European Journal of Agronomy 101 (2018) 163–173

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data

Yi Chen[a,b], Zhao Zhang[b], Fulu Tao[a,c,*]

[a] Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China

[b] State Key Laboratory of Earth Surface Processes and Resource Ecology, Key Laboratory of Environmental Change and Natural Hazards, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

[c] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:
Winter wheat
Leaf area index (LAI)
Spatial differences
Data assimilation
MCWLA-Wheat model
Yield estimation

ABSTRACT

Crop yield estimation at regional scale using crop model is generally subjected to large uncertainties from insufficient spatial information on heterogeneous growth environment and agronomic management practices. To solve this problem, we assimilated crop phenology and leaf area index (LAI) derived from remote sensing into a crop model (MCWLA-Wheat) to improve its reliability in estimating winter wheat yields at regional scale. Since the LAI magnitude was obviously underestimated however its spatial pattern was relatively well captured by remote sensing, we developed a novel spatial assimilation scheme that assimilated the spatial differences instead of the absolute values of LAI into crop model. Firstly, we retrieved the information of critical development stages of winter wheat from remote sensing data to adjust the simulation of phenology by MCWLA-Wheat model; then the spatial differences of LAI derived from remote sensing were assimilated into the MCWLA-Wheat model using a kind of constant gain Kalman Filter algorithm to improve the ability of the model in estimating winter wheat LAI and yields at regional scale in the North China Plain. This assimilation scheme extracted effective information from remote sensing LAI and meanwhile abandoned the information with obvious errors, ensuring that the assimilation variables could be close to the reality. It avoids the requirement for correction of the LAI derived from remote sensing using other high-quality ancillary data from field measurements. Using this assimilation scheme, the performance of crop model improved substantially. It successfully produced more accurate yield estimates at regional scale during the period of 2001–2008 (mean R^2 = 0.42, RMSE = 737/ha) than those without assimilation (mean R^2 = 0.26, RMSE = 1012 kg/ha) and those directly assimilating the absolute LAI values derived from remote sensing (mean R^2 = 0.30, RMSE = 1257/ha). Our findings demonstrated a reliable and promising assimilation scheme for improving yield estimation of crop model at regional scale with low data requirement.

1. Introduction

The importance of crop growth monitoring and yields forecasting at a large scale has been increasingly highlighted in recent years because accurate information on crop growth and yields is essential for timely coping with climate risk, assessing national food security, and developing suitable food trade strategies (Macdonald and Hall, 1980; Hutchinson, 1991; Lobell et al., 2003; Franch et al., 2015). In recent decades, the development of crop models has provided powerful tools to dynamically simulate crop growth, development and grain formation process, and to predict crop yields (Launay and Guerif, 2005). With the representations of key processes of crop growth and productivity, and the detailed input data, crop models perform generally well in simulating crop growth and productivity at field scale (Hadria et al., 2010; Jégo et al., 2012). However, the applications of crop model at a large scale generally confronted with insufficient spatial information on heterogeneous growth environment and agronomic management practices. (Curnel et al., 2011). Limitations in data availability and quality have restricted the ability of crop models in precisely quantifying the spatial heterogeneity of environment and crop growth, and consequently affect the accuracy of yield estimation at regional scale, particularly for the simulation of crop canopy development and soil

* Corresponding author at: Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.

E-mail address: taofl@igsnrr.ac.cn (F. Tao).

https://doi.org/10.1016/j.eja.2018.09.006
Received 9 December 2017; Received in revised form 14 September 2018; Accepted 15 September 2018
1161-0301/

moisture content (de Wit and van Diepen, 2007; Mignolet et al., 2007; Dente et al., 2008). Therefore, improving the model performance at regional scale by introducing sufficient spatial heterogeneous information on crop growth has been of key concern (Challinor et al., 2004; Tao et al., 2009).

The development of remote sensing technology provides a potential to solve this problem since it provides extensive observations on spatial characteristics of land surface dynamics (Doraiswamy et al., 2003; Launay and Guerif, 2005; Reichle, 2008; Liang and Qin, 2008; Dong et al., 2015). Assimilating the valuable information from remote sensing into crop model simulation process could help to identify the cultivation area and the values of state variables during crop growth. This makes it possible to re-estimate missing information on model parameters and complement simulation results under scenarios that have not been considered in the model (Batchelor et al., 2002; Luo et al., 2011; de Wit et al., 2012). Along this line, studies have tried to alleviate the effects of data shortage on model performance through local recalibrating the model parameters, adjusting initial status, or modifying state variables (Curnel et al., 2011). Multiple kinds of variables such as the leaf area index (LAI) (Dente et al., 2008), soil moisture (de Wit and van Diepen, 2007) and normalized difference vegetation index (NDVI) (Fang et al., 2011) have been selected as assimilation variables in previous studies according to the model characteristics. Studies with different degrees of successes have demonstrated the positive effects of these assimilation strategies (Quaife et al., 2008; de Wit et al., 2012; Ma et al., 2013). Among the variables used in assimilation, LAI is one of the most widely used variables because of its critical roles in indicating crop growth status, reflecting the comprehensive influences of growth environment and managements, and determining the estimation of biomass and yield in crop model (Curnel et al., 2011; Wang et al., 2013; Zhao et al., 2013; Li et al., 2014a,b; Huang et al., 2015, 2016). For example, Ines et al. (2013) used the ensemble Kalman Filter (EnKF) algorithm to assimilate MODIS LAI data into the DSSAT model to correct the simulated LAI values. The methods generated improved estimates on maize yield with reduction in RMSE by 500 kg/ha comparing to the results without assimilation. Vazifedoust et al. (2009) adopted an assimilation framework using the constant gain Kalman Filter algorithm and MODIS LAI data. The assimilation results reduced the biases in wheat yield estimation from 4 ~ 39% to less than 10%. Dente et al. (2008) assimilated remote sensing LAI data with multiple resolutions into CERES-Wheat model to reinitialize the model parameters (sowing date, soil wilting point and field capacity) and finally reduced the model errors from 460 kg/ha to 360 ~ 420 kg/ha. Ma et al. (2013) assimilated the bias-corrected MODIS LAI into WOFOST model to reinitialize the emergence date, initial biomass and initial available soil water, reducing the RMSE of yield estimates from 983 kg/ha to 414 ~ 667 kg/ha. Additionally, some studies also focused on the correction of phenology simulation through re-estimating the development parameters in crop model based on the phenological information derived from remote sensing LAI data. As an example, Xu et al. (2011) showed that the optimization of phenology simulation reduced RMSE of yield estimates by approximately 100 kg/ha.

Despite the positive effects of assimilation with remote sensing LAI, the application of assimilation always encountered with limitations from data availability and quality. Remote sensing data with relatively high resolution could provide more detailed information of land surface but might be limited by the scale, revisit period and availability of cloud-free image (Huang et al., 2016). Therefore the relatively coarse resolution data (e.g., LAI data based on MODIS images) is still dominantly used as data source for assimilation because of its advantages of short revisit. The reliability of remote sensing data might be weakened by the scale mismatch between remote sensing data and the agricultural landscape, particularly for coarse resolution data (de Wit and van Diepen, 2007; Curnel et al., 2011; Huang et al., 2016). Previous studies have revealed that the inaccurate LAI in remote sensing data would force the crop model to produce unrealistic estimation of LAI during data assimilation (Yang et al., 2010; Fang et al., 2012; Huang et al., 2016). Some researchers tried to correct the biased values in remote sensing data through introducing high-quality ancillary data from field measurement (Huang et al., 2015, 2016). Such corrections contributed improved assimilation results, but simultaneously brought requirements for large amount of field measurement with high spatio-temporal density (Huang et al., 2015). The huge money-, labor- and time-consuming for collecting and treating the necessary data consequently restricted studies within small spatial scale and short period (Curnel et al., 2011), making the model improvement using data assimilation over a large area remain a great challenge. Therefore, besides the efforts on obtaining more accurate LAI values, we should also focus on whether available remote sensing products could be used more effectively in data assimilation. Considering that the absolute values in coarse resolution remote sensing data may be biased in representing the actual crop growth status, thus these absolute values should not be directly applied in assimilation. Nevertheless, these remote sensing data could still depict the relative spatial differences of land surface characteristics over a large area (Garrigues et al., 2008; Duveiller et al., 2011, 2012). Moreover, recent studies have focused on improving the retrieve algorithm to make remote sensing products better represent the reality, resulting in more obvious linear relationship between remote sensing products and field measurements, providing promising datasets to depict the spatial heterogeneity of land surface (Fang et al., 2014; Li et al., 2014a,b; Xu et al., 2016). Thus, focusing on the spatial difference instead of the absolute value in remote sensing products should be an alternative scheme for developing assimilation and alleviate the impacts of biased absolute values.

In this study, unlike many previous studies that assimilated the absolute values of LAI at each grid cell, we developed a novel spatial assimilation scheme that assimilates the spatial differences of LAI into crop model. This assimilation scheme was used to assimilate an improved MODIS-based LAI product (GLASS LAI) (Xiao et al., 2014) into MCWLA-Wheat model to estimate winter wheat yields at regional scale in the North China Plain. The overall objective of this study is to demonstrate the new scheme in improving the ability of MCWLA-Wheat (Tao et al., 2009) in estimating winter wheat yields over a large area without requiring for the correction of LAI products using other high-quality ancillary data from field measurements. This will lay good foundation for regional yield estimation using the new assimilation scheme.

2. Materials and methods

2.1. Study area

The study area was the North China Plain, which is the most important production region of winter wheat in China (Fig. 1). The study area covered 139 counties in four provinces (Hebei, Shandong, Henan, and Anhui). This region extends from 32.1 °N to 38.8 °N and from 111.5 °E to 118.2 °E, with a total area of 160,731 km^2 (Fig. 1), covering most of the main cultivation areas of winter wheat in the North China Plain. This region is dominated by alluvial plains and has a typical temperate monsoon climate, with mean temperature of 15.6 °C and annual precipitation of 834.5 mm. Monthly mean temperature ranges from 0.9 °C in January to 27.9 °C in July. The dominant crop system in the study region is winter wheat - summer maize rotation system. Winter wheat is generally sown at the beginning of October and harvested at the end of May or the beginning of June in the following year. Then the maize is sown in the middle of June and harvested at the end of September (Fig. S1). The crops in this region were generally irrigated and sufficient fertilization was applied for winter wheat.

2.2. Description of the MCWLA-Wheat model

The MCWLA-Wheat model is a process-based crop growth model