A gridded dataset of consumptive water footprints, evaporation, transpiration, and associated benchmarks related to crop production in China during 2000-2018

Wei Wang, La Zhuo, Xiangxiang Ji, Zhiwei Yue, Zhibin Li, Meng Li, Huimin Zhang, Rong Gao, Chenjian Yan, Ping Zhang, Pute Wu

Dear Editor and Referees,

We thank you very much for the valuable chance of revision and such structural and detailed comments improving our study substantially. We carefully learned and addressed the comments point by point. The revised part is in RED in the revised manuscript.

Authors' responses to Referee #1' s comments

This study constructed a gridded dataset for crop water consumption in China. The advantages of this dataset are the separate estimation of blue and green evaporation (E), transpiration (T), the use of the local information on irrigation type per crop, and the long time series. Even though comparisons with other studies on ET look good, I have many concerns about methods that are not clear in the paper. Without this information, it is hard for me to conclude on the dataset quality and reliability. I suggest a major revision and reconsideration for publication unless the authors can address the following concerns/questions.

Responses: We thank Referee #1 very much for such valuable comments and suggestions improving the manuscript. We have carefully addressed all the comments and provided our detailed responses below them, responding point by point.

Major comments:

1. The aquaCrop model setting is absent.

1-A) There are different versions of AquaCrop and available on different platforms, e.g. Windows interface for FAO AquaCrop v7, v6 or before, and Python and MATLAB open source versions. Which AquaCrop version was used for this study?

Responses: Thank you for your comments and questions regarding the AquaCrop version used in our study. We apologize for not providing this information in the manuscript. We used FAO AquaCrop Plug-In program V6.0 for crop simulation because AquaCrop v7 was not yet released during the course of our study. AquaCrop v7 was developed based on AquaCrop v6.0 and includes some bug fixes, performance improvements, internal reorganization, as well as the addition of perennial forage crops and new calibration crops. However, these updates have minimal impact on the crops involved in our study. Moreover, the Plug-In program and the standard windows program share consistent simulation logic. In order to enable batch calculations at the grid scale, we decided to utilize AquaCrop Plug-In program V6.0 for the simulations.

In line 68 of the revised manuscript, we have supplemented details on the version of the model utilized. In section 2.1 of Figure 1, the model logo was updated.

1-B) As I understand, publicly available AquaCrop cannot simulate perennial crops. How did the authors deal with perennial crops using AquaCrop, e.g. tea, and apple?

Responses: Thank you for your comments. As you have indicated, the number of studies that utilize AquaCrop to simulate perennial crops is limited. Hunink and Droogers (2010, 2011) conducted simulations in Albania and Uzbekistan to assess the response of yield and water requirements of different woody plants, including apple trees, grapevines, and olive trees, to climate change. Zhuo et al. (2016b) simulated the yield and evapotranspiration of tea and apple trees in China.

To accommodate the simulation of perennial crops in Aquacrop, the model is used differently than the normal model set-up. In AquaCrop, the simulated annual crops are programmed to die at the harvest stage, signifying the completion of their life cycle, upon which their biomass is reduced to zero. This stands in contrast to perennial plants such as tea and apple trees, where the harvest of fruits does not result in the complete loss of the standing biomass.

We attempted to simulate the perennial crops by simulating the foliage, twigs and stem of the plants following Poppe (2016). These components are considered the annual portion of perennial crops within the scope of this study. The remaining biomass, including major branches, is assumed to remain constant once the tree matures. Additionally, there will also be no root development for the crop. Since yield is a direct function of biomass and harvest index, adjustments are made to the harvest index to reflect its applicability to foliage, twigs and stem biomass, rather than the whole biomass. Similar to other crops, the evapotranspiration of perennial crops is directly associated with the canopy cover.

In the revised manuscript, we have added section 2.2.2 to explain how the AquaCrop model was utilized to simulate perennial crops in this study (Lines 134-144).

1-C) Using AquaCrop needs crop characteristics as input, e.g. the maximum canopy cover, canopy cover decline coefficient, canopy growth coefficient, and many others. Where did these inputs come from and what are the inputs?

Responses: Thank you for your comments. To clarify the sensitivity of a WFCP assessment to the main parameters in a simulation, a previous study by the authors applied the one-at-a-time and sensitivity index methods to quantitatively evaluate a WFCP calculation by AquaCrop (Li et al., 2022). The results indicated that crop water consumption and production were extremely sensitive to the reference evapotranspiration, the crop transpiration coefficient (**KcTr**) and planting date (**PD**).

In the Annex of the Reference manual for the AquaCrop (Raes et al., 2018), default values of crop parameters for the crops covered in this study are given, including crop transpiration, biomass production and yield formation, and stresses, totaling 41 parameters. Furthermore, these parameters are further classified based on crop sensitivity as conservative generally applicable (including **KcTr**), conservative for a given species but can or may be cultivar specific, dependent on environment and/or management and cultivar specific. The conservative parameters are generally applicable and remain unchanged across a wide spectrum of conditions, including different climatic and geographic locations, crop cultivars and genotypes, as well as variable soil moisture stress statuses. Once calibrated, these identical parameters would be utilized without further modification.

Table R1. Regional classification.

Region	Provinces	Regional classification
North	Beijing, Tianjin, Shanxi	Temperate
Northeast	Inner Mongolia, Liaoning, Jilin, Heilongjiang	Continental temperate and temperate
Huang-huai-hai	Hebei, Henan, Shandong, Anhui	Temperate

Northwest	Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang	Continental temperate and plateau and Mountain
Southeast	Shanghai, Zhejiang, Fujian	Sub-tropics
East	Jiangsu, Hubei, Hunan, Jiangxi	Sub-tropics
South	Guangdong, Guangxi, Hainan	Sub-tropics and tropics
Southwest	Chongqing, Sichuan, Guizhou, Yunan, Tibet	Sub-tropics

According to the regional classification results in Table R1, these key parameters like **PD**, reference harvest index, crop growth stages, and maximum root depth for this study are obtained by referring to the literatures described in revision section 2.1.3(Allen et al., 1998; Vanuytrecht et al., 2014; Chen et al., 1995; Xie et al., 2011; Zhang and Zhu, 1990; Hoekstra and Chapagain, 2006). These data have been validated to be reliable and applicable in large-scale studies (Cao et al., 2014; Zhuo et al., 2016a; Wang et al., 2019). Due to data limitations, the remaining parameters such as maximum canopy cover, canopy cover decline coefficient, canopy growth coefficient were assigned the mean values within the reference range provided in the Annex of the Reference manual for the AquaCrop. Although this approach may overlook certain potential variations, the use of mean values generally captures the central tendency of the data.

In summary, the objective of this study is to investigate the response and variability of crop water footprints to different water supply and irrigation practices at a large regional scale. Similar to previous global and national scale studies (e.g., Liu and Yang, 2010; Mekonnen and Hoekstra, 2011; Chiarelli et al., 2020), it was neither practical nor feasible to calibrate crop parameters individually for each grid. Nevertheless, we have made every effort to ensure the reliability of the model input parameters within the constraints of the available data.

In section 4.2 (Lines 440-445) of the revision and Phenology selection of the Supplementary data and methods, we have provided the screening process for key parameters affecting WFCP quantification, supplemented explanation on the sources and reliability of other crop parameters.

1-D) What do the authors mean by calibrated AquaCrop? How did the authors calibrate AquaCrop? Section 2.3.1 does not explain the calibration on AquaCrop.

Responses: Thank you for your comments. The calibrated AquaCrop in this study was achieved by strictly screening the input parameters and keeping the simulated yields consistent with the provincial-scale statistics. In line with previous studies (Wang et al., 2019; Siebert and Döll, 2010; Mialyk et al., 2022), we attempt to represent the combined effect of advances in agricultural inputs (e.g., fertilisers, machinery, and chemical control of weeds and insects) via production scaling factors that scale simulated crop production to the annual statistics.

For the ambiguous expression "calibrated AquaCrop", the verbiage has been emended in line 68 of the revised manuscript.

2. Reliability in separating E and T.

2-A) Distinguishing T from ET is not traditionally in water footprint studies. One challenge is that the ratio of T to ET depends not only on crop growth, and irrigation type but also on field management, e.g. weeds, soil fertility, mulching. As for large-scale assessment, there is always lacking data for field management, thus, it is difficult to estimate T and E separately. However, total ET is more robust to field management, so lacking such data will not be a big issue for total ET. Here the authors try to explicitly distinguish E and T from ET but still have very limited ability to describe the field management in the model. In this sense, how do the authors evaluate the reliability of separating E and T?

Responses: Thank you for your comments. We acknowledge that the ratio of transpiration (T) to evapotranspiration (ET) is influenced not only by crop growth and irrigation type but also by field management factors such as weeds, soil fertility, and mulching. In this study, we specifically focused on the impact of different irrigation practices on crop water footprint, aiming to elucidate the proportion of evaporation (E) and transpiration (T) in crop water consumption under different irrigation practices.

The water consumption results of current study were validated against the dual-source (PML-V2(China)) and single-source (SEBAL) remote sensing products. Notably, the SEBAL products (Cheng et al., 2021) solely comprised aggregate evapotranspiration figures, whereas the PML-V2 products (He et al., 2022) separated land surface evapotranspiration into vegetation transpiration (Ec), soil evaporation (Es), evaporation of intercepted precipitation (Ei), and water body evaporation (Ew). In this study, Ec + Es, Ec and Es were compared with the generated ET, E and T data, respectively.

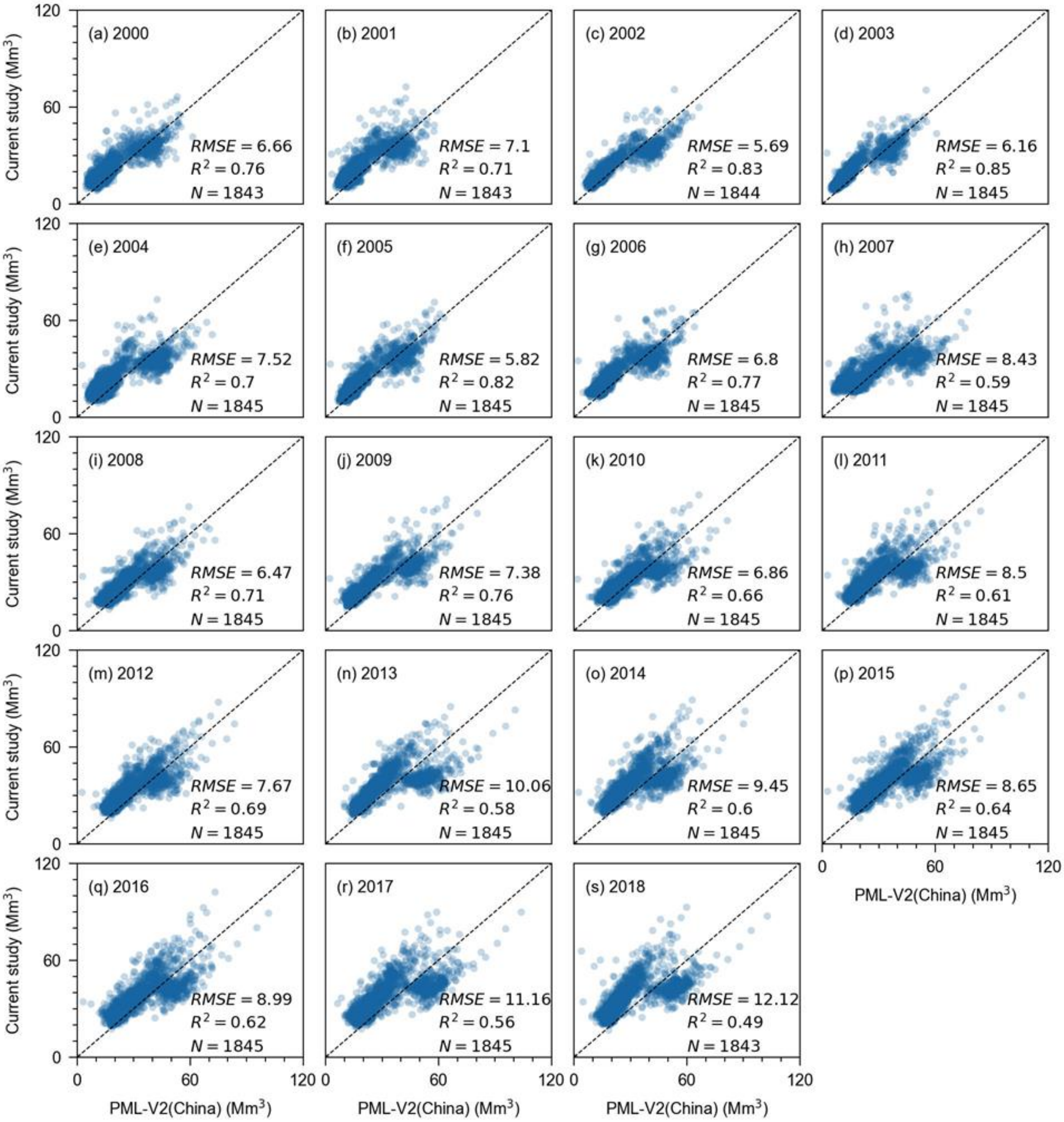

Figure R1. Validation of the evapotranspiration at croplands for the period April to August with PML-V2(China) datasets.

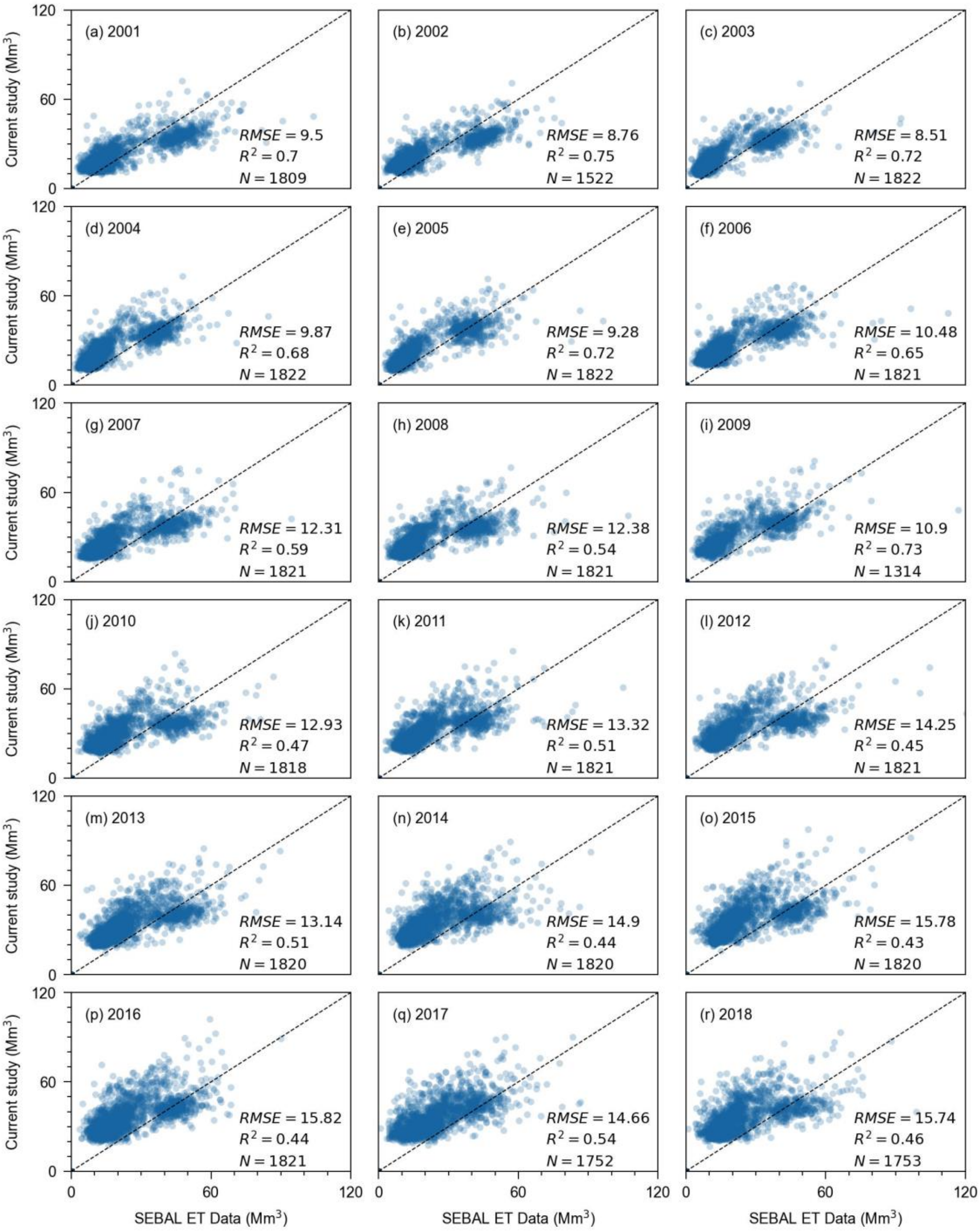

Figure R2. Validation of the evapotranspiration at croplands for the period April to August with SEBAL datasets.

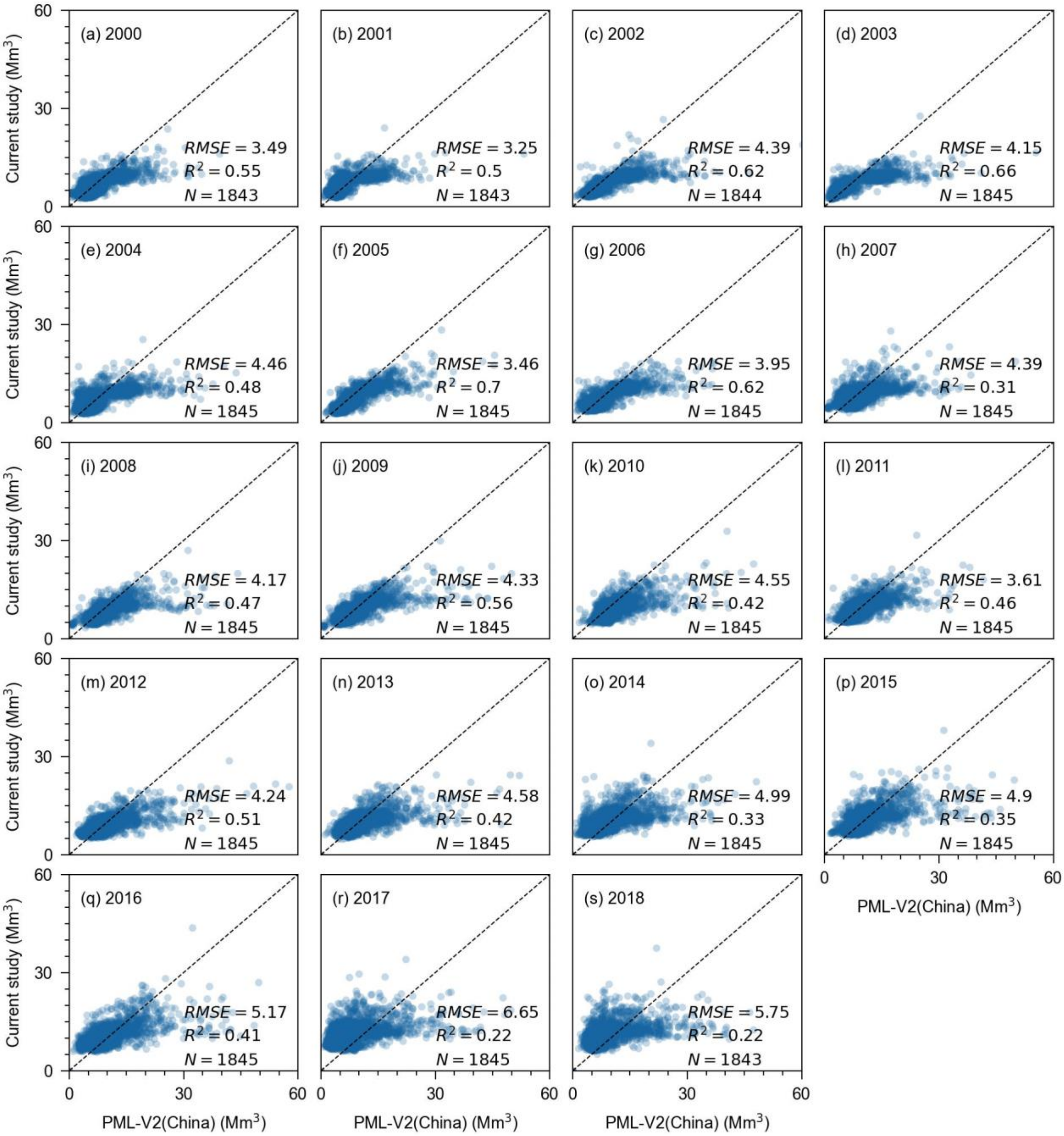

Figure R3. Validation of the evaporation at croplands for the period April to August with PML-V2(China) datasets.

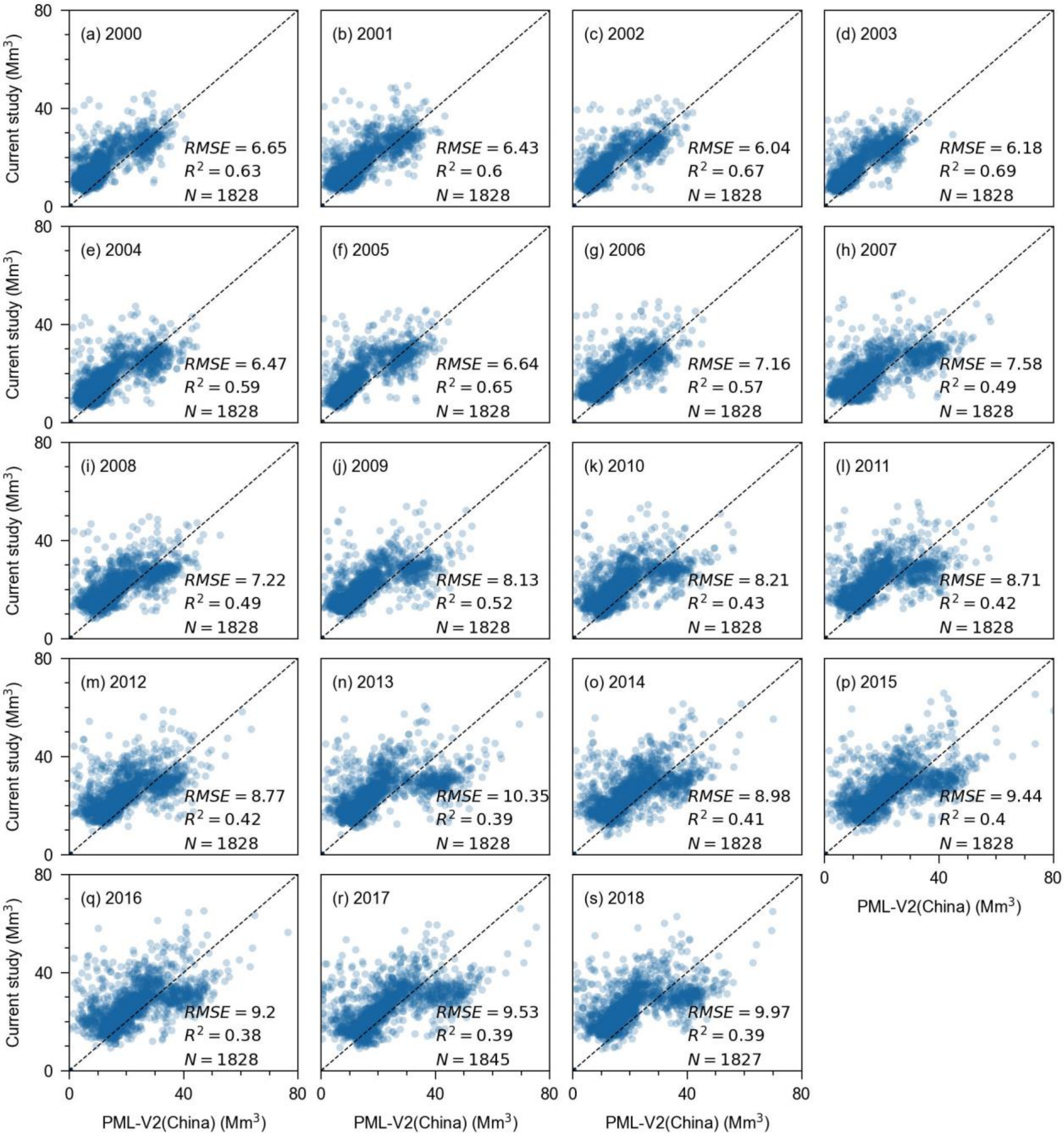

Figure R4. Validation of the transpiration at croplands for the period April to August with PML-V2(China) datasets.

Comparative analysis in Figure R1 and Figure R2 revealed stronger agreement between the simulated evapotranspiration and the PML-V2 products (R^2=0.49 – 0.85, *RMSE*= 5.82 – 12.12 Mm^3) than those with the SEBAL products (R^2= 0.44 – 0.75, *RMSE*= 8.51 – 15.82 Mm^3), although both comparisons demonstrated robust overall consistency. The validation results of soil evaporation are presented in Figure R3. The simulated E were marginally lower than the PML-V2 products (R^2=0.22 – 0.70, *RMSE*=3.25 – 6.65 Mm^3), owing to the current study calculating E exclusively for the planted regions of 21 crops, whereas the PML-V2 disregarded land use types during E estimation. Comparative analysis of crop transpiration in Figure R4 indicated that our simulated values were higher than the PML-V2 products which deducted canopy evaporation (R^2=0.38 – 0.69, *RMSE*=6.04 – 10.35 Mm^3). Overall, considering the differences in basic input data, spatiotemporal resolution and calculation methods, the evapotranspiration, evaporation, and transpiration data products produced in this study showed acceptable results when compared with various remote sensing products, given the discrepancies exhibited.

In section 2.3.2 (Lines 230-241) of the revised manuscript, we introduced the datasets and data preprocessing methods used for validation. In lines 421-435 and validation results of the Supplementary data and methods, we presented the comparison results between the simulated values in this study and existing remote sensing datasets.

2-B) Did the authors further distinguish the color of E and T? and how?

Responses: Thank you for your comments. We distinguished the color of *E* and Tr as follows (Zhuo et al., 2016b; Chukalla et al., 2015):

$$\mathrm{ET}_{\mathrm{b}[t]} = \mathrm{IRR}_{[t]} + S_{\mathrm{b}[t-1]} - S_{\mathrm{b}[t]} - \mathrm{RO}_{[t]}\left(\frac{\mathrm{IRR}_{[t]}}{\mathrm{PR}_{[t]} + \mathrm{IRR}_{[t]}}\right) - \mathrm{DP}_{[t]}\left(\frac{S_{\mathrm{b}[t-1]}}{S_{[t-1]}}\right) \tag{1}$$

$$\mathrm{ET}_{\mathrm{g}[t]} = \mathrm{PR}_{[t]} + S_{\mathrm{g}[t-1]} - S_{\mathrm{g}[t]} - \mathrm{RO}_{[t]}\left(\frac{\mathrm{PR}_{[t]}}{\mathrm{PR}_{[t]} + \mathrm{IRR}_{[t]}}\right) - \mathrm{DP}_{[t]}\left(\frac{S_{\mathrm{g}[t-1]}}{S_{[t-1]}}\right) \tag{2}$$

$$E_{\mathrm{b}[t]} = \mathrm{E}_{[t]}\left(\frac{S_{\mathrm{b}[t-1]}}{S_{[t-1]}}\right) \tag{3}$$

$$E_{\mathrm{g}[t]} = E_{[t]}\left(\frac{S_{\mathrm{g}[t-1]}}{S_{[t-1]}}\right) \tag{4}$$

$$\mathrm{Tr}_{\mathrm{b}[t]} = \mathrm{Tr}_{[t]}\left(\frac{S_{\mathrm{b}[t-1]}}{S_{[t-1]}}\right) \tag{5}$$

$$\mathrm{Tr}_{\mathrm{g}[t]} = \mathrm{Tr}_{[t]}\left(\frac{S_{\mathrm{g}[t-1]}}{S_{[t-1]}}\right) \tag{6}$$

The green and blue components in *E* and Tr are calculated per day based on the fractions of green and blue water in the total soil water content at the end of the previous day.

In lines 181-187 of the revised manuscript, we have supplemented details on how evaporation (*E*) and transpiration (Tr) were distinguished.

3. Irrigation and soil moisture assumption. In Table 5, comparing the previous studies,

this study has a higher water footprint. According to my experience using these datasets, they generally overestimate water consumption compared to hydrological models and recent crop models for some crops because of two reasons (maybe more): first, they assume irrigation once there is a water deficit, even though the water deficit would be tiny; second, when setting the initial soil moisture, they assumed the field capacity of the soil moisture at the beginning of each year. Both are unrealistic and will lead to overestimation of water consumption. How do the authors set irrigation rules and soil moisture in the model and did the authors have a spin-up for the model?

Responses: Thank you for your comments. In our study, we employed a supplementary irrigation strategy, similarly to the first possible reason as mentioned, whereby irrigation is applied when soil moisture falls below the plant wilting point to bring it up to field capacity. Different irrigation practices indirectly affect water consumption during the growth period due to differences in the fraction of the surface wetted. The irrigation strategy used in this study was added in lines 193-196 of the revised manuscript.

To establish the initial soil moisture content at the beginning of the growing season, we adopted the method and assumptions proposed by Siebert and Döll (2010). Following their approach, we generated the initial soil moisture content by using the maximum soil moisture content of rainfed fallow land in the two years preceding the planting period. The initial soil moisture at the start of the growing period is assumed as green water. Such settings and assumptions have been widely applied and with acceptable uncertainties (Chiarelli et al., 2020; Hoogeveen et al., 2015). In the revised manuscript, section 2.2.1 (Lines 128-133) has been added to elaborate on the spin-up methodology employed in this study.

4. Usability and quality of the benchmark

4-A) Figure 6 shows the benchmark at the grid level. However, if we have a close look at the graph, we find usually, the resolution is the provincial level, e.g. the whole province is efficient or not. This is because the calibration and scaling were on the provincial level. If it is the case, scaling factors seem to play a critical role here other than irrigation, climate, and so on. What do the authors think of the reliability of the benchmark in this case?

Responses: Thank you for your comments. Figure 6 presents the benchmarks for uWFCP at different production percentiles under furrow irrigation in China by 2018. A few uWFCP benchmarks show significant regional correlations at the provincial scale due to inter-annual variations in model inputs. It is essential to emphasize that crop production and water consumption simulations for the growing period in this study were conducted at the grid scale after integrating high spatial resolution soil texture, precipitation, temperature and other model inputs. The above process ensured the spatial heterogeneity of the simulated results. Given the data accessibility, the current study utilized provincial statistics to validate the simulated production by scaling each grid-based simulated result using provincial calibration coefficients, rather than forcing the simulated production of all grid within a province to a constant value (Equations 16 and 17). This method maintained the spatial variability of simulated crop production within each province. Hence, we consider the uWFCP benchmarks estimated in this study to be robust.

Additional elaborations have been provided in section 2.3.1 (Lines 214-217, 223-228) of the revised manuscript to clarify how calibrating crop production at the provincial level retains the spatial variability of uWFCP.

4-B) The whole of China was divided into two climate zones for the benchmark without considering soil type. This seems oversimplified to me. Furthermore, this dataset provides the benchmark for each year, but the idea for a benchmark is to provide a kind of efficient "reference". So, why not derive using the whole time series? Imagine one wants to use the benchmark for future analysis, which year should be selected? How do the authors suggest using their benchmark?

Responses: Thank you for your comments. Our previous study examined crop uWFCP benchmarks for rain-fed and irrigated croplands, wet and dry years, warm and cold years, different soil types, and various climate zones have indicated that different climate zones are crucial factors influencing the uWFCP benchmarks (Zhuo et al., 2016c). Therefore, this study considered distinguishing the uWFCP benchmarks for different climate zones at the grid scale. In Supplementary Table S1, we provide the uWFCP benchmark of 21 crops over the whole time series. By utilizing the multi-year average of specific crop uWFCP benchmarks in guiding production practices, the influence of extreme climate variability within the same region can be mitigated. In fact, as you pointed out, Yue et al. (2022) utilized wheat uWFCP benchmarks generated by Wang et al. (2019) for their future analysis.

In section 2.2.6 (Lines 204-211) of the manuscript, we have elaborated on why we only computed the uWFCP benchmarks for different climate zones. In section 3.3 of the revision (Lines 348-350), we point out that uWFCP benchmarks generated using whole time series should be employed for analysis in subsequent studies.

Minor comments:

1. In the title, "consumptive water footprints". By definition, water footprint is consumptive water use. Is consumptive redundant here? Or "consumptive water use"?

Responses: Thank you for your comments. The water footprint is a multi-dimensional indicator and divided into consumptive water footprint and degradative water footprint (Hoekstra, 2013), in which the consumptive water footprint refers to the use of both rainfall and ground-surface water (the green and blue water footprint, respectively), and the degradative water footprint refers to the water required to assimilate anthropogenic loads of pollutants to freshwater bodies (the grey water footprint). Here, consumptive water footprint refers to the blue and green water footprint crop production, excluding grey water footprint.

In line 37 of the manuscript, we provide the source for the definition of "consumptive water footprint of crop production".

2. Line 95-102. The authors basically only rely on the global datasets for land use from the year 2000 and then scale to each year. Are there other better datasets that have better quality/time coverage in China?

Responses: Thank you for your comments. In our study, we employed global datasets for land use from the MIRCA2000 as the primary data source due to its wide accessibility and compatibility with our research objectives. It is important to note that MIRCA2000 dataset (e.g., Hoch et al., 2023; Li et al., 2023; Ruess et al., 2022; Lutz et al., 2022; Liu et al., 2022; Chiarelli et al., 2022; Chiarelli et al., 2020; Rosa et al., 2020) and proportional scaling method (Sloat et al., 2020; Yue et al., 2022; Mialyk et al., 2022; Wang et al., 2019) have been applied in several studies within the field. As shown in section 4.1, the harvested area of specific crop obtained by the proportional scaling method show a good consistency with the published data in the same time range of the grid scale.

Table R2 presents the crop planting area and irrigated area data products. Recent years have witnessed the emergence of numerous long-term and high-resolution irrigation area datasets for China, thanks to the combined application of remote sensing technology and machine learning approaches. However, these datasets do not differentiate between irrigated and rainfed cropping systems, and do not contain crop-specific planting information. These deficiencies fail to fulfill the original intentions of this study design. Global scale data compensates the aforementioned deficiencies to some extent. However, it is worth mentioning that existing global databases have certain limitations, including a limited range of crop types and intermittent time series. For instance, the SPAM dataset is only publicly available for a few specific years 2000, 2005 and 2010, and interpolation is still required to fill in the gaps.

Table R2. Inventory of irrigated cropland data.

Source	Spatial coverage	Temporal resolution	Spatial resolution	Crop type distinction	Planting pattern
Zhang et al., 2022a	China	2000-2019	500 m	No	Only irrigated croplands
Zhang et al., 2022b	China	2000	250 m	No	Only irrigated croplands
Zhu et al., 2014	China	2000	5 arcmin	No	Only irrigated croplands
GFSAD1KCD	Globe	2007-2012	1000 m	6 crops	Irrigated and rainfed croplands
GAEZ+ (Grogan et al., 2022)	Globe	2015	5 arcmin	26 crops	Irrigated and rainfed croplands
SPAM (IFPRI 2019)	Globe	2000, 2005, 2010	5 arcmin	42 crops	Irrigated and rainfed croplands
MIRCA2000 （Portmann et al., 2010）	Globe	2000	5 arcmin	26 crops	Irrigated and rainfed croplands

As a supplement, we present a comparison between the crop planting areas of 15 crops in the SPAM2010 dataset and our research dataset at both the provincial and grid scales (Figures R5 and R6). It is evident that there is a high degree of consistency between the two datasets at the provincial scale. The differences at the grid scale can be attributed to disparities in the identification of grid-level land use between the MIRCA2000 and SPAM dataset. According to Figures R7 and R8, the planting area data for sorghum, millet, barley, and sugar beets in the GAEZ+ dataset exhibit significant deviations from the values applied in this study, both at the provincial and grid scales. However, it should be emphasized that all crop planting area data in this study have been calibrated against statistical data at the provincial scale, implying an underestimation of the planting area for the mentioned crops in the GAEZ+ dataset.

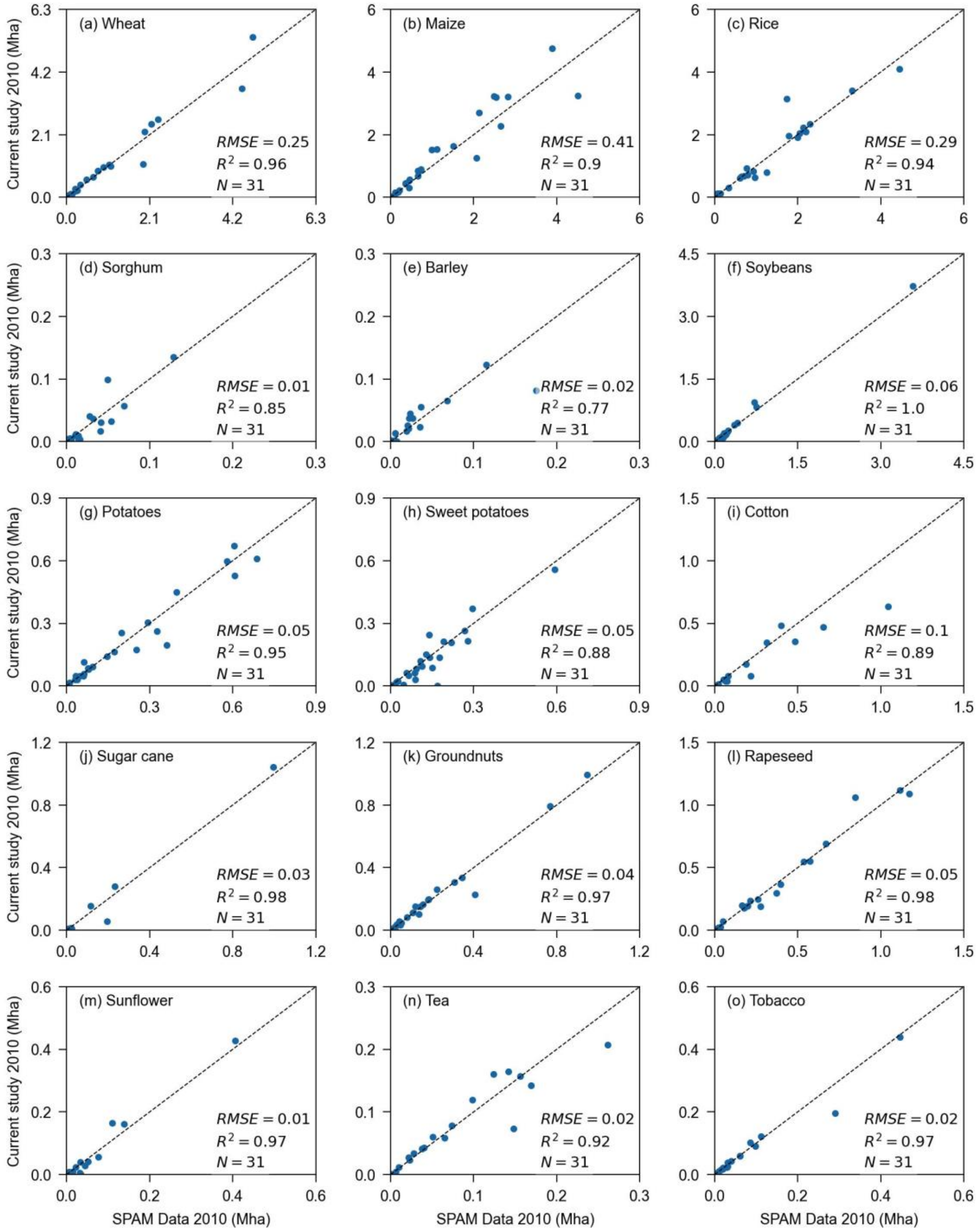

Figure R5. Comparison of the current provincial area representing land coverage with the SPAM datasets.

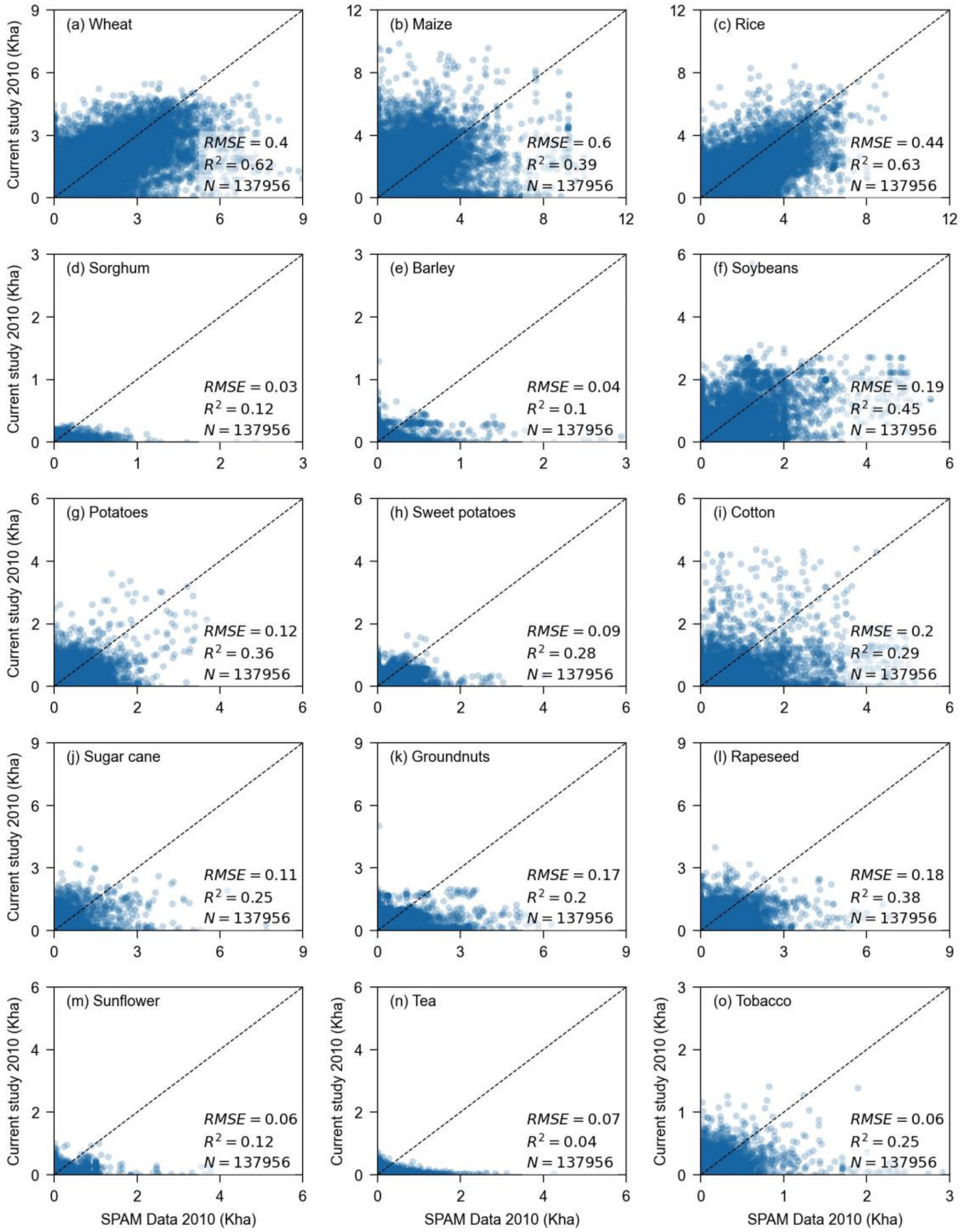

Figure R6. Comparison of the current gridded area representing land coverage with the SPAM datasets.

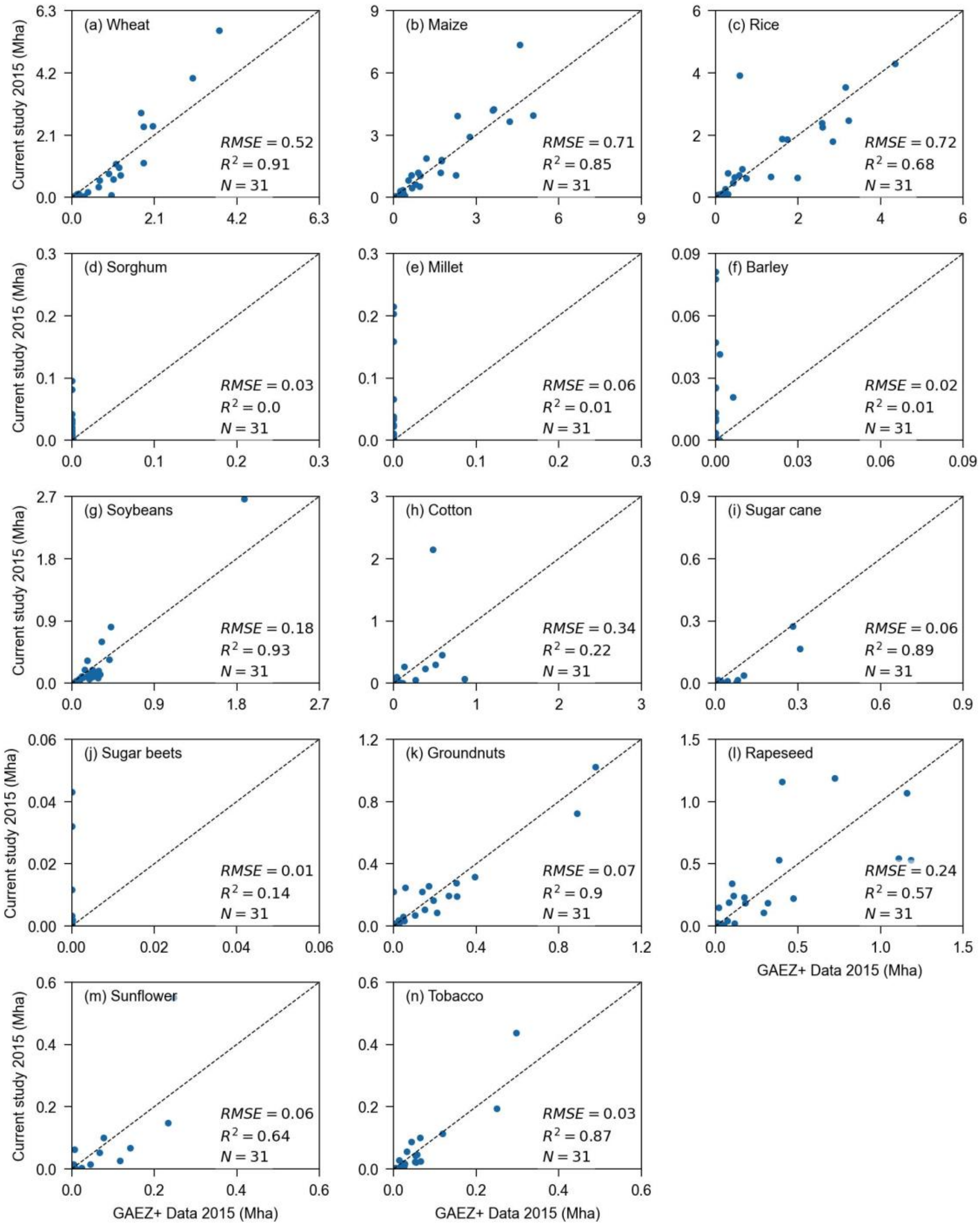

Figure R7. Comparison of the current provincial area representing land coverage with the GAEZ+ datasets.

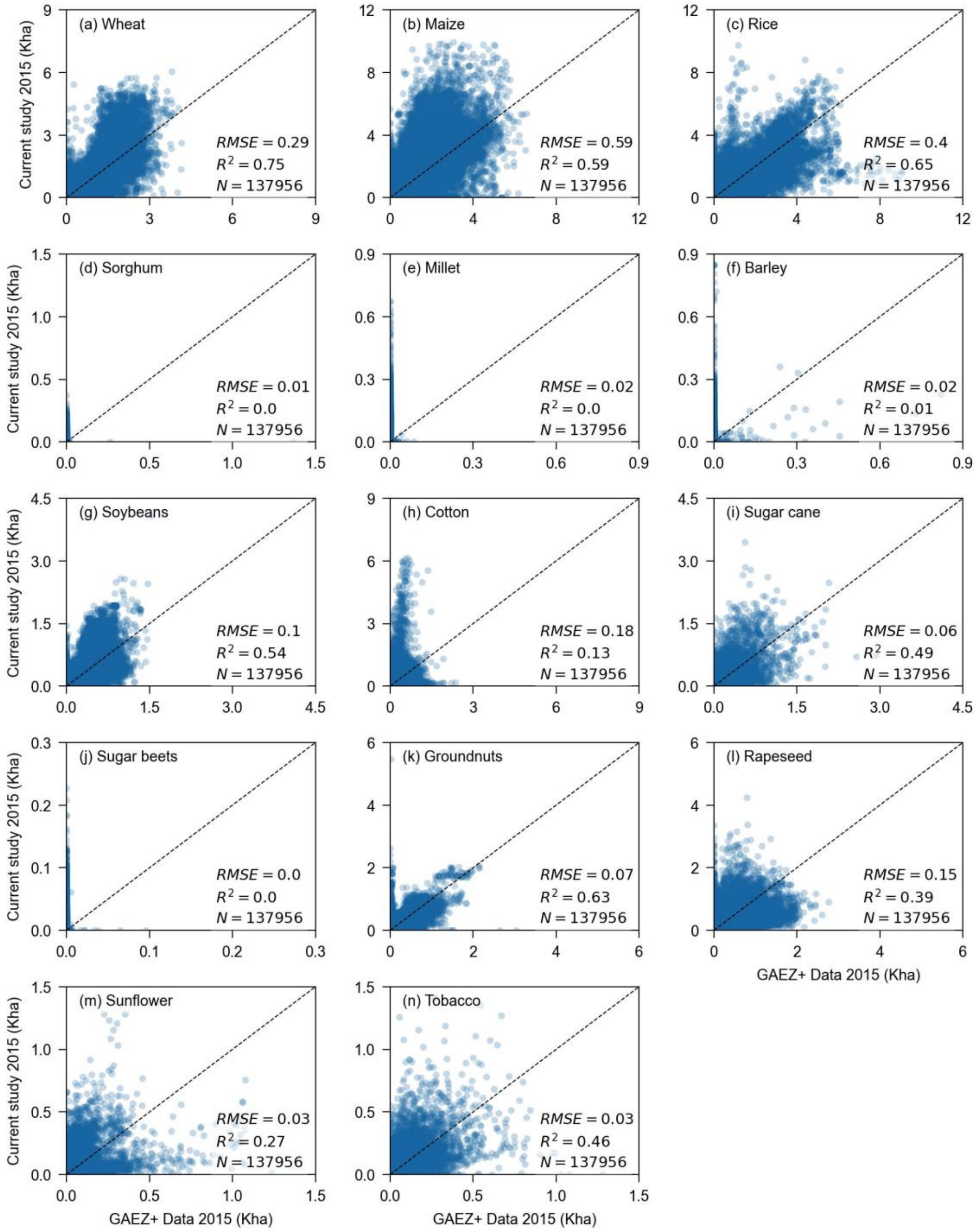

Figure R8. Comparison of the current gridded area representing land coverage with the GAEZ+ datasets.

Given the above analysis, we choose the MIRCA2000 database for estimating the planted areas of different crops under various water supply and irrigation practices. While we acknowledge the potential for alternative datasets, it is important to highlight that our study made a conscious decision based on the available data sources and applied rigorous methods to ensure reliability and comparability.

Further elaboration regarding the crop planting area data selection has been noted in lines 100-101 of the revised manuscript. Comparative analyses between the data utilized in the current study and the SPAM and GAEZ+ data products have been incorporated in section 4.1 of the revision (Lines 404-419). The inventory of the latest Chinese irrigated cropland database, detailed delineation of the crop planting area data selection, and additional comparative crop planting area data results have been provided within the Planting area selection of Supplementary data and methods.

3. What is CC in Equation 11? Canopy cover?

Responses: Thank you for your comments. Sorry for not explaining this variable in the text. CC is canopy cover (m^2 m^{-2}).

In line 202 of the revision, we provide the explanation for CC.

4. I didn't see the causal relationship between the number in lines 222-228 and why it is important to distinguish irrigation type. Lines 263-269 do explain some.

Responses: Thank you for your comments. In this paragraph, our research findings underscore the substantial increase in WFCP by 0.8 and 25 times under sprinkler and micro-irrigation, respectively, while a decreasing trend is observed under furrow irrigation. Considering the positive correlation between WFCP and the cultivated area under different water supply and irrigation practices, the above results reflect the preference for sprinkler and micro-irrigation over furrow irrigation on existing and freshly reclaimed farmland.

Given the large scale of crop cultivation in China, such a significant shift in irrigation practices will have important implications: (i) it will affect the quantification of national crop water consumption; (ii) it will create market opportunities while concurrently propelling technological innovation in the irrigation infrastructure.

Therefore, different from the comparison of uWFCP in lines 263-269 in original manuscript, this paragraph strives to elucidate the significant impact of considering different irrigation practices at the national level from the perspective of WFCP.

In lines 270-275 of the revision, we have elaborated on the causal relationships between the numbers in this paragraph and explained why distinguishing between irrigation types is important, per your suggestion.

5. In the paper, e.g. line 240, the authors discussed the monthly blue and green water consumption but is not available in the dataset. Some items are available monthly and some annually. What are the criteria to decide which data is open to the public or not? And the dataset lacks projected coordinate system information.

Responses: Thank you for your comments. The current online database only comprises monthly-scale data on WFCP under various water supply and irrigation practices. Monthly-scale data for blue and green water footprint will be uploaded to the online database in a timely manner once our ongoing work is published. If necessary, we can share the unpublished data with you. The online database adopts the "WGS 84" projection coordinate system.

In the updated data description, we have added the projected coordinate system information. See https://zenodo.org/record/7756013 for details.

6. Line 245-247, why sprinkler irrigation has the highest value? Because of crop type?

Responses: Thank you for your comments. The spatial distribution of WFCP under different water supply and irrigation practices is expressed in units of mm mon^{-1} (monthly scale) and mm yr^{-1} (annual scale), representing the water consumption per unit planted area at monthly and annual scales. As described in section 2.2.3, Different irrigation practices indirectly affect water consumption during the growth period due to differences in the fraction of the surface wetted (f_w). The f_w-values used for furrow, sprinkler, and micro-irrigation were 80%, 100%, and 40%, respectively. The fraction of the surface wetted influences the soil evaporation coefficient (K_e), resulting in increased water consumption per unit cultivated area during the crop growing period (Equations 6, 7 and 14). Therefore, the higher grid WFCP, WFCPb, and WFCPg of sprinkler irrigation at the monthly and annual scales mentioned in this paragraph refers to the higher water consumption per unit cultivated area. Since the planting area under sprinkler irrigation is small, the total WFCP, WFCPb, and WFCPg under sprinkler irrigation is not high (As shown in Table 1).

As mentioned in our previous study (Wang et al., 2019), conventional water-saving irrigation methods primarily focus on increasing water use efficiency in the field and reducing water losses during delivery. For instance, sprinkler irrigation can effectively reduce water demand by 46% compared to furrow irrigation (Xue and Ren, 2016). However, when it comes to reducing the water footprint of crop production, sprinkler irrigation is not an efficient solution. On the one hand, although furrow irrigation is less efficient with higher percolation and runoff fluxes compared to sprinkler irrigation, these fluxes return to the catchment and are not considered as actual water consumption (Hoekstra et al., 2011; Grafton et al., 2018). On the other hand, sprinkler irrigation results in a larger ET due to the large surface-wetting rate for an equal yield.

In lines 297-299 of the revision, we have explained why the grid WFCP, WFCPb, and WFCPg is higher under sprinkler irrigation.

7. Line 298-300, can the authors explain why the benchmarks of some crops are higher in arid zones and others are higher in humid zones?

Responses: Thank you for your comments. The benchmarks of crops under rainfed cultivation are higher than those under irrigation cultivation. The benchmarks of crops under different irrigation practices vary according to climatic zones and crop types. Several factors contribute to these results. Firstly, crops cultivated in arid zones are more irrigation-reliant due to scarce precipitation and undergo greater evapotranspiration, resulting in higher uWFCP versus humid zones. Secondly, certain crops like cotton possess higher benchmarks in humid zones since their yields are markedly lower than those extensively grown in arid regions. Importantly, the uWFCP benchmarks of the same crop vary depending on different irrigation practices and climatic zones. Based on the findings of this study, agricultural practitioners can guide production practices from the perspectives of irrigation practices and climatic zones.

In lines 352-355 of the revision, we have explained why the benchmarks of some crops are higher in arid zones and others are higher in humid zones.

8. in GAEZ+2015 and MapSPAM2010, they also have other crops, did the authors include them in the comparison? And GAEZ+2015 is developed for the year 2015 and MapSPAM2010 is developed for the year 2010, why not use the corresponding years other than the average between 2001 and 2008?

Responses: Thank you for your comments. The publicly available crop planting area data in GAEZ+ dataset is limited to the year 2015, whereas the SPAM dataset covers data for the years 2000, 2005, and 2010. In the main text, we have presented the comparison of major crop planting areas with the GAEZ+2015 and MapSPAM2010 datasets. The comparative results of other crops planting areas have been discussed and explained earlier in the minor comments 2.

Technical corrections

1. Table 3 doesn't seem consistent with the text. I only checked for rice. In Table 3, the total water consumption of rice is 81847+58979+4629+5540=150995 M m^3 and in the text, it is 143 G m3.

Responses: Thank you for your comments. Sorry for this expression error. The 143 G m^3 in the text refers to the multi-year average of WFCP from 2000 to 2018, while the table only presented the WFCP for the year 2018.

In lines 259, 260 and 270 of the revision, we have differentiated between annual average WFCP and WFCP for specific years in the text descriptions.

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration—Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, 300, D05109, FAO, Rome, 1998.

Cao, X., Wu, P., Wang, Y., and Zhao, X.: Assessing blue and green water utilisation in wheat production of China from the perspectives of water footprint and total water use, Hydrol. Earth Syst. Sci., 18(8), 3165-3178, https://doi.org/10.5194/hess-18-3165-2014, 2014.

Chen, Y., Guo, G., Wang, G., Kang, S., Luo, H., and Zhang, D.: Main crop water requirement and irrigation of China, Hydraulic and Electric Press, Beijing, 1995.

Cheng, M., Jiao, X., Li, B., Yu, X., Shao, M., and Jin, X.: Long time series of daily evapotranspiration in China based on the SEBAL model and multisource images and validation, Earth Syst. Sci. Data, 13, 3995-4017, https://doi.org/10.5194/essd-13-3995-2021, 2021.

Chiarelli, D. D., D'Odorico, P., Müller, M. F., Mueller, N. D., Davis, K. F., Dell'Angelo, J., Penny, G., and Rulli, M. C.: Competition for water induced by transnational land acquisitions for agriculture, Nat. Commun., 13(1), 505, https://doi.org/10.1038/s41467-022-28077-2, 2022.

Chiarelli, D. D., Passera, C., Rosa, L., Davis, K. F., D'Odorico, P., and Rulli, M. C.: The green and blue crop water requirement WATNEEDS model and its global gridded outputs, Sci. Data, 7, 273, https://doi.org/10.1038/s41597-020-00612-0, 2020.

Chukalla, A. D., Krol, M. S., and Hoekstra, A. Y.: Green and blue water footprint reduction in irrigated agriculture: effect of irrigation techniques, irrigation strategies and mulching, Hydrol. Earth Syst. Sci., 19, 4877-4891, https://doi.org/10.5194/hess-19-4877-2015, 2015.

GFSAD1KCD.: Global Food Security Support Analysis Data (GFSAD) Crop Dominance 2010 Global 1 km, available at: https://doi.org/10.5067/MEaSUREs/GFSAD/GFSAD1KCD.001. (last access: 6 September 2023), 2014.

Grafton, R. Q., Williams, J., Perry, C. J., Molle, F., Ringler, C., Steduto, P., Udall, B., Wheeler, S. A., Wang, Y., Garrick, D., and Allen, R. G.: The paradox of irrigation efficiency, Science, 361, 748-750, https://doi.org/10.1126/science.aat9314, 2018.

Grogan, D., Frolking, S., Wisser, D., Prusevich, A., and Glidden, S.: Global gridded crop harvested area, production, yield, and monthly physical area data circa 2015, Sci. Data, 9, 15, https://doi.org/10.1038/s41597-021-01115-2, 2022.

He, S., Zhang, Y., Ma, N., Tian, J., Kong, D., and Liu, C.: A daily and 500 m coupled evapotranspiration and gross primary production product across China during 2000–2020, Earth Syst. Sci. Data, 14, 5463-5488, https://doi.org/10.5194/essd-14-5463-2022, 2022.

Hoch, J. M., Sutanudjaja, E. H., Wanders, N., Van Beek, R. L., and Bierkens, M. F.: Hyper-resolution PCR-GLOBWB: opportunities and challenges from refining model spatial resolution to 1 km over the European continent, Hydrol. Earth Syst. Sci., 27(6), 1383-1401, https://doi.org/10.5194/hess-27-1383-2023, 2023.

Hoekstra, A. Y. and Chapagain, A. K.: Water footprints of nations: water use by people as a function of their consumption pattern, Water Resour. Manag., 21, 35-48, https://doi.org/10.1007/s11269-006-9039-x, 2006.

Hoekstra, A. Y., Chapagain, A. K., Aldaya, M. M., and Mekonnen, M. M.: The water footprint assessment manual: Setting the global standard, Routledge, London, 2011.

Hoekstra, A. Y.: The water footprint of modern consumer society, Routledge, London, 2013.

Hoogeveen, J., Faurès, J.-M., Peiser, L., Burke, J., and van de Giesen, N.: GlobWat–a global water balance model to assess water use in irrigated agriculture, Hydrol. Earth Syst. Sci., 19, 3829-3844, https://doi.org/10.5194/hess-19-3829-2015, 2015.

Hunink, J. and Droogers, P.: Climate Change Impact Assessment on Crop Production in Albania, Futurewater, Wageningen, The Netherlands, vol. 105, 2010.

Hunink, J. and Droogers, P.: Climate change impact assessment on crop production in Uzbekistan, Futurewater, Wageningen, The Netherlands, vol. 106, 2011.

IFPRI, International Food Policy Research Institute: Global spatially-disaggregated crop production statistics data for 2010 version 2.0, available at: https://doi.org/10.7910/DVN/PRFF8V. (last access: 7 March 2023), 2019.

Li, Z., Feng, B., Wang, W., Yang, X., Wu, P., and Zhuo, L.: Spatial and temporal sensitivity of water footprint assessment in crop production to modelling inputs and parameters, Agric. Water Manage, 271, 107805, https://doi.org/10.1016/j.agwat.2022.107805, 2022.

Li, Z., Sun, S., Zhao, J., Li, C., Gao, Z., Yin, Y., Wang, Y., and Wu, P.: Large Scale Crop Water Footprint Evaluation Based on Remote Sensing Methods: A Case Study of Maize, Water Resour. Res., e2022WR032630, https://doi.org/10.1029/2022WR032630, 2023.

Liu, J. and Yang, H.: Spatially explicit assessment of global consumptive water uses in cropland: Green and blue water, J. Hydrol., 384, 187-197, https://doi.org/10.1016/j.jhydrol.2009.11.024, 2010.

Liu, X., Liu, W., Tang, Q., Liu, B., Wada, Y., and Yang, H.: Global agricultural water scarcity assessment incorporating blue and green water availability under future climate change, Earths Future, 10(4), e2021EF002567, https://doi.org/10.1029/2021EF002567, 2022.

Lutz, A. F., Immerzeel, W. W., Siderius, C., Wijngaard, R. R., Nepal, S., Shrestha, A. B., Wester, P., and Biemans, H.: South Asian agriculture increasingly dependent on meltwater and groundwater, Nat Clim Chang, 12(6), 566-573, https://doi.org/10.1038/s41558-022-01355-z, 2022.

Mekonnen, M. M. and Hoekstra, A. Y.: The green, blue and grey water footprint of crops and derived crop products, Hydrol. Earth Syst. Sci., 15, 1577-1600, https://doi.org/10.5194/hess-15-1577-2011, 2011.

Mialyk, O., Schyns, J. F., Booij, M. J., and Hogeboom, R. J.: Historical simulation of maize water footprints with a new global gridded crop model ACEA, Hydrol. Earth Syst. Sci., 26, 923-940, https://doi.org/10.5194/hess-26-923-2022, 2022.

Poppe, M.: Simulating the water footprint of woodies in Aquacrop and Apex, University of Twente, 2016.

Portmann, F. T., Siebert, S., and Döll, P.: MIRCA2000—Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling, Glob. Biogeochem. Cycle, 24, GB1011, https://doi.org/10.1029/2008GB003435, 2010.

Raes, D., Steduto, P., Hsiao, T., and Fereres, E.: Annexes, AquaCrop version 6.0–6.1. Reference manual, FAO, Rome, 2018.

Rosa, L., Chiarelli, D. D., Rulli, M. C., Dell'Angelo, J., and D'Odorico, P.: Global agricultural economic water scarcity, Sci. Adv., 6(18), eaaz6031, https://doi.org/10.1126/sciadv.aaz6031, 2020.

Ruess, P. J., Konar, M., Wanders, N., and Bierkens, M.: Irrigation by crop in the Continental United States from 2008 to 2020, Water Resour. Res., 59(2), e2022WR032804, https://doi.org/10.1029/2022WR032804, 2023.

Siebert, S. and Döll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, J. Hydrol., 384, 198-217, https://doi.org/10.1016/j.jhydrol.2009.07.031, 2010.

Sloat, L. L., Davis, S. J., Gerber, J. S., Moore, F. C., Ray, D. K., West, P. C., and Mueller, N. D.: Climate adaptation by crop migration, Nat. Commun., 11(1), 1243. https://doi.org/10.1038/s41467-020-15076-4, 2020.

Vanuytrecht, E., Raes, D., Steduto, P., Hsiao, T. C., Fereres, E., Heng, L. K., Vila, M. G., and Moreno, P. M.: AquaCrop: FAO's crop water productivity and yield response model, Environ. Modell. Softw., 62, 351-360, https://doi.org/10.1016/j.envsoft.2014.08.005, 2014.

Wang, W., Zhuo, L., Li, M., Liu, Y., and Wu, P.: The effect of development in water-saving irrigation techniques on spatial-temporal variations in crop water footprint and benchmarking, J. Hydrol., 577, 123916, https://doi.org/10.1016/j.jhydrol.2019.123916, 2019.

Xie, G., Han, D., Wang, X., and Lü, R.: Harvest index and residue factor of cereal crops in China, Journal of China agricultural university, 16, 1-8, 2011(in Chinese).

Xue, J. and Ren, L.: Evaluation of crop water productivity under sprinkler irrigation regime using a distributed agro-hydrological model in an irrigation district of China, Agric. Water Manage., 178, 350-365, https://doi.org/10.1016/j.agwat.2016.10.003, 2016.

Yue, Z., Ji, X., Zhuo, L., Wang, W., Li, Z., and Wu, P.: Spatiotemporal responses of the crop water footprint and its associated benchmarks under different irrigation regimes to climate change scenarios in China, Hydrol. Earth Syst. Sci., 26, 4637-4656, https://doi.org/10.5194/hess-26-4637-2022, 2022.

Zhang, C., Dong, J. and Ge, Q.: Mapping 20 years of irrigated croplands in China using MODIS and statistics and existing irrigation products, Sci. Data, 9, 407, https://doi.org/10.1038/s41597-022-01522-z, 2022a.

Zhang, F. and Zhu, Z.: Harvest index for various crops in China, Scientia Agricultura Sinica, 23, 83-87, 1990 (in Chinese).

Zhang, L., Zhang, K., Zhu, X., Chen, H., and Wang, W.: Integrating remote sensing, irrigation suitability and statistical data for irrigated cropland mapping over mainland China, J. Hydrol., 613, 128413, https://doi.org/10.1016/j.jhydrol.2022.128413, 2022b.

Zhu, X., Zhu, W., Zhang, J., and Pan, Y.: Mapping irrigated areas in China from remote sensing and statistical data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(11), 4490-4504, https://doi.org/10.1109/JSTARS.2013.2296899, 2014.

Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: The effect of inter-annual variability of consumption, production, trade and climate on crop-related green and blue water footprints and inter-regional virtual water trade: A study for China (1978–2008), Water Res., 94, 73-85, https://doi.org/10.1016/j.watres.2016.02.037, 2016a.

Zhuo, L., Mekonnen, M. M., Hoekstra, A. Y., and Wada, Y.: Inter-and intra-annual variation of water footprint of crops and blue water scarcity in the Yellow River basin (1961–2009), Adv. Water Resour., 87, 29-41, https://doi.org/10.1016/j.advwatres.2015.11.002, 2016b.

Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: Benchmark levels for the consumptive water footprint of crop production for different environmental conditions: a case study for winter wheat in China, Hydrol. Earth Syst. Sci., 20, 4547-4559, https://doi.org/10.5194/hess-20-4547-2016, 2016c.

Authors' responses to Referee #2' s comments

This paper calculated the WFCP from physical model and then analyzed the water footprints of major crops in Chine between 2001 and 2018, it is a useful topic and the analysis workflow is fine. But my major concern is about the quality of the data, I think there are data with better quality can be used as input and generate more reliable dataset for analysis. Therefore, my opinion is that the paper needs a major revision and a new version of the dataset should be generated before the analysis as the current version make no sense in scientific community.

Responses: Thank you very much for the positive words on the significance of the dataset and recognition of the workflow. We are very grateful for your kind concerns about the input data quality and related issues. We totally agree with the comments. Actually, we tested carefully the feasibility of all the well-acknowledged existing and new data to ensure the best data quality as we can. Please kindly refer to our point-to-point responses as followed.

1. Some concerns about the input dataset:

(1) Irrigated and rainfed crop areas between 2001 and 2018 is generated from the combination of MIRCA 2000 with 5 arcmin resolution and the year book data, but currently, there are actually a lot of new dataset about irrigated cropland from satellite data, for example, Zhang et.al (https://www.sciencedirect.com/science/article/abs/pii/S0022169422009830) 2022, and Zhang et.al 2022 (https://www.nature.com/articles/s41597-022-01522-z). Maybe there are still other data products, I just searched in Google and found the above two datasets. All these irrigation data could be verified and then resampled to 5 arcmin, and used as input data. Basically, the spatial distribution of irrigation land generated from satellite are more reliable than a 15-year-old coarse resolution data and yearbook. As the irrigation cropland is the basement of all the following data generation, please keep the pace with new datasets in the geo-agricultural community.

Responses: Thank you for your comments. Prior to initiating this study, we screened the required crop planting area data based on the following criteria: distinguishing crop types, separating irrigated and rainfed areas, long-term temporal resolution, and high spatial resolution. MIRCA2000 was selected because it meets the objectives of this study. This is also the reason that the MIRCA2000 dataset is still the most widely used for crop water consumption or requirement dataset making (e.g., Hoch et al., 2023; Li et al., 2023; Ruess et al., 2022; Lutz et al., 2022; Liu et al., 2022; Chiarelli et al., 2022; Chiarelli et al., 2020; Rosa et al., 2020). In order to improving the reliability of the input land use data, the proportional scaling approach based on the MIRCA2000 dataset have been applied in numerous studies in this field (Sloat et al., 2020; Yue et al., 2022; Mialyk et al., 2022; Wang et al., 2019).

Table R2 presents the crop planting area and irrigated area data products, including those the Referee#2 recommended. Recent years have witnessed the emergence of numerous long-term and high-resolution irrigation area datasets for China, thanks to the combined application of remote sensing technology and machine learning approaches. However, these datasets do not differentiate between irrigated and rainfed cropping systems, and do not contain crop-specific planting information. These deficiencies fail to fulfill the original intentions of this study design.

Global scale data compensates the aforementioned deficiencies to some extent. However, it is worth mentioning that existing global databases have certain limitations, including a limited range of crop types and intermittent time series. For instance, the SPAM dataset is only publicly available for a few specific years 2000, 2005 and 2010, and interpolation is still required to fill in the gaps. The GFSAD1KCD dataset encompasses a smaller variety of crop types.

We further compared planting areas of other crops in SPAM and our data provincially and in grids (Figures R5 and R6). It is evident that there is a high R^2 at the provincial scale. The differences at the grid scale can be attributed to discrepancies in the identification of gridded land use between the MIRCA2000 and SPAM. According to Figures R7 and R8, the planting area data for sorghum, millet, barley, and sugar beets in the GAEZ+ exhibit significant deviations from this study, both at the provincial and grid scales. However, it should be emphasized that all crop planting area data in this study have been calibrated against statistical data at the provincial scale, implying an underestimation of the planting area for the mentioned crops in the GAEZ+.

Given the above analysis, we used MIRCA2000 to estimate crop planting areas by water supply and irrigation practices. Other datasets have potential but based on our screening criteria, we made assessments of available data and adopted rigorous methods to ensure the reliability of the applied data.

In lines 100-101 of the revision, we point out that further details about the planting area data selection and comparative results have been provided in the Supplementary data and methods. In section 4.1 of the revision (Lines 404-419), we have added comparisons of the data used in this study with the SPAM and GAEZ+ data products. In the Planting area selection of Supplementary data and methods, we list the inventory of the latest Chinese crop irrigated area database, explain the principles for selecting crop planting area data in detail, and supplement the comparative results of crop planting area data.

(2) The crop phenology for the major crop need to be further specific. For example, the winter wheat, the planting date of winter wheat is Oct 15, and all winter wheat are assumed to have the same phenology characters. But this is not the situation in practice, and the phenology of winter have significant difference even in North China Plain with different sowing, emergence and dormancy date. It is essential to use more detailed phenology data as input.

Responses: Thank you for your comments. As you rightly pointed out, we recognize the significance of planting date (PD). Currently, there are some phenology datasets for major Chinese crops. The dataset generated by Luo et al. (2020), which only encompasses three major crops wheat, rice and maize. As indicated on the website of China Meteorological Data Service Center, the "Ten-day Values Dataset of Crop Growth and Development and Soil Moisture Content" they published has not gone through quality control and is of average quality. Here we mainly used phenology data published by Chen et al. (1995) for model input since it is widely used and its reliability is validated (Long et al., 2010; Cao et al., 2014; Ding et al., 2020).

We discussed the effects of inputs like PD on water footprint of crop production (WFCP) estimation in the discussion, "The effect of PD differed for each crop, and advancing or delaying it exposed crops to completely different rain and heat conditions .... in future research, attention to the collection and organisation of basic data can play a positive role in the improvement of the model mechanism and accuracy of the output."
As shown in Table R3, our previous study conducted a sensitivity analysis of WFCP to PD at the site scale. The results indicated that when PD shifts ±10 days, the change in WFCP remains within 4%. With PD shifts of ±20 days, the variation in WFCP is under

8.5% (Li et al., 2022).

Table R3. Sensitivity analysis of water footprint of crop production to planting date.

Crop	-20 days	-15 days	-10 days	-5 days	5 days	10 days	15 days	20 days
Wheat (297 sites)	-5.9%	-4.5%	-3.0%	-1.4%	2.0%	3.9%	5.6%	7.5%
Maize (304 sites)	-0.4%	0.0%	0.2%	0.3%	0.2%	-0.1%	-0.6%	-1.5%
Rice (480 sites)	0.4%	0.5%	0.5%	0.4%	-0.5%	-1.1%	-2.3%	-3.6%
Soybean (299 sites)	6.3%	5.0%	3.5%	1.8%	-1.9%	-4.0%	-6.2%	-8.5%

Note: “-” means advance planting date. “+” means delay planting date. Sources: Li et al., (2022).

In short, PD's effect on WFCP estimation is acceptable since crop water consumption is primarily concentrated in crop development (L2) and mid-season (L3) stages. In this study for instance, with over 13 crops having L2 and L3 water consumption proportions exceeding 80% (Figure R9). Therefore, minor shifts in PD forward or backward have relatively small influences on WFCP.

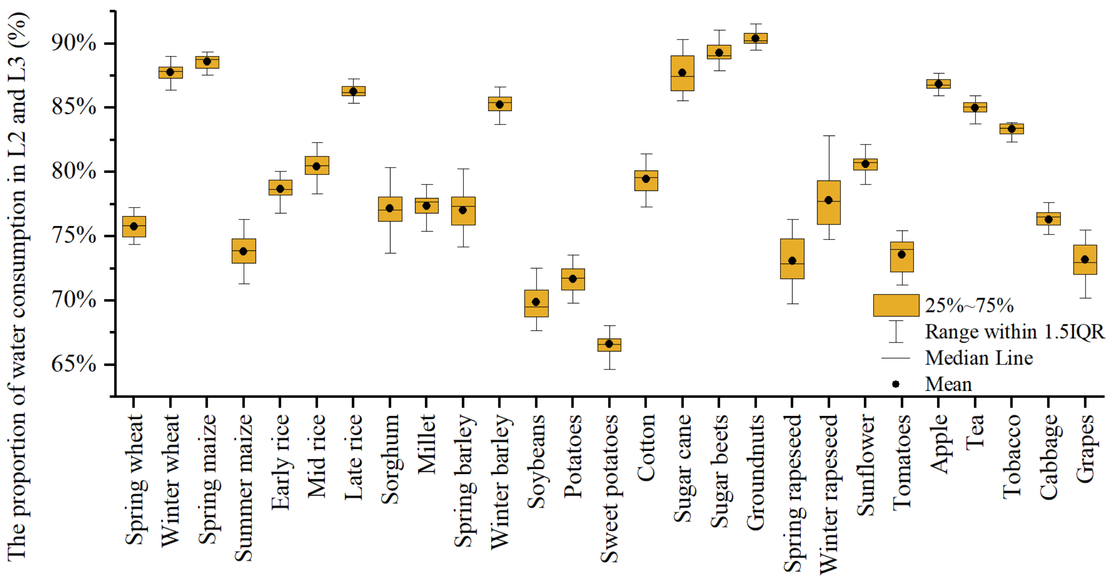

Figure R9. The proportions of crop water consumption in stages L2 and L3 for various crops.

In addition to considering fixed planting dates and crop duration, we conducted a sensitivity analysis of the effect of growing degree days (GDD) on the quantification of WFCP (Zhuo et al., 2014). The GDD measures heat units during crop growth, greatly improves the accuracy of expressing and predicting crop phenological cycles compared to other methods like calendar year or days (McMaster and Wilhelm, 1997). The results indicated that when wheat PD was shifted 30 days earlier than the reference date, yield and WFCP decreased by 0.25% and 0.3% respectively. When rice planting was delayed by 30 days, yield and WFCP reduced by 0.2% and 9.3% respectively. Therefore, under constant GDD, yield and WFCP showed low sensitivity to changes in crop PD.

In section 2.1.3 of the revision (lines 120-121), we point out that the crop phenology in this study underwent strict screening. In section 4.2 of the revision (lines 440-445), we have supplemented the conclusions from our previous study on the sensitivity analysis of WFCP to fixed planting dates, crop duration and GDD. The results of the sensitivity analysis have been appended to the phenology selection of Supplementary data and methods.

1. Data comparison

Figure 9, Please include RMSE in the Figure, when analyzing quantitative results, RMSE is more commonly used to evaluate overestimation and underestimation. Although it seems the value of R square is good here, I found the number of high ET value is small, this indicates the high R square value in this Figure do not make sense either, Please include more high ET-value pixels in the analysis.

Responses: Thank you for your comments. The lack of high ET values in Figure 9 is attributed to the preliminary screening applied when selecting data for comparison in this study. As described in Section 2.3.2 (Lines 230-236): Because of the spatially fragmented nature of crop cultivation, we conducted remote sensing validation according to the Chinese Agricultural Cropping System to reduce the interference of non-agricultural land with the validation results. We selected grids in which the sum of planted areas was greater than 5 kha (> 50% of a single grid) and greater than 10 kha (>100% of a single grid) for single- and multi-crop regions, respectively. In terms of the time span, 19 of the 21 crops studied experienced growth periods from April to August; therefore, these five months were set as the validation interval in terms of total evapotranspiration.

As suggested, the RMSE has been added in the revised Figure 9 (Line 435) and other figures (Figures 7 and 8). Correspondingly, the computational methodology for RMSE has been appended in section 2.3.4 (Lines 248-255).

1. Separation of E and T

The separation of E and T can not be verified because there is no validation data. Please address this in discussion section, pay attention to include the uncertainty analysis of E and T separation.

Responses: As mentioned in reply to reviewer#1's question 2-A, the water consumption results of current study were validated against the dual-source (PML-V2(China)) and single-source (SEBAL) remote sensing products. Notably, the SEBAL products (Cheng et al., 2021) solely comprised aggregate evapotranspiration figures, whereas the PML-V2 products (He et al., 2022) separated land surface evapotranspiration into vegetation transpiration (Ec), soil evaporation (Es), evaporation of intercepted precipitation (Ei), and water body evaporation (Ew). In this study, Ec + Es, Ec and Es were compared with the generated ET, E and T data, respectively.

Comparative analysis in Figure R1 and Figure R2 revealed stronger agreement between the simulated evapotranspiration and the PML-V2 products (R^2=0.49 – 0.85, *RMSE*= 5.82 – 12.12 Mm^3) than those with the SEBAL products (R^2= 0.44 – 0.75, *RMSE* =8.51 – 15.82 Mm^3), although both comparisons demonstrated robust overall consistency. The validation results of soil evaporation are presented in Figure R3. The simulated E were marginally lower than the PML-V2 products (R^2=0.22 – 0.70, *RMSE*=3.25 – 6.65 Mm^3), owing to the current study calculating E exclusively for the planted regions of 21 crops, whereas the PML-V2 disregarded land use types during E estimation. Comparative analysis of crop transpiration in Figure R4 indicated that our simulated values were higher than the PML-V2 products which deducted canopy evaporation (R^2=0.38 – 0.69, *RMSE*= 6.04 – 10.35 Mm^3). Overall, considering the differences in basic input data, spatiotemporal resolution and calculation methods, the evapotranspiration, evaporation, and transpiration data products produced in this study showed acceptable results when compared with various remote sensing products, given the discrepancies exhibited.

The datasets and data preprocessing procedures implemented for validation purposes were delineated in section 2.3.2 (Lines 230-241) of the revision. Comparative analysis between the simulated values from the current study and established remote sensing datasets was presented in lines 421-435 and the validation results of the Supplementary data and methods.

Line 183, Can you show the map of humid (AI > 0.5) and arid (AI < 0.5) zones?

Responses: Thank you for your comments. We sincerely apologize for the confusion caused by the unclear illustration, which inconvenienced your review. The climate zones are defined as the ratio of rainfall to reference evapotranspiration, where humid (AI > 0.5) and arid (AI < 0.5) zones are delineated.

In Figure 6 (Lines 368-369), we differentiated the arid and humid zones with wine red boundary lines, and the specific locations of arid and humid areas can be found in Figure 6a.

Line 194, What remote sensing data have you used for comparison?

Responses: Thank you for your comments. We sincerely apologize for the oversight on our part, which caused inconvenience to your review. The water consumption results of current study were validated against the dual-source (PML-V2(China)) and single-source (SEBAL) remote sensing products.

In section 2.3.2 (Lines 230-241) of the revised manuscript, we introduced the datasets and data preprocessing methods used for validation.

Line 316~319, I think the two factors addressed here do not make sense as the data used in the analysis have significant shortage, please use your new version to compare.

Responses: Thank you for your comments. Prior to the simulation, we conducted rigorous screening and reliability validation of the input data based on principles of accuracy and representativeness, using the currently optimal available data. We recognize that the accuracy of all model studies, including AquaCrop, is contingent on model mechanisms and input data. Therefore, laying emphasis on collection and collation of fundamental data in future studies will play a positive role in improving model mechanisms and output precision.

It should be noted that the WATNEEDS represents crop water requirements, while WFCP in our study refers to crop water consumption during the growing period, as has been elaborated in lines 51-52 of the Introduction. Therefore, we think the two factors addressed here are valid, that the primary factors leading to discrepancies between the datasets are the differences in simulation mechanisms (crop water requirements vs. crop water consumption) and irrigation practices (whether distinguishing irrigation practices).

In section 4.2 (lines 437-494) of the revision and the Supplementary data and methods section, the sensitivity and uncertainty analysis and the screening procedure implemented for the input data in the current study were rewritten in the interest of guaranteeing the credibility of the outcomes.

References

Cao, X., Wu, P., Wang, Y., and Zhao, X.: Assessing blue and green water utilisation in wheat production of China from the perspectives of water footprint and total water use, Hydrol. Earth Syst. Sci., 18(8), 3165-3178, https://doi.org/10.5194/hess-18-3165-2014, 2014.

Chen, Y., Guo, G., Wang, G., Kang, S., Luo, H., and Zhang, D.: Main crop water requirement and irrigation of China, Hydraulic and Electric Press, Beijing, 1995.

Cheng, M., Jiao, X., Li, B., Yu, X., Shao, M., and Jin, X.: Long time series of daily evapotranspiration in China based on the SEBAL model and multisource images and validation, Earth Syst. Sci. Data, 13, 3995-4017, https://doi.org/10.5194/essd-13-3995-2021, 2021.

Chiarelli, D. D., D'Odorico, P., Müller, M. F., Mueller, N. D., Davis, K. F., Dell'Angelo, J., Penny, G., and Rulli, M. C.: Competition for water induced by transnational land acquisitions for agriculture, Nat. Commun., 13(1), 505, https://doi.org/10.1038/s41467-022-28077-2, 2022.

Chiarelli, D. D., Passera, C., Rosa, L., Davis, K. F., D'Odorico, P., and Rulli, M. C.: The green and blue crop water requirement WATNEEDS model and its global gridded outputs, Sci. Data, 7, 273, https://doi.org/10.1038/s41597-020-00612-0, 2020.

Ding, Y., Wang, W., Zhuang, Q., and Luo, Y.: Adaptation of paddy rice in China to climate change: The effects of shifting sowing date on yield and irrigation water requirement, Agric. Water Manage, 228, 105890, https://doi.org/10.1016/j.agwat.2019.105890, 2020.

He, S., Zhang, Y., Ma, N., Tian, J., Kong, D., and Liu, C.: A daily and 500 m coupled evapotranspiration and gross primary production product across China during 2000–2020, Earth Syst. Sci. Data, 14, 5463-5488, https://doi.org/10.5194/essd-14-5463-2022, 2022.

Hoch, J. M., Sutanudjaja, E. H., Wanders, N., Van Beek, R. L., and Bierkens, M. F.: Hyper-resolution PCR-GLOBWB: opportunities and challenges from refining model spatial resolution to 1 km over the European continent, Hydrol. Earth Syst. Sci., 27(6), 1383-1401, https://doi.org/10.5194/hess-27-1383-2023, 2023.

Li, Z., Feng, B., Wang, W., Yang, X., Wu, P., and Zhuo, L.: Spatial and temporal sensitivity of water footprint assessment in crop production to modelling inputs and parameters, Agric. Water Manage, 271, 107805, https://doi.org/10.1016/j.agwat.2022.107805, 2022.

Li, Z., Sun, S., Zhao, J., Li, C., Gao, Z., Yin, Y., Wang, Y., and Wu, P.: Large Scale Crop Water Footprint Evaluation Based on Remote Sensing Methods: A Case Study of Maize, Water Resour. Res., e2022WR032630, https://doi.org/10.1029/2022WR032630, 2023.

Liu, X., Liu, W., Tang, Q., Liu, B., Wada, Y., and Yang, H.: Global agricultural water scarcity assessment incorporating blue and green water availability under future climate change, Earths Future, 10(4), e2021EF002567, https://doi.org/10.1029/2021EF002567, 2022.

Long, D., Gao, Y., and Singh, V. P.: Estimation of daily average net radiation from MODIS data and DEM over the Baiyangdian watershed in North China for clear sky days, J. Hydrol., 388(3-4), 217-233, https://doi.org/10.1016/j.jhydrol.2010.04.042, 2010.

Luo, Y., Zhang, Z., Chen, Y., Li, Z., and Tao, F.: ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products, Earth Syst. Sci. Data, 12(1), 197-214, https://doi.org/10.5194/essd-12-197-2020, 2020.

Lutz, A. F., Immerzeel, W. W., Siderius, C., Wijngaard, R. R., Nepal, S., Shrestha, A. B., Wester, P., and Biemans, H.: South Asian agriculture increasingly dependent on meltwater and groundwater, Nat Clim Chang, 12(6), 566-573, https://doi.org/10.1038/s41558-022-01355-z, 2022.

McMaster, G. S., and Wilhelm, W. W.: Growing degree-days: one equation, two interpretations, Agr. Forest. Meteorol., 87(4), 291-300, https://doi.org/10.1016/S0168-1923(97)00027-0, 1977.

Mialyk, O., Schyns, J. F., Booij, M. J., and Hogeboom, R. J.: Historical simulation of maize water footprints with a new global gridded crop model ACEA, Hydrol. Earth Syst. Sci., 26, 923-940, https://doi.org/10.5194/hess-26-923-2022, 2022.

Rosa, L., Chiarelli, D. D., Rulli, M. C., Dell'Angelo, J., and D'Odorico, P.: Global agricultural economic water scarcity, Sci. Adv., 6(18), eaaz6031, https://doi.org/10.1126/sciadv.aaz6031, 2020.

Ruess, P. J., Konar, M., Wanders, N., and Bierkens, M.: Irrigation by crop in the Continental United States from 2008 to 2020, Water Resour. Res., 59(2), e2022WR032804, https://doi.org/10.1029/2022WR032804, 2023.

Sloat, L. L., Davis, S. J., Gerber, J. S., Moore, F. C., Ray, D. K., West, P. C., and Mueller, N. D.: Climate adaptation by crop migration, Nat. Commun., 11(1), 1243. https://doi.org/10.1038/s41467-020-

15076-4, 2020.

Wang, W., Zhuo, L., Li, M., Liu, Y., and Wu, P.: The effect of development in water-saving irrigation techniques on spatial-temporal variations in crop water footprint and benchmarking, J. Hydrol., 577, 123916, https://doi.org/10.1016/j.jhydrol.2019.123916, 2019.

Yue, Z., Ji, X., Zhuo, L., Wang, W., Li, Z., and Wu, P.: Spatiotemporal responses of the crop water footprint and its associated benchmarks under different irrigation regimes to climate change scenarios in China, Hydrol. Earth Syst. Sci., 26, 4637-4656, https://doi.org/10.5194/hess-26-4637-2022, 2022.

Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: Sensitivity and uncertainty in crop water footprint accounting: a case study for the Yellow River basin, Hydrol. Earth Syst. Sci., 18(6), 2219-2234, https://doi.org/10.5194/hess-18-2219-2014, 2014.

Authors' responses to Referee #3' s comments

The article " A gridded dataset of consumptive water footprints, evaporation, transpiration, and associated benchmarks related to crop production in China during 2000–2018" established a gridded dataset of monthly green and blue water footprint of crop production (WFCP), evaporation and transpiration et al. based on AquaCrop model. On the whole, the paper has important research significance and much work was done, but there are some shortcomings in the paper, and it is suggested to modify.

Responses: We sincerely appreciate the referee #3's constructive comments and suggestions. We acknowledge the reviewer's opinion that there are areas needing improvement in our paper. In accordance with the reviewer's valuable suggestions, we will make revisions and updates to the manuscript where needed to further enhance its quality and contribution.

1. In Lines 186-187, "The statistical yearbook only has crop production statistics on the provincial level. Therefore, we calibrated crop production at the provincial scale.". The statement is not accurate. Statistics data on the city level can be found in the provincial statistical yearbook.

Responses: Thank you for your comments. You rightly pointed out that the statistical yearbooks contain crop production data at the city level. For years and regions with more comprehensive data, it would be worthwhile to utilize higher-resolution data. It should be noted that although provincial yearbooks include some city-level crop production data, considering the numerous crop types involved in this study, and the division of certain crops by harvest periods (e.g., winter wheat, spring wheat, early rice, mid rice, late rice), there are indeed many instances of missing and incomplete data at the city scale. To ensure data integrity and accuracy, production calibration was carried out at the provincial level. Such provincial calibration has been extensively applied in previous studies (Yue et al., 2022; Zhuo et al., 2016).

The meteorological and soil factors are critical factors affecting the estimation of water footprint of crop production (WFCP) (Zhuo et al., 2014; Tuninetti et al., 2015). By incorporating precipitation, temperature, soil texture and other elements into the fundamental model inputs, the simulation results of this study were still able to reflect heterogeneities in the spatial distribution of crop water consumption, production, and

WFCP well, despite crop production was calibrated at the provincial level. We have ensured these sensitive factors meet the accuracy requirements of this study at temporal and spatial scales.

In section 2.3.1 (Lines 214-228) of the revision, we provided the reason for calibrating production at the provincial level and ensured the validity of the simulated results.

2. In your research, AquaCrop model was used to simulate crop production more than 21 crops. Wheat, corn and rice are the main food crops in China. In crop production, irrigation and fertilization are both important management measure to improve crop yield. However, only irrigation was considered. In my opinion, it is necessary to consider the effects of irrigation and fertilization on crop production in different regions when simulating crop production process using AquaCrop model.

Responses: Thank you for your comments. AquaCrop here only modeled irrigation and excluded other practices like fertilization. As you pointed out, both irrigation and fertilization are critical agronomic measures to improve crop production. It should be noted that AquaCrop was developed by the FAO as a **water-driven model**. Since the focus of this study is to assess the impacts of different water supply and irrigation practices on quantifying WFCP, irrigation practice was selected as the sole simulation factor.

The AquaCrop model adopts a semi-quantitative method to evaluate fertilizer stress. That is, it cannot directly simulate crop response to fertilizer based on plant nutritional demand and soil nutrient content (Akumaga et al., 2017). Research shows AquaCrop performs better without fertilizer stress versus with stress (Adeboye et al., 2021; Wu et al., 2022). In fact, there is a serious overapplication of chemical fertilisers in Chinese farmlands (Chen et al., 2014; Cui and Shoemaker, 2018). The impact of fertilization on crop production was indirectly reflected through calibration against statistical data. Moreover, gridded data is deficient regarding fertilizer varieties and application quantities, more so for crop-specific data. So like past AquaCrop global (Mialyk et al., 2022) and national (Wang et al., 2019) studies, nutrient stress is not considered in simulations.

Certainly, the above assumption has limitations. Establishing high resolution fertilizer application databases is vital for future crop production research. When updating the

WFCP database later, we will enhance model mechanisms to improve accuracy if fertilizer data becomes available.

In section 4.2 (Lines 456-468) of the revision, we have discussed the reasons for not considering fertilization in crop production simulation, and called for establishing high-resolution fertilizer application databases to improve crop production research.

3. As we all know, AquaCrop is a farm-scale model. In your research, the crop production at the provincial scale was calibrated and validated by statistical yearbook. I don't think this is a good idea. It is recommended that some representative sites were selected to verify the AquaCrop parameters.

Responses: Thank you for your comments. Conservative parameters in the basic model inputs originate from the Annex of the Reference manual for the AquaCrop (Raes et al., 2018), while key sensitive parameters are acquired from extensively utilized databases that considered regional differences originally. With regard to your later reviewer comment 5, comprehensive delineation pertaining to the crop parameter selection procedure has been furnished.

In the context outputs validation, similar to previous large-scale footprint studies (Yu et al., 2019; Li et al., 2023), model calibration focuses on the verification of evapotranspiration (ET) against high-precision remote sensing products and production validation at administrative level. Therefore, we validated the simulated crop ET against remote sensing products over the same grids and time, which showed good consistency (Lines 230-241, 421-432). For crop production validation, on one hand it is difficult to obtain site-level measured data, and on the other hand crop production data at the city and county levels contain missing values for the numerous crops involved. As mentioned in the responses to reviewer comment 1, by incorporating precipitation, temperature, soil texture and other elements into the fundamental model inputs, the simulation results of this study were still able to reflect heterogeneities in the spatial distribution of crop water consumption, production, and WFCP well, despite crop production was calibrated at the provincial level (Lines 223-228).

4. Whether the parameters are consistent in AquaCrop model under different irrigation methods?

Responses: Thank you for your comments. As stated in the AquaCrop model user manual, different irrigation practices are distinguished by the parameter of surface wetness fraction, while other parameters remain consistent (Raes et al., 2018). This irrigation practices differentiation approach is commonly used before (Pereira et al., 2015; Wang et al., 2019; Chibarabada et al., 2020; Li et al., 2022; Yue et al., 2022).

Specifically, as elaborated in Section 2.2.5, Different irrigation practices indirectly affect water consumption during the growth period due to differences in the fraction of the surface wetted (fw) by each method. where the fw-values used for furrow, sprinkler, and micro-irrigation were 80%, 100%, and 40%, respectively.

In section 2.2.5 (Lines 193-196) of the revision, we have supplemented the applicability of the method for the model used to distinguish irrigation practices.

5. Lines 391-392. The crop parameters do not vary with simulation time and space. It is inaccurate for the whole Country to use only one set of crop parameters to simulate crop yield. In my opinion, different crop parameters should be used for each province.

Responses: Thank you for your comments. Due to inappropriate narration, our description of crop parameters in the manuscript has caused severe misunderstanding during your review. In fact, strict regional differences were considered during the initial screening of the 21 crops' parameters. The 31 provinces were divided into 8 major regions at the provincial level (Table R1). According to the regional classification results, we adjusted the crop types and planting date parameters. Based on planting and irrigation practices, we tuned the crop rooting depth.

In our previous research (Li et al., 2022), ten input variables or model parameters were selected, including reference evapotranspiration (ETo), crop transpiration coefficient (KcTr), planting date (PD), soil evaporation coefficient (KE), maximum canopy cover (MCC), precipitation (PR), canopy decline coefficient (CDC), planting density of the crop (DC), reference harvest index (HI_0), and normalised water productivity (WP*). The results showed that WFCP is generally more sensitive to **KcTr** and **PD** among the model parameters.

In the Annex of the Reference manual for the AquaCrop (Raes et al., 2018), default values of crop parameters for the 12 crops covered in this study are given, including crop transpiration, biomass production and yield formation, and stresses, totaling 41 parameters. Furthermore, these parameters are further classified based on crop sensitivity as conservative generally applicable (including **KcTr**), conservative for a given species but can or may be cultivar specific, dependent on environment and/or management and cultivar specific. The conservative parameters are generally applicable and remain unchanged across a wide spectrum of conditions, including different climatic and geographic locations, crop cultivars and genotypes, as well as variable soil moisture stress statuses. Once calibrated, these identical parameters would be utilized without further modification.

Regarding the other sensitive parameter **PD**, the phenology dataset generated by Luo et al. (2020) only included three major crops wheat, rice and maize. As indicated on the website of China Meteorological Data Service Center, the "Ten-day Values Dataset of Crop Growth and Development and Soil Moisture Content" they published has not gone through quality control and is of average quality. In this study, we primarily used the phenology data published by Chen et al. (1995) as model input, because this dataset has been widely applied and its reliability has been validated (Long et al., 2010; Cao et al., 2014; Ding et al., 2020).

Table R4. The proportions of crop water consumption in stages L2 and L3 for various crops.

Crop	Proportion	Crop	Proportion
Spring wheat	75.8%±0.8%	Cotton	79.5%±1.0%
Winter wheat	87.8%±0.6%	Sugar cane	87.7%±1.4%
Spring maize	88.6%±0.5%	Sugar beets	89.3%±0.7%
Summer maize	73.8%±1.3%	Groundnuts	90.4%±0.6%
Early rice	78.7%±0.8%	Spring rapeseed	73.1%±1.8%
Mid rice	80.4%±1.0%	Winter rapeseed	77.8%±2.2%
Late rice	86.3%±0.4%	Sunflower	80.6%±0.6%
Sorghum	77.2%±1.6%	Tomatoes	73.6%±1.3%
Millet	77.4%±0.9%	Apple	86.8%±0.5%
Spring barley	77.0%±1.4%	Tea	85.0%±0.7%
Winter barley	85.2%±0.9%	Tobacco	83.3%±0.5%
Soybeans	69.9%±1.4%	Cabbage	76.3%±1.4%
Potatoes	71.7%±1.1%	Grapes	73.2%±1.4%
Sweet potatoes	66.6%±1.0%		

In this study for example, the proportions of crop water consumption in the crop development (L2) and mid-season (L3) stages for various crops are shown in Table R4, with more than 13 crops having L2 and L3 water consumption proportions over 80%. As shown in Table R3, when PD varies ±10 days, the change in WFCP is within 4%. When PD varies ±20 days, the change in WFCP is within 8.5%. The impact of PD on WFCP estimation is acceptable, because the crop water consumption during the growing season is mainly concentrated in L2 and L3 stages. In these processes, crop water consumption increases substantially with the rise of canopy cover and crop growth demand. Therefore, based on the parameter adjustments that have been implemented at the provincial level, minor shifts in PD forward or backward have relatively small influences on WFCP.

We obtained these key parameters like reference harvest index, crop growth stages, and maximum root depth for this study by referring to the literature (Allen et al., 1998; Vanuytrecht et al., 2014; Xie et al., 2011; Zhang and Zhu, 1990; Hoekstra and Chapagain, 2006). These data have been validated to be reliable and applicable in large-scale studies (Zhuo et al., 2016; Wang et al., 2019). Due to data limitations, the remaining parameters such as maximum canopy cover, canopy cover decline coefficient, canopy growth coefficient were assigned the mean values within the reference range provided in the Annex of the Reference manual for the AquaCrop. Although this approach may overlook certain potential variations, the use of mean values generally captures the central tendency of the data.

In summary, unlike small-scale studies at site level that emphasize region-specific measured parameters for model simulation, large regional-scale studies often adopt literature-recommended parameter values during data collection, with greater focus on regional variability and wide adaptability of the parameters. (Hoekstra and Wiedmann 2014; Davis et al., 2017; Mekonnen and Hoekstra 2020; Lutz et al., 2022; Halpern et al., 2022; Liu et al., 2022; Chiarelli et al., 2022; Demay et al., 2023). It was neither practical nor feasible to calibrate crop parameters individually for each grid given the constraints of available data. Nevertheless, we have made every effort to ensure the reliability of the model input parameters within the existing limitations.

Within section 4.2 (Lines 438-455) of the revision and the Supplementary data and methods, we have provided the screening process for key parameters affecting WFCP quantification. In section 4.2 (Lines 473-481) of the revision, we have supplemented explanation on the sources and reliability of other crop parameters. In section 4.2 (Lines 482-490) of the revision, we have discussed the differences between large regional-scale studies and small regional-scale studies in crop parameter selection. It should be reemphasized that the WFCP database generated in this study is updatable annually. Should higher-resolution crop parameter products become available in future, we will update the existing WFCP database accordingly.

6. The discussion needs to be rewritten. The results of this study can be compared with those of others, and the reasons for the similarities and differences can be analyzed. For example, in section 4.1, it is inadequate just compared with the evapotranspiration result with SEBAL model, which is a single-source model. In your research, evaporation and transpiration has been estimated. It will be more meaningful to verify evaporation and transpiration based on the two-source model.

Responses: Thank you for your comments. As mentioned in reply to reviewer#1's question 2-A and reviewer#2's question about "Separation of E and T", the water consumption results of current study were validated against the two-source (PML-V2(China)) and single-source (SEBAL) remote sensing products. Notably, the SEBAL products (Cheng et al., 2021) solely comprised aggregate evapotranspiration figures, whereas the PML-V2 products (He et al., 2022) separated land surface evapotranspiration into vegetation transpiration (Ec), soil evaporation (Es), evaporation of intercepted precipitation (Ei), and water body evaporation (Ew). In this study, Ec + Es, Ec and Es were compared with the generated ET, E and T data, respectively.

Comparative analysis in Figure R1 and Figure R2 revealed stronger agreement between the simulated evapotranspiration and the PML-V2 products (R^2=0.49 – 0.85, *RMSE* =5.82 – 12.12 Mm^3) than those with the SEBAL products (R^2=0.44 – 0.75, *RMSE* =8.51 – 15.82 Mm^3), although both comparisons demonstrated robust overall consistency. The validation results of soil evaporation are presented in Figure R3. The simulated E were marginally lower than the PML-V2 products (R^2= 0.22 – 0.70, *RMSE* =3.25 – 6.65 Mm^3), owing to the current study calculating E exclusively for the planted regions of 21 crops, whereas the PML-V2 disregarded land use types during E estimation. Comparative analysis of crop transpiration in Figure R4 indicated that our simulated values were higher than the PML-V2 products which deducted canopy evaporation (R^2= 0.38 – 0.69, *RMSE* =6.04 – 10.35 Mm^3). Overall, considering the differences in basic input data, spatiotemporal resolution and calculation methods, the evapotranspiration, evaporation, and transpiration data products produced in this study showed acceptable results when compared with various remote sensing products, given the discrepancies exhibited.

In section 2.3.2 (Lines 230-241) of the revised manuscript, we introduced the datasets and data preprocessing methods used for validation. In lines 421-435 and validation results of the Supplementary data and methods, we presented the comparison results between the simulated values in this study and existing remote sensing datasets.

Minor questions:

1. In Line 40, what's the mean of "CWASI".

Responses: Thank you for your comments. We sincerely apologize for the inconvenience caused to your review work due to the use of abbreviations. "CWASI" refers to the abbreviation for the virtual water and water footprint dataset created by Tamea et al. (2021) in line 41.

In line 41 of the revised manuscript, we have rewritten this sentence to avoid confusion.

2. In Lines 130, 135 et al., no spaces before "where …".

Responses: Thank you for your comments.

In the revised manuscript, we have deleted spaces before all instances of "where", in order to ensure formatting accuracy and consistency.

References

Adeboye, O. B., Schultz, B., Adeboye, A. P., Adekalu, K. O., and Osunbitan, J. A.: Application of the AquaCrop model in decision support for optimization of nitrogen fertilizer and water productivity of soybeans, Inf. Process, 8(3), 419-436, https://doi.org/10.1016/j.inpa.2020.10.002, 2021.

Akumaga, U., Tarhule, A., and Yusuf, A. A.: Validation and testing of the FAO AquaCrop model under different levels of nitrogen fertilizer on rainfed maize in Nigeria, West Africa, Agr. Forest Meteorol., 232, 225-234, https://doi.org/10.1016/j.agrformet.2016.08.011, 2017.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration—Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, 300, D05109, FAO, Rome, 1998.

Cao, X., Wu, P., Wang, Y., and Zhao, X.: Assessing blue and green water utilisation in wheat production of China from the perspectives of water footprint and total water use, Hydrol. Earth Syst. Sci., 18(8), 3165-3178, https://doi.org/10.5194/hess-18-3165-2014, 2014.

Chen, X., Cui, Z., Fan, M., Vitousek, P., Zhao, M., Ma, W., Wang, Z., Zhang, W., Yan, X., and Yang, J.: Producing more grain with lower environmental costs, Nature, 514, 486-489, https://doi.org/10.1038/nature13609, 2014.

Chen, Y., Guo, G., Wang, G., Kang, S., Luo, H., and Zhang, D.: Main crop water requirement and irrigation of China, Hydraulic and Electric Press, Beijing, 1995.

Cheng, M., Jiao, X., Li, B., Yu, X., Shao, M., and Jin, X.: Long time series of daily evapotranspiration in China based on the SEBAL model and multisource images and validation, Earth Syst. Sci. Data, 13, 3995-4017, https://doi.org/10.5194/essd-13-3995-2021, 2021.

Chiarelli, D. D., D'Odorico, P., Müller, M. F., Mueller, N. D., Davis, K. F., Dell'Angelo, J., Penny, G., and Rulli, M. C.: Competition for water induced by transnational land acquisitions for agriculture, Nat. Commun., 13(1), 505, https://doi.org/10.1038/s41467-022-28077-2, 2022.

Chibarabada, T. P., Modi, A. T., and Mabhaudhi, T.: Calibration and evaluation of aquacrop for groundnut (Arachis hypogaea) under water deficit conditions, Agr. Forest Meteorol., 281, 107850, https://doi.org/10.1016/j.agrformet.2019.107850, 2020.

Cui, K. and Shoemaker, S. P.: A look at food security in China, NPJ Sci. Food., 2, 4, https://doi.org/10.1038/s41538-018-0012-x, 2018.

Davis, K. F., Rulli, M. C., Seveso, A., and D'Odorico, P.: Increased food production and reduced water use through optimized crop distribution, Nat. Geosci., 10(12), 919-924, https://doi.org/10.1038/s41561-017-0004-5, 2017.

Demay, J., Ringeval, B., Pellerin, S., and Nesme, T.: Half of global agricultural soil phosphorus fertility derived from anthropogenic sources, Nat. Geosci., 16(1), 69-74, https://doi.org/10.1038/s41561-022-01092-0, 2023.

Ding, Y., Wang, W., Zhuang, Q., and Luo, Y.: Adaptation of paddy rice in China to climate change: The effects of shifting sowing date on yield and irrigation water requirement, Agric. Water Manage, 228, 105890, https://doi.org/10.1016/j.agwat.2019.105890, 2020.

Halpern, B. S., Frazier, M., Verstaen, J., Rayner, P.-E., Clawson, G., Blanchard, J. L., Cottrell, R. S., Froehlich, H. E., Gephart, J. A., Jacobsen, N. S., Kuempel, C. D., McIntyre, P. B., Metian, M., Moran, D., Nash, K. L., Többen, J., and Williams, D. R.: The environmental footprint of global food production, Nat. Sustain., 5(12), 1027-1039, https://doi.org/10.1038/s41893-022-00965-x, 2022.

He, S., Zhang, Y., Ma, N., Tian, J., Kong, D., and Liu, C.: A daily and 500 m coupled evapotranspiration and gross primary production product across China during 2000–2020, Earth Syst. Sci. Data, 14, 5463-5488, https://doi.org/10.5194/essd-14-5463-2022, 2022.

Hoekstra, A. Y. and Chapagain, A. K.: Water footprints of nations: water use by people as a function of their consumption pattern, Water Resour. Manag., 21, 35-48, https://doi.org/10.1007/s11269-006-9039-x, 2006.

Hoekstra, A. Y., and Wiedmann, T. O.: Humanity's unsustainable environmental footprint, Science, 344(6188), 1114-1117, https://doi.org/10.1126/science.1248365, 2014.

Li, Z., Feng, B., Wang, W., Yang, X., Wu, P., and Zhuo, L.: Spatial and temporal sensitivity of water footprint assessment in crop production to modelling inputs and parameters, Agric. Water Manage, 271, 107805, https://doi.org/10.1016/j.agwat.2022.107805, 2022.

Li, Z., Sun, S., Zhao, J., Li, C., Gao, Z., Yin, Y., Wang, Y., and Wu, P.: Large Scale Crop Water Footprint Evaluation Based on Remote Sensing Methods: A Case Study of Maize, Water Resour. Res., e2022WR032630, https://doi.org/10.1029/2022WR032630, 2023.

Liu, X., Liu, W., Tang, Q., Liu, B., Wada, Y., and Yang, H.: Global agricultural water scarcity assessment incorporating blue and green water availability under future climate change, Earths Future, 10(4), e2021EF002567, https://doi.org/10.1029/2021EF002567, 2022.

Long, D., Gao, Y., and Singh, V. P.: Estimation of daily average net radiation from MODIS data and DEM over the Baiyangdian watershed in North China for clear sky days, J. Hydrol., 388(3-4), 217-233, https://doi.org/10.1016/j.jhydrol.2010.04.042, 2010.

Luo, Y., Zhang, Z., Chen, Y., Li, Z., and Tao, F.: ChinaCropPhen1km: a high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products, Earth Syst. Sci. Data, 12(1), 197-214, https://doi.org/10.5194/essd-12-197-2020, 2020.

Lutz, A. F., Immerzeel, W. W., Siderius, C., Wijngaard, R. R., Nepal, S., Shrestha, A. B., Wester, P., and Biemans, H.: South Asian agriculture increasingly dependent on meltwater and groundwater, Nat Clim Chang, 12(6), 566-573, https://doi.org/10.1038/s41558-022-01355-z, 2022.

Mekonnen, M. M., and Hoekstra, A. Y.: Blue water footprint linked to national consumption and international trade is unsustainable, Nat. Food, 1(12), 792-800, https://doi.org/10.1038/s43016-020-00198-1, 2020.

Mialyk, O., Schyns, J. F., Booij, M. J., and Hogeboom, R. J.: Historical simulation of maize water footprints with a new global gridded crop model ACEA, Hydrol. Earth Syst. Sci., 26, 923-940, https://doi.org/10.5194/hess-26-923-2022, 2022.

Pereira, L. S., Paredes, P., Rodrigues, G. C., and Neves, M.: Modeling malt barley water use and evapotranspiration partitioning in two contrasting rainfall years. Assessing AquaCrop and SIMDualKc models, Agric Water Manag, 159, 239-254, https://doi.org/10.1016/j.agwat.2015.06.006, 2015.

Raes, D., Steduto, P., Hsiao, T., and Fereres, E.: Reference Manual, Chapter3, AquaCrop Model, Version 6.1, FAO, Rome, 2018.

Tamea, S., Tuninetti, M., Soligno, I., and Laio, F.: Virtual water trade and water footprint of agricultural goods: the 1961–2016 CWASI database, Earth Syst. Sci. Data, 13, 2025-2051, https://doi.org/10.5194/essd-13-2025-2021, 2021.

Tuninetti, M., Tamea, S., D'Odorico, P., Laio, F., and Ridolfi, L.: Global sensitivity of high-resolution estimates of crop water footprint, Water Resour. Res., 51, 8257-8272, https://doi.org/10.1002/2015WR017148, 2015.

Vanuytrecht, E., Raes, D., Steduto, P., Hsiao, T. C., Fereres, E., Heng, L. K., Vila, M. G., and Moreno, P. M.: AquaCrop: FAO's crop water productivity and yield response model, Environ. Modell. Softw., 62, 351-360, https://doi.org/10.1016/j.envsoft.2014.08.005, 2014.

Wang, W., Zhuo, L., Li, M., Liu, Y., and Wu, P.: The effect of development in water-saving irrigation techniques on spatial-temporal variations in crop water footprint and benchmarking, J. Hydrol., 577, 123916, https://doi.org/10.1016/j.jhydrol.2019.123916, 2019.

Wu, H., Yue, Q., Guo, P., Xu, X., and Huang, X.: Improving the AquaCrop model to achieve direct simulation of evapotranspiration under nitrogen stress and joint simulation-optimization of irrigation and fertilizer schedules, Agric Water Manag, 266, 107599, https://doi.org/10.1016/j.agwat.2022.107599, 2022.

Xie, G., Han, D., Wang, X., and Lü, R.: Harvest index and residue factor of cereal crops in China, Journal of China agricultural university, 16, 1-8, 2011(in Chinese).

Yu, C., Huang, X., Chen, H., Godfray, H. C. J., Wright, J. S., Hall, J. W., Gong, P., Ni, S., Qiao, S., Huang, G., Xiao, Y., Zhang, J., Feng, Z., Ju, X., Ciais, P., Stenseth, N. C., Hessen, D. O., Sun, Z., Yu, L., Cai, W., Fu, H., Huang, X., Zhang, C., Liu, H., and Taylor, J.: Managing nitrogen to restore water quality in China, Nature, 567(7749), 516-520, https://doi.org/10.1038/s41586-019-1001-1, 2019.

Yue, Z., Ji, X., Zhuo, L., Wang, W., Li, Z., and Wu, P.: Spatiotemporal responses of the crop water footprint and its associated benchmarks under different irrigation regimes to climate change scenarios in China, Hydrol. Earth Syst. Sci., 26, 4637-4656, https://doi.org/10.5194/hess-26-4637-2022, 2022.

Zhang, F. and Zhu, Z.: Harvest index for various crops in China, Scientia Agricultura Sinica, 23, 83-87, 1990 (in Chinese).

Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: Sensitivity and uncertainty in crop water footprint accounting: a case study for the Yellow River basin, Hydrol. Earth Syst. Sci., 18(6), 2219-2234, https://doi.org/10.5194/hess-18-2219-2014, 2014.

Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: The effect of inter-annual variability of consumption, production, trade and climate on crop-related green and blue water footprints and inter-regional virtual water trade: A study for China (1978–2008), Water Res., 94, 73-85, https://doi.org/10.1016/j.watres.2016.02.037, 2016.