



OPEN

DATA DESCRIPTOR

# The annual dynamic dataset of high-resolution crop water use in China from 1991 to 2019

Minglei Wang<sup>1,2</sup> & Wenjiao Shi<sup>1,2</sup> ✉

Accurately quantifying agricultural water use is essential for protecting agricultural systems from the risk of water scarcity and promoting sustainable water management. While previous studies have innovatively provided spatially explicit analyses or datasets, they tend to have relatively coarse resolution (~8.3 km), and inadequately considered precise localization parameters. Here, we produced annual blue and green water use for 15 main crops with a resolution of 1 km for the years 1991–2019 in China. Firstly, we estimated the yearly crop blue and green water use at the site scale by incorporating more localized input parameters using a dynamic water balance model. Then, the random forest model was combined with site-scale simulation results to generate spatial predictions of blue and green water for each crop from 1991 to 2019. The resulting maps showed a high correlation with locally observed values at field stations ( $R^2 = 0.95$ ), statistics ( $R^2 = 0.77$ ), and exhibited some strengths compared with existing datasets that covered various scales. This dataset can play a key role in devising sustainable water management strategies.

## Background & Summary

With an increasing global population and rapid economic development, ensuring the sustainability of agricultural development while meeting human food demands has become a pressing global concern<sup>1</sup>. Freshwater resources, serving as the foundation of agricultural systems, play an essential role in ensuring food security and achieving sustainable agricultural development<sup>2–4</sup>. As the largest consumer of freshwater resources, global agricultural water use has surged by 124.35% over the past 40 years, constituting more than 70% of global freshwater withdrawals and nearly reaching the upper limit of blue water for agriculture on the planet<sup>5–7</sup>. The scarcity of freshwater presents a significant threat to the sustainability of agricultural systems. Consequently, there is an urgent need to accurately quantify agricultural water use to help stakeholders improve efficiency and formulate policies for sustainable water resource management.

Since the beginning of the 21st century, the gradual improvement of irrigation infrastructure has led to significant advancements in irrigated agriculture in China<sup>8</sup>. This has played a crucial role in promoting the production of agricultural products, with 75% of the grains production and over 90% of the cash crops production being produced on irrigated farmland, which accounts for 50% of cropland in China<sup>9</sup>. However, the inefficiency and expansion of irrigated agriculture have resulted in a continuous increase in agricultural water consumption, accounting for over 65% of China's total freshwater consumption<sup>10</sup>. This surpasses the threshold for maintaining the stability of the Earth's system and has caused environmental issues, such as groundwater depletion in the North China Plain and soil salinization in Northwest China<sup>11–14</sup>. Improving the efficiency of agricultural water use and managing the total amount used pose significant challenges to sustainable development in China. Additionally, the rapid development of urbanization and industrialization in China has escalated the demand for water resources, intensifying the competition between non-agricultural and agricultural water use<sup>15,16</sup>. In the future, China's agricultural sector is poised to experience heightened water stress as a result of shifting precipitation patterns due to climate change and escalating food demand<sup>12,17,18</sup>. Consequently, it is imperative to conduct a quantitative and spatiotemporal dynamic analysis of agricultural water use in China to improve irrigation efficiency, promote the reallocation of water resources, and ensure sustainability within the agricultural system.

<sup>1</sup>Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China. <sup>2</sup>College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100049, China. ✉e-mail: [shiwj@reis.ac.cn](mailto:shiwj@reis.ac.cn)

So far, traditional data on agricultural water consumption in China mainly relies on administrative-level statistics published by government departments, which rarely capture detailed information about specific water requirements for various crops and intra-administrative districts<sup>19,20</sup>. To address this issue, the water footprint framework has been proposed as a potential assessment avenue<sup>3,4</sup>. This framework distinguishes the water requirements for crop growth into blue water, which comes from surface or groundwater, and green water, which comes from rainwater stored in the soil at the crop roots<sup>21</sup>. Various techniques have been developed over the past two decades to estimate crop blue water use (BWU) and green water use (GWU), such as the global crop water model (GCWM)<sup>21</sup> and the grid-based dynamic water balance model<sup>3,4</sup>. These models provide spatially explicit information on water use based on the soil balance principle and crop coefficient approach at a resolution of  $5' \times 5'$  ( $\sim 8.3$  km). For irrigated crops, the model estimates GWU as the potential evapotranspiration under water stress (without irrigation), while the BWU is defined as the difference between potential evapotranspiration with fully satisfied irrigation and GWU. For the rainfed crops, only the GWU is estimated. Although pioneering in providing spatially clear analyses of global crop water use, these studies are primarily based on crop production data from the year 2000 and lack crop-specific and year-specific datasets for the BWU and GWU<sup>4,21</sup>. Subsequently, efforts have been made to develop maps showing crop water use at a grid scale over time. For example, a global dataset of monthly crop-specific BWU and GWU for 23 main crops was obtained using the WATNEEDS model at a  $5'$  resolution for the years 2000 and 2016<sup>22</sup>. Additionally, a crop water use map from 1990 to 2019 at a  $5'$  resolution was produced using a global process-based crop model with upgraded soil water balance techniques<sup>23</sup>.

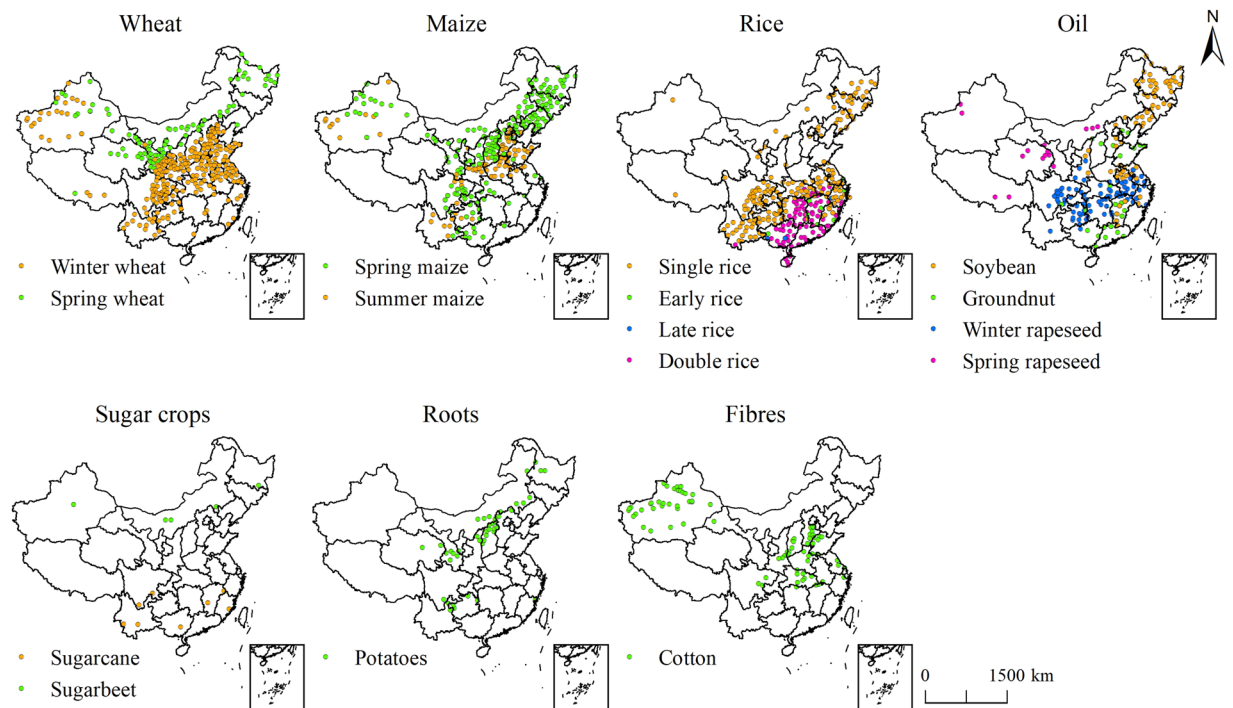
In summary, the previous studies represent a significant contribution to freshwater resource management by offering innovative solutions and datasets to the challenges related to efficient water allocation and sustainable agricultural development. Nevertheless, it is important to acknowledge the existence of several limitations and uncertainties that must be taken into account. Firstly, the current spatial and temporal dynamics of crop water use are usually represented with a coarse spatial resolution ( $\sim 8.3$  km), which may not accurately capture the spatial variability of crop water demand within grid cells. This limitation could impact the precision of water management strategies at a local level. Secondly, the integration of multiple raster data sources in these models can introduce errors that propagate across layers, potentially affecting the overall accuracy of the results. Thirdly, the requisite crop parameters used in the above models, such as crop coefficients ( $K_c$ ) and crop calendars, are often outdated or lack spatial heterogeneity. For instance, Chiarelli *et al.*<sup>22</sup> assessed the crop water requirement for the years 2000 and 2016 using the crop planting and harvesting dates around the year 2000 and internationally used  $K_c$ . Similarly, Mialyk *et al.*<sup>23</sup> estimated the crop water use annually since 1990 based on the static growing periods, i.e., the multi-year average estimates, rather than accounting for year-to-year variability. These studies did not fully account for the dynamic nature of agricultural practices and climate impacts on water use over time. Overall, while these studies offer valuable insights into agricultural water use, addressing the limitations mentioned above is important for enhancing the accuracy and relevance of water management strategies in the context of sustainable agricultural development.

Here, we presented annual blue and green water use for 15 major crops with a resolution of 1 km for the years 1991–2019. Firstly, following the methodology laid out by Siebert *et al.*<sup>21</sup>, a dynamic water balance model was applied at the site scale to estimate the crop blue and green water use for each station (Fig. S1), incorporating more localized input data for improved accuracy. Specifically, the daily reference evapotranspiration is computed by combining the modified Penman-Monteith (P-M) equation recommended by FAO with climate data observed at meteorological stations. Planting and harvested dates for the crops were determined using observational data from agro-meteorological stations in China; and  $K_c$  is adjusted by the specific climatic and growing conditions of each station based on the internationally recognized values. Secondly, we employed a random forest (RF) model to spatialize the simulation results of crop water use at site scale. We validated our outcomes by comparing them with (1) locally observed values across several crops and locations from the Chinese Ecosystem Research Network (CERN) and literature; (2) provincial or municipal statistics from the National Statistics Bureau (NSB) in China; and (3) existing datasets provided by Chiarelli *et al.*<sup>22</sup>, Mialyk *et al.*<sup>23</sup>, Zhang *et al.*<sup>24</sup> and Wang *et al.*<sup>25</sup>. The resulting dataset on crop water use holds significant potential in formulating more informed water management strategies, guiding agricultural practices, and addressing the impacts of climate change on water resources.

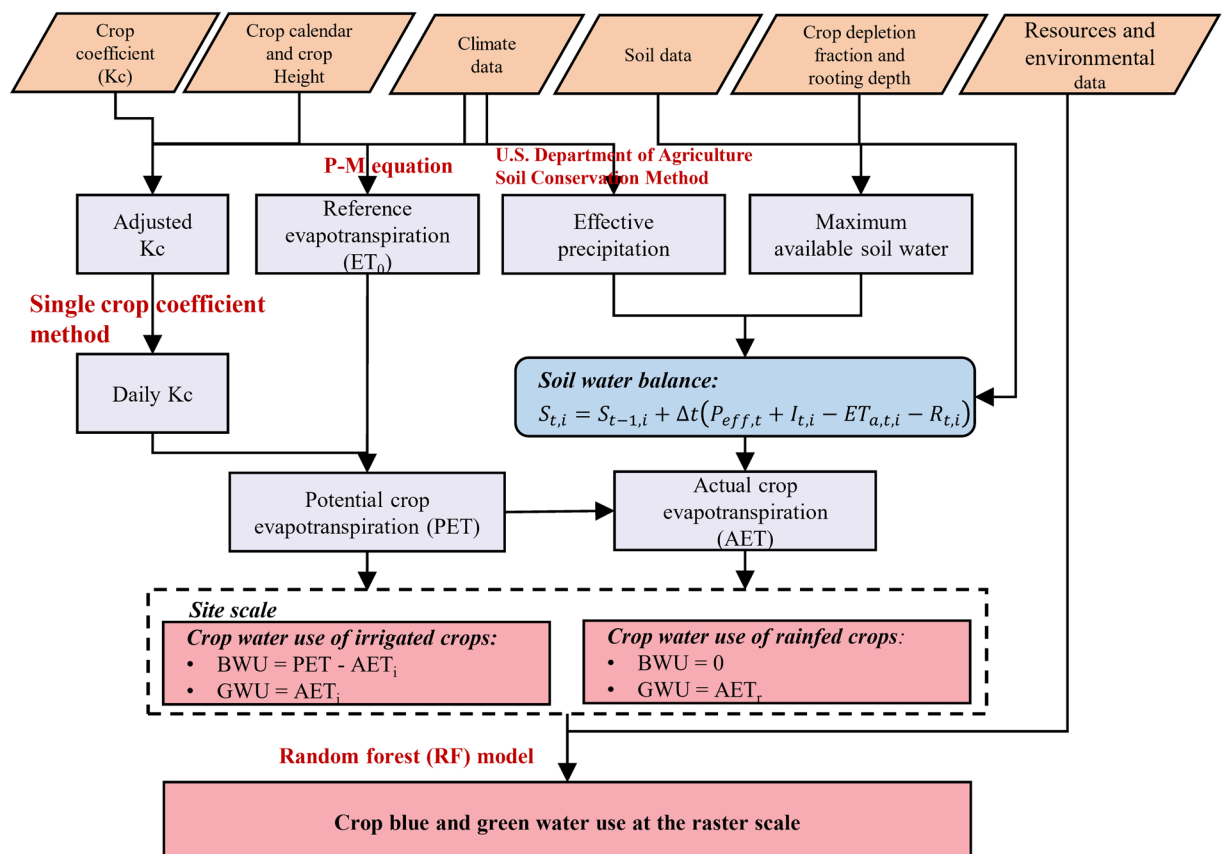
## Methods

The study covers a comprehensive scope, encompassing 31 provinces, municipalities, and autonomous regions within mainland China. Given China's vast agricultural production and diverse crop planting structure, the research estimated the annual blue and green water use for 15 primary crops. These crops include wheat (winter wheat, spring wheat), maize (spring maize, summer maize), rice (early rice, late rice, single-season rice), soybean, potato, sugar beet, sugarcane, groundnut, rapeseed (spring rapeseed, winter rapeseed), and cotton (Fig. 1). These selected crops collectively represent a significant portion of China's crop production and harvested area, accounting for approximately 71.77% of crop production and 76.39% of the harvested area in the country. The water use estimation differentiates between irrigated crops, for which both blue and green water use are calculated, and rainfed crops, for which only green water use is considered (Fig. 2). This comprehensive approach provides valuable insights into the water use patterns within China's agriculture sector, essential for sustainable water resource management and agricultural planning.

**Data collection.** The input data for our study primarily comprises two components. The first section focuses on calculating data related to crop water use, including climate data, crop parameters, crop calendars from planting to harvested dates, crop heights, and soil data. Daily meteorological data such as air temperature, precipitation, relative humidity, wind speed, sunshine hours, and atmospheric pressure of 616 national meteorological



**Fig. 1** Spatial distribution of 15 crops observed by agro-meteorological stations in China.



**Fig. 2** Schematic flow of the crop water use estimation.

Stations	Lat	Lon	Crop type	Time	Reference
Luancheng	37°53'N	114°41'E	Winter wheat, summer maize	2005–2007	CERN
Fengqiu	35°01'N	114°32'E	Winter wheat, summer maize	2005–2008	CERN
Shangqiu	34°35'N	115°34'E	Winter wheat, cotton	2000–2004	CERN
Hailun	47°27'N	126°55'E	Soybean	2004–2006	CERN
Taoyuan	28°55'N	111°27'E	Early rice, late rice	2004–2007	CERN
Changshu	31°33'N	120°38'E	Late rice	2004–2006	CERN
Yingtian	28°15'N	116°55'E	Rice	2004–2006	CERN
Shihezi	44°17'N	85°49'E	Cotton	2010	ref. <sup>34</sup>
Tongzhou	39°36'N	116°48'E	Wheat	2005–2008	ref. <sup>33</sup>
Guangli	35°09'N	114°40'E	Winter wheat	2011–2013	ref. <sup>32</sup>

**Table 1.** Description of the field stations.

stations from 1991 to 2019 were came from the China Meteorological Data Service Centre (CMDSC) (<https://data.cma.cn/>). Subsequently, daily reference evapotranspiration was calculated using the P-M equation. Information on crop growth and development from 1991 to 2013, including planting and harvesting dates and crop heights, was sourced from national agro-meteorological stations (<https://data.cma.cn/>) (Fig. 1). In the absence of pertinent data, the information for the period 2014–2019 has been substituted with the average value for the 2010–2013 interval. Soil data, specifically available water capacity (AWC) at a resolution of 250 m, was sourced from SoilGrids (<https://www.isric.org/explore/soilgrids>). Crop parameters such as crop depletion fraction and rooting depth were referenced from Allen *et al.*<sup>26</sup>. The total available soil water capacity was then calculated by multiplying the AWC and crop-specific rooting depth. Standard crop coefficients ( $K_c$ ) adjusted for China's climatic conditions were derived from Li *et al.*<sup>27</sup>, based on FAO recommendations. The proportion of each development stage (initial, crop development, mid-season, and late season) of various crops to the whole growing period was derived from Fisher *et al.*<sup>28</sup>.

The second section of the data pertains to result validation and includes information such as crop-specific harvested area from the SPAM2010 dataset<sup>29</sup> (<https://mapspam.info/data/>), which estimates crop harvested areas for 42 crops in 2010 with a spatial resolution of 5' based on the Spatial Production Allocation Model (SPAM). Historical data on crop-harvested areas in 31 provinces in mainland China from 1991 to 2019 were obtained from the National Bureau of Statistics<sup>30</sup> (<https://www.stats.gov.cn/sj/ndsj/>). Information on total agricultural water use, the proportion of agricultural irrigation to total water use, and irrigation project efficiency for 31 provinces from 2000 to 2019 were sourced from the China Water Resources Bulletin<sup>31</sup> (<https://mwr.gjzwfw.gov.cn/>). Data on consumptive blue water use at the field scale for crops were collected from various sources including the CERN (<https://www.nesdc.org.cn/>) and literature<sup>32–34</sup> (Table 1). The city-level reported irrigation water and irrigated area for wheat, maize, and rice from 2005–2013 were obtained from Zhou *et al.*<sup>19</sup>. All raster data mentioned above was resampled to a consistent 1 km resolution for analysis.

**Estimating of crop blue and green water use.** The daily potential evapotranspiration for a specific crop is defined as the amount of evapotranspiration occurring during the crop's growth period, regardless of any constraints posed by limited water availability. The primary factors influencing this quantity are the crop type and the growth stage of the crop being evaluated. Daily potential evapotranspiration for crop-specific can be calculated as follows:

$$PET_{t,i} = K_{c,t,i} ET_{0,t} \quad (1)$$

where  $PET_{t,i}$  (mm) is the potential evapotranspiration of crop  $i$  on day  $t$ .  $K_{c,t,i}$  is the crop coefficients of crop  $i$  on day  $t$ .  $ET_{0,t}$  (mm) is the reference evapotranspiration on day  $t$ , which is calculated using the P-M equation.

For the  $K_c$ , the standard values for different development stages are taken from Li *et al.*<sup>27</sup>, which were adjusted based on the  $K_c$  provided by FAO and the specific conditions in China. The standard  $K_c$  values are not universally applicable across China, as they indicate results for climatic conditions with air humidity of approximately 45%, wind speed of approximately  $2 \text{ m s}^{-1}$ , sufficient water supply, effective management, normal growth, and large areas aiming for high yields. Consequently, the crop coefficients were adjusted to reflect the actual climatic conditions and management conditions at each meteorological station. The  $K_c$  of each meteorological station is modified by the following formula:

$$K_{c,i} = K_{c(tab),i} + [0.04 \times (\mu_2 - 2) - 0.004 \times (RHU_{\min} - 45)] \times \left(\frac{h_i}{3}\right)^{0.3} \quad (2)$$

where  $K_{c(tab),i}$  is the standard  $K_c$  of crop  $i$ .  $\mu_2$  ( $\text{m s}^{-1}$ ) is the average wind speed at a height of 2 m at different growth stage.  $RHU_{\min}$  (%) is the average minimum relative humidity at different growth stage.  $h_i$  (m) is the average height of the crop  $i$  at different growth stage.

The actual evapotranspiration of crops is computed as:

$$AET_{t,i} = K_{s,t,i} PET_{t,i} \quad (3)$$

where  $AET_{t,i}$  (mm) is the actual evapotranspiration of crop  $i$  on day  $t$ .  $K_{s,t,i}$  is a dimensionless transpiration reduction factor of crop  $i$  on day  $t$  that is calculated based on a daily function of the actual and maximum soil moisture in the root zone. The equation is as follows:

$$K_{s,t,i} = \begin{cases} \frac{S_{t,i}}{(1 - p_{t,i}) S_{\max}}, & S_{t,i} < (1 - p_{t,i}) S_{\max} \\ 1, & \text{otherwise} \end{cases} \quad (4)$$

where  $S_{t,i}$  is the actual soil moisture content, and the  $S_{\max}$  is the maximum available soil water in the effective root zone, which is computed by multiplying the available water capacity by crop rooting depth, as in Allen *et al.*<sup>26</sup>.  $p_{t,i}$  is the fraction of  $S_{\max}$  that crop  $i$  can extract from the root zone without suffering water stress, and is calculated as follows:

$$p_{t,i} = P_{std,i} + 0.04(5 - PET_{t,i}) \quad (5)$$

where  $P_{std,i}$  is the depletion fraction of crop  $i$  derived from Allen *et al.*<sup>26</sup>.

The  $S_{t,i}$  is simulated by performing soil water balances, and the equation is expressed in the following form:

$$S_{t,i} = S_{t-1,i} + \Delta t (P_{eff,t} + I_{t,i} - ET_{a,t,i} - R_{t,i}) \quad (6)$$

where  $S_{t-1,i}$  is the actual soil moisture content on day  $t - 1$ .  $\Delta t$  is equal to 1.  $P_{eff,t}$  (mm) is the effective precipitation on day  $t$ , which is calculated using the U.S. Department of Agriculture Soil Conservation Method<sup>35,36</sup>:

$$P_{eff,t} = \begin{cases} \frac{Pre_t(4.17 - 0.2Pre_t)}{4.17}, & Pre_t < 8.3 \text{ mm} \\ 4.17 + 0.1Pre_t, & \text{otherwise} \end{cases} \quad (7)$$

where  $Pre_t$  is the precipitation on day  $t$ .  $I_{t,i}$  (mm) symbolizes the supplemental irrigation water applied on day  $t$  for irrigated crops exclusively. Lastly,  $R_{t,i}$  (mm) signifies the runoff volume derived using the specified formula for the day  $t$ .

$$R_{t,i} = (P_{eff,t} + I_t) \times \left( \frac{S_{t,i}}{S_{\max}} \right)^\gamma \quad (8)$$

where  $\gamma$  represents the parameter related to crop management conditions. In irrigated areas, the value is set to 3, while in rainfed areas, it is set to 2<sup>21</sup>.

The preceding steps enable the calculation of the actual evapotranspiration ( $AET_{t,i}$ ) for each crop on a daily basis under water stress, which represents the green water use ( $GWU_{t,i}$ ). Blue water use ( $BWU_{t,i}$ ) is characterized by the difference between the potential evapotranspiration of crops ( $PET_{t,i}$ ) without water stress and the  $AET_{t,i}$ , and was only considered for irrigated areas. Subsequently, the total amount of BWU and GWU for each crop's growth are determined by summing their individual values across every day during the growing season.

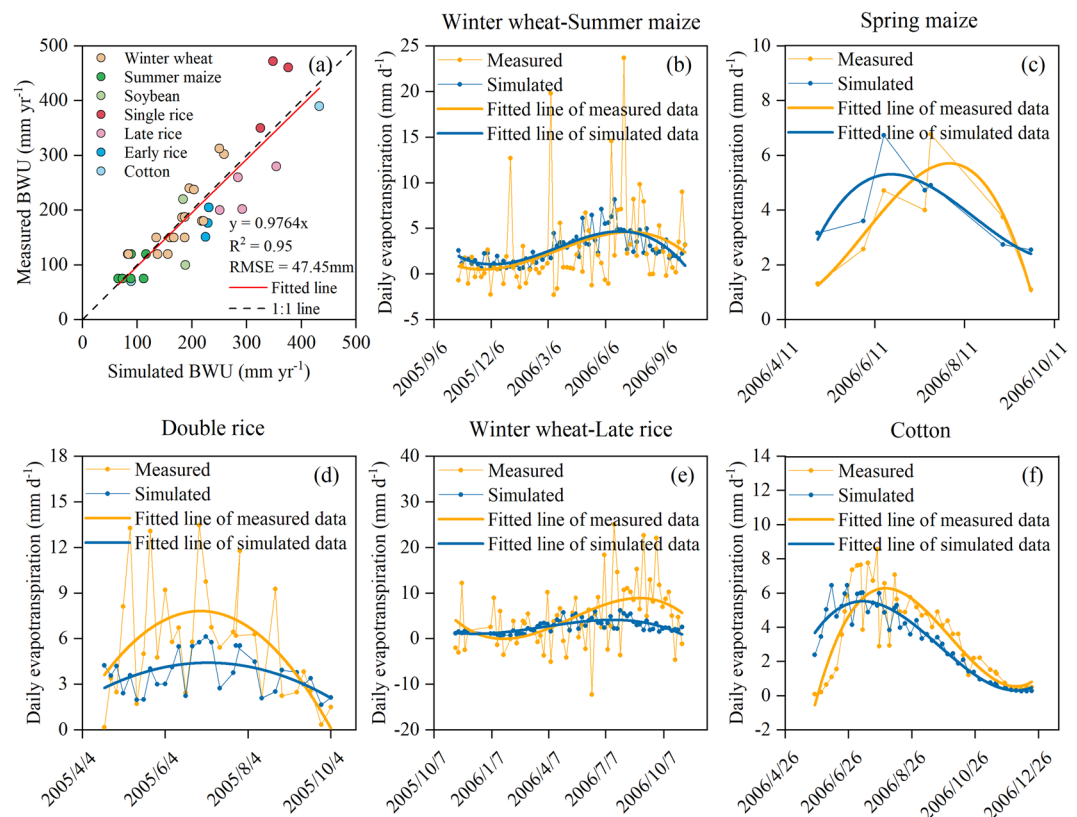
$$GWU_i = \sum_{t=1}^n AET_{t,i} \quad (9)$$

$$BWU_i = \sum_{t=1}^n (PET_{t,i} - AET_{t,i}) \quad (10)$$

where  $GWU_i$  (mm yr<sup>-1</sup>) and  $BWU_i$  (mm yr<sup>-1</sup>) are the green water use and blue water use of crop  $i$ .  $n$  indicates the number of days during the growing period of the crops.

**Spatialization.** A random forest (RF) model was developed to generate spatial predictions of crop blue and green water use at a high spatiotemporal resolution. Compared with the traditional grid-based dynamic water balance model, the proposed method effectively mitigates cross-layer propagation errors that can arise from integrating multiple grid data source, especially when large amounts of climate data with high temporal resolution (e.g., daily scale) are used in this model. In addition, as far as we know, the directly estimating spatially the blue and green water use is further limited by the fact that some climate data products with high spatial resolution have not yet been released. Therefore, we combined site-scale simulation with RF-based spatial prediction to simulate crop blue and green water requirements. The RF model, a form of binary decision rule model that incorporates machine learning technology, has been consistently validated in numerous studies as an effective tool for spatial prediction. In this study, blue and green water use were defined as the dependent variables. We selected several environmental variables, including temperature, precipitation, potential evapotranspiration, available water capacity (AWC), digital elevation model (DEM), and soil components such as sand, silt, clay, and soil type. These variables were analysed using RF modeling with 10-fold cross-validation. The model was set up with 1500 decision trees, and each node in the trees used three variables. The dataset, comprising records of blue and green water use, was split into ten equal





**Fig. 3** Comparison of crop blue water use (BWU) and evapotranspiration between measured data at field stations and corresponding values in this study. **(a)** Scatter plots of simulated crop BWU and measured BWU for 7 crops. Scatter plots of simulated evapotranspiration and measured evapotranspiration for the cropping system in **(b)** winter wheat-summer maize, **(c)** spring maize, **(d)** double rice, **(e)** winter wheat-late rice, and **(f)** cotton. The measured data at field stations were obtained from the CERN and existing literature<sup>32,33,34</sup>.

subsets. Seven of these subsets were employed to train the model, while the remaining subsets were used to validate its accuracy. The model's performance was evaluated using metrics such as the coefficient of determination ( $R^2$ ), root mean square error (RMSE), and normalized root mean square error (NRMSE). Consequently, it is now possible to generate detailed estimates of blue and green water use for 15 distinct crops at a spatial resolution of 1 km.

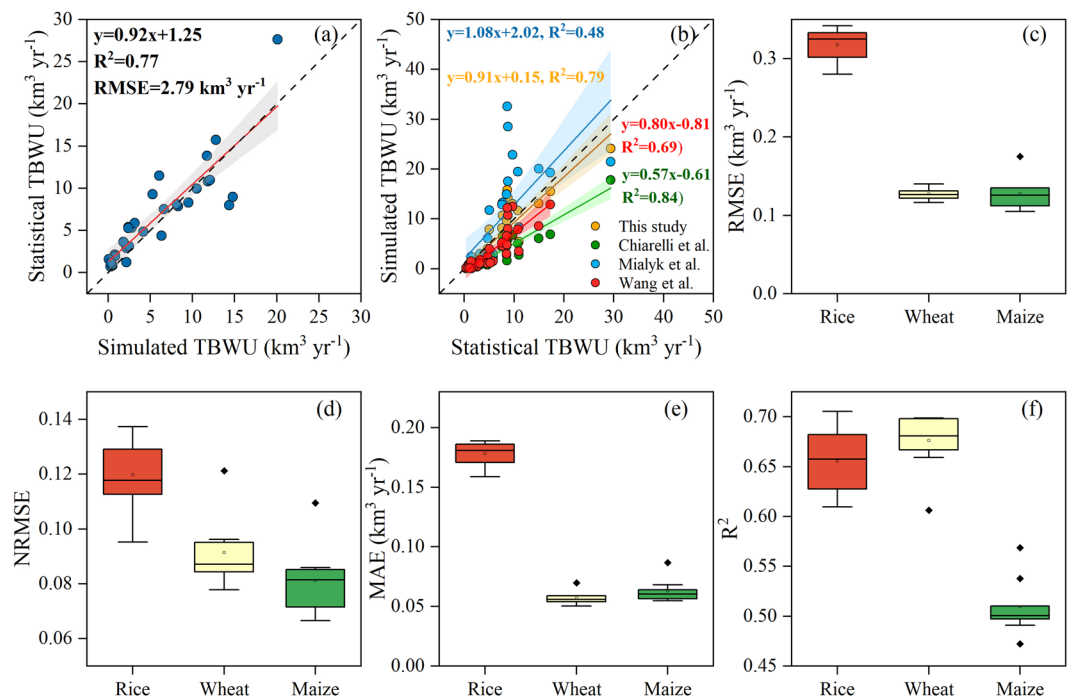
### Data Records

The dataset is available at Figshare<sup>37</sup>. We provide annual blue and green water for irrigated crops and green water for rainfed crops with a 1 km resolution in China from 1991 to 2019. The datasets are provided in the format of NetCDF4, with a spatial reference system of EPSG:4214 (Beijing1954). All the maps can be visualized in ArcMap. Specific datasets detailing the blue and green water use for each crop are available in the provided link (<https://doi.org/10.6084/m9.figshare.25980358>)<sup>37</sup>.

### Technical Validation

The accuracy of our simulation results was evaluated through three methods, including site-scale validation with measured data at field stations, comparison with statistics, and intercomparison with existing data products. In addition, we compared the annual reference evapotranspiration with the existing datasets and verified the accuracy of the RF model.

**Comparison with measured data at field stations.** To validate the accuracy of our model, we compared its output with the locally measured values across several crops and sites sourced from the CERN and existing literature (Fig. 3). Firstly, we compared the simulated blue water use of seven crops with the measured results. The findings revealed a strong consistency with an  $R^2$  of 0.95 and RMSE of 47.55 mm when the intercept was set to 0 (Fig. 3a). The linear regression line ( $y = 0.9764x$ ) closely aligns with the 1:1 line. Subsequently, we conducted a comparative analysis of evapotranspiration under various cropping systems (Fig. 3b–f). Due to variations in the geographical positioning of the field stations and meteorological stations used in this study, the daily evapotranspiration obtained from the field stations was juxtaposed with the simulated data derived from nearby meteorological stations. The comparative assessment revealed that while there were disparities between the simulated and observed values of crop evapotranspiration, the overall trends exhibited substantial consistency. The noted differences between the datasets could be attributed to the distinct geographical settings of the respective stations. Collectively, these results indicate the reliability and feasibility of estimating crop water use based on the model proposed in this study.

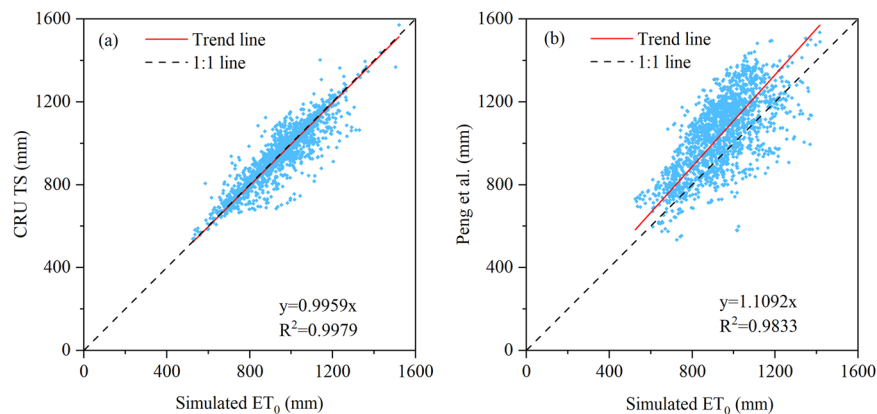


**Fig. 4** Comparison of total blue water use (TBWU) between statistics and corresponding values in this study. **(a)** Scatter plots of simulated average TBWU and corresponding statistical BWU for 31 provinces from 1991–2019. **(b)** Comparison of TBWU simulated by this study, Chiarelli *et al.*<sup>22</sup>, Mialyk *et al.*<sup>23</sup>, and Wang *et al.*<sup>25</sup> with statistical data for 31 provinces in 2016. Box plots of RMSE **(c)**, NRMSE **(d)**, MAE **(e)**, and  $R^2$  **(f)** for rice, wheat, and maize from 2005–2013 at the city scale compared with Zhou *et al.*<sup>19</sup>.

**Comparison with statistical data.** Statistical data typically provides the total amount of water use ( $\text{km}^3 \text{yr}^{-1}$ ). Therefore, the simulated crop blue water use ( $\text{mm yr}^{-1}$ ) and crop harvested area (ha) data were initially combined to calculate the total blue water use (TBWU) of crops. However, the traditional harvested area is available at the provincial scale provided by the National Bureau of Statistics and is therefore unable to capture the spatial heterogeneity of agricultural systems at the grid cells. To address this limitation, we used the SPAM2010 dataset to conduct downscaling research. Here, we assumed that the proportion of each crop harvested area on each grid cell to the total harvested area of each crop in the province remain constant based on the SPAM2010. Then the harvested area of province-level statistics over the 1991–2019 period for each province and crop is multiplied by the above fixed proportion, and SPAM2010 harvested area for each grid cell is adjusted to 1991–2019.

Subsequently, a comparative analysis of TBWU was conducted at the provincial level for all crops and at the city level for wheat, maize, and rice, respectively. According to the China Water Resources Bulletin, we collected the total agricultural water use, the proportion of agricultural irrigation to agricultural water use, and the irrigation project efficiency for 31 provinces in mainland China from 2000 to 2019. Notably, for the irrigation project efficiency, except for the provinces in the Haihe River Basin and Guangdong, which are 80% and 45% respectively, the rest of the provinces adhered to the national average<sup>38</sup>. These data were then multiplied to obtain the total amount of water used for agricultural irrigation in each province. The results highlighted a strong correlation between simulated and statistical values at the provincial level, with an  $R^2$  of 0.77, a slope of the fitted line of 0.92, and an RMSE of  $2.79 \text{ km}^3 \text{yr}^{-1}$  (Fig. 4a). Despite the overall agreement, certain provinces exhibited significant discrepancies between simulated TBWU and statistical TBWU data. In our simulations, we assumed that irrigation water is always available when the irrigation infrastructure is present. This assumption led to an overestimation of TBWU in regions with inadequate irrigation facilities, such as Hebei Province, Shandong Province, and Henan Province in the North China Plain, where farmers typically utilize less water for irrigation than required by crops<sup>38</sup>. Additionally, we omitted considerations for water-intensive cropping systems like tea, fruits, and vegetables due to the absence of observational data from national agro-meteorological stations during crop growth, potentially resulting in an underestimation of TBWU in certain provinces.

The city-level reported BWU and irrigated area for wheat, maize, and rice during the 2005–2013 period were obtained from Zhou *et al.*<sup>19</sup>. We then calculated the TBWU for three crops and compared our results with theirs. The analysis showed that wheat exhibited the highest estimation accuracy with an  $R^2$  of 0.68 (0.61–0.70), followed by rice (0.61–0.71) and maize (0.40–0.57) (Fig. 4c–f). Given the higher irrigation demands of rice cultivation, the RMSE, NRMSE, and MAE were higher, which were elevated at  $0.32 \text{ km}^3 \text{yr}^{-1}$ ,  $0.12 \text{ km}^3 \text{yr}^{-1}$ , and  $0.18 \text{ km}^3 \text{yr}^{-1}$ , respectively. For wheat, the corresponding values were  $0.12 \text{ km}^3 \text{yr}^{-1}$ ,  $0.09 \text{ km}^3 \text{yr}^{-1}$  and  $0.06 \text{ km}^3 \text{yr}^{-1}$ , while maize recorded RMSE, NRMSE, and MAE of  $0.13 \text{ km}^3 \text{yr}^{-1}$ ,  $0.08 \text{ km}^3 \text{yr}^{-1}$  and  $0.06 \text{ km}^3 \text{yr}^{-1}$ , respectively. All of these results were close to the reported accuracies by Bo *et al.*<sup>39</sup>.



**Fig. 5** Comparison of annual reference evaporation in this study with datasets from the (a) Climate Research Unit Time Series (CRU TS) v4.01 datasets<sup>43</sup> and (b) Peng *et al.*<sup>42</sup> during the period 1990–2019.

**Comparison with existing datasets.** For the maps of the year 2016, we compared our simulation results with three existing grid datasets that were widely used around the world. Overall, the blue water use of irrigated crops in this study exhibited higher values than those reported by Chiarelli *et al.*<sup>22</sup> and Wang *et al.*<sup>25</sup>, but lower than the findings of Mialyk *et al.*<sup>23</sup> (Table S1). This phenomenon may be attributed to the differences in crop planting dates. Mialyk *et al.*'s crop calendar involve early planting, which generally extends the growing period and potentially increases total water requirements due to higher evapotranspiration during that time<sup>26</sup>. When considering the harvested area of irrigated crops, the TBWU in 2016 aligned more closely with statistical data, offering a more accurate representation than the three existing data products (Fig. 4b). Overall, it is more consistent with the results of Wang *et al.*<sup>25</sup>. Our study revealed relatively lower green water use for both irrigated and rainfed crops compared to the datasets, showing more similarity to the results of Mialyk *et al.*<sup>23</sup>, particularly for maize and potatoes. A potential reason for this discrepancy may be attributed to the use of an updated crop calendar in the study of Mialyk *et al.*<sup>23</sup>, a composite data product that integrates a multitude of observational data sources. While Chiarelli *et al.*<sup>22</sup> used crop calendars from around 2000. In addition, reference evapotranspiration was used as the basis for quantifying crop water use. Both our study and Mialyk *et al.*<sup>23</sup> calculated the daily values by combining daily meteorological data with the P-M equation, while Chiarelli *et al.*<sup>22</sup> obtained daily reference evapotranspiration by dividing the monthly reference evapotranspiration products by the number of days in the corresponding month. From the perspective of spatial correlation, our study exhibited a higher correlation with Chiarelli *et al.*<sup>22</sup> and Wang *et al.*<sup>25</sup>, particularly about the blue water use of irrigated crops. This is primarily because we employed the methodology of integrating the soil water balance model and the crop coefficient, whereas Mialyk *et al.*<sup>23</sup> estimated crop water use by the process-based grid crop models.

**Comparison with remote sensing data.** In a recent study, Zhang *et al.*<sup>24</sup> estimated the global irrigation water use by the Integration of multiple satellite observations. However, the simulation results of Zhang *et al.*<sup>24</sup> for Xinjiang and Northeast China exhibited a discernible underestimation<sup>24,40</sup>. This may be attributed to the fact that the irrigated areas used in the study are static rather than dynamic, which may lead to an underestimation of the total amount of irrigation water required in areas with rapid irrigation expansion<sup>24,41</sup>. Therefore, a comparison was conducted between the simulated crop blue water use and the datasets from the remaining 27 provinces. The results demonstrated a robust correlation at the provincial scale, with an  $R^2$  of 0.9, a slope of the fitted line of 0.85, and an RMSE of  $2.76 \text{ km}^3 \text{ yr}^{-1}$  (Fig. S2).

**The validation of the reference evapotranspiration.** Two publicly available monthly reference evapotranspiration datasets at grid scale were collected<sup>42,43</sup> and then calculated the corresponding annual values. Subsequently, the annual reference evapotranspiration data for the stations in this study were then extracted and compared with these datasets. The findings of this study were in good agreement with the dataset, especially the comparison with the Climate Research Unit Time Series (CRU TS) v4.01 datasets<sup>43</sup> with an  $R^2$  of 0.99 and RMSE of 61.18 mm when the intercept was set to 0 (Fig. 5a). The correlation between this study and Peng *et al.*<sup>42</sup> was relatively weak, with an  $R^2$  of 0.98 and RMSE of 172.11 mm. This is mainly due to differences in calculation methods. Both our study and the CRU TS dataset utilized the P-M equation to calculate reference evapotranspiration, whereas Peng *et al.*<sup>42</sup> relied on the Hargreaves' method. Despite the weaker correlation with Peng *et al.*<sup>42</sup>, the strong agreement with the CRU TS dataset underscores the reliability of our methodology in assessing annual reference evapotranspiration data at the grid scale.

**Accuracy validation of random forest model.** In this study, we conducted calculations of crop blue and green water use at each station (Fig. S1) and generated a spatial map of crop water use using an RF model from 1991 to 2019. Subsequently, the model's performance was assessed yearly using 10-fold cross-validation. This process is repeated 30 times per year for each crop, and the resulting data are averaged to produce the values presented in the Table 2. The results highlighted the optimal performance of the model in estimating blue water use for irrigated crops, with  $R^2$ , RMSE, and NRMSE of 0.68, 68.17 mm, and 8.07%, respectively, especially for the



Crops	BWU			GWU_I			GWU_R		
	RMSE (mm)	R <sup>2</sup>	NRMSE (%)	RMSE (mm)	R <sup>2</sup>	NRMSE (%)	RMSE (mm)	R <sup>2</sup>	NRMSE (%)
Winter wheat	51.71 (±2.71)	0.78 (±0.02)	6.47 (±1.22)	28.58 (±1.79)	0.58 (±0.03)	10.98 (±1.13)	26.36 (±1.80)	0.57 (±0.04)	10.08 (±0.95)
Spring wheat	62.12 (±3.83)	0.68 (±0.03)	6.69 (±1.21)	24.18 (±1.99)	0.74 (±0.04)	9.81 (±1.07)	21.82 (±1.92)	0.70 (±0.04)	9.97 (±1.16)
Spring maize	70.97 (±3.55)	0.77 (±0.02)	5.94 (±0.89)	30.13 (±2.38)	0.63 (±0.05)	10.77 (±1.15)	28.13 (±2.72)	0.53 (±0.03)	11.98 (±0.97)
Summer maize	70.49 (±3.74)	0.61 (±0.04)	8.58 (±0.80)	30.49 (±2.80)	0.67 (±0.05)	9.73 (±1.50)	28.19 (±3.06)	0.58 (±0.05)	10.79 (±1.27)
Early rice	61.96 (±5.60)	0.56 (±0.09)	8.76 (±0.96)	36.49 (±2.72)	0.48 (±0.09)	13.42 (±1.65)	32.14 (±2.37)	0.44 (±0.10)	13.71 (±1.63)
Late rice	51.75 (±5.53)	0.59 (±0.07)	13.53 (±1.23)	32.14 (±2.33)	0.51 (±0.07)	11.57 (±2.91)	28.18 (±2.12)	0.47 (±0.08)	11.60 (±2.61)
Single rice	83.86 (±5.20)	0.69 (±0.04)	7.30 (±1.54)	27.87 (±1.21)	0.78 (±0.03)	9.45 (±0.86)	23.84 (±1.03)	0.74 (±0.04)	9.95 (±0.89)
Soybean	66.05 (±5.94)	0.67 (±0.07)	8.19 (±1.10)	24.87 (±1.46)	0.55 (±0.07)	11.08 (±1.88)	21.37 (±1.33)	0.57 (±0.11)	11.20 (±2.63)
Potatoes	62.17 (±4.10)	0.85 (±0.02)	4.93 (±0.67)	28.30 (±2.29)	0.67 (±0.04)	10.56 (±1.16)	25.97 (±2.03)	0.64 (±0.04)	11.04 (±1.19)
Cotton	65.17 (±4.50)	0.75 (±0.04)	6.15 (±1.12)	28.39 (±2.41)	0.75 (±0.06)	9.51 (±1.45)	26.01 (±2.25)	0.68 (±0.07)	10.30 (±1.57)
Sugarcane	91.95 (±8.49)	0.56 (±0.06)	12.52 (±1.47)	46.61 (±4.48)	0.47 (±0.09)	14.49 (±3.47)	45.47 (±3.53)	0.51 (±0.10)	14.09 (±4.03)
Sugar beet	87.61 (±5.52)	0.78 (±0.03)	5.94 (±0.79)	24.05 (±1.88)	0.80 (±0.04)	9.20 (±1.55)	21.51 (±1.75)	0.78 (±0.05)	9.79 (±1.87)
Groundnut	65.17 (±2.93)	0.74 (±0.05)	7.42 (±0.70)	28.39 (±1.75)	0.75 (±0.09)	9.51 (±2.06)	26.01 (±1.66)	0.68 (±0.09)	10.30 (±1.96)
Spring rapeseed	83.92 (±8.75)	0.59 (±0.04)	8.05 (±1.11)	21.71 (±2.07)	0.69 (±0.04)	11.00 (±1.99)	19.14 (±1.87)	0.63 (±0.05)	11.55 (±1.95)
Winter rapeseed	47.6 (±3.69)	0.64 (±0.05)	10.54 (±1.09)	30.21 (±1.70)	0.58 (±0.06)	12.82 (±1.68)	28.31 (±1.55)	0.54 (±0.07)	12.26 (±1.68)

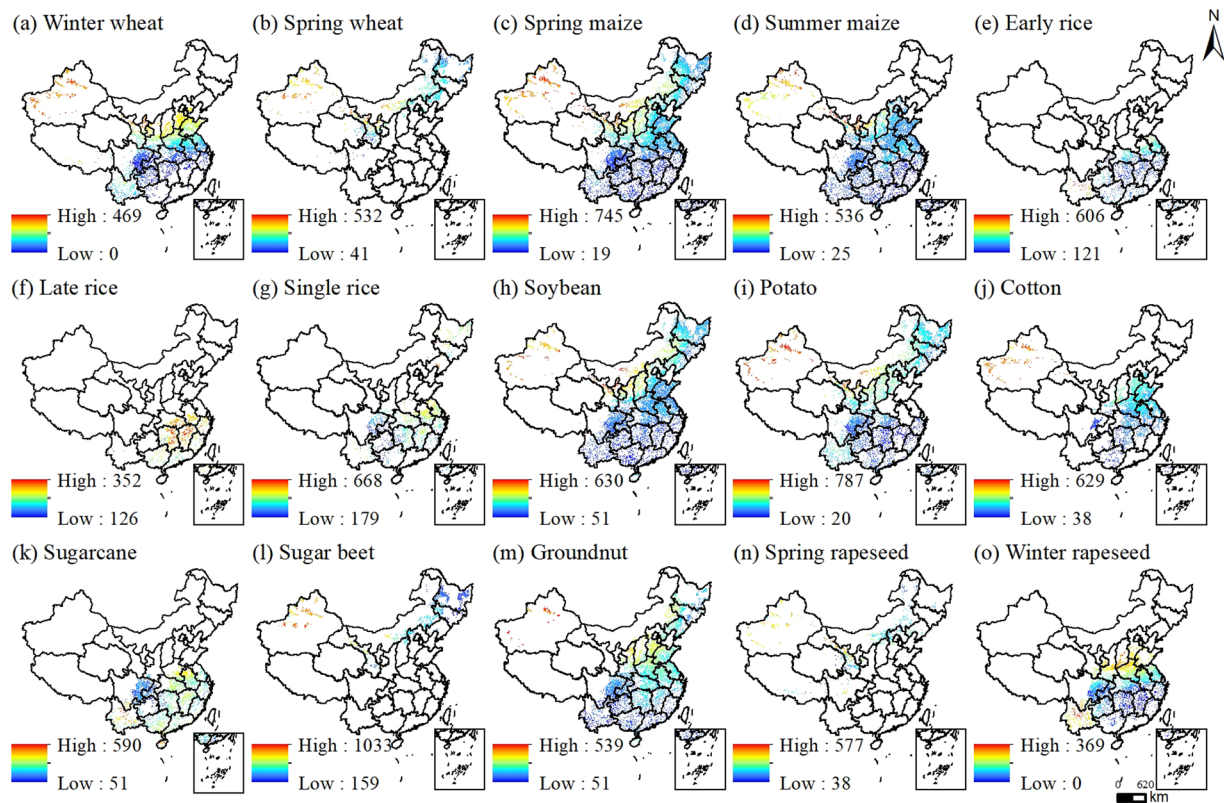
**Table 2.** Prediction performances of the random forest models derived using 10-fold cross-validation. BWU, GWU\_I and GWU\_R represent the crop blue water use for irrigated crops, green water use for irrigated crops and for rainfed crops. Error range are shown in parenthesis.

winter wheat ( $R^2 = 0.78$ ), potatoes ( $R^2 = 0.85$ ) and sugar beet ( $R^2 = 0.78$ ). For the green water use, the simulation of irrigated crops revealed a superior efficacy, with  $R^2$ , RMSE, and NRMSE of 0.64, 29.49 mm, and 10.93%, respectively, especially for the single rice ( $R^2 = 0.78$ ), and sugar beet ( $R^2 = 0.80$ ). The accuracy for rainfed crops was slightly lower at 0.60, 26.83 mm, and 11.24%, respectively. Lower accuracy was observed for crops such as sugar cane, and rapeseed, etc., primarily due to the limited number of stations dedicated to monitoring these specific crops.

**Spatial distribution.** Overall, the blue water use of various crops demonstrated an increasing trend from south to north and from east to west (Fig. 6). Crops such as rice, wheat, sugar cane, cotton and potato have relatively high blue water requirements, whereas crops such as soybean, groundnut and maize have relatively low blue water demand. In addition, the blue water use in northern China is predominantly increasing from 1991 to 2019, particularly in the North China Plain and the northwest inland regions. The decline in this indicator is primarily concentrated in southern China, especially in the provinces along the coast of China (Fig. S3). These findings are in accordance with those previously reported by Yin *et al.*<sup>40</sup>. As for green water use, its spatial distribution pattern and change trend are opposite to that of blue water use. Additionally, rainfed crops demonstrated a higher green water use compared to those that were irrigated (Figs. S4–S7).

**Uncertainties and limitations.** This study leverages localized input parameters such as meteorological data, crop coefficients, and crop growth periods to generate maps of blue and green water use for 15 crops in China using soil water balance models combined with a random forest model. Despite the comprehensive approach, the study acknowledges potential uncertainties that could impact the accuracy of the final water use maps. To assess the uncertainties associated with crop blue and green water use, the study focuses on variations in  $K_c$  values and crop calendars.  $K_c$  values adopted in the research of Zhuo *et al.*<sup>44</sup>, allowing for a  $\pm 15\%$  variability for each crop. Additionally, variations in crop calendars were explored by shifting planting dates within  $\pm 15$  days while maintaining the overall crop growth duration constant. The results indicated that changes in  $K_c$  values and crop calendars lead to modest fluctuations in green water use (within  $-9.19\%$  to  $9.12\%$ ), except for specific crops like early rice, potato, and winter rapeseed (Table 3). Conversely, crop blue water use demonstrated more significant variations ranging from  $-38.71\%$  to  $44.34\%$ , aligning with previous research findings<sup>22,44</sup>.

In addition to quantifying uncertainty, some limitations in this study also should be noted. Firstly, it should be noted that the study did not consider all crop types in China, such as tea, fruits, vegetables, tobacco, etc. Secondly, due to data limitations, planting and harvesting dates for crop-specific considerations post-2013 were approximated using average data from 2011 to 2013. Furthermore, the deep percolation is not considered in the model, as it is challenging to quantify and separate from ET<sup>45</sup>. Moreover, the study only used an RF model to produce maps of crop blue and green water use, benefiting from the model's robustness in estimating crop water requirements compared with other models<sup>39</sup>. Future research could explore integrating additional environmental factors using models like Support Vector Machines (SVM), Artificial Neural Networks (ANN), High Accuracy Surface Modelling (HASM), among others, to enhance spatial predictions of crop water use. Despite these limitations, the study, when augmented with more localized parameters, offers high-resolution data on blue and green water use for a wide range of crops in China. This dataset is crucial for safeguarding China's water resources and warrants further exploration and refinement in future studies.



**Fig. 6** Spatial distribution of crop average blue water use (mm) from 1991 to 2019. It should be noted that the spatial distribution does not represent the actual extent of crop distribution. The distribution was obtained by masking cultivated land in counties or provinces with records of crop cultivation.

Crop	BWU				GWU_I				GWU_R			
	$K_c-15$	$K_c+15$	length-15	length+15	$K_c-15$	$K_c+15$	length-15	length+15	$K_c-15$	$K_c+15$	length-15	length+15
Winter wheat	-36.62	39.27	-28.03	30.44	-2.62	4.61	-8.55	9.01	-3.18	5.03	-8.27	8.41
Spring wheat	-31.25	33.04	-16.53	14.54	-2.35	1.58	-7.70	7.59	-2.89	1.94	-6.93	6.65
Spring maize	-36.02	40.38	-12.57	15.44	-4.04	2.79	-8.08	6.83	-4.80	3.30	-7.09	6.07
Summer maize	-38.71	44.34	-15.95	19.13	-4.92	3.43	-8.40	7.08	-5.80	3.96	-7.73	6.52
Early rice	-24.92	26.20	-18.06	20.47	-2.73	9.12	-13.52	12.24	—	—	—	—
Late rice	-24.34	25.36	-8.64	7.55	-4.83	3.71	-7.37	6.11	—	—	—	—
Single rice	-22.01	22.68	-8.12	6.93	-3.63	2.76	-7.03	5.79	—	—	—	—
Soybean	-29.95	32.07	-11.60	11.21	-3.45	2.73	-9.19	8.22	-4.18	3.23	-7.82	7.05
Potatoes	-27.27	29.36	-9.19	9.14	-3.66	3.01	-11.06	10.01	-4.15	3.49	-10.48	9.61
Cotton	-33.32	36.05	-13.97	12.28	-3.60	2.61	-7.04	6.17	-3.82	2.78	-6.19	5.47
Sugarcane	-44.63	53.38	-6.16	5.25	-5.58	3.30	-2.06	1.59	-5.87	3.61	-2.08	1.61
Sugar beet	-23.55	23.93	3.16	-5.01	-1.49	1.09	-4.17	3.11	-1.92	1.44	-3.33	2.50
Groundnut	-30.26	32.46	-12.19	10.31	-4.29	3.13	-9.10	7.98	-4.49	3.31	-8.06	7.05
Spring rapeseed	-29.59	31.20	-13.34	9.71	-2.13	1.51	-5.81	5.23	-2.93	1.96	-4.97	4.26
Winter rapeseed	-38.21	42.63	-30.04	32.01	-7.22	5.44	-9.97	10.74	-8.26	6.19	-9.69	10.04

**Table 3.** Percentage changes of crop-specific blue water use (BWU) and green water use for irrigated crops (GWU\_I) and rainfed crops (GWU\_R) with different  $K_c$  and crop calendars.

Usage Notes

The 1 km resolution agricultural water consumption dataset generated in this study over the past 30 years serves as crucial data support for agricultural water resources management and ensuring the stability of the agricultural system in China. Researchers can utilize this dataset to analyse crop water consumption patterns, optimize irrigation management strategies, and evaluate water resource utilization efficiency. Government agencies can utilize this dataset to develop irrigation policies, strategize agricultural water resource allocation, advocate for sustainable agricultural methods, and manage water resources efficiently to boost agricultural productivity and

safeguard national water resource security. Nevertheless, it is important to note that users should be aware of certain issues when working with this dataset. (1) The blue water use of crops as modeled by us does not correspond precisely to the actual irrigation water use of crops. Crop blue water use represents the water required for irrigation when there is sufficient and functional irrigation infrastructure, indicating an irrigation potential. In regions with insufficient irrigation infrastructure, this can result in an overestimation of the total amount of irrigation water. (2) This study exhibits a degree of error propagation. Initially, due to the discrepancy in spatial locations between meteorological and agro-meteorological stations, we obtained data on crop growth and development from the latter and subsequently mapped it onto Thiessen polygons. Subsequently, we extracted data on crop calendars and crop heights corresponding to each meteorological station. While we anticipate that changes in crop calendars do not significantly impact our estimations of crop water use for most crops (Table 3), such considerations are crucial for research aiming to apply our methodology over extended periods.

### Code availability

R code for calculating crop water use is freely available from <https://doi.org/10.6084/m9.figshare.25980358><sup>37</sup>.

Received: 6 June 2024; Accepted: 28 November 2024;

Published online: 18 December 2024

### References

1. Damerau, K. *et al.* India has natural resource capacity to achieve nutrition security, reduce health risks and improve environmental sustainability. *Nature Food* **1**, 631–639 (2020).
2. Davis, K. F., Rulli, M. C., Seveso, A. & D'Odorico, P. Increased food production and reduced water use through optimized crop distribution. *Nature Geoscience* **10**, 919–924 (2017).
3. Mekonnen, M. M. & Hoekstra, A. Y. A global and high-resolution assessment of the green, blue and grey water footprint of wheat. *Hydrology and Earth System Sciences* **14**, 1259–1276 (2010).
4. Mekonnen, M. M. & Hoekstra, A. Y. The green, blue and grey water footprint of crops and derived crop products. *Hydrology and Earth System Sciences* **15**, 1577–1600 (2011).
5. Huang, Z. *et al.* Global agricultural green and blue water consumption under future climate and land use changes. *Journal of Hydrology* **574**, 242–256 (2019).
6. Springmann, M. *et al.* Options for keeping the food system within environmental limits. *Nature* **562**, 519–525 (2018).
7. Wang, M. L. & Shi, W. J. Research progress in assessment and strategies for sustainable food system within planetary boundaries. *Science China Earth Sciences* **67**, 375–386 (2024).
8. Cao, X., Ren, J., Wu, M., Guo, X. & Wang, W. Assessing agricultural water use effect of China based on water footprint framework. *Transactions of the Chinese Society of Agricultural Engineering* **34**, 1–8 (2018).
9. The State Council of China. China has become the world's largest irrigation country. [https://www.gov.cn/xinwen/2022-11/06/content\\_5724969.htm](https://www.gov.cn/xinwen/2022-11/06/content_5724969.htm) (2022).
10. Liang, H., Hu, K., Batchelor, W. D., Qi, Z. & Li, B. An integrated soil-crop system model for water and nitrogen management in North China. *Scientific Reports* **6**, 25755 (2016).
11. Long, D. *et al.* South-to-North Water Diversion stabilizing Beijing's groundwater levels. *Nature Communications* **11**, 3665 (2020).
12. Hu, Y. *et al.* Food production in China requires intensified measures to be consistent with national and provincial environmental boundaries. *Nature Food* **1**, 572–582 (2020).
13. Zhang, Y. *et al.* Characterization of soil salinization and its driving factors in a typical irrigation area of Northwest China. *Science of The Total Environment* **837**, 155808 (2022).
14. Shi, W. *et al.* Wheat redistribution in Huang-Huai-Hai, China, could reduce groundwater depletion and environmental footprints without compromising production. *Communications Earth & Environment* **5**, 380 (2024).
15. Wang, S. *et al.* Urbanization can benefit agricultural production with large-scale farming in China. *Nature Food* **2**, 183–191 (2021).
16. Huang, F. *et al.* Current situation and future security of agricultural water resources in North China (in Chinese). *Strategic Study of CAE* **21**, 28–37 (2019).
17. Zhao, H. *et al.* China's future food demand and its implications for trade and environment. *Nature Sustainability* **4**, 1042–1051 (2021).
18. Zhang, C., Dong, J. & Ge, Q. Mapping 20 years of irrigated croplands in China using MODIS and statistics and existing irrigation products. *Scientific Data* **9**, 407 (2022).
19. Zhou, F. *et al.* Deceleration of China's human water use and its key drivers. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 7702–7711 (2020).
20. Li, X. R., Jiang, W. L. & Duan, D. D. Spatio-temporal analysis of irrigation water use coefficients in China. *Journal of Environmental Management* **262**, 110242 (2020).
21. Siebert, S. & Döll, P. Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. *Journal of Hydrology* **384**, 198–217 (2010).
22. Chiarelli, D. D. *et al.* The green and blue crop water requirement WATNEEDS model and its global gridded outputs. *Scientific Data* **7**, 273 (2020).
23. Mialyk, O. *et al.* Water footprints and crop water use of 175 individual crops for 1990–2019 simulated with a global crop model. *Scientific Data* **11**, 206 (2024).
24. Zhang, K., Li, X., Zheng, D., Zhang, L. & Zhu, G. Estimation of Global Irrigation Water Use by the Integration of Multiple Satellite Observations. *Water Resources Research* **58**, e2021WR030031 (2022).
25. Wang, W. *et al.* A gridded dataset of consumptive water footprints, evaporation, transpiration, and associated benchmarks related to crop production in China during 2000–2018. *Earth System Science Data* **15**, 4803–4827 (2023).
26. Allan, R. G., Pereira, L. S., Rase, D. & Smith, M. *Crop evapotranspiration: Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56*. (FAO, Rome, 1998).
27. Li, X. *Agricultural irrigation water requirement and its regional characteristic in China* (in Chinese). Tsinghua University (2006).
28. Fischer, G. *et al.* *Global Agro-Ecological Zones v4 – Model documentation*. (FAO, Rome, 2021).
29. Yu, Q. *et al.* A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps. *Earth System Science Data* **12**, 3545–3572 (2020).
30. National Bureau of Statistics. *China Statistical Yearbook 2020* (China Statistics Press, 2020).
31. Ministry of Water Resources of China. *China Water Resources Bulletin* (China Water & Power Press, 2020).
32. Gao, Y. *et al.* Winter wheat with subsurface drip irrigation (SDI): Crop coefficients, water-use estimates, and effects of SDI on grain yield and water use efficiency. *Agricultural Water Management* **146**, 1–10 (2014).
33. Liu, H. J., Yu, L. P., Luo, Y., Wang, X. P. & Huang, G. H. Responses of winter wheat (*Triticum aestivum* L.) evapotranspiration and yield to sprinkler irrigation regimes. *Agricultural Water Management* **98**, 483–492 (2011).

34. Zhou, S. Q. *et al.* Evapotranspiration of a drip-irrigated, film-mulched cotton field in northern Xinjiang, China. *Hydrological Processes*. **26**, 1169–1178 (2012).
35. Döll, P. & Siebert, S. Global modeling of irrigation water requirements. *Water Resources Research* **38**, 1037 (2002).
36. Martin, S. & Land, F. *Cropwat: a computer program for irrigation planning and management*. (FAO, Rome, 1992)
37. Wang, M. & Shi, W. Data underlying the publication: The annual dynamic dataset of high-resolution crop water use in China from 1991 to 2019. *figshare* <https://doi.org/10.6084/m9.figshare.25980358.v1> (2024).
38. Liu, J. & Yang, H. Spatially explicit assessment of global consumptive water uses in cropland: Green and blue water. *Journal of Hydrology* **384**, 187–197 (2010).
39. Bo, Y. *et al.* Hybrid theory-guided data driven framework for calculating irrigation water use of three staple cereal crops in China. *Water Resources Research* **60**, e2023WR035234 (2024).
40. Yin, L. *et al.* Irrigation water consumption of irrigated cropland and its dominant factor in China from 1982 to 2015. *Advances in Water Resources*. **143**, 103661 (2020).
41. Zhang, C., Dong, J. & Ge, Q. IrriMap\_CN: Annual irrigation maps across China in 2000–2019 based on satellite observations, environmental variables, and machine learning. *Remote Sensing of Environment*. **280**, 113184 (2022).
42. Peng, S. 1 km monthly potential evapotranspiration dataset in China (1901–2022). *Center National Tibetan Plateau Data* <https://doi.org/10.11866/db.loess.2021.001> (2022).
43. Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* **7**, 109 (2020).
44. Zhuo, L., Mekonnen, M. M. & Hoekstra, A. Y. Sensitivity and uncertainty in crop water footprint accounting: a case study for the Yellow River basin. *Hydrology and Earth System Sciences* **18**, 2219–2234 (2014).
45. Shi, P., Gai, H. & Li, Z. Partitioned soil water balance and its link with water uptake strategy under apple trees in the Loess-Covered region. *Water Resources Research* **59**, e2022WR032670 (2022).

## Acknowledgements

This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA0440405), the National Key Research and Development Program of China (2022YFB3903504), and the National Natural Science Foundation of China (72221002 and 42330707).

## Author contributions

Minglei Wang: Data curation, methodology, software, visualization, validation, writing-original draft. Wenjiao Shi: Resources, methodology, writing-review & editing, funding acquisition.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-024-04185-0>.

**Correspondence** and requests for materials should be addressed to W.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024