RESPONSES TO COMMENTS FROM REVIEWERS #1

This manuscript generated a global 5km monthly potential evapotranspiration dataset (1982–2015) estimated by comprehensively considering interspecific differences in various vegetation-related parameters (e.g., plant stomatal resistance and CO_2 effects on stomatal resistance) to calibrate and parametrize the Shuttleworth-Wallace (SW) model. It can be accepted after some revisions. The authors should carefully consider my comments and suggestions below.

Responses: We thank this reviewer very much for the positive comments and the valuable suggestions, which are believed to be very useful for us to improving the study. Seriously according to these suggestions, we have revised this manuscript, and the detailed information could be found below and the revised version.

Comment 1: Lines 72. Maybe it is necessary to explain more clearly what is vegetation not "closed?

Response: Thanks for your suggestion. The concept of "closed" vegetation emphasizes the important role that vegetation plays in the vegetation-atmosphere interaction (Shuttleworth and Wallace, 1985). Soil evaporation is often overlooked due to its small proportion, such as the PM model treating crop canopy as a single uniform cover (the "big leaf" model), but neglecting soil surface evaporation (Zhou et al, 2006). The "closed" vegetation is mostly distributed in areas with abundant vegetation cover. In areas with sparse or no vegetation, soil evaporation is significant and cannot be ignored, so it's not a "closed" vegetation. Therefore, the relevant explanation has been added in the revised manuscript, i.e., "*open canopy that light can penetrate to the ground*". **(L73)**

References:

Shuttleworth, W. J. and Wallace, J. S.: Evaporation from sparse crops—an energy combination theory. Quarterly Journal of the Royal Meteorological Society, 111, 839–855, 1985.

Zhou, M. C., Ishidaira, H., Hapuarachchi, H. P., Magome, J., Kiem, A. S. and Takeuchi, K.: Estimating potential evapotranspiration using Shuttleworth-Wallace model and NOAA-AVHRR NDVI data to feed a distributed hydrological model over the Mekong River basin. Journal of Hydrology, 327(1–2), 151–173, 2006.

Comment 2: Lines 74. "Heterogeneous" should be "homogeneous"?

Response: Thanks for pointing out the mistake. We have modified the word of "heterogeneous" as “homogeneous”.

Comment 3: Line 97-98. “vegetation height” and “Leaf Area Index” should be “vegetation height datasets” and “Leaf Area Index datasets”, respectively?

Response: We have modified the dataset names of “vegetation height” and “Leaf Area Index” as “*vegetation height datasets*” and “*Leaf Area Index datasets*”, respectively.

Comment 4: Figure 1. I suggest enlarging the size of the font in the legend and adjusting the position of the legend.

Response: Thanks. We have revised Figure 1 in this revision, and please see below (Figure R1) or the revised manuscript.

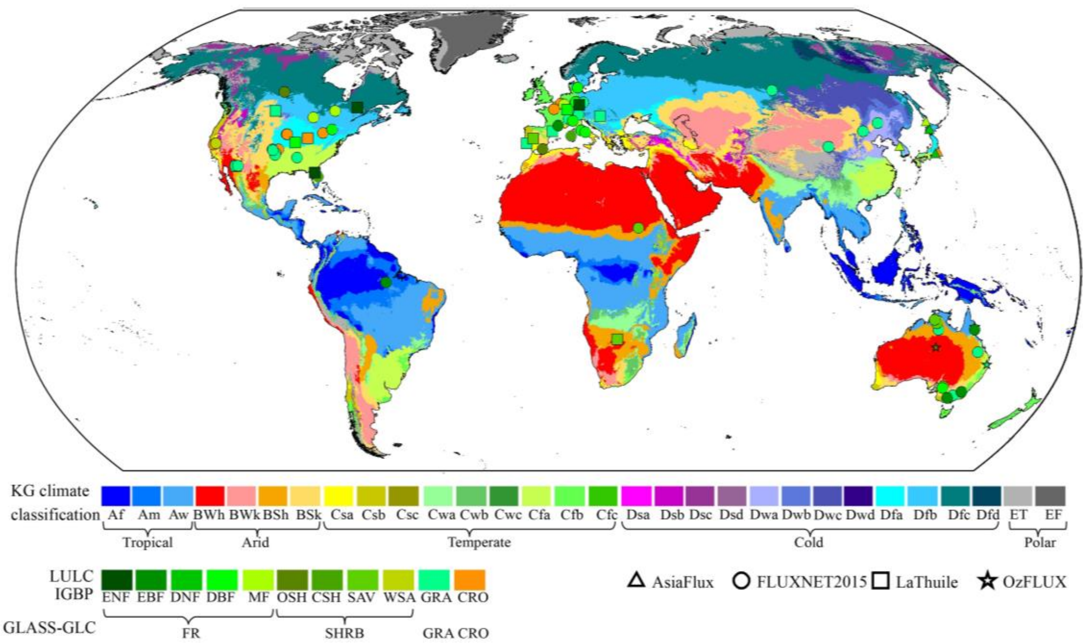

Figure R1: Locations of the used EC sites in this study over Köppen-Geiger (KG) climate regions (Beck et al., 2018). International Geosphere-Biosphere Programme (IGBP) classification system: CRO—cropland; GRA—grasslands; DBF—deciduous broadleaf forest; EBF—evergreen broadleaf forest; ENF—evergreen needleleaf forest; MF—mixed forest; CSH—closed shrubland; WSA—woody savannah; SAV—savannah; OSH–open shrubland. GLASS-GLC classification system: FR—ENF, EBF, DNF, DBF and MF; SHRB—SAV CSH, OSH and WSA; GRA; CRO.

Comment 5: Line 144-145. How did the authors consider the uncertainty from gap-filled heat fluxes? As far

as I know, the MDS method was applied to fill the gap of ET measurements in the FLUXNET2015 datasets. Recently, Jiang et al. (2022) have developed a novel physical full-factorial scheme for gap-filling of EC-measured ET. Their scheme has been demonstrated to outperform the MDS and MDV gap-filling methods. The authors should introduce this new advance in the gap-filling of ET and at least discuss the possible uncertainty caused by using the MDS gap-filled ET.

Response: We thank this reviewer for the valuable suggestions. We have checked the EC data with no water stress, and found that the gap-filling data points only accounted for less than 30% of the EC data used in this study. To answer this question, we only remained the data points without gap-filling to recalibrate and revalidate the established SW model (Figures R2 and R3). Overall, the new final r_{smin} values have no evident changes relative to our original values (Figure R2). In addition, seen from the validation results, we also found that the new results became worse but much limited compared to our original results (Figure R2). These suggested that the gap-filling data points may have limited impacts on our results.

Notably, we have contacted the authors of the novel physical full-factorial scheme, but it is unfortunate that we have not obtained the corresponding data for comparison before the deadline of this response. Moreover, it may be difficult for us to fill the data gaps using this novel physical full-factorial scheme, mainly because the $z_{0\mathrm{m}}$ (which is a critical parameter for conduct the physical full-factorial scheme of Jiang et al. (2022)) was needed for this scheme but was also unavailable. Anyway, the gap-filling data may introduce more or less into our results, and therefore the related discussions have been added to the revised manuscript, such as "*In this study, for maximizing the use of data, the marginal distribution sampling method was employed to fill the gaps in the EC LE measurements. However, it should be noted that even though the controlling thermodynamic and kinetic factors of the atmosphere are similar between the missing and retrieved moments (Jiang et al., 2022), the gap-filled LE may be of low confidence, especially when soil moisture has abrupt changes. To quantify potential impacts of the gap-filled values, the SW model was re-calibrated and re-validated against the data points without gap-filling. Relative to the SW model used in this study, the new r_{smin} and the validation metrics changed insignificantly (not shown here), suggesting that the uncertainties induced by the gap-filled LE were limited.*" **(L564-570)**

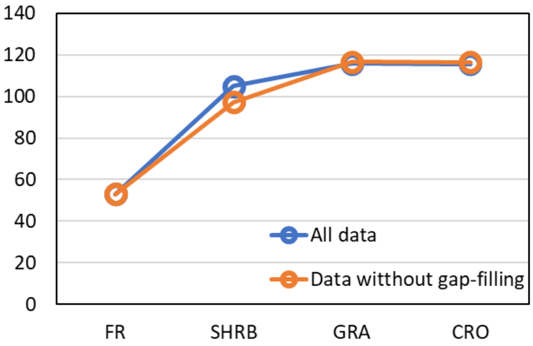

Figure R2: The final r_{smin} values for FR, SHRB, GRA and CRO.

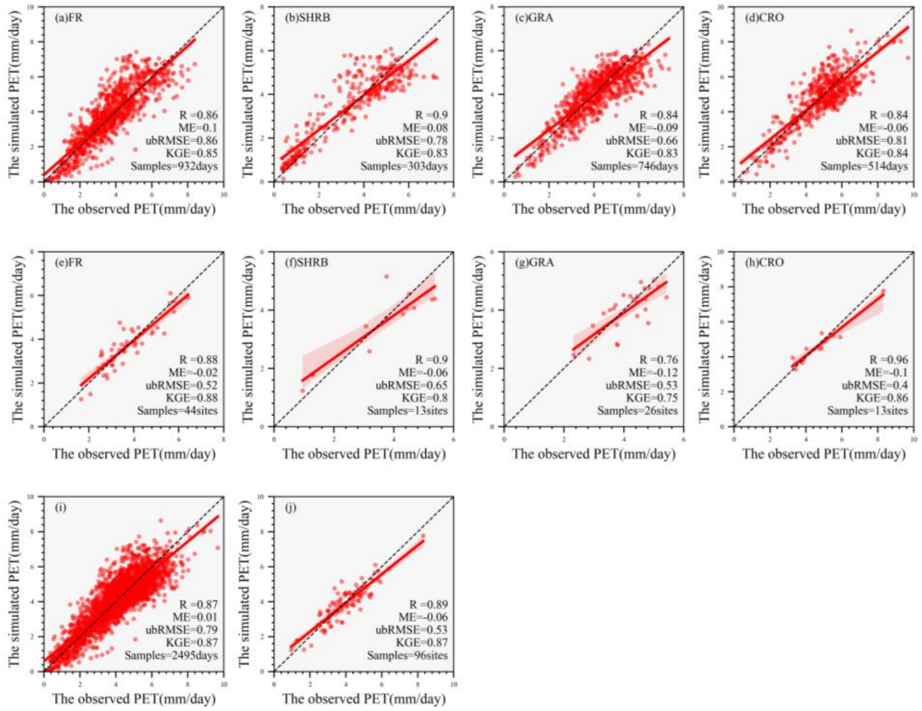

Figure R3: Validation results for the SW model calibrated using the data points without gap-filling. (a-d) Scatter-plots of daily observations against simulations for each LULC type. (e-h) Scatter-plots of site mean observations against simulations for each LULC type. (i) Scatter-plot of daily observations against simulations for all of 96 EC sites. (j) Scatter-plot of site mean observations against simulations for all of 96 EC sites.

References:

Jiang, Y., Tang, R. and Li, Z. L.: A physical full-factorial scheme for gap-filling of eddy covariance measurements of daytime evapotranspiration. Agricultural and Forest Meteorology, 323, 109087, 2022.

Comment 6: Line 149-154. How did the authors consider the effect of nighttime ET on daily ET? The

authors seemed to take the daytime ET as a surrogate of daily ET.

Response: Thank you very much for this comment. Yes, this study did not consider the nighttime ET (ET_n). Our considerations are two-fold. First, we have to admitted that ET_n does exist. For example, Padrón et al. (2020) showed that the proportion of ET_n/ET is approximately 6.3% based on the FLUXNET2015 dataset, and 7.9% based on the multiple global model simulations. Relatively, the proportion is small and may exert limited impacts on our PET estimates. Second, due to impacts of many biological and abiotic factors (Han et al., 2021), the ET_n process is very complex, and the related controlling mechanisms are still not clear to date. For example, Novick et al (2009) and Groh et al (2019) found that vapour pressure deficit (VPD) and wind speed have a significant impact on ET_n, while Groh et al. (2019) state that the contributions of night dew to ET_n cannot be ignored. As an important component of ET_n, the nighttime transpiration (Tr_n) is not only related to the incomplete stomatal closure (Dawson et al., 2007; Duursma et al., 2019) but also the circular regulation of nighttime water uses by plants (De Dios et al, 2015). To date, the factors influencing Tr_n are still disputed. Dawson et al. (2007) and Moore et al. (2008) found a positive correlation between Tr_n and VPD and soil moisture content (SWC), while Barbour and Buckley (2007), Phillips et al. (2010) and De Dios et al. (2015) pointed out that no or negative correlation existed between Tr_n and the two variables aforementioned. Meanwhile, the biological factors (e.g., plant species and ecosystem types) can also significantly influence Tr_n (O'keefe and Nippert, 2018; Zeppel et al., 2014). Just due to the complex mechanisms, to establish a common model for estimating ET_n across various ecosystems is still a challenging to date, and may lie beyond the scope of our study.

Anyway, neglecting ET_n process may result in the estimated PET to be smaller than the truth values, and therefore, we have added the related discussions in the revised manuscript, such as "*Padrón et al. (2020) showed that on average the fraction of nighttime ET accounts for approximately 6.3% of the total ET informed by the FLUXNET2015 dataset while 7.9% based on multiple global models. Notably, this fraction may exceed to 15% in mountain forest with snowy and windy winter. Ignoring vegetation canopy interception and nighttime ET may underestimate PET (Tourula and Heikinheimo, 1998; Lawrence et al., 2007; Mu et al., 2011; Padrón et al., 2020; Singer et al.; 2021). Subsequent research will be done to integrate these two processes in the SW model to further enhance the model's physical mechanism.*" **(L540-545)**

References:

Barbour, M. M., and Buckley, T. N.: The stomatal response to evaporative demand persists at night in Ricinus communis plants with high nocturnal conductance. Plant, Cell and Environment, 30(6), 711–721, 2007.

Dawson, T. E., Burgess, S. S., Tu, K. P., Oliveira, R. S., Santiago, L. S., Fisher, J. B., Simonin, K. A. and Ambrose A. R.: Nighttime transpiration in woody plants from contrasting ecosystems. Tree Physiology, 27, 561–575, 2007.

De Dios, V. R., Roy, J., Ferrio, J. P., Alday, J. G., Landais, D., Milcu, A. and Gessler, A.: Processes driving nocturnal transpiration and implications for estimating land evapotranspiration. Scientific Reports, 5(1), 10975, https://doi.org/10.1038/srep10975, 2015.

Duursma, R. A., Blackman, C. J., Lopéz, R., Martin-StPaul, K., Cochard, H. and Medlyn, B. E.: On the minimum leaf conductance: Its role in models of plant water use, and ecological and environmental controls. New Phytologist, 221(2), 693–705, 2019.

Groh, J., Pütz, T., Gerke, H., Vanderborght, J. and Vereecken, H.: Quantification and prediction of nighttime evapotranspiration for two distinct grassland ecosystems. Water Resources Research, 55(4), 2961–2975, 2019.

Han, Q, Wang, T., Wang, L., Smettem, K., Mai, M. and Chen X.: Comparison of nighttime with daytime evapotranspiration responses to environmental controls across temporal scales along a climate gradient. Water Resources Research, 57(7). https://doi.org/10.1029/2021WR029638, 2021.

Moore, G. W., Cleverly, J. and Owens, M. K.: Nocturnal transpiration in riparian Tamarix thickets authenticated by sap flux, eddy covariance and leaf gas exchange measurements. Tree Physiology, 28(4), 521–528, 2008.

Novick, K. A., Oren, R., Stoy, P. C., Siqueira, M. and Katul, G. G.: Nocturnal evapotranspiration in eddy-covariance records from three co-located ecosystems in the Southeastern US: Implications for annual fluxes. Agricultural and Forest Meteorology, 149(9), 1491–1504, 2009.

O'keefe, K. and Nippert, J. B.: Drivers of nocturnal water flux in a tallgrass prairie. Functional Ecology, 32(5), 1155–1167, 2018.

Padrón R. S., Gudmundsson, L., Michel, D. and Seneviratne, S. I.: Terrestrial water loss at night: Global relevance from observations and climate models. Hydrology and Earth System Sciences, 24(2), 793–807, 2020.

Phillips, N. G., Lewis, J. D., Logan, B. A. and Tissue, D. T.: Inter- and intra-specific variation in nocturnal water transport in Eucalyptus. Tree Physiology, 30(5), 586–596, 2010.

Zeppel, M. J. B., Lewis, J. D., Phillips, N. G. and Tissue, D. T.: Consequences of nocturnal water loss: A synthesis of regulating factors and implications for capacitance, embolism and use in models. Tree Physiology, 34(10), 1047–1055, 2014.

Comment 7: Line 163-164. In arid areas, a day may not be identified to have no water limits even when its corresponding EF exceeded the 95th percentile EF threshold.

Response: Thanks for your comment. Based on Köppen-Geiger (KG) climate regions, we selected 8 sites in arid areas, including one FR (i.e., AU-Lox), four GRA (i.e., AU-Emr, AU-Stp, US-SRG, and US-Wkg) and three SHRB (ES-Amo, AU-TTE, and SD-Dem) sites. For knowing impacts of the EF threshold on the SW model, we re-calibrated r_{smin} at each of the selected 8 sites with the 96th and 98th percentile EF thresholds, and then compared these new r_{smin} values with our used values in the manuscript (Figure R4). Overall, the r_{smin} values were around the 1:1 line with a regression coefficient above 0.96, implying slight changes r_{smin} when the threshold changed from 95th to 96th or 98th. Figure R5 shows the changes (i.e., Δx = 95th EF – 96th (98th) EF); x represents the validation metric) in validation metrics of the SW model calibrated with the 96th and 98th percentile EF thresholds relative to the original values. For most of the selected 8 sites, the ΔR, ΔME, $\Delta ubRMSE$ and ΔKGE values are between -0.1 and 0.1, between -0.02 mm/day and 0.04 mm/day, between -0.05 mm/day and 0.1 mm/day, and between -0.05 and 0.05. This suggested that the changes in the EF threshold may have limited impacts on the performance of the SW model at site scale. Additionally, using the 98th percentile EF threshold for theses 8 sites and the 95th percentile EF threshold for other 88 sites, we re-estimated the new optimal r_{smin} for each LULC and show the scatter-plots in Figure R6. Comparing Figures 4 and 5 in the manuscript vs. Figure R6, the validation metrics for each LULC have no evident changes when the EF threshold changed from 95th to 98th.

To sum up, we think that the 95th EF threshold could be used to define the condition with no water limits in arid areas, based on limited impacts of the EF threshold changes on the r_{smin} values and performance of the SW model. However, we have to admit that the possible cases may existed in arid areas, i.e., a day may not be identified to have no water limits even when its corresponding EF exceeded the 95th percentile EF threshold. Therefore, for more serious, the related discussion has been added in the revised manuscript, such as "*Second, due to usual water deficits in arid regions, the EF threshold may exceed to the 95th percentile of EF. Thus, the identified unstressed days based on this criterion may actually include the stressed days. This potentially biased the PET estimates in arid regions. Through increasing the EF threshold (e.g., 96th and 98th; not shown here) for several EC sites in arid regions, we found that selecting 95th percentile as threshold has limited impact on the PET estimates.*" **(L555-564)**

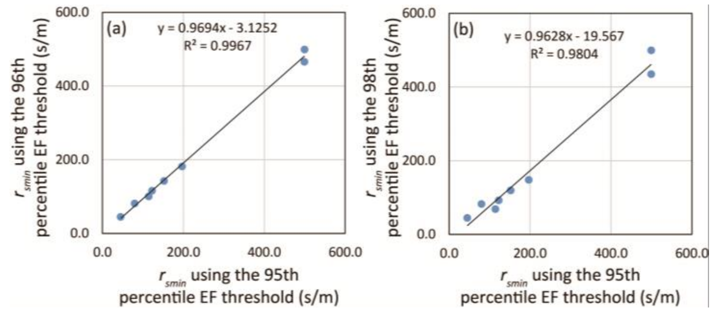

Figure R4: Scatter-plots of the recalibrated r_{smin} against with those used in the manuscript for the selected 8 EC sites in the KG arid areas

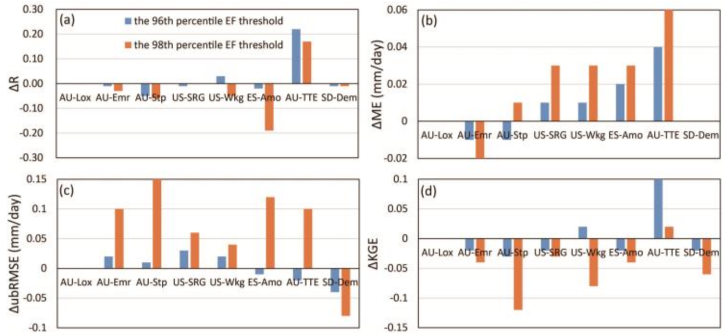

Figure R5: Changes in the validation metrics of the SW model calibrated with the 96th and 98th percentile EF thresholds for 8 EC sites in the KG arid areas, relative to those of the SW model used in the manuscript.

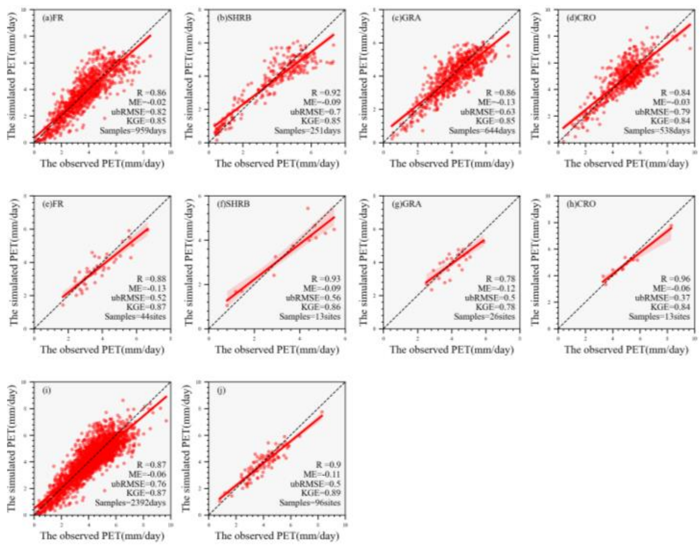

Figure R6: Validation results for the SW model calibrated with the 98th percentile EF threshold for the 8 sites in arid areas and the 95th percentile EF threshold for other 88 sites. (a-d) Scatter-plots of daily observations against simulations for each LULC type. (e-h) Scatter-plots of site mean observations against simulations for each LULC type. (i) Scatter-plot of daily observations against simulations for all of 96 EC sites. (j) Scatter-plot of site mean observations against simulations for all of 96 EC sites.

Comment 8: Line 170. The CO_2 concentration observed at Mauna Loa cannot represent the spatial heterogeneity of CO_2 concentration over the globe. How did the authors consider the uncertainty from this act?

Response: Thanks for your comments. For considering the spatial differences in CO_2, we have recalibrated the SW model and reproduced PET, PT and PE with the gridded CO_2 dataset (i.e., the monthly CO_2 concentration with a spatial resolution of 1◦ × 1◦ and a time span of 1850–2013 from https://doi.org/10.5281/zenodo.5021361 (Cheng et al., 2022), and the monthly Global CO_2 Distribution product from Japan Meteorological Agency with a spatial resolution of 2◦ × 2◦ and a time span of 1985–2021). The detailed information could be found in the revision.

Comment 9: Lines 178. I suggest specifying the full name of CRUTS like the previous dataset. Please keep CRUTS consistent with the CRU TS below.

Response: Thanks for your suggestion. The corrections have been done, and please see the revision.

Comment 10: Figure 2. I suggest adjusting the color configuration of the flow chart, such as reducing the color saturation.

Response: Thanks. We have revised this figure according to the suggestion, and please see below or the revised manuscript.

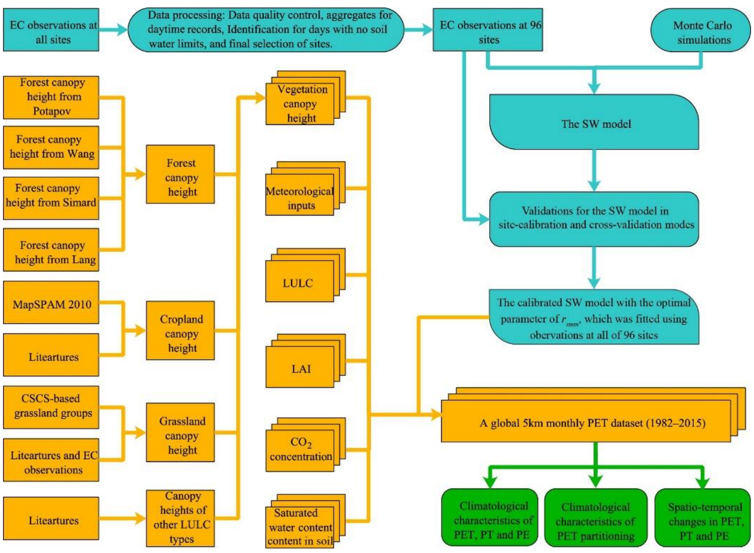

Figure 2: Workflow of this study. The blue and the yellow-green colors show operating procedures for calibrating and validating the SW model and for producing a global 5 km monthly PET dataset, respectively, while the green color presents the related analyses.

Comment 11: Lines 265. W/m2 should be corrected to W/m^2.

Response: Thanks. This mistake has been corrected in this revision.

Comment 12: Section 2.1.4. I suppose that the authors use the vegetation canopy height to estimate surface roughness length and further aerodynamic resistance. I have two questions. First, the authors have used multiple datasets as inputs or to obtain vegetation canopy height (h). How did the authors consider the difference between different datasets? Second, how to parameterize the spatially and temporally variable surface roughness length by using the time-invariant vegetation height? The authors are suggested to refer to the work of Peng et al. (2022) who developed for the first time a practical method for global estimates of 500 m daily aerodynamic roughness length with a combination of machine learning techniques, wind profile equation, observations from 273 sites and MODIS remote sensing data.

Response: Thanks for the valuable suggestions. Although the vegetation height is a critical parameter for establishing and running the SW model, the existing several forest height datasets could not fully cover the globe. Therefore, a vegetation height dataset with a global coverage was reconstructed in this study based on the several popular forest height datasets, the nearest neighbor interpolation method, and heights of the typical croplands and the typical grasslands. Additionally, to reduce uncertainties of forest height datasets (which were related to observation time, instruments, algorithms, pretreatments for the radiation signals, and spatial resolutions) and then the PET estimates, the ensemble mean forest height was used as the final values for the forests in this study. It is admitted that although such reconstructed dataset could fully cover the globe, the uncertainties still existed, potentially resulting in the biases of the estimated PET. Despite that, we believed that our PET datasets may be better than others (e.g., CRU, GLEAM, PT-JPL, and hPET) based on two considerations, such as (1) vegetation height considered here (relative to studies without considerations of vegetation height), and (2) spatial differences in the vegetation height used in this study (relative to studies which used the height of each typical LULC but ignored spatial differences in vegetation height for a given LULC). Additionally, we have shown some discussion about the uncertainties induced by the vegetation height in the manuscript, such as "*The reconstructed global vegetation canopy height also has limitations, which may raise from (1) uncertainties in the retrieval algorithms and remote sensing data (Simard et al., 2011; Wang et al., 2016; Potapov et al., 2020; Lang et al., 2021, 2022), (2) neglecting the spatial differences in CRO and GRA heights and using an alternative specific value, and (3) not considering the inter-annual changes in the FR and SHRB canopy heights and the intra-annual cycle in the CRO and GRA heights. These limitations undermine the accuracy of the PET estimates.*" **(L585-589)**

Following Zhou et al. (2006), we employed an empirical model to estimate aerodynamic roughness length (z_0) in this study, which combined effects of vegetation height and leaf area index (LAI). Seen from the model, the spatial variations of z_0 were mainly controlled by LAI and vegetation height, while LAI should be responsible for the temporal variations of z_0. Some studies have stated that besides LAI the temporal changes in vegetation height could also impact on z_0 (Peng et al., 2022; Guo et al., 2010; Zheng et al., 2014; Masseroni et al., 2015; Huang et al., 2016; Colin et al., 2006). However, due to lack of the time-varying vegetation height dataset, to consider the temporal changes in z_0 caused by the time-varying vegetation height is difficult. Therefore, without impacts of vegetation height, the temporal changes in z_0 based on the empirical model exist uncertainties.

We thank this reviewer very much for the valuable information about the 500 m daily aerodynamic roughness length developed by Peng et al. (2022). When we received the decision of this manuscript, we have tried to contact the authors, and would like to get this dataset to compare z_0 estimated in this study against z_0 from Peng et al. (2022). However, it is unfortunate that when we finished the responses, we still did not get this dataset (they replied our email but did not provide the dataset) and could not conduct comparisons. Moreover, we note that the time span of this dataset was shorter than our study period (mainly due to the availability of MODSI data). Therefore, even if we could get this dataset, it is still difficult to use this dataset to produce the PET before 2000.

References:

Colin, J., Menenti, M., Rubio, E. and Jochum, A.: Accuracy Vs. Operability: a Case Study Over Barrax In The Context Of The DEMETER Project. AIP Conference Proceedings 852, 75–83, 2006.

Guo, X., Qin, J., Zhou, D., Yang, K. and Chen, Y.: Improving the Noah land surface model in arid regions with an appropriate parameterization of the thermal roughness length. Journal of Hydrometeorology, 11, 995–1006, 2010.

Huang, Y., Salama, M.S., Su, Z., Van Der Velde, R., Zheng, D., Krol, M. S., Hoekstra, A. Y. and Zhou, Y.: Effects of roughness length parameterizations on regional-scale land surface modeling of alpine grasslands in the Yangtze River basin. Journal of Hydrometeorology, 17, 1069–1085, 2016.

Masseroni, D., Facchi, A. and Gandolfi, C.: Estimation of zero-plane displacement height and aerodynamic roughness length on rice fields. Italian Journal of Agrometeorology, 20, 67–75. 2015.

Peng, Z., Tang, R., Jiang, Y., Liu, M. and Li, Z. L.: Global estimates of 500 m daily aerodynamic roughness length from MODIS data. ISPRS Journal of Photogrammetry and Remote Sensing, 183, 336–351, 2022.

Zheng, D., Van Der Velde, R., Su, Z., Booij, M. J. and Hoekstra, A. Y.: Assessment of roughness length schemes implemented within the Noah land surface model for highaltitude regions. Journal of Hydrometeorology, 15, 921–937, 2014.

Zhou, M. C., Ishidaira, H., Hapuarachchi, H.P., Magome, J., Kiem, A. S., Takeuchi, K.: Estimating potential evapotranspiration using Shuttleworth-Wallace model and NOAA-AVHRR NDVI data to feed a distributed hydrological model over the Mekong River basin. Journal of Hydrology, 327(1–2): 151–173, 2006.

Comment 13: Line 246, Eq. (2c). The extinction coefficient 0.6 is empirical and may not be suitable for all vegetation types and fractional vegetation coverages. How did the authors consider this uncertainty?

Response: Thanks for your comments. In the original manuscript, the extinction coefficient values have no differences among vegetation. Therefore, for considering such differences, we have given different extinction coefficient values for different LULC types (Table S5 in the revision) through reviewing the existing references, and recalibrated and reproduce the PET dataset.

In this study, we did not consider changes in the extinction coefficient with fractional vegetation coverages. To that end, the related discussion about this issue has been added in the revision, such as "*The canopy light extinction coefficient of k_{ex}, which represents a partitioning of radiant energy over vegetation canopy and soil surface, is a key factor affecting ecosystem carbon, water, and energy processes (Tahiri et al., 2006; Zhang et al., 2014). As a result, the accuracy of the SW model was believed to be associated with the k_{ex} parameterization used in this study. Considering the physiological and morphological differences among terrestrial ecosystems (Emami-Bistghani et al., 2012; Zhang et al., 2014), we followed the popular biogeochemical models and remote sensing models of ET and gross primary productivity (e.g., Lund-Potsdam-Jena Dynamic Global Vegetation Model, and Vegetation Photosynthesis Model; Xiao et al., 2004; Sitch et al., 2003), and assumed this coefficient as a constant for every LULC type (Table S1). Despite that, it is notable that the k_{ex} values vary with the growth of plant and/or vegetation coverages (Lindroth and Perttu, 1981; Aubin et al., 2000; Maddoni et al., 2001; Emami-Bistghani et al., 2012; Fauset et al., 2017). Emami-Bistghani et al. (2012) stated that with an increase of vegetation coverages, the k_{ex} values decreased especially in early reproductive stage in sunflower cultivars. It suggested that the fixed k_{ex} within the SW model existed limitations, potentially leading to uncertainties of the PET estimates. Implied by Tahiri et al. (2006), relative to the variable k_{ex} values, the fixed values gave a less precise estimation of plant transpiration under irrigated maize.*" **(L522-534)**

Comment 14: Line 265-266. How did the authors consider the uncertainty from the minimum threshold (30 W/m^2 for forests and 100 W/m^2 for other vegetation)?

Response: The settings of the minimum threshold of canopy radiation ($K_{\downarrow dbl}$) for each vegetation type are mainly based on various references. For forests, some scholars recommend $K_{\downarrow dbl}$ to be 30 W/m^2 (Dickinson et al., 1984; Noilhan and Planton, 1989). Additionally, when conducting this study, we compared impacts of $K_{\downarrow dbl}$ (i.e., = 30 W/m^2 and 100 W/m^2) on performance of the SW model. Results showed that the SW model with $K_{\downarrow dbl}$ = 30 W/m^2 is slightly better than that with $K_{\downarrow dbl}$ = 100 W/m^2. This comparison confirms that the $K_{\downarrow dbl}$ = 30 W/m^2 for forests is reasonable. For other vegetation, $K_{\downarrow dbl}$ is recommended to be $100\mathrm{W/m^2}$ (Dickinson, 1984; Mu et al., 2017. Lo et al., 1997; Noilhan and Planton, 1989). Overall, the $K_{\downarrow dbl}$ values for forests and other vegetation have been proved to be reasonable by the previous studies, and we believed that the uncertainties induced by the different settings of $K_{\downarrow dbl}$ existed but limited.

References:

Dickinson, R. E.: Modeling evapotranspiration for three-dimensional global climate models. Climate processes and climate sensitivity, 29, 58–72, 1984.

Lo, S. D., Chehbouni, A., Njoku, E., Saatchi, S., Mougin, E., Monteny, B.: An approach to couple vegetation functioning and soil-vegetation-atmosphere-transfer models for semiarid grasslands during the HAPEX-Sahel experiment. Agricultural and Forest Meteorology, 83(1–2), 49–74, 1997.

Mu, Y., Li, J. and Tong, X.: Evapotranspiration simulated by Penman-Monteith and Shuttleworth-Wallace models over a mixed plantation in the southern foot of the Taihang Mountain, northern China. Journal of Beijing Forestry University, 39(11), 35–44, 2017.

Noilhan, J. and Planton, S.: A simple parameterization of land surface processes for meteorological models. Monthly Weather Review, 117, 536–549, 1989.

Comment 15: Figure 5 was not referenced in the text.

Response: Thanks for pointing out the mistake, and it has been corrected in this revision.

Comment 16: Line 428. "Figure 1b" should be "Figure 11b"?

Response: Thanks for pointing out the mistake. We have changed "Figure 1b" as "Figure 11b" in this revision.

Comment 17: Line 440. "PE was the major contributor of PE"? Perhaps the later PE should be PET.

Response: Sorry for this mistake. We have corrected "PE was the major contributor of PE" as "*PE was the major contributor of PET*" in this revision.