Evaluation of the FAO AquaCrop model for winter wheat on the North China Plain under deficit irrigation from field experiment to regional yield simulation

M. Anjum Iqbal a,⁎, Yanjun Shen a, Ruzica Stricevic b, Hongwei Pei a, Hongyoung Sun a, Ebrahim Amiri c, Angel Penas d, Sara del Río d

a Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang 050021, China
b University of Belgrade, Faculty of Agriculture, Nemanjina 6, 11080 Zemun, Serbia
c Department of Agriculture, Islamic Azad University, Lahijan Branch, Lahijan, Iran
d Department of Biodiversity and Environmental Management (Area of Botany), Mountain Livestock Farming Institute (Joint Center CSIC-ULE), Faculty of Biological and Environmental Sciences, University of Leon, Spain

A R T I C L E   I N F O

Article history:
Received 10 April 2013
Received in revised form 18 December 2013
Accepted 22 December 2013
Available online 19 January 2014

Keywords:
AquaCrop model
Winter wheat
Deficit irrigation
North China Plain

A B S T R A C T

Winter wheat is the main crop on the North China Plain (NCP), and in this region the most limiting factor for the crop is water. The objective of this study was to adapt and test the ability of the FAO-developed AquaCrop model (v3.1) to simulate winter wheat grain yield, biomass, actual evapotranspiration (ETa) and total soil water content (0–120 cm). Field experiments were conducted under deficit irrigation at the Luancheng Agro-ecosystem station (NCP) in 1998–2001, and the AquaCrop model was calibrated with treatment D (1999–2000); the rest of the data was used for validation of the model. The AquaCrop model was revalidated with data on measured grain yield from the experimental station for 1990–2010, considering actual field conditions. The second revalidation was done with the statistical grain yield for 1995–2010 in the study region. For the model validation, the significant differences between simulated and observed grain yield, biomass and ETa were in the order of: rainfed treatment > well-watered treatment > moderate water stress. Total soil water simulated by AquaCrop tends to follow closely the trend in the measured data, but with slight underestimations for irrigated treatments and significant overestimations for rainfed treatments. In general, errors in the model’s evaluation such as RMSE and Willmott’s d statistics were for grain yield (0.58 Mg ha−1, 0.92), biomass (0.87 Mg ha−1, 0.95), ETa (32.3 mm, 0.93) and soil water content (24.5–37.6 mm, 0.85–0.90). The overall results based on extensive validation and revalidation showed that AquaCrop is a valid model and can be used with a reliable degree of accuracy for optimizing winter wheat grain yield production and water requirement on the NCP.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The North China Plain (NCP) covers an area of 0.44 × 106 km2 and has a population of over 200 million; it is considered one of the most productive and intensively cultivated agricultural regions in China (Chen, 2003; Zhang et al., 2005). Due to its excellent thermal conditions and flat terrain, the North China Plain (NCP) has become one of the primary food-producing regions in the country (Zhang et al., 2008) and is known as the “breadbasket of China”. About 50–61% of the nation’s wheat, 31–33% of its maize and 42% of its cotton are produced in this region (CNS Bureau, 1999; World Bank, 2001; Wang et al., 2001, 2012; Zhang et al., 2003; China Agriculture Yearbook, 2002), although the main crop is winter wheat.

Annual precipitation on the NCP is highly variable, ranging from 300 to 1000 mm, with an average of about 500–600 mm (Zhang and You, 1996; Cao et al., 2013), more than 70% of which is concentrated in the maize-growing season (July–September). Precipitation during the wheat-growing season varies between 100 and 180 mm (Li et al., 2007, 2010) which can only meet approximately 25–40% of the water requirements throughout the wheat-growing season (October–June), leaving a water deficit of 200–300 mm (Liu et al., 2001; Zhang et al., 1995). Water demand is 400–500 mm (Liu et al., 2002; Li et al., 2008), which greatly exceeds precipitation. As a result, more than 70% of irrigation water resources are allocated for winter wheat in order to ensure maximum production. This 70% of the water used in the wheat-growing seasons comes from groundwater pumped from wells with a depth of over 40 m.
(Yuan and Shen, 2013), causing a rapid decline in the groundwater table on the NCP over an extension of more than 40,000 km² (Chen, 1999); this is considered to be the largest groundwater drawdown area in the world. Another study (Jiang and Zhang, 2004; Zhang et al., 2004; Fang et al., 2010) showed that in the Luancheng agricultural station, the most representative farmland on the NCP, there has been a 20-m decrease in groundwater over 30 years, and 10 m in 10 years on the whole of the North China Plain (World Bank, 2005). International agencies also agree that this region is facing a serious water shortage (FAO, 2003; New York Times, 2007), and this will limit wheat production on the NCP. Climate change studies on the NCP indicate that this situation will be further exacerbated with the predicted decline in water resources and the increase in agricultural water demand (Tao et al., 2003, 2005; Xiong et al., 2009; Bates et al., 2008; Falloon and Betts, 2010).

This situation is the result of excessive irrigation applied by the farmers on the NCP in pursuit of high yields, particularly in the wheat-growing season (Zhang et al., 2002; Yang et al., 2002). Water-use efficiency is still very low due to poor irrigation management practices such as flood irrigation (Wang et al., 2002; Deng et al., 2006; Shao et al., 2009).

In view of the above situation, there is an urgent need in this area for alternative strategies that ensure sustainable use of the limited groundwater resources at regional and site-specific levels (Qadir et al., 2003).

Research in the NCP and from around the world has shown that irrigation strategies based on crop responses to water stress at different growth stages or under deficit irrigation can improve water-use efficiency (WUE) (Shang and Mao, 2006; Farahani et al., 2006; Pereira, 2006; Fang et al., 2007; Ali and Talukder, 2008; Farre and Faci, 2009; Bebera and Panda, 2009). Production losses will depend on the timing, duration and extent of the deficit, and may be amply offset when the real value of water is taken into account. Furthermore, high yields can still be obtained by supplying the required amount of irrigation water during sensitive crop growth stages, and by restricting water stress to tolerant growth stages (Blum, 2009; Geerts and Raes, 2009).

A number of studies have been conducted using pot and field experiments to assess the effects of water stress, and to quantify the exact amount of irrigation for winter wheat and the yield response to different water applications on the NCP. For example, Zhang et al. (2003) studied the effect of decreasing irrigation times, and showed that in a dry year, four irrigation times may not produce as high a yield as three irrigations. Again, Zhang and Yu (2003) showed that three irrigations were the best choice, and four irrigations did not improve wheat yield at all. Lu et al. (2000) demonstrated that one or two irrigation times could achieve a similar yield to four irrigation times. Using pot experiments, Zhang et al. (1999) determined the response factor to water stress at different stages. By applying a single irrigation at different stages, their results showed that water deficit at the jointing to flowering stage reduced wheat yield the most. Lan et al. (2001) and Zhu et al. (2003) studied the effect of water deficit on plant growth parameters at different growth stages. In experiments on irrigation during the different growing stages of winter wheat, Sun et al. (2006) drew the conclusion that full irrigation to field capacity (FC) did not produce any greater yield than treatments with a degree of water stress at certain stages.

The experimental results are site- and season-specific, and may not be applicable to other growing seasons and locations with different climate and soil conditions. These types of field experiments may also have their own limitations, namely that they are laborious and expensive, and require extensive field trials and a high number of treatments. For example, agricultural activities are strongly influenced by weather conditions, especially precipitation which varies from year to year. When rainfall provides the main water input for wheat growth in the experimental years, it is difficult to assess the amount of irrigation from the field experiments. Process-oriented crop simulation models can be of great help in solving these problems, and can be pre-evaluated through a well-proven model to refine the field tests and lower their overall costs and risk uncertainties (Whisler et al., 1986).

Crop simulation models have been used for decades to analyze crop responses to environmental stresses, and to test alternate management practices (Boote et al., 1996; Sinclair and Seligman, 1996) and promising deficit irrigation strategies (Lobell and Ortiz-Monasterio, 2006; Benli et al., 2007; Heng et al., 2007; Lorite et al., 2007; Pereira et al., 2009).

Some widely known models are CERES (Jones and Kiniry, 1986) which has been inserted in the DSSAT (Jones et al., 2003) and APSIM (Keating et al., 2003), EPIC (Williams et al., 1989), ALMANAC (Kiniry et al., 1992), CropSyst (Stöckle et al., 2003), and the Wageningen models (van Ittersum et al., 2003), which are based on the crop’s physiological response to environmental factors. Most of these models require a very high number of parameters to run, and many of them are not easily available in the field and need to be determined experimentally. For this reason, detailed models may be less reliable and require more advanced skills than simpler but more robust models (Sinclair and Seligman, 1996).

To overcome these complications, in 2009 the Food and Agriculture Organization (FAO) of the United Nations developed a new water-driven crop simulation model (Steduto et al., 2009; Raes et al., 2009), named “AquaCrop”, from the basic yield response to water algorithm in Doorenbos and Kassam (1979). It has limited complexity and aims to offer a balance of accuracy, simplicity, and robustness. One of the most important parameters in AquaCrop is the normalized biomass water productivity (WPI), which is typically constant for a given crop species (Steduto et al., 2007, 2009). It uses a relatively small number of explicit and mostly intuitive parameters and input variables, requiring simple methods for their derivation (FAO, 2009). For more details on the conceptual framework of AquaCrop, see Steduto et al. (2009); and for algorithmic and software solutions, see Raes et al. (2009).

Since its release, the performance of the AquaCrop model has been evaluated for winter wheat grain yield and biomass 2009–2010 (Du et al., 2011) on the North China Plain under different irrigation systems. The present study was designed to calibrate and extensively validate the AquaCrop model for soil water content, actual evapotranspiration plus grain yield and biomass (1998–2001) under deficit irrigation conditions; and to detect past measured and statistical trends in grain yields (16–21 years) in the Luancheng comprehensive experimental station on the North China Plain.

2. Methodology

Experiments were conducted at the Luancheng agro-ecosystem station (37°53’ N, 114°41’ E, 50.1 m a.s.l.), one of 34 agricultural ecosystem stations in the Chinese Ecological Research Network. The station is located in Luancheng county on the NCP (Fig. 1), and has a fertile topsoil, plenty of organic matter in loamy soil, and a high grain yield (Zhang et al., 2002). It has a temperate semi-arid monsoon climate, with a mean annual temperature of 12.2 °C, mean annual global radiation of 524 kJ/cm², and mean annual precipitation of 481 mm, most of which occurs from late June to September. Winter wheat is the main crop in this region. The growing season for winter wheat is from early October to mid-June (Table 1). The rainfall does not meet the needs of wheat for normal growth, especially during the dry, windy spring season. Table 2 shows the distribution of precipitation in the years of the experiment. Fig. 2 gives the overall climate conditions for the wheat season from 1985 to 2010. Local soil characteristics and parameters are shown in Table 3, indicating fertile topsoil and abundant organic matter in the loamy soil.
Table 1
Growth stages of winter wheat at Luancheng Station.

<table>
<thead>
<tr>
<th>Crop</th>
<th>Sowing</th>
<th>Emergence</th>
<th>Winter dormancy</th>
<th>Spring green up</th>
<th>Stem-extension</th>
<th>Heading</th>
<th>Flowering</th>
<th>Grain-filling</th>
<th>Harvesting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter wheat</td>
<td>1 October</td>
<td>7 October</td>
<td>22 November</td>
<td>3 March 10</td>
<td>10 April</td>
<td>1 May</td>
<td>5 May</td>
<td>10 May</td>
<td>10 June</td>
</tr>
</tbody>
</table>

Table 2
Precipitation during winter wheat seasons from 1998 to 2001, at Luancheng Station (unit: mm).

<table>
<thead>
<tr>
<th>Year</th>
<th>OCT</th>
<th>NOV</th>
<th>DEC</th>
<th>JAN</th>
<th>FEB</th>
<th>MAR</th>
<th>APR</th>
<th>MAY</th>
<th>JUN</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998–1999</td>
<td>10.3</td>
<td>0</td>
<td>0.7</td>
<td>0</td>
<td>0</td>
<td>6.6</td>
<td>43.1</td>
<td>18.5</td>
<td>79.2</td>
<td></td>
</tr>
<tr>
<td>1999–2000</td>
<td>24.3</td>
<td>6.1</td>
<td>0.2</td>
<td>7.6</td>
<td>0</td>
<td>0.8</td>
<td>2.5</td>
<td>11.8</td>
<td>45.8</td>
<td>99.1</td>
</tr>
<tr>
<td>2000–2001</td>
<td>73.9</td>
<td>8.4</td>
<td>0</td>
<td>20.1</td>
<td>6.9</td>
<td>0.5</td>
<td>23.3</td>
<td>6.2</td>
<td>47.5</td>
<td>186.8</td>
</tr>
</tbody>
</table>

Table 3
Soil properties at experimental site at Luancheng Comprehensive Experimental Station.

<table>
<thead>
<tr>
<th>Soil depth</th>
<th>Texture</th>
<th>BD (g/cm³)</th>
<th>SAT (mm/mm)</th>
<th>DUL (mm/mm)</th>
<th>LL (mm/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>Sandy loam</td>
<td>1.41</td>
<td>0.44</td>
<td>0.36</td>
<td>0.10</td>
</tr>
<tr>
<td>20–40</td>
<td>Sandy loam</td>
<td>1.51</td>
<td>0.46</td>
<td>0.35</td>
<td>0.11</td>
</tr>
<tr>
<td>40–60</td>
<td>Light loam</td>
<td>1.47</td>
<td>0.43</td>
<td>0.33</td>
<td>0.14</td>
</tr>
<tr>
<td>60–80</td>
<td>Medium loam</td>
<td>1.51</td>
<td>0.43</td>
<td>0.34</td>
<td>0.14</td>
</tr>
<tr>
<td>80–100</td>
<td>Light clay</td>
<td>1.54</td>
<td>0.44</td>
<td>0.34</td>
<td>0.13</td>
</tr>
<tr>
<td>100–120</td>
<td>Light clay</td>
<td>1.64</td>
<td>0.44</td>
<td>0.39</td>
<td>0.14</td>
</tr>
<tr>
<td>120–160</td>
<td>Light clay</td>
<td>1.59</td>
<td>0.48</td>
<td>0.38</td>
<td>0.16</td>
</tr>
</tbody>
</table>

BD: bulky density; SAT: saturation; DUL: field capacity; LL: lower limit.

Experiments on winter wheat were conducted in three consecutive seasons from 1998 to 2001. Fifteen 5 m × 10 m plots were created and divided by concrete walls for five irrigation schedules. The walls are 24.5 cm thick and extend 1.5 m beneath the surface, in accordance with the specifications set by the Food and Agricultural Organization (FAO). Winter wheat, Gaoyou No. 503, was sown by hand at a rate of 150 kg ha⁻¹, 20 cm wide per row. Before sowing, each plot was irrigated with about 80 mm water

Table 4
Levels of soil water content under different treatments of winter wheat at Luancheng Station (1998–2001).

<table>
<thead>
<tr>
<th>Crop</th>
<th>Treatment</th>
<th>Winter dormancy</th>
<th>Spring green up</th>
<th>Stem-extension</th>
<th>Heading</th>
<th>Grain-filling</th>
</tr>
</thead>
</table>

a θ is average soil water content of crop root depth; θₜₑ is average field capacity of crop root depth. 1.0 or 0.8 means the ratio of θ to θₜₑ.
containing 300 kg ha$^{-1}$ ammonium phosphate and 150 kg ha$^{-1}$ urea. Five kinds of deficit irrigation schedules were tested for winter wheat based on growth stages, and each treatment was replicated three times (Table 4). Irrigation amounts in the experimental treatments were in the order of D > B > A > C > E. In our experiment, treatment E is not irrigated after spring, and is thus regarded as a rainfed treatment; all treatments were irrigated before winter dormancy to ensure the winter wheat could survive the cold winter. Irrigation amounts and timing during the experiments are shown in Table 5. The area of the grain yield (GY) measurement was a 3 m × 8 m portion in the heart of each plot, and a 1000-kernel weight was determined from the harvested grains. For dry matter, 10 plants were measured after drying for 48 h in an oven at 65 °C. Soil water content was measured using a neutron probe (Institute of Hydrology, UK) down to 160 cm depth at 20 cm intervals, approximately every 5 days. Actual evapotranspiration (ET$_{a}$) was calculated by the water balance method. Details of the field layout and experimental structure are given in Zhang et al. (2004), but a brief description is presented here to contribute to the understanding of this paper.

### 2.1. Input data and calibration of the AquaCrop model

AquaCrop version 3.1 Plus was used in this study and run under GDD (Growing Degree Days) for a more realistic consideration of the behavior of winter crops below base temperature or under cold stress during the winter wheat dormancy period, and to ensure a more realistic simulation (Steduto et al., 2009).

AquaCrop uses six input files for simulation: climate file, crop file (time to emergence, maximum canopy cover, start of senescence, maturity), soil file, management file, irrigation file, and initial soil water conditions; all these are user specific. The climate file consists of three sub-files: (i) minimum and maximum air temperature, (ii) ET$_{a}$, and (iii) rainfall, all with daily values as described by Raes et al. (2009). The crop file contains both conservative parameters (that do not change with location) and user-specific parameters (non-conservative). Gauch et al. (1999) pointed out that minor changes in initial SWC (e.g., from 8.5 to 10.5 vol%) resulted in major changes in the model output (400 kg ha$^{-1}$ additional biomass and 250 kg ha$^{-1}$ additional yield). The parameters used for the AquaCrop model were measured or estimated using experimental data; some were based on field experience, and some used the default values given in the model, regardless of the year (Table 6).

The AquaCrop model was calibrated using the values observed from the field experiment during 1999–2000, with deficit irrigation treatment D, for grain yield, biomass, ET$_{a}$ and soil water content (0–120 cm). Measured soil water content was only available for this year, and was the reason this year was selected.

Treatment D was used for the model calibration, and the model’s default values were used for stress treatments. The difference between the simulated model and the experimental data was minimized by using a trial and error approach.

### 2.2. Validation of the AquaCrop model

Validation is an important step in model verification (Addiscot et al., 1995; Power, 1993), and is done by comparing independent field measurements (data) with the outputs created by the model. Grain yield, biomass, ET$_{a}$ and soil water content were considered as the evaluation parameters for the AquaCrop model.

The crop parameters obtained from the model calibration were used in the validation. The independent data set was used to test the model’s performance from 1998 to 2001, with 14 treatments, excluding the data used for calibration.

### 2.3. Model evaluation

Since no single measure can indicate how well a simulation model performs, a combination of statistical indices are generally used to evaluate the model (Caton et al., 1999; Kobayashi and Salam, 2000; Gauch et al., 2003).

The agreement between the measured and simulated values was assessed using the following six statistics: RMSE (root mean square error), NRMSE (normalized root mean square error), MBE (mean bias error), MAE (mean absolute error), IoA (index of agreement) (Willmott, 1982). Percentage difference was determined using the following equation, where $S_{i}$ indicates simulated values and $M_{i}$ indicates measured values in all statistical indices.

\[
NRMSE = \sqrt{\frac{\sum (S_{i} - M_{i})^2}{n}} \times \frac{100}{M} \tag{1}
\]

\[
RMSE = \frac{1}{n} \sum (S_{i} - M_{i})^2 \tag{2}
\]

\[
MAE = \frac{1}{n} \sum |S_{i} - M_{i}| \tag{3}
\]

\[
MBE = \frac{1}{n} \sum (S_{i} - M_{i}) \tag{4}
\]

\[
d = 1 - \left[ \frac{\sum (S_{i} - M_{i})^2}{\sum (|S_{i} - M| + |M_{i} - M|)^2} \right] \tag{5}
\]

\[
Percent\ deviation = \frac{(simulated - measured)}{measured} \times 100 \tag{6}
\]

### Table 5


<table>
<thead>
<tr>
<th>Years</th>
<th>Treatment A</th>
<th>Treatment B</th>
<th>Treatment C</th>
<th>Treatment D</th>
<th>Treatment E – rainfed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998–1999</td>
<td>252.0</td>
<td>239.5</td>
<td>202.3</td>
<td>362.1</td>
<td>50.0</td>
</tr>
<tr>
<td>1999–2000</td>
<td>281.5</td>
<td>323.9</td>
<td>247.9</td>
<td>404.8</td>
<td>90.0</td>
</tr>
<tr>
<td>2000–2001</td>
<td>234.0</td>
<td>230.0</td>
<td>216.7</td>
<td>364.7</td>
<td>80.0</td>
</tr>
</tbody>
</table>
The simulation is considered excellent with a normalized RMSE of less than 10%, good if the normalized RMSE is greater than 10 and less than 20%, fair if the normalized RMSE is greater than 20% and less than 30%, and poor if the normalized RMSE is greater than 30% (Jamieson et al., 1991).

IoA: A value of 1.0 indicates excellent agreement between measured and simulated values.
RMSE: Close to zero indicates better model performance.
MBE: This reveals the long-term performance of the model. A positive value of MBE gives the average amount of overestimation in the estimated values and vice versa.
MAE: This measures the weighted average magnitude of the absolute errors. All these statistics were computed using the IRENE statistical software (Fila et al., 2003).

3. Results and discussion

3.1. AquaCrop model calibration

Table 7 and Fig. 3 show the simulation results for winter wheat on the North China Plain using the calibration data set (treatment D) for grain yield, biomass, ETa and soil water content up to 0–120 cm depth. The minimum difference was simulated for biomass (0.1%) while the highest difference was observed in grain yield (9.6%). Araya et al. (2010a) simulated the highest deviation for biomass, at 8.5%. Differences in ETa and soil water content were 3.7% and –9.0% respectively. The calibration results show a reasonably close match between the measured values and those simulated by the model.

3.2. AquaCrop model validation

Fourteen independent treatments were used in the validation step for AquaCrop, considering the parameters used in the calibration procedure. Table 8 shows the measured and simulated results for the validated data sets for grain yield, biomass and ETa.

3.2.1. Grain yield

The results in Table 8 show that no significant deviation (−0.6% to −7.0%) was observed between the measured and simulated values of AquaCrop for the year 1998–1999, as this year was relatively dry, and the cropping season depends mainly upon irrigation. In the second season (1999–2000), the greatest deviation was simulated for grain yield in the case of treatment E – rainfed (−43.2%, underestimated). This could possibly be due to the fact that the senescence of the canopy accelerates under severe water stress, and the underground root system may be restricted and prevented from extracting more deeply stored soil water, thereby limiting its water uptake. Several authors (Heng et al., 2009; Araya et al., 2010a,b;
Fig. 3. Comparison of modeled to observed soil water content (0–120 cm) for winter wheat cropping season 1999–2000.

Zeleke et al., 2011; Abedinpour et al., 2012) reported much greater deviations under severe water stress or rainfed treatments, as compared to well-watered treatments for maize, teff and canola crops simulated by AquaCrop. In the third year (2000–2001), overestimations (18.2%, 12.9% and 14.1%) were noted for treatments C – least irrigated, D – well-irrigated, and E – rainfed, respectively. This cropping season received a total precipitation of 186.8 mm, which is a higher amount than in the two previous

Table 8
Measured vs. simulated results for validated data sets of winter wheat from 1998 to 2001 in the North China Plain.

<table>
<thead>
<tr>
<th>Year</th>
<th>Treatment</th>
<th>Yield (Mg ha⁻¹)</th>
<th>Deviation (%)</th>
<th>Biomass (Mg ha⁻¹)</th>
<th>Deviation (%)</th>
<th>ET (mm)</th>
<th>Deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Measured</td>
<td>Simulated</td>
<td></td>
<td>Measured</td>
<td>Simulated</td>
<td>Measured</td>
<td></td>
</tr>
<tr>
<td>1998–1999</td>
<td>A</td>
<td>5.3</td>
<td>5.1</td>
<td>–3.2</td>
<td>10.0</td>
<td>10.6</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5.3</td>
<td>5.0</td>
<td>–6.0</td>
<td>10.3</td>
<td>10.7</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.9</td>
<td>5.5</td>
<td>–7.0</td>
<td>10.7</td>
<td>11.1</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>5.4</td>
<td>5.3</td>
<td>–2.4</td>
<td>11.2</td>
<td>10.6</td>
<td>–4.6</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>2.6</td>
<td>2.5</td>
<td>–1.4</td>
<td>5.2</td>
<td>5.0</td>
<td>–2.3</td>
</tr>
<tr>
<td>1999–2000</td>
<td>A</td>
<td>5.7</td>
<td>5.6</td>
<td>–0.6</td>
<td>11.7</td>
<td>11.3</td>
<td>–2.8</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5.5</td>
<td>5.8</td>
<td>5.2</td>
<td>11.6</td>
<td>11.6</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.5</td>
<td>5.5</td>
<td>–1.1</td>
<td>12.1</td>
<td>11.4</td>
<td>–6.0</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>3.6</td>
<td>2.0</td>
<td>–43.2</td>
<td>8.1</td>
<td>6.1</td>
<td>–23.6</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>3.2</td>
<td>3.1</td>
<td>–3.1</td>
<td>10.2</td>
<td>10.3</td>
<td>0.7</td>
</tr>
<tr>
<td>2000–2001</td>
<td>A</td>
<td>5.2</td>
<td>5.1</td>
<td>–3.1</td>
<td>10.2</td>
<td>10.3</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5.1</td>
<td>5.3</td>
<td>3.9</td>
<td>10.5</td>
<td>10.7</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>5.0</td>
<td>5.9</td>
<td>18.2</td>
<td>10.2</td>
<td>11.9</td>
<td>16.5</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>5.5</td>
<td>6.2</td>
<td>12.9</td>
<td>12.0</td>
<td>12.4</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>3.3</td>
<td>3.8</td>
<td>14.1</td>
<td>8.5</td>
<td>7.0</td>
<td>–17.3</td>
</tr>
</tbody>
</table>
seasons. Of this amount, 73.9 mm fell in October, and contributed to increasing initial soil water content – crucially important for better seedling emergence and a deeper root system (Li et al., 2001b), while 104.5 mm fell between January and June. In other words, it can be said to be almost uniformly distributed. This was therefore the main reason for the overestimation of the validation of AquaCrop this year.

The 2000–2001 cropping season was relatively wetter than previous years. The model overestimated the grain yield in the order of C > E > D due to irrigation plus precipitation. This was also confirmed by field experiments carried out by several authors (Zhang et al., 2004; Sun et al., 2006). The findings showed that treatment D is not effective on the NCP for a high yield and WUE for winter wheat, compared to treatment C – least irrigated.

The deviation range is considerably better for grain yield (−0.6–14.1%) in the validation of the AquaCrop model from 1998 to 2001, whereas Araya et al. (2010a) reported that the deviation range in the validation data was −13–15.1% in the case of grain yield for barley. Table 9 shows the statistical assessment of the AquaCrop model for three experimental years under deficit irrigation. The results of RMSE (0.58 Mg ha\(^{-1}\)), MAE (0.38 Mg ha\(^{-1}\)), MBE (−0.01 Mg ha\(^{-1}\)), \(d\) (0.92) and NRMSE (11.9%) are comparable with those obtained by Mkabela and Paul (2012) for winter wheat grown in Western Canada using the same AquaCrop model for grain yield simulation. Fig. 4 shows the relationship between the observed and simulated winter wheat grain yield. The straight-line equation and the coefficient of determination \(R^2\) show that the model simulated grain yield with a high degree of reliability has a \(R^2\) of 0.78. In our case we obtained a slightly lower index of agreement (0.92) compared to the abovementioned study. This could be due to differences in the cultivar, and also to the fact that our study model was assessed under deficit irrigation, while the study in question was done in conditions of no water stress.

The results of the AquaCrop model can be compared with other crop models used on the North China Plain for winter wheat yield prediction and water management strategies. For example, Lu and Fan (2012) used the EPIC model for yield gap analysis in winter wheat, with a RMSE ranging between 0.84 and 3.60 Mg ha\(^{-1}\). Mo et al. (2009) explained 18% of the relative error between simulated and statistical grain yield in the VIP (Vegetation Interface Processes) model to analyze wheat production, water consumption and water-use efficiency. Mo et al. (2009) simulated grain yield, water consumption and WUE by the SVAT model using remote sensing data and GIS in the study region, and validation explained the relative errors of 23%, with a RMSE of 1.12 Mg ha\(^{-1}\). Ma et al. (2013) used the WOFOST model with remote sensing data to reduce the uncertainties in regional yield estimation, and by assimilating HJ-1 CCD NDVI time series into the WOFOST–ACRM model with the Ensemble Kalman Filter, suggested that the error could be reduced by up to 0.77–3.0 Mg ha\(^{-1}\) with statistical data. Yu et al. (2006) calibrated and validated the Root Zone Water Quality Model (RZWQM) with both a generic plant growth module (RZWQM-G) and the CERES plant growth module (RZWQM-C) to simulate winter wheat and maize double cropping systems on the North China Plain, and showed that the RZWQM-C model simulated grain yields with a RMSE of 0.94 Mg ha\(^{-1}\), compared to a RMSE of 1.23 Mg ha\(^{-1}\) with RZWQM-G. Du et al. (2011) used the AquaCrop model to simulate grain yield and biomass under different irrigation management systems. Their findings indicate that the AquaCrop model is able to simulate grain yield and biomass well with maximum mean errors of 12.35% and 14.1%, respectively. Our results for both parameters are also similar to these, except in rainy conditions.

The results of this study therefore suggest that the AquaCrop model can be used with a considerable degree of accuracy for deficit irrigation scheduling, considering crop growth stages or timing of irrigation for modeling winter wheat grain yield on the North China Plain, except under very severe soil water deficit conditions.

### 3.2.2. Final aboveground biomass

Table 8 shows a validation of the AquaCrop model for final aboveground biomass. The highest negative deviations (−23.6, −17.3%) were simulated for the year 1999–2000, and for 2000–2001 in rainfed treatments, as in the case of grain yield. This is due to severe water stress experienced in the cropping season, as already explained for grain yield. Another significant positive deviation (16.5%) was noted for treatment C in the 2000–2001 season, where irrigation was omitted in the grain filling period, when this treatment received the lowest amount of irrigation. The additional amount of water obtained from precipitation could lead to an overestimation of the biomass. A similar trend was observed in the validation of the AquaCrop model for grain yield.

The AquaCrop model simulated aboveground biomass more accurately than grain yield, with deviations ranging from 0.4% to 5.8%. Araya et al. (2010a,b) reported a deviation for aboveground biomass of −4.3–14.6% for barley, and −0.10–8.70% for teff. For simulated biomass, Heng et al. (2009) reported a RMSE of between 0.46 and 6.51 Mg ha\(^{-1}\) for maize using the AquaCrop model, but with data from different locations. Greets et al. (2009) calculated that the model evaluation criteria between measured and simulated biomass were RMSE = 16%, \(R^2 = 0.87\). Hsiao et al. (2009) presented a deviation for maize biomass simulation in AquaCrop of between −0.4% and 21.9%. Vila et al. (2009) used the AquaCrop model to optimize the deficit irrigation of cotton, and simulated biomass
for calibration and validation data sets with a RMSE of 1.32 and 1.01 Mg ha\(^{-1}\), respectively. Zeleke et al. (2011) reported percentages of RMSE and deviations by AquaCrop in the calibration and validation data of canola of 2.10, 2.58 t ha\(^{-1}\) and 1.2%, −9.7%, respectively.

Fig. 4 shows the linear correlation between simulated and observed biomass throughout the 1998–2001 growing seasons. The linear trendline deviates slightly from the desirable straight line \(x = y\), and thus has a high coefficient of determination \(R^2\), at 0.85. Table 9 shows the overall statistical parameters for the model validation. Values of RMSE (0.87 Mg ha\(^{-1}\), MAE (0.69 Mg ha\(^{-1}\), MBE (−0.08 Mg ha\(^{-1}\), d (0.95) and NRMSE (8.62%) are comparable with those obtained by other crop models. Lu and Fan (2012) used the EPIC model to study yield gap analysis for winter wheat on the North China Plain, with a RMSE for biomass of between 1.18 and 2.0 Mg ha\(^{-1}\). Yu et al. (2006) calibrated and validated the Root Zone Water Quality Model (RZWQM) with both a generic plant growth module (RZWQM-G) and the CERES plant growth module (RZWQM-C) to simulate winter wheat and maize double cropping systems on the North China Plain, and reported a RMSE of 2.07 Mg ha\(^{-1}\) with the RZWQM-G and 2.26 t ha\(^{-1}\) with the RZWQM-C model.

The results of the present study show that the AquaCrop model simulates aboveground biomass more effectively than grain yield, which agrees with the results obtained by Araya et al. (2010a).

### 3.2.3. Actual evapotranspiration (\(ET_a\))

Table 8 shows the validation of the AquaCrop model for actual evapotranspiration from 1998 to 2001 in the experimental years. AquaCrop simulated the largest negative deviation (−21.7%) for well-watered treatment D in the year 1998–1999 – a relatively dry year when the crop water requirement depended mainly on irrigation. Sun et al. (2006) indicated – based on field experiments – that treatment D had the highest amount of drainage. lowest water-use efficiency, and produced a lower yield compared to other deficit irrigation treatments. This could therefore be the cause of the underestimation of \(ET_a\) in this treatment. Another underestimated deviation (−10.3%) in the same year was observed for treatment B, which is the second highest irrigated treatment, and also prone to drainage; in contrast, grain yield and biomass were acceptably simulated by AquaCrop for both treatments. For 1999–2000 and 2000–2001 – both wet years – rainfall treatment E showed a positive deviation (12.5%, 8.4%) by the AquaCrop model. Sun et al. (2006) reported that the lowest amount of drainage was noted in treatment E. This means the lowest drainage plus seasonal precipitation contributed to the overestimation of \(ET_a\), but failed to produce a satisfactory grain yield and biomass due to severe soil water deficit or poor root development. It has already been explained in the literature, and is supported by the findings for grain yield and biomass in this study, that the AquaCrop model cannot provide satisfactory results under severe water stress conditions. Other deviations between the measured and modeled \(ET_a\) were smaller (−0.4–6.9%). Table 9 contains all the model evaluation criteria from 1998 to 2001, with RMSE (33.2 mm), MAE (24.0 mm), MBE (−6.95 mm), d (0.93) and NRMSE (8.99%). Normalized RMSE for \(ET_a\) is in the excellent range.

Farahani et al. (2009) evaluated the AquaCrop model for cotton, and calculated a deviation for \(ET_a\) of between 2.1% and −10.2%. Heng et al. (2009) presented total measured and simulated ET. Deviations were calculated in the range of −1.23% to −8.4%. Fang et al. (2010) simulated \(ET_a\) using RZWQM2 with a RMSE of 41.5 mm for a wheat-maize cropping system on the North China Plain.

The model acceptably simulates the value of \(ET_a\) under deficit irrigation for winter wheat in the arid-semiarid conditions of North China. These results thus confirm that the AquaCrop model can be used to study water balance, further estimate water consumption, and for planning winter wheat on the North China Plain.

### 3.2.4. Soil water content

The AquaCrop model was validated for the soil water content of the total profile (0–120 cm), and the results are shown in Table 9. Statistical parameters such as RMSE ranged from 24.5 to 37.6 mm, NRMSE from 7.74 to 13.9%, \(d\) varied from 0.86 to 0.89, MAE was in the range of 20.9–33.7 mm, and MBE from −5.97 to −17.1 mm under different irrigation treatments. Data were available for a single winter wheat season (1999–2000), as shown (Fig. 3). Thus treatment D was used in the model calibration, and the rest of the four treatments were used for validation. The highest deviation was simulated by the AquaCrop model, as compared to the measured soil water content (−9.0%) for well-watered treatment D, while other deviations were smaller (−1.88% to −6.70%). In general, the total soil water content simulated by the AquaCrop model closely follows the trend of the values observed, although there are some slight mismatches with the measured data. The modeling of the soil water content revealed a common trend in this study: i.e. a slight underestimation – but initial soil water content was well simulated, except in the case of the rainfed treatment. This trend is more pronounced in the order of D > A > C > B treatments. A closer analysis of Fig. 3 shows that the model does not allow the soil water content to evolve towards field capacity after irrigation, but instead decreases rapidly towards the permanent wilting point. This behavior of the model leads to an underestimation of the soil water content.

In the case of treatment E – rainfed (Fig. 3), the AquaCrop model significantly underestimated initial soil water content, while after the mid season there were some good matches between measured and modeled soil water content. This trend coincides with grain yield, biomass and \(ET_a\) in this study, indicating that mismatches
largely occur in the model’s performance under conditions of severe soil water deficit.

In line with this, Mkhabela and Paul (2012) highlighted the underestimation of modeled soil water content but in reverse of the present study: i.e. the model does not allow soil water depletion below the permanent wilting point. However no irrigation water was applied in this study. Farahani et al. (2009) and Hussein et al. (2011) reported that the AquaCrop satisfactorily simulated the soil water content of the whole profile; however it consistently tended to overestimate total soil water content, particularly in the deficit irrigation plots. Hsiao et al. (2009) also found that the AquaCrop model overestimated the soil water content by a significant amount – about 80 mm – although the general trend of the measured data set was captured for deficit irrigation. Araya et al. (2010b) verified the underestimation of the soil water content simulated by the AquaCrop model under severe water stress conditions, although overall there was a perfect match between the simulated and measured data sets. Zeleke et al. (2011) also confirmed that AquaCrop accepted simultaneously both the amount and trend of soil water content, but tended to overestimate in the cropping season. The results of the measured and simulated soil water content obtained by Mkhabela and Paul (2012) for winter wheat were as follows: RMSE (49.4 mm), MAE (40.5 mm), MBE (10.5 mm), and d (0.99); while Andarzian et al. (2011) reported that the AquaCrop model can simulate soil profile water content accurately with a RMSE of 18 mm, NRMSE of 3.5%, and d of 0.84.

The performance of the AquaCrop model in simulating soil water content for irrigation strategies for winter wheat in North China can be compared with other models. For example, Yang et al. (2006) used the DSSAT–CERES-Wheat model for better irrigation management to overcome groundwater depletion for winter wheat in the study region. Their simulated soil water content showed deviations of 0.43% to 20.4% from the measured data. Fang et al. (2010) simulated total soil water content using RZWQM2 from 0 to 120 cm in a wheat-maize cropping system, with a RMSE of 28.5–34.8 mm in this area.

Despite the slight mismatching, the overall results of this study suggest that the AquaCrop model can be used with a reliable degree of accuracy to simulate soil water content in winter wheat on the North China Plain under deficit irrigation, except in rainfed conditions.

4. Revalidation of the AquaCrop model

The revalidation of the model depends upon its successful calibration based on field experimental data, and the accurate estimation of the specific model’s coefficients in a given environment. Eitzinger et al. (2003) revalidated the CERES-Wheat model using the grain yield observed over 9 years (1985–1993) at the experimental site, and compared it with model outputs. This study did not include deficit irrigation experimentation.

In the current study for the AquaCrop revalidation, we used 21 years (1990–2010) of grain yield data measured at the experimental site using daily weather data for the same period obtained from a nearby meteorological station.

4.1. Modifications in the AquaCrop model for revalidation based on per-annual experimental data

Our 1998–2001 experimental data was based on deficit irrigation and small experimental plots divided by concrete walls according to the FAO standard. Five kinds of deficit irrigation scheduling were tested, and the model was calibrated and validated against this data. Neutron probe access tubes were installed in the center of each plot to measure soil water content approximately every 5 days, and after irrigation or a precipitation event. This may cause a degree of soil compaction or soil disturbance. The seed rate used during experimentation was 150 kg ha⁻¹, as described in Section 3.1.

Sun et al. (2007) conducted field experiments (2002–2005) in the study area to examine the response of winter wheat yields to sowing and harvest times. They concluded that winter wheat development occurred during emergence to heading, and that yield components such as kernel weight and kernels per spike were mainly determined by thermal time: there is a 0.5% reduction in yield per day⁻¹ if the sowing date is delayed after 10 October, while the maturity time is almost the same.

To address this uncertainty in the sowing date, and to ensure accurate detection of past yield trends by the AquaCrop model, sowing time was considered from 1 October (based on field experiments) until 10 October.

However, the farmers’ actual ongoing practices in the field differ compared to these small experimental plots. Firstly, the seed rate is 180 kg ha⁻¹, no soil water deficit is applied, and watering is by means of flood irrigation. These two conditions resulted in a high harvest index, crop canopy and increased root length to extract water from deeper soil layers, as compared to the experimental plots. Therefore in the revalidation of the AquaCrop model, the following adjustments were made in the non-conservative section, while other parameters were maintained as shown (Table 6). These modifications are based on actual field conditions.

1. Seed rate = 180 kg ha⁻¹
2. Maximum canopy cover in soil cover fraction = 94%
3. Harvest index = 0.49
4. Maximum rooting depth = 1.4 m
5. Under basin irrigation, 60% depletion of RAW, return to field capacity to avoid any water stress.

According to our revalidation of AquaCrop, statistical analysis shows that the value of RMSE obtained was 0.61 Mg ha⁻¹, NRMSE (9.92%), d (0.84), MAE (0.34 Mg ha⁻¹), and MBE (0.08 Mg ha⁻¹), while the mean revalidation difference over 21 years between observed and simulated grain yield with the AquaCrop model was 3.3%. Fig. 5 shows clearly that the model very accurately predicted climate variability in order to detect the yield trend of the past 21 years.

![Fig. 5. Relationship between simulated and observed winter wheat grain yield (Mg ha⁻¹) at the experimental station from 1990 to 2010.](image-url)
Eitzinger et al. (2003) reported that the mean revalidation in the CERES-Wheat model over 9 years for the difference between observed and simulated grain yield was 5.1%. Yang et al. (2006) used the DSSAT Wheat model from 1987 to 2001 with measured data for the estimation of groundwater use by winter wheat on the North China Plain, and the simulated statistical errors between measured and predicted winter wheat yields were RMSE (0.85 Mg ha\(^{-1}\)), NRMSE (15.2%) and \(d (0.70)\).

4.2. Second revalidation of the AquaCrop model based on the statistical grain yield of Shijiazhuang

This revalidation of AquaCrop is based on the statistical grain yield (1995–2010, 16 years) obtained from the Agricultural Department of Shijiazhuang region (Fig. 6). The location of this area is shown in Fig. 1. In this revalidation, there is no further change in the parameters of the AquaCrop model. The same parameters were used as for the previous revalidation in this study. The model evaluation criteria show that RMSE, NRMSE, \(d\), MAE and MBE were 0.41 Mg ha\(^{-1}\), 6.0%, 0.60, 0.29 Mg ha\(^{-1}\) and \(-0.28\), respectively. Overall mean deviation from the statistical yield and modeled yield by the AquaCrop was \(-4\%\).

The results of this study very clearly indicate that the AquaCrop model can be used with a high degree of accuracy for regional yield simulation and yield gap analysis of winter wheat on the North China Plain, and could even also serve as a useful tool for assessing national food security in the agricultural sector.

5. Conclusions

The calibration of the AquaCrop model version 3.1 plus and its validation were tested in the 1998–2001 winter wheat seasons under deficit irrigation in the arid and semiarid conditions of the North China Plain. Moderately satisfactory agreements were obtained for grain yield, biomass, \(E_T\) and soil water content in the validation process. Major deviations were observed under severe stress conditions. The model prediction slightly underestimated soil water content. This could be due to an over-extensive simplification in order to make the model less data intensive (Steduto et al., 2009; Raes et al., 2009). When the model was revalidated against long-term measured yield data and statistical yield for this area, the results of the modeled grain yield obtained were well matched with minimum error ranges (3.3\% to \(-4.0\%\), respectively).

We can conclude from this study that the AquaCrop model can be used with a reliable degree of accuracy under mild water stress to determine the least sensitive and stress-tolerant stages of the growing period. This makes it very useful for the design and evaluation of deficit irrigation strategies, preventing unnecessary losses from runoff, drainage and soil evaporation, in addition to enhancing water-use efficiency. The particular feature that distinguishes AquaCrop from other crop models is its focus on water, especially under water-limiting conditions. One important application of AquaCrop would be for use in yield gap analysis, and to identify the constraints limiting crop production and water productivity. Its application could also be evaluated under climate-change scenarios for regional or national food and water security. Our revalidations, and the detection of yield trends and statistical evaluation criteria for the past 16–21 years, confirm the validity of the AquaCrop model.

Acknowledgments

This research was supported by a Young International Scientists Fellowship to Muhammad Anjum Iqbal from the Chinese Academy of Sciences (CAS), China and the Knowledge Innovation Project (KICSP, EW-J-5). This paper has been revised by Ms Pru Brooke-Turner.

Authors would like to thank the editor and anonymous reviewers for their helpful comments.

References


