Authors' responses to RC's comments

Revision Notes

For

Manuscript ID: **egusphere-2026-2275**

Quantifying key parameter sensitivities for water table depth in hydrological schemes of CoLM-PSUADE at a mountain forest site

Dear Editor,

We thank the editor and the anonymous reviewers for their insightful comments on our manuscript. In response to the comments, we have carefully revised the manuscript to address the remaining minor points raised by Reviewer #3, which have helped us further improve the manuscript.

Key revisions include:

- Clarifying the temporal coverage of the data;
- Explaining the inconsistency between the default value and prescribed range of parameter $r_{sub,max}$;
- Clarifying the descriptions of the screening results;
- Clarifying the statements regarding the performance of the optimization methods;
- Conducting a comprehensive check of language and consistency;
- Standardizing parameter names, symbols, and terminology.

Below, we provide a point-by-point response. The reviewer's comments are **Relisted in black**, followed by our **Responses in blue** and **Revisions in purple**.

Reviewer #3:

The authors developed a CoLM2024–PSUADE framework to identify and optimize key parameters controlling water table depth at the Gongga Mountain forest site. The revised manuscript provides a clearer definition of the study scope, a more detailed

description of the sensitivity-analysis framework, and an improved discussion of parameter robustness and unresolved WTD dynamics. Below are some suggestions.

Response:

We thank the reviewer for his/her careful review. According to the comments, we have revised the manuscript, particularly by clarifying the temporal coverage of the datasets, explaining the distinction between the prescribed parameter search ranges and the original CoLM default values, refining the interpretation of sensitivity robustness across different random-seed realizations, explicitly acknowledging the limitations associated with the unequal computational budgets when comparing the optimization methods, and standardizing parameter names and terminology throughout the manuscript, tables, and figures.

All comments have been carefully considered and **Relisted in black**, followed by our **Responses in blue** and **Revisions in purple**. The point-by-point responses are provided below.

Specific comments:

1. The manuscript states that the optimized parameters were validated over 2007–2014, whereas Table 1 reports that the meteorological forcing and LAI data cover only 2005–2007. Please clarify the actual temporal coverage.

Response:

We apologize for the typo and thank the reviewer for bringing this out. The temporal coverage of meteorological forcing and LAI data has been corrected to 2005–2014 in Table 1.

Revision:

Table 1 (Page 6, Lines 149–150)

The revised parts are indicated in *italics*:

Table 1: The datasets used in this study.

Data Type	Source	Coverage	Resolution
Soil texture	National Ecosystem Science Data Center (https://www.nesdc.org.cn/)	Point; Aug 2005, Sep 2015	1 day
Atmosphere	National Tibetan Plateau Data Center Third Pole Environment Data Center (https://data.tpdc.ac.cn/)	Tibetan Plateau; *2005–2014*	1/30°; Hourly
LAI	Land-Atmosphere Interaction Research Group at Sun Yat-sen University (http://globalchange.bnu.edu.cn)	Global; *2005–2014*	15 "; Monthly
Soil hydraulic and thermal parameters	Land-Atmosphere Interaction Research Group at Sun Yat-sen University (http://globalchange.bnu.edu.cn)	Global	15 "

WTD	National Ecosystem Science Data Center (https://www.nesdc.org.cn/)	Point; 2005–2014	Daily[a]

[a]WTD observations were recorded at irregular intervals ranging from 1–9 days, with most measurements spaced 4–6 days apart. During model evaluation, simulated WTD values were matched with the corresponding observation dates.

2. The default value of rsubmax in Table 6 (5.5 ×10-3) exceeds its limit range in Table 2 ([1×10-5, 5×10-3]). Please explain why the default value falls outside the prescribed range.

Response:

We thank the reviewer for pointing out this inconsistency. We agree that the default value of rsubmax (5.5×10^{-3}) is slightly higher than the upper bound of the parameter range specified in Table 2 (5×10^{-3}). The range reported in Table 2 represents the prescribed parameter search range in the CoLM-PSUADE experiments.

For rsubmax, the upper bound was set to 5×10^{-3} based on the order of magnitude of the parameter in the CoLM implementation, while maintaining a relatively broad search space for parameter analysis. In the reference study, Niu et al. (2005) explored rsubmax (denoted $R_{sb,max}$ in their study) across values extending down to the order of 10^{-5} (Figure 6 in their study) and found that values of approximately 1×10^{-4} to 3×10^{-4} produced the highest model efficiencies. Therefore, when defining the candidate range for the present analysis, we extended the range substantially beyond the magnitude of the values examined in the reference SIMTOP study to cover the parameter space relevant to the CoLM implementation. The upper bound was then manually rounded to 5×10^{-3}, which resulted in the CoLM default value being slightly outside the prescribed search range.

Importantly, the CoLM default value was used as the baseline for comparison, whereas the sensitivity and optimization experiments were conducted within the prescribed range of 1×10^{-5} to 5×10^{-3}. The optimization results further showed that the rsubmax values associated with the low-RMSE solutions were mainly concentrated at the order of 10^{-4} (Fig. 18), which is consistent in magnitude with relatively high model efficiency reported by Niu et al. (2005). To avoid ambiguity, we have clarified the distinction between the prescribed PSUADE search ranges and the original CoLM default values in the text before Table 2 and the note to Table 6.

Revision:

(1) Sect. 2.2.2 of the Data and methods (Page 9, Lines 200–203)

The revised descriptions are indicated in *italics*:

When coupled with PSUADE, the execution of CoLM is initiated by the simulator specified within PSUADE, which is implemented as a Python script in this study (Wu

et al., 2026). Prior to execution, the script replaces the corresponding parameters in the CoLM source files using the sampled parameter values, after which the model is compiled and executed. The selection of adjustable model parameters is primarily guided by their relevance to WTD simulations, with careful consideration of the strong linkages between the simulation process and key hydrological components such as infiltration, evaporation, runoff, freeze-thaw dynamics, and the vertical movement of soil moisture. Accordingly, a total of 56 key parameters associated with these processes were identified. Their physical meanings, categories, value ranges, and the references for determining these ranges are summarized in Table 2. The parameter categories were determined based on the physical processes, with some classifications adapted from Li et al. (2013). *The parameter ranges were determined based on the default parameter values implemented in CoLM, the values reported in previous studies, and the need to provide a broad parameter space for analysis. Therefore, the ranges in Table 2 were interpreted as the parameter space explored in this study rather than as universal physically valid limits.*

(2) Sect. 3.3 the Results (Page 34, Lines 568–572)

The revised descriptions are indicated in *italics*:

Table 6: Comparison of the default and optimized values for the eight stably sensitive parameters.

Parameter	Default value	Optimized value		
		SCE	PSO	BOBYQA
alpha_rain	2.500×10^{-1}	3.232×10^{-1}	7.170×10^{-1}	3.310×10^{-1}
rhol_nir	3.500×10^{-1}	4.626×10^{-1}	5.245×10^{-1}	4.116×10^{-1}
rsubmax[a]	**5.500×10^{-3}**	**1.020×10^{-4}**	**4.286×10^{-4}**	**2.451×10^{-4}**
krmax	**3.981×10^{-9}**	**6.013×10^{-8}**	**4.185×10^{-8}**	**7.008×10^{-8}**
ck0	3.950	3.701	2.724	2.602
beta	-1.623	-2.383	-2.864	-1.028
soil_n[b]	1.255	8.994	9.609	5.556
soil_alpha[b]	**2.245×10^{-2}**	**1.475×10^{-3}**	**2.408×10^{-3}**	**1.847×10^{-3}**

[a]Default values represent the original parameter settings implemented in CoLM and are provided as the baseline for comparison. The optimized values were obtained by searching within the parameter ranges specified in Table 2. For rsubmax, the CoLM default value (5.500×10^{-3}) slightly exceeds the prescribed upper bound (5.000×10^{-3}), because the latter was manually rounded; [b]The original values of the two parameters are different across the 10-layer soil profile. Here, they are represented as the average values.

3. Table 7 indicates that 40-rsubmax was not selected as a sensitive parameter in any of the four summarized screening results for the two additional random-seed realizations. However, Abstract (Line 18-20) indicated that all eight parameters are robustly sensitive. Please clarify the inconsistency.

Response:

We thank the reviewer for this suggestion. We agree that the original statement in the Abstract that all eight parameters were "robustly sensitive" was inaccurate. We have revised the manuscript to distinguish between parameters identified in the initial screening and their robustness across different random-seed realizations. The corresponding descriptions have been mainly revised in the Abstract, Results, and Conclusions.

Revision:

(1) Abstract (Page 1, Lines 18, 20–22)

The revised descriptions are indicated in *italics*:

Accurately representing groundwater dynamics in land surface models (LSMs) is crucial for understanding water-energy cycles and assessing water resources. However, most LSMs lack systematic evaluations of parameter sensitivities governing water table depth (WTD) simulations. This study coupled the Common Land Model (CoLM) with Problem Solving environment for Uncertainty Analysis and Design Exploration (PSUADE) in a single-point framework to facilitate systematic parameter analysis and optimization for improving WTD simulation. The CoLM-PSUADE framework was then applied to evaluate groundwater-related parameters using WTD observations from the Gongga Mountain site. A comprehensive analysis integrating qualitative screening, quantitative validation, and parameter optimization techniques was conducted to evaluate the sensitivities of 56 parameters associated with key hydrological processes and identify their optimal ranges. *The results identified eight sensitive parameters in the initial screening*, including those controlling unsaturated soil water movement (56-soil_alpha, 53-soil_n), subsurface runoff (40-rsubmax), plant hydraulic processes (49-beta, 45-krmax, 46-ck0), and net surface water infiltration (4-alpha_rain, 10-rhol_nir). *Additional experiments with independent random seeds revealed varying sensitivity rankings. In particular, the subsurface runoff parameter $r_{sub,max}$ exhibited sampling-dependent sensitivity but showed relatively high importance in the analysis.* It exhibited a well-defined optimal range (on the order of 10^{-4}) and regulated both the magnitude of subsurface runoff and its decay with increasing WTD when combined with another empirical parameter in the SIMTOP (Simple TOPMODEL-based) scheme, f_{drai}, thereby exerting strong control on WTD. The soil hydraulic parameter α showed the highest sensitivity. It regulated unsaturated hydraulic conductivity and soil water retention, thereby exerting a dominant influence on the variability and lagged response of WTD. Based on these findings, a stepwise calibration strategy was recommended, in which the subsurface runoff parameters ($r_{sub,max}$ and f_{drai}) are first adjusted to constrain the mean WTD, followed by optimization of other key parameters, such as α, to improve the temporal dynamics of WTD. This study demonstrated that CoLM-PSUADE provides a useful tool for sensitivity-guided parameter optimization in high-dimensional LSMs and hydrological models.

(2) Sect. 3.1 of the Results (Page 23, Line 400)

The revised descriptions are indicated in *italics*:

Based on the comprehensive ranking results from the two aforementioned perspectives, eight parameters *in this screening* were identified as stable highly sensitive parameters: 56-soil_alpha, 45-krmax, 4-alpha_rain, 53-soil_n, 49-beta, 10-rhol_nir, 46-ck0, and 40-rsubmax. The remaining highly sensitive parameters were classified as unstable. This process effectively identified key parameters that significantly influence WTD from the total set of 56 parameters. Given that the number of stably sensitive parameters was notably small relative to the total parameter count, the screening procedure demonstrated good applicability for this study.

(3) Sect. 4.1 of the Discussion (Page 37, Line 602)

The revised descriptions are indicated in *italics*:

The eight stably sensitive parameters *identified from the previous screening* including 56-soil_alpha, 53-soil_n, 49-beta, 40-rsubmax, 45-krmax, 4-alpha_rain, 46-ck0 and 10-rhol_ for WTD simulations were comprehensively identified through both qualitative and quantitative sensitivity analyses. The key processes and calculation formulas related to the parameters are detailed in Sect. S5 of the supplementary material.

(4) Conclusions (Page 47, Lines 786–798)

The initial qualitative sensitivity analyses identified eight stably sensitive parameters (56-soil_alpha, 45-krmax, 4-alpha_rain, 53-soil_n, 49-beta, 10-rhol_nir, 46-ck0, and 40-rsubmax) governing soil water movement, subsurface runoff, plant hydraulics, and surface infiltration processes. Subsequent RSMSobol-based quantitative analysis provided a complementary assessment of these eight parameters and identified a core group (56-soil_alpha, 53-soil_n, 49-beta, 45-krmax, 40-rsubmax, and 4-alpha_rain). Building on this reduced parameter space, parameter optimization using SCE, PSO, and BOBYQA demonstrated that targeted calibration can substantially improve WTD simulation performance, with SCE identifying the lowest-RMSE solution most rapidly. To further assess the robustness of the initial qualitative sensitivity results, two additional experiments using independent random seeds were conducted. These experiments demonstrated that qualitative sensitivity rankings were generally consistent for the dominant parameters but exhibited variability among different random-seed realizations. In particular, 40-rsubmax was identified in the initial screening but was not selected as sensitive in the summarized screening results of either additional random-seed realization. In contrast, 56-soil_alpha maintained its top sensitivity ranking across all experiments, confirming its robust and dominant influence on WTD simulations.

4. The optimization budgets differ among the three methods, with 1000 evaluations for SCE and BOBYQA but 496 valid evaluations for PSO. However, the authors should acknowledge more explicitly that the comparison was conducted under unequal computational budgets. Statements such as “SCE outperformed PSO and BOBYQA” or “SCE is more suitable for CoLM” should preferably be limited to the present

experimental configuration.

Response:

We thank the reviewer for this suggestion. We agree that the unequal computational budgets should be explicitly acknowledged when comparing the three optimization methods. We have revised the Results to limit the comparison of convergence speed and solution quality to the experimental configuration used in this study. We have also expanded the Discussion to acknowledge the limitations of comparing methods under unequal computational budgets and to highlight the need for future matched-budget experiments.

Revision:

(1) Sect. 3.3 of the Results (Page 29, Lines 516–520)

The revised descriptions are indicated in *italics*:

A comparison of the SCE, PSO, and BOBYQA methods revealed clear differences in their convergence characteristics and optimization performance (Fig. 15). In terms of overall behaviour (Fig. 15a), most $RMSE$ values obtained by the SCE method were concentrated between 0.4 and 1 m, with occasional sharp peaks corresponding to failed solutions. The low-$RMSE$ range was stable across evaluations. For the PSO method, $RMSE$ values were primarily clustered between 0.6 and 0.9 m and exhibited evident high-frequency oscillations across evaluations when multiple particles were simultaneously searched. The BOBYQA method showed a markedly different changing pattern. Specifically, during the first 90–140 evaluations, $RMSE$ values fell into the relatively poor range because the change of trust region. However, after approximately 200 evaluations, $RMSE$ remained stable within the range of 0.5–1.0 m. Regarding optimization performance (Fig. 15b), SCE found a better solution during the early stage. As the number of evaluations increased, population of parameter sets evolved in the solution space were progressively constrained to good-quality zones. Ultimately, the SCE method achieved the lowest $RMSE$ value of 0.419 m, demonstrating the best overall performance. Although the $RMSE$ obtained by PSO exhibited considerable variability, its fluctuation range remained within a reasonable interval (0.5–1.2 m), and it was able to identify an optimal solution (0.531 m) within the early stage of the optimization process (around the first 50 evaluations), suggesting relatively fast initial exploration capability. The minimum $RMSE$ obtained by BOBYQA was 0.517 m, lying between the results of the other two methods. *The cumulative minimum RMSE curves exhibited relatively stable over the respective evaluation ranges of the three methods, although the comparison was conducted under unequal computational budgets (1000 evaluations for SCE and BOBYQA and 496 valid evaluations for PSO).* Overall, SCE outperformed PSO and BOBYQA in terms of *the speed of identifying the best solution and the minimum RMSE achieved*, indicating that it is more suitable for parameter optimization in the CoLM *under present experimental configuration*.

(2) Sect. 4.4 the Discussion (Page 46, Lines 762–764)

The revised descriptions are indicated in *italics*:

Considering the cost of computation, a fixed optimization cost (about 500–1000 runs) was adopted for the evaluation. Such comparisons of fixed costs have also been conducted in previous studies (Kavetski et al., 2018), and they are divided into two types of computational budgets. The evaluations ranging from 200 to 500 times are classified as the "tight" type, while those ranging from 1000 to 5000 times are classified as the "large" type (Qin et al., 2022). The convergence behaviours of different methods are different. Some methods can still find satisfactory solutions even when the budget is limited. However, for some methods, even in the case of a large budget, there is still a significant variability in the convergence behaviour. This might be related to the complexity of the model, especially when dealing with highly parameterized models, and the speed of convergence also depends on various factors such as the sampling strategy of the algorithm (Qin et al., 2022). In addition, the optimization process was also controlled by the convergence criteria. The detailed stopping criteria of PSO, SCE, and BOBYQA implemented in PSUADE are provided in Sect. S7 of the supplementary material. The primary purpose of the optimization experiments in this study was not to benchmark the absolute global optimization capability of different algorithms, but rather to evaluate whether the sensitive parameters identified from the sensitivity-analysis framework could effectively improve WTD simulations. Therefore, a moderate computational budget in this study was considered appropriate for assessing the effectiveness of the selected parameters. Nevertheless, the use of fixed-cost evaluation for parameter optimization still has limitations. Specifically, performance metrics can be subject to sampling error under the low number of algorithm invocations (Qin et al., 2022). In addition, the relative performance may vary under larger evaluation budgets (Zhang et al., 2009). *Thus, the comparison of the three optimization methods should be interpreted in the context of their practical but unequal computational budgets used in this study. Further experiments with matched evaluation budgets would be useful for a more rigorous comparison of their performance.*

5. A final language and consistency check would be helpful. In particular, parameter names such as soil_n/soil_n_vgm, soil_alpha/soil_alpha_vgm, and rhol_nir should be standardized throughout the text, tables, and figures.

Response:

We thank the reviewer for this suggestion. We have revised the wording in Table 2 to improve the consistency of parameter names and checked the terminology against the corresponding descriptions in the manuscript. In addition, we conducted a final language and consistency check throughout the manuscript.