Supplement of Hydrol. Earth Syst. Sci., 30, 6867–6905, 2026
https://doi.org/10.5194/hess-30-6867-2026-supplement
© Author(s) 2026. CC BY 4.0 License.



Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

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

Tingting Wu et al.

Correspondence to: Xiaofan Yang (xfyang@bnu.edu.cn)

The copyright of individual parts of the supplement might differ from the article licence.

S1: Representativeness of selected water years

The period 2005–2007 was used for sensitivity analysis and parameter optimization. Specifically, the 2005–2006 water year was treated as the spin-up period to allow the model states to approach dynamic equilibrium before analysis. The climatic conditions during this year were close to the long-term mean at the GGF site, with an annual mean air temperature difference of approximately 0.09 °C and an annual precipitation difference of approximately −64 mm relative to the long-term average. The 2006–2007 water year was selected for analysis because it represented a hydrological period with relatively shallow observed WTD and enhanced groundwater dynamics. Compared with the long-term mean conditions, the annual mean air temperature and precipitation during the 2006–2007 water year differed by approximately −0.20 °C and +84 mm, respectively. Such conditions provided sufficient variability in hydrological processes, including infiltration, runoff generation, groundwater recharge, and drainage, facilitating the diagnosis of parameters controlling WTD simulations. The subsequent validation period (2007–2014) included diverse hydroclimatic conditions, including relatively wet and dry years and different groundwater states, allowing an independent evaluation of the identified parameters under varying environmental conditions.

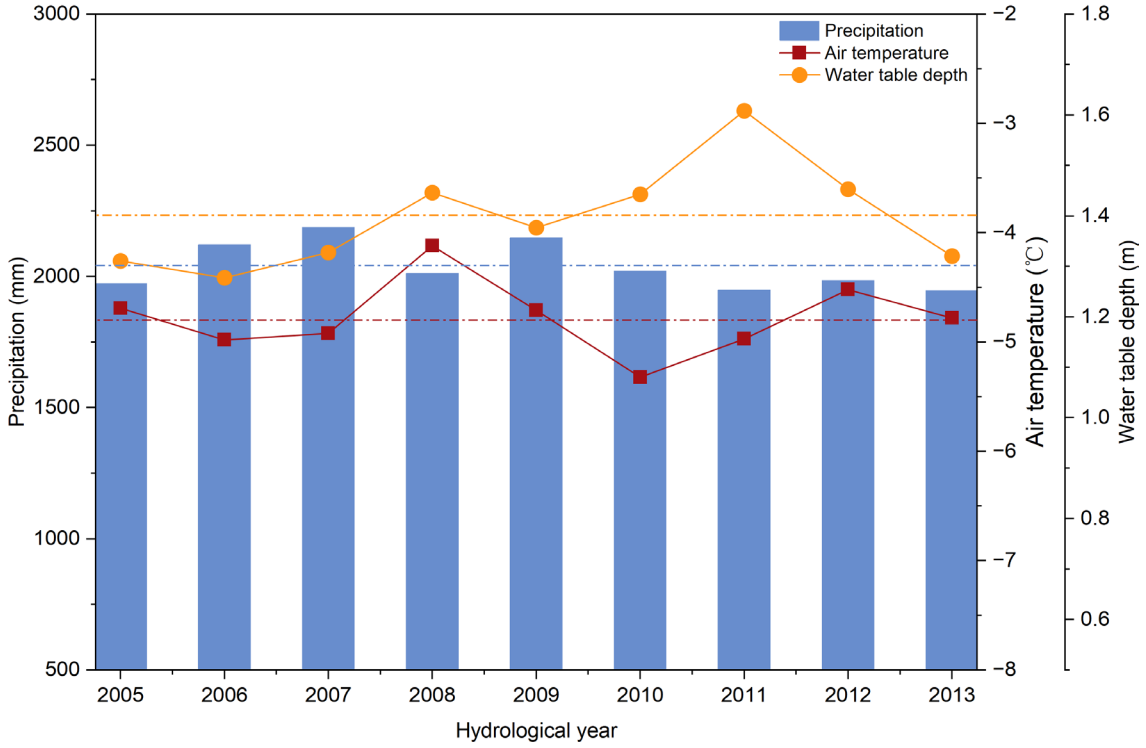


Figure S1: The total annual precipitation, annual average air temperature and annual average water table depth in the water years from 2005 to 2013. The dotted lines in the figure correspond to the average values over these years.

S2: Results of qualitative sensitivity analysis

S2.1 Qualitative screening results

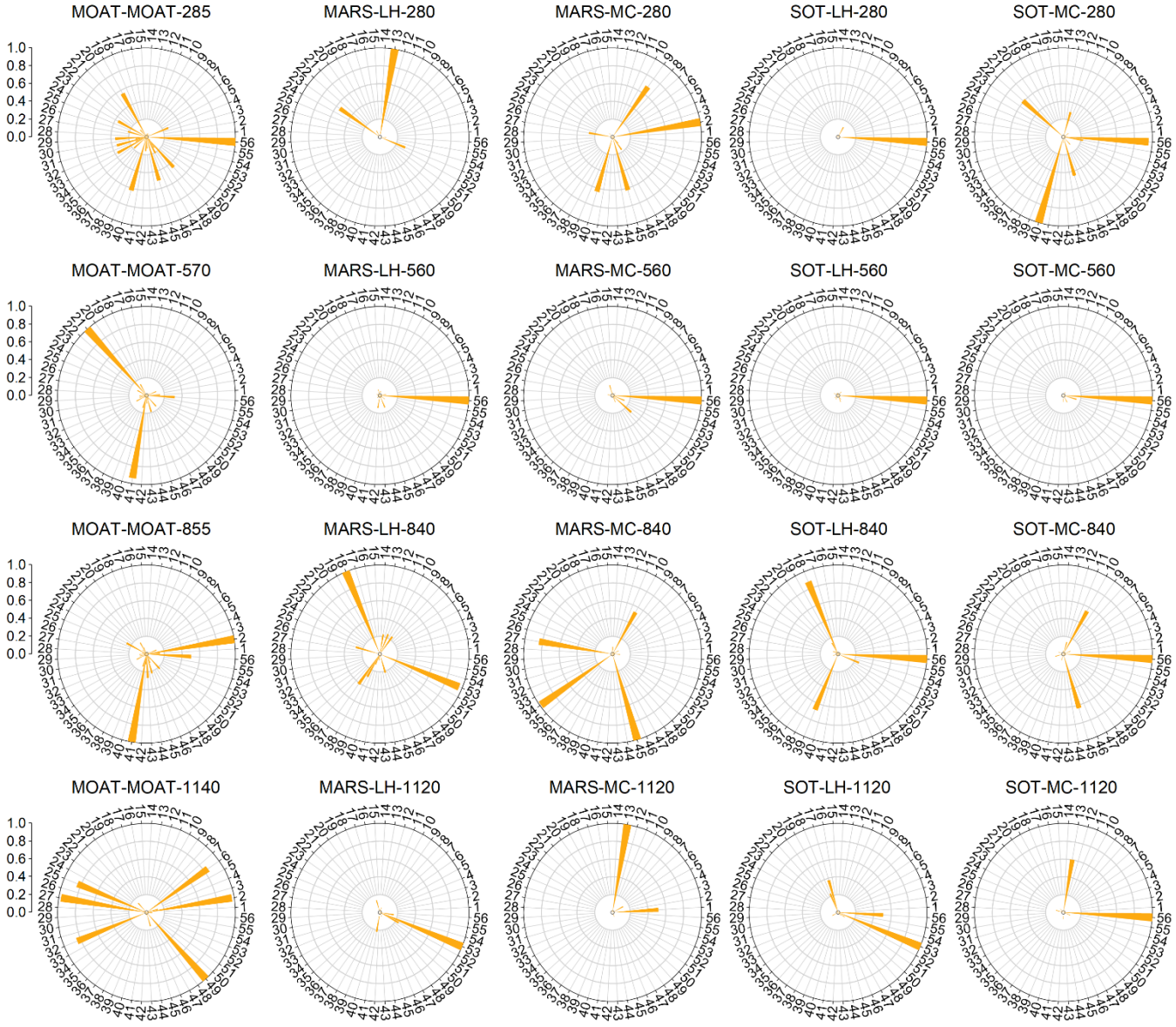


Figure S2: Qualitative sensitivity analysis results based on the *MAE* objective function. The numbers on the outer ring represent parameter indices (refer to Table 2 for the corresponding parameter definitions). The length of each pointer indicates the normalized sensitivity index, ranging from 0 to 1, with longer pointers denoting higher sensitivity. Each subplot title follows the format "sensitivity analysis method-sampling technique-sample size".

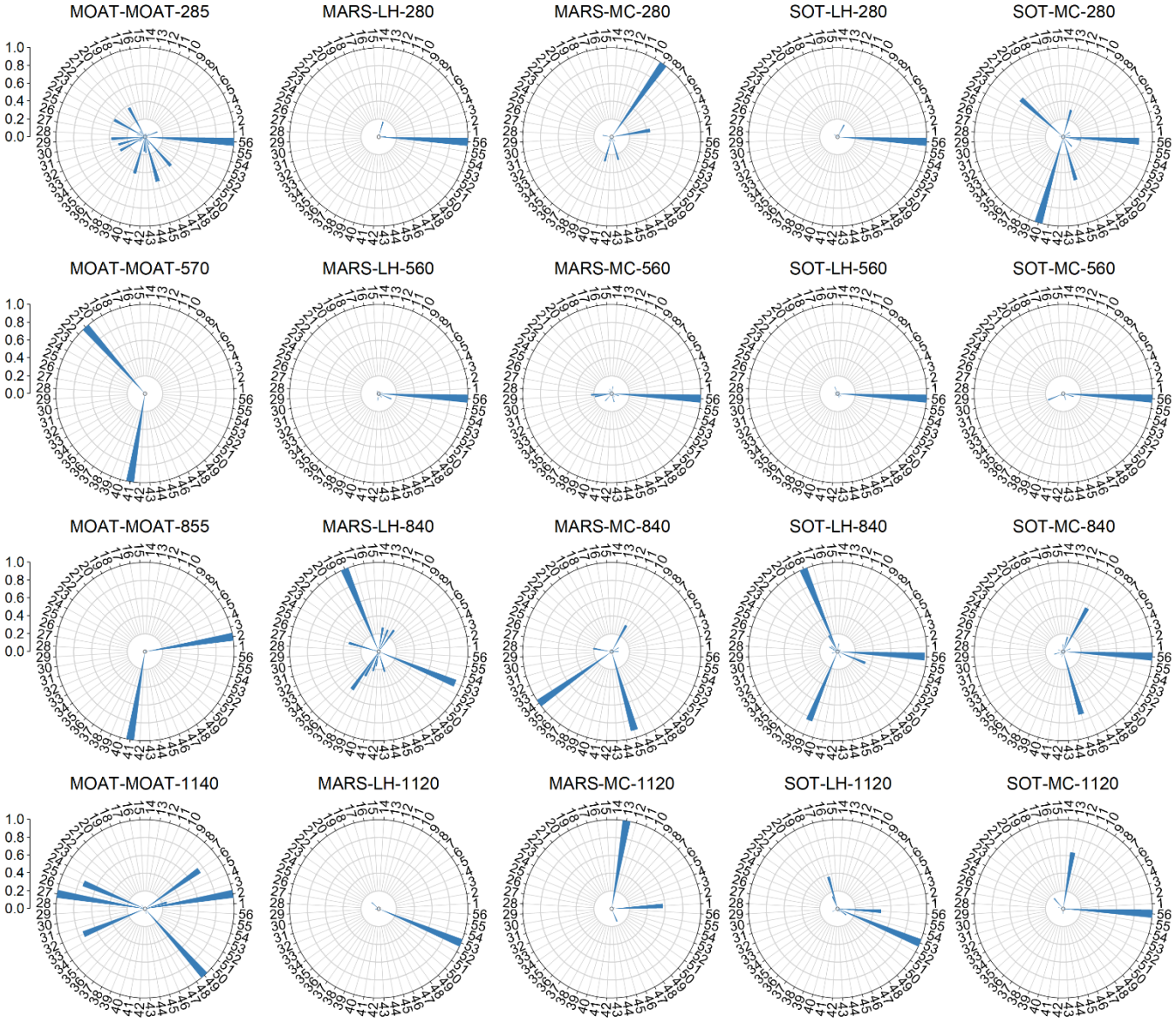


Figure S3: Qualitative sensitivity analysis results based on the *NSE* objective function. The numbers on the outer ring represent parameter indices (refer to Table 2 for the corresponding parameter definitions). The length of each pointer indicates the normalized sensitivity index, ranging from 0 to 1, with longer pointers denoting higher sensitivity. Each subplot title follows the format "sensitivity analysis method-sampling technique-sample size".

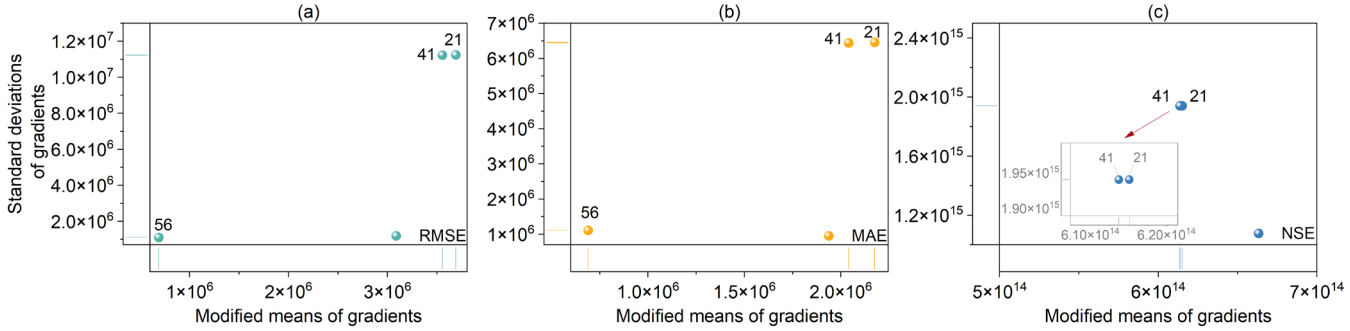


Figure S4. Sensitivity and interaction results from the MOAT method with a sample size of 570. The horizontal axis denotes the modified mean (μ_i^*), and the vertical axis denotes the standard deviation (σ_i^*). Subplots (a)–(c) correspond to the *RMSE*, *MAE*, and *NSE* objective functions, respectively.

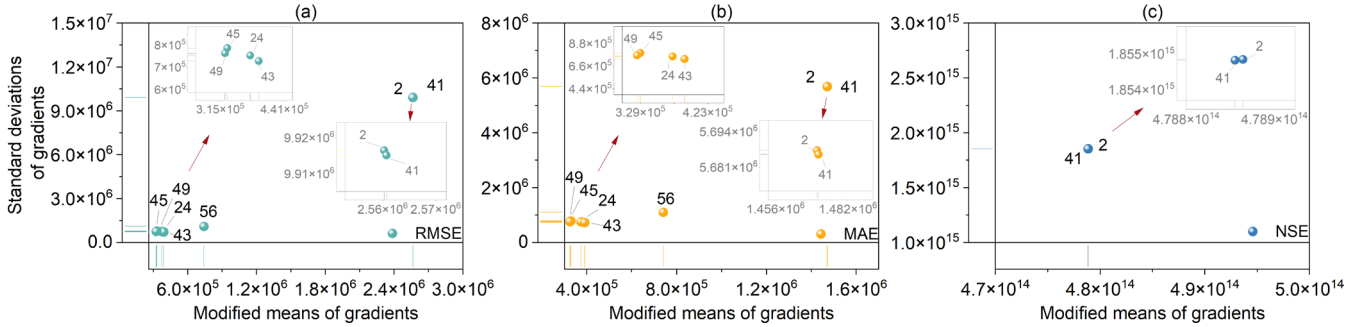


Figure S5. Sensitivity and interaction results from the MOAT method with a sample size of 855. The horizontal axis denotes the modified mean (μ_i^*), and the vertical axis denotes the standard deviation (σ_i^*). Subplots (a)–(c) correspond to the *RMSE*, *MAE*, and *NSE* objective functions, respectively.

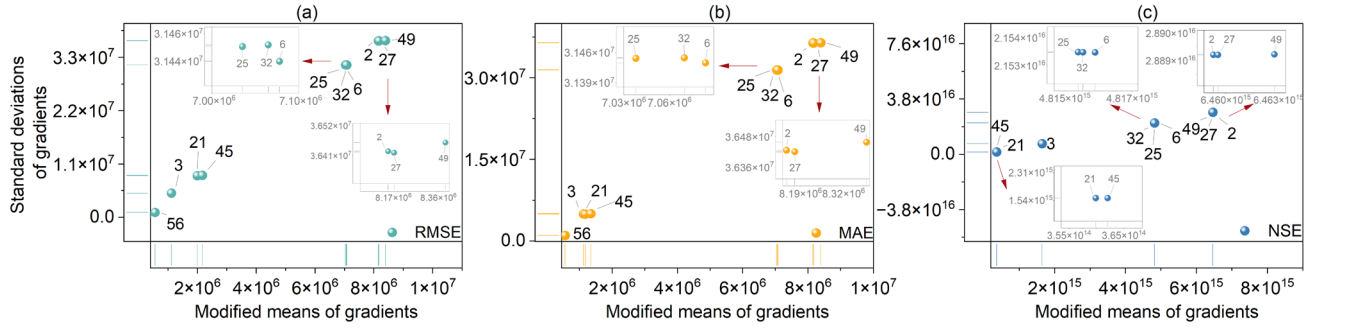


Figure S6. Sensitivity and interaction results from the MOAT method with a sample size of 1140. The horizontal axis denotes the modified mean (μ_i^*), and the vertical axis denotes the standard deviation (σ_i^*). Subplots (a)–(c) correspond to the *RMSE*, *MAE*, and *NSE* objective functions, respectively.

S2.2 Sample quality check

To quantify the space-fillingness of a sample, PSUADE provides the command named 'sqc', which computes a few metrics such as the minimum distance between all pairs of sample points (the larger the better), the average minimum distance for all sample points (the larger the better) and the minimum distance between the sample points and the corners (the smaller the

better). However, the average minimum distances between all points and corners are not calculated when the number of input parameters is greater than 30. For the qualitative sensitivity experiment, since the input dimension is 56, only the first two distance indicators were calculated based on the normalized sampling results (Table S1).

Table S1: Sample quality check for qualitative sensitivity analysis.

Method	Sampling technique	Avg-min distance	Min-min distance
MARS-5times	MC	3.041	2.063
MARS-10times		3.043	1.890
MARS-15times		3.043	1.954
MARS-20times		3.040	1.844
SOT-5times	LH	3.051	2.051
SOT-10times		3.048	1.981
SOT-15times		3.047	2.011
SOT-20times		3.047	1.936
MOAT-5times	MOAT	3.932	0.667
MOAT-10times		4.005	0.667
MOAT-15times		4.073	0.667
MOAT-20times		4.062	0.667

The sampling quality metrics exhibit clear differences among the three sampling strategies. For the MARS and SOT methods, the average minimum distance (Avg-min distance) remains nearly constant at approximately 3.04–3.05 as the sample size increases, indicating that both approaches maintain a stable level of global space-filling capability. In contrast, their minimum pairwise distance (Min-min distance) shows a slight decreasing trend with increasing sample size, reflecting the expected densification of sample points in high-dimensional space as more samples are added. In contrast, the Avg-min distance for MOAT is significantly larger (ranging from 3.93 to 4.07) than those of MARS and SOT, indicating that sample points are more widely separated on average. Meanwhile, its Min-min distance remains constant at 0.667 regardless of sample size, reflecting the deterministic and grid-based nature of its trajectory design. This behaviour underscores that MOAT prioritizes structured one-at-a-time perturbations rather than achieving uniform coverage of the multidimensional input space.

Overall, the results demonstrate that MARS and SOT achieve a better balance between global coverage and local separation, making them more suitable for space-filling experimental design, while MOAT is inherently limited by its trajectory-based structure.

S3: Results of quantitative validation of key sensitivity parameters

According to the response surface validation results across different experiments (Table S2), the error metrics Err_{rmss} and Err_{maxs} showed relatively high values. However, they were still 1–2 orders of magnitude smaller than the overall variation range of the sample outputs (Table S3). Further evaluation using Err_{maxs2} was therefore necessary. After min-max normalization, most Err_{maxs2} values were less than 1, and most of the average errors Err_{avgs} did not deviate significantly from 1, indicating that the accuracy of the response surfaces was acceptable. Furthermore, the coefficient of determination R^2_{rsm} from cross-validation was above 0 (0.122–0.804) when a GP was used as the response surface model. After applying a logarithmic transformation to the outputs (logGP), R^2_{rsm} increased to above 0.7 (0.726–0.837). For the parameter sets with input dimensions of 12 and 21, the remaining errors decreased by orders of magnitude. These results indicated that the response surfaces trained on the loaded samples (constructed from input parameters and $RMSE$ outputs) were of good quality, and also suggested that the output range influenced the fitting accuracy. In addition, the data points in the corresponding diagnostic plots were closely distributed along the diagonal (Fig. S7–Fig. S9), further supporting this conclusion. Based on the above evaluation metrics, both GP and logGP response surfaces can be considered reliable and suitable for subsequent quantitative sensitivity analysis.

Table S2: Response surface cross-validation metrics.

Experiment[a]	Response surface	Err_{avgs}	Err_{rmss}	Err_{maxs}	Err_{maxsb}	Err_{maxs2}	R^2_{rsm}
8-50times	GP	1.623	16.820	174.700	0.610	0.711	0.372
	logGP	1.669	29.360	633.700	0.005	0.687	0.786
8-100times	GP	6.079	64.540	1531.000	0.478	0.678	0.203
	logGP	11.310	236.400	7401.000	6.712×10^{-5}	0.687	0.726
8-150times	GP	5.542	73.020	937.600	1.302	0.900	-
	logGP	10.950	304.400	1.173×10^{4}	6.712×10^{-5}	0.625	0.792
12-40times	GP	1. 284	41.610	435.800	1.100	0.902	0.122
	logGP	0.449	5.700	78.810	0.036	0.734	0.737
12-80times	GP	1.713	37.850	568.900	1.100	0.705	0.548
	logGP	0.412	6.750	103.700	0.040	0.748	0.758
12-100times	GP	1.466	44.420	616.900	0.659	0.483	0.804
	logGP	0.539	7.261	159.900	0.017	0.620	0.818
21-20times	GP	299.200	4966.000	1.043×10^{5}	0.637	0.894	0.658
	logGP	1.382	9.868	118.800	0.008	0.670	0.832
21-40times	GP	2.655×10^{4}	4.782×10^{5}	7.188×10^{6}	0.365	1.000	-

	logGP	0.618	7.354	129.800	0.008	0.962	0.746
21-100times	GP	1.672×10^{4}	3.869×10^{6}	9.006×10^{7}	0.514	1.201	-
	logGP	2.328	53.590	2166.000	7.869×10^{-5}	0.921	0.837

[a]"$\boldsymbol{n}$-$\boldsymbol{k}$ times" in the column of "Experiment" denotes an experimental configuration with $\boldsymbol{n}$ input parameters and a sample size equal to $\boldsymbol{k}$ times the number of parameters. In the software implementation, the actual number of model runs (sample sizes) is $\boldsymbol{k} \times (\boldsymbol{n} + \boldsymbol{2})$.

Table S3: Range of RMSE outputs in the RSMSobol analysis experiments.

Experiment[a]	Response surface	Sample $\text{Output}_{\text{min}}$	Sample $\text{Output}_{\text{max}}$
8-50 times	GP	4.783662×10^{-1}	1.192372×10^{3}
	logGP	-7.373787×10^{-1}	7.083700
8-100 times	GP	4.783662×10^{-1}	1.467529×10^{3}
	logGP	-7.373787×10^{-1}	7.291335
12-40 times	GP	3.924328×10^{-1}	1.625459×10^{3}
	logGP	-9.353899×10^{-1}	7.393545
12-80 times	GP	3.827627×10^{-1}	2.079594×10^{3}
	logGP	-9.603402×10^{-1}	7.639928
12-100 times	GP	3.827627×10^{-1}	3.835408×10^{3}
	logGP	-9.603402×10^{-1}	8.252031
21-20 times	GP	3.812231×10^{-1}	1.847612×10^{5}
	logGP	-9.643704×10^{-1}	1.212682×10^{1}
21-40 times	GP	3.190905×10^{-1}	3.820296×10^{7}
	logGP	-1.142281	1.745842×10^{1}
21-100 times	GP	3.190905×10^{-1}	3.856754×10^{7}
	logGP	-1.142281	1.746792×10^{1}

[a]"$\boldsymbol{n}$-$\boldsymbol{k}$ times" in the column of "Experiment" denotes an experimental configuration with $\boldsymbol{n}$ input parameters and a sample size equal to $\boldsymbol{k}$ times the number of parameters. In the software implementation, the actual number of model runs (sample sizes) is $\boldsymbol{k} \times (\boldsymbol{n} + \boldsymbol{2})$.

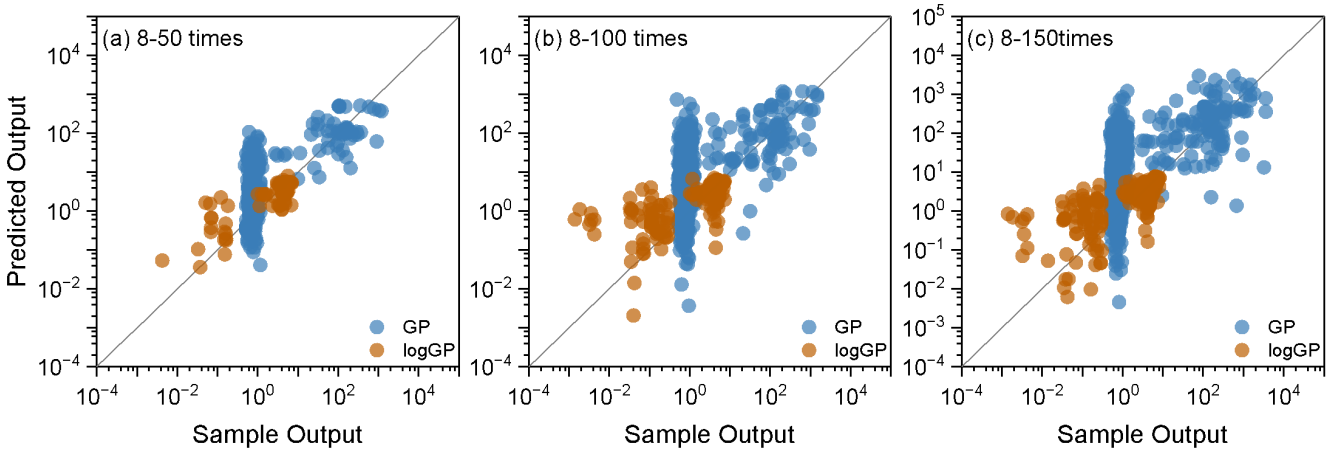


Figure S7: Response surface cross-validation scatter plot for the case where the input dimension is 8.

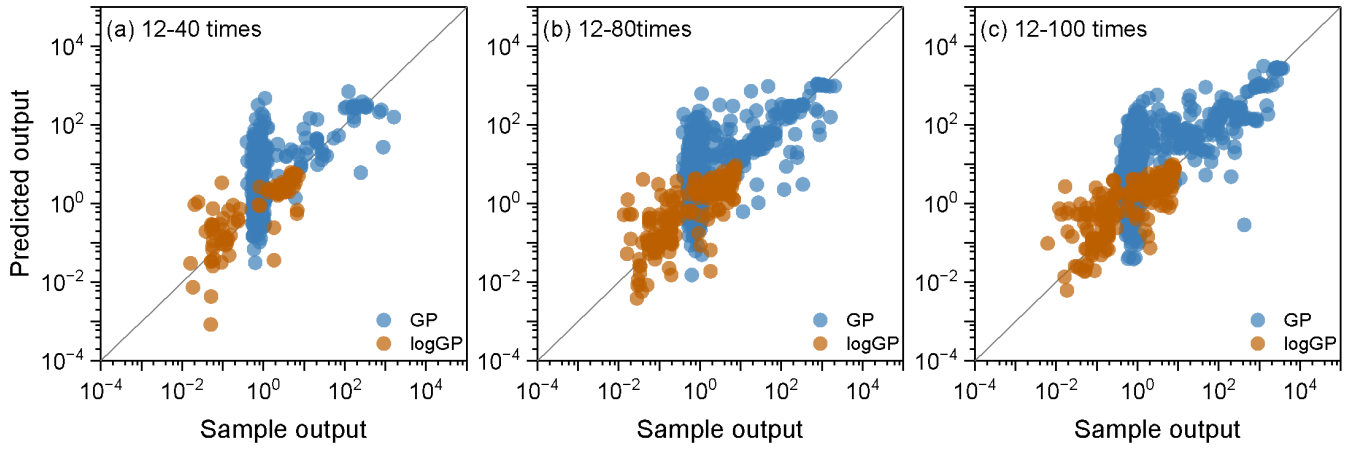


Figure S8: Response surface cross-validation scatter plot for the case where the input dimension is 12.

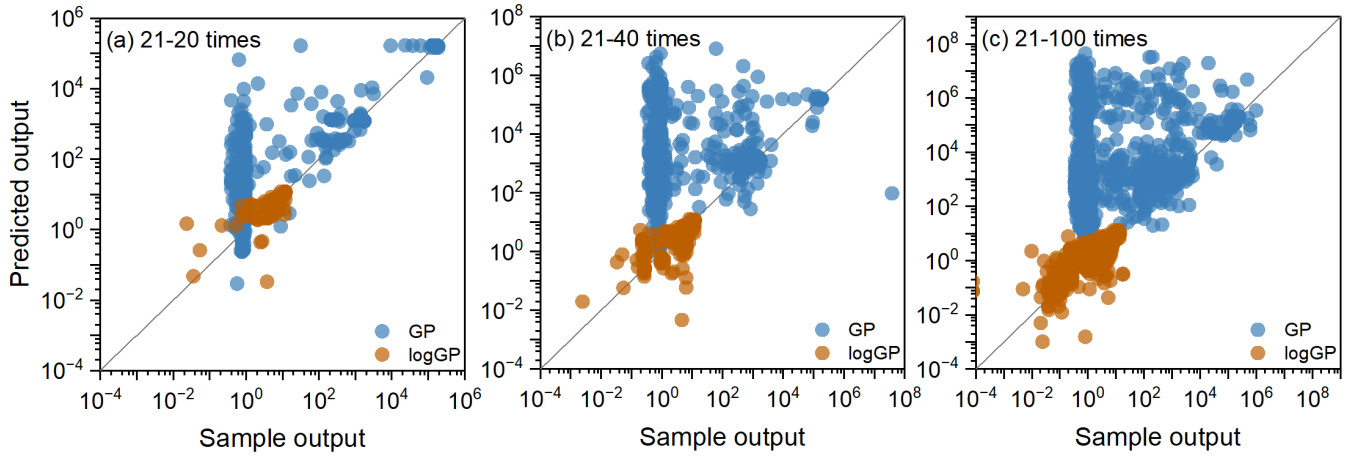


Figure S9: Response surface cross-validation scatter plot for the case where the input dimension is 21.

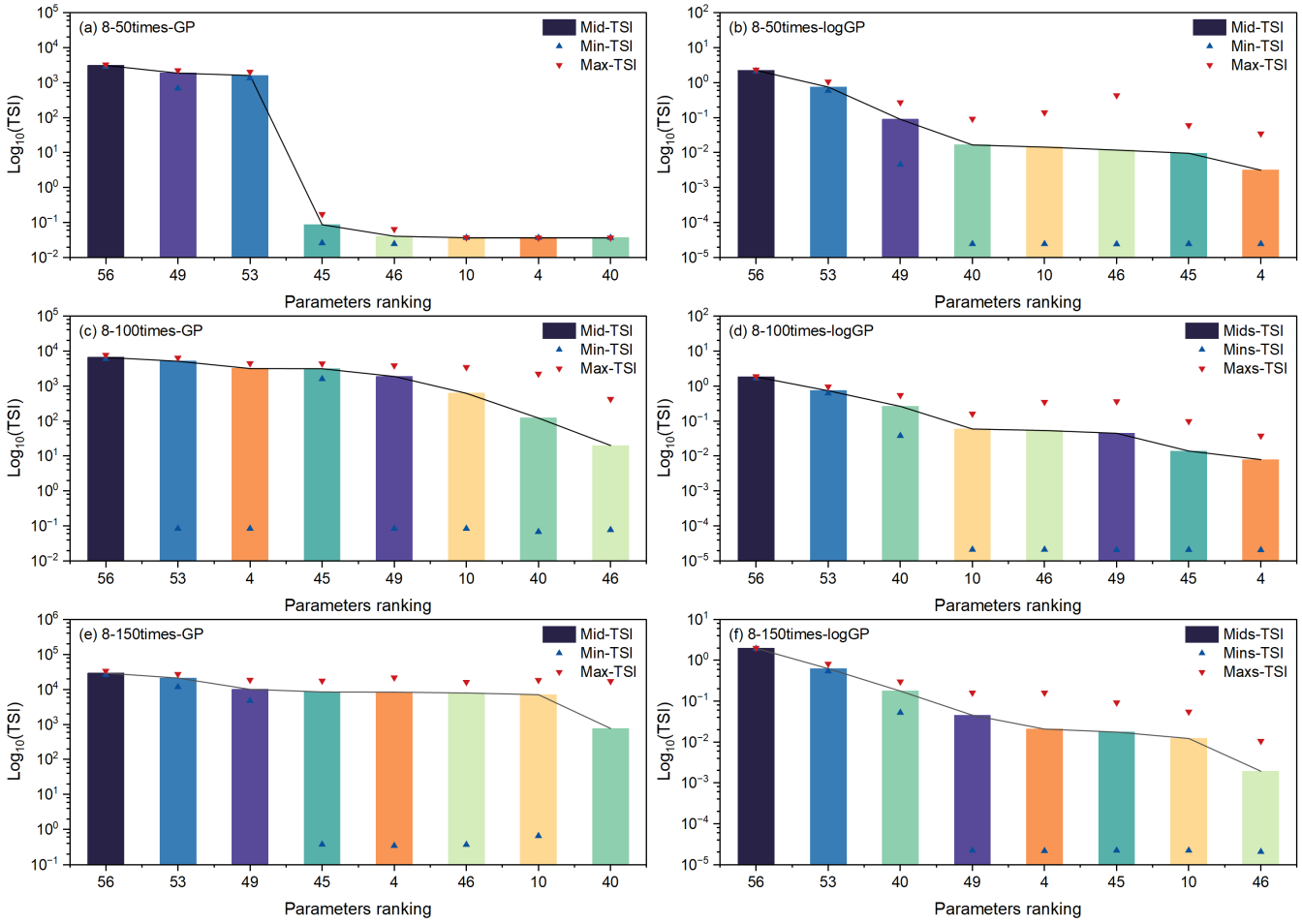


Figure S10: The ranking results of quantitative sensitivity indicators when the input dimension is 8.

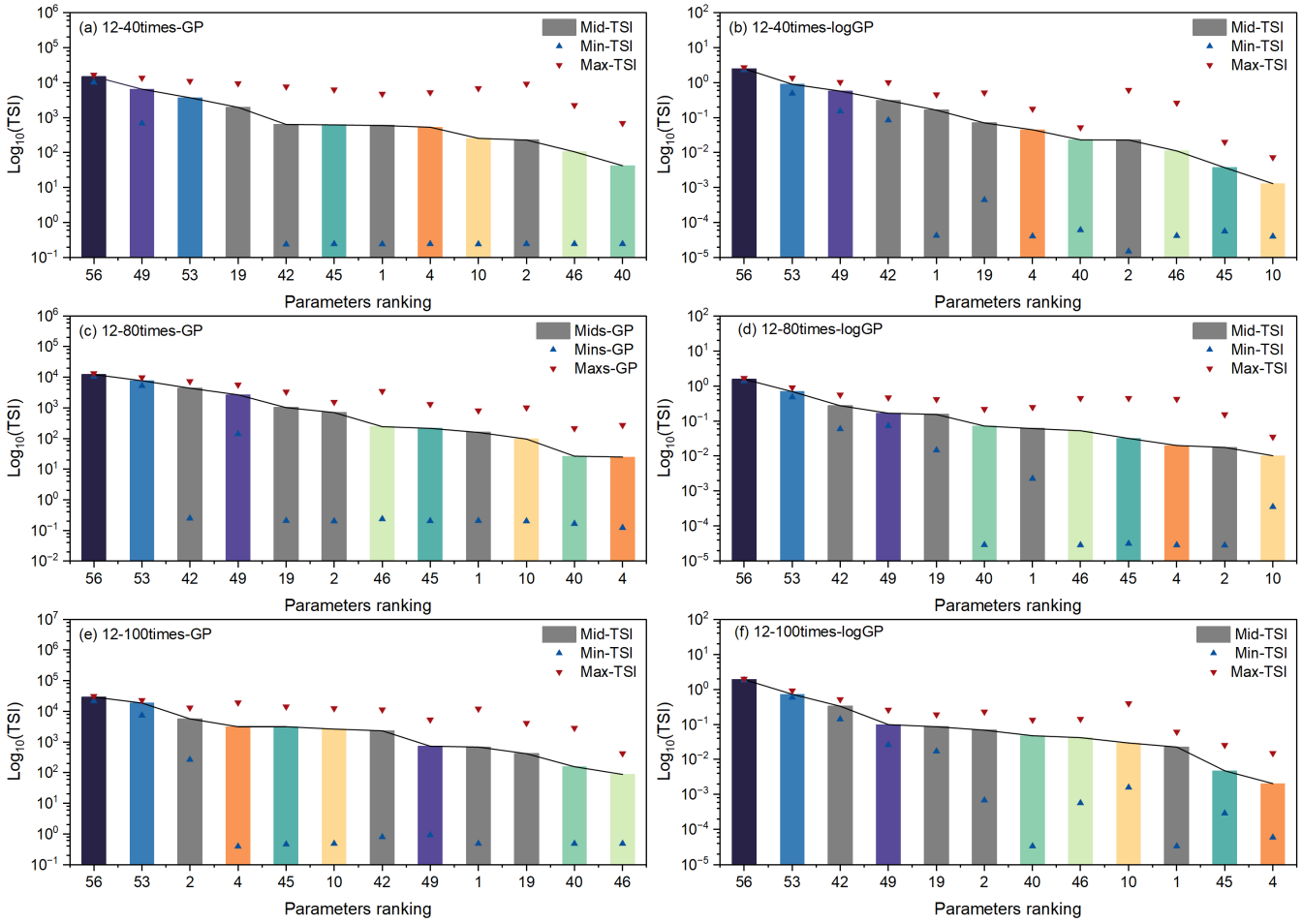


Figure S11: The ranking results of quantitative sensitivity indicators when the input dimension is 12.

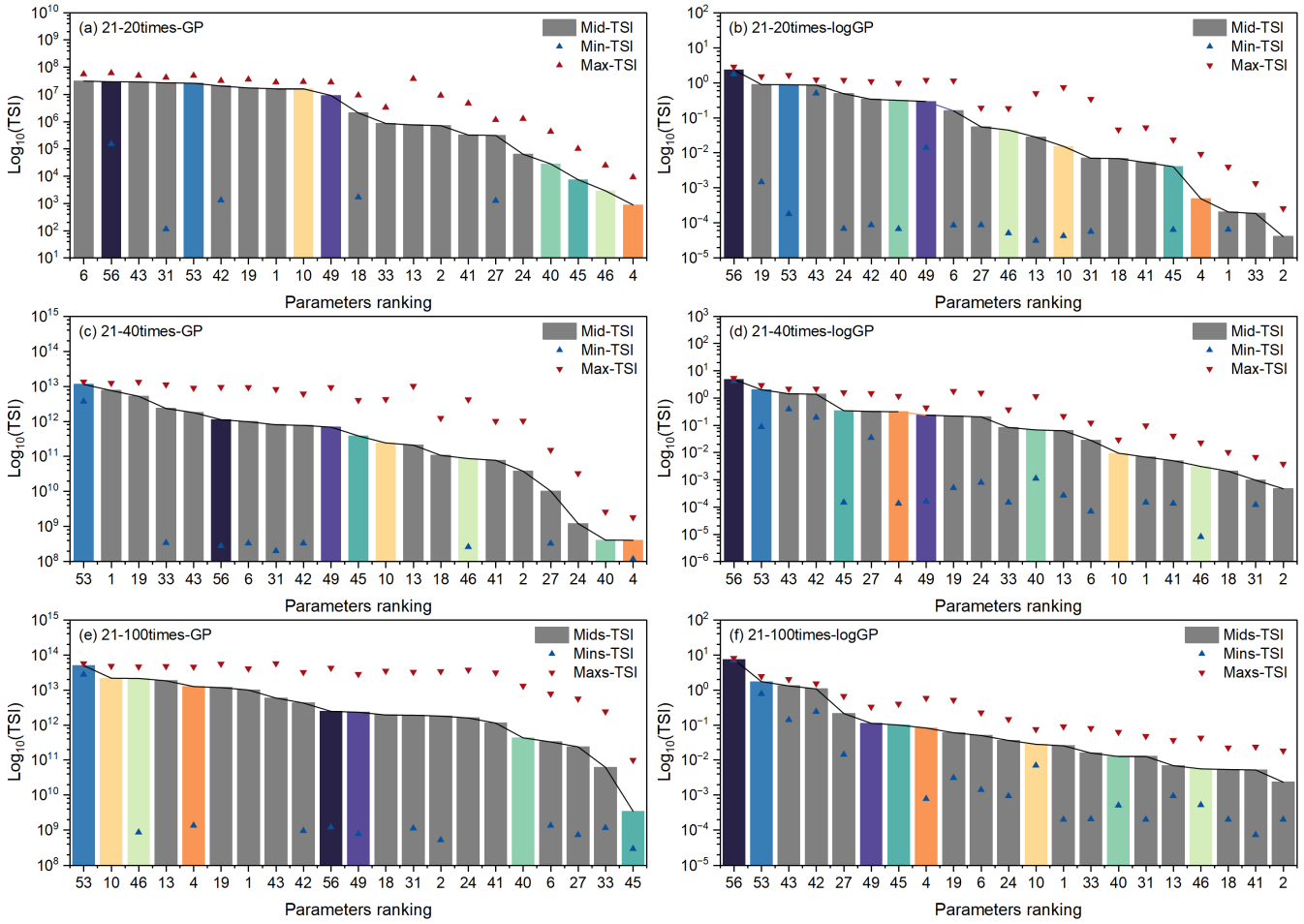


Figure S12: The ranking results of quantitative sensitivity indicators when the input dimension is 21.

S4: Results of optimal estimation of key sensitivity parameters

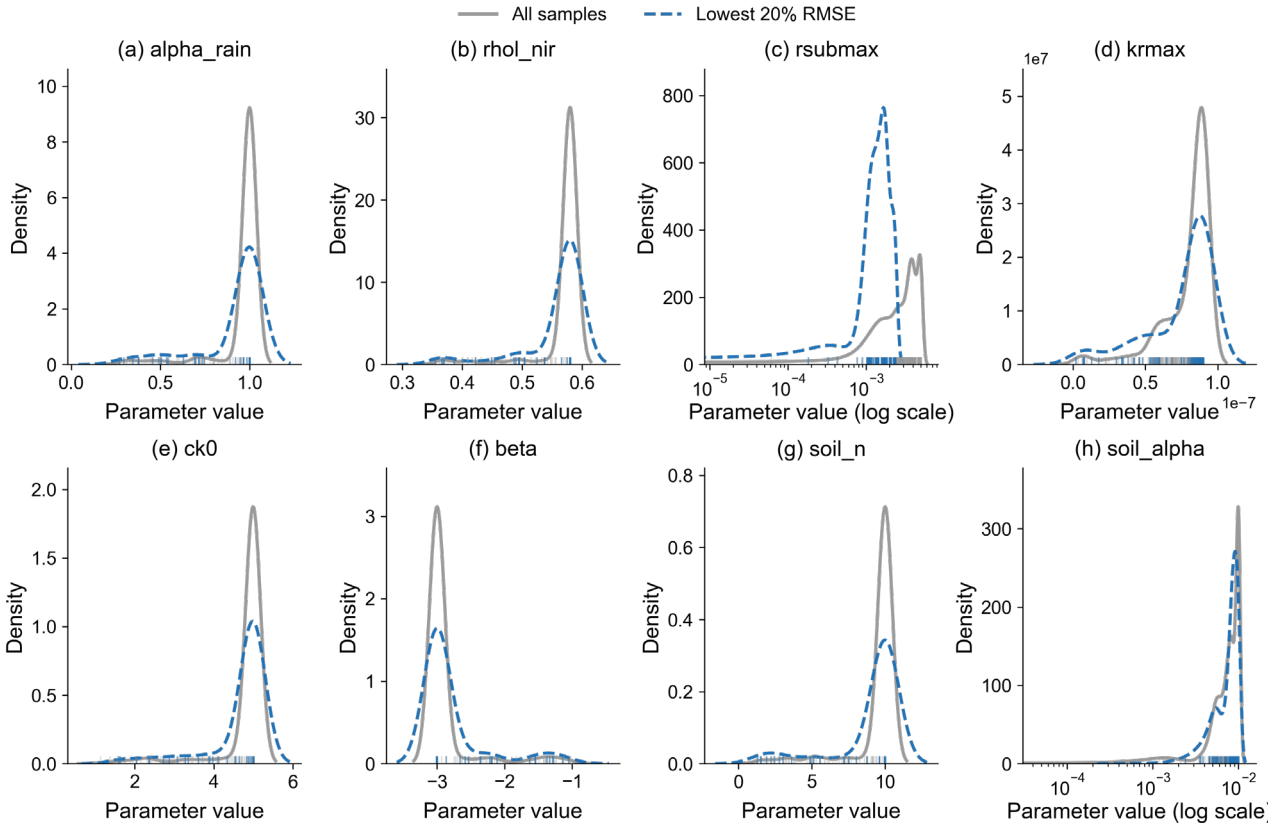


Figure S13: Comparison of KDE for the top 20% low-*RMSE* parameter sets (blue) versus the entire set of 496 evaluations (grey) obtained by the PSO method: (a) alpha_rain, (b) rhol_nir, (c) rsubmax, (d) krmax, (e) ck0, (f) beta, (g) soil_n, and (h) soil_alpha.

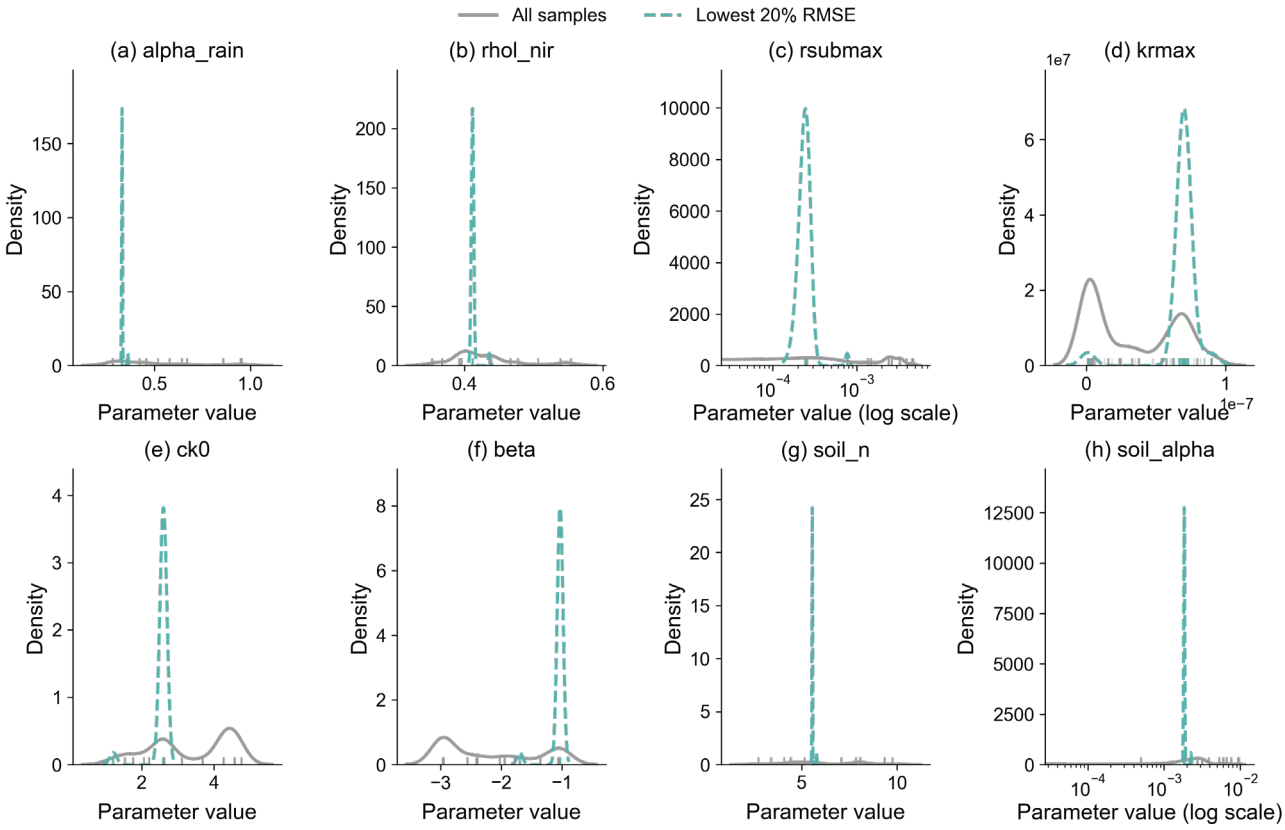


Figure S14: Comparison of KDE for the top 20% low-*RMSE* parameter sets (green) versus the entire set of 1000 evaluations (grey) obtained by the BOBYQA method: (a) alpha_rain, (b) rhol_nir, (c) rsubmax, (d) krmax, (e) ck0, (f) beta, (g) soil_n, and (h) soil_alpha.

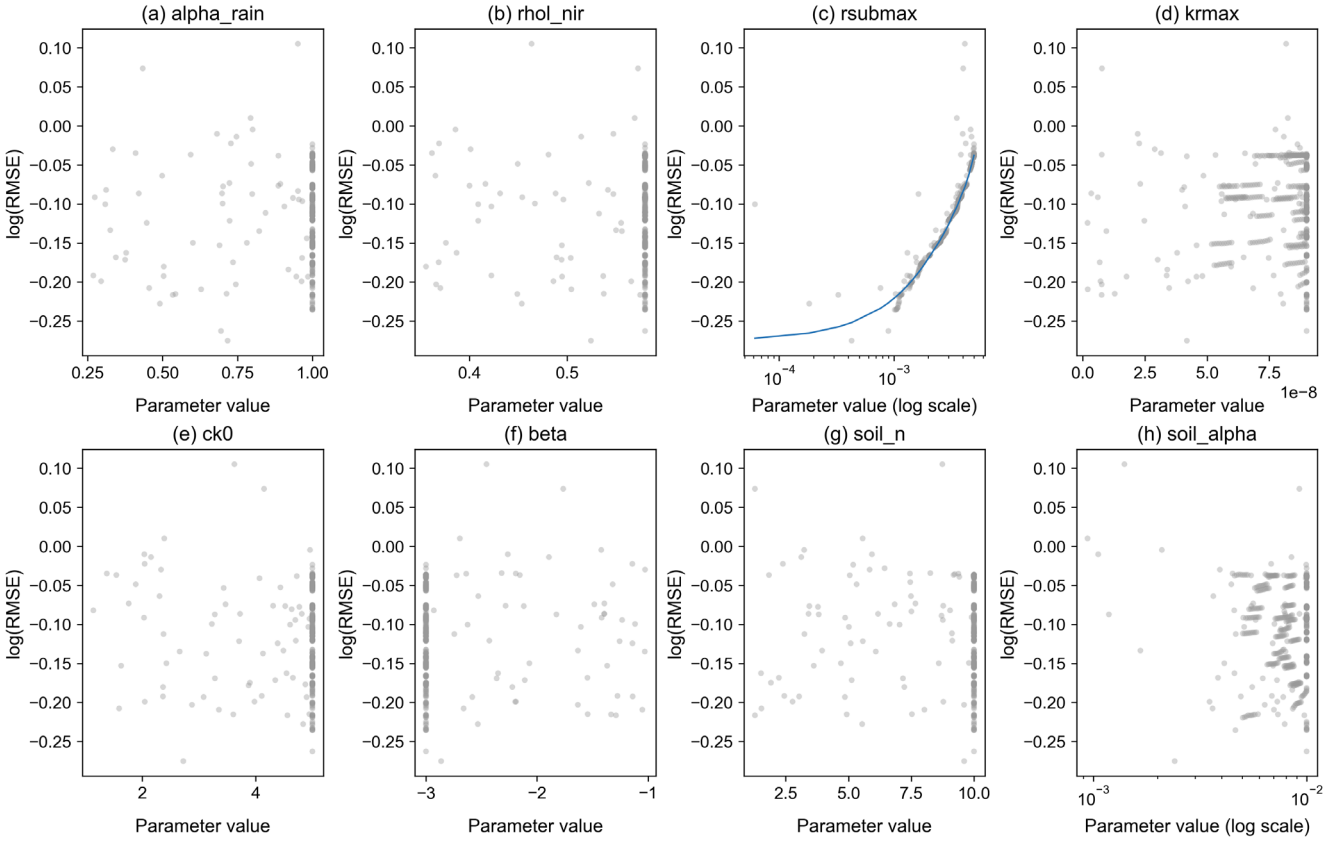
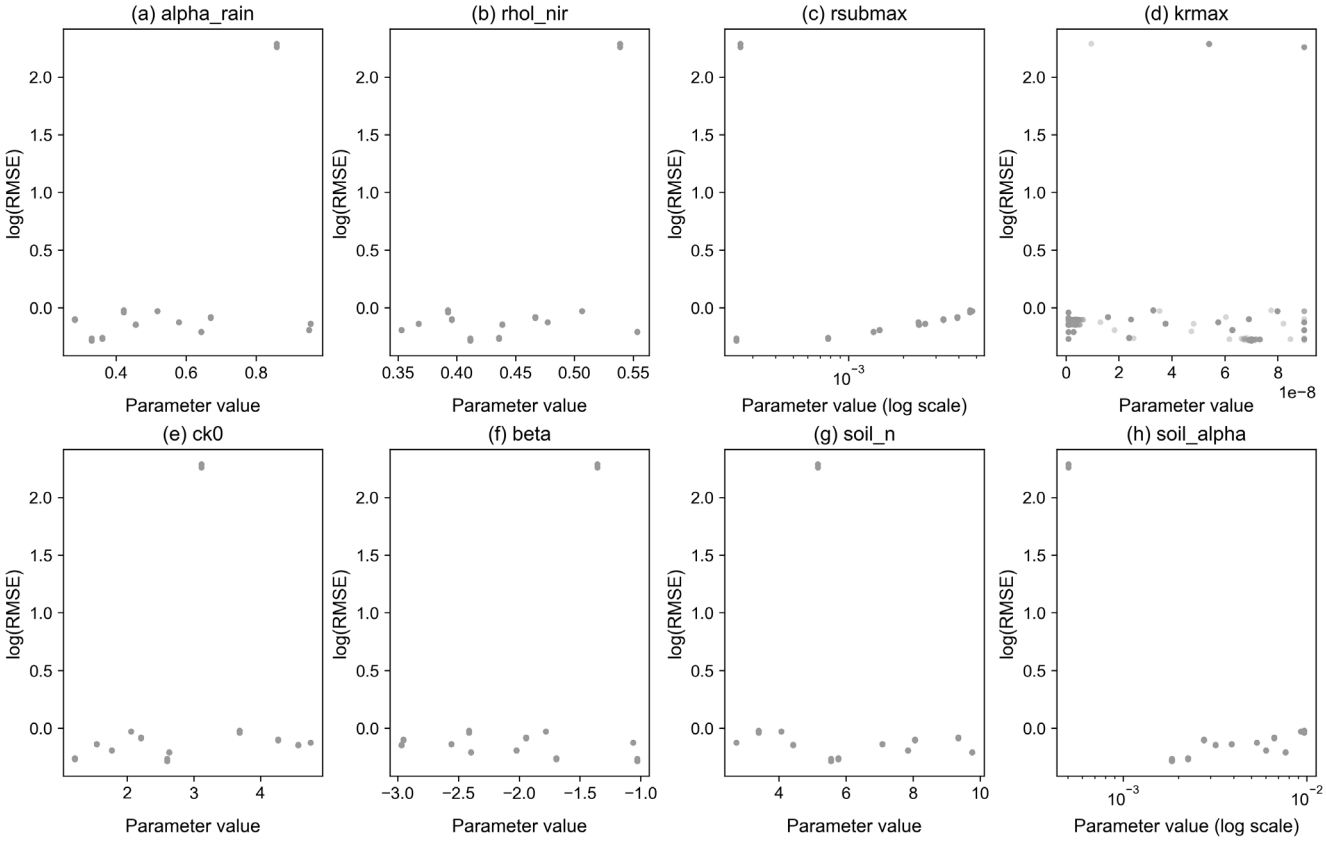


Figure S15: Scatter plots of *RMSE* against parameter values based on 496 evaluations obtained during the PSO optimization process: (a) alpha_rain, (b) rhol_nir, (c) rsubmax, (d) krmax, (e) ck0, (f) beta, (g) soil_n, and (h) soil_alpha.



FigureS16: Scatter plots of $RMSE$ against parameter values based on 1000 evaluations obtained during the BOBYQA optimization process: (a) alpha_rain, (b) rhol_nir, (c) rsubmax, (d) krmax, (e) ck0, (f) beta, (g) soil_n, and (h) soil_alpha.

S5: Mathematical formulations of eight stably sensitive parameters

This section provides the detailed governing equations for the parameters identified as stably sensitive to water table depth (WTD) dynamics in CoLM. Table S4 lists these parameters.

Table S4: Summary of the eight sensitive parameters.

Variable Name	Parameter Symbol	Reference Section
56-soil_alpha	α	Sect S5.1
53-soil_n	n	Sect S5.1
40-rsubmax	$r_{sub,max}$	Sect S5.2
49-beta	β	Sect S5.3

45-krmax	$k_{soil-soot,max}$	Sect S5.3
46-ck0	ck_0	Sect S5.3
4-alpha_rain	α_{rain}	Sect S5.4
10-rhol_nir	$\rho_{l,nir}$	Sect S5.4

S5.1 Unsaturated water movement

Water flux between adjacent soil layers is calculated as (Dai et al., 2019):

$$q_{hm}(\Delta z, h_u, h_l) = -K_{hm} \cdot \left(\frac{h_l - h_u}{\Delta z} - 1\right), \quad \text{(S1)}$$

where q_{hm} is a function for calculating water flux within a homogeneous soil medium (mm s^{-1}), Δz is the thickness of the soil layer (mm), , h_u and h_l are the soil water potentials at the upper and lower boundaries of the soil layer, respectively (mm), and K_{hm} is the equivalent hydraulic conductivity (mm s^{-1}), which is the function of hydraulic conductivity K of soil layer and parameter r_0 associated with soil hydraulic model and parameters of soil hydraulic properties.

The soil hydraulic properties are then described using the van Genuchten–Mualem (VGM) model (Ippisch et al., 2006):

$$S_e = \begin{cases} \frac{1}{S_c} \cdot [1 + (\alpha\psi)^n]^{-m}, & \psi < \psi_s \\ 1, & \psi \geq \psi_s \end{cases}, \quad \text{(S2)}$$

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r}, \quad \text{(S3)}$$

$$S_c = [1 + (\alpha\psi_s)^n]^{-m}, \quad \text{(S4)}$$

$$m = 1 - \frac{1}{n}, \quad \text{(S5)}$$

$$K = K_s \cdot S_e^l \cdot \left[\frac{1 - \left(1 - (S_e S_c)^{1/m}\right)^m}{1 - \left(1 - S_c^{1/m}\right)^m}\right]^2, \quad \text{(S6)}$$

$$r_0 = \frac{1}{l(n-1)+2n}, \quad \text{(S7)}$$

where S_e is the effective saturation (–), S_c is the saturation at the cut-off point (–), α is a shape parameter related to the inverse of the air-entry suction (mm^{-1}), n is a shape parameter representing the pore-size distribution (–), ψ is the unsaturated soil water potential (mm), ψ_s is the saturated matric potential (mm), θ is the volumetric water content (mm^3 mm^{-3}), θ_r is the residual soil water content (mm^3 mm^{-3}), θ_s is the saturated volumetric water content (mm^3 mm^{-3}), K is the unsaturated hydraulic conductivity (mm s^{-1}), K_s is the saturated hydraulic conductivity (mm s^{-1}), and l is the pore-connectivity parameter (–).

S5.2 Subsurface runoff parameterization

Subsurface runoff is parameterized as (Niu et al., 2005):

$$r_{subsurface} = r_{sub,max} e^{-f_{drai} \times z_{wt}}, \quad \text{(S8)}$$

where $r_{subsurface}$ is the subsurface runoff (mm s^{-1}), $r_{sub,max}$ is the maximum subsurface runoff when the max grid cell mean water table depth is zero (with a model default value of 5.5 $\times 10^{-3}$ mm s^{-1}) (mm s^{-1}), f_{drai} is the decay factor (with a model default value of 2.5 m^{-1}) (m^{-1}), and z_{wt} is water table depth (m).

When soil ice is present, the impedance factor is applied (Niu and Yang, 2006):

$$f_{impd,ice} = 1 - \frac{e^{-\alpha_{ice} \times \left(1 - f_{ice,sum}\right)} - e^{-\alpha_{ice}}}{1 - e^{-\alpha_{ice}}}, \quad \text{(S9)}$$

$$r_{subsurface} = f_{impd,ice} \times r_{sub,max} e^{-f_{drai} \times z_{wt}}, \quad \text{(S10)}$$

where $f_{impd,ice}$ is the impedance factor accounting for the effect of ice on soil permeability (–), α_{ice} is an adjustable scale-dependent parameter with a prescribed value of 3 in the model (–), $f_{ice,sum}$ represents the total volumetric ice content within and below the soil layer where the water table is located (–).

S5.3 Plant hydraulics and transpiration processes

S5.3.1 Root transpiration

Root transpiration is the primary component of the source-sink term in the soil moisture governing equation. Root water uptake is driven by the water potential gradient between the soil and the roots (Van Den Honert, 1948; Kennedy et al., 2019; Amenu and Kumar, 2008; Zhu et al., 2017):

$$q_{root,i} = k_{soil-soot,i}\left(\psi_{soil,i} - \psi_{root,i}\right), \quad \text{(S11)}$$

where $q_{root,i}$ represents the radial water transport rate (s^{-1}), $k_{soil-soot,i}$ is the soil-and-root conductance (mm s^{-1}), $\psi_{soil,i}$ and $\psi_{root,i}$ denote the water potentials of the i-th soil node and the root node, respectively (mm).

The conductance is calculated as (Kennedy et al., 2019):

$$k_{soil-soot,i} = \frac{k_{root,i} \cdot k_{soil,i}}{k_{root,i} + k_{soil,i}}, \quad \text{(S12)}$$

$$k_{root,i} = \frac{k_{soil-soot,max}}{l_{root,lateral} + z_{soil,i}} \cdot RAI \cdot 2^{-\left(\frac{\psi_{soil,i}}{p50_{root}}\right)^{c_k}}, \quad \text{(S13)}$$

$$RAI = (LAI + SAI) \cdot r_i \cdot f_{root-leaf}, \quad \text{(S14)}$$

$$k_{soil,i} = \frac{\min\left(k_{sat,i}, k_{s,i}\right)}{1000 dx_{root,i}}, \quad \text{(S15)}$$

where $k_{root,i}$ and $k_{soil,i}$ are the root and soil hydraulic conductance, respectively (s^{-1}), $k_{soil-soot,max}$ is the root conductance per area and per length (s^{-1}), $l_{root,lateral}$ is the lateral root length (mm), $z_{soil,i}$ is the soil layer depth (mm), RAI is the Root Area Index (–), $p50_{root}$ is the water potential at which 50% of root hydraulic conductivity is lost (mm), c_k is a shape factor

fitting the vulnerability curve (–), SAI and LAI are the Stem and Leaf Area Indices (–), r_i is the root fraction (–), $f_{root-leaf}$ is a scaling ratio (–), $k_{sat,i}$ and $k_{s,i}$ represent the saturated and unsaturated hydraulic conductivity of the i-th layer (mm s^{-1}), $dx_{root,i}$ is the distance between roots (m).

For root fraction, the model utilizes the scheme ROOTFR_SCHEME = 1. For each land cover type n, the root fraction for each soil layer is defined as (Schenk and Jackson, 2002):

$$r_{i,n} = \begin{cases} \frac{1}{\left(1+\frac{100z_{soil,1}}{d_{50,n}}\right)^{\beta_n}}, i = 1 \\ \frac{1}{\left(1+\frac{100z_{soil,i}}{d_{50,n}}\right)^{\beta_n}} - \frac{1}{\left(1+\frac{100z_{soil,i-1}}{d_{50,n}}\right)^{\beta_n}}, i = 2, nl_soil \\ 1 - \frac{1}{\left(1+\frac{100z_{soil,nl_{soil}-1}}{d_{50,n}}\right)^{\beta_n}}, i = nl_soil \end{cases}, \quad \text{(S16)}$$

where z_{soil} is the interface depth (m), d_{50} is the depth containing 50% of the root system (m), β is a dimensionless shape factor (–).

S5.3.2 Plant Hydraulic

By default, CoLM sets DEF_USE_PLANTHYDRAULICS = .true., The vegetation hydraulic module calculates transpiration based on the Soil-Plant-Atmosphere Continuum (SPAC) concept, where water transport is driven by water potential gradients.

Assuming a steady-state water transport within the plant, the governing equations are as follows (Kennedy et al., 2019):

$$q_{sun\leftarrow stem} = k_{sun\leftarrow stem}(\psi_{sun} - \psi_{stem}), \quad \text{(S17)}$$

$$q_{sha\leftarrow stem} = k_{sha\leftarrow stem}(\psi_{sha} - \psi_{stem}), \quad \text{(S18)}$$

$$q_{stem\leftarrow root} = k_{stem\leftarrow root}\left(\psi_{stem} - \psi_{root,0}\right), \quad \text{(S19)}$$

where q denotes the water flux (mm s^{-1}), K represents the hydraulic conductance (s^{-1}), ψ is the water potential (mm), the subscripts sun, sha, $stem$, $root$, and root,0 represent sunlit leaves, shaded leaves, stem, root and root at surface, respectively.

Hydraulic conductance varies with changes in plant water potential. It is based on a sigmoidal vulnerability curve (Kennedy et al., 2019):

$$k_{i\leftarrow j} = k_{max} \cdot 2^{-\left(\frac{\psi_j}{p50}\right)^{c_k}}, \quad \text{(S20)}$$

where k_{max} is the maximum hydraulic conductance (s^{-1}), ψ_j is the water potential at node j (mm), $p50$ represents the water potential at which hydraulic conductance is reduced by 50% (mm), c_k is the shape parameter of the vulnerability curve (–).

Simultaneously, the above-ground water fluxes satisfy the following mass conservation equations (Kennedy et al., 2019):

$$E_{sun} = q_{sun\leftarrow stem}, \quad \text{(S21)}$$

$$E_{sha} = q_{sha\leftarrow stem}, \quad \text{(S22)}$$

$$q_{sha\leftarrow stem} + q_{sun\leftarrow stem} = q_{stem\leftarrow root}, \quad \text{(S23)}$$

$$q_{stem\leftarrow root} = q_{root,0}, \quad \text{(S24)}$$

where E_{sun} and E_{sha} represent the transpiration rates for sunlit and shaded leaves, respectively (mm s^{-1}). These rates are determined by the potential transpiration and a transpiration reduction function controlled by leaf water potential (Kennedy et al., 2019):

$$E_{sun} = E_{sun,max} \cdot 2^{-\left(\frac{\psi_{sun}}{p50}\right)^{c_k}}, \quad \text{(S25)}$$

$$E_{sha} = E_{sha,max} \cdot 2^{-\left(\frac{\psi_{sha}}{p50}\right)^{c_k}}, \quad \text{(S26)}$$

where $E_{sun,max}$ and $E_{sha,max}$ are the transpiration rates under non-water-stressed conditions (mm s^{-1}).

S5.4 Surface water input and energy balance

S5.4.1 Precipitation after interception

The model employs the CoLM2014 scheme by default (DEF_Interception_scheme = 1), where the net precipitation reaching the ground p_{rain} is defined as:

$$p_{rain} = \frac{x_{sc} + thru_{rain}}{\Delta t}, \quad \text{(S27)}$$

where p_{rain} is the net precipitation reaching the ground (mm s^{-1}), x_{sc} represents the excess water exceeding the maximum canopy storage capacity (mm), $thru_{rain}$ denotes the canopy throughfall from effective precipitation areas (mm), Δt is the time duration (s).

When calculating $thru_{rain}$ (with DEF_VEG_SNOW = .false.), it is assumed that only liquid water can be shed from the canopy (Lawrence et al., 2019):

$$thru_{rain} = (R_c + R_l)\Delta t\left(1 - \delta_p\right)\frac{a_p}{b}\left(1 - e^{-bx_s}\right) + c_p x_s + L_{dew} x_s - S_c x_s, \quad \text{(S28)}$$

$$\delta_p = 1 - \alpha_{rain} \times [1 - e^{-K_p \times LSAI}], \quad \text{(S29)}$$

where R_c and R_l are the convective and large-scale rainfall rates (mm s^{-1}), δ_p is the canopy throughfall coefficient (–), x_s represents the fraction of the saturated area (–), L_{dew} is the depth of water on the foliage (mm), S_c is the maximum canopy storage capacity (mm), α_{rain} is an empirical coefficient reflecting the water-holding capacity of leaves (–), K_p represents the extinction coefficient (–), and $LSAI$ represents the sum of LAI and SAI (–).

S5.4.2 Snowmelt and energy balance

The snowmelt rate s_m is driven by the energy surplus or deficit:

$$s_m = \max(0, \frac{H_i \Delta t / L_f}{\Delta t}), \quad \text{(S30)}$$

$$H_i = f\left(h_s, T, T_f, \alpha_{Grank}, F, c, \Delta t, \Delta z\right), \quad \text{(S31)}$$

where s_m is the snowmelt rate (mm s^{-1}), H_i is the energy surplus or deficit (W m^{-2}), L_f is the latent heat of fusion (J kg^{-1}), h_s is the net heat flux entering the surface (W m^{-2}), T is the temperature of the snow/soil (K), T_f is the freezing temperature (K), α_{Grank} is the Crank-Nicholson factor (–), F is the heat flux (W m^{-2}), c is the volumetric heat capacity (J m^{-3} K^{-1}), Δt is the times duration, and Δz is the soil layer thickness.

The net heat flux entering the ground surface, h_s, is expressed as follows:

$$h_s = S_g + L_g(T_g) - H_g(T_g) - \lambda E_g(T_g) + H_{prcg}(T_g), \quad \text{(S32)}$$

where S_g and L_g represent the net solar (shortwave) radiation and net longwave radiation absorbed by the ground surface (W m^{-2}), H_g is the sensible heat flux (W m^{-2}), λE_g is the latent heat flux (W m^{-2}), H_{prcg} is the sensible heat flux from precipitation (W m^{-2}), and T_g is the soil surface temperature (K).

The net solar radiation, S_g, is expressed as follows:

$$S_g = S_{vg} - S_{sun} - S_{sha}, \quad \text{(S33)}$$

where S_{vg} is the total solar shortwave radiation flux absorbed by the surface (including both ground and vegetation) (W m^{-2}), S_{sun} and S_{sha} represent the radiation absorption by sunlit and shaded leaves, respectively (W m^{-2}).

The total radiation flux absorbed by the vegetation-soil system is given by:

$$S_{vg} = S_{vis,dir}(1 - \alpha_{vis,dir}) + S_{vis,dif}(1 - \alpha_{vis,dif}) + S_{nir.dir}(1 - \alpha_{nir.dir}) + S_{nir.dif}(1 - \alpha_{nir.dif}), \quad \text{(S34)}$$

where S on the right-hand side represents the incoming solar shortwave radiation flux across different spectral bands (visible, vis; near-infrared, nir) (W m^{-2}) and radiation types (direct, dir; diffuse, dif) (W m^{-2}), and the term α is the albedo (–).

In the radiative transfer scheme, α_{veg} is a function of the leaf albedo ω (–), the Leaf Area Index LAI (–), the solar zenith angle μ (–), and the leaf angle distribution factor χ_L (–). Considering the canopy nir albedo, the albedo is defined as:

$$\alpha_{veg,nir} = f(w_{nir}, \mu, LAI, \chi_L), \quad \text{(S35)}$$

The albedo w_{nir} is defined as:

$$w_{nir} = \rho_{l,nir} + \tau_{l,nir}, \quad \text{(S36)}$$

where $\rho_{l,nir}$ is the leaf reflectance (–) and $\tau_{l,nir}$ is the leaf transmittance (–).

S6: Results of the qualitative sensitivity parameter screening based on additional random seed experiments

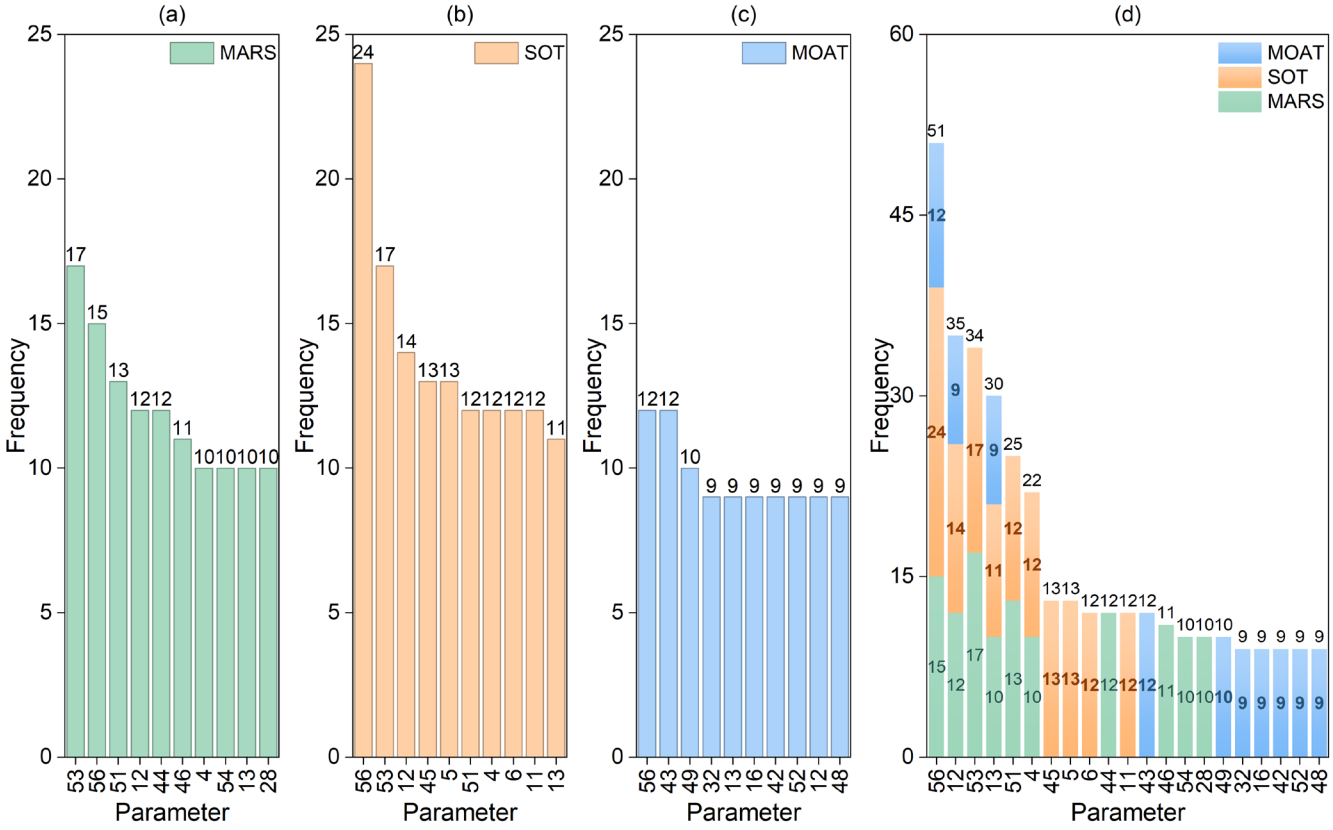


Figure S17: Sensitive parameters identified across various objectives and sample sizes under different sensitivity analysis methods based on experiment s1. The horizontal axis denotes parameter indices, while the vertical axis represents the frequency of occurrence. Labels above the bars indicate the total frequency for each parameter. Subplots (a)–(c) display the 10 most frequently identified sensitive parameters under the MARS, SOT, and MOAT analysis methods, respectively. Subplot (d) shows the integrated results of (a)–(c), where labels within the bars indicate the frequency counts for each specific analysis method.

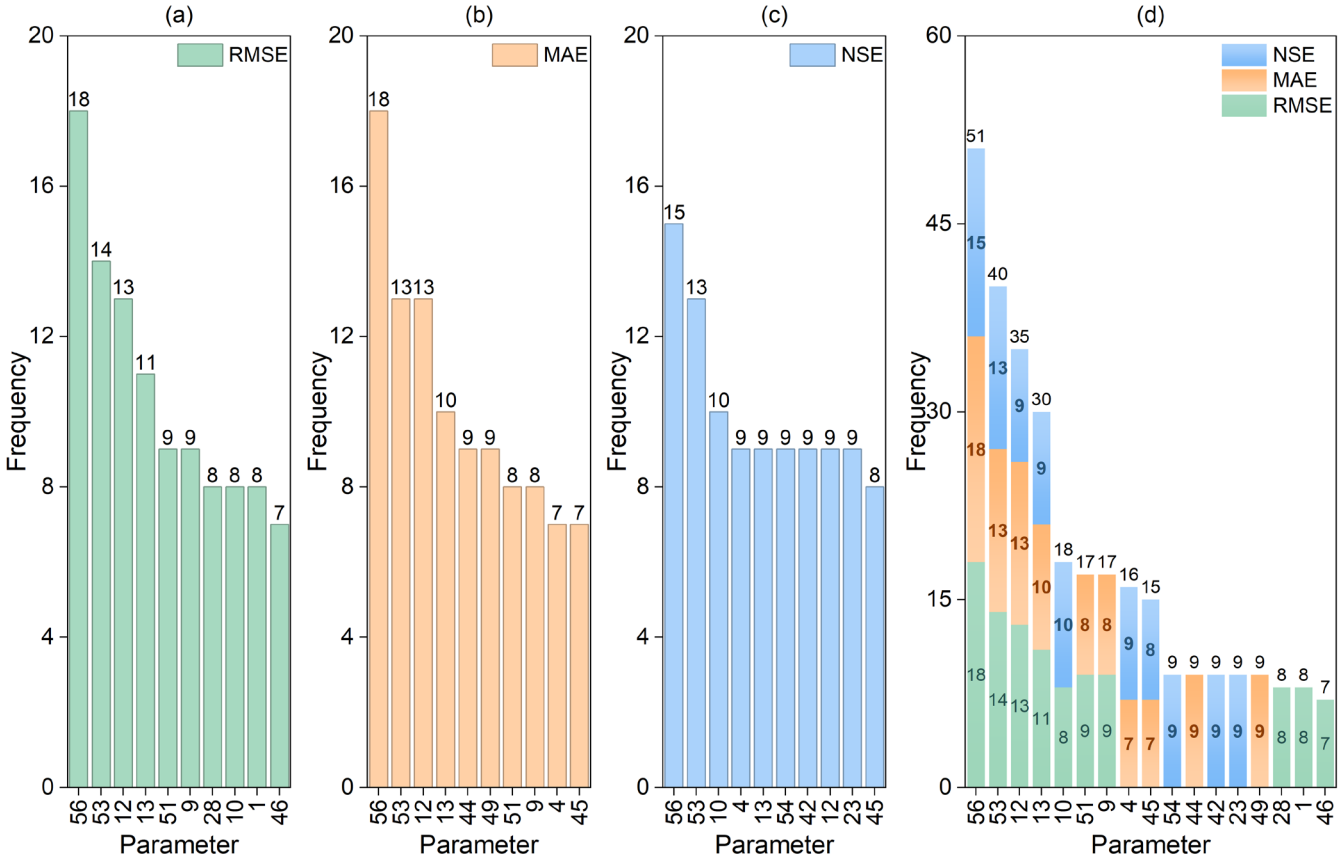


Figure S18: Sensitive parameters identified across various methods and sample sizes under different objective functions based on experiment s1. The horizontal axis denotes parameter indices, while the vertical axis represents the frequency of occurrence. Labels above the bars indicate the total frequency for each parameter. Subplots (a)–(c) display the 10 most frequently identified sensitive parameters under the *RMSE*, *MAE*, and *NSE* objective functions, respectively. Subplot (d) shows the integrated results of (a)–(c), where labels within the bars indicate the frequency counts for each specific objective function.

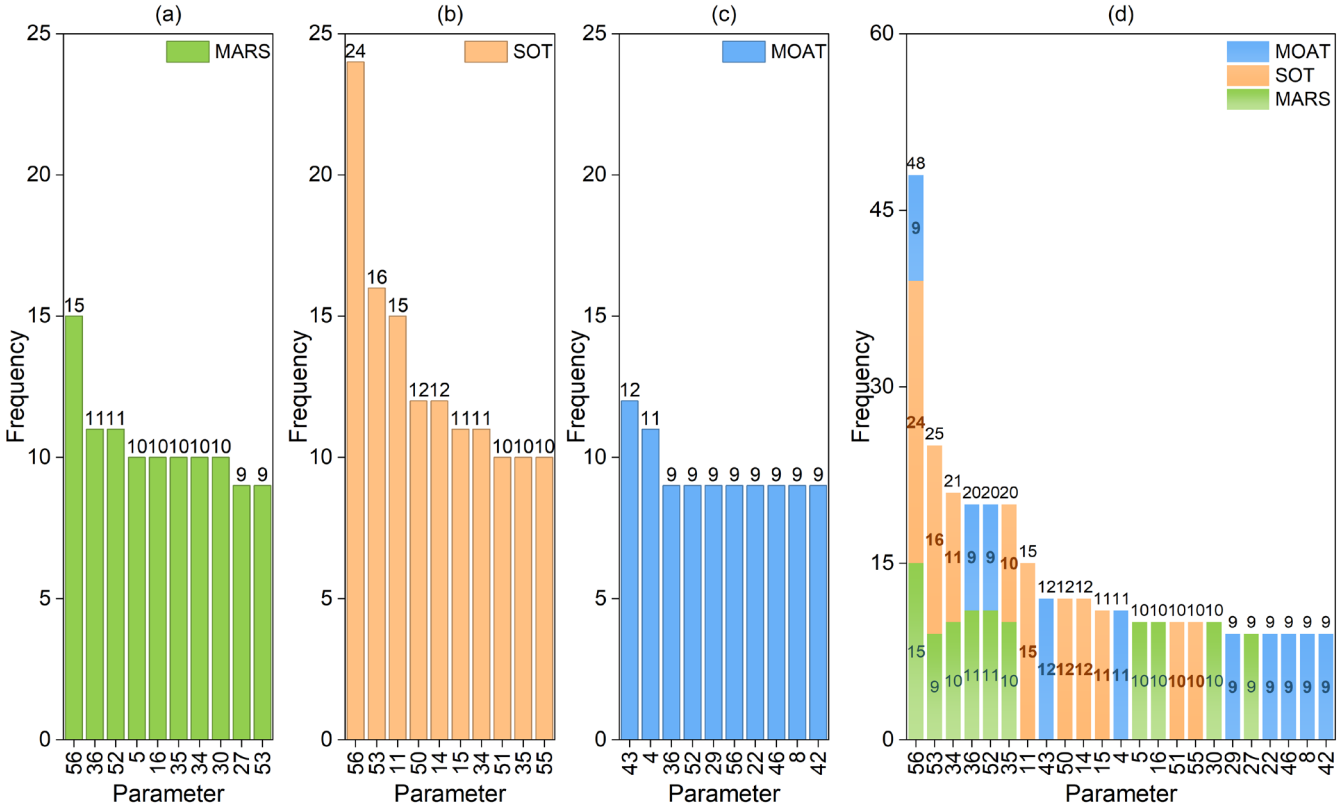


Figure S19: Sensitive parameters identified across various objectives and sample sizes under different sensitivity analysis methods based on experiment s2. The horizontal axis denotes parameter indices, while the vertical axis represents the frequency of occurrence. Labels above the bars indicate the total frequency for each parameter. Subplots (a)–(c) display the 10 most frequently identified sensitive parameters under the MARS, SOT, and MOAT analysis methods, respectively. Subplot (d) shows the integrated results of (a)–(c), where labels within the bars indicate the frequency counts for each specific analysis method.

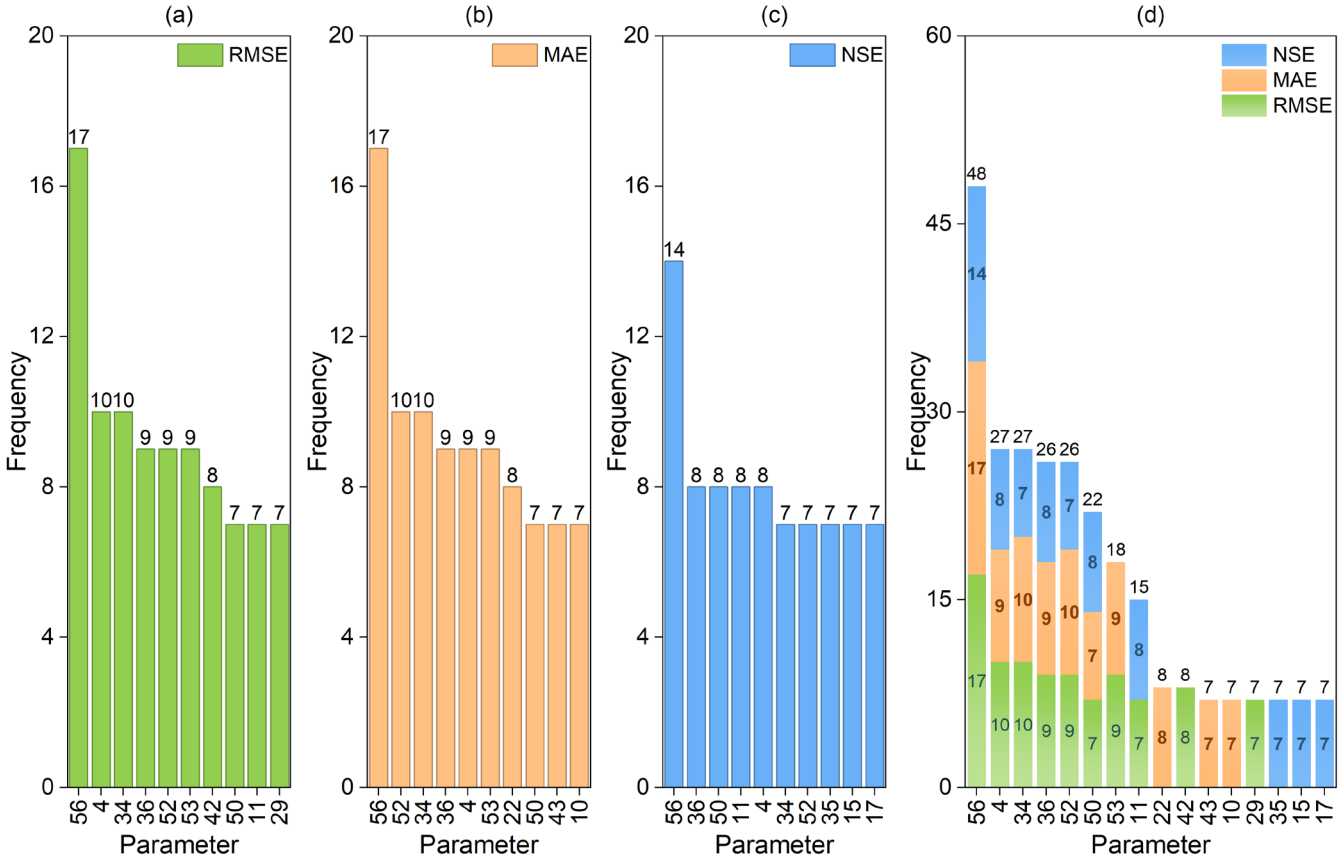


Figure S20: Sensitive parameters identified across various methods and sample sizes under different objective functions based on experiment s2. The horizontal axis denotes parameter indices, while the vertical axis represents the frequency of occurrence. Labels above the bars indicate the total frequency for each parameter. Subplots (a)–(c) display the 10 most frequently identified sensitive parameters under the *RMSE*, *MAE*, and *NSE* objective functions, respectively. Subplot (d) shows the integrated results of (a)–(c), where labels within the bars indicate the frequency counts for each specific objective function.

S7: Convergence criteria of the parameter optimization methods

Among the convergence criteria, the relevant user-defined settings in the PSUADE input file include *optimization max_feval*, *optimization* f_{min}, and *optimization tolerance*, which specify the maximum number of objective-function evaluations, the target objective-function value, and the convergence tolerance, respectively. Following the example settings provided in PSUADE and considering the actual performance of the WTD simulations, *optimization max_feval*, *optimization* f_{min}, and *optimization tolerance* were set to 10000, 0.4, and 1×10^{-4}, respectively.

S7.1 Implementation of SCE optimization convergence criteria in PSUADE

The convergence criteria for the SCE optimization are explicitly defined according to the implementation of the *SCEOptimizer.cpp* source code in PSUADE. Each SCE optimization run is terminated when any of the following conditions are satisfied:

(1) The maximum number of objective-function evaluations is reached:

$$N_{FE} \geq N_{FE,max}, \tag{S37}$$

where N_{FE} is the number of objective-function evaluations performed during the SCE optimization process, and $N_{FE,max}$ is the prescribed maximum allowable number of objective-function evaluations.

(2) The normalized parameter-space geometric range falls below an internal threshold:

$$GR < GR_{threshold}, \tag{S38}$$

$$GR = \exp\left(\frac{1}{m}\sum_{i=1}^{m} ln\left(\frac{x_{i,max}-x_{i,min}}{U_i-L_i}\right)\right), \tag{S39}$$

where GR denotes the normalized geometric range, $GR_{threshold}$ is the parameter space convergence threshold specified in PSUADE (0.0001), $x_{i,max}$ and $x_{i,min}$ are the maximum and minimum sampled values of parameter i, respectively, m is the number of parameters, and U_i and L_i are the upper bound and lower bound of the value of parameter i, respectively.

(3) The relative improvement of the best objective value over the recent evolution history becomes smaller than the user-specified tolerance:

$$C_f = \frac{100\left|f_{k-1}^{best}-f_{k-n}^{best}\right|}{max\left(1,\frac{1}{n}\sum_{j=k-n}^{j=k-1}\left|f_j^{best}\right|\right)} < \varepsilon, \tag{S40}$$

where f_k^{best} is the best objective value after the k-th evolution loop, k is the number of evolution loops, n is the number of recent evolution loops considered in the convergence test, and ε is the user-defined optimization tolerance.

In the PSUADE input file, *optimization max_feval* specifies the maximum number of evaluations for each SCE optimization run, whereas *optimization tolerance* defines the tolerance for the objective-function improvement criterion.

S7.2 Implementation of BOBYQA optimization convergence criteria in PSUADE

The convergence criteria for the BOBYQA optimization are explicitly defined according to the implementation of the *BobyqaOptimizer.cpp* and *bobyqa.f* source code in PSUADE. Each BOBYQA search is terminated when either of the following criteria is satisfied:

(1) The maximum number of objective-function evaluations is reached:

$$N_{FE} \geq N_{FE,max}, \tag{S41}$$

where N_{FE} is the number of objective-function evaluations performed during the BOBYQA optimization process, and $N_{FE,max}$ is the prescribed maximum allowable number of objective-function evaluations.

(2) The trust-region radius has been progressively reduced to the prescribed final trust-region radius:

$$\rho < \rho_{end}, \tag{S42}$$

$$\rho_{end} = \rho_{beg} \times \varepsilon, \tag{S43}$$

$$\rho_{beg} = \frac{1}{2}\min_i(U_i - L_i), \tag{S44}$$

where ρ is the current trust-region radius, ρ_{end} is the final trust-region radius, ε is the user-defined optimization tolerance specified in the PSUADE input file, ρ_{beg} is the initial trust-region radius, and U_i and L_i denote the upper and lower bounds of parameter i, respectively.

In the PSUADE input file, *optimization max_feval* specifies the maximum number of objective-function evaluations, whereas *optimization tolerance* controls the final trust-region resolution.

S7.3 Implementation of PSO optimization convergence criteria in PSUADE

The convergence criteria for the PSO optimization are explicitly defined according to the implementation of the *PSOptimizer.cpp* source code in PSUADE. The PSO iterations are continued only while all of the following conditions are satisfied:

(1) The maximum number of objective-function evaluations is reached:

$$tN_p \geq N_{FE,max}, \tag{S45}$$

where t is the number of completed epochs, N_p is the swarm size, and $N_{FE,max}$ is the prescribed maximum allowable number of objective-function evaluations.

(2) The current global-best objective-function value is sufficiently close to the prescribed target value:

$$\left|f_{gbest} - f_{target}\right| \leq \varepsilon, \tag{S46}$$

$$f_{gbest} = \min_{\substack{0\leq\tau\leq t \\ 1\leq i\leq N_p}} f\left(x_i^{(\tau)}\right), \tag{S47}$$

where f_{gbest} is the current global optimal objective function value, f_{target} is the prescribed target value, ε is the prescribed tolerance, t is the current number of completed PSO iterations, τ is the iteration index, i is the particle number, and N_p is the total number of particles in the particle swarm.

(3) No new global-best solution is found for consecutive particle evaluations:

$$N_{stag} \geq 20N_p, \tag{S48}$$

$$N_{stag}^{(k+1)} = \begin{cases} N_{stag}^{(k)} + 1, & f(x_k) \geq f_{gbest}^{(k)} \\ 0, & f(x_k) < f_{gbest}^{(k)} \end{cases}, \tag{S49}$$

where N_{stag} is the stagnation counter, N_p is the number of particles, k denotes the k-th particle evaluation, $f(x_k)$ is the current particle's objective function value, and f_{gbest} is the current global optimal objective function value.

In the PSUADE input file, *optimization max_feval* specifies the maximum number of objective-function evaluations, whereas *optimization tolerance* controls the allowable absolute difference between the current global-best objective value and the prescribed target value. The prescribed target value is specified by *optimization* f_{min}, which defines the expected minimum objective-function value and is passed to the PSO optimizer. The stagnation multiplier of 20 is internally defined in *PSOptimizer.cpp*, while the number of particles is 50 by default and can be modified in optimization expert mode.

S8: Diagnosis of unresolved WTD fluctuation events

To further diagnose the factors contributing to the unresolved event-scale WTD fluctuations, the 1000 WTD simulations generated during the SCE search were additionally evaluated using the Kling–Gupta efficiency (KGE), followed by analyses of meteorological forcing uncertainty and freeze–thaw-related limitations. KGE was calculated as (Kling et al., 2012):

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2}, \quad \text{(S50)}$$

$$\beta = \frac{\mu_s}{\mu_o}, \quad \text{(S51)}$$

$$\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o}, \quad \text{(S52)}$$

where KGE is the modified KGE-statistic, r is the correlation coefficient between simulated and observed values, β is the bias ratio, γ is the variability ratio, μ is the mean value, CV is the coefficient of variation, σ is the standard deviation, and the indices s and o represent simulated and observed values, respectively.

S8.1 Influence of objective functions on event-scale WTD dynamics

To diagnose the influence of the objective function, the 1000 WTD simulations generated during the SCE search were evaluated using additional performance metrics, including overall KGE and RMSE/KGE values during the two groundwater-rise periods (Fig. S21). KGE is commonly used during model calibration because it considers correlation, variability, and bias components, whereas RMSE is frequently used for evaluating and comparing model accuracy (Williams, 2025). Comparison between the minimum RMSE solution and the maximum KGE solution showed that the minimum RMSE solution achieved a KGE of 0.435 (Fig. S21a), whereas the maximum KGE among the 1000 simulations reached 0.529, with only a 0.123 m difference in RMSE. This indicates that parameter combinations with substantially improved temporal dynamics existed within the RMSE convergence region, but they were not favoured when RMSE was used as the sole optimization objective.

Similarly, the parameter combination achieving the lowest global RMSE did not correspond to the lowest RMSE during the two groundwater-rise events (Fig. S21b). This occurred because the event periods accounted for only a limited proportion of the entire simulation period. Therefore, optimization based on global RMSE preferentially reduced overall errors at the expense of event-scale deviations. When different evaluation metrics were used to select optimal simulations from the RMSE-driven parameter search results (Fig. S21c), RMSE tended to favour simulations with lower absolute errors and reduced variability. This finding is consistent with Gupta et al. (2009), who demonstrated that error-based metrics may favour models with underestimated variability. In contrast, KGE explicitly incorporates correlation and variability components and therefore better preserves temporal dynamics and overall fluctuations (Kling et al., 2012). It should be noted that the parameter sets selected according to KGE, event-scale RMSE, and event-scale KGE were obtained from the existing simulations rather than from independent optimization experiments using these metrics as objective functions. Therefore, the potential benefits of adopting multi-objective optimization strategies or alternative objective functions for improving event-scale groundwater dynamics require further investigation.

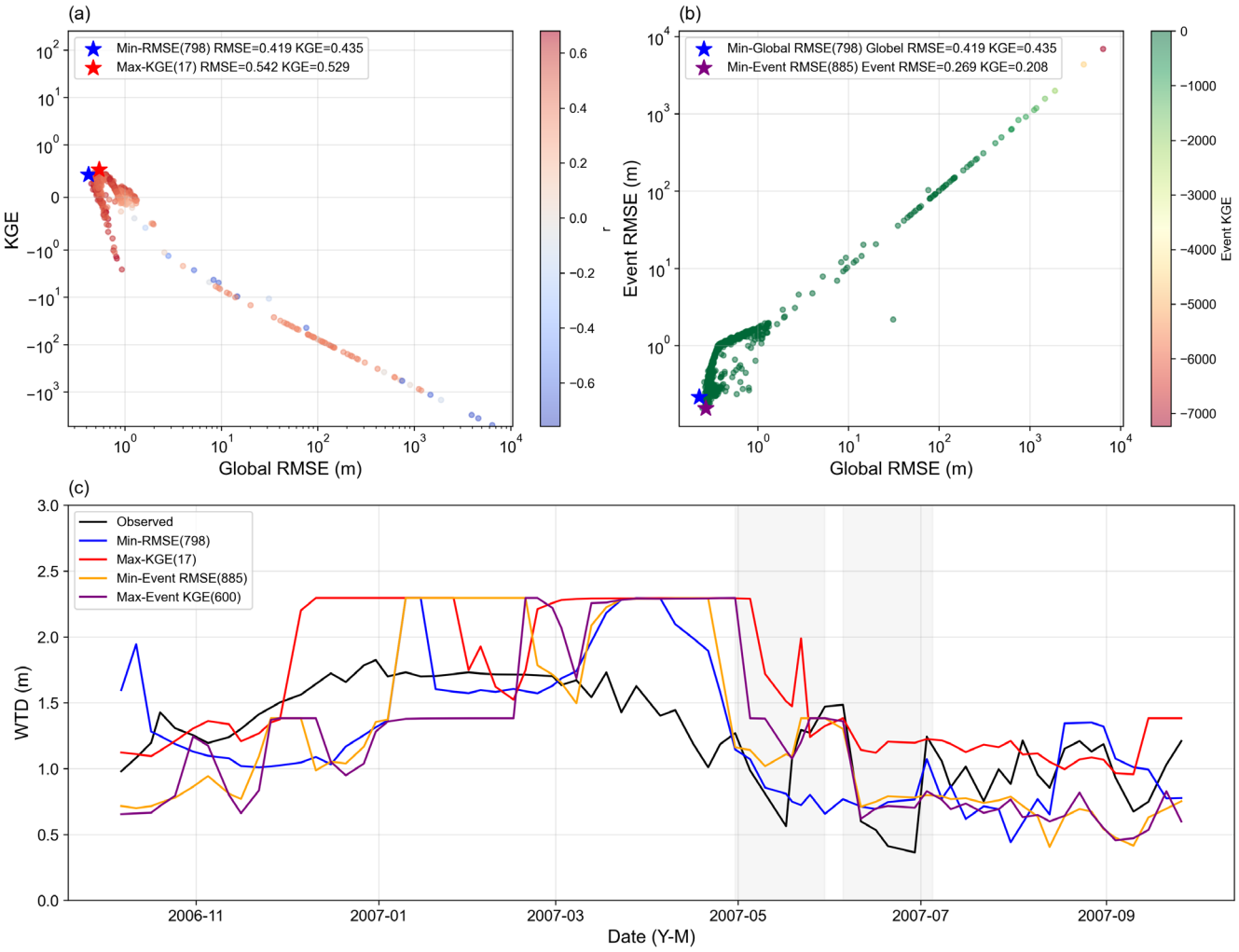


Figure S21: Comparison of model performance and simulated WTD hydrographs selected by different optimization metrics based on 1000 simulations: (a) KGE versus Global RMSE; (b) Event RMSE versus Global RMSE; and (c) Observed and simulated WTD time series under optimal parameter sets for different evaluation targets. The shaded areas in (c) indicate the periods of pronounced declines in the observed WTD, which were also used for calculating the event-based metric. The numbers in parentheses in the legends indicate the specific simulation ID (out of 1000 runs) that achieved the corresponding optimal metric.

S8.2 Influence of meteorological forcing uncertainty

Besides the influence of objective functions, uncertainties in meteorological forcing and observations may also contribute to the discrepancies in simulated WTD dynamics. As shown in Fig. S22a, during the months corresponding to major groundwater-rise events, the monthly precipitation from the TPMFD dataset was comparable with observations in June but lower in May (Yang et al., 2020). Although TPMFD has demonstrated good performance over the Tibetan Platea (Liu et al., 2025), discrepancies may still exist at individual sites, resulting in underestimated WTD responses under insufficient precipitation forcing. Furthermore, monthly air temperature from TPMFD was consistently lower than observations (Fig. S22b), which may delay soil thawing and reduce effective groundwater recharge during the freeze–thaw transition period.

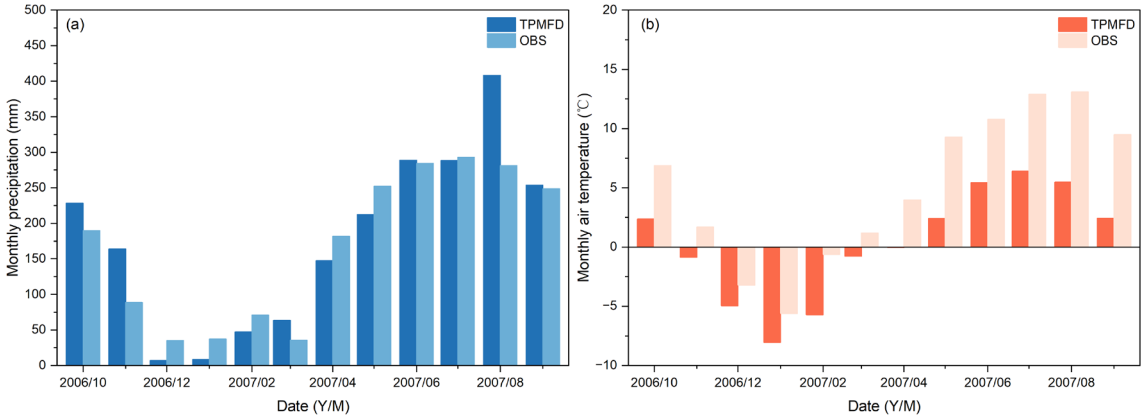


Figure S22: Comparison of monthly precipitation (a) and mean air temperature (b) between the TPMFD dataset and in-situ observations.

S8.3 Influence of freeze–thaw processes on groundwater recharge

Around the first observed groundwater-rise event, the optimal simulation indicated relatively large net infiltration. However, soil temperature in the eighth layer remained below the freezing point, resulting in considerable ice content and reduced liquid water availability (Fig. S23a). Consequently, soil matric potential and hydraulic conductivity remained low, limiting effective groundwater recharge. The subsurface runoff component, which is a deterministic function of WTD in CoLM, remained relatively stable when WTD was stable. Combined with evaporative losses, the soil–aquifer exchange flux remained positive, indicating water loss from the aquifer; however, the simulated stable WTD suggested a dynamic balance between limited recharge from the unsaturated zone and groundwater drainage. Across the soil profile (Fig. S23b), ice below approximately 0.6 m depth resulted in low hydraulic conductivity and weak water transport capacity. During the second groundwater-rise event, the soil profile gradually thawed, and liquid water content, matric potential, and hydraulic conductivity increased. However, net infiltration and liquid soil water at about 0–0.3 m remained relatively limited, resulting in insufficient groundwater recharge.

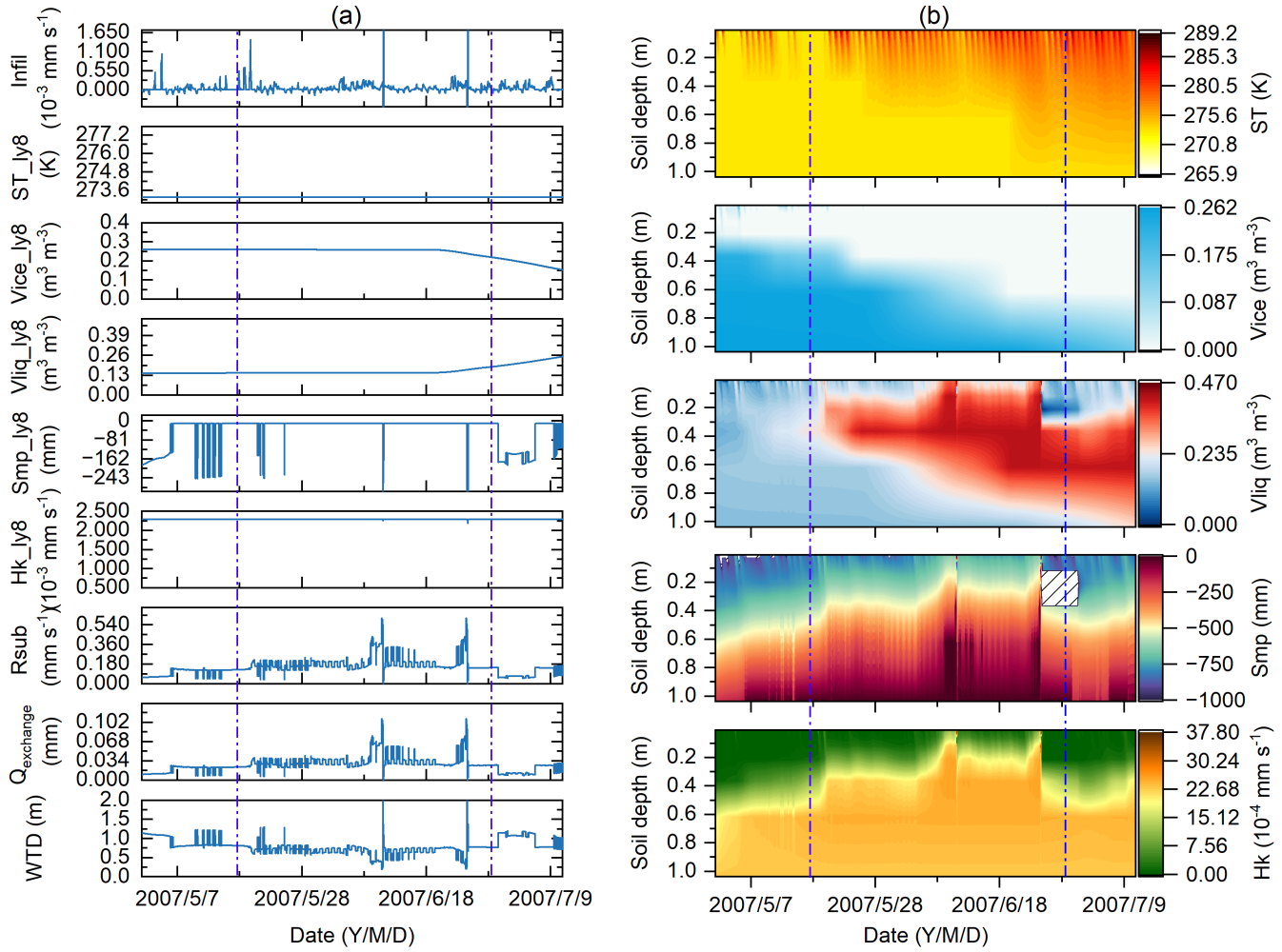


Figure S23: The key variables obtained through simulation corresponding to the optimal result of the SCE algorithm during the period when the measured WTD rose significantly: (a) Time series of infiltration (Infil), 8th-layer soil temperature (ST_ly8), 8th-layer partial volume of ice lens in layer (Vice_ly8), 8th-layer liquid volumetric water content (Vliq_ly8), 8th-layer soil matric potential (Smp_ly8), 8th-layer soil hydraulic conductivity (Hk_ly8), subsurface runoff (R_{sub}), water exchange between soil and aquifer ($Q_{exchange}$), and water table depth (WTD); (b) Spatial and temporal variations of soil temperature (ST), partial volume of ice lens in layer (Vice), liquid volumetric water content (Vliq), soil matric potential (Smp) and soil hydraulic conductivity (Hk). The two dashed lines in the plots mark the timing of the observed WTD peaks.

References

Amenu, G. G. and Kumar, P.: A model for hydraulic redistribution incorporating coupled soil-root moisture transport, Hydrol. Earth Syst. Sci., 12, 55-74, https://doi.org/10.5194/hess-12-55-2008, 2008.

Dai, Y., Zhang, S., Yuan, H., and Wei, N.: Modeling Variably Saturated Flow in Stratified Soils With Explicit Tracking of Wetting Front and Water Table Locations, Water Resour. Res., 55, 7939-7963, https://doi.org/10.1029/2019wr025368, 2019.

Ippisch, O., Vogel, H. J., and Bastian, P.: Validity limits for the van Genuchten–Mualem model and implications for parameter estimation and numerical simulation, Adv. Water Resour., 29, 1780-1789, https://doi.org/10.1016/j.advwatres.2005.12.011, 2006.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80-91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Kennedy, D., Swenson, S., Oleson, K. W., Lawrence, D. M., Fisher, R., Lola da Costa, A. C., and Gentine, P.: Implementing Plant Hydraulics in the Community Land Model, Version 5, J. Adv. Model. Earth Syst., 11, 485-513, https://doi.org/10.1029/2018MS001500, 2019.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424-425, 264-277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., Collier, N., Ghimire, B., van Kampenhout, L., Kennedy, D., Kluzek, E., Lawrence, P. J., Li, F., Li, H., Lombardozzi, D., Riley, W. J., Sacks, W. J., Shi, M., Vertenstein, M., Wieder, W. R., Xu, C., Ali, A. A., Badger, A. M., Bisht, G., van den Broeke, M., Brunke, M. A., Burns, S. P., Buzan, J., Clark, M., Craig, A., Dahlin, K., Drewniak, B., Fisher, J. B., Flanner, M., Fox, A. M., Gentine, P., Hoffman, F., Keppel-Aleks, G., Knox, R., Kumar, S., Lenaerts, J., Leung, L. R., Lipscomb, W. H., Lu, Y., Pandey, A., Pelletier, J. D., Perket, J., Randerson, J. T., Ricciuto, D. M., Sanderson, B. M., Slater, A., Subin, Z. M., Tang, J., Thomas, R. Q., Martin, M. V., and Zeng, X.: The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty, J. Adv. Model. Earth Syst., 11, 4245-4287, https://doi.org/10.1029/2018ms001583, 2019.

Liu, R., Su, J., Zheng, D., Lü, H., Zhu, Y., Pan, Y., Shi, C., Sun, S., and He, J.: Comprehensive assessment of various meteorological forcing datasets on the Tibetan Plateau: Insights from independent observations and multivariate comparisons, J. Hydrol., 656, 133025, https://doi.org/10.1016/j.jhydrol.2025.133025, 2025.

Niu, G.-Y. and Yang, Z.-L.: Effects of Frozen Soil on Snowmelt Runoff and Soil Water Storage at a Continental Scale, J. Hydrometeorol., 7, 937-952, https://doi.org/10.1175/JHM538.1, 2006.

Niu, G. Y., Yang, Z. L., Dickinson, R. E., and Gulden, L. E.: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models, J. Geophys. Res.-Atmos., 110, https://doi.org/10.1029/2005jd006111, 2005.

Schenk, H. J. and Jackson, R. B.: The Global Biogeography of Roots, Ecol. Monogr., 72, 311-328, https://doi.org/10.1890/0012-9615(2002)072[0311:TGBOR]2.0.CO;2, 2002.

van den Honert, T. H.: Water transport in plants as a catenary process, Discuss. Faraday Soc., 3, 146-153, https://doi.org/10.1039/DF9480300146, 1948.

Williams, G. P.: Friends don't let friends use Nash-Sutcliffe Efficiency (NSE) or KGE for hydrologic model accuracy evaluation: A rant with data and suggestions for better practice, Environ. Modell. Softw., 194, 106665, https://doi.org/10.1016/j.envsoft.2025.106665, 2025.

Yang, Y., Wang, K., Hu, Z., and Wang, G.: Meteorological dataset observed by Alpine Ecosystem Observation and Experiment Station of Gongga Mountain, during 1998–2018, China Scientific Data [dataset], https://doi.org/10.11922/csdata.2020.0004.zh (last access: 6 November 2026), 2020.

Zhu, S., Chen, H., Zhang, X., Wei, N., Shangguan, W., Yuan, H., Zhang, S., Wang, L., Zhou, L., and Dai, Y.: Incorporating root hydraulic redistribution and compensatory water uptake in the Common Land Model: Effects on site level and global land modeling, J. Geophys. Res.-Atmos., 122, 7308-7322, https://doi.org/10.1002/2016JD025744, 2017.