Crop-Growth Driven Forward-Modeling of Sentinel-1 Observables Using Machine-Learning

Nikaein, Tina; Kummer, Vineet; Steele-Dunne, Susan; Lopez Dekker, Paco

DOI
10.1109/IGARSS46834.2022.9883154

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium

CROP-GROWTH DRIVEN FORWARD-MODELING OF SENTINEL-1 OBSERVABLES USING MACHINE-LEARNING

Tina Nikaein, Vineet Kummer, Susan Steele-Dunne and Paco Lopez-Dekker

Geoscience and Remote Sensing, Delft University of Technology, 2628 CN Delft, The Netherlands

IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-2792-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/IGARSS46834.2022.9883154

ABSTRACT

This paper presents an approach to implement a forward model for Sentinel-1 co-pol and cross-pol backscatter and coherence using crop bio-geophysical parameters namely leaf area index, biomass, canopy height, soil moisture and root zone moisture as inputs for the maize. These required input parameters are generated using Decision Support System for Agrotechnology Transfer (DSSAT), one of the state-of-the-art crop growth models. The predicted SAR signal is generated using Support Vector Regression (SVR) over all the maize fields in an agricultural region, Flevoland, Netherlands. The correlation between simulated signal and observed signal is evaluated.

***Index Terms*—** Crop, DSSAT, Sentinel-1, SAR, simulation, forward-model

1. INTRODUCTION

In general, crop models are simplified representations of the vegetation in real world. Models help us understand the dynamic interaction between environment variables and crop development. Crop growth models can be classified into two categories: statistical and dynamic models. The applicability of statistical models is limited to cases in which the operating conditions are consistent with those used during the development of the model. Dynamic models need more input parameters but have the advantage of being able to simulate the evolution of the crop, and in particular its yield, in situations that have not been previously observed [1]. Crop models play an important role for sustainable management, irrigation, fertilization, etc. The accuracy of the input data determines the certainty of the model. In this paper, decision support system for agrotechnology transfer (DSSAT) that is a dynamic model, has been used to estimate a number of variables, including leaf area index (LAI), dry biomass, soil moisture and canopy height.

Remote sensing satellites provide the potential to monitor vegetated areas and soil conditions at a range of temporal and spatial resolutions. Several studies have shown value of combining crop growth models and remote sensing data. For example, the assimilating soil water index (SWI) derived from coarse resolution satellite into crop model using Ensemble Kalman filter (EnKF) has been studied in [2]. Also, optical and radar data assimilation into crop growth model to improve the simulated parameters were explored in [3].

The objective of this study is to have a forward operator in order to simulate SAR observations (backscatter and coherence) from vegetation and soil characteristics provided by a crop growth model. The vegetation characteristics such as LAI, canopy height, biomass and soil properties such as surface and root zone moisture simulated by DSSAT model are used as input to support vector regression (SVR) to estimate the SAR observables.

2. STUDY AREA

Maize is a crucial fodder crop in the Dutch agricultural landscape and accounts for approximately 10% of the total agricultural land. Maize parcels in the Flevopolder region of the Netherlands are selected for this study. Flevopolder region is a reclaimed flat land dominated by clay loam soil type with a high water holding capacity.Typical summer crop cultivation season starts with the field preparation activities in March, sowing in between mid-April to mid-May. Harvesting takes place between mid-September and early October. Maize parcel information of the region is attributed from Basisregistratie Gewaspercelen (BRP) of the Netherlands. During the 2017, maize was grown in around 539 parcels of the study area. Figure 1 shows all agricultural fields in Flevoland with maize fields marked red.

3. DATA AND METHODS

3.1. Input data for DSSAT

The minimum required data to execute the DSSAT model is weather, soil and crop field management practices. This study uses CERES (Crop-Environment Resource Synthesis)-Maize crop module of DSSAT v4.7 to simulate the maize growth variables from the sowing to the harvest period in a daily step. The standard weather inputs, namely daily solar radiation (SRAD), precipitation, and maximum and minimum air temperature, were prepared using Royal Dutch Meteorological Institute (KNMI) weather data. KNMI collects meteorological information from 50 weather stations over the Nether-

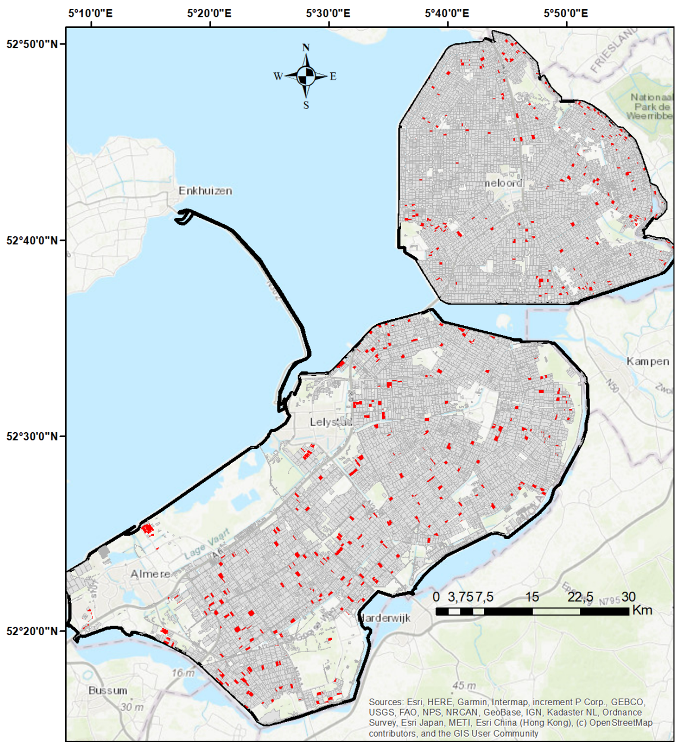

Fig. 1. Spatial distribution of maize parcels in the Flevoland province (Black polygon) of the Netherlands. Red polygons represent maize parcels. Gray polygons are other crop type parcels.

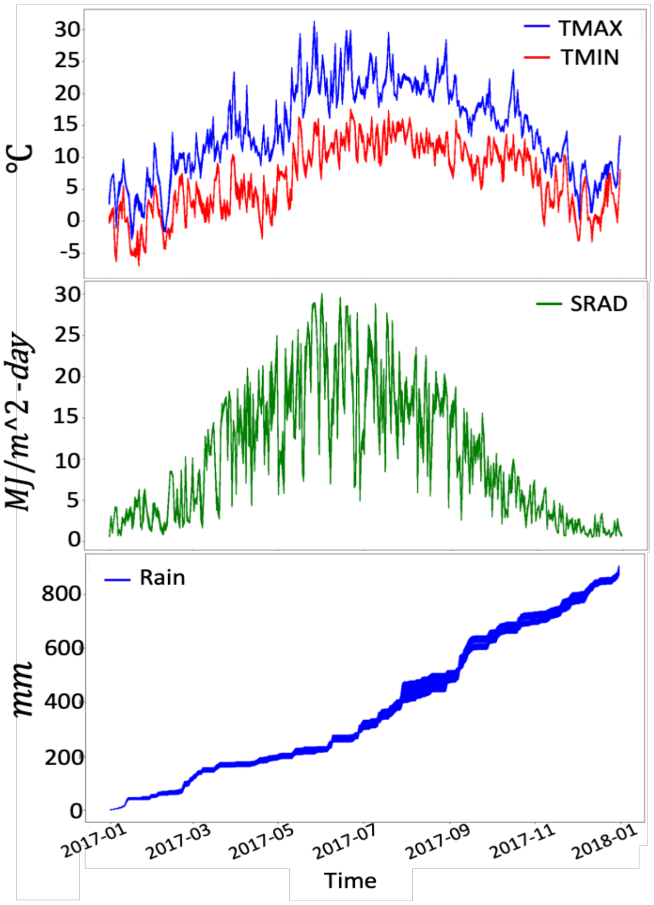

Fig. 2. DSSAT input data. (Top) Maximum and minimum of temperature, (Middle) Solar radiation and (bottom) cumulative rainfall.

lands. Two weather stations are located in the studied area which cover north and south part of the province. Inverse Distance Weighting (IDW) spatial interpolation was applied to extract parcel-level weather information. Figure 2 shows the time series of weather data ingested into DSSAT simulation.

Gridded soil data were obtained at 250-m resolution from ISRIC's global Soil Information System (SoilGrids) [4]. Three soil layers are considered, 0-5, 5-15 and 15-30 cm. Several required soil data that was not available in SoilGrids250m, was obtained from HarvestChoise HC27 which includes 27 different soil profiles that are generated based on only three criteria: soil texture, soil organic carbon content and rooting depth [5]. In order to provide information about soil water movement, soil hydraulic properties need to be estimated. The soil hydraulic parameters are computed by pedo-transfer functions in [6].

Crop management inputs such as planting, emergence and harvest information are provided based on in-situ information collected during the field campaign. In the Flevopolder region, silage maize is typically grown under rainfed conditions. The default model values are considered for the remaining crop management inputs such as tillage, fertilizer, chemical applications and organic amendments.

3.2. Sentinel-1 SAR data

In this study, parcel-level Sentinel-1 SAR interferometric wide (IW) swath mode observations from Agricultural SandboxNL [7] database were used. This database consists of spatially averaged parcel-level backscatter (VV and VH) and interferometric coherence values over the Netherlands. The backscatter and interferometric coherence values are extracted after applying standard pre-processing on Ground Range Detected (GRD) and Single-look Complex (SLC) data, respectively. 6-day interval Sentinel-1 observations from relative orbit no-88, which is an ascending pass, are used.

3.3. Support Vector Regression

In order to map the vegetation states to the observables, we use a supervised machine learning SVR [8] as forward operator. SVR has been used in studies for different application, for instance in [9], the result shows that SVR outperforms a semiempirical water cloud model (WCM) in predicting the backscatter. Additionally, in [10] the theoretical integral equation (IEM) and the semi-empirical Oh models were compared with SVR for soil moisture retrieval and they showed potential of SVR to retrieve soil moisture.

SVR is used with LAI, height, biomass, and soil moisture as input data to model SAR observable as the outputs. SVR is first trained using 80% of data and then tested with the remaining data. SVR is implemented in Python using the scikit-learn package [11] and data normalized in pre-processing step. To address the concern of overfitting, hyper-parameter tuning was applied by grid search and k-fold cross validation to separate the observation to training set and a validation set. Radial basis kernel (RBF), C=100 are considered as the optimum parameters. We tried to maximize the R-Squared value between simulated and observed SAR features, defined as

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \overline{y})^2}, \quad (1)$$

where y_i is observed values, $\overline{y}$ is the mean of the observed data and f_i represents the modeled value. In an ideal but unrealistic case of a perfect model R^2 would be 1.

4. RESULTS AND DISCUSSION

Weather, soil and management data are provided in daily steps for each maize field in a readable format for DSSAT software as it is discussed in 3.1 by Python and it runs through batch processing. Canopy height and LAI outputs are presented in Figure 3 for four arbitrary maize fields based on day of the year (DoY). LAI values as we expected are between 3 to 4 over the Flevoland area, with the height about 1.2 to 1.4 meter. The figure illustrates that fields with higher height also have higher LAI values.

The simulated results including backscatter and coherence in both polarization from DSSAT outputs are compared to Sentinel-1 observations in Table 1. R^2 score and Pearson correlation represents desirable fit between simulated and observed signals. Figure 4 shows satellite-based observations and the simulated ones to evaluate the performance of SVR. The values are averaged over all maize fields for each time and the transparent buffer shows the 20th-80th percentiles.

5. CONCLUSION

SAR observables (backscatter and coherence) in co-pol and cross-pol are modeled by combining crop growth model with machine learning. We use weather, soil and management data as inputs into DSSAT to simulate maize growth variables. The simulated LAI, biomass, height and soil moisture from DSSAT have been used to model SAR observable using SVR.

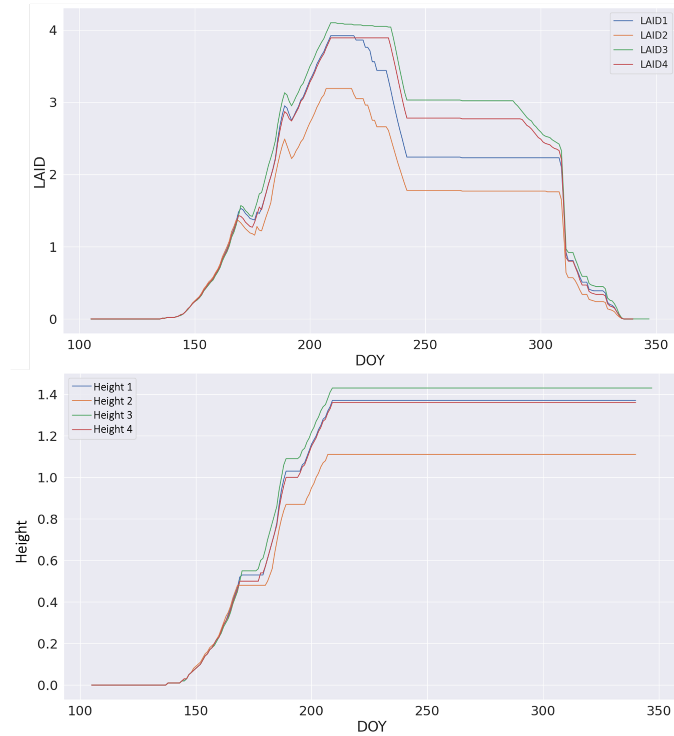

Fig. 3. LAI and canopy height from DSSAT model. Four lines show four different maize fields.

Table 1. R^2 score and Pearson correlation between simulated and observed SAR features.

Simulated feature	R^2	Pearson
Amplitude VH	0.75	0.87
Amplitude VV	0.68	0.83
Coherence VH	0.67	0.83
Coherence VV	0.64	0.80

The performance is accessed by statistical parameters which shows the capability of combining DSSAT and SVR. This approach can be used to assimilate SAR observables into crop-growth models in order to improve the simulation results.

6. REFERENCES

[1] Dimitrios A. Kasampalis, Thomas K. Alexandridis, Chetan Deva, Andrew Challinor, Dimitrios Moshou, and Georgios Zalidis, "Contribution of Remote Sensing on Crop Models: A Review," *Journal of Imaging*, vol. 4, no. 4, pp. 52, Apr. 2018.

[2] A. J. W. de Wit and C. A. van Diepen, "Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts," *Agricultural and*

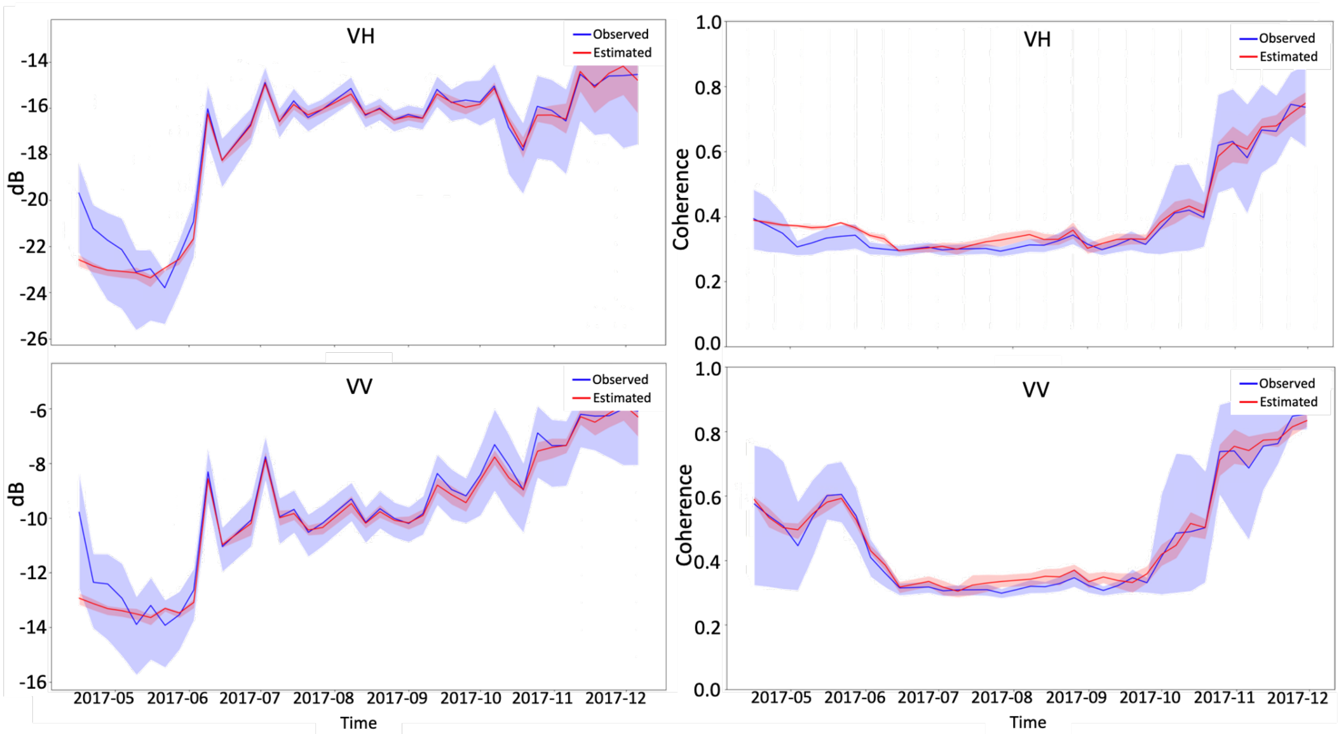

Fig. 4. Comparison of SAR features between SVR simulated and Sentinel-1 observations.(Left) Amplitude and (Right) Coherence in both polarizations. Solid lines represent the average value of the feature and bounded area shows 20th-80th percentiles.

Forest Meteorology, vol. 146, no. 1, pp. 38–56, Sept. 2007.

[3] L. Prévot, H. Chauki, D. Troufleau, M. Weiss, F. Baret, and N. Brisson, "Assimilating optical and radar data into the STICS crop model for wheat," 2001.

[4] Tomislav Hengl et al., "SoilGrids250m: Global gridded soil information based on machine learning," *PLOS ONE*, vol. 12, no. 2, pp. e0169748, Feb. 2017.

[5] "HC27 Generic Soil Profile Database — IFPRI : International Food Policy Research Institute," https://www.ifpri.org/publication/hc27-generic-soil-profile-database.

[6] K. E. Saxton and W. J. Rawls, "Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions," *Soil Science Society of America Journal*, vol. 70, no. 5, pp. 1569–1578, 2006.

[7] Vineet Kumar, Manuel Huber, Bjorn Rommen, and Steele-Dunne, Susan C, "Agricultural SandboxNL Database V1.0," July 2021.

[8] Michael C. Mozer, Michael I. Jordan, and Thomas Petsche, *Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference*, MIT Press, 1997.

[9] Dominik Rains, Hans Lievens, Gabriëlle J. M. De Lannoy, Matthew F. McCabe, Richard A. M. de Jeu, and Diego G. Miralles, "Sentinel-1 Backscatter Assimilation Using Support Vector Regression or the Water Cloud Model at European Soil Moisture Sites," *IEEE Geoscience and Remote Sensing Letters*, pp. 1–5, 2021.

[10] Jamal Ezzahar et al., "Evaluation of Backscattering Models and Support Vector Machine for the Retrieval of Bare Soil Moisture from Sentinel-1 Data," *Remote Sensing*, vol. 12, no. 1, pp. 72, Jan. 2020.

[11] Fabian Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.