Questions & Answers Part 3

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Sean McCartney (sean.mcartney@nasa.gov), Erika Podest (erika.podest@jpl.nasa.gov), Heather McNairn (heather.mcnairn@AGR.GC.CA), Emily Lindsay (emily.lindsay@AGR.GC.CA), or Xianfeng Jiao (Xianfeng.jiao@agr.gc.ca)

Question 1: Could you explain how this data was collected? https://appliedsciences.nasa.gov/sites/default/files/2023-04/Carmen_CornForRegression.csv.
Answer 1: The "Corn.csv" file contains the extracted pseudo-polarimetric variables derived from Sentinel-1 imagery, as well as a corresponding NDVI derived from Sentinel-2 data. The file can be used to run the Random Forest Regression.

Question 2: How can we distinguish crop and forest types using SAR data? Can any indices available like a Radar Vegetation Index (RVI), etc. or any polarimetric combination be used to do so?
Answer 2: We have provided a previous ARSET training on the use of SAR to classify crop types. Although we did not specifically address land cover classification (for example, crops versus forest) during that training, the same principles can be applied. A time series of SAR will be very important in any classification. We have found that in addition to a dense time series of SAR, polarimetric and multi-frequency SAR can be very helpful in classifications. The role of polarimetry is still evolving in the research but we have found, for example, that the 4 Stokes vectors are very helpful in classification. In terms of frequency, it would be very helpful to integrate L-band with C-band given the biomass differences between crops and forests. Look towards access to NiSAR L-band as integration with Sentinel-1 will be very helpful to this type of classification.

Question 3: Please help me to interpret any polarimetric decomposition result generated using PolSARPro. My doubt is its (e.g., volumetric) ranges in 0 – 255 like 8 bit data. How do you convert it into dB scale or linear as generated in SNAP?

Answer 3: If you are getting a data range of 0–255, please review your process chain. If processing is done correctly, the polarimetric parameters calculated in PolSARPro should be in db. Please contact us for further clarification if needed.

Question 4: What are the limitations of SAR for agriculture and water resources monitoring? Is SAR Data available for researchers?
Answer 4: The answer depends on what you mean by limitations. In terms of data access, look towards Sentinel-1 as these data are free and open, and easy to access. Access to Sentinel-1 will be covered in the hands-on section of the webinar. Data from other sensors will also be available in the future, including L-band from NiSAR. If you have other questions regarding limitations, please reach out to our team at AAFC.

Question 5: Is it possible to estimate crop phenology only from a SAR image?
Answer 5: Yes. Our team has collaborated with a Canadian company who has been able to estimate the phenology of crops (corn, canola, soybeans, and wheat) using Growing Degree Days, and SAR (C- and X-band). Please refer to: McNairn, H., Jiao, X., Pacheco, A., Sinha, A., Tan, W., and Li, Y. (2018). Estimating canola phenology using synthetic aperture radar, Remote Sensing of Environment, 219: 196-205. The same method has been applied to these other crops and a publication is in preparation.

Question 6: Please provide some insights into dual pol Entropy Alpha and full pol Entropy Alpha.
Answer 6: Cloude developed a Dual pol entropy/Alpha decomposition method, which can be applied to Sentinel-1 SLC data. It is different from entropy/anisotropy/alpha decomposition for quad pol data developed by Cloude-Pottier. From our experience these polarimetric parameters (from dual or QP) are correlated, but values may not be exactly the same/interchangeable. Evaluating the difference between QP and dual pol derived parameters is still something we are investigating and this is why we chose to refer to dual pol parameters as “quasi” or “psuedo” to remind us that they are not equivalent/interchangeable to those derived from QP.

Please refer to these papers,

“THE DUAL POLARISATION ENTROPY/ALPHA DECOMPOSITION: A PALSAR CASE STUDY”

S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 2, pp. 498-518, March 1996, doi: 10.1109/36.485127.

Question 7: Can this workflow also be used for a mixed cropping or grasslands with mixed grassland field?
Answer 7: The workflow can be applied to these land covers. However, you will need to create your own calibration equation for these particular cropping systems and test how well a SAR VI would work in assessing the condition of these cover types. In addition, there may be limitations to using C-band SAR wavelength for other land cover types. Additional SAR wavelengths (i.e., X or L Band) may yield better results for different canopy structures of grassland or mixed annual crop types, etc. We started our research with canola, but have found in subsequent research that this method works well for other crop types (corn, soybeans, wheat).

Question 8: Is there a different processing channel for retrieving SAR data from a flooded paddy field? How is it different from collecting SAR data from dryland crops?
Answer 8: The same workflow can be used. However, you will need to develop a calibration specific to flooded paddy fields, and test to see how well a SAR VI can estimate rice condition.

Question 9: I would like to know if the polarimetric data can also be used in small tillages (0.5–2.0 ha) to discriminate between crops and small tillages with multiple crops.
Answer 9: This would be an interesting research project. In theory, the scattering that comes from vegetation is very different than scattering from bare soils (vegetation will have a larger presence of volume scattering and greater entropy, as examples). Our team is also looking at phase information and using coherent change detection to flag when fields are tilled. This has been successful because the phase (distance from sensor to field) changes when the height of the soil changes due to tillage activities.

Question 10: I imagine wind can affect the radar backscatter return. How do you differentiate genuine changes in crop structure from changes caused by wind?
Answer 10: This is an interesting observation. We know, for example, that phase can change due to movement of the crops by wind. Our research, however, has demonstrated that there is a strong correlation between phenology (which is very

dependent on structure) and polarimetric parameters. Although wind may create some change in how the SAR sees the canopy, polarimetric parameters are overwhelmingly sensitive to overall structure. It may be, however, that some of the errors in our calibration equation are due to this type of effect. We had a recent paper that may be of interest as it looked at the relationship between crop structure/phenology and polarimetric responses. Dingle Robertson, L., McNairn, H., Jiao, X., McNairn, C., and Ihuoma, S.O. (2022). Monitoring crops using compact polarimetry and the RADARSAT Constellation Mission, Canadian Journal of Remote Sensing, doi: 10.1080/07038992.2022.2121271.

Question 11: What does the VV - decimal gamma0 and VV - linear gamma0 mean in Sentinel-Hub EO. How can we get the interferogram and understand the heights?
Answer 11: We did not use interferometry to get vegetation height for this project. We are calibrating the SAR-derived polarimetric variables to Sentinel-2 NDVI to estimate crop condition, not crop height.

Question 12: Can SAR data be used to identify a spreading of invasive species among the crop?
Answer 12: We have not examined this, but in theory, this may be possible. If the invasive species has a structure different from the crop and the spatial extent of the invasive species is significant enough, this may be possible.

Question 13: Originally, NDVI is an optically derived index based on NIR and Red wavelengths. How can a SAR equivalent index be correlated with optical NDVI?
Answer 13: The link is through the canopy structure and the link between structure and biomass/LAI. Both optical reflectance and SAR scattering are correlated with LAI and biomass. We described this in detail in the introduction of the following paper: Jiao, X., McNairn, H., and Dingle Robertson, L. (2021). Monitoring crop growth using a canopy structure dynamic model and time series of Synthetic Aperture Radar (SAR) data, International Journal of Remote Sensing, 42:6437-6464, doi: 10.1080/01431161.2021.1938739.

Question 14: If we use SAR for crop differentiation (i.e., crop mapping), will different stages of their development confuse the model?
Answer 14: It can be the case. This is why a time series of SAR is important. However, you can also think of these changes in the growth stage as helpful. These changes in

canopy structure can be informative to the SAR in differentiating crops. As with optical sensing, a crop may look similar to the SAR at one point in time (because the structure is the same) but as the structure develops, this helps the SAR separate crops.

Question 15: What were the differences you found in accuracy between Artificial Neural Networks (ANN) and Random Forest?
Answer 15: These results are described in this paper: Jiao, X., McNairn, H., Yekkehkhany, B., Dingle Robertson, L., and Ihuoma, S. (2022). Integrating Sentinel-1 SAR and Sentinel-2 optical imagery with a crop structure dynamics model to track crop condition, International Journal of Remote Sensing, 43:6509-6537, doi: 10.1080/01431161.2022.2142077. The correlations (R2) were very similar (0.869 versus 0.888). As such, either could be used. We settled on RFR as it was simpler for us to implement. However, this is still an area of active research and as more advanced Machine Learning algorithms are developed, these should be tested as improvements may be possible.

Question 16: What can be the possible range of a particular scattering type (e.g., for an agricultural area)? Having Full Pol data, what is the possible range of each Yamaguchi 4 Component Decomposition?
Answer 16: We would need this question rephrased in order to answer it. We are unclear what scattering type you are referring to. Please reach out to our team to clarify your question.

Question 17: How do you evaluate the performance of window size using any tool instead of visualization only?
Answer 17: Window size of any tool can be tested by implementing the workflow for 2-3 different window sizes, and compare the results of the random forest regression. If there is a statistically significant difference in R^2 the best window size can be used. Remember that the window size is very specific to your site and cropping system. The size we might use in our sites in Canada, may or may not be appropriate for your site.

Question 18: The demo shows a great way to process a single image and uses several different software packages to do so (SNAP desktop, PolSARPro, Python/Google Colab, and MATLAB). What would be the best way to create a single continuous workflow/pipeline for processing large amounts of data?
Answer 18: This is a great question. In this case, there are strengths to each software selected for each processing stage as we were in the research stage of the project. We

are exploring new processing methods to simplify and implement this work-flow in operational contexts. There is ongoing work with PolSARPro and SNAP to develop a dedicated bridge between the two software packages, but was not operational at the time of this research. Batch processing is also available in both SNAP and PolSARPro. You can use python with SNAP and batch process in PolSARpro to speed up and automate the workflow.

Question 19: Is PolSARPro open source software?
Answer 19: It is an open source software, can be downloaded here: https://ietr-lab.univ-rennes1.fr/polsarpro-bio/

Question 20: Is there any similar method for Normalized Difference Moisture Index (NDMI; e.g., soil moisture) using both Sentinel-1 and Sentinel-2?
Answer 20: We have not tried this. It is likely that the polarimetric parameters that would be helpful for soil moisture, would be different from the ones we used for SAR VI.

Question 21: When we randomly split training and test datasets, how would you manage spatial autocorrelation?
Answer 21: Spatial autocorrelation was managed by selecting a single sub-field object for each “Field” of each crop type in the reference dataset itself. This reduces any bias linked to spatial autocorrelation.

Question 22: Considering that everything is related to the analysis of vegetation cover, I wonder if you could use PolSARpro to analyze other types of cover, for example, buildings?
Answer 22: Our analysis is focused on crop type, but PolSARpro would be a good tool to test the application of SAR to map other land cover types. However, scattering parameters are very different among land cover types. Polarimetry offers very rich information. The way a wave scatters from a building versus other land use types is very different and could be exploited for this application.

Question 23: I have a question about Sentinel-1 data selection. Do you select only dry dates and filter with rainy dates in order to avoid the surface moisture in the analysis?
Answer 23: Yes we always check the meteorological conditions before we use SAR data. If rain is occurring at the time of acquisition, the scattering from raindrops in the

atmosphere usually makes the SAR image unusable. Even heavy rain before a SAR acquisition, will change the backscatter due to water present on the canopy. As a matter of good practice, you should investigate the weather conditions at the time of the SAR acquisition and remove those that occurred during rain events.

Question 24: How do you calculate cumulative Growing Degree Days (GDD)?
Answer 24: GDD is calculated using a local weather station where we can download temperature data. But there may be other sources of temperature data that you can investigate. The formula is given in the webinar (slide 27):

$$GDD = (T_{max} + T_{min})/2 - T_{base}$$

Then accumulated GDD is calculated by summing GDDS for each day.

Question 25: Could non-timber products be classified the same as crops?
Answer 25: Yes this would be possible to use non-timber land cover types with this methodology, however, at this time, the research is limited to the 4 crop types tested.

Question 26: Are there ready-to-run standard cloud-based resources to process the NDVI and SAR data in this manner together? If not, do you know of any efforts under development?
Answer 26: We are unaware of any cloud based resources available at this time.

Question 27: Is it possible to do the preprocessing and/or the parameters derivation solely in Python without SNAP and PolSARPro? Are there any open-source libraries or scripts?
Answer 27: To my understanding, SNAP can be run in a python environment. There is currently no substitute for PolSARPro for advanced processing of polarimetric data, but a bridge is being developed to implement PolSARPro in SNAP/Python.

Question 28: If I understood correctly SAR-NDVI is crop dependent? If so, how does the model account for crop type?
Answer 28: Yes, SAR-NDVI is calibrated to each crop type using the crop structure dynamics model. However, we are also assessing a global calibration model. To date, we have seen that the global model is less accurate. To implement a crop specific model, would require knowledge of crop type early in the season. We believe it may be

possible with further testing and training, that the calibration of a global model may improve.

Question 29: After identifying the highly correlated parameters from SAR, can we proceed with utilizing just those parameters for SARcal-NDVI?
Answer 29: Yes, a variable reduction strategy can be used to remove variables from the model that do not have a high variable importance. This may result in improved accuracy of subsequent Random Forest Regression scenarios by reducing the correlation and dimensionality, and potential for overfitting the model. For more information about the Variable reduction strategy tested please see the following Publication: Jiao, X., McNairn, H., Yekkehkhany, B., Dingle Robertson, L., and Ihuoma, S. (2022). Integrating Sentinel-1 SAR and Sentinel-2 optical imagery with a crop structure dynamics model to track crop condition, International Journal of Remote Sensing, 43:6509-6537, doi: 10.1080/01431161.2022.2142077. The correlations (R2) were very similar (0.869 versus 0.888)

Question 30: Can we process two bursts at the same time while running split?
Answer 30: Yes, this step is optional if you want to reduce the area of the Sentinel-1 SLC image to a smaller area, or you may not wish to do this because your study area is large or you do not need to cut down on processing time. Anywhere from 1-9 bursts can be included in this step.

Question 31: How does a SAR based vegetation index compare with NDVI in terms of accuracy? Are there any limitations to take note of?
Answer 31: The SAR-calibrated VI for corn and canola had correlations (R2) greater than 85% when compared to NDVI. We have these results for our 4 crop types and a publication is forthcoming. We are also linking the SAR VI to biomass measured in fields, and for example, correlations (R^2) between SAR VI and biomass for canola were 0.88 until peak biomass and then 0.42 during the period of senescence. Please refer to Jiao, X., McNairn, H., Yekkehkhany, B., Dingle Robertson, L., and Ihuoma, S. (2022). Integrating Sentinel-1 SAR and Sentinel-2 optical imagery with a crop structure dynamics model to track crop condition, International Journal of Remote Sensing, 43:6509-6537, doi: 10.1080/01431161.2022.2142077.

In terms of limitations, more research is needed to test the robustness of these methods over different growing seasons and cropping systems. A challenge will be to develop a universal/global model that could be applied to all crop types. Otherwise

knowledge of crop type would be needed early in the season in order to select the crop-specific model. Note that our next step is to test the value of the SAR VI in our operations to monitor crop yield.

Question 32: Is there any open source alternative available to MatLab to do the SARcal-NDVI to Canopy Structure Dynamics Model (CDSM) part from this exercise?
Answer 32: Yes, The Crop Structure Dynamics Model calibration equation could be implemented in Python if needed.

Question 33: Are the SAR-NDVI calibration coefficients published somewhere?
Answer 33: Not yet, but please reach out to us. Note however, that more testing would be needed to determine if our calibration coefficients would work for your site/crops. It is likely that some adaptation would be needed and perhaps a new calibration would need to be developed for your site.

Question 34: What's the optimal number of SAR data points needed to get accurate Canopy Structure Dynamics Model (CDSM) predictions of the temporal NDVI?
Answer 34: We used 500 segments (objects) to train our model. However, the exact optimal number is not known. Recall that we are not creating the model using pixels, but rather objects (collection of pixels created by segmentation). What is also important is to ensure that the data used to create the model cover the range of crop development (from beginning to end of the growing season).

Question 35: Is it possible to fit the SARcal-NDVI to CSDM step in the processing methodology – or in other software (e.g., Python or R)? Why do you use MATLAB?
Answer 35: See answer to question 32.

Question 36: Are the weekly Crop Condition Assessment Program (CCAP) products available for download or as a consumable REST service?
Answer 36: CCAP products are downloadable. See this url:

To see these open products, https://geoprod.statcan.gc.ca/ccap/en/index

Question 37: Having the SAR-NDVI, can we replace the Backscatter with NDVI for these crops? And use it for any other health assessment?

Answer 37: Backscatter is sensitive to incidence angle, soil moisture, early morning dew on the plants, and other background conditions. As such, polarimetric parameters which describe the characteristics of the scattering may be more robust and less sensitive to these other site factors, when compared to backscatter intensity. As well, our initial research created a simple statistical model between SAR parameter and NDVI. Subsequently we developed a machine learning model which allowed us to integrate multiple SAR parameters in the calibration. This improved results because various SAR parameters add different information on how the canopy is developing. Given these results, using multiple SAR responses (perhaps backscatter intensity as well as other SAR scattering parameters) will yield the best results.

Question 38: What would be a likely scenario of satellite crop mapping with the help of Artificial Intelligence (AI) / Machine Learning (ML) in 5-10 years time? What research focus should we dig into now to make sure the future of Satellite Crop Mapping keeps on getting better?

Answer 38: Advances are occurring with respect to technology and methods. SAR satellites (current and future) offer fully polarimetric or quasi polarimetric modes and as demonstrated in this webinar, a great deal of important information about how crops develop, can be extracted from these complex data. As well, integrating data from satellites that collect data in different SAR frequencies (X +C + L) will also advance the use of SAR for agriculture. It is a very exciting time to be involved in SAR research. Finally, algorithm development is advancing and we have found in our own research that AI and machine learning methods can yield exceptionally good results if trained with robust data. This should encourage the international research community to come together and share satellite and field data to train these models.

Question 39: In the SAR-NDVI model, is the SAR data in linear scale or in dB scale? Will there be any effect if we use them interchangeably?

Answer 39: We have not tested this, In theory there should be minimal effect if one or the other calibration type is used. We recommend using db, and recommend avoiding using them interchangeably.