RESEARCH ARTICLE

Crop biophysical parameter retrieval from Sentinel-1 SAR data with a multi-target inversion of Water Cloud Model

Dipankar Mandal[a], Vineet Kumar[a,b], Juan M. Lopez-Sanchez[c], Avik Bhattacharya[a], Heather McNairn[d] and Y. S. Rao[a]

[a]Microwave Remote Sensing Lab, Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Mumbai, India; [b]Department of Water Resources, Delft University of Technology, Delft, The Netherlands; [c]Institute for Computer Research, University of Alicante, Alicante, Spain; [d]Ottawa Research and Development Centre, Agriculture and Agri-Food Canada, Ottawa, Canada

ARTICLE HISTORY
Compiled February 27, 2020

ABSTRACT
Estimation of bio-and geophysical parameters from Earth observation (EO) data is essential for developing applications on crop growth monitoring. High spatio-temporal resolution and wide spatial coverage provided by EO satellite data are key inputs for operational crop monitoring. In Synthetic Aperture Radar (SAR) applications, a semi-empirical model (viz., Water Cloud Model (WCM)) is often used to estimate vegetation descriptors individually. However, a simultaneous estimation of these vegetation descriptors would be logical given their inherent correlation, which is seldom preserved in the estimation of individual descriptors by separate inversion models. This functional relationship between biophysical parameters is essential for crop yield models, given that their variations often follow different distribution throughout crop development stages. However, estimating individual parameters with independent inversion models presume a simple relationship (potentially linear) between the biophysical parameters. Alternatively, a multi-target inversion approach would be more effective for this aspect of model inversion compared to an individual estimation approach. In the present research, the multi-output support vector regression (MSVR) technique is used for inversion of the WCM from C-band dual-pol Sentinel-1 SAR data. Plant Area Index (PAI, $m^2\,m^{-2}$) and wet biomass (W, $kg\,m^{-2}$) are used as the vegetation descriptors in the WCM. The performance of the inversion approach is evaluated with in-situ measurements collected over the test site in Manitoba (Canada), which is a super-site in the Joint Experiment for Crop Assessment and Monitoring (JECAM) SAR inter-comparison experiment network. The validation results indicate a good correlation with acceptable error estimates (normalized root mean square error–nRMSE and mean absolute error–MAE) for both PAI and wet biomass for the MSVR approach and a better estimation with MSVR than single-target models (support vector regression–SVR). Furthermore, the correlation between PAI and wet biomass is assessed using the MSVR and SVR model. Contrary to the single output SVR, the correlation between biophysical parameters is adequately taken into account in MSVR based simultaneous inversion technique. Finally, the spatio-temporal maps for PAI and W at different growth stages indicate their variability with crop development over the test site.

KEYWORDS
model inversion; ill-posed problem; biomass; PAI; JECAM.

CONTACT Dipankar Mandal. Email: dipankar_mandal@iitb.ac.in

1. Introduction

Remote sensing data assimilation for crop yield models often employ vegetation biophysical parameters (e.g., Leaf Area Index (LAI), vegetation water content (VWC), biomass, etc.) as key state variables (Baruth et al. 2008; Chipanshi et al. 2012; Boogaard et al. 2013). In a dynamic system for such crop growth model, it is fundamental to regulate crop biophysical parameter distributions during the plant growth period. However, these distributions of the biophysical parameters do not change in the same way along the cultivation cycle (Kross et al. 2015). For example, the dry biomass increases during the grain filling stage of wheat, while the LAI barely changes at this stage. Conversely, during the tillering stage, both these biophysical parameters advance almost proportionally. Therefore a compromise between the functional relationship (e.g., between LAI and biomass) is often made while estimating only one vegetation parameter from remote sensing data. These estimates from the remote sensing data are used to replace crop model simulations at a particular time step in a forcing strategy. Hence, direct ingestion of these crop estimates with a presumed relationship (often linear) in a crop model may lead to instability in the overall model setting.

Alternatively, the simultaneous estimation of biophysical parameters from remote sensing data would be essential considering their correlations and differences in variation (Borchani et al. 2015). Among several Earth observation (EO) systems, Synthetic Aperture Radar (SAR) data have been recognized for vegetation monitoring due to its unique characteristics and sensitivity to geometric and dielectric properties of the target (Ulaby 1975; Steele-Dunne et al. 2017). Radar backscatter intensity simulation for vegetation canopies often relies on radiative transfer models (Ulaby et al. 1990; Prevot, Champion, and Guyot 1993; Karam et al. 1995; De Roo et al. 2001). Such

models characterize vegetation cover with a couple of direct and indirect measurable crop biophysical parameters (Prevot, Champion, and Guyot 1993). To date, a number of studies have been carried out to retrieve these biophysical parameters from SAR data. The semi-empirical Water Cloud Model (WCM) has been extensively utilized to estimate these crop descriptors (Attema and Ulaby 1978), given its relative simplicity to model and invert these parameters (Graham and Harris 2003). Thus, this approach is expected to be well adapted for the operational monitoring of agricultural crops. A large number of experiments (Prevot, Champion, and Guyot 1993; Inoue, Sakaiya, and Wang 2014; Chakraborty et al. 2005; Dabrowska-Zielinska et al. 2007; Bériaux et al. 2015; Hosseini et al. 2015; Fieuzal and Baup 2016; Hosseini and McNairn 2017) reported the potentiality of WCM for biophysical parameter estimation for several crop types. These experiments suggest the possible approaches to invert the WCM for estimating crop biophysical parameters with acceptable accuracies and its scalability.

Operational monitoring of agricultural areas benefits from remote sensing data with high temporal revisit and wide spatial coverage. In this context, the C-band Sentinel-1 SAR constellation offers global high-resolution imagery at an unprecedented spatio-temporal coverage. After the launch of Sentinel-1A in 2014, Satalino et al. (2015) demonstrated the retrieval of above ground biomass (AGB) of wheat. The regression results indicate that the linear cross-pol ratio (VH/VV) is well correlated up to $3.0\,\mathrm{kg\,m^{-2}}$ of wheat biomass, with an root mean square error (RMSE) of $0.6\,\mathrm{kg\,m^{-2}}$. The biomass map generated from the regression model clearly captured the spatial variability within the wheat fields. In more recent studies, the VV and VH backscatter coefficients and the VH/VV ratio of a dense time series of Sentinel-1 data showed sensitivity to crop dynamics at critical phenological stages (Nguyen, Gruber, and Wagner 2016; Veloso et al. 2017; Mandal et al. 2018b).

Estimating vegetation parameters from dual-pol SAR systems might hold significant interest from an operational perspective for agricultural applications based on time series of satellite data. These could be globally obtained from multiple SAR satellites considering the rapid expansion of constellations of satellites such as Sentinel-1 A/B, Canadian RADARSAT Constellation Mission (RCM), SAOCOM (SAtélite Argentino de Observación COn Microondas) and NASA-ISRO SAR (NISAR) mission. Nevertheless, the estimation of vegetation and soil descriptors via the WCM inversion may often lead to an ill-posed inversion problem, which arises because more than one combination of LAI, biomass, and soil parameters can produce an identical backscatter intensity (Bériaux et al. 2015). It may lead to unstable and inaccurate inversion performance. In literature, the iterative optimization (IO) and look-up table (LUT) search techniques are aptly used to solve the ill-posed nature of such inversion problems (Prevot, Champion, and Guyot 1993; McNairn et al. 2012; Hosseini et al. 2015). However, the iterative method produce accurate estimates at the expense of high computational resources when optimizing such inversion problems. Mandal et al. (2019a) indicated the highest computation-intensive nature of IO approach in memory-time performance analysis compared with other inversion approaches. Besides, the intrinsic problem associated with such an optimization of non-linear multi-variate merit function is the possibility to get confined into a local minimum instead of reaching the global minimum (Perez, Jansen, and Martins 2012). Conversely, the LUT search techniques provide an alternative for estimating crop parameters from WCM inversion. However, the LUT approach lacks good generalization capability (Bériaux et al. 2015) which might increase the prerequisite of high computational resources for large areas. It is likely due to the search method need to be run through the entire LUT to get the optimum solution for each resolution cell. It might pose a challenge for estimation of

vegetation descriptors using new generation and upcoming operational SAR missions providing wide area observations.

Recognizing the potential issues of the traditional approaches (IO and LUT search techniques) for applications to larger areas, ill-posed inversion problems in remote sensing are often solved by data-driven nonparametric models (Durbha, King, and Younan 2007; Bériaux, Lambot, and Defourny 2011; Verrelst et al. 2012; Mandal et al. 2019a), which provides a stable and optimum solution. Also, machine learning regression models can provide an optimum solution with a lower computational cost (Mandal et al. 2019a). Machine learning regression approaches such as the support vector regression (SVR) can cope with the nonlinearity of the functional dependence between crop descriptors (in WCM) and the SAR backscatter intensities. Hence, SVR can serve as a suitable technique for model inversion in operational scale applications (Caicedo et al. 2014; Mandal et al. 2019a). However, the formulation of standard SVR involves a single target output. Hence, to estimate multiple biophysical parameters (e.g., Plant Area Index–PAI, and wet biomass), we have to run several (in this case two) independent SVRs which often ignore the correlation between the estimated biophysical parameters (Tuia et al. 2011). Indeed, for plant biophysical parameter retrieval, it is then necessary to contemplate multivariate outputs. This strategy takes care of the functional dependency between the distribution of biophysical parameters, which may not be preserved while estimating them individually. A preliminary attempt in this line was made by Mandal et al. (2018a), which utilized the Multi-output SVR (MSVR) to extend the standard SVR to a multi-target problem to simultaneously estimate vegetation descriptors.

The objective of the current research is to appraise the WCM inversion capability using the MSVR method for simultaneous retrieval of biophysical parameters (viz.,

PAI and wet biomass). The estimation of biophysical parameters is achieved by using the dual-pol VV and VH data acquired from the C-band Sentinel-1 SAR system. The proposed MSVR method is compared with standard SVR to demonstrate how the correlation between retrieved biophysical variables is preserved and their relative performances. The performance of MSVR is investigated for structurally different crops so that the model can be used for various cropping systems. Furthermore, the model transferability is evaluated from point to a regional-site scale by generating spatio-temporal maps of PAI and W.

The rest of the paper is organized as follows: Section 2 illustrates the test site description and the dataset utilized. Section 3 details the methodology proposed in this work. Section 4 discusses the results and their analysis, and finally, the work is summarized and concluded in Section 5.

2. Study area and dataset

The test site is located at south and west of Winnipeg, Manitoba, Canada. This area is considered as one of the super-sites of the Joint Experiment for Crop Assessment and Monitoring (JECAM) network. The average extent of the test area is $26 \times 48\,\mathrm{km}^2$. This area is characterized by major cultivable lands and diverse crop types. The soil texture varies across the site with heavy clay in the east, changing sharply to loamy-sand in the west. Clay and loamy soils account for 76% of the test site, whereas coarse loam, and sandy soils account for 14%. This contrast in soils across the test site leads to significant spatial variability in soil moisture and farming operations across the region. According to the Manitoba Agricultural Services Corporation (MASC) crop insurance data, more than 85% of the site is dominated by annual crops. The major annual crops grown in this area include wheat, canola, soybeans, corn, and oats.

Only a small fraction (<5%) is under grassland and pasture. According to Manitoba Agriculture (Agriculture 2016), the seeding of annual crops starts at the end of April to mid-May with harvesting commencing in August. The field photographs indicating differential vegetation growth stages are shown in Fig. 2 for major crop types.

[Figure 1 about here.]

[Figure 2 about here.]

2.1. *Sampling strategy*

The in-situ measurements were carried out over the test area in 2016 during the Soil Moisture Active Passive Validation Experiment 2016 (SMAPVEX16-MB) (Bhuiyan et al. 2018). During the campaign, in-situ measurements of soil and crops were collected in two separate periods, i.e., Intensive Observation Periods (IOP) (June 08 to June 22 and July 8 to July 22, 2016) from 50 fields. During the field campaign, the majority of crops advanced from an early stage of development to a peak accumulation of biomass at the full vegetative stage. The soil and vegetation sampling were conducted in the individual fields with an average size of 800 m×800 m. In each sampling field, soil moisture measurements were collected from 16 sampling locations, which are placed in two parallel transects along the row direction, as shown in Fig. 1. Two sampling points in a transact were separated by 75 m, while the distance between two transects was 200 m. In each soil sampling location, soil moisture was taken with 3 replicate measurements using Stevens Hydra probes during both IOPs of the campaign.

In each field, 3 points (i.e., points 2, 11, 14 in the first week and 3, 10, 13 in the second week of the campaign) were selected as the vegetation sampling locations as shown in Fig. 1. The vegetation sampling included measurement of PAI, wet biomass, plant

height, and phenological stage (McNairn et al. 2016). A comprehensive illustration of vegetation and soil sampling strategies during the field campaign is provided in the SMAPVEX16-MB report (McNairn et al. 2016). In addition, the annual crop inventory map (Davidson et al. 2017) prepared by the Agriculture and Agri-Food Canada (AAFC) is used as ancillary data for this study to generate crop specific map.

2.2. *Satellite data and data pre-processing*

We have utilized C-band Sentinel-1 data for the current study. The details of the dual-pol (VV and VH) Sentinel-1A data are given in Table 1. The selection of Sentinel-1 images is solely based on the coincidence of the campaign and satellite acquisition dates.

[Table 1 about here.]

The Sentinel-1 provides data in the Interferometric Wide (IW) swath mode in single-look complex (SLC) format. In general, the IW mode comprises three sub-swaths. By following the standard protocol (Mandal et al. 2019b) for processing the Sentinel-1 data, the sub-swaths are split and de-bursted using SNAP toolbox (ESA 2015) for all individual acquisitions. The SLC images have range and azimuth spatial resolution of 5 m$\times$ 20 m. These SLC images are then multi-looked by 4$\times$1 to form the 2$\times$2 covariance matrix, $\langle[\mathbf{C}]\rangle$. Subsequently, despeckle operation is performed for all the elements of $\langle[\mathbf{C}]\rangle$, using the refined Lee filter with a 3$\times$3 window. These multi-temporal images (elements of the diagonal of the $\langle[\mathbf{C}]\rangle$ matrix) are then co-registered using ground control points with an RMSE < 1.02 m. Subsequently, terrain correction and geocoding are applied to these images. Then backscatter intensities $\sigma^{\circ}_{\mathrm{VV}}$ and $\sigma^{\circ}_{\mathrm{VH}}$ are derived from the elements of the $\langle[\mathbf{C}]\rangle$ matrix for an individual date. The in-situ measurement

points (vector file) are overlayed on these σ°_{VV} and σ°_{VH} images. Here it is important to note that the plot size is comparatively bigger (approx. 800 m×800 m) than the size of the image pixel (approx. 15 m×15 m). Hence, the backscattering intensities for each sampling location are calculated as the average over a 3×3 window centered on each site.

2.3. *Calibration and validation dataset*

The extracted backscatter intensities (VV and VH) are tabulated with corresponding in-situ measurements available for each acquisition date as given in Table 1. These tabulated datasets are further utilized for calibration and validation of the Water Cloud Model.

According to SMAPVEX16-MB experimental plan (McNairn et al. 2016), in each field, 3 points were selected for vegetation sampling. However, the preprocessed data provided by the SMAPVEX16-MB team to the end-users have some discrepancy for different crop types when compared with the experiment plan. For wheat, both the PAI and biomass measurements are provided for each three sampling points; while for other crops, the biomass measurements are not available for all these three points. Hence, we split the dataset differently for wheat and other crops. For all wheat fields, data from sites 11 and 14 are used in validation and site 2 in calibration. As such, partitioned data from sites 11 and 14 are not available from the SMAPVEX16-MB data for other crop types. So, data is split in terms of field numbers. From this whole feature set, the calibration data split is performed by selecting approximately half of the data randomly, while the remainder of data is used as independent validation dataset as given in Table 2.

The first dataset is used in the WCM calibration and utilized to generate the LUT.

The validation data are reserved for testing to assess the performance of the inversion approach. Notably, all the 50 fields are not sampled on each acquisition day as given in Table 1. The sampling strategy is quite different than the general approach of repeat in-situ measurements over the same sampling site on each acquisition. It is well described in SMAPVEX16-MB campaign planning report (McNairn et al. 2016). Measurements were performed in selected sampling fields of a particular crop on 1st week of IOP1, but in second-week measurements were skipped and performed in other selected fields. In addition, Sentinel-1 images do not cover all the fields in the test site; hence, some measurements are not considered in the current work.

[Table 2 about here.]

3. Methodology

3.1. *Multi-output Support Vector Regression (MSVR)*

For a regression problem, the Support vector machine (SVM) uses the feature dataset to derive a continuous-valued function between a set of inputs and an output (Vapnik 2013). In general, SVR generates a mapping function between the input feature $\boldsymbol{x}$ and a target $\boldsymbol{y} \in \mathbb{R}$ for a one-dimensional regression case. However, for a multi-target case, several independent SVRs need to be utilized to get functional relationships between the input feature and the target variables. In a multi-target regression problem, the target variables can be represented as a vector with Q variables, i.e., $\boldsymbol{y} \in \mathbb{R}^Q$. In this context, the standard SVRs solves the regression problem by ignoring the correlation among the target variables (Tuia et al. 2011). Conversely, the MSVR solves a multi-dimensional regression problem by evaluating the regressor w^j and b^j $(j = 1, \cdots, Q)$ for every target variable. An extensive description of MSVR formulation can be found

in Tuia et al. (2011). In the present work, the target vector is generated with PAI, and wet biomass. The input features are kept as the corresponding SAR backscatter intensities. The implementation of MSVR in the present inversion chain is discussed in Sec. 3.3.

3.2. *Vegetation modeling*

In SAR theory, the radar backscatter models for vegetation canopy have evolved from the physics-driven complex functions to describe volume scattering throughout the continuous canopy layer (Graham and Harris 2003; Steele-Dunne et al. 2017). The well known physical models (e.g., the Michigan Microwave Canopy Scattering model (MIMICS) or Radiative transfer model (RTM) based approaches) can provide an excellent agreement of the estimated backscatter with the observed values. However, the complexity of these models and the requirement of intensive in-situ measurements make them difficult to implement for any operational scale. Alternatively, semi-empirical models derived from the concept of physics of scattering and experiments, are used efficiently in literature. An acceptable form of a semi-empirical model is the Water Cloud Model (Attema and Ulaby 1978). The WCM simulates the radar backscatter intensities from the vegetation-soil system as an incoherent sum of the contributions from the vegetation and the underlying soil layer components (1). Attenuation of the soil backscatter intensity component is represented by the two way attenuation factor τ^2.

$$\sigma^\circ = \sigma^\circ_{veg} + \tau^2 \sigma^\circ_{soil} \tag{1}$$

$$\sigma^{\circ}_{veg} = AV_1 \cos\theta\Big(1 - \exp\Big(-\frac{2BV_2}{\cos\theta}\Big)\Big) \tag{2}$$

$$\tau^2 = \exp(-2BV_2/\cos\theta) \tag{3}$$

where V_1 and V_2 are vegetation parameters. The radar incidence angle is indicated by the term θ. The backscatter intensity due to the soil component σ°_{soil} is expressed as proposed in (4):

$$\sigma^{\circ}_{soil} = CM_v + D \tag{4}$$

where M_v is volumetric soil moisture. The model coefficient C is a representation of the sensitivity of radar signal to soil moisture and D indicates the backscatter intensity due to surface roughness.

Here it is important to note that different canopy descriptors have been used in SAR literature for realizing the complex vegetation structure. A number of experimental research on the combination of different vegetation descriptors were conducted in several studies (Graham and Harris 2003; Lievens and Verhoest 2011; Kumar, Suryanarayana Rao, and Arora 2015), which appointed LAI, wet and dry biomass, fraction cover etc. to describe the vegetation canopy. It is interesting to observe that some selective variables are widely used due to their physical significance and availability of straight forward in-situ measurement techniques. In the present work, the following form (5) of WCM is adapted to simulate the vegetation-radar signal interaction.

$$\sigma^{\circ} = AL^E \cos\theta \left(1 - \exp\left(-\frac{2BW^F}{\cos\theta}\right)\right) + (CM_v + D) \times \exp\left(-\frac{2BW^F}{\cos\theta}\right) \tag{5}$$

where A, B, C, D, E, and F are the model coefficients. L, and W are the PAI and wet biomass, respectively.

3.3. *Model inversion*

The estimation of vegetation descriptors from the WCM model as defined in (5) is treated as an ill-posed inversion problem. The direct inversion of this model is challenging, given the observed backscatter intensities. Hence, several steps need to be followed to achieve the inversion of WCM, which essentially includes four steps: a) calibration of the WCM, b) forward modeling and LUT generation, c) MSVR model training, and d) simultaneous estimation of crop biophysical parameters using MSVR model.

[Figure 3 about here.]

Semi-empirical WCM models require the in-situ datasets that allow the fitting of the model coefficients which are derived from the concept of physics of scattering and experiments. At first, the WCM parameterization is performed by estimating the WCM model parameters (A, B, C, D, E, and F) as (5). This step is typically called as a calibration phase. The calibration dataset is used to estimate the model parameters using the non-linear least square optimization of the Levenberg-Marquardt algorithm (Moré 1978). The WCM is calibrated for both the VV and VH backscatter intensities for each crop using the in-situ measurements as given in Table 2. The performance of the WCM calibration is assessed in terms of the correlation coefficient (r) and RMSE between the simulated and the observed backscatter intensities for individual crop and each polarization channel.

Theoretically, the forward modeling denotes the generation of response values from a

model using a set of input data, which can be utilized for training the regression model. In case of the forward WCM, the backscatter intensities in both polarizations can be simulated for extrapolated points in the vegetation parameter space. This dataset is often treated as the synthetic dataset. To make this kind of synthetic data resemble natural measurements of vegetation space, some noise is usually added (Durbha, King, and Younan 2007; El Hajj et al. 2016). However, these datasets often contain uncertainties, and the realistic nature of synthetic data is not guaranteed. Hence, in the present study, we have used the combinations of vegetation parameters from the calibration data to derive the associated backscatter intensities by the forward WCM. For each crop type, the backscatter intensities are simulated from the calibrated WCM at each calibration point, and subsequently the LUT is derived.

The LUT elements are then utilized as training data to build the MSVR model. Hosseini et al. (2015) indicated that a combination of polarization channels can improve crop parameter estimation when compared with a single channel. Thus, backscatter intensities in both the co-pol and cross-pol channels (VV+VH) are included in the MSVR target vector. On the other hand, crop biophysical parameters (PAI, and wet biomass) are used as the MSVR model responses. It is important to note that all biophysical parameters used in WCM can be retrieved simultaneously using the MSVR.

The MSVR models are trained using the LUT elements for each crop separately. The meta-parameters of the MSVR (the insensitive parameter ϵ, kernel parameter γ, and margin parameter C') are selected using a k-fold cross-validation technique for an individual crop. Finally, the validation dataset is used to estimate the crop biophysical parameters simultaneously from the trained MSVRs for each crop. Here it is important to note that for each crop there exists one MSVR. As such, three MSVR models are generated separately for this study. The inversion accuracies are tested with

independent validation points as given in Table 2. The estimation accuracies between observed and estimated crop descriptors (D_c) are performed for the validation dataset with the correlation coefficient (r), Mean Absolute Error (MAE) and normalized root mean square error (nRMSE) as (6):

$$\text{nRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(D_c^{\text{Observed}} - D_c^{\text{Predicted}})^2}{n}}}{\frac{\sum_{i=1}^{n} D_c^{\text{Observed}}}{n}} \tag{6}$$

These three error measurement are estimated for each descriptor for all individual crops. Subsequently, PAI and biomass maps over the test area are generated in *on the fly mode* (prediction phase) using Sentinel-1 acquisitions and the crop inventory map. During this prediction phase, trained MSVR models are selected as per the crop map. Once selected, these models simultaneously estimate PAI and biomass for each resolution cell, as shown in Fig. 3.

3.4. *Experimental design*

This experiment tests the proposed MSVR approach and the standard single-target SVR approach for crop biophysical parameter retrieval for individual crops. We have used the same training and validation dataset for the comparison. This comparison has two objectives: 1) to test the retrieval accuracies for MSVR and SVR approaches, and 2) to test the correlation between estimated parameters (PAI and W estimates) from two approaches and compare them with the actual correlation between these parameters in ground measurement space.

Apart from generating geophysical parameter estimates simultaneously from the proposed MSVR approach (Sec. 3.3), separate SVR models are used. The biophysical parameters, i.e., PAI and W, are estimated separately by constructing two sets of the

SVR model. In one set, both the radar backscatter intensities (VV and VH) are used as predictors, while PAI as a response. In another set, W is considered as response with the same training data. The PAI and W are retrieved individually using two separate SVR models for the validation dataset. The correlation between PAI and W are assessed subsequently with scatter plot for both the SVR and MSVR. The in-situ measurements of PAI and W are used to determine if the correlation between these biophysical parameters is preserved during the model inversion for individual crops.

4. Results and discussions

This section describes the results of the WCM inversion for each crop type and compares these estimates to in-situ measurements. Subsequently, the accuracy of the WCM model calibration and inversion is assessed in Sec. 4.1 and Sec. 4.2. Further, the competitive performances between the multi and single output SVM regression are evaluated with a correlation analysis between the PAI and W.

4.1. *Calibration of the WCM*

The parameterization of the WCM for an individual crop is performed using the calibration data as discussed in Sec. 3.3. WCM is parametrized individually for both the VV and VH polarization channels for wheat, soybean, and canola. This process results in 6 different model equations. The coefficients for each model (crop and polarization combinations) along with the F-statistics and the p-values (level of significance) are given in Table 3.

[Table 3 about here.]

Prior to evaluating errors in biophysical parameter estimation, the goodness of fit of

the parameterized WCM model is evaluated using the calibration dataset. The accuracies of the calibrated WCMs are assessed by comparing the observed and the simulated backscatter intensities for the calibration dataset. This evaluation is presented in Table. 4.

[Table 4 about here.]

The correlation between the observed and the simulated backscatter intensities is 0.80 (VV) and 0.62 (VH) for wheat. The RMSE is low ($0.007\,\mathrm{m^2\,m^{-2}}$) for the cross-pol channel as compared to the co-pol VV channel ($0.026\,\mathrm{m^2\,m^{-2}}$). This disparity is likely due to the unique morphological structure of wheat plant. Wheat belongs to the graminaceous family which is characterized by vertical stems and erectophile leaf distribution. Moran et al. (2012) reported a similar differential sensitivity of σ°_{HV} and $\sigma^{\circ}_{\mathrm{VV}}$ for barley (graminaceous family) for C-band RADARSAT-2 data.

The correlation coefficients (r) between the observed and the estimated backscatter coefficients are 0.87 (VV) and 0.83 (VH) for canola. However, the RMSE is lower for VH ($0.010\,\mathrm{m^2\,m^{-2}}$) than for VV ($0.051\,\mathrm{m^2\,m^{-2}}$) possibly due to the highly random structure of canola canopies. The sensitivities of backscatter intensities in both the VV and VH polarizations is observed through the entire phenological evolution of canola. The comparison of simulated and observed backscatter intensity values are notably different for low biomass crops like soybean. The r values for soybean are found to be 0.64 (VV), and 0.68 (VH). The RMSE is comparatively lower for the cross-pol channel ($0.006\,\mathrm{m^2\,m^{-2}}$) for soybean indicating a better calibration for cross-pol channel. It may be due to the sensitivity of the VH backscatter intensity to diffused scattering from the crop canopy as the biomass and the PAI increases. Alike sensitivity of VH backscatter intensity with changes in crop biomass of soybean was indicated

for C-band RADARSAT-2 acquisitions during the SMAPVEX-12 campaign (Wiseman et al. 2014; Hosseini et al. 2015).

4.2. *Validation for crop biophysical parameter estimation*

The MSVR model inversion is evaluated with the validation dataset for each crop individually. The estimated PAI, and W are compared with the in-situ measurements on a 1:1 plot to investigate the performance of the inversion approach using the validation points. Results are shown in Fig. 4 for all individual crops.

[Figure 4 about here.]

4.2.1. Wheat

For wheat, the in-situ measured PAI varies from $1.10\,\mathrm{m^2\,m^{-2}}$ to $8.95\,\mathrm{m^2\,m^{-2}}$ and advanced from leaf development to fruiting stages. The correlation (r) between the estimated and in-situ measured PAI is 0.83 with nRMSE and MAE values of 0.246 and $0.893\,\mathrm{m^2\,m^{-2}}$, respectively as shown in Fig. 4a. The proposed MSVR technique with the VV+VH polarization combination as input produces estimation of PAI over the entire range. However, underestimation of PAI values is apparent as the plant area has reached about $7.0\,\mathrm{m^2\,m^{-2}}$ at the end of the heading stage. Conversely, an overestimation in PAI is observed during the early tillering stages (when PAI was $<2.5\,\mathrm{m^2\,m^{-2}}$), which is likely due to soil contribution to the radar backscatter intensities.

The in-situ measured wet biomass varies from $0.40\,\mathrm{kg\,m^{-2}}$ to $6.0\,\mathrm{kg\,m^{-2}}$ during the field campaign period. The $r-$value between observed and estimated wet biomass is 0.75, which is similar to PAI estimation (Fig. 4d). The errors of estimate (nRMSE and MAE) for biomass are 0.333 and $0.707\,\mathrm{m^2\,m^{-2}}$, respectively, higher than the PAI estimation errors. The wide margin of wet biomass estimation is apparent ($1.2\,\mathrm{kg\,m^{-2}}$)

across the 1:1 line for high biomass conditions ($> 2.5\ \mathrm{kg\,m^{-2}}$). It is likely due to saturation of C-band when wheat biomass is high at the end of the heading stage. Similar accuracies were also declared in Hosseini and McNairn (2017) for retrieval of total biomass for wheat using combinations of VV and HV polarization channels.

4.2.2. Canola

The estimated and in-situ measured PAI are highly correlated for canola ($r = 0.88$) with nRMSE of 0.466 and MAE of $1.144\,\mathrm{m^2\,m^{-2}}$) (Fig. 4b). However, the retrieval results during the inflorescence emergence to flowering stages indicated underestimation when the PAI reaches about $5\,\mathrm{m^2\,m^{-2}}$. This underestimation is likely due to saturation of C-band radar wave when canola accumulates considerable biomass after flowering and pod development. A considerable amount of depolarization is possible for a dense canopy with stems, leaves, and pods. Pacheco et al. (2016) also reported approximately four times increment of HV/VV differential reflectivity with the advancement of plant growth from stem elongation to the flowering stage for canola.

The simultaneous estimation using the MSVR inversion method indicates high estimation accuracy for wet biomass as well as PAI, as shown in Fig. 4e. The r and MAE between estimated and in-situ measured wet biomass are 0.89 and $0.841\,\mathrm{kg\,m^{-2}}$, respectively. Similar to other crop types, the overestimation and underestimation of the wet biomass are also apparent during the early stage of vegetative growth and the flowering stage of canola, respectively. The sensitivity of the SAR signal to canola biomass accumulation from leaf development up to the flowering stage is apparent, and saturation of C-band signal is possible due to the high volume of plant material at the time of pod development (Wiseman et al. 2014). The overall estimation result is marginally better in case of wet biomass as compared to PAI.

4.2.3. Soybean

The ground measured PAI for soybean varies from 0.2 to 4.3 $m^2\,m^{-2}$, covering phenological stages from leaf development to flowering. The correlation coefficient for PAI estimation is 0.73 with nRMSE and MAE of 0.517 and 0.798 $m^2\,m^{-2}$, respectively (Fig. 4c). An overestimation of PAI retrieval is observed during the leaf development stage, which is likely due to sparse vegetation cover. Early in its development, canopy closure is very small (PAI<1.45 $m^2\,m^{-2}$) in between soybean rows, which allows the radar wave to interact more with the exposed soil (Wiseman et al. 2014; Ratha et al. 2019). Nevertheless, estimated and observed PAI follow the 1:1 line when PAI>1.6 $m^2\,m^{-2}$. This higher plant area indicates the end of side-shoot formation stage, with dense canopy cover. Nonetheless, a significant spread (about 1.2 $m^2\,m^{-2}$) is observed for high PAI.

The behavior of W estimates (Fig. 4f) results in a higher correlation coefficient ($r = 0.79$) relative to PAI. Early in the season, soybean biomass is low. This sparse canopy along with wide row spacing make modeling of this canopy more difficult. The in-situ measured wet biomass of soybean changed from 0.02 $kg\,m^{-2}$ to 2.2 $kg\,m^{-2}$ during the field campaign window. For this crop, radar backscatter would be significantly affected by the underlying soil. A sparse soybean canopy could not attenuate much the radar signal leading to surface scattering from the underlying soil layer, as also reported with Sentinel-1 observations by Veloso et al. (2017). Error of estimates for wet biomass are nRMSE<0.491 and MAE 0.260 $kg\,m^{-2}$ with a slightly increased spread around the 1:1 line. These variations in biomass estimates are possibly due to the variations in the plant density in several fields. It is evident from in-situ measurements that the number of plants along the row was higher for a few fields ($\approx$ 50 plants m^{-1}) than for other fields (12 plants m^{-1}).

4.3. *Comparison of inversion results for MSVR and SVR*

The performance of the MSVR based inversion approach is compared with the single target SVR model for the same validation dataset. Comparative results of retrieval accuracies of biophysical parameters for each crop types are presented in Table 5 with the correlation coefficient (r), normalized RMSE, and mean absolute error (MAE).

Highest accuracy ($r = 0.83$) in the PAI estimation is obtained with the MSVR approach (nRMSE = 0.246 and MAE = $0.893\,\mathrm{m^2\,m^{-2}}$) as compared to SVR ($r = 0.75$, nRMSE = 0.306 and MAE = $1.012\,\mathrm{m^2\,m^{-2}}$) for wheat. For wet biomass, a difference in both the nRMSE and MAE is obtained for SVR as compared to the MSVR. For the given three crop types, the MSVR has outperformed the SVR estimates in all biophysical parameters for each crop types. It is noticeable from the comparative analysis that the inversion technique with MSVR, which takes into account the correlations between the plant biophysical parameters, produced acceptable inversion results for all crops. It can be further confirmed from the correlation analysis between estimated biophysical parameters, as shown in Fig. 5.

[Table 5 about here.]

4.4. *PAI and wet biomass relationship*

In addition to comparative analysis between retrieval results from the MSVR and SVR approaches, it is imperative to analyze the correlation between the estimated biophysical parameters and the observed ones for both methods. To achieve this analysis, the experiment design is set up as discussed in Sec. 3.4. Apart from generating biophysical parameters simultaneously from the proposed MSVR approach (Sec. 3.3), separate SVR models are used. The correlation between PAI and W are then evaluated with

scatter plots for both the SVR and MSVR runs. The in-situ measurements of PAI and W are used to determine whether the correlation between these two biophysical parameters is preserved during the model inversion for individual crops. The functional relationships between the PAI and W are shown in Fig. 5 for individual crops.

[Figure 5 about here.]

For wheat, a non-linear function is apparent for the in-situ measured PAI and W. The relationships between the estimated biophysical parameter are also similar in the case of MSVR, and SVR approaches. However, MSVR performed better (R^2 >0.79) as compared to the SVR based single target output (R^2 <0.72). Of note, an exponential model fits the data well in both cases. However, for canola, a logarithmic function fit is found superior. It is evident from Fig. 5 that, for canola, the correlation between the MSVR-based estimates of PAI and W (R^2 =0.88) is preserved better when compared with the SVR model (R^2 =0.66). However, for soybean, a non-linear functional dependence (possibly exponential function) is found better to fit in-situ measured PAI and W. The overall analysis indicates that the non-linear relationship between the biophysical parameters is successfully retained by the MSVR method (Fig. 4.4), while the performance of the single-output SVR is comparatively low.

4.5. *PAI and biomass maps*

The MSVR approach is used to invert the WCM to simultaneously produce PAI and biomass maps for the test site using the Sentinel-1 acquisitions. The high resolution (20 m) maps for PAI and W are shown in Fig. 6 for three selected acquisition dates. Both the spatial as well as the temporal variability in plant growth is observed for various fields in these map products. Once the regression model is established, the algorithm took a computational time of about 70 sec. to process 2500 × 2500 pixels

and generate both the PAI and biomass maps for a single date Sentinel-1 data.

[Figure 6 about here.]

The Manitoba weekly crop reports (Agriculture 2016) suggest that seeding was completed in the central region of Manitoba during the second week of June. Wheat plants were growing rapidly, and most fields were at the tillering stage on 13 June. A few of the most advanced fields were close to the flag leaf stage. Data from the SMAPVEX16-MB campaign confirms that the majority of the wheat and oat fields were tillering with average wet biomass of $1.02\,\mathrm{kg\,m^{-2}}$ and PAI of $3.44\,\mathrm{m^2\,m^{-2}}$. However, the over-estimates the PAI and W values by the inversion algorithm are pronounced in several wheat fields on 13 June during the early vegetative stage, as shown in Fig. 6. These results are consistent with the validation results shown in Fig. 4.

On 13 June, the estimated PAI and W in the majority of canola fields are low (about $0.7\,\mathrm{m^2\,m^{-2}}$ and $0.5\,\mathrm{kg\,m^{-2}}$) as compared to cereal crops. From the Manitoba weekly crop reports (Agriculture 2016) canola seeding was almost completed and the crop was emerging rapidly. Earlier seeded fields were reaching the rosette stage. Thus, the effect of the soil component is expected to dominate rather than the vegetation. Similar results are observed for the low biomass soybean crop. Soybean seeding was completed during this period and plants were in their unifoliate to third trifoliate growth stage. This results in very low PAI and biomass values in the biophysical maps.

During the first week of July, good crop growth was reported in the Manitoba weekly crop reports (Agriculture 2016). Nevertheless, excess moisture was a concern in fields with late seeded crops due to heavy rainfall events in the previous week. However, more advanced crops had been better able to handle higher rainfall, although stand thinning and yellowing of crops was evident. Cereals were at their stem elongation

to flag leaf stages. Some fields advanced to the heading stage. Thus increases in PAI and biomass are observed in Fig. 6 up to $0.7\,\mathrm{m^2\,m^{-2}}$ and $0.5\,\mathrm{kg\,m^{-2}}$, respectively. In contrast, for soybean, increases in PAI and W are negligible in the map products. Nonetheless, the majority of the soybeans were at the seventh trifoliate stage with PAI and W of $0.7\,\mathrm{m^2\,m^{-2}}$ and $0.5\,\mathrm{kg\,m^{-2}}$, respectively.

The canola fields were bolting, forming early flowers; podding started in the most advanced fields. A rapid increase in PAI and W is apparent in the PAI and W maps of canola. A rapid growth is evident in canola up to PAI of $0.7\,\mathrm{m^2\,m^{-2}}$ and wet biomass of $0.5\,\mathrm{kg\,m^{-2}}$ (Fig. 6).

During the third week of July, most of the crops were at the end of their vegetative growth and were initializing their reproductive stages. Cereal crops were at an advanced heading stage with high PAI and biomass (about $0.7\,\mathrm{m^2\,m^{-2}}$ and $0.5\,\mathrm{kg\,m^{-2}}$). However, lodging was reported in several fields as a result of high winds and thunderstorm activity (Agriculture 2016). The canola was growing rapidly and most was in full flowering. Podding was observed in the most advanced fields where flowering was completed. The PAI and biomass map clearly capture the relative growth of canola in several fields (Fig. 6). Rapid growth in soybean canopies is evident on 19 July as soybeans continue to flower and pod (Agriculture 2016). Temperature or moisture stress can limit the quality of soybean pods and seeds during these critical growth stages.

5. Conclusion

Standard inversion techniques with a single target estimation framework seldom preserve the correlation amongst the target parameters. Nevertheless, it is often indispensable to preserve correlations between the biophysical parameters for better estimation of crop productivity. Hence, a novel inversion framework for the WCM with multi-

target output technique has been proposed in this research for simultaneous retrieval of biophysical parameters (PAI, and wet biomass). The applicability of this inversion approach is assessed for several major crops using C-band Sentinel-1 dual-pol data.

Results exhibit high correlation coefficients and low estimation errors for simultaneous retrieval of biophysical parameters. Notably, the relationship between the estimated PAI and W indicates that the MSVR successfully preserves the correlation between the crop biophysical parameters during the inversion process. Besides, biophysical parameters maps capture the spatial variability among and within the crop fields. These biophysical maps would enable continuous monitoring at large spatial scales throughout the season, hence supporting yield forecasting and productivity monitoring.

Downstream users might be interested in weekly products from an operational mission like Sentinel-1. In fact, the frequent revisit of SAR satellites is necessary to monitor critical phenological stages during the crop season. With the increased capability of this constellation, the in-situ measurements can be enriched to reconstruct the MSVR at other phenological stages as well, which are not observed in the present research. With the synergy of Sentinel-1A and 1B, the inversion would be improved as integrating data from other growth stages could allow better inter-correlation between PAI and biomass; which could lead to a better estimation.

Although the formulation of MSVR model with kernel tricks often condemned to be computationally intensive, the problem of computational efficiency can be solved by employing a parallel computation strategy during the training and prediction phases of WCM inversion model. Nevertheless, this inversion strategy needs to be further investigated for different cropping systems for the applicability of WCM at cross-site validation and for a dense time series data cube through several test site under JECAM

SAR Inter-Comparison Experiment to be applied for an operational scale.

This research demonstrates a significant potential of Sentinel-1 data for simultaneous estimation of crop biophysical parameters, with promising operational accuracy levels. This is illustrated by producing analysis-ready biophysical maps (ARD products). The MSVR based model inversion approach can be transferred from a point scale to a regional test site. With the synergy between PAI and biomass products derived from Sentinel-1, data could contribute significantly to monitor crop development.

Acknowledgements

Authors would like to thank ground team members for support during data collection and ESA for providing Sentinel-1 through Copernicus Open Access Hub. The tremendous contribution of the 2016 SMAPVEX aircraft and field crews is also acknowledged. This research was supported in part by Shastri Indo-Candian Institute, New Delhi, India and the Spanish Ministry of Economy, Industry and Competitiveness, in part by the State Agency of Research (AEI), in part by the European Funds for Regional Development under project TEC2017-85244-C2-1-P.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Agriculture, M. B. 2016. "Agriculture—Province of Manitoba." http://www.gov.mb.ca/agriculture/crops/seasonal-reports/crop-report-archive/index.html.

Attema, EPW, and Fawwaz T Ulaby. 1978. "Vegetation modeled as a water cloud." *Radio science* 13 (2): 357–364.

Baruth, Bettina, A Royer, Anja Klisch, and G Genovese. 2008. "The use of remote sensing within the MARS crop yield monitoring system of the European Commission." *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci* 37: 935–940.

Bériaux, Emilie, Sébastien Lambot, and Pierre Defourny. 2011. "Estimating surface-soil moisture for retrieving maize leaf-area index from SAR data." *Canadian Journal of Remote Sensing* 37 (1): 136–150.

Bériaux, Emilie, François Waldner, François Collienne, Patrick Bogaert, and Pierre Defourny. 2015. "Maize Leaf Area Index Retrieval from Synthetic Quad Pol SAR Time Series Using the Water Cloud Model." *Remote Sensing* 7 (12): 16204–16225.

Bhuiyan, Hassan AKM, Heather McNairn, Jarrett Powers, Matthew Friesen, Anna Pacheco, Thomas J Jackson, Michael H Cosh, et al. 2018. "Assessing SMAP Soil Moisture Scaling and Retrieval in the Carman (Canada) Study Site." *Vadose Zone Journal* 17 (1). Doi: 10.2136/vzj2018.07.0132.

Boogaard, Hendrik, Joost Wolf, Iwan Supit, Stefan Niemeyer, and Martin van Ittersum. 2013. "A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union." *Field crops research* 143: 130–142.

Borchani, Hanen, Gherardo Varando, Concha Bielza, and Pedro Larrañaga. 2015. "A survey on multi-output regression." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 5 (5): 216–233.

Caicedo, Juan Pablo Rivera, Jochem Verrelst, Jordi Munoz-Mari, Jose Moreno, and Gustavo Camps-Valls. 2014. "Toward a semiautomatic machine learning retrieval of biophysical parameters." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (4): 1249–1259.

Chakraborty, M, KR Manjunath, S Panigrahy, N Kundu, and JS Parihar. 2005. "Rice crop parameter retrieval using multi-temporal, multi-incidence angle Radarsat SAR data." *ISPRS*

Journal of Photogrammetry and Remote Sensing 59 (5): 310–322.

Chipanshi, Aston, Yinsuo Zhang, Nathaniel Newlands, Harvey Hill, and David Zamar. 2012. "Canadian Crop Yield Forecaster (CCYF): a GIS and statistical integration of agro-climates and remote sensing information." *The Application of Remote Sensing and GIS Technology on Crops Productivity* 30: 3.

Dabrowska-Zielinska, K, Y Inoue, W Kowalik, and M Gruszczynska. 2007. "Inferring the effect of plant and soil variables on C-and L-band SAR backscatter over agricultural fields, based on model analysis." *Advances in Space Research* 39 (1): 139–148.

Davidson, AM, T Fisette, H Mcnairn, and B Daneshfar. 2017. "Detailed crop mapping using remote sensing data (crop data layers)." In *Handbook on remote sensing for agricultural statistics*, edited by J. Delincé, Chap. 4, 91–129. Rome: Global Strategy to improve Agricultural and Rural Statistics (GSARS).

De Roo, Roger D, Yang Du, Fawwaz T Ulaby, and M Craig Dobson. 2001. "A semi-empirical backscattering model at L-band and C-band for a soybean canopy with soil moisture inversion." *IEEE Transactions on Geoscience and Remote Sensing* 39 (4): 864–872.

Durbha, Surya S, Roger L King, and Nicolas H Younan. 2007. "Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer." *Remote Sensing of Environment* 107 (1): 348–361.

El Hajj, Mohammad, Nicolas Baghdadi, Mehrez Zribi, Gilles Belaud, Bruno Cheviron, Dominique Courault, and François Charron. 2016. "Soil moisture retrieval over irrigated grassland using X-band SAR data." *Remote Sensing of Environment* 176: 202–218.

ESA. 2015. "User Guides - Sentinel-1 SAR." `https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar/acquisition-modes/interferometric-wide-swath`.

Fieuzal, R, and F Baup. 2016. "Estimation of leaf area index and crop height of sunflowers using multi-temporal optical and SAR satellite data." *International Journal of Remote Sensing* 37 (12): 2780–2809.

Graham, AJ, and R Harris. 2003. "Extracting biophysical parameters from remotely sensed

radar data: a review of the water cloud model." *Progress in Physical Geography* 27 (2): 217–229.

Hosseini, Mehdi, and Heather McNairn. 2017. "Using multi-polarization C-and L-band synthetic aperture radar to estimate biomass and soil moisture of wheat fields." *International Journal of Applied Earth Observation and Geoinformation* 58: 50–64.

Hosseini, Mehdi, Heather McNairn, Amine Merzouki, and Anna Pacheco. 2015. "Estimation of Leaf Area Index (LAI) in corn and soybeans using multi-polarization C-and L-band radar data." *Remote Sensing of Environment* 170: 77–89.

Inoue, Yoshio, Eiji Sakaiya, and Cuizhen Wang. 2014. "Capability of C-band backscattering coefficients from high-resolution satellite SAR sensors to assess biophysical variables in paddy rice." *Remote Sensing of Environment* 140: 257–266.

Karam, Mostafa A, Faouzi Amar, Adrian K Fung, Eric Mougin, Armand Lopes, David M Le Vine, and André Beaudoin. 1995. "A microwave polarimetric scattering model for forest canopies based on vector radiative transfer theory." *Remote Sensing of Environment* 53 (1): 16–30.

Kross, Angela, Heather McNairn, David Lapen, Mark Sunohara, and Catherine Champagne. 2015. "Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops." *International Journal of Applied Earth Observation and Geoinformation* 34: 235–248.

Kumar, Kamal, Hari Prasad Suryanarayana Rao, and MK Arora. 2015. "Study of water cloud model vegetation descriptors in estimating soil moisture in Solani catchment." *Hydrological processes* 29 (9): 2137–2148.

Lievens, Hans, and Niko EC Verhoest. 2011. "On the retrieval of soil moisture in wheat fields from L-band SAR based on water cloud modeling, the IEM, and effective roughness parameters." *IEEE Geoscience and remote sensing Letters* 8 (4): 740–744.

Mandal, Dipankar, Mehdi Hosseini, Heather McNairn, Vineet Kumar, Avik Bhattacharya, YS Rao, Scott Mitchell, Laura Dingle Robertson, Andrew Davidson, and Katarzyna

Dabrowska-Zielinska. 2019a. "An investigation of inversion methodologies to retrieve the leaf area index of corn from C-band SAR data." *International Journal of Applied Earth Observation and Geoinformation* 82: 101893.

Mandal, Dipankar, Vineet Kumar, A Bhattacharya, YS Rao, and Heather McNairn. 2018a. "Crop Biophysical Parameters Estimation with a Multi-Target Inversion Scheme using the Sentinel-1 SAR Data." In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 6611–6614. IEEE.

Mandal, Dipankar, Vineet Kumar, Avik Bhattacharya, Yalamanchili Subrahmanyeswara Rao, Paul Siqueira, and Soumen Bera. 2018b. "Sen4Rice: A Processing Chain for Differentiating Early and Late Transplanted Rice Using Time-Series Sentinel-1 SAR Data With Google Earth Engine." *IEEE Geoscience and Remote Sensing Letters* 15 (12): 1947–1951. Doi: 10.1109/LGRS.2018.2865816.

Mandal, Dipankar, Divya Sekhar Vaka, Narayana Rao Bhogapurapu, VSK Vanama, Vineet Kumar, Yalamanchili S Rao, and Avik Bhattacharya. 2019b. "Sentinel-1 SLC preprocessing workflow for polarimetric applications: A generic practice for generating dual-pol covariance matrix elements in SNAP S-1 toolbox." *Preprints* 2019110393. Doi: 10.20944/preprints201911.0393.v1.

McNairn, H, J Shang, X Jiao, and B Deschamps. 2012. "Establishing crop productivity using RADARSAT-2." *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci* 39: B8.

McNairn, Heather, Jackson Tom, J., Jarrett Powers, Stephane Bélair, Aaron Berg, Paul Bullock, Andreas Colliander, et al. 2016. "Experimental Plan SMAP Validation Experiment 2016 in Manitoba, Canada (SMAPVEX16-MB)." https://smap.jpl.nasa.gov/internal_resources/390/.

Moran, M Susan, Luis Alonso, Jose F Moreno, Maria Pilar Cendrero Mateo, D Fernando De La Cruz, and Amelia Montoro. 2012. "A RADARSAT-2 quad-polarized time series for monitoring crop and soil conditions in Barrax, Spain." *IEEE Transactions on Geoscience and Remote Sensing* 50 (4): 1057–1070.

Moré, Jorge J. 1978. "The Levenberg-Marquardt algorithm: implementation and theory." In *Numerical analysis*, 105–116. Springer.

Nguyen, Duy Ba, Alexander Gruber, and Wolfgang Wagner. 2016. "Mapping rice extent and cropping scheme in the Mekong Delta using Sentinel-1A data." *Remote Sensing Letters* 7 (12): 1209–1218.

Pacheco, Anna, Heather McNairn, Yifeng Li, George Lampropoulos, and Jarrett Powers. 2016. "Using RADARSAT-2 and TerraSAR-X satellite data for the identification of canola crop phenology." In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*, Vol. 9998, 999802. International Society for Optics and Photonics.

Perez, Ruben E, Peter W Jansen, and Joaquim RRA Martins. 2012. "pyOpt: a Python-based object-oriented framework for nonlinear constrained optimization." *Structural and Multidisciplinary Optimization* 45 (1): 101–118.

Prevot, L, I Champion, and G Guyot. 1993. "Estimating surface soil moisture and leaf area index of a wheat canopy using a dual-frequency (C and X bands) scatterometer." *Remote Sensing of Environment* 46 (3): 331–339.

Ratha, Debanshu, Dipankar Mandal, Vineet Kumar, Heather McNairn, Avik Bhattacharya, and Alejandro C Frery. 2019. "A Generalized Volume Scattering Model-Based Vegetation Index From Polarimetric SAR Data." *IEEE Geoscience and Remote Sensing Letters* .

Satalino, Giuseppe, Anna Balenzano, Francesco Mattia, Michele Rinaldi, Carmen Maddaluno, and Giovanni Annicchiarico. 2015. "Retrieval of wheat biomass from multitemporal dual polarised SAR observations." In *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, 5194–5197. IEEE.

Steele-Dunne, S. C., H. McNairn, A. Monsivais-Huertero, J. Judge, P. Liu, and K. Papathanassiou. 2017. "Radar remote sensing of agricultural canopies: A review." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10 (5): 2249–2273.

Tuia, Devis, Jochem Verrelst, Luis Alonso, Fernando Pérez-Cruz, and Gustavo Camps-Valls. 2011. "Multioutput support vector regression for remote sensing biophysical parameter es-

timation." *IEEE Geoscience and Remote Sensing Letters* 8 (4): 804–808.

Ulaby, F. 1975. "Radar response to vegetation." *IEEE Transactions on Antennas and Propagation* 23 (1): 36–45.

Ulaby, Fawwaz T, Kamal Sarabandi, KYLE Mcdonald, Michael Whitt, and M Craig Dobson. 1990. "Michigan microwave canopy scattering model." *International Journal of Remote Sensing* 11 (7): 1223–1253.

Vapnik, Vladimir. 2013. *The nature of statistical learning theory*. Springer science & business media.

Veloso, Amanda, Stéphane Mermoz, Alexandre Bouvet, Thuy Le Toan, Milena Planells, Jean-François Dejoux, and Eric Ceschia. 2017. "Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications." *Remote Sensing of Environment* 199: 415–426.

Verrelst, Jochem, Jordi Muñoz, Luis Alonso, Jesús Delegido, Juan Pablo Rivera, Gustavo Camps-Valls, and José Moreno. 2012. "Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3." *Remote Sensing of Environment* 118: 127–139.

Wiseman, Grant, Heather McNairn, Saeid Homayouni, and Jiali Shang. 2014. "RADARSAT-2 polarimetric SAR response to crop biomass for agricultural production monitoring." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (11): 4461–4471.

List of Figures

List of Tables

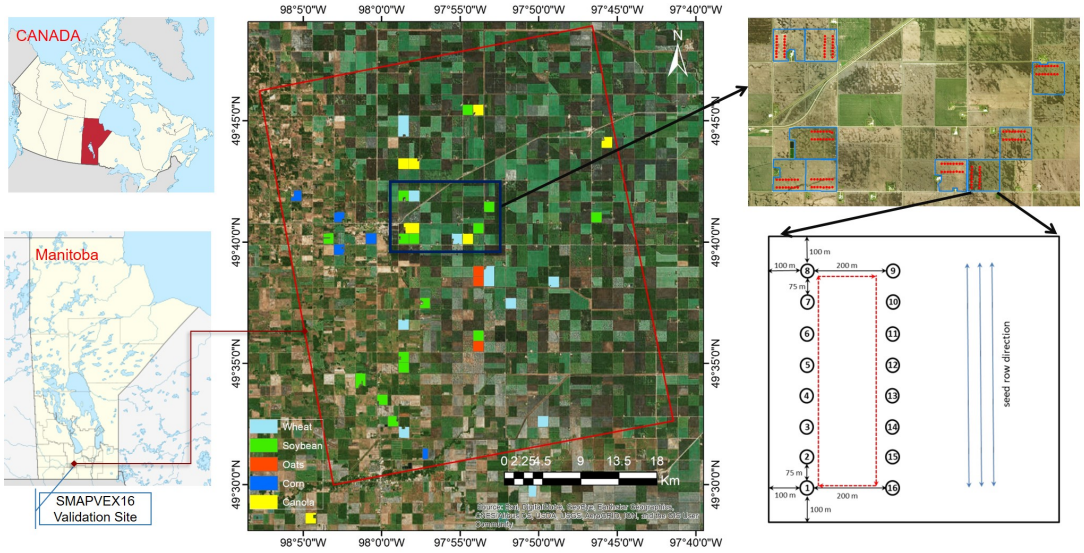

Figure 1. Study area and sampling locations over the JECAM-Manitoba (Canada) test site. The layout of the 16 sampling locations within each field is highlighted. In each field, 3 points were selected for vegetation sampling. Soil samples were collected from all the 16 sampling locations.

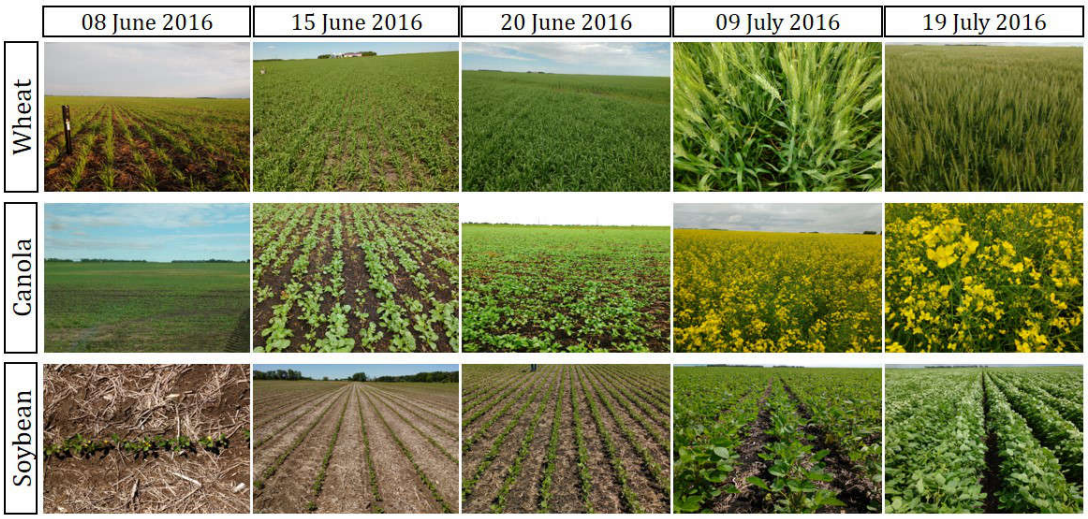

Figure 2. Field conditions of different crops during the campaign.

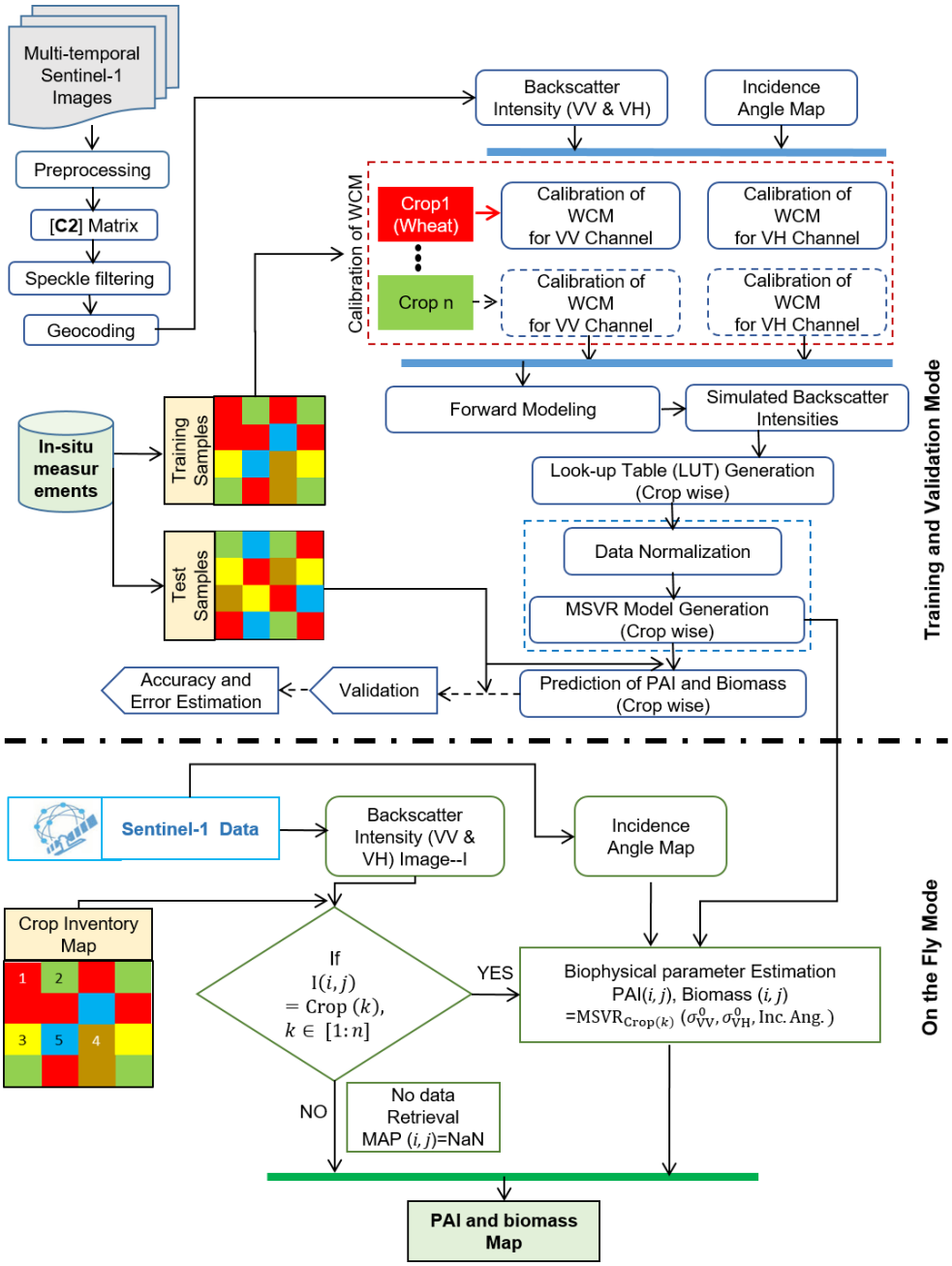

Figure 3. Schematic work-flow for retrieval of crop biophysical parameters using MSVR.

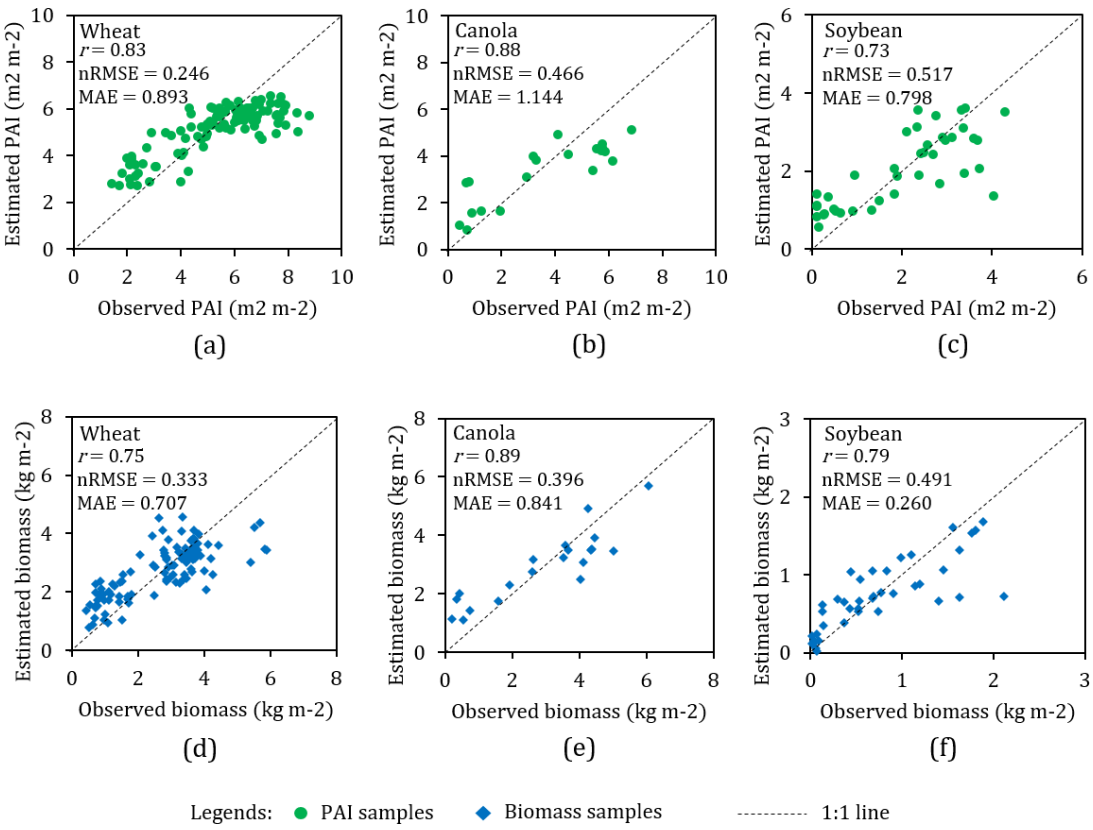

Figure 4. Validation plots of PAI ($m^2\,m^{-2}$) and wet biomass ($kg\,m^{-2}$) for wheat (a, d), canola (b, e), and soybean (c, f).

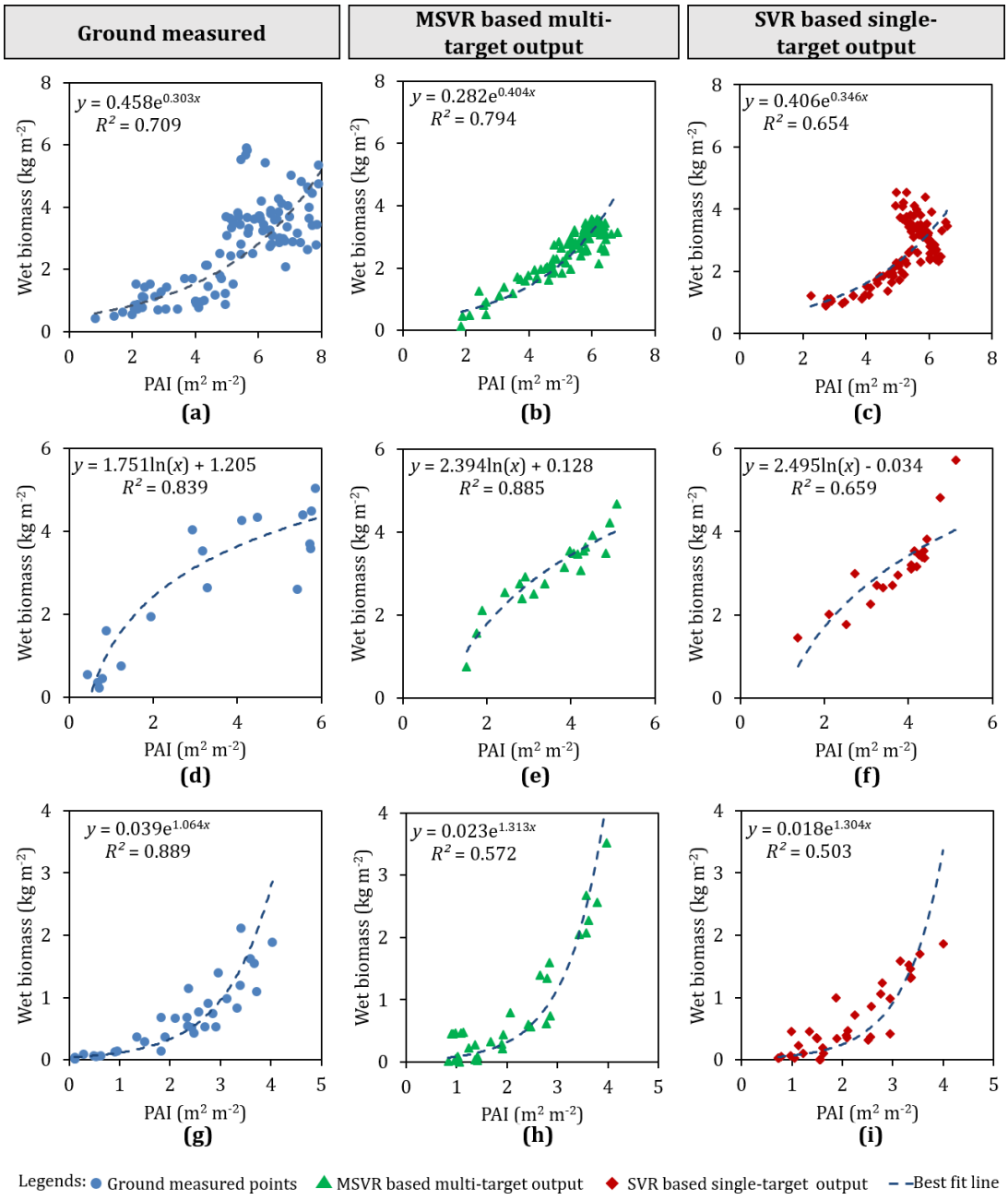

Figure 5. The relationship between PAI and wet biomass based on ground measured, MSVR, and SVR retrievals for wheat (a-c), canola (d-f), and soybean (g-i).

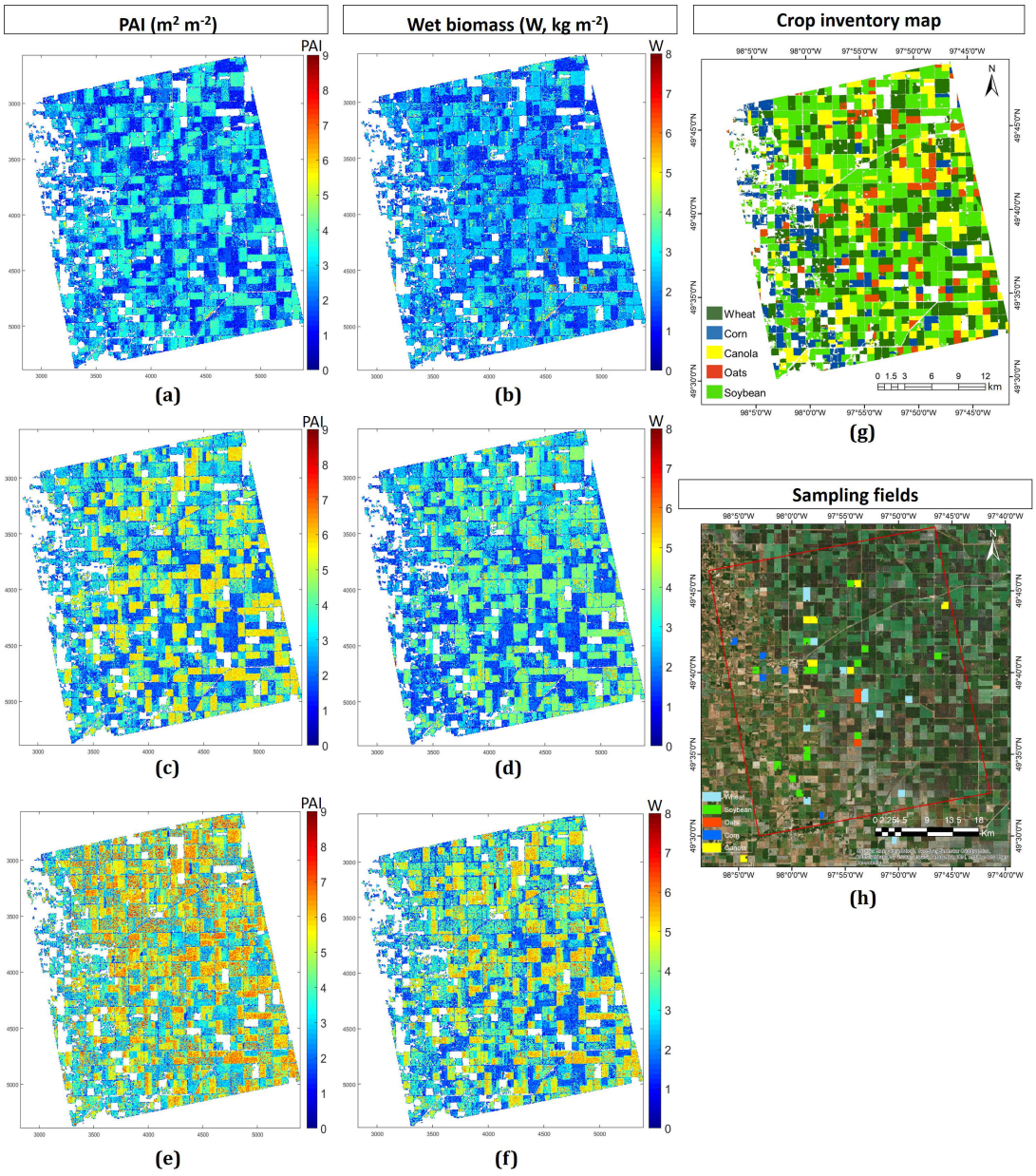

Figure 6. Plant Area Index (PAI $m^2\,m^{-2}$) and wet biomass ($kg\,m^{-2}$) maps over the Manitoba test site for three acquisitions dates, (a-b) 13 June, (c-d) 07 July and (e-f) 19 July 2016. The crop inventory map and sampling field locations are presented in (g) and (h) respectively.

Table 1. Sentinel-1A acquisitions over Manitoba (Canada) during the SMAPVEX16-MB campaign.

Satellite data acquisition date	Beam mode	Incidence angle range (°)	Orbit	Campaign window
06 June 2016	IW	39.87-41.84	Ascending	08 June 2016
13 June 2016	IW	30.22-32.47	Ascending	13 June 2016 15 June 2016
30 June 2016	IW	39.87-41.84	Ascending	27 June 2016 28 June 2016
07 July 2016	IW	30.22-32.44	Ascending	05 July 2016 06 July 2016
19 July 2016	IW	30.22-32.44	Ascending	17 July 2016 20 July 2016
24 July 2016	IW	39.82-41.79	Ascending	21 July 2016 22 July 2016

Table 2. The number of independent calibration and validation points for each crop combined from different growth stages.

Crop	Growth stage	Number of calibration feature point	Number of validation feature point
Wheat	Leaf development to fruit development	54	108
Canola	Leaf development to flowering	21	20
Soybean	Early leaf development to flowering	41	38

Table 3. Model parameters and statistics for different crops and polarizations

Crop	Polarization	Model parameter						F-Statistic	Level of significance
		A	B	C	D	E	F		
Wheat	VV	0.038	0.002	0.14	0.032	0.162	1.259	172	2.03×10^{-24}
	VH	0.024	1.283	0.026	-0.016	0.302	-1.18	173	1.46×10^{-24}
Canola	VV	0.282	-2.236	0.287	-0.018	0.004	-0.051	34.4	1.33×10^{-7}
	VH	-0.028	-0.027	0.039	-0.067	-0.516	0.412	39.1	5.84×10^{-8}
Soybean	VV	0.045	1.949	1.833	5.725	0.727	0.012	36.1	2.10×10^{-10}
	VH	0.099	0.062	-0.079	0.564	0.440	0.378	29.4	4.02×10^{-9}

Table 4. Comparison of simulated and observed backscatter (σ°) for the calibration data

Crop	VV		VH	
	r	RMSE	r	RMSE
Wheat	0.80	0.026	0.62	0.007
Soybean	0.64	0.034	0.68	0.006
Canola	0.87	0.051	0.83	0.010

Table 5. Comparison of MSVR and SVR based retrieval accuracy of PAI ($m^2\,m^{-2}$) and wet biomass (W, $kg\,m^{-2}$) for various crop types.

Crop	Biophysical parameter	MSVR			SVR		
		r	nRMSE	MAE	r	nRMSE	MAE
Wheat	PAI	0.83	0.246	0.893	0.75	0.306	1.012
	W	0.75	0.333	0.707	0.70	0.425	0.824
Canola	PAI	0.88	0.466	1.144	0.85	0.612	1.452
	W	0.89	0.396	0.841	0.81	0.486	1.054
Soybean	PAI	0.73	0.517	0.798	0.66	0.842	0.912
	W	0.79	0.491	0.260	0.70	0.674	0.387