

USING C-BAND SAR AND TEMPERATURE TO MONITOR TROPICAL AGRICULTURAL FIELDS

Cristian Silva-Perez

University of Stirling, Scotland, United Kingdom

Armando Marino

University of Stirling, Scotland, United Kingdom

Iain Cameron

Environment systems Ltd, Aberystwyth, United Kingdom

This manuscript presents the analysis and a methodology for monitoring asparagus crops from remote sensing observations in a tropical region, where the meteorological conditions change considerably between production cycles. We use data provided by the Sentinel-1 satellite and temperature from a ground station to show how particularly the VH polarisation can be used for crop monitoring in order to visualise the canopy formation, the growth rate and canopy biomass, revealing high dependencies on temperature. We also present a multi-output machine learning regression algorithm trained on a rich spatio-temporal dataset in which each output estimates the number of asparagus stems that are present in each of the pre-defined crop phenological stages. We present the results of two separate scenarios: Using a single SAR image plus temperature as input for the algorithm and using multitemporal SAR data. Results show that the methodology presented is able to retrieve each individual monitored variable when using temperature as predictor with coefficients of determination (R^2) above 0.85. Further research is currently investigating the added value of multitemporal SAR data to complement the predictions and potentially replace the temperature feature.

Keywords: Tropical agricultural monitoring, canopy development analysis, phenology retrieval, Sentinel-1, Multitemporal SAR, Multi-task machine learning.

Introduction

Accurate and continuous crop monitoring from satellite remote sensing in the tropics can be challenging due to high cloud coverage caused by water cycles and high precipitation levels. An additional distinctive operational characteristic in tropical and subtropical regions for several crop types is the uninterrupted production cycles, with more than one cycle per year. Each of these production cycles or campaigns may be under slightly different meteorological conditions due to a "soft seasonality", e.g. mild winters, thus modifying to a certain extent the crop growth rate and structure (as will be shown later in this paper) and therefore the response captured by the satellites at each campaign. This paper considers the above-mentioned conditions for agricultural monitoring, particularly analysing the case of asparagus crops in Peru. Given the importance of monitoring crop development during the campaign without interruptions, synthetic aperture radar (SAR) emerges as a potential technology for this task. The capabilities to acquire images at day and night and under nearly all-weather conditions of SAR satellites offer significant opportunities for systematic monitoring regardless of cloud coverage (Jong-Sen Lee).

In this context, the contributions of this manuscript are the analysis of the SAR sensitivity to the asparagus crop development and its potential for monitoring at a level that to our knowledge has not been reported before for this crop-type. We also present an example of how the seasonal climatological conditions influence the crop development, how this is experienced in tropical regions and how the radar sensor is sensitive to these conditions.

Finally, we present the implementation of a data-driven methodology that captures the recurrent patterns in the SAR response and the climate conditions to estimate a proxy of crop development at every new SAR acquisition. It consists of a Multi-output regression machine learning model that estimates the number of stems that have reached or are present

in each of the five pre-defined phenological stages.

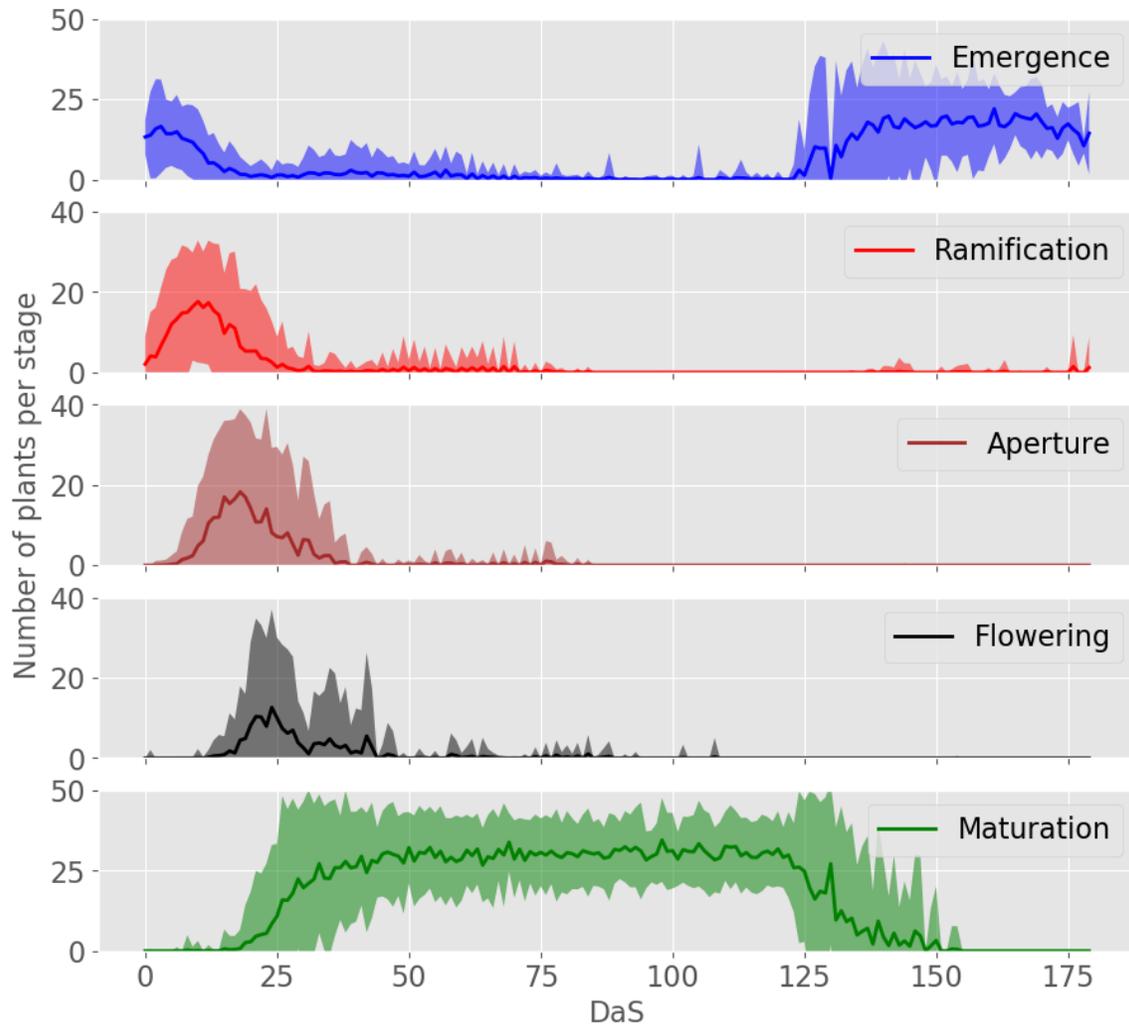


Figure 1. Number of stems in each crop stage as a function of the number of days after the season started (DaS). The solid lines represent the averages of all the ground measurements collected in 440 parcels during the first 8 months of 2019. The shaded regions represent plus and minus two standard deviations.

Test site, ground truth and SAR datasets

Test site and Climatological conditions

The asparagus fields are located in the north of Peru in a dry coastal zone with sandy soil, where approximately 442 parcels with sizes in average of two hectares have been considered in the analysis. The parcels are located at the sea level with temperatures that vary between 16 degrees Celcius in winter and 26 degrees in summer, being more humid during the summer season.

Ground truth

A survey campaign was carried on in the field covering a period of 8 months, from January to August of 2019. Surveyors randomly selected two transects to assess, with each transect being a metre in length. Within each transect the surveyors identified the phenological stage of each stem and counted the total number of stems in each of the following stages: Emergence, ramification, aperture, flowering or maturation. Results for the two transects were averaged and recorded. In this way, the proportion of stems per parcel in any of the phenological stages can be recorded, and a proxy for the evolution of crop stage over time can be established. The average temporal evolution of asparagus growth stages is shown in Figure 1 as a function of the number of days after the cultivation started (DaS) for all the production cycles covered in the ground truth.

SAR datasets

The Sentinel-1 dataset used for the analysis was built using the Level-1 Ground Range Detected (GRD) and the Interferometric Wide swath (IW) acquisition mode, with VV and VH polarisations. The data was collected using the Google Earth Engine platform (2) in which the data pre-processing steps include orbit file, GRD border noise removal, thermal noise removal, Radiometric calibration and Terrain correction. Figure 2 shows the average evolution the backscatter as a function of the number of days after the cultivation started (DaS) for all the production cycles covered in the ground truth.

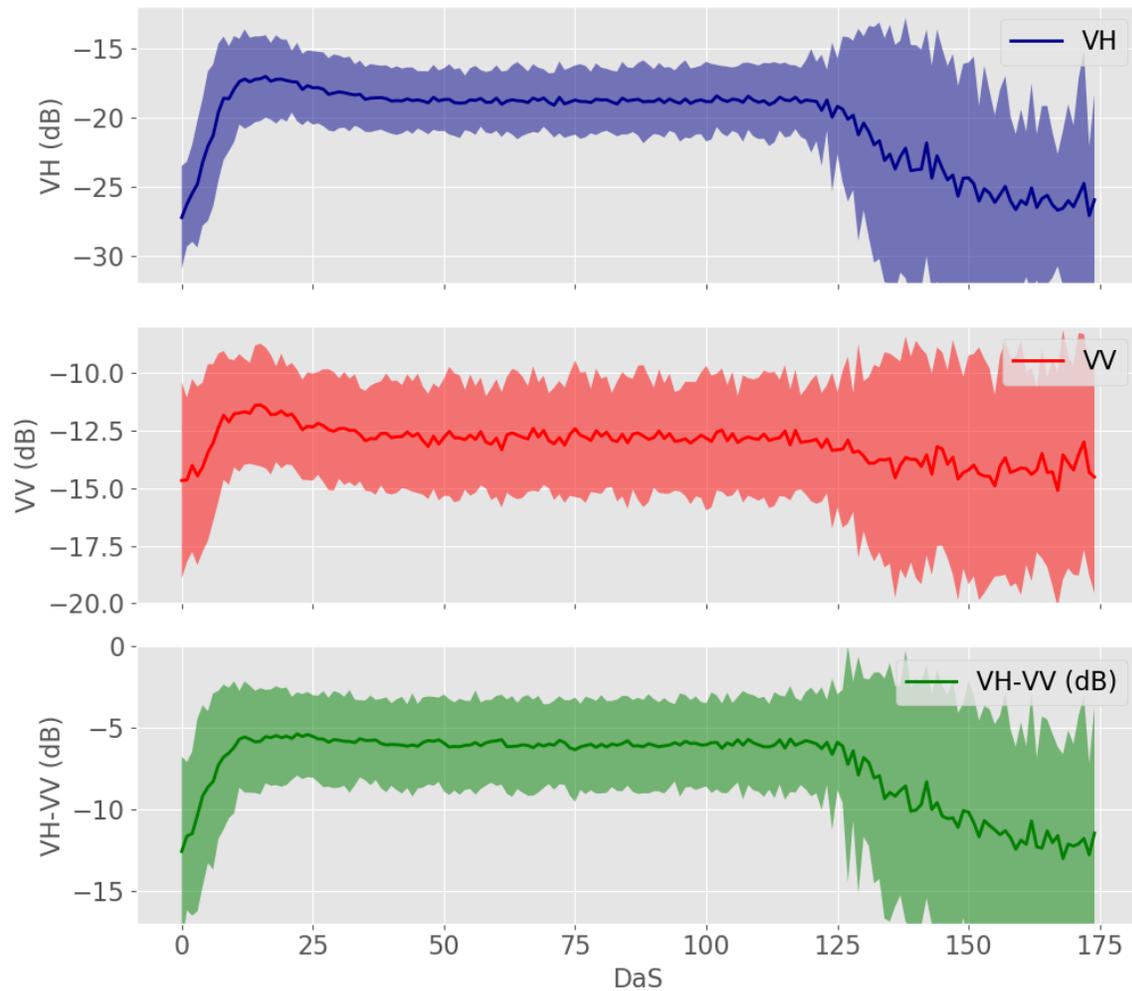


Figure 2. Sentinel-1 backscatter evolution as a function of the number of days after the start of the production cycle (DaS). This chart shows the mean backscatter plus/minus two standard deviations of 440 parcels).

SAR sensitivity to crop evolution

SAR response to crop evolution

Due to the mild winters with minimum temperatures of around 16 degrees Celsius, the plants do not reach senescence naturally and it is possible to have asparagus production

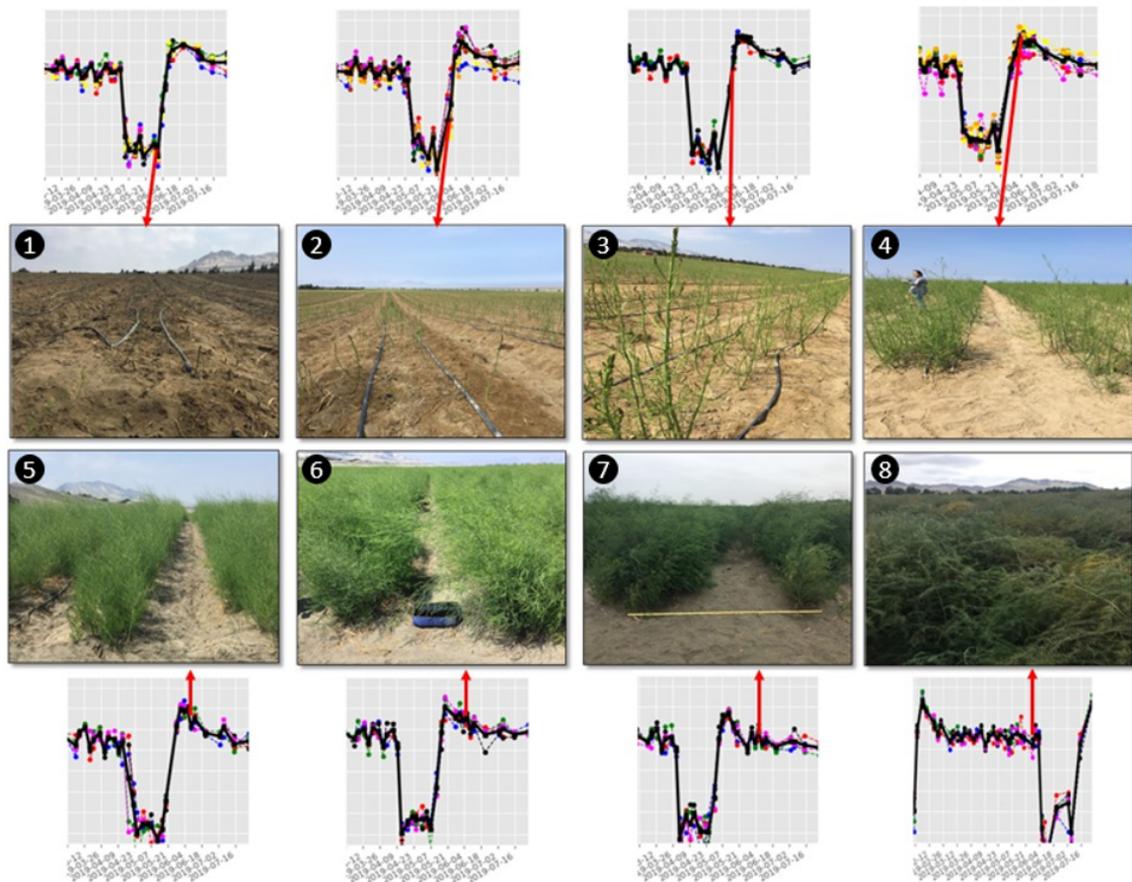


Figure 3. Images taken in the fields with the corresponding SAR backscatter time series at the date of acquisition. To be read by rows from left to right.

at any time of the year with plots in all phenological stages simultaneously. This allows growers to have two productive cycles per year with starting and ending dates driven by agri-climatological conditions but and management factors i.e. market needs and contractual planning. Given this, at a single SAR acquisition there are plots at almost all possible crop stages and we can reconstruct an approximation of the SAR response to the entire season.

The photographs shown in figure 3 were taken at the same date as a SAR acquisition (28/05/2019). Based on this, we can locate in the time series of each plot, what the SAR response to the crop is and compare it with the crop status recorded in the footage.

In the images 1 and 2 of figure 3 we can see that the lowest backscatter of the season is present during the harvest periods, since the fern has been mechanically removed and the new emerging spears are being harvested. At this point the SAR signal interacts only with the soil with a predominant surface scattering mechanism. Both polarisations, but particularly the VH polarisation increase significantly as the asparagus spears start emerging and vertically elongating. This increase is presumably a consequence of the double bounce created with the SAR signal reaching the soil and bouncing in the vertical spears back to the satellite (images 3 and 4 of figure 3). However, when the crop reaches approximately the peak of the aperture stage (see figure 1), the backscatter also reaches the peak in the entire cultivation period. At this point, the fern already have developed branches and the leaf-like structures in the stems are developing. From this moment, the contact of the SAR signal with the soil decreases thus also reducing the backscatter measured.

Subsequently, at the flowering stage as shown on image 5 of figure 3, the fern is fully developed and more dense covering the soil and creating a volume scattering response of less intensity, causing a decrease in the backscatter. Once the the crop has reached a mature stage (see figure 1), no significant changes happen in the biomass of the canopy hence the SAR signal remains at approximately the same level until the end of the season.

Impact of temperature on the crop and the SAR response

Based on empirical observations, growers have noticed in the field that the crop evolution during a "warm" and a "cold" campaign in the same year are different, in terms of canopy volume and development rate, possibly due to the different climatological conditions. To analyse this effect we used Sentinel-1 imagery, ground truth measurements and accumulated growing degree days.

Three effects of the temperature on the crop evolution were identified (images not shown here due to space restrictions):

1. The first one, corresponds to the canopy volume developed, being less biomass during hotter temperatures (e.g. summer campaigns). The SAR backscatter signal is sensitive to this by measuring lower backscatter intensity during the maturation periods.

2. The second effect is associated with the growth rate, since temperature causes the stages at the beginning of the season to develop faster in summer campaigns. This effect is also visible in the temporal trajectory described by the backscatter response.
3. The third effect is related to the season length depending on the accumulated temperature during the cultivation period. This accumulated temperature in turn depends on the month of the year when the campaign started.

Estimation of number of asparagus stems in each crop stage

This section presents the methodology used for monitoring asparagus development as described by the number of asparagus stems in each of the five stages of figure 1.

Note that the ground truth used for training corresponding to phenological information is given by multiple and correlated variables (figure 1), that is, variables that do not evolve in time independently but rather they have temporal co-variation since they are produced by the same underlying process, i.e. the crop growth. To exploit this structure in the output data, we use a multi-output regression algorithm that considers this dependence of the individual outputs before making predictions. In this context, multitask learning (MTL) has been used in several applications precisely with this objective (1).

Model development

We chose initially a multi-task Random Forest Regressor (4) from the available MTL algorithms due to its power for capturing non-linear relationships in the data and the fact that it provides importance analysis of each of the input features in which is based the model decision-making. Several scenarios have been considered in this paper as follows.

Inputs

An initial scenario was considered without using remote sensing but only temperature and cultivation days in order to analyse the added value of the Sentinel-1 data (scenario A). The scenario B uses a single SAR image, including the VH and VV polarisation channels with the corresponding ratio between them and the scenario C uses both the crop accumulated growing degree days (temperature) as well as SAR data.

Table 1: Summary of correlation coefficients for each predicted crop status output for the scenarios A to D.

Stage	A	B	C
Emergence	0.9	0.65	0.86
Ramification	0.9	0.14	0.91
Aperture	0.87	0.16	0.83
Flowering	0.75	0.05	0.76
Maturation	0.95	0.54	0.93
Overall	0.87	0.31	0.86

Training and testing data and Model hyper parameters

In order to separate training and testing data, we randomly select plots based on the ID to complete approximately 70% of them in a training dataset and 30% in a testing dataset. Tuning of the optimal model hyperparameters was done using 5-fold cross-validation with grid search.

Results

Table 1 presents the model performance based on the coefficient of determination (R^2) between the predicted and test data-points. It can be seen that in general the model has satisfactory predictive capabilities with overall coefficients of determination R^2 above 0.85 for scenarios A and C whereas the R^2 for the scenario B which uses a single SAR image, only achieves 0.31. Particularly for the scenario C, the model has the best performance predicting number of stems in maturation stage achieving 0.93 and flowering the lowest accuracy with 0.76.

Prediction maps

Figure 4 shows for all the parcels in the test site, the estimation of asparagus stems present in each phenological stage, obtained using the trained multi-task random forest of the scenario C, for the Sentinel-1 image acquired the 2018/10/12. Figure 5 shows for 5 different Sentinel-1 acquisition dates, an RGB composite of the same information using the predicted asparagus

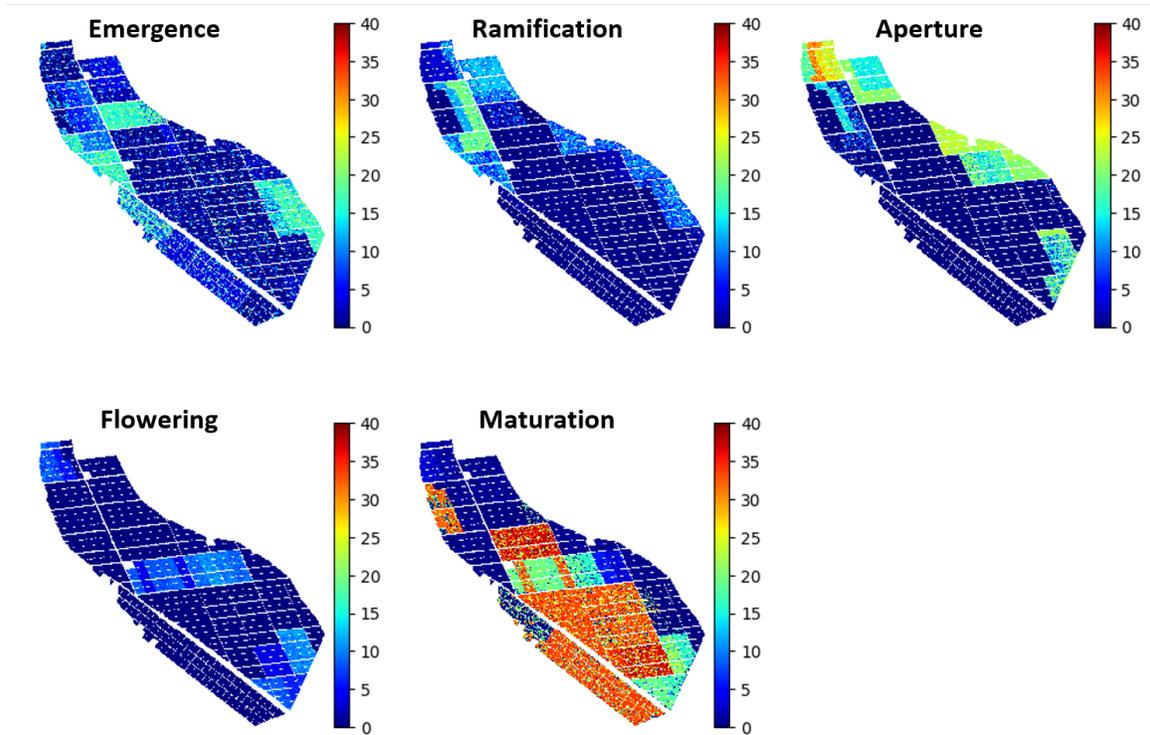


Figure 4. Number of asparagus stems estimated for each of the crop stages for the 2018/10/12 Sentinel-1 image. The RGB composite of this image is the top-right plot of figure 5

stems in emergence in the blue channel, the predicted asparagus stems in maturation in the green channel and the sum of the predicted asparagus stems in ramification, aperture and flowering in the red channel (given their short duration). Note that five different dates are shown to provide intuition of the crop evolution during this time. This composites reveal the crop stage of each parcel in an intuitive and fast way while the number of asparagus stems predictions map of figure 4 shows more detailed information for every individual crop stage.

Conclusions

In this abstract we provide an interpretation of the SAR backscatter response to asparagus crop growth and analyse the impact that temperature conditions have on the canopy

volume, its development rate and the cultivation length. We show how the sentinel-1 satellite is sensitive to all these effects. Similarly, we provide a multi-output machine learning regression algorithm trained on a temporally and spatially rich dataset to retrieve an approximation of crop stage. We tested the importance of using temperature as crop growth predictor as complement to the backscatter.

Acknowledgement

This research was funded by the Project EO4cultivar, led by Environment Systems Ltd and co-funded by the UK Space Agency (<https://www.envsys.co.uk/projects/eo4cultivar/>).

References

- [1] Caruana, R. (1997). Multitask learning. *Machine Learning*, 28(1):41–75.
- [2] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*.
- [Jong-Sen Lee] Jong-Sen Lee, E. P. *Polarimetric Radar Imaging: From Basics to Applications*.
- [4] Segal, M. and Xiao, Y. (2011). Multivariate random forests. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1):80–87.

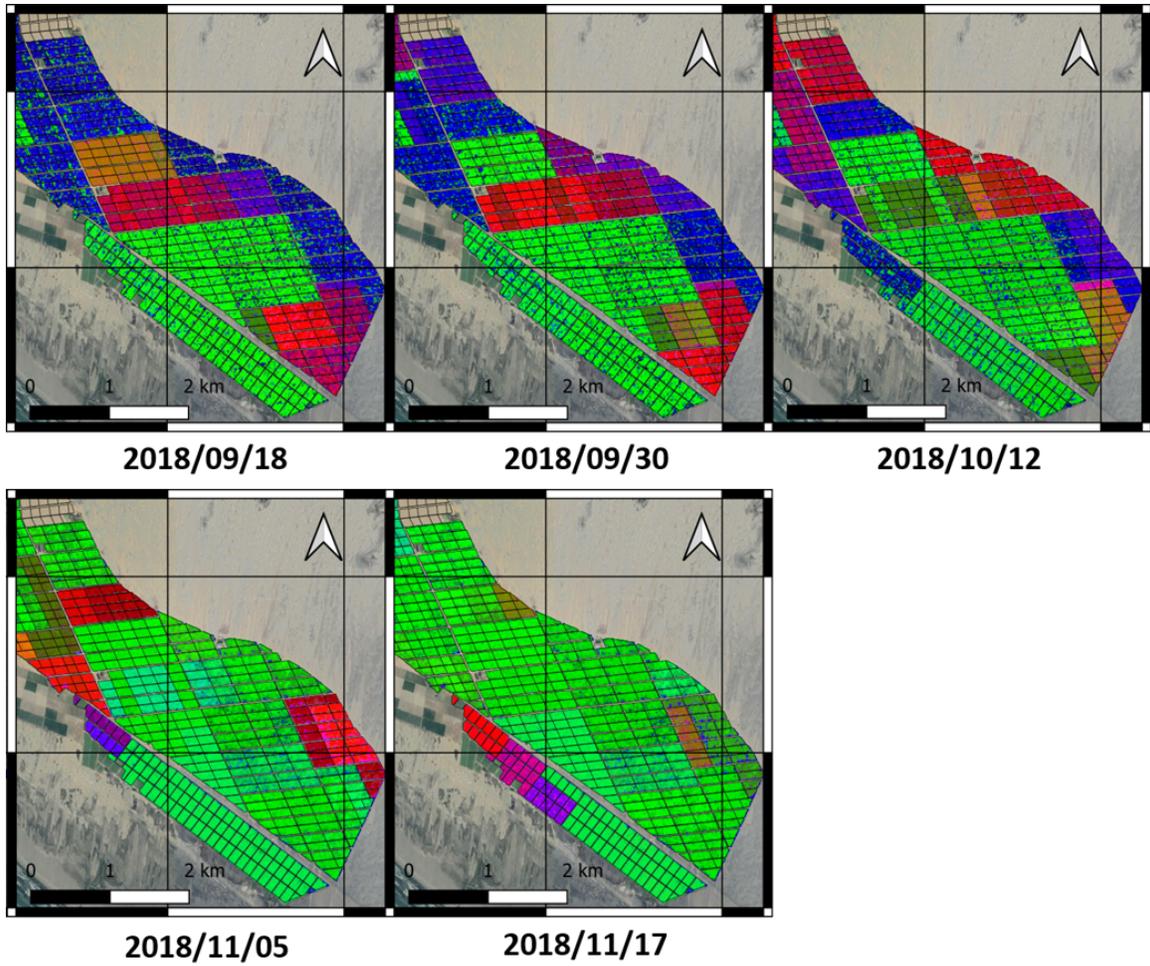


Figure 5. RGB composites of the estimated crop stage. Red: (Ramification+Aperture+flowering), Green: Maturation, Blue: Emergence.