Argentina Food Security & Agriculture

Crop Monitoring and Forecasting for Argentina using NASA Satellite Observations

 **Technical Report**

Final – August 12th, 2021

Ryan Lam (Project Lead)

Brooke Egley

Tyler Pantle

Sienna Templeman

Caroline Williams

Dr. John D. Bolten, NASA Goddard Space Flight Center (Science Advisor)

Dr. Nazmus Sazib, NASA Goddard Space Flight Center (Science Advisor)

# 1. Abstract

Early harvest information helps guide agricultural commodity assessments in Argentina, providing valuable planning information to identify potentially food-insecure regions, anticipate transportation and storage demands, predict price fluctuations, and project commodity trends. However, crop yield estimates are currently subjective, based on interviews with qualified informants (i.e., farmers, agribusiness actors). In partnership with the Buenos Aires Grain Exchange, we leveraged Terra Moderate Resolution Imaging Spectroradiometer (MODIS), Soil Moisture Active Passive (SMAP), and Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM IMERG) NASA Earth observations to develop a Google Earth Engine (GEE) toolset to monitor vegetation growth. The first component of the toolset produces spatial and temporal maps of temperature, precipitation, soil moisture, and the Normalized Difference Vegetation Index (NDVI), allowing users to visualize the influence of the region’s climate and weather. Next, we developed an autoregressive model to predict NDVI several months in advance**.** Lastly, we created a linear regression model of crop yield and NDVI for soybeans, corn, and wheat, and input the forecasted NDVI to generate a predicted crop yield output. The NDVI forecasting model produced accurate predictions at two, four, and six months when examining the most recent growing season. In the crop yield model, soybeans exhibited moderately strong correlation, wheat had consistent weak correlation, and corn varied from weak to strong correlation depending on zone. This information is vital for vegetation growth monitoring by identifying areas of high growth and allocating resources to areas of lower growth to efficiently maximize crop yields.

**Key Terms**

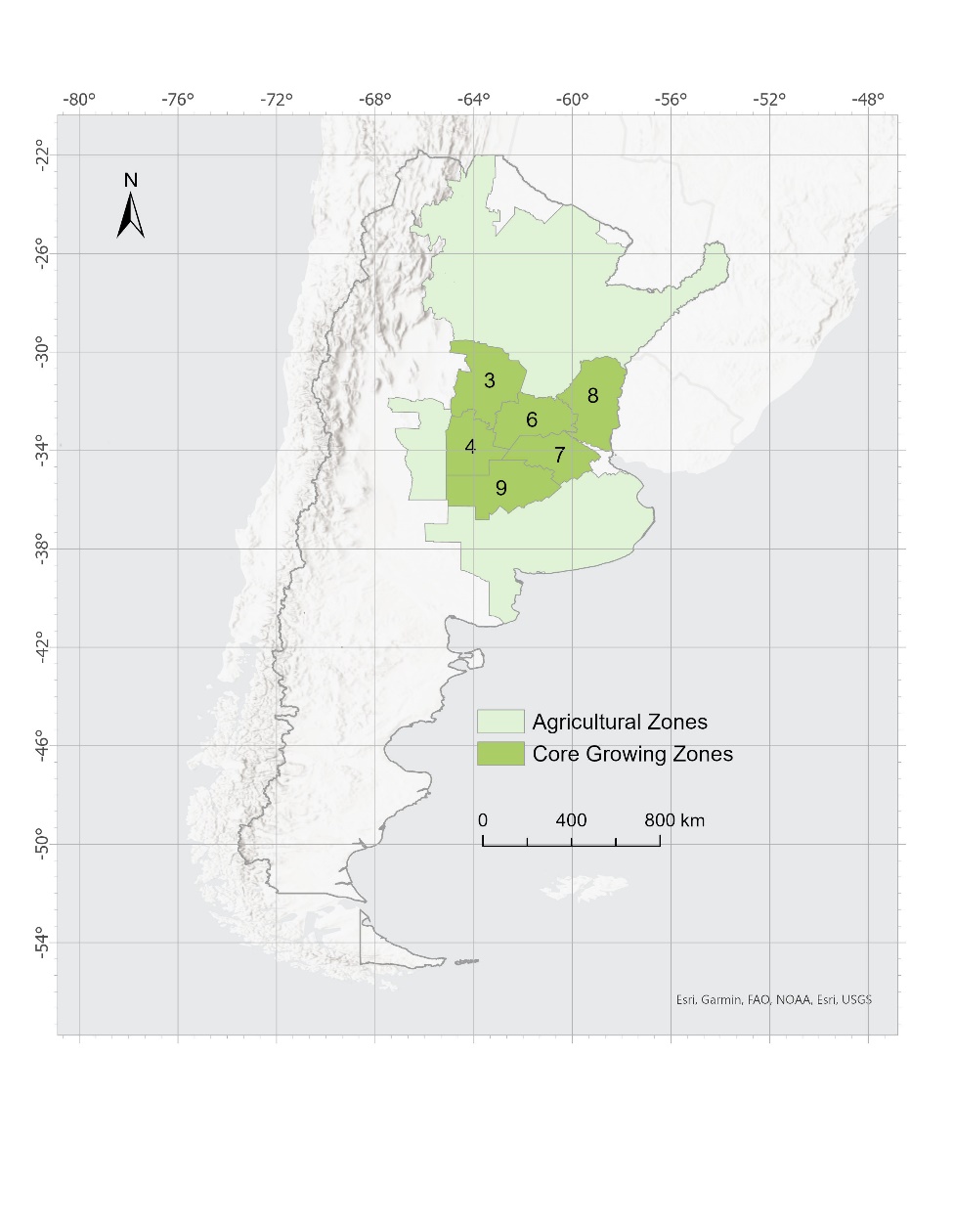
Normalized Difference Vegetation Index, Soil Moisture Active Passive, crop yield monitoring and forecasting, Google Earth Engine

# 2. Introduction

***2.1 Background Information***

Agriculture is a vital part of the economy in Argentina. The fertile soils and variety of growing zones in Argentina allow for diverse agricultural practices such as raising livestock and cultivating grains, fruits, and vegetables (Food and Agriculture Organization of the United Nations, 2017). Soybeans, corn, and wheat are Argentina’s main crop exports and account for 78% of Argentina’s farmland (Merlos et al., 2015). Soybeans are widely cultivated for the production of biodiesel, which makes them a valuable commodity for the Argentinian economy (Ministerio de Hacienda, 2019b). Soybeans and corn are the top-ranking crops respectively, with 53 million tons of soybeans and 40 million tons of corn cultivated from 2014 to 2019 (Ministerio de Hacienda, 2019a). As a major crop exporter, Argentina is crucial to both regional and global food security, as well as financial stability for regional farmers and agribusiness workers. Our partners at the Buenos Aires Grain Exchange are interested in forecasting crop yield to assess production costs and profitability. The ability to monitor growing conditions and forecast crop yield through remote sensing is valuable to Argentina’s agriculture and economy.

As remote sensing applications have advanced in recent decades, the utilization of empirical modeling for crop yield forecasting and agricultural assessments has become widely prevalent. Due to the subjectivity of survey-based yield estimates, recent studies have recommended the use of remotely sensed observations to complement in situ data, thereby improving the accuracy of the results (Lopresti et al., 2015). In combining Earth observation and in situ measurements, past studies have acquired satellite data from a variety of sensors. In particular, the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation indices like the Normalized Difference Vegetation Index (NDVI) are frequently used in crop yield estimations. Methodologies often feature linear regression and autoregressive techniques to examine the relationship between NDVI and crop yield data (Thieme et al., 2020). More recently, studies have incorporated Machine Learning (ML) and Deep Learning (DL) algorithms to estimate crop yield (Cao et al. 2021). To forecast these indices and estimate crop yield, Google Earth Engine (GEE) is a commonly used application with high storage capabilities and efficient software that is often beneficial to users that lack data and computational power (Sazib et al., 2018). While GEE has often been used in agricultural applications including drought monitoring and soil moisture assessments, fewer studies have predicted NDVI and estimated crop yield on the platform. This project examined the feasibility of NDVI forecasting and crop yield estimations in GEE and conducted a time series analysis of temperature, precipitation, soil moisture, and NDVI from November 2012 to July 2021. The Grain Exchange focuses on the main agricultural areas located in the midlatitudes and east of the Andes Mountains (Figure 1).



*Figure 1.* Study area map displaying the agricultural zones and core growing zones of Argentina. The six core zones—3, 4, 6, 7, 8, and 9—contribute to the majority of the country’s agricultural production.

***2.2 Project Partners & Objectives***

The Buenos Aires Grain Exchange is a nonprofit organization based in Argentina that provides weekly agricultural information, forecasts, and data to farmers, insurance companies, decision makers, and other various end users. The Grain Exchange gathers large amounts of observed crop yield data through surveys but has lacked the necessary infrastructure to analyze and process the observation data. The end products from this work provided the Grain Exchange with tailored satellite data and analysis tools to monitor and forecast agricultural crop conditions and improved the content, accuracy, and depth of the weekly analysis reports and their other estimation and observation data products. The overarching project goal was to decrease the limitations regarding time, technical skill, and data management capabilities the Grain Exchange faced by providing ready-made tools and customized datasets for agricultural forecasting and observation capacity.

This project aimed to analyze the relationship between agricultural yield and NDVI, and produce supplementary timeseries of climatological variables using NASA Earth observations. The project produced a GEE toolset that allowed the Grain Exchange to forecast crop yield in Argentina with an approximate lead time of one to six months using NDVI data. The GEE toolset generated maps display temperature, precipitation, soil moisture, and vegetation indices. Time series plots and geospatial maps were primary outputs with the ability to select spatial and temporal resolution as desired for regional and temporal analysis.

# 3. Methodology

***3.1 Data Acquisition***

We acquired data on NDVI, land surface temperature (LST), precipitation, and surface soil moisture from the GEE data repository (Table 1). These data were produced by NASA Earth observations including Terra Moderate Resolution Imaging Spectroradiometer (MODIS), Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM-IMERG), and Soil Moisture Active Passive (SMAP). The soil moisture dataset is produced by integrating remote-sensed SMAP data with a Palmer model driven by precipitation and temperature. We acquired Level 3 soil moisture data from 1 April 2015 to 15 June2020 to coincide with the beginning of data production from the SMAP mission. We acquired Level 3 NDVI, Level 2 LST, and Level 3 precipitation data from 1 January 2012 to 15 June 2021. We also received in situ data from the Grain Exchange including Argentina crop yield data for the 2012/2013 through the 2018/2019 growing seasons and shapefiles of agricultural zones monitored by the Grain Exchange.

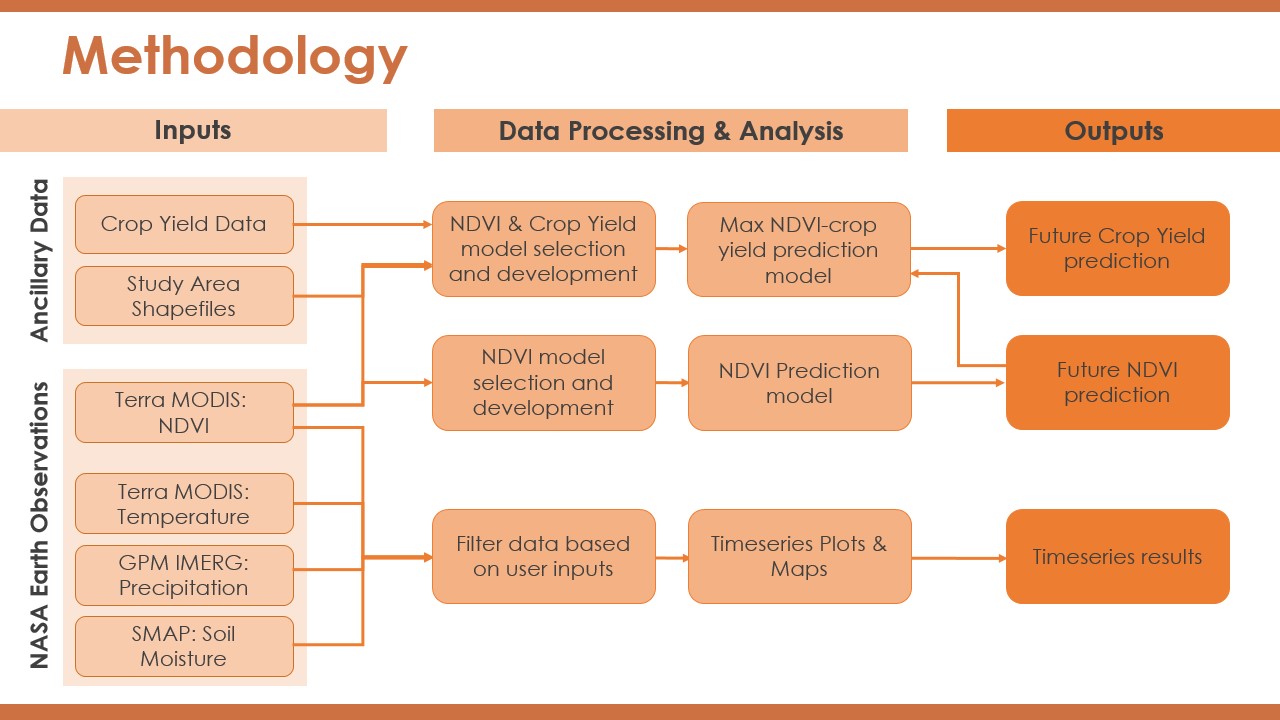
Table 1

*List of satellite-derived datasets*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Platform & Sensor** | **Data Product** | **Provider** | **Spatial Resolution** | **Digital Object Identifier** | **Acquisition Method** |
| Terra MODIS | NDVI | NASA | 1km | 10.5067/MODIS/MOD13A2.006 | GEE |
| Terra MODIS | LST | NASA | 1km | 10.5067/MODIS/MOD11A1.006 | GEE |
| GPM-IMERG | Precipitation | NASA | 10km | 10.5067/GPM/IMERG/3B-MONTH/06 | GEE |
| SMAP | Surface Soil Moisture | NASA | 10km | https://gimms.gsfc.nasa.gove/SMOS/SMAP | GEE |

***3.2 Data Processing***

The datasets required minimal processing because they were designated products for their respective data and pre-processed datasets (Figure 2). We applied a scale factor of 0.0001 to the MODIS NDVI band to output values within the -1 to 1 range commonly accepted for NDVI (Lopresti et al., 2015). We also applied a recommended scale factor of 0.02 to the MODIS LST band and converted units from Kelvin to Celsius. For the precipitation dataset, we converted units from millimeters (mm) per hour to mm per month and included a function to calculate cumulative precipitation during a growing season. We clipped all GEE datasets to a subset of six agricultural zones out of the 15 total zones monitored by the Grain Exchange (Figure 1). The subset of ‘core zones’ included Zones 3, 4, 6, 7, 8, and 9, and was identified by the Grain Exchange because they comprised the majority of crop harvest and provided the most useful data.



*Figure 2.* Flowchart of project methodology including data inputs, processing and analysis, and final outputs.

The climate variable script within the toolset automatically applies further processing measures to datasets based on required inputs from the user. The user inputs their desired year, month, agricultural zone, or coordinates, and then selects the combination of the climate variables for which they would like to output results. The script then filters datasets of the selected climate variables based on the user’s inputs, and outputs the relevant maps and time series graphs. Average and maximum values are calculated for both the input month and year, and then the results are shown on the GEE map window. Average monthly values are calculated across the input agricultural zone and then plotted for the input year and all other years during the study period for comparison. The same average monthly value graphs are also created for data at the input coordinates. The average monthly values for the input zone are also plotted alongside the average monthly values of all other zones for comparison.

***3.3 Data Analysis***

*3.3.1 Climate Variable Analysis*

The climate variable script is designed to output maps and time series graphs that are supplementary to the NDVI and crop yield forecasting models. The script is capable of producing outputs for all years and months during the study period, each of the 15 agricultural zones—including the six core zones, and any of the four climate variables. These outputs provide end users with climatological context for NDVI and crop yield predictions and allow the user to analyze all of the data concurrently. This tool provides the ability to analyze multiple climate variables and draw conclusions from the data visualization results.

*3.3.2 NDVI Analysis*

*3.3.2.1 Autoregression Model Development and Forecasting*

The NDVI forecasting analysis featured two main components. This included an autoregression model and forecasting model that produced an NDVI image collection at a 16-day time interval. The resulting image collection combined both observed MODIS NDVI and predicted future NDVI images across the study region. The model development for the first component of the NDVI forecasting analysis involved the construction of three autoregression models. An autoregression model was developed in GEE due to its ability to incorporate lag. For example, a lag of 1, referred to an autoregression model that examines two NDVI images, a present image, and an observed image 16-days prior. Therefore, a lag of x examines x+1 number of NDVI images in its calculation.

The following calculations were used in developing the autoregression models, Autoregression Model 1 (AR1) (Equation 1), Autoregression Model 2 (AR2) (Equation 2), and Autoregression Model 3 (AR3) (Equation 3), respectively:

Pt = B0 + B1Pt-1 + et(1)

Pt = B0 + B1Pt-1 + B2Pt-2 + et  (2)

Pt = B0 + B1Pt-1 + B2Pt-2 + B3Pt-3 + et (3)

where B0, B1, B2, and B3 are the coefficients at time t, t-1 t-2, and t-3. P is NDVI at time t, t-1 t-2, and t-3, and et is the error at time t.

The Root Mean Square Error (RMSE) was used to measure the accuracy of the autoregression models. The model with the highest accuracy was then automatically selected for the NDVI forecasting model. The forecasting model had the functionality to predict NDVI up to one year at 16-day time intervals which corresponded to the temporal resolution of the observed NDVI data. The forecasting model then used the coefficients at different lags derived from the autoregression model. Additionally, the prediction model also incorporated an image collection with the observed MODIS NDVI images and blank images. These blank images are set to future time stamps and were automatically overwritten by the prediction model in GEE. This resulted in a time series of historical NDVI and future NDVI. This model is ideally used to forecast NDVI, however the model can also be utilized at previous time periods to examine the prediction model’s accuracy.

*3.3.2.2 Examining Previous Growing Seasons*

The forecasting model accuracy was tested using data from the most recent growing season (16 November 2020 to 9 May 2021) for the study area and making predictions at different forecasting lengths (2-Months, 4-Months, and 6-Months). Each forecasting period incorporated 4, 8, and 12 predicted NDVI images respectively, since the model ran on a 16-day time period. This led to differing accuracies across each forecasting length. The accuracy of the observed versus predicted NDVI values for different forecasting lengths (2-Months, 4-Months, and 6-Months) was determined using the R-squared value.

*3.3.3 NDVI and Crop Yield Data Analysis*

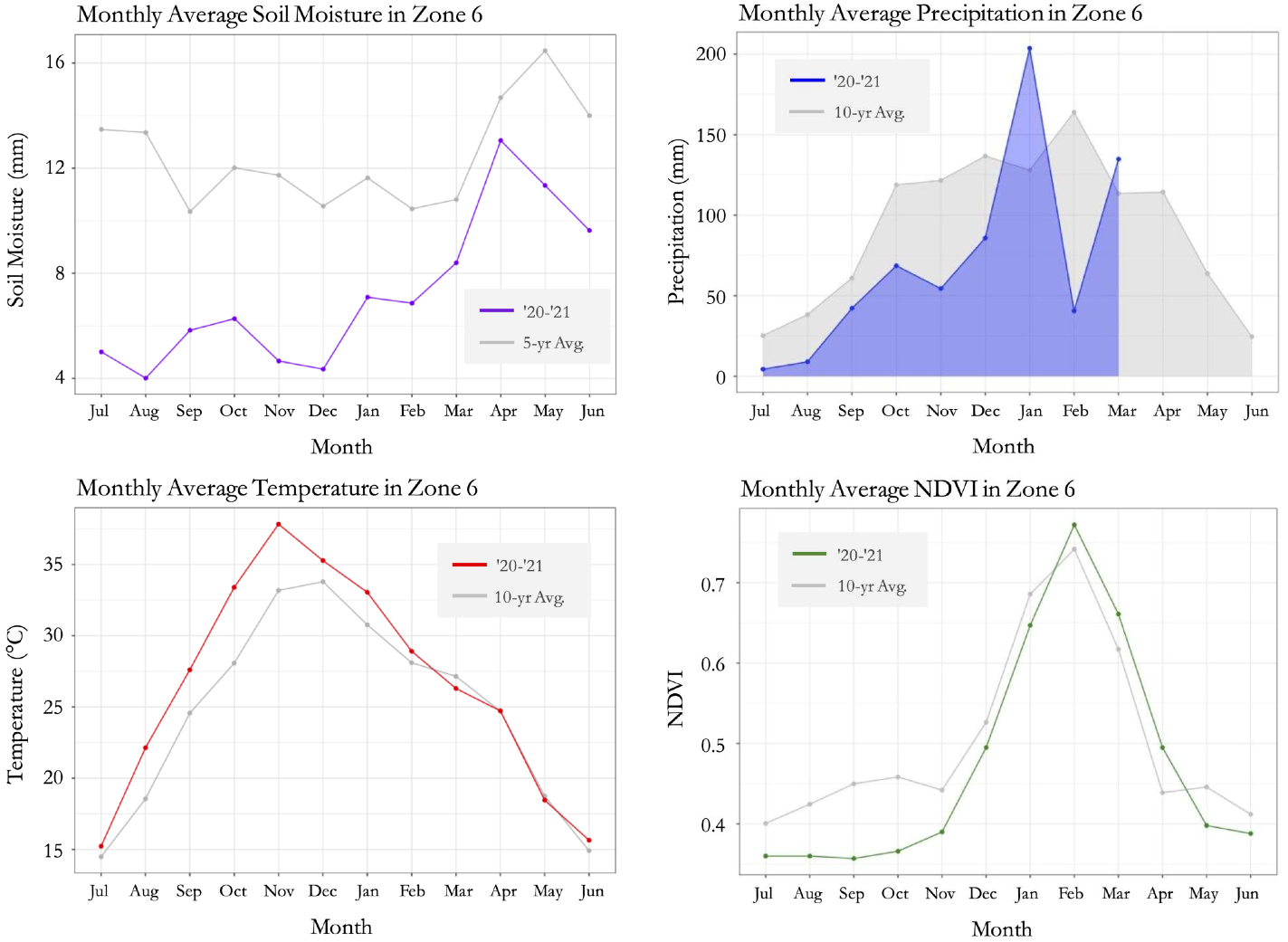
The crop yield forecasting script begins with a function that calculates the average maximum NDVI in a specified growing zone for each growing season. It returns a list of calculated NDVI for the years 2012 to 2019, which is assigned to ordered pairs with the corresponding annual crop yield data. We also calculated the natural log of crop yields, to transform the data to be normally distributed. Then, a linear fit was performed on the NDVI and crop yield data to return the linear regression equation for the specified zone and crop type with NDVI set as the independent variable and crop yields as the dependent variable. The function returns 18 unique linear regression equations for soybeans, corn, and wheat across the six core growing zones. We then used our forecasting model to output a forecasted, average maximum NDVI for the specified growing zone. This output of forecasted NDVI is then input into the respective linear regression equation to produce the forecasted crop yield for the specified crop and zone. The anti-log of the output is calculated to obtain the true crop yield in kilograms per hectare.

# 4. Results & Discussion

***4.1 Analysis of Results***

4.1.1 Climate Variables

The 2020 to 2021 growing season has been significantly impacted by the La Niña phase of El Niño Southern Oscillation (ENSO) for the former half of the season. In a La Niña event, Argentina’s major crop regions generally receive less rainfall and experience higher temperatures. From 2019 to 2020, there was also a La Niña event, causing a lag impact in some of the climate variables like soil moisture. This can be noted in *Figure 3, Figure A1,* and *Figure A2*. However, as is shown in *Figure 3, Figure A3,* and *Figure A4*, in comparison to the other variables,the NDVI values were unexpectedly high even with very low soil moisture, low precipitation, and high temperatures. This requires further exploration, as the expected result would be a decrease in NDVI values, indicating a decrease in crop health. The ability to analyze climate variables alongside each other has a tangible benefit in identifying potential climatological impacts for study such as this situation.



*Figure 3.* Climate variable graphs of temperature, precipitation, soil moisture, and NDVI in Zone 6 for the 2020 to 2021 growing season, compared to the historical averages.

4.1.2 NDVI Forecasting

The AR3 model performed with the highest accuracy across all zones and was thereby chosen for the NDVI forecasting model. Using the NDVI forecasting AR3 model, the predicted NDVI versus actual NDVI mean values by zone were compared through linear regression analysis. Table 2 demonstrates the resulting R2 values and statistical significance using an F-statistic test. Additionally, the linear regression equations are depicted in Table A1. The zones with the highest accuracy in the 2-month forecast are Zones 6 (R2 = 0.9967) and Zone 7 (R2 = 0.9833). Zones 3, 4, and 8 also showed highly accurate predictions in the 2-month forecast for the most recent growing season with R2 values ranging from 0.6041 to 0.8988. Zone 9 showed a low accuracy for the 2-month forecast. However, this was not statistically significant. The 4-month forecast similarly had a high accuracy in the predicted versus actual R2 values ranging from 0.6626 to 0.8913 across all zones. In the final forecasting period of six months, the R2 values occurred between 0.504 to 0.79. Forecasts beyond six months decreased in accuracy greatly, however, the 2- to 4-month range typically depicted highly accurate forecasts. This decrease in accuracy as the lead time increases is explained by the functionality of the forecasting model. This model incorporates the most recent values with the specified lag in predicting the next forecasted value in the series. As the prediction increases, the most recent forecasted values will begin to predict new values rather than using the most recent observed NDVI values.

Table 2

*Predicted NDVI versus Actual NDVI R*2 *values from the most recent growing season predicted at 2-, 4-, and 6-month time intervals from the start date using the AR3 model.* \* p < 0.01, \*\* p < 0.001, \*\*\* p < 0.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **2-Month Forecast** | **4-Month Forecast** | **6-Month Forecast** |
|  | **16 Nov 2020 to 1 Jan 2021** | **16 Nov 2020 to 6 Mar 2021** | **16 Nov 2020 to 9 May 2021** |
| **Zone** | R2 | R2 | R2 |
| 3 | 0.824 | 0.787\*\* | 0.659\*\* |
| 4 | 0.6041 | 0.7927\*\* | 0.6369\*\* |
| 6 | 0.9967\*\* | 0.8913\*\*\* | 0.5361\*\* |
| 7 | 0.9833\*\* | 0.8907\*\*\* | 0.5036\*\* |
| 8 | 0.8988 | 0.8292\*\* | 0.79\*\*\* |
| 9 | 0.2897 | 0.6626\* | 0.5325\*\* |

In using the GEE NDVI forecasting tool, we suggest forecasting in a 1- to 4-month period due to the high accuracy of nearly all zones. Predicting 4- to 6- months from the starting period can be accurately indicated by the results. However, a lead time beyond a 6-month prediction is not recommended as the values begin to plateau in long-term forecasting periods. The predicted values in long-term forecasts tend to underestimate NDVI. This may be due to the model’s functionality which incorporates recent NDVI values, actual or predicted, to forecast future points. Overall, these results present that using autoregressive modeling to develop an NDVI forecast tool in GEE is feasible using Earth observations. Additionally, the NDVI forecasting model shows potential for further development. This includes adding climate variables into the model’s calculation.

4.1.3 Crop Yield Forecasting

The relationship between NDVI and crop yields was highly varied for each zone and crop type. All linear regression equations are listed in *Table A2*, and their respective R2 values are listed in Table 3. Soybeans consistently exhibited a moderately strong positive correlation for each zone, with R2 values ranging from 0.64 in Zone 6 to 0.95 in Zone 9. Corn is the most variable crop, with an R2 value of 0.90 in Zone 3, but R2 values of 0.0013 and 0.21 in Zones 6 and 8. Wheat consistently exhibited very poor correlation for each zone, with R2 values as low as 0.0002 in Zone 3, with the highest value being 0.1992 in Zone 6.

Table 3

*List of R2 values for each crop type and zone.*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Soybeans** | **Corn** | **Wheat** |
| **Zone** | R2 | R2 | R2 |
| 3 | 0.8036 | 0.9014 | 0.0002 |
| 4 | 0.7559 | 0.4694 | 0.0308 |
| 6 | 0.647 | 0.0013 | 0.1992 |
| 7 | 0.7613 | 0.0961 | 0.0031 |
| 8 | 0.9282 | 0.021 | 0.1854 |
| 9 | 0.9536 | 0.6331 | 0.0644 |

Due to the variation and uncertainty in correlation exhibited in Table 3, we recommend that crop yield forecasting is only performed on the crop types and zones with strong correlations. For soybeans, Zones 3, 8, and 9 have correlations above 0.8, and would likely produce accurate forecasts. Zones 3 and 9 are also acceptable to use for corn. We suggest that no crop forecasting is performed for wheat due to the overall weak correlation values. The weak correlation may be a result of insufficient data to create a strong linear regression model. Having seven years of annual crop yields led us to reduce NDVI to a single annual value per region. This single, average value includes areas outside agricultural land, such as bodies of water or developed areas, that would negatively influence the mean NDVI. The inclusion of a crop mask in future work could solve this issue and specific masks for corn, soybeans, and wheat should lead to even stronger relationships between NDVI and yield Differences in growing periods for each crop are another potential factor contributing to our observed varying correlations. In Argentina, soybean and corn are planted in the spring and summer months, while wheat is planted in the fall and winter months (United States Department of Agriculture, n.d.). The varying schedules result in each crop relying on different periods during which weather conditions are critical for yield. Time constraints of the project did not allow for seasonal model calibration for each crop. Crop-specific modeling with seasonal calibration should also contribute to stronger relationships between NDVI and yield.

***4.2 Future Work***

A way to enhance the GEE software script is the development of a Graphical User Interface (GUI) which bypasses the need for end users to interact directly with the code. The inclusion of a crop mask in the processing phase can also improve the performance of the GEE outputs by only calculating and extracting values of agricultural land cover and exclude data such as water bodies (Lopresti et al., 2015). Model performance may increase if the spatial resolution is downscaled from 10 km. It is possible to increase the comprehensiveness of the model by including climate variables such as soil moisture, precipitation, and temperature. The crop yield forecast currently uses NDVI as a model input and could benefit significantly from the inclusion of a greater number of variables. The Grain Exchange provided crop yield data for each province, but this data was aggregated to each growing zone for simplicity given time restraints. Future work could include creating individual forecasting models for the most agriculturally productive provinces. Further study is required to fully determine the correlation between different phases and teleconnections of climatological events like ENSO, climate variables, and NDVI forecasting. Economic data that impact yields, such as domestic grain prices and input prices of fertilizer, fuel, herbicide, seed, is another area of exploration with its potential integration as inputs to the crop yield model.

# 5. Conclusions

The climate variable models output plots of temperature, precipitation, soil moisture, and NDVI for user-selected zones and timescales. The user-friendly climate variable script has the capability to output thousands of different plots depending on the desired data. These outputs allow the Grain Exchange to view climate data alongside NDVI and crop yield forecasts, providing more context to any predicted values. The autoregressive NDVI forecasting model predicts NDVI two to four months in advance with high accuracy. Additionally, the crop yield forecasting model contributes to the study of the relationship between NDVI and crop yield. By using the forecasted NDVI, the model is able to predict yields for soybeans, corn, and wheat in the six core growing zones. Forecasting crop yields is only recommended for the zones that exhibit strong correlation and is discouraged for other zones until a more accurate model is developed. Providing our partners with the ability to forecast NDVI and crop yield will improve their reports, as well as reduce the time and effort needed to collect and process large data. As a major producer and global exporter of soybeans, corn, and wheat, these two scripts and their products will benefit the agriculture and economy of Argentina.

# 6. Acknowledgments

The team would like to thank everyone that contributed to the project’s progress over the course of this term: our Science Advisors at Goddard Space Flight Center, Dr. John Bolten and Dr. Nazmus Sazib. We would like to thank our DEVELOP Fellow/Center Lead, Dr. Nicole Ramberg-Pihl for her continued support. Lastly, we appreciate the collaboration and input from our partners at the Buenos Aires Grain Exchange.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract NNL16AA05C.

# 7. Glossary

**DL** – Deep Learning

**GEE** – Google Earth Engine

**GPM** – Global Precipitation Model

**LST** – Land Surface Temperature

**ML** – Machine Learning

**MODIS** – Moderate Resolution Imaging Spectroradiometer

**NDVI** – Normalized Difference Vegetation Index

**SMAP** – Soil Moisture Active Passive

**USDA** – United States Department of Agriculture

# 8. References

Cao, J., Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., & Tao, F. (2021). Wheat yield predictions at a county and field scale with deep learning, machine learning, and Google Earth Engine. *European Journal of Agronomy, 123*, 126204. doi:https://doi.org/10.1016/j.eja.2020.126204

Food and Agriculture Organization of the United Nations. (2017). *Country Fact Sheet on Food and*  *Agriculture Policy Trends - Argentina*. http://www.fao.org/3/i7752e/i7752e.pdf

Lopresti, M. F., Di Bella, C. M., & Degioanni, A. J. (2015). Relationship between MODIS-NDVI data and

wheat yield: A case study in Northern Buenos Aires province, Argentina. *Information Processing in*  *Agriculture, 2*(2), 73–84. https://doi.org/10.1016/j.inpa.2015.06.001

Merlos, F.A., Monzon, J.P., Mercau, J.L., Taboada, M., Andrade, F.H., Hall, A.J., Jobbagy, E., Cassman, K.G., & Grassini, P. (2015). Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Research, 184,* 145-154*.* https://doi.org/10.1016/j.fcr.2015.10.001

Ministerio de Hacienda, Argentina. (2019a). Reports from Chains of Value – Cereals: Corn. ISSN 2525-0221. https://www.argentina.gob.ar/sites/default/files/sspmicro\_cadenas\_de\_valor\_maiz.pdf

Ministerio de Hacienda, Argentina. (2019b). Reports from Chains of Value – Oilseeds: Soy. ISSN: 2525-0221. https://www.argentina.gob.ar/sites/default/files/sspmicro\_cadenas\_de\_valor\_soja.pdf

Sazib, N., Mladenova, I., & Bolten, J. (2018). Leveraging the Google Earth Engine for drought assessment

using global soil moisture data. *Remote Sensing, 10*(8). https://doi.org/10.3390/rs10081265

Thieme, A., Yadav, S., Oddo, P. C., Fitz, J. M., McCartney, S., King, L. A., Keppler, J., McCarty, G. W., & Hively, W. D. (2020). Using NASA Earth observations and Google Earth Engine to map winter cover crop conservation performance in the Chesapeake Bay watershed. *Remote Sensing of Environment, 248*, 111943. <https://doi.org/10.1016/j.rse.2020.111943>

United States Department of Agriculture. (n.d.). Crop Calendars for Southern South America. Retrieved September 29, 2021, from https://ipad.fas.usda.gov/rssiws/al/crop\_calendar/ssa.aspx.

# 9. Appendix



*Figure A1.* Mean temperature map encompassing the 2020 to 2021 growing season (June 2020 to July 2021). Data exported from GEE Climatology tool. Map created in ArcGIS Pro.



*Figure A2.* Mean precipitation encompassing the 2020 to 2021 growing season (June 2020 to July 2021). Data exported from GEE Climatology tool. Map created in ArcGIS Pro.



*Figure A3.* Mean soil moisture encompassing the 2020 to 2021 growing season (June 2020 to July 2021). Data exported from GEE Climatology tool. Map created in ArcGIS Pro.



*Figure A4.* Maximum NDVI encompassing the 2020 to 2021 growing season (June 2020 to July 2021). Data exported from GEE Climatology tool. Map created in ArcGIS Pro.

Table A1

*Predicted NDVI versus Actual NDVI values in the form of a linear regression and R-squared values. This is a prediction for the most recent growing season at 2, 4, and 6-month time intervals from the start date using the AR3 model. \* p < 0.01, \*\* p < 0.001, \*\*\* p < 0.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **2-Month Forecast** | | **4-Month Forecast** | | **6-Month Forecast** | |
|  | **16 Nov 2020 to 1 Jan 2021** | | **16 Nov 2020 to 6 Mar 2021** | | **16 Nov 2020 to 9 May 2021** | |
| **Zone** | Linear Regression | R2 | Linear Regression | R2 | Linear Regression | R2 |
| 3 | y = 0.763x + 0.0843 | 0.824 | y = 0.2663x + 0.2778 | 0.787\*\* | y = 0.2352x + 0.3015 | 0.659\*\* |
| 4 | y = 0.3967x + 0.2537 | 0.6041 | y = 0.2199x + 0.3221 | 0.7927\*\* | y = 0.1821x + 0.3516 | 0.6369\*\* |
| 6 | y = 0.4903x + 0.214 | 0.9967\*\* | y = 0.2739x + 0.3164 | 0.8913\*\*\* | y = 0.1865x + 0.382 | 0.5361\*\* |
| 7 | y = 0.1557x + 0.4338 | 0.9833\*\* | y = 0.0655x + 0.4783 | 0.8907\*\*\* | y = 0.0447x + 0.4937 | 0.5036\*\* |
| 8 | y = 0.3358x + 0.3723 | 0.8988 | y = 0.2347x + 0.4304 | 0.8292\*\* | y = 0.2235x + 0.4382 | 0.79\*\*\* |
| 9 | y = 0.2305x + 0.3812 | 0.2897 | y = 0.1001x + 0.4388 | 0.6626\* | y = 0.0843x + 0.4491 | 0.5325\*\* |

Table A2

*Linear regression equations and their respective R2 values for each zone and crop type.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Soybeans** | | **Corn** | | **Wheat** | |
| **Zone** | Linear Regression | R2 | Linear Regression | R2 | Linear Regression | R2 |
| 3 | y = 8.701x - 0.111 | 0.8036 | y = 3.418x + 4.628 | 0.9014 | y = 0.0146x + 5.744 | 0.0002 |
| 4 | y = 8.4373x - 1.597 | 0.7559 | y = 4.201x + 2.864 | 0.4694 | y = -1.077x + 5.76 | 0.0308 |
| 6 | y = 8.344x - 0.571 | 0.647 | y = 0.182x + 6.908 | 0.0013 | y = -4.053x + 9.1501 | 0.1992 |
| 7 | y = 8.825x - 0.422 | 0.7613 | y = 1.862x + 6.302 | 0.0961 | y = 0.391x + 6.692 | 0.0031 |
| 8 | y = 14.631x - 4.94 | 0.9282 | y = 3.538x + 4.346 | 0.021 | y = 2.637x + 4.0497 | 0.1854 |
| 9 | y = 8.066x + 0.596 | 0.9536 | y = 4.183x + 4.722 | 0.6331 | y = 2.038x + 5.208 | 0.0644 |