The added value of SMAP soil moisture in crop yield forecasting over Argentina

Argentina is one of the major producers and exporter of soybeans, corn, and wheat to the world market; therefore, the accurate and timely forecasting of those crops yield is crucial to national crop management and global food security. Previous studies have mainly focused on developing forecasting models for a specific crop type and location using a single source of data (e.g., vegetation indices), thus providing little insight into the forecasting models' performance on different crop types and regions. Besides, these models are based on traditional statistical regression algorithms, while more advanced machine learning approaches have not been explored. This study investigated the estimation of crop yields of three major crops (corn, soybean, and winter wheat) using Multiple Linear Regression (MLR) and Support Vector Machine (SVM), over major growing provinces in Argentina. Our models were trained and evaluated on data from 2015 to 2020, where three remote sensing products (Normalized difference vegetation index (NDVI), SMAP soil moisture, and MODIS evapotranspiration) were used as predictors. Our results indicated that accurate crop yield forecasts using the developed regression models could be made one to two months before harvest. The MLR and SVM model performance varied among different crop types, where soybean and corn exhibited better predictability compare to the wheat. In most cases, the SVM outperformed the multiple linear regression model due to its ability to capture the nonlinear and complex features of the crop-production process. The forecasted model that combines data from multiple sources outperformed single-source satellite data. The highest accuracy was obtained when the three data sources were all considered in the model development. Results also indicated that the inclusion of SMAP soil moisture improved crop yield forecasting in most provinces, and the most significant improvements occurred in the drier region.