Reply to interactive comments on "Estimation of Hourly Land Surface Heat Fluxes over the Tibetan Plateau by the Combined Use of Geostationary and Polar Orbiting Satellites" by Lei Zhong et al.

Anonymous Reviewer #1

Observations of land surface heat fluxes over the QTP are essential for understanding the land-atmosphere interactions. However, limited by the small amount of land-atmosphere monitoring stations and sparse spatial coverage, it is difficult to quantify the responses of the land-atmosphere interactions under the condition of climate warming on the QTP. This study aims to provide a plateau-scale product with a notable advantage of hourly-resolution using the SEB model in conjunction with the observations from polar and geostationary satellites. As we know that the temporal resolution of land surface heat fluxes is highly dependent on the forcing in various modelling approaches. In general, temperature and wind speed are two key input variables for the latent heat flux and the turbulent flux, respectively. The input variables in this study use the hourly temperature observations and other observations with a three-hour resolution. As a result, the reliability of the turbulent flux might be problematic when using the energy balance equation for calculation, and its accuracy is even worse than the 3-hourly product using data assimilation approach (e.g., GLDAS). A rigorous analysis of the accuracy is required to consolidate the proposed method. Given the present analysis, the current conclusion of hourly-resolution is not convincing for me. Considering other issues, a substantial revision is needed for this manuscript.

Author Response: We would like to thank Reviewer #1 for the insightful and constructive comments. All your comments and suggestions are very helpful for improving our manuscript. We have carefully considered and addressed all of these comments, and significantly revised our manuscript. Please find our point-by-point response below.

Major issues:

(1) *Since forcing data is lack of homogeneity in temporal and spatial resolution, the authors should discuss their impacts on the accuracy of the product. The authors declaimed a spatial resolution of 5 km, but it has been changed to 10 km in the new version (no rational explanation in the text).I think the authors should cope with the similar problem*

for the temporal resolution. As mentioned above, the methodology needs a rigorous analysis of the accuracy of the estimated land surface fluxes. Besides, I did not find the description of how to use the 3 hour-forcing in the SEBS model to produce the hourly product.

Author Response: Thank you for the above comments. The lack of homogeneity in the temporal and spatial resolution, mainly exists in the meteorological forcing data because its spatial and temporal resolution are lower than those of other satellite derived inputs. In addition to some remarks about this issue in Section 5 (P11, L27-30; P12, L1-4) , we also performed some sensitivity tests to verify how the RMSEs of forcing data can affect the sensible heat flux and latent heat flux. As shown in Figure 4, three sites located in the northern, western and southeastern part of the TP were randomly selected to perform the sensitivity analysis. All input meteorological forcing parameters in Table 3 (R_{swd}, R_{lwd} , u, T_a, SH, P_s) are selected. The original sensible heat flux and latent heat flux from the SEBS model are used as reference values. The RMSEs of different forcing data were used as perturbations. As shown in Table 5, the sensible heat flux is highly sensitive to R_{swd}, u and T_a, while the latent heat flux is very sensitive to R_{swd}, R_{lwd} and T_a. Both sensible heat flux and latent heat flux are not sensitive to errors of SH and P_s. As the R_{swd} has a variation from –68.5 Wm^{-2} to 68.5 Wm^{-2}, the induced latent heat flux uncertainty ranges from -29.75 Wm^{-2} to 35.86 Wm^{-2}. Similarly, the sensible heat flux is very sensitive to T_a. When T_a has an uncertainty from -2.08 K to 2.08 K, the induced sensible heat flux uncertainty ranges from 14.64 Wm^{-2} to -16.94 Wm^{-2}. All the above works has been added to the revised manuscript. (P9, L6-16)

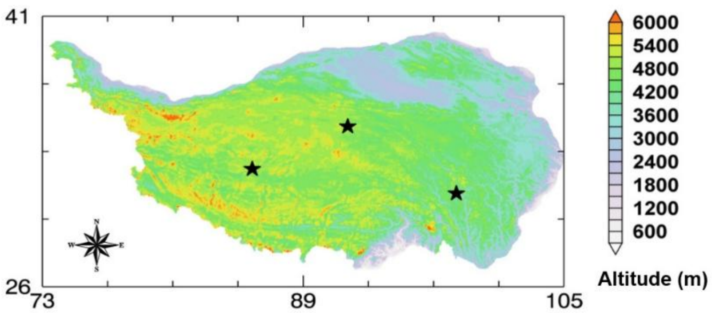

Figure 4: Locations of the three sites (marked by pentagrams) used to carry out sensitivity tests of the meteorological forcing input data. The legend of the color map indicates the elevation above mean sea level in meters.

Table 5: Uncertainties for each meteorological forcing variable and the induced changes for H_s and LE.

Variables	Assumed Uncertainty	Induced Uncertainty of H_s	Induced Uncertainty of LE
R_{swd} ($W \cdot m^{-2}$)	-68.50~68.5	-12.34~6.22 (-8.05%~4.06%)	-29.75~35.86 (-17.92%~21.60%)
R_{lwd} ($W \cdot m^{-2}$)	-20.98~20.98	-2.50~2.50 (-1.63%~1.63%)	-15.54~15.54 (-9.36%~9.36%)
u ($m \cdot s^{-1}$)	-1.71~1.71	-9.47~7.31 (-6.18%~4.77%)	9.47~-7.31 (5.71%~-4.41%)
T_a (K)	-2.08~2.08	14.64~-16.94 (9.55%~-11.05%)	-14.64~16.94 (-8.82%~10.20%)
SH ($kg \cdot kg^{-1}$)	-0.56×10^{-3}~0.56×10^{-3}	-0.01~0.01 (-0.01%~0.01%)	0.01~-0.01 (0.01%~-0.01%)
P_s (hPa)	-8.51~8.51	-0.01~0.01 (-0.01%~0.01%)	0.01~-0.01 (0.01%~-0.01%)

The spatial resolution of the final flux products should be determined by the lowest input of the source data, which was mentioned in the revised manuscript (P5, L8-9). Thus, the final surface heat flux product should be 10 km. We have corrected this mistake in the manuscript after the quick review and mentioned this issue in the response letter to the quick reviewer comments.

For the temporal resolution, a linear statistical downscaling method was used to derive the hourly meteorological forcing data based on the original 3-hour forcing data and in situ measurements in this study. The general idea is to establish an empirical relationship between each 3-hour in situ measurement. Then, this relationship is applied to meteorological forcing data (P5, L17-21). For example, T_{a00}, T_{a01} and T_{a03} represent the in situ air temperature measurements from six stations at 00h, 01h and 03h, respectively. Thus $T_{a00} = [a_1, a_2, a_3, a_4, a_5, a_6]$, $T_{a01} = [b_1, b_2, b_3, b_4, b_5, b_6]$, and $T_{a03} = [c_1, c_2, c_3, c_4, c_5, c_6]$. Then, the linear equation $T_{a01} = k_1 T_{a00} + k_2 T_{a03}$ can be solved. According to the meteorological forcing data at 00h and 03h, the plateau scale T_a at 01h can be achieved by the following formula.

$$\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = k_1 \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + k_2 \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix}$$

where a, b and c represent meteorological forcing data at 00h, 01h and 03h respectively; and m and n represent total rows and columns, respectively, of the grid data. The meteorological forcing data at other times can be achieved similarly determined.

(2) *The major supporting for the conclusion of a better performance of the proposed product than the GLDAS produce is based on the comparison with the observational data. The authors use the Bowen ratio calibration method to improve the observed data. We know the validity of the Bowen ratio method varies distinctly in different environments due to the different fulfillment of assumptions. As a result, certain biases will be brought into the observational data, and this can mislead the comparison. First, it is not clear in the text that if the comparison is under the same condition that the observational data are all corrected with the Bowen ratio method. Second, even if using the similar observational data for the comparison, the biases from the correction can still distort the RMSE. Hence, I would suggest directly using the observed data for comparison. Besides, since the data quality of eddy covariance measurements may vary at the 6 stations, comparison on the indicators like RMSE at each station separately may provide more information.*

Author Response: It should be noted that the in situ flux data have been flagged by steady state tests and developed conditions tests according to Foken and Wichura (1996) and Foken et al. (2004). Steady conditions mean that all statistical parameters do not vary with time. The flux-variance similarity was used to test the development of turbulent conditions. A data quality of only QA<5 was chosen to make the comparison. Therefore, the comparison is under similar conditions. The above information has been included in the text (P7, L27-28; P8, L1-2).

The Bowen ratio correction method was only used to correct the in situ latent heat flux measurements and was not used for the other three energy balance components (radiation heat flux, sensible heat flux and soil heat flux). The following equation was used to perform the Bowen ratio correction.

$$BRLE = \frac{1}{1+\beta}(R_n - G_0)$$

where $BRLE$ is the latent heat flux after correction and Bowen ratio $\beta = \frac{H}{LE}$. R_n and G_0 are net radiation flux and soil heat flux, respectively.

As you mentioned, the validity of the Bowen ratio method varies distinctly in different environments due to the different assumptions. We try to use the original latent heat flux measurements to perform the validation. The following figure shows the comparison between surface latent heat fluxes estimated by the SEBS model with in situ measurements. All corresponding values for the latent heat flux comparison have been replaced in Table 4 (P18). It can be seen that that indicators for latent heat flux have been changed but the they will not influence the general results of this paper. All information on the Bowen ratio correction has been deleted from the original manuscript.

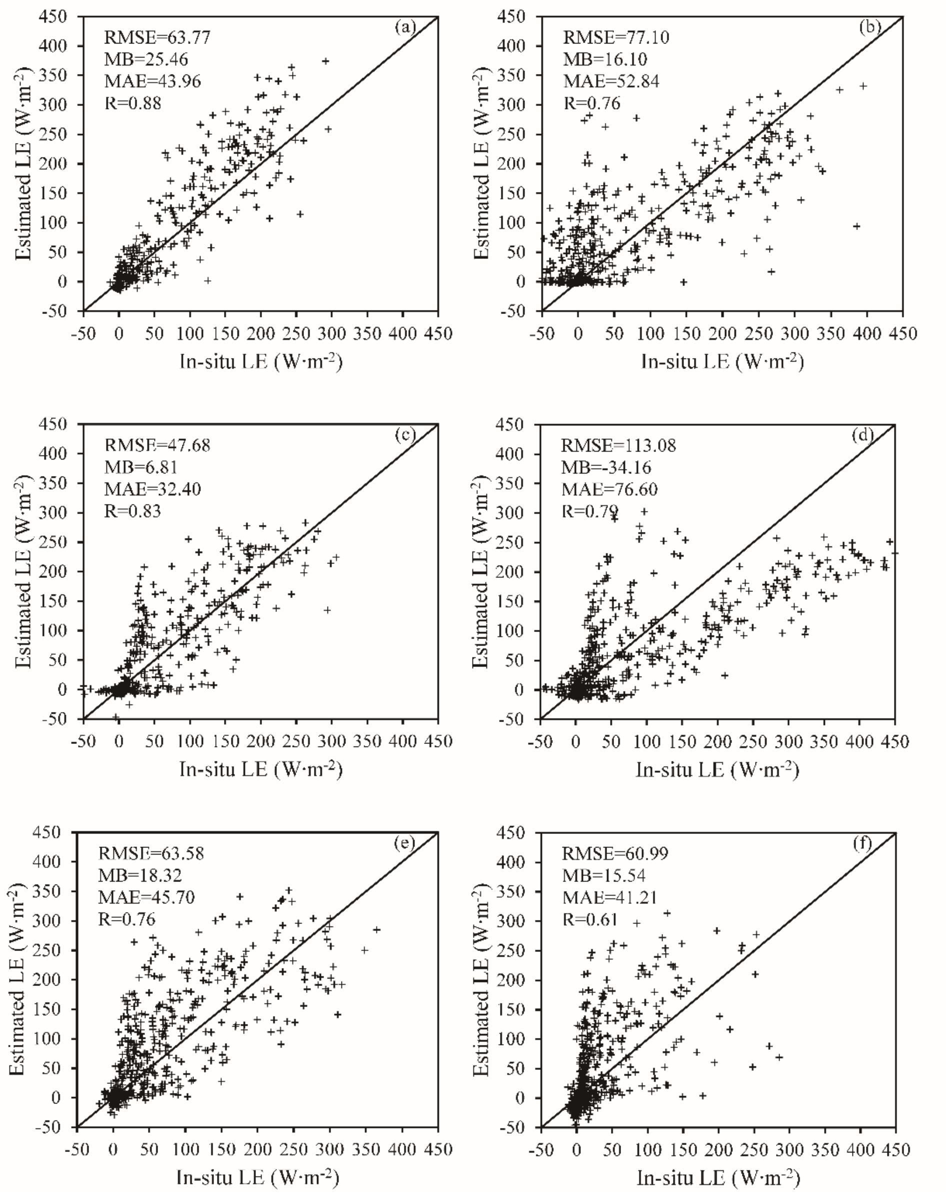

Figure: Validation of surface latent heat fluxes estimated by the SEBS model with in situ measurements (a. BJ station; b. D105 station; c. Linzhi station; d. MS3478 station; e. Nam Co station; f. QOMS station).

For your last question, since the data quality of eddy covariance measurements may vary at the 6 stations, a separate comparison of the indicators, such as RMSE, at each station may provide more information. For the data quality, QA<5 was chosen to ensure the flux measurements are under similar steady state and developed conditions; thus, it is not necessary to make a comparison at each station separately. There will be some differences among those stations, but most of these differences can be explained by the quality of the input forcing data, as shown in the sensitivity test.

(3) *The product provided by the authors is produced based on the input data with a spatial resolution no less than 10 km. The authors compare it with a product with a spatial resolution of 25 km. While the scale of the stations normally represents a scale of about less than 1 km. The authors should give some explanation about their comparability.*

Author Response: The scale problem you mentioned is an important and difficult problem to be solved in the quantitative remote sensing and atmospheric research field. First, the datasets generated by our methods need to be validated by comparison with the observation dataset. The eddy covariance system is widely accepted as a direct measurement of energy heat fluxes and has been used to validate satellite estimations (Fisher et al, 2008, Ma et al, 2006, Su 2002). Some errors can be caused due to scale mismatch among the stations between the SEBS product and GLDAS product. This issue has been discussed in Section 4.1 (P9, L16-19) and Section 5. Because of the relatively homogeneous land surface conditions of the field stations, this effect should have been minimized in our study. Scintillometry is possibly the most convenient method to measure fluxes at a 1-10 km scale. Unfortunately, this device is lacking over the TP. If we have enough in situ measurements within a grid scale of 10 km or 25 km, an average or weighted average of the measurements can be directly used to reduce some uncertainties caused by scale mismatch. However, for well-known reasons, it is very difficult to carry out such measurements in the TP with the harsh environment and climate conditions. The above discussions have been added to the revised manuscript. (P11, L27-30; P12, L1-4)

Minor issues:

(1) *P5-line 13-25: the authors validate the forcing data and find the notable variance. These differences can further propagate to the product. Please discuss its relation to the final product.*

Author Response: Yes, we totally agree with you. According to your suggestion, a sensitivity test was carried out to test how the RMSE of forcing data can affect the sensible heat flux and latent heat flux. As shown in Figure 4, three sites located in the northern, western and southeastern part of the TP were randomly selected to perform the sensitivity analysis. All input meteorological forcing parameters in Table 3 (R_{swd}, R_{lwd}, u, T_a, SH, P_s) are selected. The original sensible heat flux and latent heat flux from the SEBS model are used as reference values. The RMSEs of different forcing data were used as perturbations. As shown in Table 5, the sensible heat flux is highly sensitive to R_{swd}, u and T_a, while the latent heat flux is very sensitive to R_{swd}, R_{lwd} and T_a. Both sensible heat flux and latent heat flux are not sensitive to errors of SH and P_s. As R_{swd} varies from –68.5 Wm^{-2}to 68.5 Wm^{-2}, the induced latent heat flux uncertainty ranges from -29.75 Wm^{-2} to 35.86 Wm^{-2}. Similarly, the sensible heat flux is very sensitive to T_a. When $T_a$has an uncertainty from -2.08 K to 2.08 K, the induced sensible heat flux uncertainty ranges from 14.64 Wm^{-2} to -16.94 Wm^{-2}. All the above information has been added to the revised manuscript. (P9, L6-16)

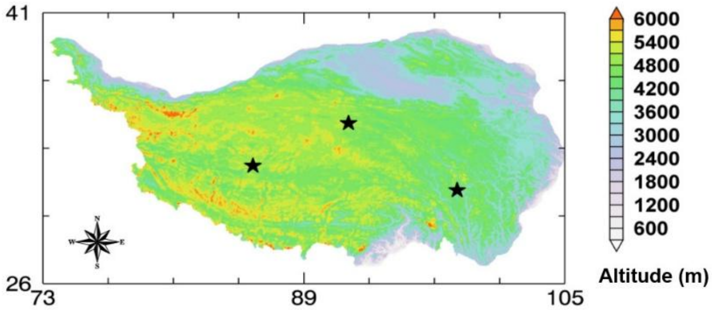

Figure 4: Locations of the three sites (marked by pentagrams) used to carry out sensitivity tests of the meteorological forcing input data. The legend of the color map indicates the elevation above mean sea level in meters.

Table 5: Uncertainties for each meteorological forcing variable and the induced changes in H_s and LE.

Variables	**Assumed Uncertainty**	**Induced Uncertainty of H_s**	**Induced Uncertainty of LE**
R_{swd} (W·m^{-2})	-68.50~68.5	-12.34~6.22 (-8.05%~4.06%)	-29.75~35.86 (-17.92%~21.60%)
R_{lwd} (W·m^{-2})	-20.98~20.98	-2.50~2.50 (-1.63%~1.63%)	-15.54~15.54 (-9.36%~9.36%)
u (m · s^{-1})	-1.71~1.71	-9.47~7.31 (-6.18%~4.77%)	9.47~-7.31 (5.71%~-4.41%)
T_a (K)	-2.08~2.08	14.64~-16.94 (9.55%~-11.05%)	-14.64~16.94 (-8.82%~10.20%)

SH (kg·kg^{-1})	-0.56×10^{-3}~0.56×10^{-3}	-0.01~0.01 (-0.01%~0.01%)	0.01~-0.01 (0.01%~-0.01%)
P_s (hPa)	-8.51~8.51	-0.01~0.01 (-0.01%~0.01%)	0.01~-0.01 (0.01%~-0.01%)

(2) P8-line 2-6: The introduction of the GLDAS dataset should not belong to Result. The authors should introduce it more in light of its importance for comparison.

Author Response: The introduction of the GLDAS dataset has been moved to section 2 (P5, L26-30). We also introduce the importance of a comparison with the GLDAS product as follows (P5, L22-26).

The high-quality, global land surface fields provided by GLDAS support weather and climate prediction, water resources applications, and water cycle investigations. Since the GLDAS data have been widely used, it is meaningful to compare our satellite estimation with these high-quality data to further prove the accuracy of our estimations.

(3) P8-line 5: what high accuracy?

Author Response: This term has been replaced with 'high quality'. (P11, L18)

(4) P9-line 15-27: the authors describe the feature of diurnal variation of hourly flux map. Are there any special in comparison on our general understanding?

Author Response: The qualitative description of diurnal variation in the hourly flux map is aligned with general knowledge. This alignment can further prove the effectiveness of our estimation method and validate the final estimation results. In the revised manuscript, we also added some quantitative expressions to improve the content. (P10)

(5) Table 4: add values of the same indicators for all sites.

Author Response: For the data quality, QA<5 was chosen to ensure the flux measurements are under similar steady state and developed conditions; thus, it is not necessary to make a comparison at each station separately. Furthermore, if we list the same indicators for all sites, two additional pages will be needed to show this content. Adding this information may make the text difficult to read. Therefore, we would like to keep the original Table 4, but the indicators for latent heat flux have been replaced by the values before the Bowen ratio correction, as you suggested.

(6) Figure 1: the caption is too brief. The same problems for other plots. What is the right plot?

Author Response: Thank you for this comment. The right plot illustrates the location of the Tibetan Plateau in China. We have improved all figure captions in the revised manuscript. (P19-24)

(7) Figure 4: the scale of the axis is misleading. Besides, how do you choose the representative days for each month? Choose the nice one? Please describe what they are in subpanels.

Author Response: Figure 4 (now Fig. 5) has been redrawn to improve its clarity. We also added additional explanations in the figure caption to prevent ambiguity. We did not select the nice days. Instead, the monthly mean value is shown in Figure 5. The subpanels are described in the caption of the new Figure 5. (P23)