Comment:

Thanks for the revised paper and the efforts you have done to consider all remarks raised by the reviewers. However, the most important element that came forward was related to the weights assigned in the multi-objective optimization. The issue has been taken seriously by giving more explanations in the paper and also the report. However, these explanations are not convincing. Let me elaborate.

Weights are assigned in multi criteria analysis for various reasons. This can be done to give more weight on variables that are more important, that may have better or more data or so. These are 'subjective', defined by the modeler and are fine if well justified/described.

But weights may have a more fundamental role: to normalize against 'strengths' of the signals. Eg if very different signals are combined, they may have very different units or order of magnitudes. The weights should balance out these differences in units.

I do not see this taking place in your study. Or, in other words, the importance of the different variables in the multi-objective function depend on the units. If, for one reason, you would change the units of one or more variable, you will also obtain a different result (you will not get different weights, but the overall strength/importance of one variable in the full optimization may change).

I would accept the reasoning if the function would be the summation of wi*NRMSE but not for wi*RMSE, as this equation does not accounts for the 'units' or 'order of magnitude' of RMSE.

I see 2 solutions to get the paper published:

1) Proof that I am wrong in my statements (ie that your method is not sensitive to units/orders of magnitudes of the individual variables)

2) Adjust the weighting scheme

Response:

Thanks very much for pointing out a very important problem. After double checking the original code of weighting function, we found it's a mistake in equation 2. Actually NRMSE was used, not RMSE. Moreover, we compared the 'subjective' weighting function in the original draft, the average weighting in [Liu et al. 2005], and the Global Optimization Criterion (GOC) in [van Griensven and Meixner, 2007]. Consequently, this paragraph (line 3-8 in page 6725) was revised as follows.

In this study, we use three weighting functions to convert the multi-objective optimization into a single objective optimization. ***Case 1:*** *Assigning more weight if the output is simulated more poorly as compared to the other outputs. The summed up objectives should have the same unit, so we use NRMSE as the objective function. The weighting function is:*

$$F = \sum_{i=1}^{n} w_i NRMSE_i \quad (2)$$

in which the $NRMSE_i$ *is the Normalized Root Mean Squared Error of each output variable that defined in equation 1,* w_i *is the weight of each output, and* $\sum_{i=1}^{n} w_i = 1$. *Table 4 shows the RMSE and NRMSE of CoLM using default parameterization scheme, and the weight of each output is proportional to the NRMSE.* ***Case 2:*** *Liu et.al [2005] normalized the RMSE of each output with the RMSE of simulation result given by default parameters. The weighting function is:*

$$F = \sum_{i=1}^{n} w_i \frac{RMSE_i}{RMSE_{i,default}} \quad (3)$$

and assign equal weights to each normalized output. ***Case 3:*** *van Griensven and Meixner [2007] defined the Global Optimization Criterion (GOC) based on Bayesian theory for multi-objective optimization. If the number of observations of each output are the same, the GOC is defined as:*

$$F = \sum_{i=1}^{n} \frac{SE_i}{SE_{i,min}} \quad (3)$$

where $SE_i = \sum_{j=1}^{N} \left(y_{i,j}^{sim} - y_{i,j}^{obs} \right)^2$ *is the Squared Error, and* $SE_{i,min}$ *is the Squared Error of optimal solution.* $SE_{i,min}$ *is dynamically updated during the optimization procedure.*

The paragraph is added to page 6726.

We carried out multi-objective optimization with ASMO using weighting functions defined in equation 2, 3 and 4. The optimization results are shown in table 5. The RMSEs of each case were compared with that given by the default parameterization scheme, and the relative improvements were calculated. Obviously, for all the three cases, all of the six outputs were significantly improved except soil temperature. All the three cases sacrificed the performance of soil temperature, but case 2 ([Liu et.al., 2005]) decreased least (only 0.78%), case 3 ([van Griensven and Meixner, 2007]) decreased most, and the case 1 (weights proportional to NRMSE) is the moderate one. The results indicated that all the three kinds of weighting functions can balance the conflicting requirements of different objectives and effectively give an optimal parameter set with ASMO algorithm. In the following studies, we only involve the moderate case (case 1).

To keep consistency, we also updated other parts of the manuscript. Please also refer to the marked manuscript for more details.

Table 5: Inter-comparison of different weighting systems.

Flux name (Units)	default parameters	Case 1 $F=\sum_{i=1}^{n} w_i NRMSE_i$ $w_i \propto NRMSE_i$		Case 2 $F=\sum_{i=1}^{n} w_i \frac{RMSE_i}{RMSE_{i,default}}$ $w_i = 1/n$		Case3 $F=\sum_{i=1}^{n} \frac{SE_i}{SE_{i,min}}$	
	RMSE	RMSE	improvement	RMSE	improve	RMSE	improvement
Sensible heat (W/m^2)	49.1424	44.7400	8.96%	44.2571	9.94%	43.0176	12.46%
Latent heat (W/m^2)	43.5944	36.8158	15.55%	36.6070	16.03%	39.1792	10.13%
Upward longwave radiation (W/m^2)	19.4317	16.3837	15.69%	15.8426	18.47%	16.4160	15.52%
Net radiation (W/m^2)	42.7769	38.8834	9.10%	38.7710	9.36%	39.2156	8.33%
Soil temperature (K)	2.6584	2.9011	-9.13%	2.6793	-0.78%	3.0305	-13.99%
Soil moisture (kg/m^2)	21.1371	18.7408	11.34%	19.7590	6.52%	19.5655	7.44%