Considering different water supplies can improve the accuracy of the WOFOST crop model and remote sensing assimilation in predicting wheat yield**

Xin Xu1,2, Shuaijie Shen2, Feng Gao2, Jian Wang1, Xinming Ma1,2*, Shuping Xiong2, and Zehua Fan2

1College of Information and Management Science, Henan Agricultural University, Zhengzhou 450002, China
2College of Agronomy, Henan Agricultural University, Zhengzhou 450002, China

Received January 5, 2022; accepted September 23, 2022

Abstract. The study was carried out in order to clarify the effects of different water and irrigation conditions on crop models and remote sensing assimilation results. It involved taking winter wheat from 17 test sites in Henan Province as the research object and calibrating the World Food Studies model. The ensemble Kalman filter algorithm was used to calibrate the two modes and Moderate-resolution Imaging Spectroradiometer-Leaf Area Index of the calibrated world food studies model. The study found that the average error of the world food studies model for simulating flowering and maturity periods is within 2 days, the R² of the leaf area index calibration results is between 0.87-0.98, and the R² and root mean square error of the verification results are 0.77 and 1.06 respectively. Under the latent model, the R² of the world food studies model taking account of the water supply situation and the assimilation results without taking account of the water supply situation are 0.50 and 0.48, respectively. In the water restriction mode, the R² increased from 0.79 to 0.86 compared with the assimilation results where the water supply was not considered. The results show that: depending on the water supply of different regions, selecting the corresponding assimilation parameters can effectively improve the prediction accuracy of crop models and remote sensing assimilation for wheat yields under different water and irrigation conditions.

Keywords: wheat, crop model, remote sensing, data assimilation, yield forecast, water restriction

INTRODUCTION

The dynamic monitoring and yield prediction of crop growth on a regional scale are of great significance for ensuring food security and formulating agricultural policies (Lipper et al., 2014). As a mechanism model, the crop growth model can simulate the physiological processes of crop photosynthesis, respiration, dry matter distribution, etc., in a specific environment. It has been widely used in crop yield prediction (Morell et al., 2016), field management decision-making (Zhang et al., 2018), agricultural production potential evaluation (Tang et al., 2018), climate change impact assessment (Vanli et al., 2019) and other fields. However, due to the heterogeneity of the surface environment, it is difficult to obtain some parameters when the model is applied to a particular region, which limits the application of the model on the regional scale (Pan et al., 2019). With the development of remote sensing technology, monitoring vegetation growth has become increasingly viable (Houborg et al., 2015). Remote sensing information can be used to monitor crop growth and also to determine physiological and biochemical indicators, remote sensing may be used to obtain instantaneous crop growth conditions, but
it cannot entirely explain the interaction between crops and the environment, and there are often discontinuities in time (Jin et al., 2016).

The assimilation of remote sensing data and crop growth models can be used to monitor crop growth and development in time and space on a continuous basis, and to improve the prediction accuracy and scope of application of the model (Dorigo et al., 2007; Huang et al., 2018). Data assimilation originated in the field of meteorology and is used in the practice of weather forecasting, oceanography, and hydrology (Charney et al., 1969). In 1979, the assimilation method was first used to couple remote sensing information and crop models to predict crop yields (Wiegand et al., 1979). In recent years, more and more studies have proved that using data assimilation methods to fuse remote sensing information with crop growth models is an important way to improve regional crop yield prediction (De Wit and van Diepen, 2007; Jiang et al., 2017). The purpose of data assimilation is to combine all possible information (models, observations, prior data, statistical data, etc.) in order to obtain the best estimate of the state of the agricultural system. Currently, the methods of remote sensing and crop model assimilation are for the most part parameter optimization and update methods (Wu et al., 2021).

The parameter optimization method uses all of the remote sensing data in a time window to re-adjust the initial parameters of the crop model through an optimization algorithm in order to minimize the difference between the model's simulated value and the remote sensing observation value, this achieves the purpose of optimizing the crop model. Commonly used optimization algorithms are the maximum likelihood method (Dente et al., 2008), compound hybrid evolutionary algorithm (Shen et al., 2009), simulated annealing algorithm (Jin et al., 2016), particle filter algorithm (Li et al., 2015), variational algorithm (Huang et al., 2019), etc. However, the accuracy of the parameter optimization method is affected by the assimilation variables, optimization algorithms, cost functions, the number of effective remote sensing observations, and the inversion accuracy of remote sensing on the assimilation variables (Wu et al., 2021), also, parameter optimization requires a considerable time span for iterative operations (Bai et al., 2019). The update method uses remote sensing observations to continuously correct the state variable trajectory of the model in order to bring the model closer to the real situation. It is an assimilation method that is continuous with time and may be applied to real-time prediction (Huang et al., 2016). The ensemble Kalman filtering algorithm (EnKF) has become a potent method among the various update methods due to the favourable processing ability of nonlinear observation operators (Li et al., 2011; Ma et al., 2013).

At present, the focus of ongoing research is concerned with remote sensing and crop model assimilation (Weiss et al., 2020; Jin et al., 2018). The uncertainty of model simulation (Tang et al., 2018) combined with uncertainty concerning the inversion of crop traits by remote sensing (Houborg et al., 2015), different data assimilation algorithms (Bai et al., 2019), and the choice of assimilation parameters (Ma et al., 2013) will all affect the assimilation results. Researchers have already used crop models under latent variable conditions to reduce model uncertainty (Huang et al., 2019; Huang et al., 2012), and they also use more efficient and accurate assimilation algorithms to reduce errors in the assimilation process (Wu et al., 2021; Xing et al., 2017), as well as using multi-source remote sensing data to reduce the uncertainty of remote sensing inversion (Pan et al., 2019) to improve the accuracy of assimilation. However, the effects of different water and irrigation conditions on crop models and remote sensing assimilation results are not clear, this research is based on the ensemble Kalman filter algorithm, and uses the potential mode and water limitation mode of the WOFOST crop model assimilation with MODIS remote sensing data to predict the wheat yield under different climate and irrigation conditions, with a view to forecasting the regional wheat yield under different production and management conditions to provide a theoretical basis for crop management.

MATERIALS AND METHODS

Henan Province is a major wheat-producing province in China. According to data from the National Bureau of Statistics (http://www.stats.gov.cn/), Henan's wheat output accounts for about 26-28% of the country's total wheat output. Henan has a vast territory (31°23'–36°22'N, 110°21’–116°39’E), and belongs to the eastern monsoon region of China, straddling the two natural regions of the northern subtropical zone and the warm temperate zone. The four seasons are distinct throughout the year, the climate is mild, and the transition from the northern subtropical zone to the warm temperate zone is obvious. The annual average temperature of the province from north to south is 283.65-289.85 K, the average annual precipitation value is 407.7-1295.8 mm, and rainfall reaches its highest value from June to August. The frost-free period of the year lasts from 201-285 days, which is suitable for growing a variety of crops. The soil types of the province, its climatic characteristics, production conditions, farming systems, and yield levels are quite varied.

The experiment was conducted from 2019 to 2020 and Huaxian, Neihuang, Yuanyang, Jiaozuo, Xiwu, Xuchang, Shangqiu, Luhe, Zhoukou, Kaifeng, Luoning, Ruyang, Pingyu, Luoshan, and Xixia, Dengzhou, and Fangcheng counties were selected in Henan Province. Sampling was carried out at 17 test sites such as Dengzhou, and Fangcheng. The specific distribution of the test sites is shown in Fig. 1, and the specific sowing, harvesting, and irrigation conditions of each test site are shown in Table 1. The leaf area index (LAI) of the examined wheat was determined using the canopy analyser LAI-2200C (PCA;
LI-COR Inc., Lincoln, NE, USA) during the four overwintering, jointing, flowering, and filling periods of wheat. The final output is measured at the time.

The weather data comes from the NASA Power database (https://power.larc.nasa.gov/). This database has been widely used in agricultural modelling, crop yield simulation, etc. (Bai et al., 2010). The soil data came from the 1:1 000 000 Chinese soil data set (Shangguan et al., 2013). Soil Water Characteristics software (Saxton and Rawls, 2006) was used to convert the required soil parameters such as the maximum field water-holding capacity, permanent wilting point, and field saturated water content. MODIS remote sensing data are obtained from the Earth Observation System portal website (http://eospso.gsfc.nasa.gov/). In this study, MODIS-LAI data products MCD15A3/MCD15A3H were used, and MODIS-LAI was corrected using S-G filtering (Huang et al., 2015; Xu et al., 2011) and measured values.

When calibrating the model, it is necessary to calibrate the growth period. In the WOFOST model, the development of crop phenology is mainly controlled by the accumulated temperature. For wheat, it is also affected by the photoperiod and vernalization (De Wit et al., 2019). Based on a study of the sensitivity of WOFOST model parameters by Xu et al.

Table 1. Production conditions of sites

<table>
<thead>
<tr>
<th>Sites</th>
<th>Sowing date (2019)</th>
<th>Maturity date (2020)</th>
<th>Irrigation date (70 mm each time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dengzhou</td>
<td>10/29</td>
<td>5/26</td>
<td>2019/12/20 2020/4/29</td>
</tr>
<tr>
<td>Fangcheng</td>
<td>10/27</td>
<td>5/26</td>
<td>2019/12/5 2020/2/25</td>
</tr>
<tr>
<td>Huaxian</td>
<td>10/19</td>
<td>6/1</td>
<td>2020/3/22 2020/5/12</td>
</tr>
<tr>
<td>Luoshan</td>
<td>10/20</td>
<td>5/23</td>
<td></td>
</tr>
<tr>
<td>Luoning</td>
<td>10/16</td>
<td>6/5</td>
<td></td>
</tr>
<tr>
<td>Pingyu</td>
<td>10/23</td>
<td>5/25</td>
<td></td>
</tr>
<tr>
<td>Ruyang</td>
<td>10/20</td>
<td>5/29</td>
<td></td>
</tr>
<tr>
<td>Shangqiu</td>
<td>10/18</td>
<td>5/28</td>
<td></td>
</tr>
<tr>
<td>Xixian</td>
<td>10/20</td>
<td>5/23</td>
<td></td>
</tr>
<tr>
<td>Xuchang</td>
<td>10/18</td>
<td>5/27</td>
<td>2019/10/28 2020/3/14 2020/5/4</td>
</tr>
<tr>
<td>Zhoukou</td>
<td>10/22</td>
<td>5/27</td>
<td>2020/3/14 2020/5/5</td>
</tr>
</tbody>
</table>
It was found that the dry matter conversion efficiency (CVO), the maximum light energy utilization rate of a single leaf (EFFTB40), the maximum CO\(_2\) assimilation rate of a single leaf (AMAXTB1.3) and the extinction coefficient (KDIFTB2.0) all have a great influence on the yield. The parameter specific leaf area (SLATB0 and SLATB0.5), root and leaf dry matter partition coefficient (FRTB0 and FLTB0), extinction coefficient (KDIFTB0 and KDIFTB2.0) and the leaf life (life span of leaves growing at 35°C, SPAN) are sensitive to LAI, and it was also found that the sensitive parameters of LAI will change with the growth period. When the moisture level is insufficient, priority should be given to parameters related to early light interception such as KDIFTB0 and SLATB0. Parameters sensitive to LAI and output (such as KDIFTB2.0) need to be adjusted simultaneously.

This study is based on the measured data of Huaxian XH, Kaifeng KF, Luoning LN, and Xixian XX, using the SUBPLEX optimization algorithm (Rowan, 1990) to construct the cost function of the measured and simulated values in order to optimize the sensitive parameters, and the calibrated model was verified in other regions.

When assimilation parameters are being selected, it is important to select those that are sensitive to LAI and yield for assimilation. A sensitivity analysis of the WOFOST model shows that water stress is a key factor affecting the sensitivity of the parameters. Therefore, when assimilating models and remote sensing data in different regions, it is necessary to fully consider the local water supply such as rainfall and irrigation. Based on the results of the parameter sensitivity analysis (Xu et al., 2021), the sensitivity of the WOFOST model parameters can be divided into the following categories: under the condition of sufficient moisture, the parameters sensitive to LAI are TDWI, SLATB0, FRTB0, FLTB0, SLATB0.5, KDIFTB0, KDIFTB2.0, SPAN. Yield-sensitive parameters are CVO, AMAXTB1.3, EFFTB40, KDIFTB2.0, SLATB0.5. Parameters sensitive to both LAI and yield are SLATB0.5 and KDIFTB2.0, while in the case of insufficient moisture, the parameters sensitive to LAI are TDWI, SLATB0, FRTB0, FLTB0, SLATB0.5, KDIFTB0, KDIFTB2.0, SPAN, and PERDL. Yield-sensitive parameters are SLATB0, FRTB0, FLTB0, KDIFTB0, and CVO. Parameters that are sensitive to both LAI and yield are FRTB0, FLTB0, SLATB0, and KDIFTB0. The specific parameter selection is shown in Table 2.

The SUBPLEX algorithm is a seed spatial search algorithm. It is based on the Nelder-Mead (NMS) simplex search algorithm which identifies an improved set of subspaces and then searching for each subspace. This method has a higher computational efficiency than the simplex search or forced search method. In the picture alignment process of the model, it is first necessary to build a cost function, the cost function \( f(x) \) generally uses the root mean square error (RMSE), see formula:

\[
 f(x) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 },
\]

where: \( n \) is the number of observations, and \( x_i \) and \( \bar{x} \) represent the simulated and observed values, respectively.

In addition, the algorithm sets a relative convergence tolerance (\( \varepsilon \)) or the maximum number of iterations at runtime in order to determine the termination condition for the optimization. \( \varepsilon \) determines the threshold for the convergence of the cost function, and theoretically, the smaller the \( \varepsilon \) value, the higher the assimilation accuracy. However, a lower convergence tolerance can greatly increase the computational cost.

The Ensemble Kalman filter is a sequential assimilation algorithm, which introduces the concept of an ensemble based on the Kalman filter, and it is capable of estimating the covariance of model prediction using the ensemble method. The Ensemble Kalman filter is based on the implementation of the Monte Carlo Kalman filter, which can be used for extremely high-dimensional, nonlinear and non-Gaussian state estimation problems (Roth et al., 2017; Huang et al., 2016; Ma et al., 2013). The basic process is to establish a set of state variables, the set contains all of the possible values of the state variables, and the average value of each set member is used as the best estimate of the state variable. All members of the set move forward through the model. When there are additional observations, all members in the set are updated through the observations. If there are no additional observations, it is dynamically updated through the running results of the model.

The basic principle of Kalman filtering is the assumption of an n-dimensional state variable \( x \) and an m-dimensional state variable \( y \), the state variable error at time \( k \) is \( v_k \) and the observation error is \( e_k \), by setting \( \text{cov}(v_k) = Q, \text{cov}(e_k) = R \), then:

\[
 x_{k+1} = Fx_k + Gu_k \, ,
\]

\[
 y_k = Hx_k + e_k \, .
\]

The update and prediction of the mean value and the covariance matrix of the state variable \( x \) are \( \hat{x}_{k|k} \) and \( P_{k|k} \):

\[
 \hat{x}_{k+1|k} = F \hat{x}_{k|k} \, ,
\]

\[
 P_{k+1|k} = FP_{k|k}F^T + GQG^T \, .
\]

For a set with \( N \) members, \( \hat{x}_{k|k} \) and \( P_{k|k} \) may be expressed as:

\[
 \bar{x}_{k|k} = \frac{1}{N} \sum_{i=1}^{N} x_{k|k}^{(i)} \approx \hat{x}_{k|k} \, ,
\]

\[
 \bar{P}_{k|k} = \frac{1}{N-1} \sum_{i=1}^{N} (x_{k|k}^{(i)} - \bar{x}_{k|k})(x_{k|k}^{(i)} - \bar{x}_{k|k})^T \approx P_{k|k} \, .
\]

Table 2. Assimilation parameters under different conditions

<table>
<thead>
<tr>
<th>Water conditions</th>
<th>Assimilation parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent mode</td>
<td>SLATB0.5, KDIFTB2.0, SPAN</td>
</tr>
<tr>
<td>Water restriction mode</td>
<td>SLATB0, KDIFTB0, SPAN</td>
</tr>
</tbody>
</table>
In the formula, \( N \) is the number of set members, and \( i \) is the member index.

According to the characteristics of the vegetation growth curve, Chen et al. (2004) proposed an envelope filtering method based on S-G filtering to reconstruct higher-quality remote sensing data. The specific operation is shown in Fig. 2, the data with cloud pollution is linearly interpolated, followed by an S-G filtering formula as follows:

\[
g_j = \sum_{i=-r}^{i=r} c_i f_{i+j},
\]

where: \( g \) is the filtered data, \( f \) is the raw data, \( c \) is the weight coefficient, \( r \) is the 1/2 filter window, \( j \) is the original data index, and \( i \) is the index within the data window.

After S-G filtering, it is assumed that the local minimum value of the original data is uncertain, and the local maximum value is the true value. The local minimum is replaced by the filtered value. The iteration process continues until the standard deviation between the filtered value and the pre-filtered value reaches the set threshold, and then it ends:

\[
N_i^t = \begin{cases} 
O_i^{t-1} & \text{if } O_i^{t-1} > N_i^{t-1} \\
N_i^{t-1} & \text{if } O_i^{t-1} < N_i^{t-1}
\end{cases},
\]

where: \( O \) and \( N \) are the initial and filtered values, \( t \) is the index of the iteration, and \( i \) is the index of the data.

This study uses LAI as the assimilation variable to assimilate the WOFOST model and MODIS remote sensing data. The assimilation parameters of the model which are coupled with remote sensing must be selected with a greater degree of uncertainty in different regions, and parameters that have a close correlation with the assimilation variable (LAI) are chosen, preferably using the target output variable (yield) (Wu et al., 2021). Based on Xu et al. (2021) the sensitivity of the WOFOST model parameters under different water and irrigation conditions were analysed. It was found that water stress is an important factor affecting the sensitivity of the WOFOST model parameters. It was also found that water stress is an important factor affecting the sensitivity of the WOFOST model parameters. Under water stress conditions, the parameters related to early light interception (specific leaf area SLATB0 when DVS=0, extinction coefficient KDIFTB0 when DVS=0, etc.), where DVS represents the phenological development stage of the WOFOST model, 0 represents seedling emergence, 1 represents flowering, and 2 represents maturity) will become the main factor affecting the model results. Therefore, using the EFAST sensitivity analysis algorithm (Saltelli et al., 2008), the parameters of SLATB0, KDIFTB0, and the specific leaf area SLATB0.5 when DVS=0.5 and the extinction coefficient KDIFTB2.0 when DVS=2.0 were selected for sensitivity analysis at each location. With regard to the water conditions in different regions, assuming that the sensitivity of SLATB0 and KDIFTB0 is greater than that of SLATB0.5 and KDIFTB2.0, respectively, it is considered that the local water supply is insufficient, and vice versa.

Specific leaf area (Specific leaf area at growth, SLATB) is easily affected by factors such as variety, soil, planting density, etc. (Meziane and Shipley, 2002; He et al., 2019), and also has a large spatial variability. The extinction coefficient is closely related to the leaf angle (Wang et al., 2007), row spacing (Flénet et al., 1996), and other factors. In addition, leaf life (SPAN) is the determinant of LAI in the later stage of wheat growth, and SPAN reflects the influence of water, nutrients, weeds, diseases, and insect pests on LAI to a certain extent, also, there is a significant degree of uncertainty in large range application (Bai et al., 2019; De Wit et al., 2012), so it is also considered to be an assimilation parameter. Therefore, SLATB0.5, KDIFTB2.0, and SPAN were selected to be the assimilation parameters when the water supply is sufficient. When the water supply is insufficient, SLATB0, KDIFTB0, SPAN should be selected as the assimilation parameters. When the water supply status is not considered, SLATB0.5, KDIFTB2.0, SLATB0, KDIFTB0, SPAN should be selected in this case.

In order to explore the influence of different production conditions and also different parameter choices on data assimilation two modes of WOFOST model, latent and water restriction (the water conditions in all regions under the latent

---

**Fig. 2.** Flow chart of S-G envelope filter.
model are sufficient) were selected for assimilation. Model and remote sensing assimilation of different water supply conditions were carried out in each mode. The LAI of remote sensing observations is determined by matching the GPS position information of the sampling points with the MODIS position information. The specific process is shown in Fig. 3.

RESULTS AND ANALYSIS

The WOFOST model can better simulate the flowering and maturity of wheat in Henan Province, and also, the model’s average calibration error and average verification error are both within the two-day range (Table 3). The calibration error of the model in the Huaxian, Kaifeng, Luoning, and Xixian areas was 0.75 days during the flowering and maturity periods, and the errors verified in the other areas were 1.62 and 1.08 days, respectively. Among them, simulation errors of the flowering period in Shangqiu and Dengzhou are relatively large, and are 3 and 4 days earlier than the observed value, respectively. The simulation errors of the mature periods of Xiuwu and Dengzhou are relatively large, they are 3 and 4 days earlier than the observed values, respectively. The simulation of Ruyang’s flowering period was 3 days later than the observed value. The simulation errors of the flowering and maturity periods in other regions were within 2 days of the actual times.

Remote sensing data are susceptible to factors such as clouds and atmospheric sols, several observational drops are prone to occur in a given time series, which obviously violates the growth law of crops. In this study, S-G envelope filtering was used to de-dry the MODIS data. Using the measured LAI values in the field, the ratio coefficient was established from the measured data of different growth periods of wheat, and the MODIS data was corrected, the results are shown in Fig. 4. For comparison, the original MODIS-LAI has several sharp drops in the time series, and the data resulting from S-G envelope filtering can better remove the noise of the MODIS data thereby making the curve of the LAI smoother while retaining the change trend of LAI with the fertility period. After the measured value correction, the LAI curve maintains the LAI change trend after filtering, and increases the value of LAI, so that the value of LAI is closer to the measured value. The calibrated MODIS data basically conforms to the actual change curve of LAI, thereby reducing the systematic error of the MODIS data.

The calibrated WOFOST model is capable of reflecting the exponential growth of wheat LAI in Henan Province after the emergence of seedlings. Wheat stops growing during the overwintering period. The calibrated WOFOST model is capable of reflecting the exponential growth of wheat LAI in Henan Province after the emergence of seedlings. Wheat stops growing during the overwintering period. The calibrated WOFOST model is capable of reflecting the exponential growth of wheat LAI in Henan Province after the emergence of seedlings. Wheat stops growing during the overwintering period.

<table>
<thead>
<tr>
<th>Site</th>
<th>Anthesis</th>
<th>Mature anthesis</th>
<th>veri. date</th>
<th>veri. date</th>
<th>Anthesis error</th>
<th>Maturity error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huaxian</td>
<td>4/25</td>
<td>4/24</td>
<td>6/1</td>
<td>5/30</td>
<td>–1</td>
<td>–2</td>
</tr>
<tr>
<td>Kaifeng</td>
<td>4/22</td>
<td>4/22</td>
<td>5/28</td>
<td>5/28</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Luoning</td>
<td>5/3</td>
<td>5/1</td>
<td>6/5</td>
<td>6/6</td>
<td>–2</td>
<td>1</td>
</tr>
<tr>
<td>Xixian</td>
<td>4/15</td>
<td>4/15</td>
<td>5/23</td>
<td>5/23</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dengzhou</td>
<td>4/17</td>
<td>4/14</td>
<td>5/26</td>
<td>5/22</td>
<td>–3</td>
<td>–4</td>
</tr>
<tr>
<td>Fangcheng</td>
<td>4/18</td>
<td>4/18</td>
<td>5/26</td>
<td>5/26</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jiaozuo</td>
<td>4/26</td>
<td>4/28</td>
<td>6/1</td>
<td>6/2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Luoshan</td>
<td>4/15</td>
<td>4/15</td>
<td>5/23</td>
<td>5/23</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Luoke</td>
<td>4/18</td>
<td>4/19</td>
<td>5/26</td>
<td>5/26</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Neihuang</td>
<td>4/24</td>
<td>4/25</td>
<td>6/1</td>
<td>5/30</td>
<td>1</td>
<td>–2</td>
</tr>
<tr>
<td>Pingyu</td>
<td>4/17</td>
<td>4/17</td>
<td>5/25</td>
<td>5/24</td>
<td>0</td>
<td>–1</td>
</tr>
<tr>
<td>Ruyang</td>
<td>4/21</td>
<td>4/24</td>
<td>5/29</td>
<td>5/30</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Shangqiu</td>
<td>4/26</td>
<td>4/22</td>
<td>5/28</td>
<td>5/29</td>
<td>–4</td>
<td>1</td>
</tr>
<tr>
<td>Xuchang</td>
<td>4/19</td>
<td>4/21</td>
<td>5/27</td>
<td>5/27</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Yuanyang</td>
<td>4/21</td>
<td>4/23</td>
<td>5/29</td>
<td>5/29</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Verification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.62</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Table 3. Observed phenology and the simulated phenology of the model in various sites, 2020
CONSIDERING DIFFERENT WATER SUPPLIES CAN IMPROVE THE ACCURACY OF THE WOFOST

Considering different water supplies can improve the accuracy of the WOFOST period, the wheat grows rapidly and reaches its maximum value around the booting stage, and the LAI begins to decrease during the reproductive growth stage. The characteristics of this change are shown in Fig. 5. Moreover, the performance results in Huaxian, Kaifeng, Luoning and Xixian counties are better. The $R^2$ of the simulated and measured values in the four regions is between 0.87-0.98, and the RMSE is between 0.34-0.79. Among them, the LAI in the jointing period of Luoning was underestimated, and the LAI in the flowering period in Huaxian and Xixian was overestimated. The accuracy of the LAI verification result is lower than that of the calibration result, the verified $R^2$ is 0.77, and the RMSE is 1.06.

The calibration results of the crop yield are shown in Fig. 6a. The calibration results of this yield in the four locations of Huaxian, Kaifeng, Luoning, and Xixian perform

Fig. 4. Comparison of remote sensing data before and after correction: Ori – original data, S-G – data after S-G filter, Cor – data after correction, Obs – observation data.

Fig. 5. Calibration and validation results of LAI: a – HX, b – KF, c – LN, d – XX, e – validation.
well. The simulation errors of the four regions respectively are 0.29, 0.22, 0.34 and 0.10 t ha$^{-1}$. The results of the verification at 13 other locations are shown in Fig. 6b. The $R^2$ of the verification results is 0.66 and the RMSE is 1.43 t ha$^{-1}$. Compared with the simulation results, the model's error which was verified in other regions is larger. This shows that when the model is applied in different regions, the use of only one set of crop parameters will lead to obvious simulation errors. Therefore, it is necessary to correct the model in combination with remote sensing data when it is applied to a particular region.

The sensitivity of the parameters SLATB0, KDIFFT0, SLATB0.5, and KDIFFT2.0 at each test site is shown in Fig. 7. Among them, Shangqiu, Pingyu, Luoning and Ruyang had less rainfall than average and no irrigation (Table 1), the sensitivity of SLATB0 and KDIFFT0 is much greater than that of SLATB0.5 and KDIFFT2.0, thus indicating that they are lacking in some respect. Although Xuchang, Jiaozuo, Huaxian, and Yuanyang were irrigated during the wheat growth period (Table 1), the sensitivity of SLATB0 and KDIFFT0 was still greater than that of SLATB0.5 and KDIFFT2.0. Although the areas with a greater degree of precipitation in the south (such as Luoshan and Xixian) had not been irrigated, the sensitivity of SLATB0.5 and KDIFFT2.0 is greater than that of SLATB0 and KDIFFT0. Overall, Jiaozuo, Pingyu, Shangqiu, Xuchang, Luoning, Ruyang, Yuanyang and Huaxian regions have an insufficient water supply. When selecting the assimilation parameters according to the water supply situation in the water restriction mode, select SLATB0, KDIFFT0 and SPAN as the assimilation parameters. For other regions SLATB0.5, KDIFFT2.0 and SPAN should be chosen as the assimilation parameters.

The 17 test sites were simulated and assimilated using models under potential production conditions, using two assimilation schemes: the first scheme (Assim1) did not distinguish between moisture conditions and took into account all parameters under the two moisture supplies (SLATB0, KDIFFT0, SLATB0.5, KDIFFT2.0 and SPAN) for assimilation. The second assimilation protocol (Assim2) is assimilated using assimilation parameters (SLATB0.5, KDIFFT2.0, and SPAN) under moisture-well-adequate conditions, as shown in Fig. 8. The simulation results (Sim) under potential conditions do not reflect the differences between the regions, and the simulation results are generally higher than the measured results, especially in areas with insufficient moisture (Neihuang, Yuanyang, Ruyang, etc.). After the assimilation of LAI information with remote sensing, the results of the simulation of the two assimilation schemes were significantly improved.

Seventeen test sites were simulated and assimilated using a moisture-constrained model, as shown in Fig. 9. Models under moisture limitations have been used successfully to reflect the yield differences between the regions. The results of assimilation in most areas are closer to the measured values than the simulation results.
The relationship between the measured results of the WOFOST model using the latent mode and the simulation results and assimilation results is shown in Fig. 10. The model results under the potential production mode are not limited by moisture, and the simulation results have small differences between the different regions, and therefore cannot reflect the differences between the various locations. The simulated yield under the potential mode is distributed between 10-11.6 t ha\(^{-1}\), and the correlation with the measured value is weak, \(R^2\) is 0, and the RMSE is 2.5 t ha\(^{-1}\). Taking into account the assimilation with LAI-sensitive parameters (SLATB0, KDIFTB0, SLATB0.5, KDIFTB2.0, and SPAN), the results have been significantly improved in comparison with the simulation results, \(R^2\) increased to 0.48, and RMSE decreased to 1.56 t ha\(^{-1}\). The assimilation results using the assimilation parameters (SLATB0.5, KDIFTB2.0 and SPAN) under sufficient water conditions were slightly improved, \(R^2\) increased to 0.50, RMSE was reduced to 1.47 t ha\(^{-1}\), but the improvement effect was minor.

In water restriction mode, the relationship between the measured results of the WOFOST model and both the simulation and assimilation results are shown in Fig. 11. It may be observed from the figure that among the different assimilation strategies, the assimilation results of selecting the corresponding assimilation parameters according to the water supply status of the different regions perform best. There is a good correlation between the model simulation results and the measured values in the water restriction mode, with an \(R^2\) value of 0.68 and an RMSE of 1.05 t ha\(^{-1}\). Both of the assimilation methods can be used to improve the simulation accuracy of the model. The \(R^2\) and RMSE of the assimilation results without taking into account the water difference between regions were 0.79 and 0.64 t ha\(^{-1}\), respectively. According to the water conditions in
In previous studies, in order to reduce the model's uncertainty, the model's production mode under latent conditions was often assimilated with remote sensing data (Ma et al., 2013; Huang et al., 2015; Curnel et al., 2011). Since the potential model does not take into account the stresses of water, nutrients, etc., when this study uses the model under certain potential conditions to simulate, the results in the region are not satisfactory (Fig. 10a). However, after using remote sensing information to assimilate LAI, the simulation results of the model show a good correlation with the measured results (Fig. 10b, c). In addition, when simulation and assimilation were performed using the model in a water restriction mode in this study, the R² of the simulation result reached 0.68, and the R² of the assimilation result reached 0.79 or more (Fig. 11), this indicates that considering water supply assimilation can serve to improve the simulation accuracy, this is similar to the research results of (Pan et al., 2019) using remote sensing to monitor soil moisture and LAI. Based on remote sensing information and the WOFOST model (Jin et al., 2015) performed an assimilation simulation of heavy metal stress in rice and simulated rice growth. These studies have shown that it is necessary to consider the various stresses that may be encountered during the actual growth cycle of crops when the model is applied to a particular region. In recent years, model simulation studies concerning high temperature, freezing damage and other adversity stresses have also been continuously strengthened (De Wit et al., 2019; Wang et al., 2017). Therefore, water stress was considered in the process of model assimilation, as it will provide a favourable foundation for the regional application of remote sensing and model assimilation.

When performing remote sensing and crop model assimilation, the selection of appropriate assimilation parameters is also very important (Jin et al., 2018). Assimilation parameters must be selected and also assimilation variables (such as assimilation yield and simulated yield) should be selected for the assimilation process. The assimilation results were compared with the measured results, and the RMSE was reduced by 39.0% and 51.4%, respectively. Therefore, when the model is applied to a particular region, it is necessary to consider the water limit, and the use of remote sensing data to assimilate the model can effectively improve the simulation accuracy of the model.

DISSCUSSION

The WOFOST model is better at simulating the phenological period of wheat (Table 3), but the simulated values of the flowering and mature periods in Dengzhou and Xiuwu occur earlier than the observed values, while the simulated values of the flowering and mature periods in Ruyang both occur later than the observed value. This may be because the WOFOST model did not take into account the influence of water conditions on the growth period of wheat when simulating the development of wheat phenology. In Dengzhou and Xiuwu, excessive rainfall and irrigation water caused late ripening of wheat, while in Ruyang, the flowering time was advanced due to water stress (McMaster et al., 2019). At present, research concerning crop phenology simulation is continuously using a wider scope, and many studies have begun to pay attention to the effects of nutrient stress (Singh and Wilkens, 2000) and even high-temperature stress (Hossain and Teixeira da Silva, 2012) on the development of crop phenology. The calibrated WOFOST model performed well in the simulation of LAI and production, but the simulation accuracy of the model was significantly reduced when it was verified in other regions (Figs 5, 6). This may be because the crop model is a simplified simulation of the crop growth process, and the degree of parameterization of the model for crop growth needs to be improved (Yin et al., 2020). In addition, most models currently place a lower priority on factors such as diseases, pests, weeds, temperature stress, etc. with regard to crop growth (Rasche and Taylor, 2019), so there is a greater degree of uncertainty in the application of the model to a regional scale (Ramirez-Villegas et al., 2017). Therefore, it is necessary to improve the regional simulation of the model in combination with the regional monitoring capabilities of remote sensing.

In previous studies, in order to reduce the model's uncertainty, the model's production mode under latent conditions was often assimilated with remote sensing data (Ma et al., 2013; Huang et al., 2015; Curnel et al., 2011). Since the potential model does not take into account the stresses of water, nutrients, etc., when this study uses the model under certain potential conditions to simulate, the results in the region are not satisfactory (Fig. 10a). However, after using remote sensing information to assimilate LAI, the simulation results of the model show a good correlation with the measured results (Fig. 10b, c). In addition, when simulation and assimilation were performed using the model in a water restriction mode in this study, the R² of the simulation result reached 0.68, and the R² of the assimilation result reached 0.79 or more (Fig. 11), this indicates that considering water supply assimilation can serve to improve the simulation accuracy, this is similar to the research results of (Pan et al., 2019) using remote sensing to monitor soil moisture and LAI. Based on remote sensing information and the WOFOST model (Jin et al., 2015) performed an assimilation simulation of heavy metal stress in rice and simulated rice growth. These studies have shown that it is necessary to consider the various stresses that may be encountered during the actual growth cycle of crops when the model is applied to a particular region. In recent years, model simulation studies concerning high temperature, freezing damage and other adversity stresses have also been continuously strengthened (De Wit et al., 2019; Wang et al., 2017). Therefore, water stress was considered in the process of model assimilation, as it will provide a favourable foundation for the regional application of remote sensing and model assimilation.

When performing remote sensing and crop model assimilation, the selection of appropriate assimilation parameters is also very important (Jin et al., 2018). Assimilation parameters must be selected and also assimilation variables (such as assimilation yield and simulated yield) should be selected for the assimilation process. The assimilation results were compared with the measured results, and the RMSE was reduced by 39.0% and 51.4%, respectively. Therefore, when the model is applied to a particular region, it is necessary to consider the water limit, and the use of remote sensing data to assimilate the model can effectively improve the simulation accuracy of the model.

Fig. 11. The relationship between measured yield, simulated yield, and assimilated yield: a – simulated value and measured value, b – the assimilation value and the measured value of LAI-sensitive parameters were selected, c – the results of using assimilation.

The relationship between measured yield, simulated yield, and assimilated yield: a – simulated value and measured value, b – the assimilation value and the measured value of LAI-sensitive parameters were selected, c – the results of using assimilation.
as LAI), preferably at the same time, and the assimilation targets (such as output) should have strong correlation parameters, and these parameters are often difficult to obtain on a regional scale (Wu et al., 2021). At the regional scale, the climatic conditions and differences in management measurements between different regions will affect the sensitivity of the model's parameters (Liu et al., 2019). Based on the sensitivity analysis results of the WOFOST model under different climate and production conditions (Xu et al., 2021), this study puts forward a selection strategy for assimilation parameters under different water supply conditions, these changes amount to a significant difference and an improvement in comparison with previous studies. The assimilation results that did not distinguish between the water supply conditions according to the different regions and the corresponding assimilation parameters which were selected according to the water supply conditions between the regions are analysed. The study found that selecting the corresponding assimilation parameters based on the water supply characteristics of different locations can serve to improve the assimilation effect to a greater extent. Without distinguishing between the assimilation of water supply conditions, despite the fact that there are other assimilation parameters, the accuracy of the assimilation results is not improved (Figs 10b, c, and 11b, c), this is consistent with the research results of Ma (Ma et al., 2013).

This study used a sensitivity analysis of the parameters sensitive to water stress in order to determine the water supply under different climatic and production conditions. Based on the water supply situation, the corresponding assimilation parameter selection principle is proposed, which in turn improves the application of remote sensing and crop model assimilation in different climate and production conditions. The data and research scope of this study is mainly focused on the Henan Province of China, therefore there may be some uncertainty in other areas, and a wider range of research data should be considered in future studies. This method can also be extended to the assimilation simulation of other crop models or other crop growth stress factors (such as nutrient stress, etc.). With the development of geographic information systems, big data, remote sensing and other information technologies, enhancing the analysis of climate and production conditions will greatly promote the application of remote sensing and model assimilation on a regional scale.

CONCLUSIONS

In this study, the world food studies model was calibrated. The calibrated model can be used to improve the simulation of the wheat growth period in Henan Province, with the average simulation error occurring within the time period of 2 days. When leaf area index and yield are calibrated, the model produced better calibration results. The verification results originating from other regions are slightly worse. This shows that when the model is applied within a particular region, the model needs to be corrected with the appropriate remote sensing data.

1. The assimilation of wheat yield predictions based on Moderate-resolution imaging spectroradiometer remote sensing data and the world food studies crop model may be realized using the ensemble Kalman filter algorithm.

2. Based on a sensitivity analysis, the water supply situation of different test points was determined, and a suitable selection strategy of assimilation parameters under different water supply conditions was proposed.

3. The results show that the use of the model under a water restriction mode and the use of the corresponding assimilation parameters for assimilation modelling in combination with the water supply of different locations can effectively improve the application capabilities of crop models and remote sensing assimilation in different climate and management conditions.

ACKNOWLEDGEMENTS

The authors would like to thank the graduate students at the College of Information and Management Science and the College of Agronomy at Henan Agricultural University for their continued support of this research.

Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES


CONSIDERING DIFFERENT WATER SUPPLIES CAN IMPROVE THE ACCURACY OF THE WOFOST


