Agricultural and Forest Meteorology 247 (2017) 207–220

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches

Zhenong Jin*, George Azzari, David B. Lobell

Department of Earth System Science and Center on Food Security and the Environment, Stanford University, Stanford, CA, 94305, USA

A R T I C L E I N F O

Keywords:
Yield estimation
Maize
Landsat
Google earth engine

A B S T R A C T

Fast, accurate and inexpensive estimates of crop yields at the field scale are useful for many applications. Based on the Google Earth Engine (GEE) platform, we recently developed a Scalable satellite-based Crop Yield Mapper (SCYM) that integrates crop simulations with satellite imagery and gridded weather data to generate 30 m resolution yield estimates for multiple crops in different regions. Existing versions of SCYM typically capture one-third to half of the variation in reported county-scale yields. Using rainfed maize in the US Midwest as an example, this study tested multiple approaches for improving SCYM's accuracy, including (i) calibrating the phenology parameters of the crop model (APSIM) used to generate training samples for SCYM; (ii) using an ensemble of three crop models (APSIM-Maize, CERES-Maize, and Hybrid-Maize) instead of a single model; (iii) using simulated biomass from the crop models instead of simulated yields to train SCYM, with the former assuming a constant harvest index (HI). Results show substantial improvement in performance, as assessed using reported county yields by USDA-NASS, both from calibrating APSIM phenology parameters and from training SCYM on simulated biomass rather than yields. Using a multi-model ensemble further improves SCYM, although the benefit is limited. The proposed preferred version of SCYM on average captures 75% of the yield variation for 2001–2015 in the 3I states (i.e. Illinois, Indiana and Iowa) where SCYM is trained, with RMSE typically less than 1 t/ha, and explains 41% to 83% of multi-year yield variations when tested across nine Midwestern US states for 2008–2015. This level of accuracy is particularly notable given that only data from 2014 were used to calibrate phenology parameters. The yield estimates for multiple years in multiple states utilized 1184 Landsat tiles, but could be completed in about 2 h per year by using the GEE platform. All approaches tested in this study do not require any site-specific measurements, and thus can be readily extended to other regions and crops.

1. Introduction

Global demand for agricultural crops for use as food, feed and bioenergy will continue to grow in the coming decades, thereby increasing competition for land and water (Alexandratos and Bruinsma, 2012). Agricultural production thus has to be intensified through more efficient farm management, which in turn will require better knowledge of crop yield variation across a range of spatial scales and over time. Yield estimates at field level, in particular, can be useful for investigating the spatially variable causes of yield gaps (Lobell, 2013); estimates over multiple years can be used to generating the spatial variability of expected field productivity, and hence the required information for varying management inputs (Diker et al., 2004), as well as for insurance or land markets (Lobell et al., 2015).

Along with the advance of earth-observing using satellites, a number of remote sensing approaches have been proposed to predict yield across a variety of crops and geographic span (see reviews by (Atzberger, 2013; Lobell, 2013; Mulla, 2013)), yet substantially more examples focus on providing estimates at regional scale rather than for individual field (Lobell et al., 2015). Accurate, scalable and cost-effective tools of field-level yield mapping remains a challenge, because of (i) the limited availability of satellite data with fine spatial, spectral and temporal resolution, (ii) the substantial cost associated with fetching and processing massive satellite data, and (iii) the lack of scalable approaches to obtain yields from the imagery. For the latter, existing approaches often rely on the construction of empirical relationships between satellite observed vegetation indices (VIs) and yields (Franch et al., 2015; Panda et al., 2010; Sakamoto et al., 2014), thus requiring new ground-measurement for recalibrating the empirical model before applying to new areas or other years. Researchers have also used remote sensing measurements to constrain the inputs or parameters of crop models that can better accommodate new settings (Hank et al., 2015; Huang et al., 2015; Ines et al., 2013; Lobell, 2013; Machwitz et al., 2014). However, the computational cost and input data

* Corresponding author at: 473 Via Ortega, Stanford, CA 94305, USA.
E-mail address: jinzn@stanford.edu (Z. Jin).

http://dx.doi.org/10.1016/j.agrformet.2017.08.001
Received 14 April 2017; Received in revised form 18 July 2017; Accepted 2 August 2017

required for these approaches hinder their application on a large scale.

In a previous study (Lobell et al., 2015), we developed an approach called SCYM (a scalable satellite-based crop yield mapper) that bypasses the traditional crop model calibration procedure that requires a range of site-specific ground data. The SCYM approach first generates several hundred pseudo-observations of daily crop attributes, such as leaf area index (LAI), as well as final biomass and yield, from crop simulations that span a realistic range of growing conditions by varying soil, weather, cultivar, fertilizer applications, and sowing date. Next, simulated daily observations are used to train date-specific statistical models that relate end-season yield to weather and satellite-observable VIs (derived from LAI). Finally, these statistical models can be applied to remote sensing (e.g. Landsat and the Moderate Resolution Imaging Spectroradiometer (MODIS)) and gridded weather data within the Google Earth Engine (GEE), a cloud-based platform that can efficiently access and process large volumes of data, to generate yield estimates on a pixel-by-pixel basis. Applications of SCYM to maize and soybean in the US Midwest show that this approach on average captured one-third of the yield variation at farm level in all state-years (Lobell et al., 2015), and nearly 50% of the variation when compared with the USDA National Agricultural Statistics Service (NASS) reported county-level yield data (Azzari et al., 2017). Moreover, estimates of spatial variability within counties agree well with estimates derived from field-level insurance data (Lobell and Azzari, 2017). These initial results are fairly encouraging, given that all data required to implement the SCYM approach can be easily obtained and that "ground-truth" yields have not been used in any way to do the calibration (Lobell et al., 2015).

Meanwhile, many avenues exist for improving SCYM's accuracy (Lobell et al., 2015), among which reducing the uncertainty and bias introduced by crop model simulations is the first thing to consider. In crop simulations, the seasonal progress of LAI determines the solar radiation intercepted by the canopy, which drives photosynthesis and dry matter production, and hence the final grain yield (Huang et al., 2015). One underpinning assumption of the SCYM approach is that the crop model can well reproduce the seasonal curve of LAI and yield corresponding to growing conditions over a realistic range. This assumption, however, is often violated as shown in several crop model intercomparison studies (Bassu et al., 2014; Battisti et al., 2017; Martre et al., 2015), because of uncertainties in model structures, initial conditions and input parameters (Asseng et al., 2013). For example, many of the widely used crop models still employ physiological parameters and response functions that were derived from calibrations using old varieties dating back to 1980s, and thus mismatch the growth of contemporary varieties (Rötter et al., 2011). This issue also applies to the APSIM (Agricultural Production Systems Simulator) model (Holzworth et al., 2014) employed by the earlier version of SCYM.

A second potential issue is that relying on a single crop model, even a well-calibrated one, might perform worse than using a combination of independent models. Several recent crop model intercomparison studies suggested that the ensemble mean of multiple models is often closer to the field measurement than any single model, even without model calibration (Bassu et al., 2014; Battisti et al., 2017; Martre et al., 2015). Thus incorporating simulations from multiple crop models rather than APSIM alone could potentially be beneficial as well.

A third lesson from the crop modeling literature is that mechanistic simulations of grain formation and growth within most models often underperform simpler approaches that predict yield by multiplying simulated biomass by a constant harvest index (HI). For example, the simpler biomass × HI method exhibited better agreement with NASS county-level yield statistics for U.S. maize than did yield simulations from the same models (Jin et al., 2016b). A possible explanation is that mechanistic methods for grain development need to estimate a substantial number of parameters beyond those needed for biomass simulation, and that values for these parameters have not been sufficiently calibrated.

The objective of this study is to improve the accuracy of SCYM through three scalable ways (i.e. not requiring additional ground-level measures), with application to US rainfed maize as an example. Specifically, we investigate improvements by: (i) calibrating APSIM phenology parameters based on high-frequency Landsat-8 observations; (ii) using multiple crop model structures to simulate LAI and yield instead of APSIM alone; (iii) using model simulations of biomass rather than yields to train SCYM (i.e., first simulating biomass and then multiplying it with a constant HI). In all cases, SCYM was trained using simulations for sites within Illinois, Indiana, and Iowa (3I states hereafter), and improvements were measured based on agreement with NASS reported county-level yield data across a broader region comprising nine Corn Belt states (Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, South Dakota and Wisconsin). These states represent the largest producers of rainfed maize, and together contribute over 65% of total national production (NASS, 2017).

2. Methods

The SCYM approach for yield estimation was documented in details in our recent work (Lobell et al., 2015; Azzari et al., 2017), and a schematic overview of this approach along with different avenues we have tested in this study to improve the existing SCYM was given in Fig. 1. Here, we mainly focused on describing modifications that were made in this study to improve the existing SCYM approach.

2.1. Phenology module calibration

2.1.1. Obtaining high-frequency LAI observations from landsat-8

Obtaining high frequency (typically weekly or bi-weekly) time series of LAI observations over the growing season is a prerequisite for calibrating the crop phenology module. LAI can be estimated from various remotely sensed VIs (Nguy-Robertson et al., 2012). Following Lobell et al. (2015), we estimated LAI from the green chlorophyll vegetation index (GCVI) based on an empirical relationship that was derived from a large sample of field measurements (Nguy-Robertson et al., 2012):

$$LAI = (GCVI - 0.93)/1.4^{0.97}$$

GCVI is defined as (Gitelson et al., 2003):

$$GCVI = \frac{\rho_{NIR}}{\rho_{GRN}} - 1$$

in which ρ_{NIR} and ρ_{GRN} are the near-infrared and green wavelength reflectance, respectively.

We used Google Earth Engine (GEE) to generate GCVI images from Landsat 8 Operational Land Imager (OLI) Surface Reflectance (SR) images acquired over the US Corn Belt in 2014. The GEE team routinely ingests all SR images in the platform's servers directly from the United States Geological Survey (USGS), which handles the satellite's data acquisition and preprocessing. The USGS generates SR images using the LEDAPS algorithm (Masek et al., 2006) to correct for atmospheric effects and computes a quality band that can be used to remove cloud- or shadow-contaminated pixels.

Due to the 16-day revisit time of Landsat satellites and frequent cloud cover, most locations had a limited number of non-contaminated observations during the growing season (Fig. 2), making it difficult to resolve a complete phenological curve. To overcome the limited number of non-contaminated satellite observations, existing studies often used data fusion techniques to generate synthetic observations of sufficient temporal resolution at the finer spatial scale by combining the Landsat data with high temporal frequency yet coarse resolution data (e.g. MODIS) (Amorós-López et al., 2013; Gao et al., 2017; Huang et al., 2015). Despite these examples, data fusion methods require assumptions about how Landsat pixels vary with similar neighboring MODIS pixels (Gao et al., 2017), which can be problematic in agricultural regions where fields are often sown on different dates. Furthermore, to