Benchmark levels for the consumptive water footprint of crop production for different environmental conditions: a case study for winter wheat in China

La Zhuo[1], Mesfin M. Mekonnen[1], Arjen Y. Hoekstra[1]

[1]Twente Water Centre, University of Twente, Enschede, 7500AE, The Netherlands

Correspondence to: L. Zhuo (l.zhuo@utwente.nl; zhuo.l@hotmail.com)

1 **Abstract.**

2 Meeting growing food demands while simultaneously shrinking the water footprint (WF) of agricultural production is one of

3 the greatest societal challenges. Benchmarks for the WF of crop production can serve as a reference and be helpful in setting

4 WF reduction targets. The consumptive WF of crops, the consumption of rainwater stored in the soil (green WF) and the

5 consumption of irrigation water (blue WF) over the crop growing period, varies spatially and temporally depending on

6 environmental factors like climate and soil. The study explores which environmental factors should be distinguished when

7 determining benchmark levels for the consumptive WF of crops. Hereto we determine benchmark levels for the consumptive

8 WF of winter wheat production in China for all separate years in the period 1961-2008, for rain-fed versus irrigated croplands,

9 for wet versus dry years, for warm versus cold years, for four different soil classes and for two different climate zones. We

10 simulate consumptive WFs of winter wheat production with the crop water productivity model AquaCrop at a 5 by 5 arc min

11 resolution, accounting for water stress only. The results show that (i) benchmark levels determined for individual years for the

12 country as a whole remain within a range of ±20% around long-term mean levels over 1961-2008; (ii) the WF benchmarks for

13 irrigated winter wheat are 8-10% larger than those for rain-fed winter wheat; (iii) WF benchmarks for wet years are 1-3%

14 smaller than for dry years, (iv) WF benchmarks for warm years are 7-8% smaller than for cold years, (v) WF benchmarks

15 differ by about 10-12% across different soil texture classes; and (vi) WF benchmarks for the humid zone are 26-31% smaller

16 than for the arid zone, which has relatively higher reference evapotranspiration in general and lower yields in rain-fed fields.

17 We conclude that when determining benchmark levels for the consumptive WF of a crop, it is useful to primarily distinguish

18 between different climate zones. If actual consumptive WFs of winter wheat throughout China were reduced to the benchmark

19 levels set by the best 25% of Chinese winter wheat production ($1224 \, \mathrm{m^3 \, t^{-1}}$ for arid areas and $841 \, \mathrm{m^3 \, t^{-1}}$ for humid areas), the

20 water saving in an average year would be 53% of the current water consumption at winter wheat fields in China. The majority

21 of the yield increase and associated improvement in water productivity can be achieved in southern China.

22

1 Introduction

Half of the large river basins in the world face severe blue water scarcity for at least one month a year (Hoekstra et al., 2012). Agriculture is the largest consumer of water in the world and therefore responsible for a large part of the water scarcity in the world. Still, global food demand continues to increase, due to growing populations and changing diets. Meeting growing food demands and simultaneously reducing the water footprint (WF) of agricultural production is therefore one of the greatest societal challenges of our time (Foley et al., 2011;Hoekstra and Wiedmann, 2014). In crop production, individual farmers generally aim to maximize their economic return through raising their productivity per unit of input such as capital, labour, land, and fertilizer. When water is scarce, raising production per unit of water (i.e. increasing water productivity in terms of t m^{-3} or reducing the WF in $m^3 t^{-1}$) is a key challenge in order to save water and achieve sustainable water use at catchment level. Even when water is not scarce, it makes sense to have a reasonable level of water productivity, i.e. a good amount of crop per drop. Farmers, however, generally lack incentives for saving water, since they pay little for their water use compared to other input factors, even under conditions of high water scarcity. In order to provide producers with an incentive to reduce the WF of their products to reasonable levels, Hoekstra (2014, 2013) has proposed to develop WF benchmarks, which can be used by governments, farmers and customers (crop traders and retailers) for setting WF reduction targets. Setting WF benchmarks for different products, particularly water-intensive products like crops, is fundamental for wise water allocation and fair sharing of water resources among different sectors and users (Hoekstra, 2013). WF benchmarks of crop production could be global, but would preferably be context-specific, given the fact that the WF of growing a crop varies as a function of environmental factors such as climate and soil (Mekonnen and Hoekstra, 2011; Siebert and Döll, 2010; Tuninetti et al., 2015).

The WF of a crop is determined by both environmental conditions (e.g. climate, soil texture, CO_2 concentration in the air) that cannot be controlled by humans and managerial factors (e.g. application of fertilizers and pesticides, irrigation technology and strategy, mulching practice) (Zwart et al., 2010; Mekonnen and Hoekstra, 2011; Brauman et al., 2013). Benchmarks for the WF of growing a crop can, for example, be set by looking at what WF level is not exceeded by the best 20-25% of the total production in an area. Alternatively, benchmarks can be determined by estimating the WF associated with best-available technology and management practice (Hoekstra, 2014, 2013). Mekonnen and Hoekstra (2014) followed the first approach and developed global benchmarks for both the consumptive (green plus blue) WF and the degradative (grey) WF for a large number of crops, based on estimated WF values for 1996-2005 at a spatial resolution of 5 by 5 arc minute. Chukalla et al. (2015) followed the second approach and explored reduction potentials of consumptive WFs for a few crops by applying different alternative irrigation techniques and strategies and different alternative mulching practices. They found that the highest reduction (~29%) in the consumptive WF of a crop could be achieved when applying drip/subsurface drip irrigation in combination with deficit irrigation and synthetic mulching.

1 Research in developing benchmark levels for the consumptive WF of crop production is still in its infancy. An important

2 question that has been insufficiently addressed is which environmental factors should play a role when developing WF

3 benchmarks. It is nice to have one global benchmark for the consumptive WF per crop, as a global reference, like the ones

4 developed by Mekonnen and Hoekstra (2014), but it remains unclear whether it is reasonable to expect the same water

5 productivity under different environmental conditions. In their global analysis, Mekonnen and Hoekstra (2014) found that a

6 crop in a temperate climate generally has a smaller WF than the same crop in a tropical climate, but this can still be due to

7 other factors (e.g. better management practices in temperate climates), so that this is not a sufficient finding to diversify

8 benchmark levels based on the distinction between temperate and tropical. Besides, even though Mekonnen and Hoekstra

9 (2014) found a difference between different climates, for each crop considered it was found that the 10% best global production

10 (e.g. with smallest WFs) were always at least partly in the tropics as well. In other words, a WF benchmark developed in the

11 temperate part of the world still offers a reference value that can be achieved in the tropics as well. Next to climate also soil

12 affects evapotranspiration and yield and thus the WF of a crop. Tolk and Howell (2012), for example, analyse the variation of

13 consumptive WFs of sunflower in relation to different types of soils. There has not been yet, though, a systematic study looking

14 at how environmental factors influence the consumptive WFs of crops and to which extent it makes sense to diversify WF

15 benchmark levels based on specific environmental factors.

16

17 The current study aims to contribute to this discussion through an explorative study for winter wheat in China. We explore

18 which environmental factors should be distinguished when determining benchmark levels for the consumptive WF of crops.

19 We subsequently determine benchmark levels for the consumptive WF of winter wheat production in China for all separate

20 years in the period 1961-2008, for rain-fed versus irrigated croplands, for wet versus dry years, for warm versus cold years,

21 for four different soil classes and for two different climate zones. Winter wheat in China accounts for 95% of total wheat

22 production in China, which is the world biggest wheat producer (FAO, 2014). Winter wheat covers 96% of China's harvested

23 wheat area and occurs across China's different climate zones (NBSC, 2013). In order to avoid interference from managerial

24 factors that cause differences in evapotranspiration and yield, we simulate WFs by means of FAO's water productivity model

25 AquaCrop (Hsiao et al., 2009; Raes et al., 2009; Steduto et al., 2009), at a resolution of 5 by 5 arc minute, considering only

26 water stress and not taking into account other stresses such as from soil fertility, salinity, frost, or pest and diseases.

2 Method and data

2.1 Estimating consumptive WF of growing a crop

29 The consumptive (green and blue) WF of growing a crop ($m^3 t^{-1}$) equals the total actual evapotranspiration (ET, $m^3 ha^{-1}$) over

30 the cropping period divided by the crop yield (Y, t ha^{-1}). In current study, the ET and Y of growing winter wheat in China were

31 simulated at daily basis, at 5 by 5 arc min resolution, with FAO's crop water productivity model AquaCrop (Hsiao et al., 2009;

32 Raes et al., 2009; Steduto et al., 2009), run for the whole period 1961-2008. Compared to other crop growth models, AquaCrop

has a significantly smaller number of parameters and better balances between simplicity, accuracy and robustness (Steduto et al., 2007; Confalonieri et al., 2016). The model performance on simulating crop growth and water use has been well tested for a variety of crop types under diverse environmental conditions (e.g. Kumar et al., 2014; Jin et al., 2014; Abedinpour et al., 2012; Mkhabela and Bullock, 2012; Andarzian et al., 2011; Stricevic et al., 2011; Heng et al., 2009; Farahani et al., 2009; García-vila et al., 2009). AquaCrop has been applied in WF accounting at field (Chukalla et al., 2015), river basin (Zhuo et al., 2016a) and national level (Zhuo et al., 2016b) at high spatial resolution.

AquaCrop simulates water-driven crop water productivity with a dynamic daily soil water balance:

$$S_{[t]} = S_{[t-1]} + PR_{[t]} + IRR_{[t]} + CR_{[t]} - ET_{[t]} - RO_{[t]} - DP_{[t]} \qquad (1)$$

where $S_{[t]}$ (mm) refers to the soil water content at the end of day t, $PR_{[t]}$ (mm) the precipitation on day t, $IRR_{[t]}$ (mm) the irrigation water applied on day t, $CR_{[t]}$ (mm) the capillary rise from groundwater, $RO_{[t]}$ (mm) daily surface runoff and $DP_{[t]}$ (mm) deep percolation. $CR_{[t]}$ is assumed to be zero because the ground water depth is considered to be much larger than 1m (Allen et al., 1998).

The green and blue WFs are determined by green and blue ET over the cropping period, respectively, divided by Y. Following Chukalla et al. (2015) and Zhuo et al. (2016a, b), the daily green and blue ET (mm) were separated by tracking the daily incoming and outgoing green and blue water fluxes at the boundaries of the root zone:

$$\begin{cases} S_{green[t]} = S_{green[t-1]} + \left(PR_{[t]} + IRR_{[t]} - RO_{[t]}\right) \times \dfrac{PR_{[t]}}{\left(PR_{[t]} + IRR_{[t]}\right)} - \left(DP_{[t]} + ET_{[t]}\right) \times \dfrac{S_{green[t-1]}}{S_{[t-1]}} \\[2em] S_{blue[t]} = S_{blue[t-1]} + \left(PR_{[t]} + IRR_{[t]} - RO_{[t]}\right) \times \dfrac{IRR_{[t]}}{\left(PR_{[t]} + IRR_{[t]}\right)} - \left(DP_{[t]} + ET_{[t]}\right) \times \dfrac{S_{blue[t-1]}}{S_{[t-1]}} \end{cases} \qquad (2)$$

where S_{green} and S_{blue} refer to the green and blue soil water content, respectively. The initial soil water moisture at the start of the growing period is assumed as green water. The contribution of precipitation (green water) and irrigation (blue water) to surface runoff was calculated based on the respective magnitudes of precipitation and irrigation to the total green plus blue water inflow. The green and blue components in DP and ET were 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.

Y was determined by multiplying the above-ground biomass (B) and the harvest index (HI, %). HI was adjusted to water and temperature stress depending on timing and extent of the stress by an adjustment factor (f_{HI}) from the reference harvest index (HI_0) (Raes et al., 2011):

$$HI = f_{HI} \times HI_0 \qquad (3)$$

Only water stress is considered in modelling, which is determined by the water availability in the root zone, thus leaving out the effects of non-environmental factors (e.g. technology, fertilization) on crop growth. For irrigated fields, we assume that the applied irrigation volumes are equal to the net irrigation requirement. We used the same input crop parameters, including a fixed crop calendar, reference harvested index and maximum root depth as calibrated for China's winter wheat, as Zhuo et al. (2016b). We simulated winter wheat production per grid cell over the years based on the irrigated and rain-fed harvested areas of around the year 2000, as obtained from Portmann et al. (2010) (Fig. 1) in order to avoid in the simulations the effects of changes in where how much wheat is grown.

Data on monthly precipitation, reference evapotranspiration (ET_0) and temperature at 30 arc min resolution were taken from the CRU-TS 3.10 dataset (Harris et al., 2014). Soil texture data were obtained from Dijkshoorn et al. (2008). For hydraulic characteristics for each type of soil, the indicative values provided by AquaCrop were used. Data on total soil water capacity were obtained from Batjes (2012).

2.2 Benchmarking consumptive WF of growing a crop

Following Mekonnen and Hoekstra (2014), benchmark levels for the consumptive WF of crop production were determined by ranking the grid-level WF values from the smallest to the largest against the corresponding cumulative percentage of total crop production. As in the earlier study, we did not distinguish between green and blue WF benchmarks for two reasons. Firstly, the ratio of green to blue WF of a crop heavily depends on local green water resources availability, which is defined by the climate of certain time in a certain location. Location-specific blue WF benchmarks can be developed as a function of the overall consumptive WF benchmarks and local green water availability (Mekonnen and Hoekstra, 2014). Secondly, the purpose of the current study is to find out to which environmental factor the consumptive WF benchmark is most sensitive.

In order to analyse differences in consumptive WFs in relatively dry versus relatively wet years, we evenly group the forty-eight considered years (1961-2008) into relative dry, average and relatively wet years. We ranked the years based on the annual precipitation over the cropping area of winter wheat in China (Fig. 2a) and classified the sixteen years with the lowest precipitation into the group of dry years and the sixteen years with the highest precipitation into the group of wet years, with the other sixteen years remaining for the group of average years. The average annual precipitation levels of the relatively dry, average and relatively wet years are 760, 799 and 850 mm y^{-1}, respectively.

1 We also grouped the years considered into relatively cold, average and relatively warm years based on annual mean

2 temperature (Fig. 2b) and into years with relatively low, average and high ET_0 (Fig. 2c). The average annual mean temperatures

3 of the relative cold, average and warm years are 10.7, 11.2 and 11.8 ℃, respectively. The average annual ET_0 values in the

4 three categories of years are 874, 896 and 927 mm y^{-1}.

5

6 For determining WF benchmarks for different soil texture classes, the soil types in the USDA (U.S. Department of Agriculture)

7 soil texture triangles were grouped into four soil classes (Raes et al., 2011): sandy soils, loamy soils, sandy clayey soils, and

8 silty clayey soils. Each soil class has different ranges of field capacity, permanent wilting point and saturated water content

9 (Table 1). The difference between soil water content and permanent wilting point defines the total available soil water content

10 in the root zone. Given certain soil water content, a soil with a higher field capacity has less deep percolation. With the same

11 water input from precipitation or irrigation and the same soil water content, soils with a smaller saturated soil water content

12 will generate more surface runoff (Raes et al., 2011). Figure 3 shows the spatial distribution of the four soil classes across

13 mainland China.

14

15 For determining WF benchmarks for different climate zones, we classify climate based on UNEP's aridity index (AI)

16 (Middleton and Thomas, 1997, 1992). The AI is an indicator of dryness, defined as the ratio of precipitation to reference

17 evapotranspiration, with five levels of aridity: hyper-arid (AI< 0.05), arid (0.05 < AI < 0.2), semi-arid (0.2 < AI < 0.5), dry

18 sub humid (0.5 < AI < 0.65), and humid (AI > 0.65). To determine the geographic spread of the five climate zones in China

19 we used the data on annual precipitation and ET_0 averaged over the period 1961-2008 at 30 by 30 arc min resolution (Harris

20 et al., 2014) (Fig. 4). In the current study, we group the five climate zones into two broad zones: the arid-semi-arid (Arid) zone

21 (AI < 0.5) and the humid-semi-humid (Humid) zone (AI >0.5).

22 **3 Result**

23 **3.1 Benchmark levels for the consumptive WF as determined for different years and for rain-fed and irrigated**
24 **croplands separately**

25 We calculated the benchmark levels at different production percentiles for the consumptive WF of winter wheat ($m^3 t^{-1}$) for the

26 country as a whole, year by year, for the period 1961-2008. The results are summarized in Fig. 5. The benchmarks, determined

27 per year and per production percentile, generally vary within ±20% of the long-term mean value over the period 1961-2008.

28 We find that the best 10% of winter wheat production in China (with smallest WFs) has a maximum long-term average

29 consumptive WF of 777 $m^3 t^{-1}$, which is larger than the maximum consumptive WF of the best 10% of wheat production

30 globally (592 $m^3 t^{-1}$) that was reported by Mekonnen and Hoekstra (2014). We note here that the figures are not fully

31 comparable, because Mekonnen and Hoekstra (2014) consider total wheat (both spring and winter wheat), use another model

32 and consider another period. We find that the best 20% of winter wheat production in China has a maximum long-term average

1 consumptive WF of 825 m^3 t^{-1}, which is *smaller* than the reported maximum consumptive WF of the best 20% of wheat

2 production globally (992 m^3 t^{-1}). Finally, we find that the best 25% of winter wheat production in China has a maximum long-

3 term average consumptive WF of 849 m^3 t^{-1}, which is again *smaller* than the maximum consumptive WF of the best 25% of

4 wheat production globally (1069 m^3 t^{-1}).

5

6 The national average consumptive WF of rain-fed winter wheat (1120 m^3 t^{-1}) is larger than the national average consumptive

7 WF of irrigated winter wheat (1075 m^3 t^{-1}). However, the benchmark levels determined by the best 10%, 20% and 25% of

8 production for rain-fed winter wheat are lower than for irrigated winter wheat. The reason is that the yields in rain-fed

9 production are generally higher than the yields in irrigated production at the same benchmark percentile. The highest rain-fed

10 yields occur in the southern wet area with sufficient precipitation over the cropping period, so that little water stress results in

11 high rain-fed yields. The WF benchmarks for irrigated winter wheat are 8% (for the 10th production percentile) to 10% (for the

12 25th production percentile) higher than for rain-fed winter wheat.

3.2 Benchmark levels for the consumptive WF for dry versus wet years

14 In a relatively dry or wet year, when considering winter wheat areas in China as a whole, we do not find typically different

15 consumptive WFs in winter wheat production (Table 2). The WF benchmarks are consistently higher in dry than in wet years

16 (1-3%), but the differences between benchmark levels for the consumptive WF for dry versus wet years are small compared

17 to the variations within the dry and wet year categories (±11-14%).

3.3 Benchmark levels for the consumptive WF for warm versus cold years

19 Overall, considering irrigated and rain-fed croplands together, WF benchmarks for relatively warm years are 7-8% smaller

20 than for relatively cold years, which is not much when seen in the context of fluctuations in the WFs within the three

21 temperature categories (Table 3). In irrigated areas, WF benchmarks for warm years are 11% smaller, on average, than for

22 cold years. In rain-fed areas, WF benchmarks for warm years are smaller than for cold years as well, but WF benchmarks in

23 average years are not in between the WF benchmarks found for cold and warm years but higher than both. The lower values

24 in cold years relate to lower ET, while the lower values in warm years relate to higher yields.

25

26 The findings when considering different ET$_0$ classes are similar when looking at the different temperature classes (Table 4).

27 Overall, considering irrigated and rain-fed croplands together, WF benchmarks for years with high ET$_0$ are on average 5%

28 smaller than for years with average ET$_0$ and only 2% smaller than for years with low ET$_0$. Again, differences between

29 consumptive WFs for years with relatively low or high ET$_0$ are small when seen in the context of fluctuations in the WFs

30 within the three ET$_0$ categories (±3-6%).

3.4 Benchmark levels for the consumptive WF for different soil classes

Tables 5 shows the consumptive WFs of winter wheat at different production percentiles in four soil classes in China. The simulated winter wheat production in sandy clayey soils accounts for 60% of national total, followed by the production in sandy soils (24%), silty clayey soils (8%) and loamy soils (8%) in average over the studied period. No consistent trends can be observed when we compare the benchmarks across the different soil classes. Overall, when we take irrigated and rain-fed fields together, the WF benchmarks for sandy soils are 10-12% lower than the WF benchmarks for loamy soils. More specifically, we find that the WF benchmarks for irrigated winter wheat in sandy soils are about 15% smaller than the WF benchmarks for the other three soil classes, due to relatively low ET. Without water stress, as is the case in the irrigated croplands, soil evaporation from sandy soils is less than from the other soil types because of the fast percolation of water below the root zone in the sandy soils, causing lower ET over the cropping period (Asseng et al., 2001). At rain-fed fields with limited water availability, crop yields are mainly affected by the soil water holding capacity. Therefore, consumptive WFs in sandy soils are larger than in the other three soils, due to the smaller crop yield in case of poorer water holding capacity. The observed differences in WFs of winter wheat in different soil classes agree with the experimental observations by Tolk and Howell (2012) for the case of irrigated sunflower in a semiarid environment as well as with the fieldwork-based simulations by Asseng et al. (2001) for irrigated and rain-fed wheat in the Mediterranean climatic region of Western Australia.

3.5 Benchmark levels for the consumptive WF for different climate zones

Consumptive WFs of winter wheat at different production percentiles in arid and humid zones in China are shown in Table 6. Significant differences between the benchmarks for different climate zones can be observed. Overall, considering irrigated and rain-fed croplands together, WF benchmarks for the humid zone are 26% (for the 10th production percentile) to 31% (for the 25th production percentile) smaller than for the arid zone. The WF benchmarks for winter wheat in China as a whole (when we take the arid and humid zones together) are close to the benchmarks for the humid zone, caused by the fact that most (96% in average over the study period) of the simulated winter wheat production in China occurs in the humid zone.

In the irrigated areas, WF benchmarks for the humid zone are 26-30% smaller than for the arid zone; in the rain-fed areas, they are 29-43% smaller. The relatively large WFs in rain-fed fields in the arid zone logically follow from the water stress and resultant low yields. For the irrigated fields, the larger WFs in the arid zone are caused by the relatively high ET_0 and ET. The results confirm the findings from previous studies that the WF of crops, especially rain-fed crops, is negatively correlated with precipitation and positively correlated with ET_0 (Zwart et al., 2010;Zhuo et al., 2014). The differences between the WF benchmarks for irrigated and rain-fed winter wheat are 7-9% in the humid zone and 3-11% in the arid zone.

Figure 6 shows, for both the humid and arid part of China, for the various winter wheat production areas whether they contribute to the best 10% of national winter wheat production in that climate zone (in the sense of having smallest WFs), to

the next best 10%, to the best 5% after that, or to the worst 75% (with WFs beyond the 25th percentile benchmark). Within the arid zone, consumptive WFs below the 25th percentile benchmark level were mostly located in Xinjiang province, with relatively high irrigation density (~98% of the harvested area). In the humid zone, consumptive WFs below the 25th percentile benchmark level were gathered in the southwest, where ET_0 is smaller than in other places (Fig. 4b).

3.6 Water saving potential by reducing WFs to selected benchmark levels

The WF benchmarks for different climate zones differ much more significantly (26-31%) than for different soils (10-12%). WF benchmarks differ even less if we compare irrigated versus rain-fed fields (8-10%), warm versus cold years (7-8%), or wet versus dry years (1-3%). Therefore, when determining benchmark levels for the consumptive WF of a crop, it seems most useful to primarily distinguish between different climate zones, at least in the case of winter wheat in China. In this section, we analyse the potential water saving if actual consumptive WFs of winter wheat throughout China were reduced to the climate-specific benchmark levels set by the best 10% of Chinese winter wheat production (1042 $m^3 t^{-1}$ for arid areas and 776 $m^3 t^{-1}$ for humid areas), the best 20% of Chinese winter wheat production (1170 $m^3 t^{-1}$ for arid areas and 819 $m^3 t^{-1}$ for humid areas), or the best 25% of Chinese winter wheat production (1224 $m^3 t^{-1}$ for arid areas and 841 $m^3 t^{-1}$ for humid areas).

Taking the estimated actual consumptive WFs of winter wheat in 2005, an average climatic year, as calibrated by the provincial statistics on yield of winter wheat (NBSC, 2013), we find that consumptive WFs in 75% of the planted grids in arid zones and in 96% of the planted grids in humid zones are over the 25th percentile benchmarks. This is largely due to low actual versus potential yields. Figure 7 shows differences between actual provincial yields of winter wheat and the simulated yield potentials from the current study (assuming no crops stresses except water stress in rain-fed areas). The largest yield gaps occur in the southern provinces in the humid zone. The largest yield gap was observed in Fujian province. South China has 81% of national blue water resources (Jiang et al., 2015). However, the risk of water shortage is increasing in the wet South with the operation of the South-to-North Water Transfer Project and the increasing competition for water resources between different sectors (Barnnet et al., 2015). Therefore, water saving and benchmarking WF for the South China are as equally important as for the drier North.

Table 7 shows the (green plus blue) water saving that would be achieved if actual consumptive WFs of winter wheat everywhere in China were reduced to the climate-differentiated WF benchmark levels set by the 10th, 20th and 25th percentiles of production, in an average year (2005). We find that if in both the arid and humid zone the actual consumptive WFs were reduced to the respective 25th percentile benchmark level, the water saving in an average year would be 53% of the current water consumption at winter wheat fields in China, which is 201 billion $m^3 y^{-1}$ in absolute terms. We further find that the water saving potential in the arid zone is substantially higher than in the humid zone.

10

3.7 Discussion

The consumptive WF of a crop in $m^3\ t^{-1}$ most strongly depends on the crop yield in $t\ ha^{-1}$ and much less on the evapotranspiration from the crop over the growing period in $m^3\ ha^{-1}$ (Tuninetti et al., 2015;Mekonnen and Hoekstra, 2011). The simulated consumptive WFs of winter wheat in China have been based on modelling under a hypothetical condition without effects of managerial factors on crop growth. For evaluating our simulations of crop growth, we compared the simulated averaged yields of winter wheat of Chinese provinces for 1961-1990 to the corresponding agro-climatic attainable yields at different agricultural input levels in the GAEZ database (FAO/IIASA, 2011) (Fig. 8). The GAEZ agro-climatic attainable yields account for different levels of yield constraints from four factors in addition to water stress: (i) pest, diseases and weed damages on plant growth, (ii) direct and indirect climatic damages on quality of produce, (iii) efficiency of farming operations, and (iv) frost hazards. Current simulated yields of irrigated winter wheat are closest to the agro-climatically attainable yields with intermediate input levels and the yields of rain-fed winter wheat are closest to the agro-climatically attainable yields with high input levels. The simulated national average yield in the current study ($6.5\ t\ ha^{-1}$) is 23% higher than the attainable wheat yield for China in the year 2000 ($5.3\ t\ ha^{-1}$) estimated by Mueller et al. (2012).

The study shows that climate is the primary factor to be considered when setting consumptive WF benchmarks. This finding is probably little sensitive to the model used; the precise WF benchmark figures found per climate zone, however, will be more sensitive to the model used. Subsequent studies, comparing WF benchmark estimates per climate zone using different models, are necessary to quantify the uncertainty in the WF benchmarks presented in this study.

Further research could also explore whether crop varieties used should play a role when developing WF benchmarks, given the fact that some crop varieties may inherently be more productive than others. On the other hand, one could also consider that choosing a productive crop variety is part of the managerial choices. Since crop variety is not a given environmental condition but a choice, one could argue that accepting a less strict WF reference level for a less productive crop variety cannot be justified.

An important remaining research question is also how combinations of specific techniques and practices can actually lead to the WF reductions that will be necessary in different locations if Chinese government would adopt certain WF benchmarks as targets to achieve greater water productivity. Suppose, for example, that two WF benchmarks for winter wheat were adopted in China: $1224\ m^3\ t^{-1}$ for arid areas and $841\ m^3\ t^{-1}$ for humid areas. Although the simulations suggest that these levels are feasible throughout the arid and humid zone, respectively, whatever is the soil, whether fields are rain-fed or irrigated, whether it is a cold or warm year, and whether it is a dry or wet year, in some places it will be harder and more would need to be done than in other places.

We studied benchmarks for combined green and blue WFs and did not look at each colour separately. For rain-fed lands, the benchmark levels presented in this study are obviously green WF benchmarks. For irrigated lands, the presented benchmark levels for overall consumptive WFs would need further specification into green and blue. Further research would need to be done to translate a certain benchmark level for the overall consumptive WF of a crop into a specific blue WF benchmark level per specific location as a function of the amount of rain per location, recognizing that the blue ratio in the WF will need to be larger if less green water is available.

4 Conclusions

Based on the case of winter wheat in China we find that (i) benchmark levels for the consumptive WF determined for individual years for the country as a whole remain within a range of ±20% around long-term mean levels over 1961-2008; (ii) the WF benchmarks for irrigated winter wheat are 8-10% larger than those for rain-fed winter wheat; (iii) WF benchmarks for wet years are on average 1-3% smaller than for dry years, (iv) WF benchmarks for warm years are on average 7-8% smaller than for cold years, (v) WF benchmarks differ by about 10-12% across different soil texture classes; and (vi) WF benchmarks for the humid zone are 26-31% smaller than for the arid zone, which has relatively higher ET_0 in general and lower yields in rain-fed fields. Therefore, we conclude that when determining benchmark levels for the consumptive WF of a crop, it is useful to primarily distinguish between different climate zones. We estimated that when in both the arid and humid zone the actual consumptive WFs are reduced to climate-specific benchmark levels set by the 25th percentile of production, the water saving in an average year would be 53% of the current water consumption at winter wheat fields in China, with greatest relative savings in the arid zone.

Author contribution

A. Y., L. Z. and M. M. M. designed the study. L. Z. carried it out. L. Z. prepared the manuscript with contributions from all co-authors.

References

Abedinpour, M., Sarangi, A., Rajput, T.B.S., Singh, M., Pathak, H. and Ahmad, T.: Performance evaluation of AquaCrop model for maize crop in a semi-arid environment, Agr. Water Manage., 110, 55-66, doi: 10.1016/j.agwat.2012.04.001, 2012.

Andarzian, B., Bannayan, M., Steduto, P., Mazraeh, H., Barati, M.E., Barati, M.A. and Rahnama, A.: Validation and testing of the AquaCrop model under full and deficit irrigated wheat production in Iran, Agr. Water Manage., 100(1), 1-8, doi: 10.1016/j.agwat.2011.08.023, 2011.

Asseng, S., Turner, N. C., and Keating, B. A.: Analysis of water- and nitrogen-use efficiency of wheat in a Mediterranean climate, Plant Soil, 233, 127-143, doi: 10.1023/A:1010381602223, 2001.

Batjes, N.: ISRIC-WISE derived soil properties on a 5 by 5 arc-minutes global grid (ver. 1.2), Wageningen, The Netherlands, 2012.

Barnett, J., Rogers, S., Webber, M., Finlayson, B. and Wang, M.: Sustainability: Transfer project cannot meet China's water needs, Nature, 527 (7578), 295-297, doi: 10.1038/527295a, 2015.

Brauman, K. A., Siebert, S., and Foley, J. A.: Improvements in crop water productivity increase water sustainability and food security—a global analysis, Environ. Res. Lett., 8, 024030, doi:10.1088/1748-9326/8/2/024030, 2013.

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

Confalonieri, R., Orlando, F., Paleari, L., Stella, T., Gilardelli, A., Alberti, L., Alberti, P., Atanassiu, S., Bonaiti, M., Cappelletti, G., Ceruti, M., Confalonieri, A., Corgatelli, G., Corti, P., Dell'Oro, M., Ghidoni, A., Lamarta, A., Maghini, A., Mambretti, M., Manchia, A., Massoni, G., Mutti, P., Pariani, S., Pasini, D., Pesenti, A., Pizzamiglio, G., Ravasio, A., Rea, A., Santorsola, D., Serafini, G., Slavazza, M. and Acutis, M.: Uncertainty in crop model predictions: what is the role of users? Environ. Modell. Softw., 81, 165-173, doi: 10.1016/j.envsoft.2016.04.009, 2016.

Farahani, H., Izzi, G. and Oweis, T.Y.: Parameterization and evaluation of the AquaCrop model for full and deficit irrigated cotton, Agron. J., 101(3), 469-476, doi: 10.2134/agronj2008.0182s, 2009.

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., Mueller, N. D., O'Connell, C., Ray, D. K., West, P. C., Balzer, C., Bennett, E. M., Carpenter, S. R., Hill, J., Monfreda, C., Polasky, S., Rockstrom, J., Sheehan, J., Siebert, S., Tilman, D., and Zaks, D. P. M.: Solutions for a cultivated planet, Nature, 478, 337-342, doi: 10.1038/nature10452, 2011.

García-vila, M., Fereres, E., Mateos, L., Orgaz, F. and Steduto, P.: Deficit irrigation optimization of cotton with AquaCrop, Agron. J., 101(3), 477-487, doi:10.2134/agronj2008.0179s, 2009.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623-642, doi: 10.1002/joc.3711, 2014.

Heng, L.K., Hsiao, T.C., Evett, S., Howell, T. and Steduto, P.: Validating the FAO AquaCrop model for irrigated and water deficient field maize, Agron. J., 101(3), 488-498, doi: 10.2134/agronj2008.0029xs, 2009.

Hoekstra, A. Y., Mekonnen, M. M., Chapagain, A. K., Mathews, R. E., and Richter, B. D.: Global monthly water scarcity: blue water footprints versus blue water availability, PLoS ONE, 7, e32688, doi: 10.1371/journal.pone.0032688, 2012.

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

Hoekstra, A. Y.: Sustainable, efficient, and equitable water use: the three pillars under wise freshwater allocation, Wiley Interdisciplinary Reviews: Water, 1, 31-40, doi: 10.1002/wat2.1000, 2014.

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

Hsiao, T. C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D., and Fereres, E.: AquaCrop-The FAO Crop Model to Simulate Yield Response to Water: III. Parameterization and Testing for Maize, Agron J, 101, 448-459, doi: 10.2134/agronj2008.0218s, 2009.

Jin, X. L., Feng, H. K., Zhu, X. K., Li, Z. H., Song-S. N., Song, X. Y., Yang, G. J., Xu, X. G. and Guo, W. S.: Assessment of the AquaCrop model for use in simulation of irrigated winter wheat canopy cover, biomass, and grain yield in the North China Plain, PLoS ONE, 9(1), e86938, doi: 10.1371/journal.pone.0086938, 2014.

1 Kumar, P., Sarangi, A., Singh, D.K. and Parihar, S.S.: Evaluation of AquaCrop model in predicting wheat yield and water
2   productivity under irrigated saline regimes, Irrig. Drain., 63(4), 474-487, doi:10.1002/ird.1841, 2014.
3 Mekonnen, M. M., and Hoekstra, A. Y.: The green, blue and grey water footprint of crops and derived crop products, Hydrol.
4   Earth Syst. Sc., 15, 1577-1600, doi: 10.5194/hess-15-1577-2011, 2011.
5 Mekonnen, M. M., and Hoekstra, A. Y.: Water footprint benchmarks for crop production: A first global assessment, Ecol.
6   Indic., 46, 214-223, doi: 10.1016/j.ecolind.2014.06.013, 2014.
7 Middleton, N., and Thomas, D. S. G.: World atlas of desertification, Arnold, London, UK, 1992.
8 Middleton, N., and Thomas, D. S. G.: World atlas of desertification, Ed. 2, Arnold, London, UK, 1997.
9 Mkhabela, M.S. and Bullock, P.R.: Performance of the FAO AquaCrop model for wheat grain yield and soil moisture
10   simulation in Western Canada, Agr. Water Manage., 110, 16-24, doi:10.1016/j.agwat.2012.03.009, 2012.
11 Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A.: Closing yield gaps through nutrient
12   and water management, Nature, 490, 254-257, doi: 10.1038/nature11420, 2012.
13 Portmann, F. T., Siebert, S., and Doll, P.: MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000:
14   A new high-resolution data set for agricultural and hydrological modeling, Global Biogeochem. Cy., 24, GB1011, doi:
15   Artn Gb101110.1029/2008gb003435, 2010.
16 Raes, D., Steduto, P., Hsiao, T. C., and Fereres, E.: AquaCrop-The FAO Crop Model to Simulate Yield Response to Water:
17   II. Main Algorithms and Software Description, Agron. J., 101, 438-447, doi: 10.2134/agronj2008.0140s, 2009.
18 Raes, D., steduto, P., Hsiao, T. C., and Fereres, E.: Reference manual AquaCrop version 4.0, Rome, Italy, 2011.
19 Siebert, S., and Doll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential
20   production losses without irrigation, J. Hydrol., 384, 198-217, doi: 10.1016/j.jhydrol.2009.07.031, 2010.
21 Steduto, P., Hsiao, T. C., and Fereres, E.: On the conservative behavior of biomass water productivity, Irrigation Sci., 25, 189-
22   207, 2007.
23 Steduto, P., Hsiao, T. C., Raes, D., and Fereres, E.: AquaCrop-The FAO Crop Model to Simulate Yield Response to Water: I.
24   Concepts and Underlying Principles, Agron. J., 101, 426-437, doi: 10.2134/agronj2008.0139s, 2009.
25 Stricevic, R., Cosic, M., Djurovic, N., Pejic, B. and Maksimovic, L.: Assessment of the FAO AquaCrop model in the simulation
26   of rainfed and supplementally irrigated maize, sugar beet and sunflower, Agr. Water Manage., 98(10),1615-1621, doi:
27   10.1016/j.agwat.2011.05.011, 2011.
28 Tolk, J. A., and Howell, T. A.: Sunflower water productivity in four Great Plains soils, Field Crop. Res., 127, 120-128, doi:
29   10.1016/j.fcr.2011.11.012, 2012.
30 Tuninetti, M., Tamea, S., D'Odorico, P., Laio, F., and Ridolfi, L.: Global sensitivity of high‐resolution estimates of crop
31   water footprint, Water Resour. Res., 51, 8257-8272, doi: 10.1002/2015WR017148, 2015.
32 Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: Sensitivity and uncertainty in crop water footprint accounting: a case study
33   for the Yellow River Basin, Hydrol. Earth Syst. Sc., 18, 2219-2234, doi:10.5194/hess-18-2219-2014, 2014.
34 Zhuo, L., Mekonnen, M. M., Hoekstra, A. Y. and Wada, Y.: Inter- and intra-annual variation of water footprint of crops and
35   blue water scarcity in the Yellow River basin (1961–2009), Adv. Water Resour., 87, 29-41, doi:
36   10.1016/j.advwatres.2015.11.002, 2016a.
37 Zhuo, L., Mekonnen, M. M., and Hoekstra, A. Y.: The effect of inter-annual variability of consumption, production, trade and
38   climate on crop-related green and blue water footprints and inter-regional virtual water trade: A study for China (1978-
39   2008). Water Res., 94, 73-85, doi: 10.1016/j.watres.2016.02.037, 2016b.
40 Zwart, S. J., Bastiaanssen, W. G. M., de Fraiture, C., and Molden, D. J.: A global benchmark map of water productivity for
41   rainfed and irrigated wheat, Agr. Water Manage., 97, 1617-1627, doi: 10.1016/j.agwat.2010.05.018, 2010.
42
43

44 **Tables**

45

  Table 1. Soil classes.

Soil class	Soil types	Soil water content (vol %)		
		Field capacity	Permanent wilting point	Saturation
Sandy	Sand, loamy sand, sandy loam	9 - 28	4 - 15	32 - 51
Loamy	Loam, silt loam, silt	23 - 42	6 - 20	42 - 55
Sandy clayey	Sandy clay, sandy clay loam, clay loam	25 - 45	16 - 34	40 - 53
Silty clayey	Silty clay loam, silty clay, clay	40 - 58	20 - 42	49 - 58

Source: Raes et al. (2011).

1

2

Table 2. Benchmark levels for the consumptive water footprint (WF) benchmarks ($m^3 \ t^{-1}$) of winter wheat for relative dry, average and wet years in China.

Crop		Consumptive WF ($m^3 \ t^{-1}$) at different production percentiles*			
		10th	20th	25th	Average
Winter wheat	Dry years	787±69	837±70	858±71	1103±82
	Average years	763±107	826±72	849±74	1073±97
	Wet years	770±68	813±60	838±50	1048±77
Irrigated winter wheat	Dry years	822±118	862±110	876±112	1095±110
	Average years	814±97	856±97	881±98	1078±93
	Wet years	799±97	850±100	870±96	1052±96
Rain-fed winter wheat	Dry years	757±44	802±57	812±56	1121±97
	Average years	736±62	771±70	783±70	1074±133
	Wet years	755±96	784±103	794±104	1164±561

* Data are mean ± S.D. for the years 1961-2008.

1

2

Table 3. National consumptive water footprint (WF) benchmarks ($m^3 t^{-1}$) of winter wheat for relative cold, warm and average years in China.

Crop		Consumptive WF ($m^3 t^{-1}$) at different production percentiles*			
		10th	20th	25th	Average
Winter wheat	Cold years	795±101	848±63	870±67	1103±96
	Average years	794±79	840±66	864±58	1087±82
	Warm years	732±42	788±58	811±57	1033±70
Irrigated winter wheat	Cold years	862±86	902±87	924±87	1121±86
	Average years	810±107	863±102	878±96	1083±93
	Warm years	763±96	804±93	824±96	1022±98
Rain-fed winter wheat	Cold years	760±59	791±68	798±69	1088±144
	Average years	772±95	821±99	831±100	1218±553
	Warm years	716±31	744±40	761±44	1053±63

* Data are mean ± S.D. for the years 1961-2008.

1

2

3

Table 4. National consumptive water footprint (WF) benchmarks ($m^3 t^{-1}$) of winter wheat for relative low-, high- and average-ET_0 years in China.

Crop		Consumptive WF ($m^3 t^{-1}$) at different production percentiles*			
		10th	20th	25th	Average
Winter wheat	Low-ET_0 years	774±99	822±64	841±62	1065±82
	Average years	806±80	846±73	866±76	1095±107
	High-ET_0 years	741±51	808±62	839±58	1065±70
Irrigated winter wheat	Low-ET_0 years	831±111	874±108	892±106	1089±98
	Average years	820±105	868±96	887±96	1073±103
	High-ET_0 years	784±93	827±97	847±97	1064±102
Rain-fed winter wheat	Low-ET_0 years	749±55	774±56	781±54	1038±100
	Average years	784±90	828±98	841±98	1249±550
	High-ET_0 years	716±72	755±59	767±58	1072±78

* Data are mean ± S.D. for the years 1961-2008.

1

2

3

Table 5. Benchmark levels for the consumptive water footprint (WF) ($m^3 t^{-1}$) of winter wheat for different soil classes in China.

Crop	Soil class	Consumptive WF ($m^3 t^{-1}$) at different production percentiles*			
		10th	20th	25th	Average
Winter wheat	Sandy	748±143	814±115	834±116	1017±125
	Loamy	846±53	912±77	928±73	1108±74
	Sandy clayey	788±76	848±61	881±66	1071±48
	Silty clayey	822±48	895±43	912±46	963±22
Irrigated winter wheat	Sandy	767±158	782±177	846±128	1000±126
	Loamy	931±91	937±93	996±70	1189±107
	Sandy clayey	879±98	932±98	969±102	1164±100
	Silty clayey	920±68	942±72	958±66	1070±52
Rain-fed winter wheat	Sandy	785±58	834±88	850±96	1151±272
	Loamy	757±77	822±73	843±73	1040±160
	Sandy clayey	764±66	799±68	818±70	1096±129
	Silty clayey	769±62	814±60	837±60	931±103

* Data are mean ± S.D. for the years 1961-2008.

1

2

Table 6. Benchmarks for the consumptive water footprint (WF) (m^3 t^{-1}) of winter wheat for different climate zones in China.

Crop	Climate zones	Consumptive WF (m^3 t^{-1}) at different production percentile*			
		10th	20th	25th	Average
Winter wheat	Arid	1042±100	1170±130	1224±125	1757±200
	Humid	776±70	819±66	841±66	1044±83
	Overall	777±72	825±67	849±65	1075±87
Irrigated winter wheat	Arid	1088±66	1205±73	1245±84	1399±163
	Humid	807±104	853±100	872±99	1055±97
	Overall	812±103	856±100	875±100	1075±99
Rain-fed winter wheat	Arid	1058±310	1311±406	1399±415	2919±1004
	Humid	749±70	784±78	795±79	1076±338
	Overall	750±70	785±78	796±78	1120±332

* Data are mean ± S.D. for the years 1961-2008.

1

2

Table 7. Water saving if actual consumptive water footprint (WF) of winter wheat everywhere in China were reduced to the climate-differentiated WF benchmark levels set by the 10th, 20th and 25th percentiles of production, in an average year (2005).

Climate zones	Water saving when actual consumptive WF of winter wheat everywhere in China were to be reduced to a certain percentile benchmark level		
	10th	20th	25th
Arid	83%	81%	80%
Humid	49%	46%	45%
Overall	56%	54%	53%

* Data are mean ± S.D. for the years 1961-2008.

1

2

1 **Figures**

2
3

Fraction of irrigated area of winter wheat

Legend:
- 0 - 20%
- 20% - 40%
- 40% - 60%
- 60% - 80%
- 80% - 100%
- Harvested area
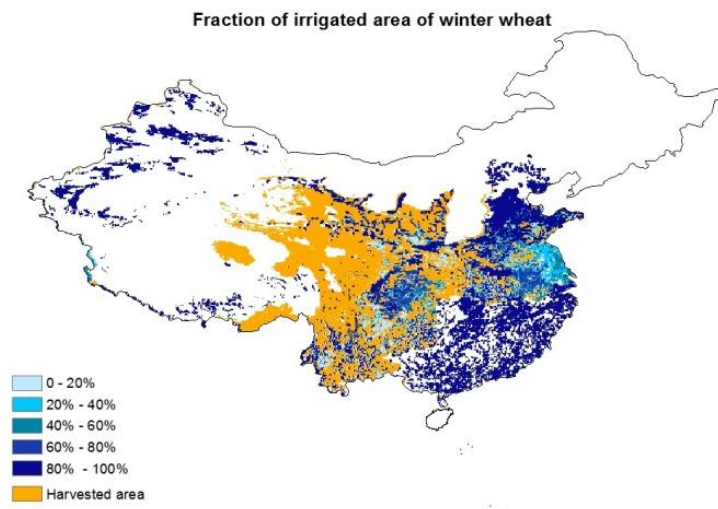
4
5 Figure 1. Harvested winter wheat areas in China in the year 2000 and fractions of the harvested areas irrigated. Data source:
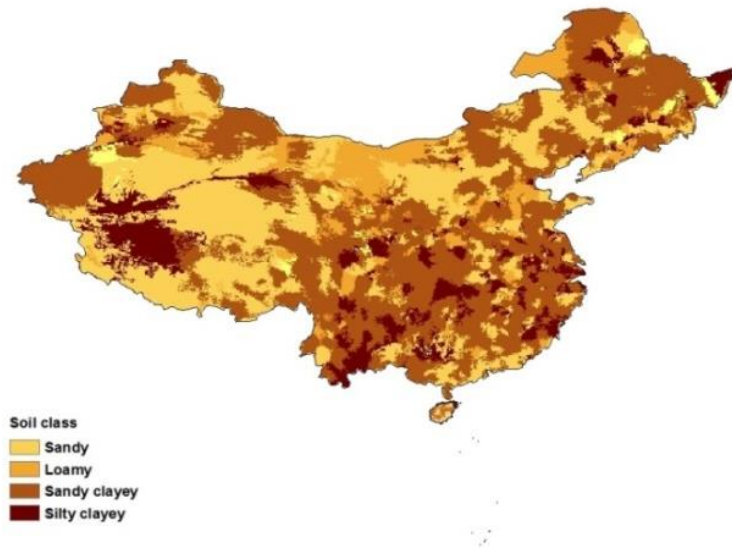6 Portmann et al. (2010).

7
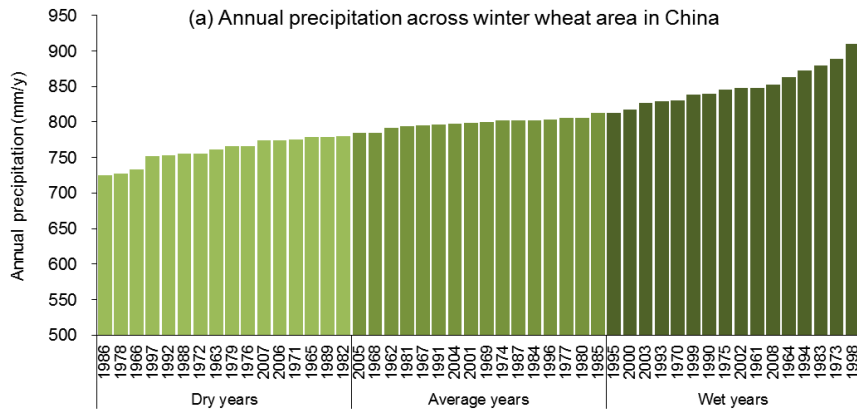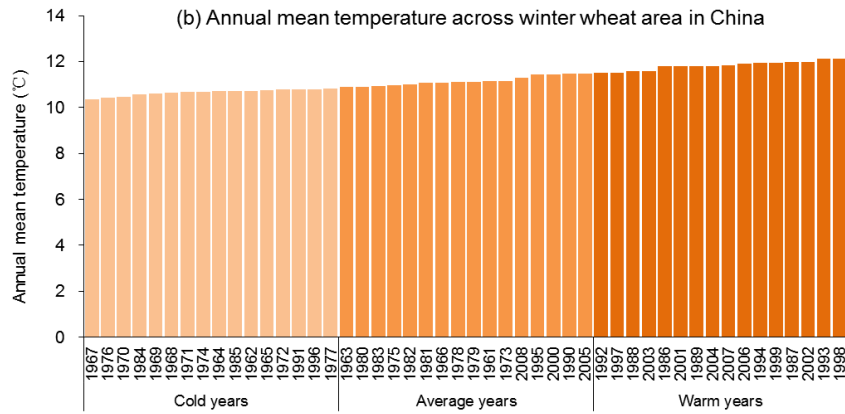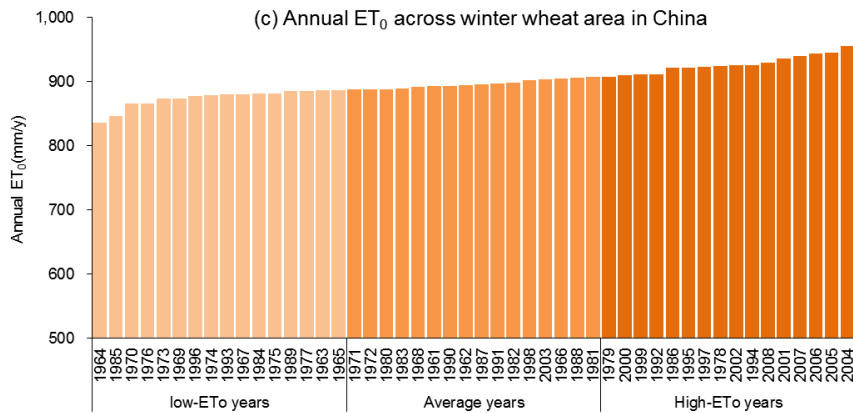
Figure 2. Annual precipitation (a), mean temperature (b), and ET_0 (c) over the cropping area of winter wheat in China for the years in the period 1961-2008, ranked from lowest to highest values. Data source: Harris et al. (2014).

1
2 Figure 3. Soil classes across mainland China, generated from the ISRIC Soil and Terrain database for China. Data source:
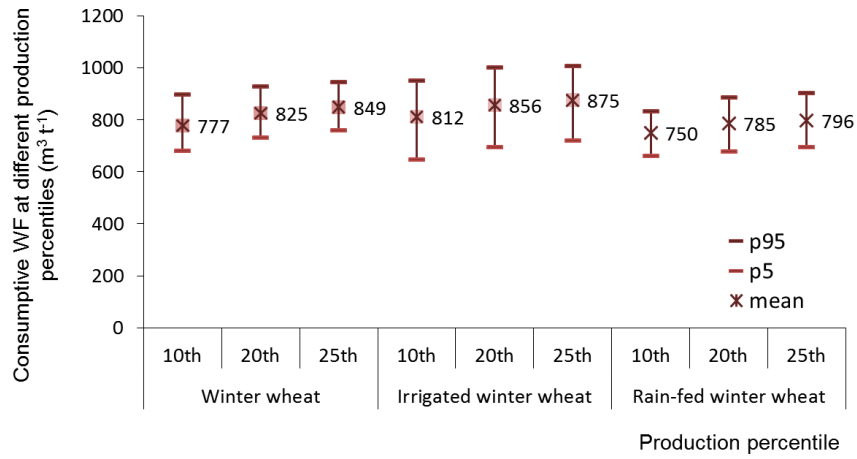3 Dijkshoorn et al. (2008).

4

Annual precipitation

[mm]
< 50
50 - 200
200 - 400
400 - 600
600 - 1200
> 1200

1　　　　(a)

Annual reference evapotranspiration

[mm]
< 600
600 - 800
800 - 1000
1000 - 1200
> 1200

(b)

Hyperarid
Arid
Semi arid
Dry subhumid
Humid
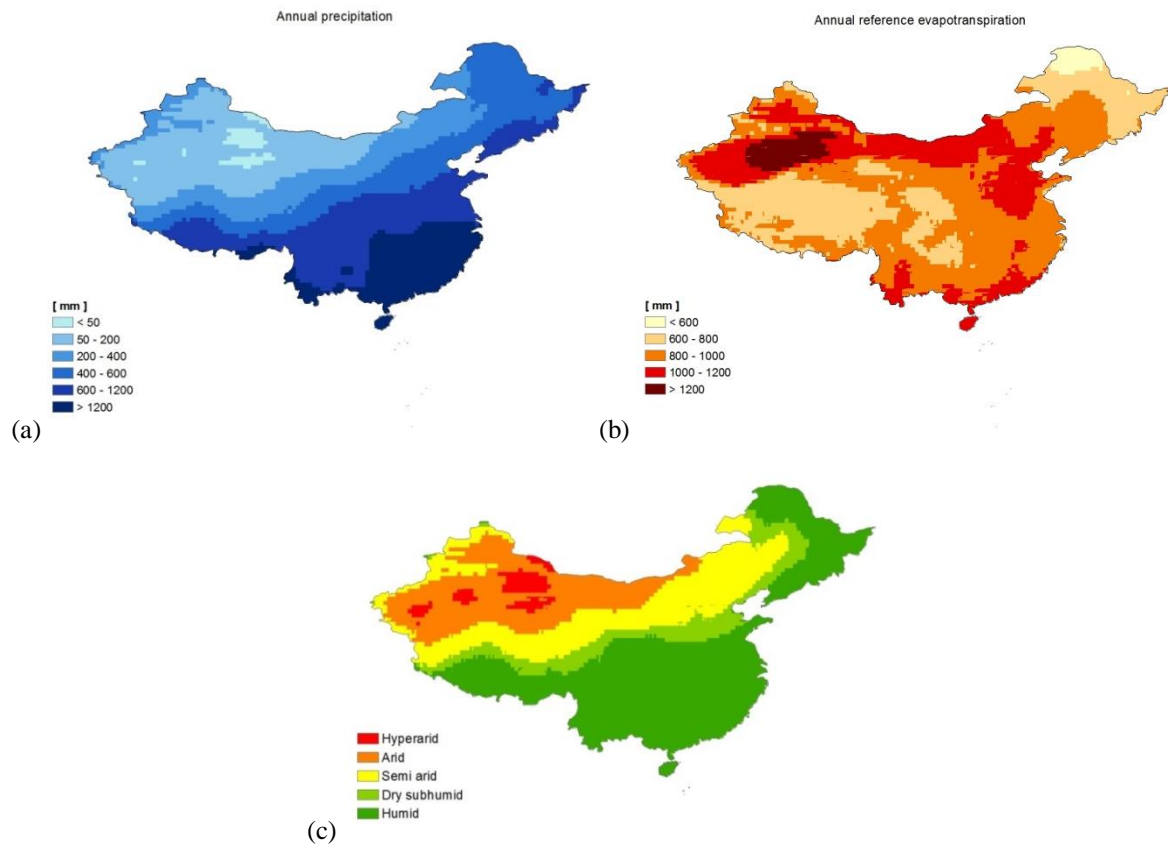
2　　　　　　　　　　(c)

3

4 Figure 4. Zoning of annual precipitation (a), annual reference evapotranspiration (b), and aridity (c) in China (1961-2008).
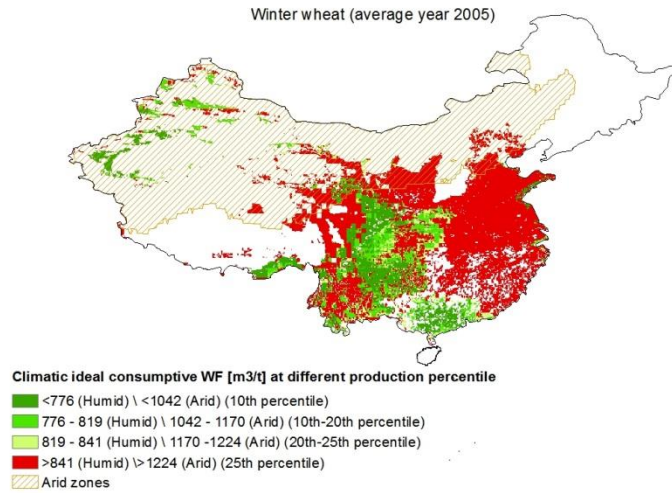
5 Data source: Harris et al. (2014).

6

1

2 Figure 5. Benchmark levels for the consumptive water footprint (WF) of winter wheat in China at different production

3 percentiles, considering all separate years in the period 1961-2008. Cross marks refer to the mean values; ranges refer to the

4 5% - 95% of accumulative frequencies.

5

Figure 6. Simulated consumptive water footprints (WFs) of winter wheat, categorized into four classes (the best 10% of production, the next best 10%, the second next best 5% and the worst 75% of production), accounting for different benchmark levels for humid versus arid part of China, for the year 2005 (climatic average year).

Actual yield vs. climatic ideal yield of winter wheat in China (2005)

Difference as percentage of climatic ideal yield [%]

- 0 - 10
- 10 - 20
- 20 - 40
- 40 - 60
- 60 - 80
- 80 - 93
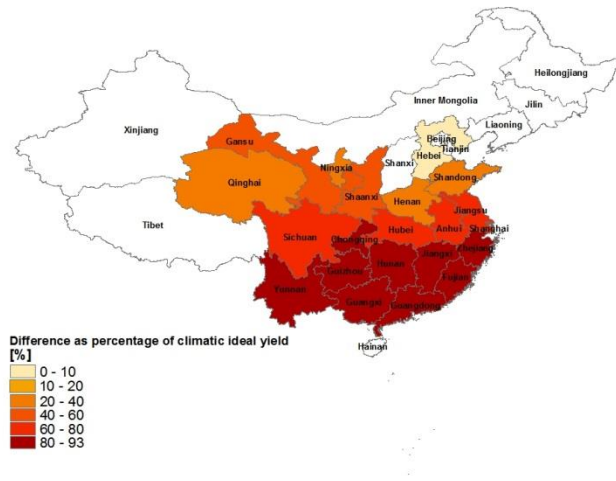
Figure 7. Differences between actual provincial yields of winter wheat in China in 2005 (NBSC, 2013) and simulated yields from the current study (assuming no crop stress except for water stress in rain-fed areas), expressed as percentage of the simulated yield.
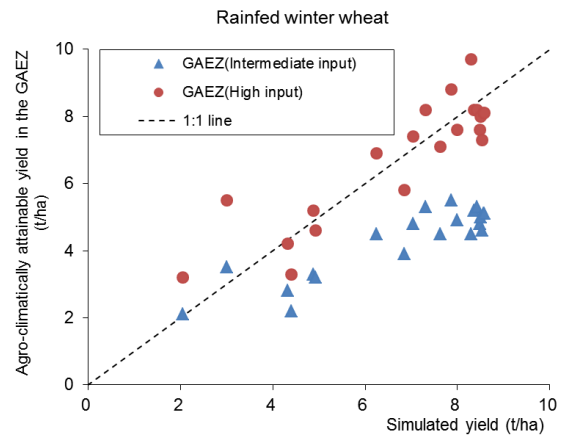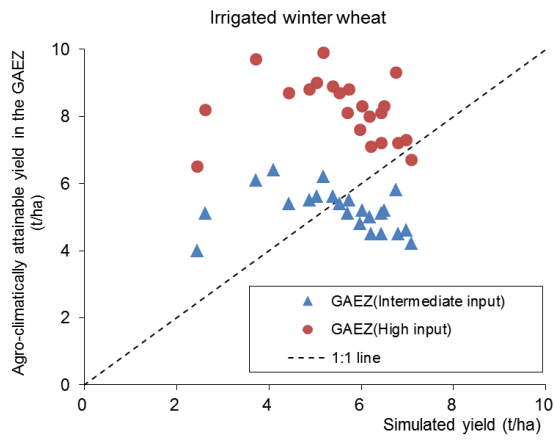
1

2 Figure 8. Comparison between the simulated yield of winter wheat and the agro-climatically attainable yield according to

3 (FAO/IIASA, 2011) at provincial level in China. Averaged over the period 1961-1990.

4

5