Assimilating Remotely Sensed Information with the WheatGrow Model Based on the Ensemble Square Root Filter for Improving Regional Wheat Yield Forecasts

Yan Huang, Yan Zhu, Wenlong Li, Weixing Cao and Yongchao Tian

(National Engineering and Technology Center for Information Agriculture, Nanjing Agricultural University, Nanjing, Jiangsu 210095, China)

Abstract: In this study, a deterministic algorithm named Ensemble Square Root Filter (EnSRF), an algorithm significantly improved from the Ensemble Kalman Filter (EnKF), was used to integrate remotely sensed information (ASD spectral data, HJ-1 A/B CCD and Landsat-5 TM data) with a wheat (Triticum aestivum L.) growth model (WheatGrow). The analyzed values of model variables, leaf area index (LAI) and leaf nitrogen accumulation (LNA), were calculated based on EnSRF without perturbed measurements. Independent datasets were used to test EnSRF and the root mean square error (RMSE) values were 0.81 and 0.82 g m$^{-2}$, with relative error (RE) values of 0.15 and 0.13, for LAI and LNA, respectively. RMSE values for LAI and LNA were 1.39 and 1.70 g m$^{-2}$, respectively (RE, 0.28 and 0.34) based on EnKF, 1.17 and 1.80 g m$^{-2}$ (RE, 0.24 and 0.35), respectively, based on the WheatGrow model alone, and 0.97 and 1.25 g m$^{-2}$ (RE, 0.21 and 0.24), respectively, based on the remote sensing models. These results indicated that the LAI and LNA values based on EnSRF matched the measured values well compared with the EnKF, WheatGrow and remote sensing models. In addition, the predicted results are consistent with the temporal and spatial distribution of winter wheat growth status and grain yields in the study area, with RE values of less than 0.2 and 0.1 for LAI and LNA, respectively. These results provide an important approach for simulating winter wheat growth status based on combining remote sensing and crop growth models.

Key words: Data assimilation, Ensemble Square Root Filter, Remote sensing, Updating, WheatGrow model.

Crop growth models, which incorporate data regarding biological and physiological parameters of crops, can describe the interactions between the genetic potential, environment and field management strategies of individual crops by simulating the dynamic rhythm patterns of crop growth. Crop models are initially derived from the field scale experiments, in which many conditions are supposedly homogeneous and a parameter is often identified by the fact that it takes the same fixed value for all site-years or possibly in a group of site-years (Wallach et al., 2001). When applied to a large area, the need to account for the intrinsic spatial variations present in crop ecosystems (for example, in soil properties, management practices, etc.) has lead to the development of spatializing crop models. However, spatialized models require a large amount of inputted variables that is virtually impossible to gather with a sufficient degree of confidence, which introduces uncertainties and limits the simulation and predictive accuracy of crop modeling (Launay and Guerif, 2005). Remote sensing can provide information on crop and soil conditions over large areas and effectively resolve these problems (Dorigo et al., 2007; Li et al., 2008). Therefore, the combined use of remote sensing-derived biophysical/biochemical state variables and crop models is expected to improve their predictive performance at the regional scale. It is highly important to study the methods that integrate remote sensing and crop growth models.

Various methods have been developed to integrate remotely sensed observations into crop models. In general, three different strategies can be effectively applied...
(Delécolle et al., 1992; Dorigo et al., 2007), including the forcing strategy, which directly replaces stated variables in the model using observed data; the initialization/parameterization strategy (calibration), which obtains an optimal agreement between the simulated variables and the observed variables by adjusting model parameters of initial states; and the process updating strategy, which optimizes the crop simulation value using remotely sensed observations at a specific moment and changes the simulated variables after the remote sensing footprint. Both the forcing and calibration strategies are based on the assumption that remote sensing observations can reflect the objective reality well and avoid introducing error into the model (Launay and Guerif, 2005; Zhao et al., 2005). However, the updating strategy has received increased attention as it simultaneously removes the errors of both remote sensing observation and crop growth simulation. For example, de Wit and van Diepen (2007) improved the water balance simulation model of the World Food Studies (WOFOST) model and the biomass predictive accuracy for the winter wheat and corn models by combining the remotely sensed and simulated soil moisture inversions through the updating approach. Wang (2010) also coupled the Moderate-resolution Imaging Spectroradiometer (MODIS) reflectance data with the Scattering by Arbitrary Inclined Leaves (SAIL) radiance transfer model (Verhoef, 1984) and improved the leaf area index (LAI) simulation accuracy of wheat predicted by the Crop Environment Resource Synthesis (CERES)-Wheat model (Godwin et al., 1990).

The ensemble Kalman filter (EnKF) considered to be an effective algorithm for the data assimilation approach to quantify the relative weight between the observed variable and the model variable (Evensen, 1994; Crow and Wood, 2003; Reichle et al., 2008), can estimate forecast error covariance based on the Monte Carlo short-term ensemble forecasting methods, and has been generally used in data assimilation (Evensen, 1994). EnKF uses the error covariance matrix of the observed and state variables to update the ensembles in the forward integrating process, and obtains the ensembles of the analysis field for the next time point. Then, the mean value is considered as the optimal estimate of the model state variable at that moment. However, in the analysis process with EnKF, eventually filter divergence occurs by underestimating the analysis error covariance matrix, without absorbing measurement information in the assimilation process. Burgers (1998) tried to solve this problem by adding a random perturbation to the measurements, but the filtering performance was lowered by the introduction of additional error to the standard Kalman filter analysis. Whitaker and Hamill (2002) proposed a new deterministic method based on a Kalman filtering scheme, the ensemble square root filter (EnSRF). EnSRF has been used in assimilating observations in both ecological and meteorological models. Das et al. (2008) successfully estimated soil moisture via assimilating the aircraft-based remotely sensed surface soil moisture into the Soil-Water-Atmosphere-Plant model. Slater and Clark (2009) improved the snow model (SNOW-17) estimate of snow water equivalent by merging the uncertainties associated with meteorological forcing data and snow water equivalent observations within the model based on EnSRF. Accordingly, EnSRF should also perform well in assimilating remotely sensed information for crop growth models in order to improve regional crop yield forecasts.

The objectives of the present study were: (1) to introduce the EnSRF algorithm to the assimilation process of remote sensing and the WheatGrow model, which was developed to predict wheat growth status and grain yield (Cao et al., 2002), and to compare it with the EnKF algorithm; (2) to develop an updating method based on assimilating two remotely sensed crop growth parameters, the leaf area index (LAI) and leaf nitrogen accumulation (LNA); and (3) to establish a remote sensing-WheatGrow integration model for wheat growth status and grain yield prediction at both the field and regional scales.

Materials and Methods

1. Study sites

The data used in the present study were obtained from four wheat field experiments and one regional scale field observation in the study area.

(1) Experimental fields

The four wheat field experiments involved different sites and years, winter wheat cultivars, nitrogen rates and density treatments (Table 1). A detailed description of the experimental design is given below.

Experiment 1 was undertaken at the experiment station of the Jiangsu Academy of Agricultural Sciences at Nanjing in the season from 2004 to 2005 (32°02′N, 118°01′E). Two winter wheat cultivars, Ningmai 9 and Yangmai 12, were sown on 7 November. Nitrogen (N) fertilizer was applied at 0 (N1), 75 (N2), 150 (N4) and 225 (N7) kg N ha$^{-1}$ with 50% applied at the pre-planting stage and 50% at jointing. The row space was 25 cm and sowing rate was 84 kg ha$^{-1}$ (as monocalcium phosphate [Ca(H$_2$PO$_4$)$_2$]) and 150 kg K$_2$O ha$^{-1}$ (as KCl) were applied prior to seeding. The experiment was performed with a randomized complete block design, with three replications and 18 m$^2$ area for each plot. Other management strategies followed local standard practices in wheat production.

Experiment 2 was conducted in the season from November 2007 to June 2008 at Jiangpu experiment station of Nanjing Agricultural University (32°02′N, 118°37′E). The winter wheat cultivar Ningmai 9 was sown
on 3 November. N fertilizer was applied at 90 (N3), 180 (N5) and 270 (N8) kg N ha\(^{-1}\) with 50\% applied at pre-planting and 50\% at jointing. The experiment was performed with a randomized complete block design with three replications and a 4 m × 4 m = 16 m\(^2\) area was used for each plot. The row space and sowing rate was 25 cm and 79.5 kg ha\(^{-1}\) (D3), respectively. Monocalcium phosphate and potassium chloride were applied at basal dose of 150 kg P\(_2\)O\(_5\) ha\(^{-1}\) and 150 kg K\(_2\)O ha\(^{-1}\), respectively, in each treatment (plot).

Experiment 3 was conducted in the season from October 2008 to June 2009 at Haian County (32º30′ N, 120º 23′ E). The winter wheat cultivar Ningmai 13 was sown on 27 October. N fertilizer was applied at 150 (N4), 210 (N6) and 270 kg N ha\(^{-1}\) (N8) with 40\% at the pre-planting stage, 40\% at the seedling stage and 20\% at jointing. The experiment was established as a randomized complete block design with two replications and each research plot was larger than 90 m × 90 m = 0.81 ha area in total. The sowing rate was 67.5 kg ha\(^{-1}\) (D2). Monocalcium phosphate and potassium chloride were applied at basal dose of 120 kg P\(_2\)O\(_5\) ha\(^{-1}\) and 190 kg K\(_2\)O ha\(^{-1}\) for each treatment (plot).

Experiment 4 was carried out between November 2008 and June 2009 at Rugao City (32º30′ N, 120º35′ E). The winter wheat cultivar Yangmai 13 was sown on 13 November. N fertilizer was applied at 225 (N7), 292.5 (N9) and 360 kg N ha\(^{-1}\) (N10) with 40\% at the pre-planting stage, 40\% at the seedling stage and 20\% at jointing. The two sowing rates were 60.8 kg ha\(^{-1}\) (D1) and 82.5 kg ha\(^{-1}\) (D4). The design of the experimental plots and the applications of P and K fertilization were the same as described for Experiment 3.

(2) Sampling points in the study region

A total of 25 and 44 winter wheat cultivation sampling points in the farmers’ fields, which were relatively dispersed and covered the study region, were set up (using GPS locations) in the 2007–2008 and 2008–2009 seasons, respectively; the locations of the points in the two seasons partially overlapped. All fields were larger than 30 m × 30 m. The sampling points were located in Haian County and
Rugao City of Jiangsu Province, China, which lie in the southeastern region of the Yangtze-Huai plain. This region belongs to the main weak-gluten wheat production zone of China, with a total cultivation area of $1.31 \times 10^5$ ha. The terrain of the study area was flat, with an altitude of approximately 1 meter above sea level. In the study area, the agronomic parameters and grain yield were measured using the same method in Experiments 3 and 4. Information about the regional scale observation is shown in Table 2.

2. Data acquisition

(1) Acquisition and pre-processing of satellite imagery

A Landsat-5 TM image and a SPOT-5 HRG image at the grain filling stage (2 May 2008) and jointing stage were acquired for study (1 April 2009), and three further HJ-1 A/B CCD images were acquired at the jointing stage (6 April 2009), heading stage (22 April 2009) and grain filling stage (5 May 2009), respectively. Images were pre-processed by ENVI software (RSI, 2006). The basic parameters of the HJ-1 A/B CCD sensor are shown in Table 3. Firstly, geometric correction for the HRG images was carried out using 40 GPS ground control points. Then, CCD and TM images were calibrated against the corrected HRG image. Both HRG and TM images were calibrated against the HRG image. Both HRG and TM images were corrected by the FLAASH model of the ENVI software, while CCD images were corrected by the empirical linear method (Shen et al., 2007). Combined with the field survey and ground-based spectral data, the supervised classification and reflectance threshold segmentation of the TM and HRG images were carried out to extract the winter wheat planting area (Yang et al., 2008). The obtained TM image classification was used to mask the CCD images to obtain the canopy reflectance data of the winter wheat in the study area.

(2) Canopy spectral reflectance measurements

Ground-based wheat canopy spectral reflectance was measured using a FieldSpec Pro FR 2500 spectroradiometer (Analytical Spectral Devices, Boulder, CO, USA). This instrument records reflectance between 350 and 1,000 nm, with a sampling interval of 1.40 nm and a resolution of 3 nm, and reflectance between 1,000 and 2,500 nm, with a sampling interval of 2 nm and a resolution of 10 nm. All spectral measurements were made during cloud-free periods at midday, between 1000 and 1400. The instrument has a 25° field of view, and was placed at nadir 1 m above the wheat canopy resulting in a view area diameter of approximately 0.44 m$^2$. A white Spectralon reference panel (Labsphere, North Sutton, NH, USA) was used under the same illumination conditions to convert the spectral radiance measurements to reflectance. In Experiments 2, 3 and 4, there were five spectra sampling points for each plot established with a 75 m line drawn through each plot, and a 15 m interval between two neighboring points. A single-point measurement from ten scans was used for the ground-based analysis, while the average of five points was used to create the final data for each plot. In Experiment 1 and the study area of the farmers’ field which covered all of Haian County and Rugao City of Jiangsu Province, China, ten scans were made for each sample point to produce the final canopy spectra.

(3) Determination of agronomic parameters

The agronomic parameters were measured at the corresponding spectral sampling points and sampling times, as detailed in Table 1. After each measurement of the canopy spectral reflectance, 20 plants (4 plants × 5...
Table 4. Algorithms for estimating agronomic parameters based on remote sensing.

| Sensor         | Agronomic parameter | Vegetation index | Regression model               | $R^2$  | Reference         |
|----------------|---------------------|------------------|--------------------------------|--------|------------------|
| LSD            | LAI                 | RVI (810, 560)   | $y = 0.5613x + 0.1833$         | 0.767  | Feng et al., 2009|
|                | LNA                 | RVI (870, 660)   | $y = 0.3614x + 0.392$          | 0.850  | Zhu et al., 2008 |
| Landsat-5 TM   | LAI                 | RVI (4,3)        | $y = 0.216x - 0.356$           | 0.778  |                  |
|                | LNA                 | RVI (4,2)        | $y = 0.458x - 0.347$           | 0.682  |                  |
| HJ-1 CCD       | LAI                 | RVI (4,5)        | $y = 0.292x - 0.093$           | 0.699  | This paper        |
|                | LNA                 | RVI (4,2)        | $y = 0.748x - 1.170$           | 0.639  |                  |

LAI, LNA and RVI are leaf area index, leaf nitrogen accumulation and ratio vegetation index, respectively. For ASD spectral data, the numbers in parentheses for the vegetation indices indicate the wavelength of selected bands, while for Landsat and HJ-1 data, the numbers in parentheses indicate the band numbers of the images.

3. Data utilization and analysis

(1) Monitoring models for LNA and LAI

For the space-borne data, the 1-nm interval ASD canopy reflectance data were fitted to the wide bands of HJ-1 A/B CCD and Landsat-5 TM by the spectral resampling method using ENVI software (RSI, 2006) according to the spectral responding functions of the sensors (Numata et al., 2008; Yuan et al., 2009). In addition, vegetation indices (VIs) for estimating LAI and LNA were extracted using these resampling data. Then, VIs were correlated to the measured LNA and LAI to develop space-borne estimation models for both the LNA and LAI; the ratio vegetation index (RVI) performed well and was selected for estimating the LNA and LAI for the Landsat-5 TM and HJ-1 CCD images. Ground-based estimation models for the LNA and LAI were developed based on spectral vegetation indices (Zhu et al., 2008; Feng et al., 2009) (Table 4).

(2) Data utilization and statistical analysis

The ground-based canopy spectral and agronomic parameters of Experiments 1 – 4 were used for testing the prediction accuracy of the remote sensing-WheatGrow coupling model based on EnSRF that was developed in the present paper and compared with the model based on EnKF. Ground-based canopy spectral and agronomic parameters from Experiments 3 – 4 were used to develop the estimation models for LAI and LNA based on the TM and CCD sensors, respectively. Space-borne remote sensing data and agronomic data from the sampling fields in the study zones were used for testing the validity of the remote sensing-WheatGrow coupling model on a regional scale. The root mean square error (RMSE, as formula 2) and mean relative error (RE (%), as formula 3) were used to calculate the fitness between the simulated and measured values and to evaluate the reliability of the assimilation technique (Jamieson et al., 1991).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(S_i - O_i)^2}{N}}$$

$$\text{RE} = \frac{\sum_{i=1}^{N}(S_i - O_i)}{O_i} \times 100\%$$
overbar represents the ensemble mean as formula 8.

\[
R = \frac{\sum_{i=1}^{N} (S - O)}{N} \tag{3}
\]

Where \(O\) is the measured value, \(S\) is simulated value and \(N\) is the number of samples.

**Data Assimilation Algorithm and Model Selection**

1. **Ensemble Square Root Filter (EnSRF)**

The algorithm of EnSRF (Whitaker and Hamill, 2002; Das et al., 2008) is described as follows. Based on the notation of Ide et al. (1997), \(\hat{X}^b\) is an \(m\)-dimensional background model forecast; \(y^b\) is a \(p\)-dimensional set of observations; \(H\) is the operator that converts the model space to the observation space (here, identity matrix, as the remote sensing measurement is the external one); \(P^b\) is the \(m \times m\)-dimensional background-error covariance matrix; and \(R\) is the \(p \times p\)-dimensional observation-error covariance matrix. The minimum error-variance estimate of the analyzed (or updated) state \(\hat{X}^a\), is given by the traditional Kalman filter update equation (Lorenc, 1986).

\[
\hat{X}^a = \hat{X}^b + K(y^b - H\hat{X}^b) \tag{4}
\]

The analysis (or update)-error covariance \(P^a\) is

\[
P^a = (1 - KH) P^b \tag{5}
\]

Where

\[
K = P^b H^T (H P^b H^T + R)^{-1} \tag{6}
\]

The EnKF forecast and analysis error covariance can be estimated directly using the ensemble (Evensen, 1994; Houtekamer and Mitchell, 1998).

\[
P^a H^T = \frac{1}{n - 1} \sum_{i=1}^{n} (\hat{X}^b - \bar{X}^b)(H \hat{X}^b - H \bar{X}^b)^T \tag{7}
\]

Where \(n\) is the number of ensemble members and the overbar represents the ensemble mean as formula 8.

\[
\bar{X}^b = \frac{1}{n} \sum \hat{X}^b \tag{8}
\]

Expressing the variables as an ensemble mean (denoted by an overbar) and a deviation from the mean (denoted by a prime), the update equations for EnKF may be written as

\[
\bar{X}^a = \bar{X}^b + K(y^b - H\bar{X}^b) \tag{9}
\]

\[
X^a = X^b + K(y^a - HX^b) \tag{10}
\]

Where the prime represents a deviation from the mean, \(\bar{K}\) is the traditional Kalman gain calculated from formula 6 and \(K\) is the gain used to update deviations from the ensemble mean.

Burgers et al. (1998) demonstrated that \(P^a\) is underestimated when observations were not treated as random variables, which could cause the EnKF to reject observations in favor of the ensemble forecast and lead the analysis incrementally further away from its reality, resulting in filter divergence. Furthermore, Whitaker and Hamill (2002) showed that adding random noise to observations further skewed the distribution of \(P^a\), which results in a more erroneous analysis, even though the covariance was increased. Therefore, they suggested using EnSRF, which simplified formula 10 to the following:

\[
X^a = X^b - \bar{K} HX^b \tag{11}
\]

Where

\[
\bar{K} = P^b H^T \left[ (\sqrt{H P^b H^T + R}) \right]^{-1} \times \left[ \sqrt{H P^b H^T + R} + \sqrt{R} \right]^{-1} \tag{12}
\]

For an individual observation,

\[
\bar{K} = \alpha K \tag{13}
\]

\(\alpha = (1 + \sqrt{R (H P^b H^T + R)})^{-1}\)

\(y^a = 0\) means perturbed observations are no longer necessary. This scalar version of EnSRF for the assimilation of a single observation was implemented in the present paper.

2. **Wheat growth model (WheatGrow)**

The WheatGrow model, a process based wheat growth model, was developed from systematic and comprehensive studies on various aspects of the wheat cultivation system, such as the mechanisms and relationships between wheat growth, climate, soil and management technology (Fig. 1). The physiological development time (PDT) serves as the quantitative measurement of the wheat development process. The main functional modules of WheatGrow include phasic development and phenology (Yan et al., 2000), photosynthesis and dry matter production, dry matter distribution and organ formation (Cao et al., 2002), water and nutrient balance (Hu et al., 2004; Zhuang et al., 2004). The genetic characteristics and cultivar ecotype parameters of WheatGrow were adjusted by the trial and error method according to characteristics of wheat cultivars and experimental treatments. These parameters were analyzed to ensure that they could preferentially simulate the dynamic process of wheat growth, development and grain yield. The key model state variables, LAI and LNA, were selected as the coupling parameters for WheatGrow and remote sensing, which could not only express wheat growth conditions, such as canopy coverage characteristics and individual leaf nitrogen content, but were also directly related to the final grain yield. In addition, the data for both of these parameters could be successfully retrieved by remote sensing (Feng et al., 2008, 2009; Zhu et al., 2008).
Fig. 1. Modular structure of the WheatGrow model. The solid and dotted lines denote material and information flow, respectively.

Fig. 2. Integration flowchart of remotely sensed information and the wheat growth model based on EnSRF algorithm and updating method.
3. Process of coupling remote sensing information and the WheatGrow model

The process updating strategy was adopted to achieve the integration of remote sensing information with the WheatGrow model (Fig. 2). During the updating process, remotely sensed LAI and LNA series were integrated into the WheatGrow running process, which is similar to the method of Inoue et al. (1998). The analysis values (i.e., filtering results or update values at k moment, denoted as $X^a$) for LAI and LNA were estimated from those observed by remote sensing (denoted as $y_o$) and background model forecasted (denoted as $X^b$, which is the sum of model simulated values $M_k$ and the random model simulating errors $w_k$) at each remote sensing footprint by simultaneously considering the predictive errors of both remote sensing and the model based on the EnSRF algorithm. The mean value of $X^a$ ($X^a$) was again assimilated into the WheatGrow model to update corresponding state variables, such as dry matter weight and leaf N content, synchronously. Thus, the WheatGrow model could continuously run forward based on these updated variables and achieved the improved final model simulation accuracy for growth parameters and grain yield. In addition, the dimension of the ensemble in this context was determined accordingly as 50 (de Wit et al., 2007). Based on the dataset of Experiments 2 – 4, the mean values and the standard deviation of the WheatGrow simulating error were estimated to be $-0.77 \pm 0.56$ and $1.29 \pm 1.99$ g m$^{-2}$ for LAI and LNA, respectively. Those of the ground-based remote sensing were $-0.53 \pm 0.36$ and $0.84 \pm 0.86$ g m$^{-2}$ for LAI and LNA, while the space-borne remote sensing model predicted $0.63 \pm 1.23$ and $1.63 \pm 2.07$ g m$^{-2}$ for LAI and LNA, respectively. The error covariance matrixes of the model simulation and remote sensing observations were then developed correspondingly.

### Results

#### 1. Field scale validation runs

The main agronomic parameters (LAI, LNA and grain yield) used for testing the integrating technique at different growth stages of Experiments 1 – 4 are summarized in Table 5. In general, the LAI and LNA at different growth stages, and the grain yield of different cultivars, tended to increase with increasing N rate. However, some treatments did not follow this trend. For instance, the LAI and LNA of N7D5 at the booting stage were lower than those of N4D5. The high basic soil fertility in Experiment 1, which led to earlier senescence of the leaves under the plant canopy in response to high N treatments, may have contributed to the insignificant grain yield difference between N4 and N7 in Experiment 1. In addition, the values of LAI and LNA changed to varying degrees during different growth stages due to variations in weather conditions, soil and cultivar types at different eco-
The overall results of LAI, LNA and grain yield analysis indicate that varied N fertilization rates, along with varied soil and weather conditions, affected LAI, LNA and grain yield. The wide range of variation in growth status of wheat can be effectively used to test the integration technique developed in this study.

The retrieved LAI/LNA from Experiments 1 – 4 based on field spectral measurements were used to test the integrating technique of the remote sensing and WheatGrow model based on this algorithm, and different assimilation algorithms (EnSRF and EnKF) were compared (Fig. 3). The results showed that the best LAI/LNA analysis values were obtained when the EnSRF algorithm was adopted, which were more consistent with the measured values compared with those simulated by the WheatGrow model or estimated directly by remote sensing, with RE ± RMSE values of 0.15 ± 0.80 and 0.13 ± 0.82 g m⁻² for LAI and LNA, respectively, despite the fact that there was no evident improvement in prediction performance compared with remote sensing in the middle ranges of LAI and LNA, respectively. Meanwhile, the corresponding values of RMSE and RE for LAI and LNA were 0.24 ± 1.17 and 0.35 ± 1.80 g m⁻² based on the WheatGrow model; these values were 0.21 ± 0.97 and 0.24 ± 1.25 g m⁻² for ground-based remote sensing. Furthermore, the EnKF algorithm produced worse results than EnSRF with the RE and RMSE values of 0.28 ± 1.39 and 0.34 ± 1.70 g m⁻², respectively. Figure 4 showed the LAI and LNA curves of different N treatments and different wheat varieties simulated by the WheatGrow model and updated by the EnSRF analysis values. These values had a better level of consistency with the measured values than those simulated directly by WheatGrow. In addition, the yields simulated by the updated WheatGrow and WheatGrow were compared with the measured yields shown in Figure 5 based on Experiments 1 – 4. The results showed that the updated WheatGrow model based on the EnSRF algorithm could simulate the grain yield accurately with the RE ± RMSE value of 0.03 ± 350.98 kg ha⁻¹; however, those values were 0.04 ± 436.31 kg ha⁻¹ based on the WheatGrow model alone.

Fig. 3. Comparisons of the measured and simulated growth parameters values by WheatGrow (a), remote sensing (b), EnKF (c) and EnSRF (d) based on Experiments 1, 2, 3 and 4.
Fig. 4. Comparisons of the predicted results of LAI (a) and LNA (b) for Ningmai 9, LAI (c) and LNA (d) for Yangmai 12 by WheatGrow model after updating based on EnSRF, WheatGrow and measured values, respectively. N7D5, N1D5, N7D6 and N1D6 indicate different treatments (N1, 0 kg N ha\(^{-1}\); N7, 225 kg N ha\(^{-1}\); D5, 84 kg ha\(^{-1}\) of sowing rate; D6, 84 kg ha\(^{-1}\) of sowing rate).

Fig. 7. Spatial distributions of LAI (a), LNA (b) and grain yield (c) values of wheat simulated by WheatGrow model based assimilating remotely sensed values with EnSRF algorithm in study area.
2. Regional scale validation runs

The remote sensing-WheatGrow integration model based on the EnSRF algorithm was further tested using the data acquired from two-year sampling points on the entire study area with LAI, LNA and yield values estimated by satellite images. The results showed that EnSRF-based analysis values for LAI and LNA at each remote sensing footprint were more consistent with the measured values compared to values that were simulated or estimated directly by the WheatGrow or space-borne remote sensing models, with RE ± RMSE values of 0.19 ± 1.17 and 0.22 ± 1.53 g m\(^{-2}\), respectively, which were generally lower than the RE ± RMSE values of 0.18 ± 1.04 and 0.38 ± 1.84 g m\(^{-2}\), based on the WheatGrow model alone; RE ± RMSE values of 0.30 ± 1.86 and 0.40 ± 2.29 g m\(^{-2}\) were estimated for LAI and LNA based on space-borne remote sensing, respectively. Then, LAI and LNA at the heading stage and grain yield values in the study area were simulated by the updated WheatGrow model based on EnSRF (2009 as example). The simulated LAI and LNA values were consistent with the actual field surveys (Fig. 6), and grain yield values also correlated well with the measured values with the RE ± RMSE values of 0.06 ± 487.6 kg ha\(^{-1}\), which were lower than the values of 0.06 ± 542.0 kg ha\(^{-1}\) based on the WheatGrow model alone. The spatial variability of the predicted results was demonstrated in Figure 7.

Discussion

The ability to assimilate the remote sensing and crop growth models based on both forcing and initialization/parameterization strategies is based on the assumption that remote sensing observations can reflect the objective reality well (Zhao et al., 2005). However, uncertain errors are usually introduced during the treatment process of remote sensing data or by the predicting errors of crop growth parameters based on remote sensing. Meanwhile, the simulating error of crop growth model itself was not considered by both assimilation strategies (forcing and calibration). However, the process updating strategy can solve this issue by taking errors caused by remote sensing and the crop growth model into account simultaneously through data assimilation algorithms, such as EnKF. This improves the simulation and prediction accuracy of the assimilation model based on remote sensing and the specific crop model (Dorigo et al., 2007). Although the EnKF algorithm has been widely used to couple remote sensing and crop models based on a processing updating strategy (de Wit et al., 2007; Curnel et al., 2011; Xiao et al., 2011), its main drawback is that the analysis-error covariance is usually underestimated, which causes EnKF to reject observations in favor of the ensemble forecast, and lead the analysis incrementally further away from reality.

EnSRF, a deterministic method with an improved performance over EnKF in updating model simulations (Das et al., 2008; Han and Li, 2008), was introduced to assimilate remote sensing and the crop growth model in the present paper. The main advantage of EnSRF is that it can obtain the analysis values of the crop model state variables, such as LAI and LNA, at each remote sensing footprint and subsequently update the running process of the crop model without requiring adding additional disturbances to the measurements (Whitaker and Hamill, 2002). Therefore, the assimilation technique by updating LAI/LNA values modeled by the WheatGrow model with those values estimated by remote sensing based on EnSRF was developed and tested at both the field and regional scales. The results showed that the analysis values of LAI and LNA based on EnSRF at each moment were more
consistent with the corresponding actual measured values than values simulated, estimated and analyzed by the WheatGrow model, remote sensing and EnKF, in field scale validation runs. EnSRF significantly improved the prediction performance in the full ranges of LAI and LNA compared with the WheatGrow model and EnKF, which was also true for the low and high extreme ranges of LAI and LNA by remote sensing. However, there was no evident improvement in the middle ranges of LAI and LNA compared with remote sensing. This may be due to the fact that remote sensing provided good estimation accuracy for the LAI and LNA within the middle ranges, and EnSRF overestimated the error of remote sensing observation in the middle ranges, which was derived from the full ranges of the LAI and LNA.

In addition, the combined use of remote sensing-derived biophysical/biochemical state variables and crop models is expected to improve their predictive performance, especially at the regional scale. The major problem with crop growth models may be an oversimplified description of the agricultural system, inaccurate parameterization and uncertainty, and hence a low prediction performance when used at regional scales where model input parameters have to be gathered from scattered point locations, such as weather stations (de Wit et al., 2005). The present study showed that the updated model also performed well in simulating the spatial distributions of winter wheat growth parameters including LAI, LNA and grain yield. This confirmed that the complement of spatial and temporal advantages of the remote sensing and crop growth coupled model could be achieved by this effective assimilation method (Dorigo et al., 2007). These findings also indicated that the EnSRF algorithm was a good deterministic method for assimilating remote sensing and the crop model to improve the predictive accuracy of this crop model and its practicability in regional to continental scales.

However, this method had limitations for determining the errors of model simulation and remote sensing observations, due to the limited number of samples. Therefore, effective techniques for estimating errors, such as the adaptive filter method (Reichle et al., 2008), should be tested in future studies. The adaptive filter method could be used to identify model and observation error variances and provide generally improved assimilation estimates compared with the non-adaptive system, by continually adjusting model and observation error parameters in response to internal diagnostics, although it is currently in the theoretical validation and application research phases at present. However, more emphasis should be concentrated on improving the accuracy of crop modeling and remote sensing, and further extensive validation and improvements should be carried out for different wheat production areas.

**Acknowledgments**

This research was supported by grants from the National 863 High-tech Program (2013AA102301 and 2011AA100703), the National Natural Science Foundation of China (3137135), the Special Fund for Agro-scientific Research in the Public Interest (201303109), the Science and Technology Support Program of Jiangsu (BE2010395, BE2011351 and BE2012302) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), China.

**References**

Burgers, G., van Leeuwen, P.J. and Evensen, G. 1998. Analysis
scheme in the ensemble Kalman filter. *Mon. Wea. Rev.* 126: 1719-1724.

Cao, W.X., Liu, T.M., Luo, W.H., Wang, S.H., Pan, J. and Guo, W.S. 2002. Simulating organ growth in wheat based on the organ-weight fraction concept. *Plant Prod. Sci.* 5: 248-256.

Crow, W. and Wood, E. 2003. The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using Ensemble Kalman filtering: a case study based on ESTAR measurements during SGP97. *Adv. Water Resour.* 26: 137-149.

Curnel, Y., de Wit, A., Duveiller, G. and Defourny, P. 2011. Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS Experiment. *Agric. For. Meteorol.* 151: 1845-1855.

Das, N.N., Mohanty, B., Cosh, M. and Jackson, T. 2008. Modeling and assimilation of root zone soil moisture using remote sensing observations in Walnut Gulch Watershed during SMEX04. *Remote Sens. Environ.* 112: 415-429.

de Wit, A.J.W., Boogaard, H.L. and van Diepen, C.A., 2005. Spatial resolution of precipitation and radiation: the effect on regional crop yield forecasts. *Agric. For. Meteorol.* 135: 156-168.

de Wit, A.J.W. and van Diepen, C.A. 2007. Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts. *Agric. For. Meteorol.* 146: 38-56.

Delgado-Rodriguez, R., Maas, S., Guérif, M. and Baret, F. 1992. Remote sensing and crop production models: present trends. *ISPRS J. Photogramm.* 47: 145-161.

Dorigo, W.A., Zurita-Milla, R., de Wit, A.J., Brazil, J., Singh, R. and Schaepman, M.E. 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. Applied Earth Observation Geoinf.* 9: 165-193.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* 99: 10413-10452.

Feng, W., Yao, X., Zhu, Y., Tian, Y.C. and Cao, W.X. 2008. Monitoring leaf nitrogen status with hyperspectral reflectance in wheat. *Eur. J. Agron.* 28: 394-404.

Feng, W., Zhu, Y., Yao, X., Tian, Y.C. and Cao, W.X. 2009. Monitoring leaf dry weight and leaf area index in wheat with hyperspectral remote sensing. *Chinese J. Plant Ecol.* 33: 34-44*.

Godwin, D., Ritchie, J., Singh, U., Hunt, L. 1990. A user’s guide to CERES wheat-V2.10. Muscle Shoals: International Fertilizer Development Center.

Han, X.J. and Li, X. 2008. Review of the nonlinear filters in the land data assimilation. *Adv. Earth Sci.* 23: 813-820*.

Houtekamer, P. and Mitchell, H. 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.* 126: 796-811.

Hu, J.C., Cao, W.X., Zhang, J.B., Jiang, D. and Feng, J. 2004. Evaluation of hyperspectral data for pasture estimate in the Brazilian Amazon using field and imaging spectrometers. *Remote Sens. Environ.* 112: 1569-1583.

Pan, J., Zhu, Y., Cao, W.X., Dai, T.B. and Jiang, D. 2006. Predicting the protein content of grain in winter wheat with meteorological and genotypic factors. *Plant Prod. Sci.* 9: 323-333.

Pan, J., Zhu, Y. and Cao, W.X. 2007. Modeling plant carbon flow and grain starch accumulation in wheat. *Field Crops Res.* 101: 270-284.

Reichle, R.H., Crow, W.T. and Keppenne, C.L. 2008. An adaptive ensemble Kalman filter for soil moisture data assimilation. [Online]. Available at http://onlinelibrary.wiley.com. Wiley Online Library. doi: 10.1029/2007WR006357. *Water Resour. Res.* 44: W03425.

Shen, Y., Niu, Z., Chen, F. and Wang, C.Y. 2007. Rational consideration about construction land expansion of Changsha urban area in the last ten years. *Geogr. Geoinfo. Sci.* 23: 27-30, 54.

Slater, A. and Clark, M. 2009. Snow data assimilation via an Ensemble Kalman Filter. *J. Hydrometeor.* 7: 478-493.

RSL. 2006. ENVI: Environment for Visualizing Images, Version 4.3. Research Systems Inc., Boulder, CO, USA.

Verhoef, W. 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. *Remote Sens. Environ.* 16: 125-141.

Wallach, D., Goffinet, B., Bergez, J., Debaeke, P., Leenhardt, D. and Aubertot, J. 2001. Parameter estimation for crop models: a new approach and application to a corn model. *Agron. J.* 93: 757-766.

Wang, D.W., Wang, J.D., Liang, S.L. 2010. Retrieving crop leaf area index by assimilation of MODIS data into crop growth model. *Sci. China Earth Sci.* 53: 721-730*.

Whitaker, J.S. and Hamail, T.M. 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.* 130: 1913-1924.

Xiao, Z.Q., Liang, S.L., Wang, J.D., Jiang, B. and Li, X.J. 2011. Real-time retrieval of leaf area index from MODIS time series data. *Remote Sens. Environ.* 115: 97-106.

Yan, M.C., Cao, W.X., Luo, W.H. and Jiang, H.D. 2000. A mechanistic model of phasic and phenological development of wheat. I. Assumption and description of the model. *Chinese J. Appl. Ecol.* 11: 355-359*.

Yang, S.B., Shen, S.H., Li, B.B., Toan, T.L. and He, W. 2008. Rice mapping and monitoring using ENVISAT ASAR data. *IEEE Geosci. Remote Sens. Lett.* 5: 108-112.

Yuan, J.G., Niu, Z. and Wang, X.P. 2009. Atmospheric correction of Hyperion hyperspectral image based on FLAASH. *Spectroscopy Spectral Anal.* 29: 1181-1185*.

Zhao, Y. X., Zhou, J.X. and Liang, S.L. 2005. Methods and application of coupling remote sensing data and crop growth models advance in research. *J. Nat. disasters* 14: 103-109*.

Zhu, Y., Yao, X., Tian, Y.C., Liu, X.J. and Cao, W.X. 2008. Analysis of common canopy vegetation indices for indicating leaf nitrogen accumulations in wheat and rice. *Int. J. Appl. Earth Obs. Geoinf.* 10: 1-10.

Zhuang, H.Y., Cao, W.X., Jiang, S.X. and Wang, Z.G. 2004. Monitoring on nitrogen uptake and partitioning in crops. *Syst. Sci. Comprehend. Stud. Agric.* 20: 5-8*.

* In Chinese with English abstract.