Combining time series of Sentinel-1 and -2 with in-situ data for estimating soil moisture at crop field scale

Marc Padilla, Ana Pérez, and Mirta Pinilla
Starlab, Barcelona, Spain (marc.padilla@starlab.es)

Soil moisture is one of the key variables for crop modeling and scheduling farm operations. Current available soil moisture products are generated at global or regional scales and its spatial resolution (~1 km$^2$) is usually too coarse for common small farms. Within the framework of the EU Horizon2020 funded TWIGA project, we intended to provide improved soil moisture estimates at crop field scale. The advantage of focusing at the scale of a single crop is that the algorithm selection can be based more on the retrieval accuracy rather than on the computing performance.

Time series of Sentinel-1 SAR backscatter (at VH and VV polarizations) and Sentinel-2 NDVI observations, on each crop field, were assimilated with a semi-empirical polarimetric backscattering model for bare soil surfaces (Oh) coupled with a Water Cloud model (WCM). Some of the model parameters are the actual variables of interest to be estimated, in our case the daily surface soil moistures. They were estimated by a Bayesian inversion approach. The key advantage of using WCM, is that the effects of vegetation on backscatter can be taken into account, and therefore soil moisture estimates are available even when vegetation is present. The empirical model parameters (surface roughness, and A and B parameters of WCM) were calibrated with in-situ data from four stations in Ghana, with observations every 30 minutes from May to October 2019 at 10 cm depth. The calibration was based on a hierarchical Bayesian regression, to take into account that model parameter distributions might vary across land cover types and across in-situ stations themselves. The validation was based on the comparison between the soil moisture observations of one in-situ station and estimates from the model couple calibrated with the data from the other three in-situ stations. That procedure was repeated for each station. Correlation coefficients were above 0.64 and root mean square error bellow 0.065 m$^3$/m$^3$ in two out of the four stations. Accuracy tended to be dependent on field size, due to the well known SAR speckle noise. The station with the lowest accuracy was locate on a 30x30m$^2$ field. Accuracy was additionally affected by likely sudden changes on the surface soil or vegetation during the analysis time windows. Correlation coefficients were higher (~0.85) on the time periods without such sudden changes.

Given the results of the current study, we would recommend that the location of eventual future in-situ stations should be preferably placed on larger fields, larger than 30x30m$^2$. Further research would be needed to improve the model and understand better its limitations for an eventual operational implementation.