SMHL contributes the best improved.

Figure 4 presents the improvement of the entire radiation variables. It owes to the specific humidity and cloud improvement in the middle and upper troposphere. The EXP consumes the more water vapor and gets the better simulation than CNTL in this vertical height, shown as figure 5. The decrease of the atmospheric water vapor reduces the its greenhouse effect. Therefore, it emits the more outgoing long-wave radiation in clean sky, reducing the simulation error of clear sky long wave upward flux at TOA (FLUTC), shown as figure 7(a).

Compared with CNTL, the middle and high cloud significantly increase, as a result of the reducing rhminh parameter, shown as figure 6. Consequently, it enhances the blocking effect on the long wave upward flux at TOA (FLUT), reducing the FLUT in 30°-60° of the southern and northern hemisphere in cloud, shown as figure 7(b). Taken into account the long wave cloud forcing (LWCF) computed as $FLUTC - FLUT$, the LWCF should increase, because of the increasing FLUTC and the decreasing FLUT, shown as figure 7(c).

6 Conclusions

In this paper, we presents the “three-steps” strategy of uncertain physical parameters in GAMIL2.The initial high parameters space is reduced by the global sensitivity methods, Morris and Sobol. To improve the convergence speed of down-hill simplex, a initial values pre-processing optimal algorithms is used based on the full factor sampling to evaluate the likely area of the optimal solution. The experiment results show the “three-steps” outperforms the PSO and DE in both of effective and efficiency. Though the “two-steps” has an advantage in efficiency, the optimal solution is worse than the “three- steps”. Besides, if the parameters dimension increases, it may not get the best performance. Taken into account the complex configuration and manipulation, we design and implement the automatic parameters calibration workflow to further enhance operation efficiency and to support multi uncertainty quantification analysis and calibration strategy. The optimal results of the “three-steps” demonstrate most of the variables are improved compared with the CNTL, especially for the radiation variables, which present the entirely improved. The mechanism analysis explores that the reduced simulation errors of water vapor has direct effect on the FLUT, and via cloud fraction indirectly influences the FLUTC. The LWCF is improved under the improved FLUT and FLUTC.

For future work, we plan to test the DE and PSO in the “three-steps”. We will also develop the computation cheap surrogate model, because the high computational cost is the one of the biggest challenge for calibrating the climate system model. Besides, the optimal results need more mechanism analysis to deeply understand the physical parameterization processes.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., et al.: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present), Journal of Hydrometeorology, 4, 1147–1167, 2003.

Aksoy, A., Zhang, F., and Nielsen-Gammon, J. W.: Ensemble-based simultaneous state and parameter estimation with MM5, Geophysical research letters, 33, 2006.

Allen, M. R., Stott, P. A., Mitchell, J. F., Schnur, R., and Delworth, T. L.: Quantifying the uncertainty in forecasts of anthropogenic climate change, Nature, 407, 617–620, 2000.

Barkstrom, B. R.: The earth radiation budget experiment (ERBE), Bulletin of the American Meteorological Society, 65, 1170–1185, 1984.

Cameron, D., Beven, K. J., Tawn, J., Blazkova, S., and Naden, P.: Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty), Journal of Hydrology, 219, 169–187, 1999.

Campolongo, F., Cariboni, J., and Saltelli, A.: An effective screening design for sensitivity analysis of large models, Environmental modelling & software, 22, 1509–1518, 2007.

Carrassi, A. and Vannitsem, S.: State and parameter estimation with the extended Kalman filter: an alternative formulation of the model error dynamics, Quarterly Journal of the Royal Meteorological Society, 137, 435–451, 2011.

DelSole, T. and Yang, X.: State and parameter estimation in stochastic dynamical models, Physica D: Nonlinear Phenomena, 239, 1781–1788, 2010.

Gill, M. K., Kaheil, Y. H., Khalil, A., McKee, M., and Bastidas, L.: Multiobjective particle swarm optimization for parameter estimation in hydrology, Water Resources Research, 42, 2006.

Gleckler, P. J., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, Journal of Geophysical Research: Atmospheres (1984–2012), 113, 2008.

Hack, J. J., Boville, B., Kiehl, J., Rasch, P., and Williamson, D.: Climate statistics from the National Center for Atmospheric Research community climate model CCM2, Journal of Geophysical Research: Atmospheres (1984–2012), 99, 20 785–20 813, 1994.

Hakkarainen, J., Ilin, A., Solonen, A., Laine, M., Haario, H., Tamminen, J., Oja, E., and Järvinen, H.: On closure parameter estimation in chaotic systems, Nonlinear Processes in Geophysics, 19, 127–143, 2012.

Hararuk, O., Xia, J., and Luo, Y.: Evaluation and improvement of a global land model against soil carbon data using a Bayesian Markov chain Monte Carlo method, Journal of Geophysical Research: Biogeosciences, 119, 403–417, 2014.

Jackson, C., Sen, M. K., and Stoffa, P. L.: An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions, Journal of Climate, 17, 2828–2841, 2004.

Jackson, C. S., Sen, M. K., Huerta, G., Deng, Y., and Bowman, K. P.: Error reduction and convergence in climate prediction, Journal of Climate, 21, 6698–6709, 2008.

Li, L., Wang, B., Dong, L., Liu, L., Shen, S., Hu, N., Sun, W., Wang, Y., Huang, W., Shi, X., et al.: Evaluation of grid-point atmospheric model of IAP LASG version 2 (GAMIL2), Advances in Atmospheric Sciences, 30, 855–867, 2013.

Morris, M. D.: Factorial sampling plans for preliminary computational experiments, Technometrics, 33, 161–174, 1991.

Morrison, H. and Gettelman, A.: 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, 3642–3659, 2008.

Murphy, J. M., Sexton, D. M., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Nelder, J. A. and Mead, R.: A simplex method for function minimization, The computer journal, 7, 308–313, 1965.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B.: Numerical Recipes in Fortran, Cambridge Univ, Press, Cambridge, p. 70, 1992.

Reichler, T. and Kim, J.: How well do coupled models simulate today's climate?, Bulletin of the American Meteorological Society, 89, 303–311, 2008.

Rucong, Y.: A two—step shape—preserving advection scheme, Advances in Atmospheric Sciences, 11, 479–490, 1994.

Severijns, C. and Hazeleger, W.: Optimizing parameters in an atmospheric general circulation model, Journal of climate, 18, 3527–3535, 2005.

Simmons, A., Uppala, S., Dee, D., and Kobayashi, S.: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, ECMWF newsletter, 110, 25–35, 2007.

Snyder, C., Bengtsson, T., Bickel, P., and Anderson, J.: Obstacles to high-dimensional particle filtering, Monthly Weather Review, 136, 4629–4640, 2008.

Sobol, I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and computers in simulation, 55, 271–280, 2001.

Sun, Y., Hou, Z., Huang, M., Tian, F., and Ruby Leung, L.: Inverse modeling of hydrologic parameters using surface flux and runoff observations in the Community Land Model, Hydrology and Earth System Sciences, 17, 4995–5011, 2013.

Swiler, L. P., Wildey, T. M., and Dalbey, K.: Uncertainty Assessment in Atmospheric Component of Climate Models.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres (1984–2012), 106, 7183–7192, 2001.

Wang, B., Wan, H., Ji, Z., Zhang, X., Yu, R., Yu, Y., and Liu, H.: Design of a new dynamical core for global atmospheric models based on some efficient numerical methods, Science in China Series A: Mathematics, 47, 4–21, 2004.

Warren, S. G. and Schneider, S. H.: Seasonal simulation as a test for uncertainties in the parameterizations of a Budyko-Sellers zonal climate model, Journal of the Atmospheric Sciences, 36, 1377–1391, 1979.

Williams, P. D.: Modelling climate change: the role of unresolved processes, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 363, 2931–2946, 2005.

Xu, K.-M. and Krueger, S. K.: Evaluation of cloudiness parameterizations using a cumulus ensemble model, Monthly weather review, 119, 342–367, 1991.

Yang, B., Qian, Y., Lin, G., Leung, L. R., Rasch, P. J., Zhang, G. J., McFarlane, S. A., Zhao, C., Zhang, Y., Wang, H., et al.: Uncertainty quantification and parameter tuning in the CAM5 Zhang-McFarlane convection scheme and impact of improved convection on the global circulation and climate, Journal of Geophysical Research: Atmospheres, 118, 395–415, 2013.

Yang, B., Zhang, Y., Qian, Y., Huang, A., and Yan, H.: Calibration of a convective parameterization scheme in the WRF model and its impact on the simulation of East Asian summer monsoon precipitation, Climate Dynamics, pp. 1–24, 2014.

Zhang, G. J. and Mu, M.: Effects of modifications to the Zhang-McFarlane convection parameterization on the simulation of the tropical precipitation in the National Center for Atmospheric Research Community Climate Model, version 3, Journal of Geophysical Research: Atmospheres (1984–2012), 110, 2005.

Zhang, T., Xie, F., Xue, W., Li, L., Xu, H., and Wang, B.: Quantification and optimization of parameter uncertainty in the Grid-point Atmospheric Model GAMIL2, Chinese Journal of Geophysics.

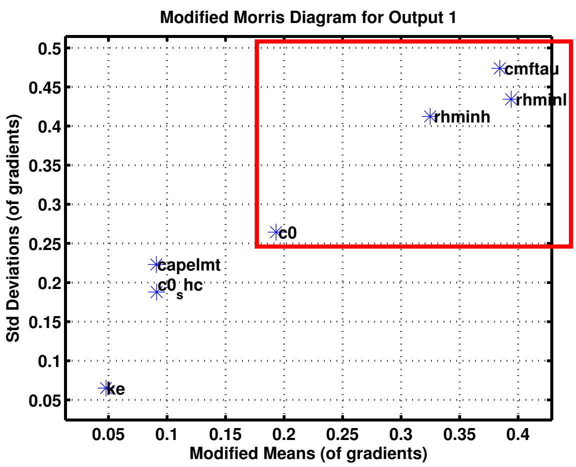

Figure 1. Morris results

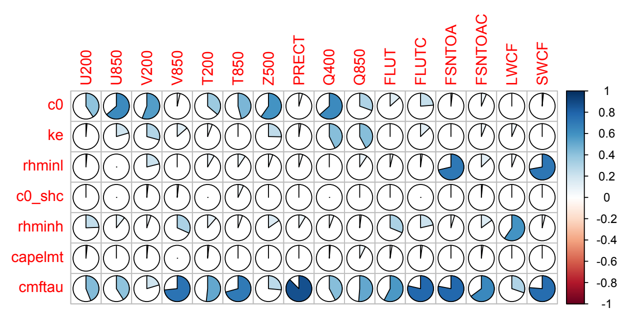

Figure 2. Sobol results

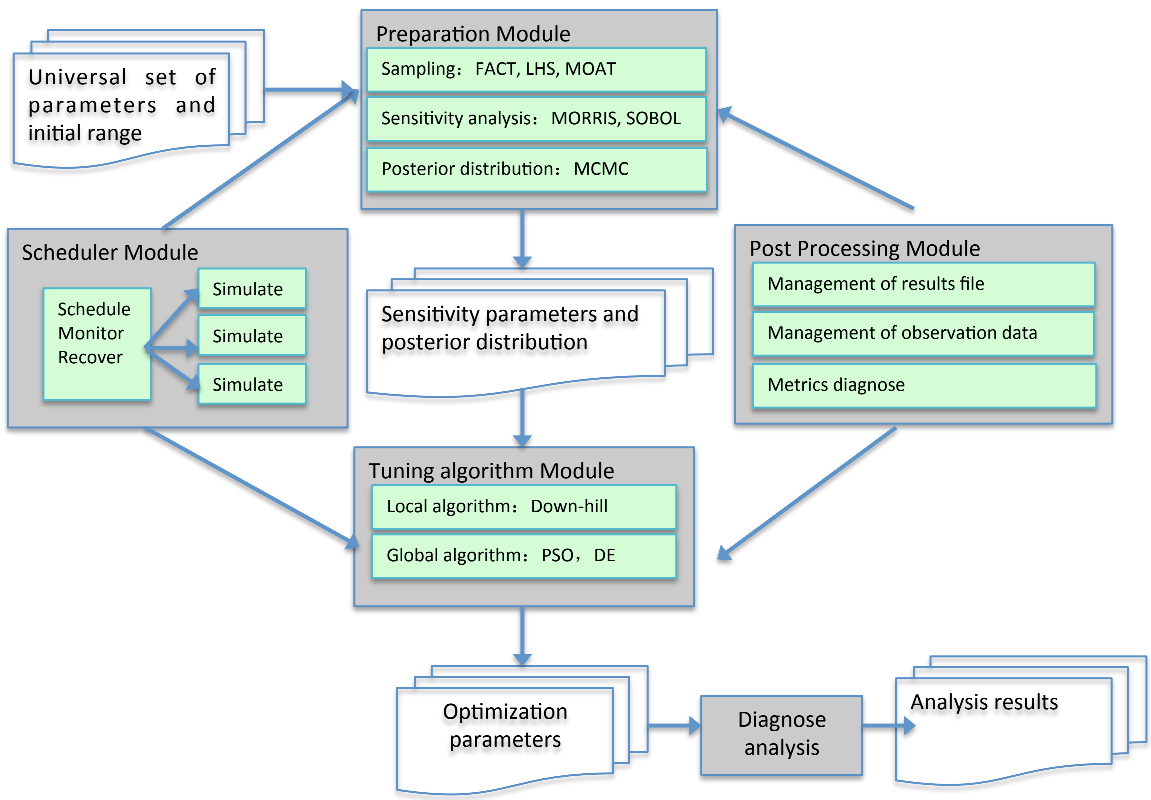

Figure 3. calibration workflow

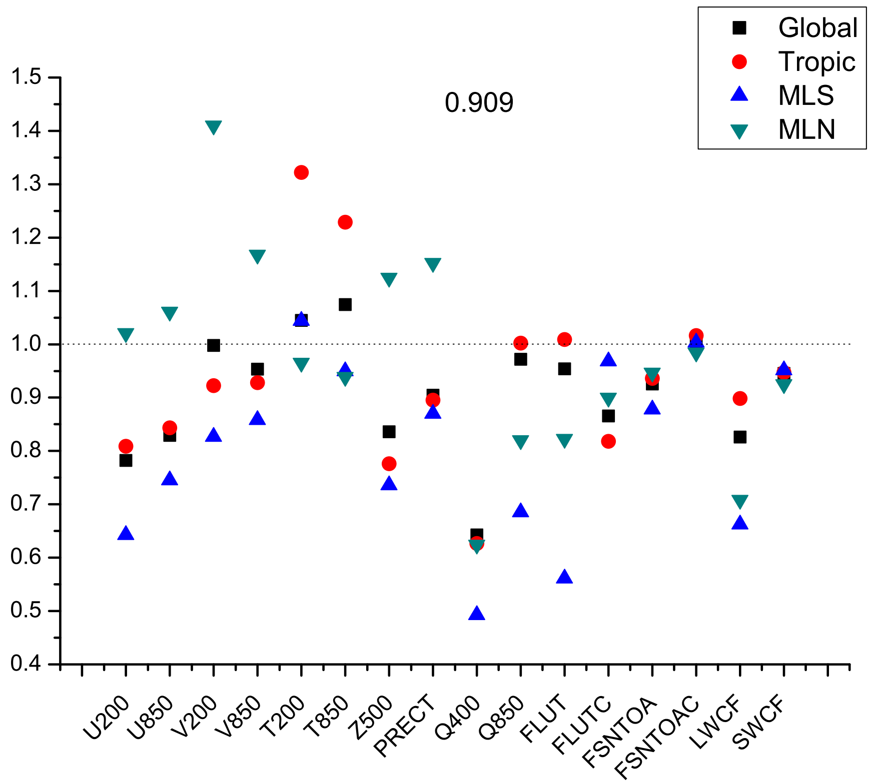

Figure 4. Metrics of each variables with the global, tropic, and northern / southern middle and high latitudes

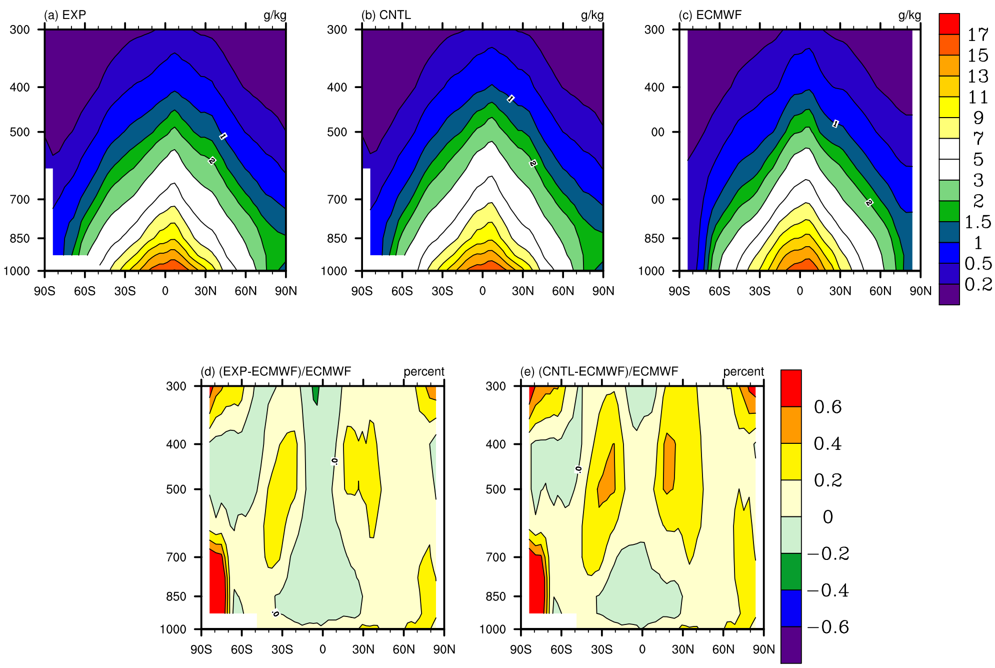

Figure 5. Pressure-latitude distributions of specific humidity of EXP (a), CNTL (b), observation (c), EXP-observation (d), and CNTL-observation (e).

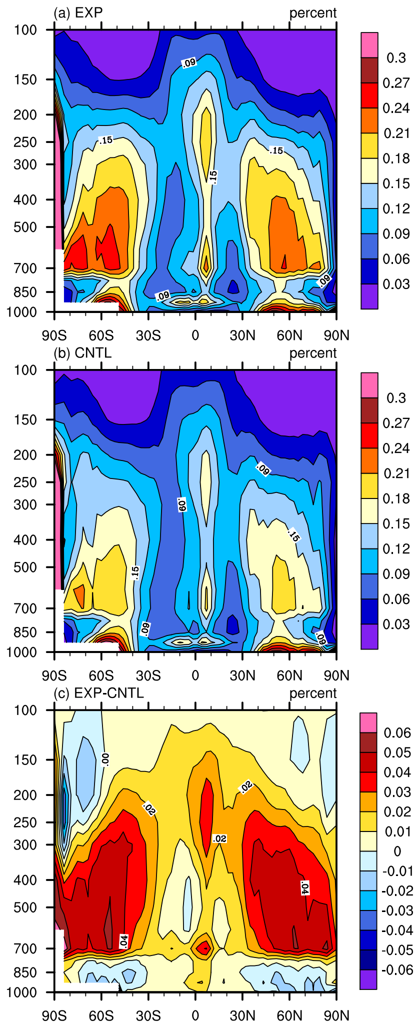

Figure 6. Pressure-latitude distributions of cloud fraction of EXP (a), CNTL (b), and EXP-CNTL (c).

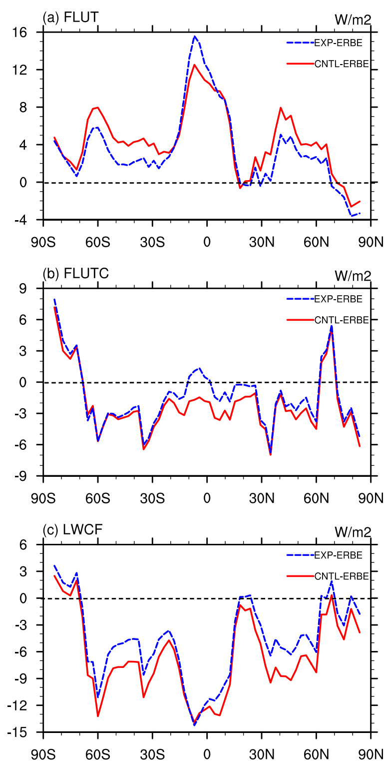

Figure 7. Meridional distributions of annual mean difference between EXP / CNTL and observation of FLUT (a), FLUTC (b), and LWCF (c).

Table 1. Initial selected uncertain parameters in GAMIL2

Parameter	Description	Default	Range
c0	rain water autoconversion coefficient for deep convection	3.0e-4	1.e-4 $\sim$ 5.4e-3
ke	evaporation efficiency for deep convection	7.5e-6	5e-7 $\sim$ 5e-5
capelmt	threshold value for cape for deep convection	80	20 $\sim$ 200
rhminl	threshold RH for low clouds	0.915	0.8 $\sim$ 0.95
rhminh	threshold RH for high clouds	0.78	0.6 $\sim$ 0.9
c0_shc	rain water autoconversion coefficient for shallow convection	5e-5	3e-5 $\sim$ 2e-4
cmftau	characteristic adjustment time scale of shallow cape	7200	900 $\sim$ 14400

Table 2. Model output variables in the metrics

Variable	Observation	Variable	Observation
Meridional wind at 850hPa	ECMWF	Geopotential Z at 500hPa	ECMWF
Meridional wind at 200hPa	ECMWF	Total precipitation rate	GPCP
Zonal wind at 850hPa	ECMWF	Longwave cloud forcing	ERBE
Zonal wind at 200hPa	ECMWF	Shortwave cloud forcing	ERBE
Temperature at 850hPa	ECMWF	Long wave upward flux at TOA	ERBE
Temperature at 200hPa	ECMWF	Clearsky long wave upward flux at TOA	ERBE
Specific Humidity at 850hPa	ECMWF	Short wave net flux at TOA	ERBE
Specific Humidity at 400hPa	ECMWF	Clearsky short wave net flux at TOA	ERBE

Table 3. Fall factor samplings of parameters and metrics

ID	c0	rhminl	rhminh	cmftau	metrics	ID	c0	rhminl	rhminh	cmftau	metrics
1	1.00E-04	0.915	0.78	7200	1.152	14	3.00E-04	0.875	0.78	7200	1.019
2	3.00E-04	0.915	0.78	7200	1	15	3.00E-04	0.913	0.78	7200	1.007
3	3.04E-04	0.915	0.78	7200	1.054	16	3.00E-04	0.95	0.78	7200	1.094
4	5.08E-04	0.915	0.78	7200	1.017	17	3.00E-04	0.915	0.6	7200	1.00547
5	7.13E-04	0.915	0.78	7200	0.987	18	3.00E-04	0.915	0.675	7200	1.027676
6	9.17E-04	0.915	0.78	7200	1.01	19	3.00E-04	0.915	0.75	7200	1.023358
7	1.12E-03	0.915	0.78	7200	1.04	20	3.00E-04	0.915	0.825	7200	1.028264
8	1.33E-03	0.915	0.78	7200	1.044	21	3.00E-04	0.915	0.9	7200	1.160479
9	2.55E-03	0.915	0.78	7200	1.075	22	3.00E-04	0.915	0.78	900	1.22922
10	3.78E-03	0.915	0.78	7200	1.084	23	3.00E-04	0.915	0.78	4275	1.064064
11	5.00E-03	0.915	0.78	7200	1.09	24	3.00E-04	0.915	0.78	7650	1.004806
12	3.00E-04	0.8	0.78	7200	1.223	25	3.00E-04	0.915	0.78	11025	1.077167
13	3.00E-04	0.838	0.78	7200	1.054	26	3.00E-04	0.915	0.78	14400	1.148265

Table 4. Comparison with effective and efficiency

	Optimal solution	N_{step}	N_{size}	Core hours
Down-hill simplex	0.9585	80	1	14400
PSO	0.911537	24	12	51840
DE	0.942148	33	12	71280
Dowmhill_2_steps	0.9256899	25+34	1	10620
Downhill_3_steps	0.9098545	80+25+50	1	27900

Algorithm 1 Preprocessing the initial values of Downhill Simplex Algorithm

```
sampling_sets=full_factor_sampling(parameters_range)
for each initial V_i of N+1 vertexes do
    candidate_init_sets += min(i, sampling_sets)
end for
while one parameter have the same values in the N+1 sets do
    j=1
    remove_parameter_set(the parameter set with higher metrics, candidate_init_sets)
    candidate_init_sets += min(N+1+j, sampling_sets)
    j+=1
end while
```