Reply to Reviewer 2

General Comments:

He et al. constructed daily and 500m ET and GPP datasets in China using PML-V2. Compared with previous products, this model outputs improved in several aspects, including 26 EC sites being used for model calibration and validation, country-specific meteorological forcing, daily data, and intra-annual dynamics for multiple ecosystems. This ambitious work provides valuable data products for assessing the carbon and water cycles in China. They may also provide guidance in agricultural production and ecosystem management. The authors may consider the following suggestions to improve the robustness of this manuscript.

Response:

Thank you for appreciating our work and considering that the products are very valuable. We have carefully checked and re-edited the original manuscript. In the following, we reply to all comments in a point-by-point response. All comments are shown in blue. Sentences from the manuscript are in italics and the revised contents are indicated in red.

Specific Comments:

1. Line 99, the whole name for CMFD should be provided when it is first mentioned in the text.

Response:

We have added the whole name - the China Meteorological Forcing Dataset for the CMFD dataset.

2. Line 144-146, this sentence is not appropriate. You may use the MODIS land cover product, but it is debatable if it has the highest accuracy in China since there are many recently released land use/covered datasets with a high spatial resolution (30m and 10m). Many MODIS products based on the MODIS land use dataset may have low credibility in regions with complex terrain such as in the Loess Plateau.

Response:

We revised the sentence as follows:

Here we used the International Geosphere-Biosphere Program (IGBP) layer of MCD12Q1.006 land cover product (Sulla-Menashe et al., 2019) during 2000-2020 since IGBP classification is annually continuous and has acceptable accuracy in China when compared with other land cover products (Feng and Bai, 2019).

3. Did you test the continuity between GLDAS-2.1 and CMFD?

Response:

Currently, we compared the magnitude and variability of the products using different meteorological forcing inputs, i.e., PML-V2(China)$_{GLDAS\text{-}2.1}$ and PML-V2(China)$_{CMFD}$, at the grid and national scale in section 4.4.2, as follows:

To extend the simulation period, we used GLDAS-2.1 meteorological forcing data during 2019-2020 since the CMFD dataset is only up to 2018. To check if using these two datasets *generates a systematic bias, we reran the PML-V2(China) in 2001-2018 using GLDAS-2.1 and compared the modelling results with those obtained using CMFD (Fig. S1). At the national scale, the mean difference, calculated by* $(PML\text{-}V2(China)_{GLDAS\text{-}2.1} - PML\text{-}V2(China)_{CMFD})/PML\text{-}V2(China)_{CMFD}$, *varied from -1.22% to 1.62% among* E_i, E_c, *and GPP, and was 13.72% for* E_s *and 7.78% for ET. The difference is within -25% ~ 25% in more than 66% of the research region for all five variables (Fig. S1b2-e2), specifically 100% for GPP, 95% for* E_c, *84% for ET, 73% for* E_i, *and 66% for* E_s *(Fig. S1b3-e3). This illustrates that PML-V2(China) using the GLDAS-2.1 in 2019-2020 does not generate a noticeable systematic deviation.*

The PML-V2(China) product of 2019-2020 is the interim data as the supplement of PML-V2(China) after 2018. We suggest that users do spatial variability analysis instead of trend analysis if they want to use the PML-V2(China) of 2019-2020. With the release of the meteorological dataset, we will continue to update the PML-V2(China) using the CMFD inputs. Moreover, we have removed the description and figures about the trend analysis of PML-V2(China) for 2019-2020 in the manuscript.

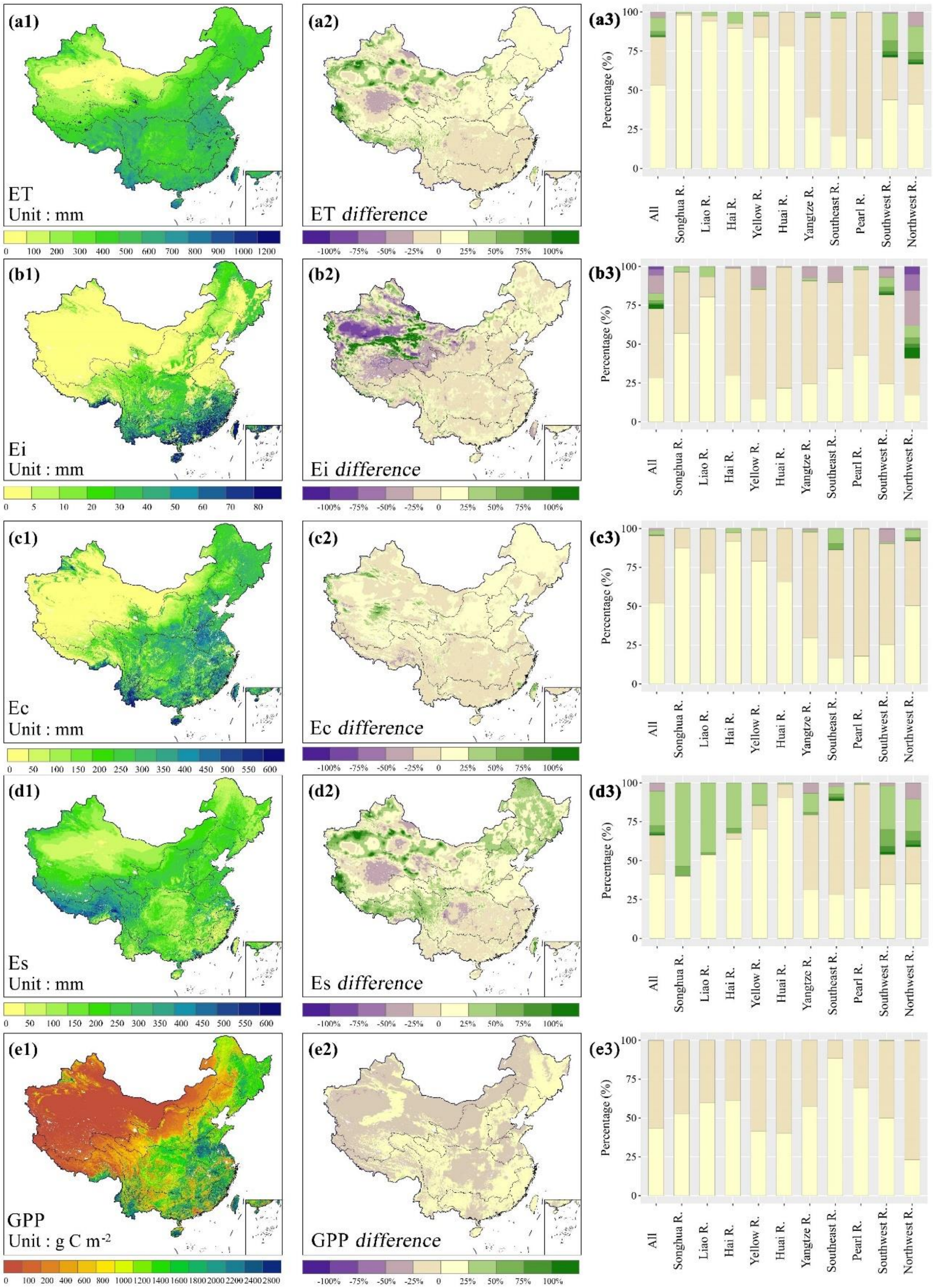

Figure S1: The modelling results using GLDAS-2.1 meteorological forcing data during 2001-2018 and comparison with the PML-V2(China) product using CMFD: (a1-e1) Spatial distribution of the 18-year mean of five variables; (a2-e2) Spatial distribution of the difference using two forcing datasets, calculated by $(PML\text{-}V2(China)_{GLDAS\text{-}2.1} - PML\text{-}V2(China)_{CMFD})/PML\text{-}V2(China)_{CMFD}$***; and (a3-e3) Proportion of difference in each river basin. 'ALL' represents the whole study area. The legends for (a3-e3) are the same as***

that for (a2-e2). Taking Fig.(a3) as an example, the area percentage of ET difference in 0 ~ 25% in the Songhua River Basin is about 99%.

References

He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., and Li, X.: The first high-resolution meteorological forcing dataset for land process studies over China, Sci Data, 7, 25, https://doi.org/10.1038/s41597-020-0369-y, 2020.

4. Section 2.6, the simulated model outputs were validated at the EC site-level, and compared with the publicly available dataset. How did you get the parameter set for a certain land use type? Did all the land use types have a unique parameter set? Did you run the model at each site?

Response:

(1) For each land use type, we used a global optimization method - genetic algorithm to gain the optimal solution by setting population size 1000 and number of generations 50 by minimizing an objective function including ET_{obs} and GPP_{obs}. (2) Each of the nine land use types has a unique parameter set, so they are nine parameter sets. (3) Yes, we run the model at each site. We revise sections 2.5 and 2.6 to make the model calibration and model validation parts clearer, as follows:

2.5 Model calibration and model validation

The 11 parameters of the PML-V2 model for each PFT were calibrated and cross-validated against 26 EC sites by a global optimization method - genetic algorithm (GA). The GA generates a randomly initialized population and then evaluates the fitness of solutions according to its objective function. As generations iterate, the population includes more appropriate solutions, and eventually, it will converge (Holland, 1992; Konak et al., 2006). Specifically, we applied the GA algorithm with population size 1000 and number of generations 50. All EC-observed ET and GPP data within a PFT are used to minimize the following objective function (F_{opt}):

$$F_{opt} = 2 - NSE_{ET} - NSE_{GPP} = \frac{\sum_{i=1}^{N}(ET_{est}-ET_{obs})^2}{\sum_{i=1}^{N}\left(ET_{obs}-\overline{ET_{obs}}\right)^2} + \frac{\sum_{i=1}^{N}(GPP_{est}-GPP_{obs})^2}{\sum_{i=1}^{N}\left(GPP_{obs}-\overline{GPP_{obs}}\right)^2} \quad (8)$$

where NSE_{ET} and NSE_{GPP} are the Nash-Sutcliffe Efficiency of the daily ET and the daily GPP, respectively. The subscripts est and obs stand for the estimated and the observed, respectively. In this way, each of the nine PFTs gained a unique set with 11 calibrated parameter values, illustrated in Table S1

The 'leave-one-out' cross-validation method was utilized to evaluate the robustness of the PML-V2 model (Zhang et al., 2019). For each PFT, the data from one "ungauged" observation was excluded from the optimization while the data from all other observations at the same PFT were used for model calibration to obtain the simulated at the "ungauged" position. All nine PFTs were actualized in this way. Note that the PFT including EBF, MF, OSH, and SAV only has one ground site (Table 2). Therefore, it is appropriate to divide the data in each of the four sites into two sub-groups for cross-validation. The CF-CBF and the CF-HBG_S01 covering from 2003 to 2010, were divided into two sub-groups, each of which had 4 years: 2003-2006 and 2007-2010. While both the BNXJL and YJGRHG only covered *one year and were divided into two sub-groups by a two-day time step, separately. After that, the daily estimates in the cross-validation mode were against the daily observation from the 26 stations to explore the model transferability from known observations to any location.*

2.6 Model performance metrics

We assessed the performance of calibration and cross-validation of PML-V2 (and other seven mainstream ET and GPP products) against the observed sites or water-balance basins utilizing the following four metrics:

$$NSE_X = 1 - \frac{\sum_{i=1}^{N}(X_{est}-X_{obs})^2}{\sum_{i=1}^{N}\left(X_{obs}-\overline{X_{obs}}\right)^2}, \quad (9)$$

$$R_X = \frac{\sum_{i=1}^{N}\left(X_{est}-\overline{X_{est}}\right)\left(X_{obs}-\overline{X_{obs}}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_{est}-\overline{X_{est}}\right)^2 \times \sum_{i=1}^{N}\left(X_{obs}-\overline{X_{obs}}\right)^2}}, \quad (10)$$

$$RMSE_X = \sqrt{\frac{\sum_{i=1}^{N}(X_{est}-X_{obs})^2}{N}}, \quad (11)$$

$$Bias_X = \frac{\sum_{i=1}^{N}(X_{est}-X_{obs})}{N \times \overline{X_{obs}}}, \quad (12)$$

where NSE, R, $RMSE$, and $Bias$ are the Nash-Sutcliffe Efficiency, the correlation coefficient, the Root Mean Square Error, and the ratio of the difference between the estimated and the observed to the observed average. The subscript X represents ET or GPP; the subscripts est and obs stand for the estimated and the observed, respectively.

References here are the same as those in the manuscript.