Science of the Total Environment 650 (2019) 1707–1721

Contents lists available at ScienceDirect

Science of the Total Environment

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

Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean

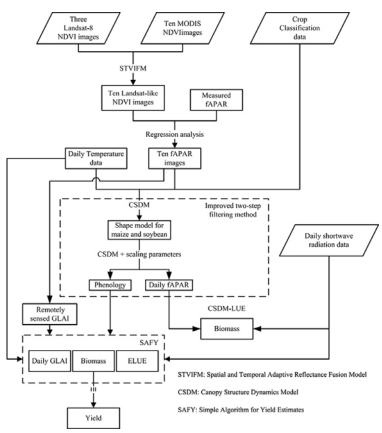
Chunhua Liao [a,*], Jinfei Wang [a], Taifeng Dong [b], Jiali Shang [b], Jiangui Liu [b], Yang Song [a]

[a] *Department of Geography, The University of Western Ontario, London N6A 5C2, ON, Canada*
[b] *Science and Technology Branch, Agriculture and Agri-Food Canada, Ottawa K1A 0C6, ON, Canada*

HIGHLIGHTS

- Improves the Two-Step Filtering method for phenology detection
- Calibrates the Simple Algorithm for Yield estimates model for corn and soybean
- Estimates the biomass and yield accurately at a subfield scale
- A good correlation is found between effective light use efficiency and *f*APAR$_{max}$.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 26 June 2018
Received in revised form 20 September 2018
Accepted 23 September 2018
Available online 26 September 2018

Editor: Ralf Ludwig

Keywords:
Crop yield model
Phenology
Corn
Soybean
Fusion
SAFY

ABSTRACT

The Simple Algorithm for Yield estimates (SAFY) is a crop yield model that simulates crop growth and biomass accumulation at a daily time step. Parameters in the SAFY model can be determined from literature, in situ measurements, or optical remote sensing data through data assimilation. For effective determination of parameters, optical remote sensing data need to be acquired at high spatial and high temporal resolutions. However, this is challenging due to interference of cloud cover and rather long revisiting cycles of high resolution satellite sensors. Spatio-temporal fusion of multi-source remote sensing data may represent a feasible solution. Here, crop phenology-related parameters in the SAFY model were derived using an improved Two-Step Filtering (TSF) model from remote sensing data generated through spatio-temporal fusion of Landsat-8 and Moderate Resolution Imaging Spectroradiometer (MODIS) data. Remaining parameters were determined through an optimization procedure using the same dataset. The SAFY model was then used for dry aboveground biomass and yield estimation at a subfield scale for corn (*Zea mays*) and soybean (*Glycine* max). The results show that the improved TSF method is able to determine crop phenology stages with an error of <5 days. After calibration, the SAFY model can reproduce daily Green Leaf Area Index (GLAI) effectively throughout the growing season and estimate crop biomass and yield accurately at a subfield scale using three Landsat-8 and 10 MODIS images acquired for the season. This approach improves the accuracy of biomass estimation by about 4% in relative Root Mean Square Error (RRMSE), compared with the SAFY model without forcing the phenology-related parameters. The RMSE of yield

* Corresponding author.
E-mail address: cliao33@uwo.ca (C. Liao).

https://doi.org/10.1016/j.scitotenv.2018.09.308

estimation is 146.33 g/m^2 for corn and 82.86 g/m^2 for soybean. The proposed framework is applicable for local-scale or field-scale phenology detection and yield estimation.

1. Introduction

To ensure food security for the growing world population, crop growth needs to be monitored and crop production needs to be estimated. Crop phenology is an indicator to understand agricultural response to environmental conditions and is essential to estimate crop yield (Sakamoto et al., 2010). Crop production related statistics are also key indicators to understand seasonal ecosystem carbon exchange and environmental impacts (Marshall and Thenkabail, 2015; Paruelo et al., 2000).

Remote sensing provides spatially continuous information on crop growth conditions, hence contributes to an improved monitoring of crop production and management (Claverie et al., 2012). Statistical models (Tucker and Sellers, 1986) and process-based models (Jamieson et al., 1998) have been applied to estimate crop yield using remotely sensed data. Statistical models typically attempt to build relationships between remotely sensed vegetation indices and in situ measurements. However, they may be applicable only to specific growth stages or specific regions (Cheng et al., 2016), and the accuracy may vary with environmental conditions (Kuwata and Shibasaki, 2016). Process-based models, such as the AquaCrop (Steduto et al., 2009), CERES-Maize (Jones et al., 1986), STICS (Brisson et al., 2003) and WOFOST (van Dipen et al., 1989), were originally developed to simulate the key physical and physiological processes of the plant-soil-atmosphere system to obtain daily dry aboveground biomass and final grain yield (Marshall and Thenkabail, 2015; Sellers, 1985). The process-based models need a large set of agro-environment variables and model parameters, which may not be available or difficult to obtain over large areas (Battude et al., 2016; Betbeder et al., 2016; Claverie et al., 2012). The Simple Algorithm for Yield estimates (SAFY) model (Duchemin et al., 2008) through combination of Monteith's light use efficiency (LUE) theory (Monteith, 1972) and Maas's leaf partitioning function (Maas, 1990), simulates daily Green Leaf Area Index (GLAI) and Dry Aboveground Mass (DAM) (i.e., dry aboveground biomass) from the date of emergence. The parameters in the SAFY model can be found in literature, obtained through in situ measurements, derived from remote sensing data or determined through optimization. It has been demonstrated that the SAFY model could estimate crop biomass effectively when model parameters were determined from remotely sensed phenological dates and optimized by time-series GLAI derived from high spatial and temporal resolution data (Claverie et al., 2012; Battude et al., 2016; Dong et al., 2016; Betbeder et al., 2016).

Some of the parameters in SAFY are linked with crop phenology information, which can be detected from time-series remote sensing data. Traditional remote sensing-based phenology models define a few typical phenology stages according to the growth curve represented by time-series of remotely sensed vegetation indices. For example, the dates when the Normalized Difference Vegetation Index (NDVI) is greater than a specific threshold are considered key phenological stages (Sakamoto et al., 2005), or the inflection points (minimum/maximum value of first derivative) of the NDVI curve are considered as the start of season (SOS) and end of season (EOS) (Jeong et al., 2011). However, phenology detection at a subfield scale using such methods is sensitive to data noise induced by atmospheric constituents (Sakamoto et al., 2010). Optimization of the SAFY model parameters at a subfield scale requires remote sensing data at high spatial and temporal resolution. High spatial resolution satellite data, such as Formosat-2, SPOT-4, and Deimos-1, are costly, whereas free high spatial resolution data, such as Landsat and Sentinel-2, may be unavailable on important dates due to low temporal resolution and frequent cloud contamination over the study site. Therefore, a research gap exists concerning how a sufficient number of high spatial-temporal resolution remote sensing data can be provided for calibration of process-based yield models.

The objective of this study is to propose a strategy to estimate subfield-scale crop phenology, crop biomass and yield based on the SAFY model calibrated with a remote sensing dataset generated through the spatio-temporal fusion of Landsat-8 and MODIS images. First, a recently proposed spatio-temporal vegetation index image fusion method (STVIFM) developed by Liao et al. (2017) was used to generate a NDVI time series with high spatial and-temporal resolution. Then the fraction of Absorbed Photosynthetically Active Radiation (APAR) absorbed by green canopy (*f*APAR) was derived from the NDVI, and the Two-Step Filtering (TSF) method proposed by Sakamoto et al. (2010) was improved by using daily *f*APAR simulated by the Canopy Structure Dynamics Model (CSDM), to derive phenological stages. Lastly, the phenology information was linked to the parameters in the SAFY model, and GLAI derived from the remote sensing data was used to optimize the SAFY model in order to estimate pixel level biomass and effective light use efficiency (ELUE), defined as equivalent LUE under all environmental stresses excluding temperature stress.

2. Materials

2.1. Study area

The study area is located in the Mixedwood Plains Ecozone in Southwestern Ontario, characterized by abundant water supply and a relatively mild climate during the growing season but harsh winters. The area has productive soils for agriculture and a longer growing season than most of the country. The common cropping practice in this region is one harvest per year for annual field crops. The dominant crops are winter wheat, corn and soybean. Generally, the winter wheat is seeded in October of the previous year and harvested in July, while the corn and soybean are seeded in May and harvested in September or October. The study area is about 14 km by 8 km, near the city of London, Ontario (Fig. 1).

2.2. Field data collection

Field work was conducted weekly from 23 May to 21 September in 2015. Field data, including digital hemispherical photos (DHP), crop phenology, and crop height, were collected for a total of 27 soybean sample sites and 6 corn sample sites each time. For each sample site, 7 photographs were taken along one transect and then another 7 photographs along a parallel transect 1-m apart (Shang et al., 2014). Effective LAI and *f*APAR were derived from the photographs using the CAN-EYE software (Weiss and Baret, 2017).

Crop type information was also surveyed in September 2015. Crop biomass was measured on 25 September and 2 October using a destructive method in relatively homogeneous locations for soybean (17 samples) and corn (15 samples). For corn, 5 plants were randomly selected and manually harvested within a 5 m by 5 m area around each sampling site. The number of corn plants was counted within this area to determine corn plant density. For soybean, plants were harvested from two blocks of 0.25-m by 0.25-m within a 5 m by 5 m area. The number of soybean plants was counted within this area to determine soybean plant density. The harvested plants were placed in large plastic bags and transferred back to the lab. The plants were separated into stems, leaves and seeds and were weighed separately to obtain the fresh