NASA DEVELOP National Program
North Carolina – NCEI

Spring 2024

Northern Brazil Agriculture

Measuring Soybean Yields in Northern Brazil During El Niño-Southern Oscillation Conditions to Evaluate Trends in Agricultural Production and Support Crop Forecasting, 1984 – 2023

DEVELOP Technical Report

March 29th, 2024

Devon V. Maloney (Project Lead)
Manpreet K. Singh
Grayson Shanley Barr
Sofya Goncharenko

Advisors:
Dr. Garrett Graham, North Carolina Institute for Climate Studies (Science Advisor)
Dr. Boyin Huang, NOAA National Centers for Environmental Information (Science Advisor)
Molly Woloszyn, NOAA National Integrated Drought Information System (Lead Science Advisor)

Lead:
Kathryn Caruso (North Carolina - NCEI)

1. Abstract

As one of the largest agricultural exporters in the world, Brazil's crop productivity is highly influential to the world's food supply. Crop productivity across Brazil is heavily influenced by climatic factors, such as the El Niño-Southern Oscillation (ENSO), which drives spatially heterogenous effects on growing conditions. To understand the relationship between ENSO conditions and staple crop productivity, the team used 1984 – 2023 climatic data and vegetation indices for four Brazilian states: Bahia, Mato Grosso, Pará, and Tocantins. The team focused on soybean growing areas in each state derived from annual land use/land cover classifications. Monthly mean Normalized Difference Vegetation Index (NDVI) per state were calculated using multispectral imagery from the Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imagery (OLI) sensors. In addition, the team used crop production indices, ENSO anomalies, temperature, and precipitation data in this study. The team found that although monthly ENSO anomaly did have a positive relationship between temperature and a mild negative relationship with precipitation, ENSO and monthly soy NDVI were not significantly correlated. Results find no association between ENSO conditions and NDVI in the study area at the state-level spatial resolution during the study period. Furthermore, the team found that there was no correlation between cumulative growing season soy NDVI and detrended soy production yield for this study region at the spatial and temporal scale of the analysis. Findings from this project will be used by the United States Department of Agriculture's Foreign Agriculture Service to inform crop productivity forecasting.

Key Terms

Remote sensing, Landsat, soybean crop productivity, Brazil, El Niño, NDVI, El Niño Southern Oscillation (ENSO)

2. Introduction

2.1 Background Information

Brazil serves as a top five global agricultural producer with over thirty agriculture commodities grown on its arable land (Valdes, 2022). Soybean is one of the key crops produced in Brazil in terms of annual production by weight and annual area harvested, and consequently has a significant local economic impact (USDA FAS IPAD, 2024; Cattelan and Dall'Agnol, 2017). From market year 2013–2014 to market year 2022–2023, soybean production increased from 2.9 tons/hectare to 3.4 tons/hectare (FAS IPAD, 2024). In addition, in recent years, soybean production has expanded northward into the northeastern interior of Brazil (Colussi & Schnitkey, 2021). Although there are many factors that impact Brazilian agriculture production—including the implementation of agricultural technology, access to financing, and urbanization—climate variability serves as a significant factor that impacts soybean harvest (Reis et al., 2020).

One phenomenon that contributes to increased climatic variability in Brazil is the El Niño-Southern Oscillation (ENSO), which results from irregular variation in winds and sea surface temperatures over the tropical Pacific Ocean (Cirino et al., 2015). ENSO conditions affect the climate of the tropics and subtropics (Cirino et al., 2015). Northeastern Brazil is particularly vulnerable to ENSO effects, which result in drier than normal conditions and lower than normal precipitation (Cirino et al., 2015). As weather is a primary determinant of agricultural productivity, ENSO-related events can lead to a loss in crop production yields due to flooding and severe droughts (Cirino et al., 2015). Crop production declines are linked to socioeconomic costs including an increase in food prices, both locally and globally, and a decrease in rural income that negatively impacts farmers (Cirino et al., 2015). As ENSO conditions drive global patterns of temperature and precipitation, multiple simultaneous crop failures due to climatic variability can lead to global scale crises including food insecurity (Anderson et al., 2017). Consequently, due to ENSO conditions, Brazil faces potential reductions in agricultural productivity, which may impact the global food supply.

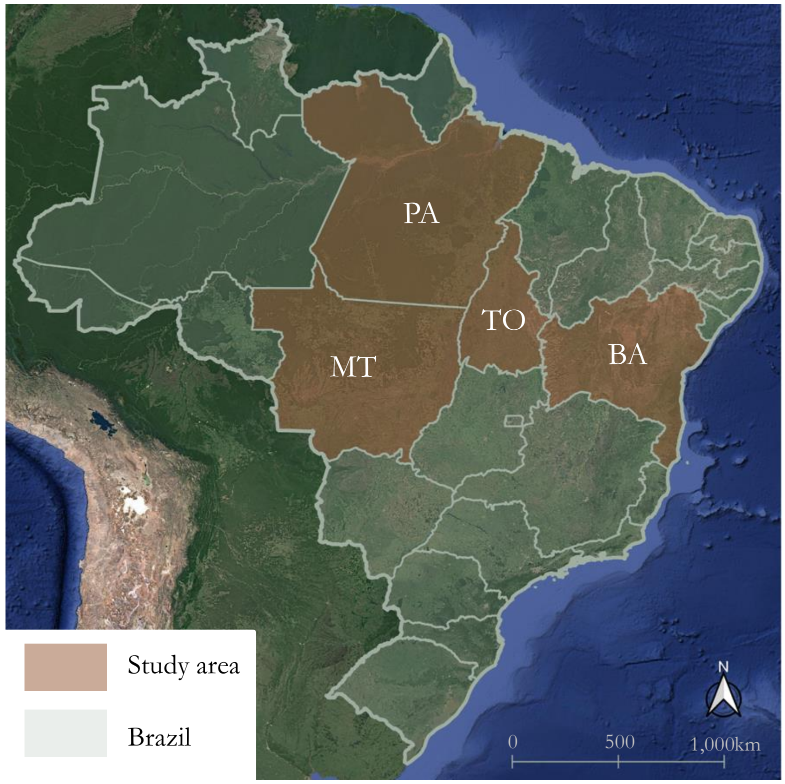


Figure 1. The study site in northern Brazil encompassed the states of Bahia (BA), Mato Grosso (MT), Pará (PA), and Tocantins (TO). Base map layer: Google Satellite.

2.2 Study Area and Period

Upon the recommendation of project partners, the analysis focuses on four states in Brazil: Bahia, Mato Grosso, Pará, and Tocantins (Figure 1). Each state is unique in its environment, plays a crucial role in Brazil's agricultural sector, and is considerably impacted by ENSO conditions (Cirino et al., 2015). The study area encompasses multiple eco-regions, including the Amazon rainforest, cerrado (savannah), and caatinga (dry forest). In addition, the study area states account for the largest production of soybean in Brazil, making the region important to food supply in both Brazil and globally (FAS IPAD, 2024). The study period for this project falls between 1984 and 2023 and analyzes each year for relationships between climatic variables and measures of soybean production.

Pará, located in northern Brazil, experiences tropical weather and strong rains (Martorano et al., 2016). A dominant hydrological feature of the state is the Amazon River, which crosses the state for about 500 miles (800 km) before entering the Atlantic Ocean (Martorano et al., 2016). Over the past 40 years, the state has experienced high rates of deforestation for development and agricultural purposes (Carvalho et al., 2019).

Mato Grosso is unique in that the state contains both rainforest and savannah ecosystems (Arvor et al., 2015). The northern region of the state is part of the transition between the Amazon and cerrado biomes, while the southern region is entirely part of the cerrado biome (Heinemann et al., 2021). Similar to Pará, Mato Grosso has also experienced deforestation of convert forested areas to cropland (Richards et al., 2015). Mato Grosso produces the largest quantity of soybeans in Brazil in addition to substantial amounts of corn and cotton (USDA, 2024; Richards et al., 2015; Arvor et al., 2014).

The climate in Tocantins is generally subtropical, and the state is considered to be Brazil's newest agricultural frontier (Avanzi et al., 2019). Economic activities in Tocantins are focused on cattle breeding, forestry, and grain production of soybean, corn, and rice (Avanzi et al., 2019). Rainfall in Tocantins is quite variable as the northern part of the state is located within the Amazon biome whereas the southern part of the state is located within the cerrado biome, which experiences a substantially longer dry season between May and September (Avanzi et al., 2019).

Located on the eastern coast of Brazil, Bahia is situated in the transition between cerrado and caatinga biomes (Batistella & Valladares, 2009). Since the 1980's, farmers in the region have heavily invested in the production of grains, primarily soybean and corn (Batistella & Valladares, 2009). Bahia has two well-defined seasons: a dry season with mild temperatures between May and September and a hot and rainy season between October and April which are crucial months for soybean production (Batistella & Valladares, 2009). Like Tocantins and Mato Grosso, the soybean planting season occurs between October and December, while the harvest season typically occurs between March through May (USDA, 2024).

2.3 Project Partners & Objectives

This project collaborated with the United States Department of Agriculture (USDA) Foreign Agricultural Services (FAS) International Production Assessment Division (IPAD) for Brazil and the USDA Office of the Chief Economist and World Agricultural Outlook Board. The USDA FAS plays a crucial role in international activities through trade, development of the global market, and the collection and analysis of market information with a goal to inform US decision making related to agricultural exports. The work involves using geospatial data, statistical analysis, and crop classification to study the impact of climatic variability on agriculture production yield. IPAD assesses two key aspects of international agriculture: global agricultural production outlook and conditions that impact international food security.

With the return of El Niño conditions in 2023, the project partners were interested in assessing the impact of ENSO on soybean crop production yield in Brazil. Using geospatial data and data analysis, the project aimed to support partner's efforts to produce crop production outlooks for northern and central Brazilian states. Modeling crop production outlooks involved the following objectives: 1) produce a geospatial analysis displaying the relationship between crops and ENSO conditions, 2) conduct a time series analysis from 1984 to 2023 comparing crop yield and environmental conditions during ENSO phases, and 3) conduct statistical analysis to examine relationships between soybean crop yield, normalized difference vegetation index (NDVI), and temperature, precipitation, and ENSO conditions. The findings from this project can be used by partners to evaluate the impact of ENSO on crop production at the state level from 1984 to 2023 and to inform crop production outlooks during future El Niño years.

2.4 Scientific Basis

Previous literature analyzing the impacts of ENSO conditions on crop production utilized a variety of methods, such as remote sensing and data analysis, to examine the influence of soil type and meteorological variables on agriculture production in Brazil.

When examining ENSO anomalies and conditions, past studies primarily used sea surface temperature (SST) anomalies to identify El Niño or La Niña conditions. The Oceanic Niño Index defines El Niño as a phenomenon that occurs in the equatorial Pacific Ocean when the 3-month running mean of SST anomalies is above or below the threshold of ±0.5º C (NOAA, 2024). More specifically, values lower than -0.5º C are classified as La Niña conditions whereas values higher than 0.5º C are classified as El Niño conditions. Neutral values are between -0.5º and 0.5º C (Sentelhas and Pereira, 2019). NOAA's definition of ENSO anomalies served as the foundation that defined thresholds of ENSO anomalies and it impacts on production yields for this analysis.

Previous literature has utilized NOAA's definition of ENSO conditions to conduct geospatial analyses on the influence of ENSO conditions on agriculture production. The primary motivations of such research projects have been to adjust cropping calendars for specific ENSO phases where crop rotations occur, assess the influence of ENSO on the spatial and temporal off-season yield of corn and soybean, and examine the rapid expansion of agriculture products in the project areas of focus (Heinemann et al., 2021; Júnior et al., 2019; Arvor et al., 2013). For example, when conducting an inter-annual trend analysis for vegetation greenness and climate anomalies in Brazil, Erasmi et al. (2014) found a strong relationship between ENSO warm events and periods of reduced vegetation greenness for corn and soybean production with a 12-month temporal lag. Erasmi et al.'s (2014) findings underscore the need to examine temporal lags in the onset of ENSO periods and when the impacts could be detected in precipitation and temperature in northern and central Brazil.

Previous research regarding ENSO impacts on agricultural production in Brazil has guided project scope and methodology. Literature review assisted in identifying geospatial and statistical tools to conduct geospatial analysis on ENSO anomalies and NDVI; for example, utilizing NDVI as a measure of crop health and productivity, and conducting principal component analysis (PCA) to reduce dimensionality and identify meaningful relationships between variables in large datasets (Ji & Peters, 2003; Arvor et al., 2013; Paul et al., 2013). Furthermore, these papers served as a foundation for methods to examine the relationship between ENSO anomalies and crop production.

3. Methodology

3.1 Data Acquisition

To compile data for this analysis, the team acquired optical and climatic Earth observation data products, administrative boundary features, and land use/land cover (LULC) classifications (Table 1) for the study area and period using the Google Earth Engine (GEE) Catalog and NOAA National Centers for Environmental Information (NCEI) products. The team compiled image data from 1984 – 2011 from the Landsat 5 TM sensor, 2012 – 2013 from the Landsat 7 ETM+ sensor, and 2014 – 2023 from the Landsat 8 OLI sensor. The team also acquired rainfall and precipitation data from the ERA5 Monthly Aggregates dataset compiled by the European Centre for Medium Range Weather Forecasts (ECMWF). Gaps in precipitation data were filled by acquiring data from the daily Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) dataset by the United States Geological Survey (USGS) and the University of California Santa Barbara Climate Hazards Center (CHC).

The team then isolated the study region by extracting the state boundaries for Bahia, Mato Grosso, Pará, and Tocantins from the Level-1 Global Administrative Unit Layers (GAUL) feature collection in the Google Earth Engine Catalog. Finally, the team downloaded tabular data on monthly equatorial Pacific SST anomalies from the El Niño 3.4 dataset provided by NOAA NCEI.

Table 1
Earth observations, data products and source overview

Provider	Products	Spatial Resolution	Source
Copernicus Climate Change Service Climate Data Store (CDS)	ERA5 monthly aggregated total precipitation and mean 2m air temperature data provided by the European Centre for Medium Range Weather Forecasts (ECMWF)	27 km x 27 km	Google Earth Engine Catalog
Landsat Earth Observing System (EOS)	Landsat 5 Level 2 Collection 2, Tier 1, 1985 – 2013 Landsat 7 Level 2 Collection 2, Tier 1, 1999 – Present	30 m x 30 m	Google Earth Engine Catalog

	Landsat 8 Level 2 Collection 2, Tier 1, 2013 – Present		
MapBiomas	MapBiomas annual land use/land cover dataset, collection 8 1985 – 2021	30 m x 30 m	Google Earth Engine Catalog
NOAA	El Niño 3.4 monthly SST anomaly (5N-5S, 170W-120W)	N/A	NOAA NCEI
United States Geologic Survey (USGS) **University of California Santa Barbara Climate Hazards Center (CHC)**	Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)	0.05 x 0.05 degrees	Google Earth Engine Catalog
UN Food and Agriculture Service	Global Administrative Unit Layers (GAUL) provided by the United Nations Food and Agriculture Service	N/A	Google Earth Engine Catalog

3.2 Data Processing

The team conducted all image and raster processing in GEE. To process the Landsat imagery, the team filtered pixels for clouds and quality using the Quality Assessment (QA) bands, clipped to the four study states, and mosaiced. The team used the red bands (Band 3 in TM and ETM+, Band 4 in OLI) and the near-infrared (NIR) bands (Band 4 in TM and ETM+, Band 5 in OLI) to calculate NDVI, as photosynthetically active vegetation is highly reflective of NIR light and absorptive of red light (Equation 1) (Ji & Peters, 2003). As El Niño 3.4 data are compiled monthly, the team aggregated NDVI into monthly means over the study period.

Equation 1:

$$NDVI = (NIR - Red) / (NIR + Red) \tag{1}$$

To isolate soybean growing areas, the team used the MapBiomas LULC classification and created a mask that restricted analyses to pixels classified as soybean (value = 39) for each year in the study period (Souza et al., 2020). Using the annual soybean growing area mask, NDVI monthly mean values of soybean were extracted per year within the study region. The team then calculated mean monthly NDVI for soybean growing areas per state.

Finally, cumulative growing season NDVI was calculated by summing monthly NDVI values from planting through harvest season for each state (Table 2). As some monthly image composites throughout the time series were too cloudy to derive NDVI, missing NDVI data were temporally gap-filled by calculating the average NDVI of the two adjacent months.

Table 2

Peak NDVI curves for soybean growing season per state

State	Planting	Harvest	Peak NDVI curves for soybean grown areas per state
Bahia	Oct – Dec	Jan – May	Jan 1 – Feb 29
Mato Grosso	Oct – Dec	Jan – May	Dec 1 – Jan 31
Pará	Nov – Jan	March – Jun	Jan 1 – Feb 29
Tocantins	Oct – Dec	Jan – May	Dec 1 – Jan 31

Temperature and precipitation data were filtered to the study dates and aggregated into monthly means per study state. As some gaps in the ERA5 dataset presented temporal continuity issues, the team filled in gaps with the CHIRPS daily precipitation dataset for the year of 1984, July – December 2020, and 2021 – 2023. CHIRPS precipitation was aggregated into monthly means to correspond with data derived from ERA5. The team chose to use monthly aggregated ERA5 data over most of the study period due to computational challenges associated with processing daily CHIRPS data. Similarly, there were gaps presented in the temperature data that were initially extracted from ERA5, which included missing data for Pará 1984 and July – December 2020, and 2021 – 2023 for all four states. To fill in the gaps, the team utilized Climate Engine, which hosts a complete dataset for ERA5, and extracted the daily mean average for temperature for the missing dates. The team then calculated the monthly average for the acquired data to fill in the gaps that were presented in the initial data acquisition. When calculating the average monthly temperature, additional data gaps were found, as all states were missing December 28 – 31, 2022 temperature data. The ENSO anomaly data from the El Niño 3.4 dataset required no additional processing for this project.

3.3 Data Analysis

To analyze the relationship between ENSO conditions and crop productivity, the team first identified the lag time between El Niño SST anomalies and the impacts of these anomalies on NDVI, temperature, and precipitation. As conditions in the Pacific can take several months to impact temperature and precipitation in northern and central Brazil, and up to 12 months to affect NDVI, the data required temporal adjustment to account for the delayed effect (Erasmi et al., 2014). To determine the lag time, the team first conducted a cross-correlation function test in R with ENSO and the dependent variables: temperature, precipitation, and NDVI. This test uses Pearson's correlation coefficient across time series data to determine the dominant cross-correlation at each time interval. Pearson's correlation coefficient measures linear correlation between sets of data (Sedgwick, 2012). The team compared the results for ENSO against the dependent variables and selected the lag time with the most significant Pearson's correlation coefficient (*r*) (Figure 2). The team then applied a temporal lag adjustment to the ENSO anomalies to reflect locally experienced ENSO conditions.

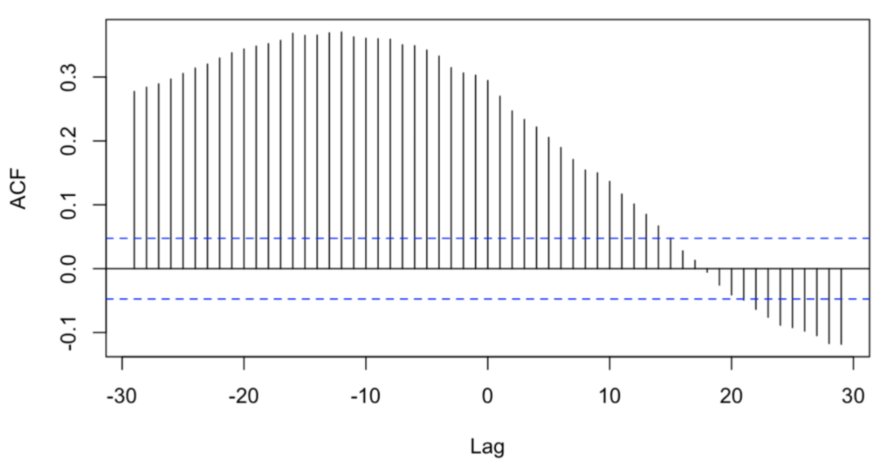


Figure 2. Autocorrelation function (ACF) representing the temporal relationship between ENSO and temperature anomalies.

The team then conducted a series of statistical analyses to examine the relationship between NDVI of soybean planted areas and lagged ENSO anomalies, temperature, and precipitation. The team performed linear regressions, principal component analysis (PCA), and a multivariable time series analysis in R to explore the relationship between the variables.

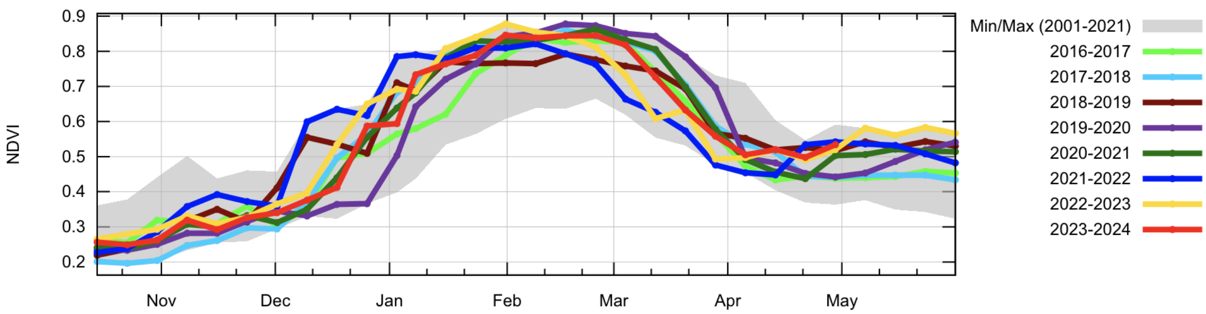


Figure 3. Annual soybean NDVI curves 2016-2024 in Western Bahia from the Global Inventory Monitoring and Modeling Studies (GIMMS) Global Agricultural Monitoring (GLAM). Figure derived from GLAM based on data from the Terra 8-day MODIS NDVI dataset and the 2020 GeoSpatial Data Analysis Corperation's Brazil Soybean crop mask.

4. Results & Discussion

4.1 Analysis of Results

In assessing the impact of ENSO on soybean crop production, this project found that El Niño and La Niña conditions vary per state and per year (1984 – 2023) along with SST anomalies, which were also detected in the study area and period. El Niño years were observed during the late 1990s, 2015, and the beginning of 2023, all of which featured high SST anomalies (Appendix 1). On the other hand, La Niña conditions were observed during the late 1980s, early 2000s, and the onset of 2020 (Appendix 2). These findings were crucial to serve as the baseline for understanding ENSO conditions in the study area during the study period.

Although there were significant variations in temperature between states throughout the calendar year, temperatures increased on average in all states over the study period (Appendix 4). When analyzing the temporal trend by month, the team found the most significant increases in temperature were in April, September, and October. Tocantins, in particular, experienced significant warming in September and October.

The analysis found strong seasonal trends in precipitation, with the lowest precipitation rates in the summer months across all the study states (Appendix 5). Total precipitation steadily decreased for every study state aside from Pará over the study period (Appendix 5). Pará experienced an increase in precipitation from February through April and a slight decrease in precipitation in July through September (Appendix 5). Meanwhile, Tocantins and Mato Grosso experienced a significant decrease in precipitation in December over the study period, which is important to note as December is a crucial month for the soy growing season (Appendix 5).

The cross-correlation analysis found higher ENSO anomalies correlated with higher temperatures and lower precipitation rates (Appendix 6). This relationship was found to be strongest with an eight-month lag. Furthermore, the results suggest that ENSO impacts local weather conditions in the study region eight months after SST were measured in the Pacific Ocean. The cross-correlation tests between monthly temperature/precipitation and monthly soy NDVI found the strongest correlation with a 0-month lag, meaning the study area experienced a real-time correlation between soy NDVI and climatic conditions when aggregated monthly (Appendix 7).

Due to the lagged relationship between ENSO and temperature/precipitation and the non-lagged relationship between temperature/precipitation and soy NDVI, the team expected to similarly find a correlation between ENSO anomaly and soy NDVI anomaly at an 8-month lag. However, a cross-correlation analysis between monthly ENSO anomaly and monthly soy NDVI found little to no relationship at various lag times (Appendix 8). When tested on each study state, the team found no statistically significant linear relationship between 8-month lagged ENSO anomaly and monthly soy NDVI in any of the states (Appendix 9).

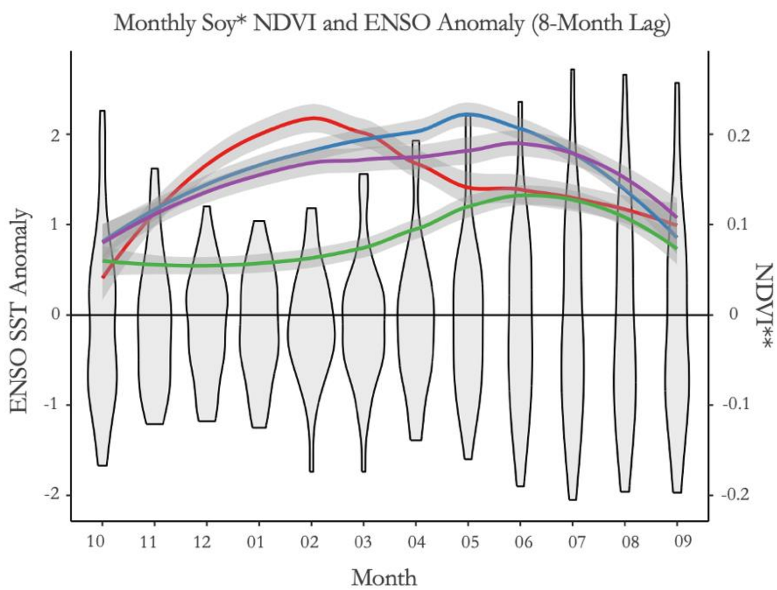


Figure 4. Violin plot of monthly ENSO Anomaly (8-Month Lag) and soy NDVI. *Soy crop mask by MapBiomas. **Scale of -1 to 1, values between 0 and 1 represent vegetation greenness, with higher values corresponding with increased greenness.

Figure 4 shows the monthly ENSO anomalies with an 8-month lag to reflect locally experienced ENSO effects which tend to be strongest in June through September. When compared with monthly soybean NDVI curves, which tend to peak in December through February as shown in Figure 4, there is a difference between periods with the strongest ENSO anomalies and the peak soy growing season. The results indicate that ENSO's impact on soybean production is limited in the region due to seasonal misalignment (Figure 4).

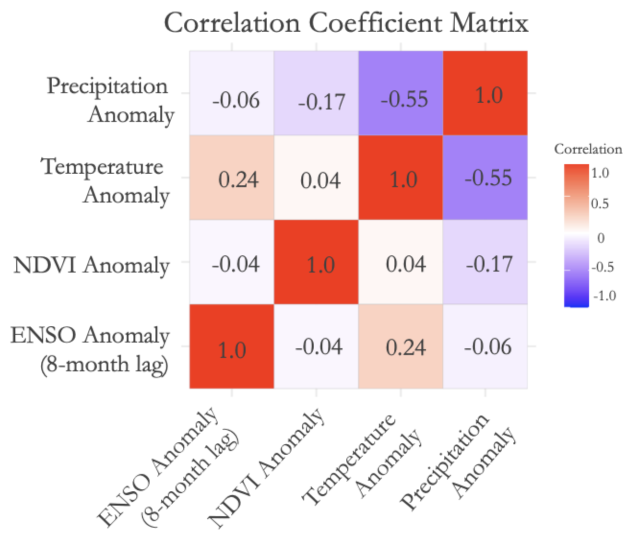


Figure 5. Correlation Coefficient Matrix depicting relationships among precipitation, temperature, NDVI, and ENSO anomalies.

The PCA results identified a positive relationship between ENSO anomaly and temperature (*r* = 0.24) and a slight negative relationship between 8-month lagged ENSO and precipitation (*r* = -0.06). It also identified little to no relationship between temperature anomaly and NDVI anomaly (*r* = 0.04) and a negative relationship between precipitation anomaly and NDVI anomaly (*r* = -0.17). Finally, the PCA analysis identified little to no relationship between 8-month lagged ENSO anomaly and NDVI anomaly (*r* = -0.04). While there was a positive correlation between 8-month lagged ENSO and temperature and a negative correlation between NDVI and precipitation, there was not a clear relationship between 8-month lagged ENSO and NDVI (Figure 5).

The team additionally conducted a state-level PCA analysis for each study area and found variations in all states in terms of relationships between ENSO anomalies, temperature, and precipitation. For example, ENSO anomalies had little to no effect on precipitation in both Bahia (Appendix 10) and Mato Grosso (Appendix 11) and slight negative effects in Pará (Appendix 12) and Tocantins (Appendix 13). Furthermore, Bahia had the weakest ENSO effects on temperature compared to the study areas, while Mato Grosso, Tocantins, and Pará exhibited positive effects of ENSO anomalies and temperature. Pará was the only state that exhibited a slight positive correlation between ENSO anomalies and soy NDVI which further suggests variations at the state level.

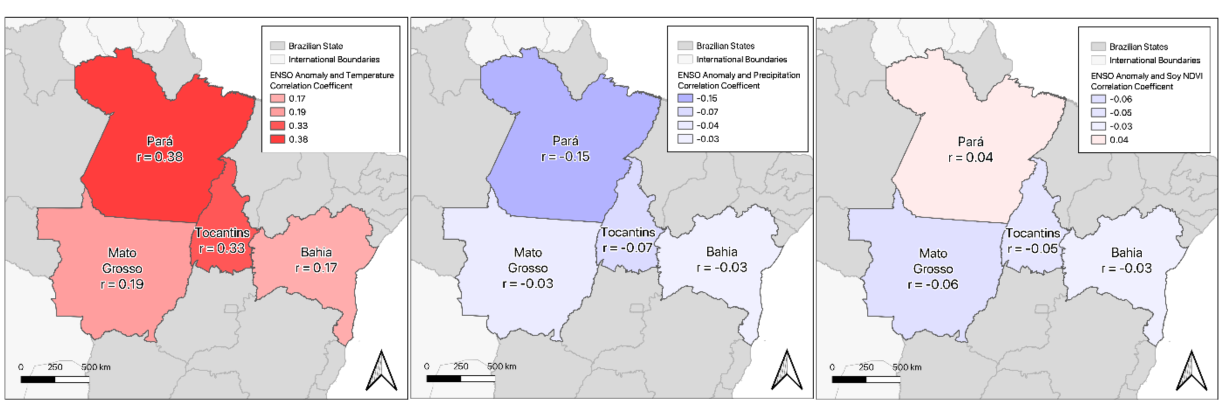


Figure 6. Spatial trends depicting ENSO Anomaly and Temperature Correlation Coefficient (left), ENSO Anomaly and Precipitation Correlation Coefficient (center), ENSO Anomaly and Soy NDVI Correlation Coefficient (right).

Figure 6 spatially visualizes results of correlation coefficients from components of the PCA analysis for each state showcasing ENSO anomalies and its effects on temperature, precipitation, and soy NDVI. The team found Pará to have slightly higher correlation coefficients in all relationships when compared to the other study states (Figure 6). Finally, a linear regression analysis between cumulative soy growing season NDVI and annual detrended production yield data found no statistically significant relationship between these variables (Appendix 14).

4.2 Feasibility for Partner Use
The project evaluated the use of ENSO conditions to forecast soybean crop productivity. Given the results that NDVI is not correlated with ENSO at state-level aggregations in the states of Bahia, Mato Grosso, Pará, Tocantins during the study period, ENSO conditions may have minimal impact on the health of soybean crops in this region of Brazil (Appendix 10). The methodology developed in this project may be useful for examining the impact of ENSO, temperature, and precipitation on crop productivity. Furthermore, a combination of temperature and precipitation data is viable to ascertain relationships among NDVI, seasonality, and development of changes in climate over time. Finally, correlation analyses producing time series and NDVI or crop yield maps provide a means of obtaining accurate trends in assessing which factors affect crop production yield. These results may be used by project partners in their future work to prepare crop outlooks for Brazil.

4.3 Methodological Limitations
The team found methodological limitations through the use of MapBiomas LULC classification which presented unverifiable information. Since the analysis relied on an external, machine-learning produced classification, limitations arise from being unable to validate its classification. Additionally, changing crop seasonality in the region creates uncertainty around the cumulative growing season NDVI metric. Further limitations faced in this project include missing NDVI values due to cloud coverage and missing temperature and precipitation data in the ERA5 dataset. The team addressed gaps in the NDVI data by calculating the mean NDVI value of the two adjacent months. Meanwhile to address missing temperature and precipitation years in the ERA5 dataset, the team supplemented the data with the CHIRPS dataset, which included taking the average monthly calculation for every state and missing months manually. In sum, potential human error from manually gap-filling NDVI and ERA5 datasets and methodological uncertainties from using an external land cover classification contribute to possible shortcomings of the project.

Future analyses can address these limitations by refining the methodology and including additional ENSO indices like those that account for atmospheric pressure. Forthcoming projects can also conduct a machine learning land cover classification using remote sensing data to ensure that information produced is verifiable and can be repeated in subsequent studies. Furthermore, future work can address the data limitations by incorporating additional datasets that contain temperature, precipitation, and other climatic data. Additional limitations include the lack of crop yield reference data at a sub-state level to conduct finer spatial resolution analysis of sub-state trends. Due to limited time, the team did not produce maps of LULC change trajectories or a time series maps of soy-masked NDVI which would have provided additional context into the spatial distribution and sub-state spatial heterogeneity in trends. Finally, uncertainties associated with the spatial resolution of the imagery and LULC datasets and the aggregation methods may also result in additional uncertainties and limitations.

4.4 Future Recommendations
The findings from this project lay the foundation for better understanding the impacts of ENSO conditions on crop production yield for four states in Brazil: Bahia, Mato Grosso, Pará, and Tocantins. In addition, this project identifies climatic relationships that connect soybean production yield to ENSO conditions including

temperature, precipitation, temporal lags, NDVI, seasonality, and SST anomalies. These relationships establish the groundwork for future work examining the impact of variable weather conditions on agriculture production as climate change proses significant threats to food security.

Although the team's findings indicate a strong temporal relationship among climatic conditions, NDVI, and seasonality, we found no association between ENSO and NDVI. Nevertheless, findings from this analysis establish methods for forthcoming research that can provide insights into the impact of ENSO on crop production yield. When conducting future research, it will be beneficial to refine the methodology and conduct an analysis using a multivariate index for SST and ENSO conditions. Using a multivariate SST index will potentially strengthen the relationship between climatic conditions and crop yield while also providing further insights that can aid USDA FAS in decision-making and crop forecasting. Moreover, the methods presented in this project will be useful to test on additional crops in Brazil, such as second planting corn and cotton – both of which Brazil is a leading producer – to examine if the findings are consistent across crops planted during different parts of the growing season. To resolve issues with cloud cover in the Landsat dataset, further research with finer temporal-scale multispectral imagery such as MODIS could provide additional insights into seasons affected by frequent cloud cover. Finally, the methodology can be applied to a finer geographic scale, such as a single state or at the sub-state level, as there can be significant climatic variation within states which can impact crop production. Future work on this issue could enhance final results interpretation through further spatial data visualizations of LULC change trajectories or a time series maps of soy-masked NDVI.

5. Conclusions

The project's finding provides a methodology to assess the impacts of climatic conditions on crop productions yield in each of the study areas: Bahia, Mato Grosso, Pará, and Tocantins. By utilizing multispectral imagery, a diverse array of datasets, and conducting a correlation analysis, the team showed the relationship of ENSO conditions on crop production. More specifically, the team's findings indicate that temperature and precipitation data are especially viable for determining how NDVI, seasonality, and time correlate with one another in relation to changes in climatic conditions. Furthermore, the methods presented in this project can be applied to future projects to understand the impacts of changes in climate on crop production in tropical regions. However, the limitations in available crop yield reference data, accurate long-term crop masks, and high-temporal resolution imagery presented obstacles to the analysis. Results can support the USDA Foreign Agriculture Service, International Production Assessment Division and the USDA Office of the Chief Economist and World Outlook Board by providing a comprehensive assessment of climatic changes impacting an agricultural commodity that can be used to ensure global food security and enhance US trade and foreign policy.

6. Acknowledgements

The team would like to thank project partners, Dr. Sunita Yadav-Pauletto and Mr. Mark Brusberg, and science advisors, Dr. Garrett Graham, Dr. Boyin Huang, Molly Woloszyn for their advice and guidance throughout the project. The team would also like to thank Center Lead, Katie Caruso, and Impact Analysis Fellow, Maya Hall for their edits and feedback on the report.

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 80LARC23FA024.

7. Glossary

ACF – Autocorrelation function
Anomalies – Variation from a 30-year climatic normal
CDS – Climate Data Store by the Copernicus Climate Change Service
CHC –Climate Hazards Center at the University of California Santa Barbara
CHIRPS – Climate Hazards Group InfraRed Precipitation with Station data
Earth observations – Satellites and sensors that collect information about the Earth's physical, chemical, and biological systems over space and time
ECMWF – European Centre for Medium Range Weather Forecasts
ENSO – El Niño-Southern Oscillation
EOS – Earth Observing System
ERA5 – fifth generation ECMWF atmospheric reanalysis
ETM+ – Enhanced Thematic Mapper sensor on the Landsat 7 satellite
FAS – Foreign Agriculture Service
GAUL – Global Administrative Unit Layers
GEE – Google Earth Engine
GIMMS – Global Inventory Monitoring and Modeling Studies
GLAM – Global Agricultural Monitoring
IPAD – International Production Assessment Division
NCEI – National Centers for Environmental
NDVI – Normalized difference vegetation index
NIR – Near infrared
NOAA – National Oceanic and Atmospheric Administration
LULC – Land use/land cover classification
MODIS – Moderate Resolution Imaging Spectroradiometer sensor mounted on the Terra and Aqua satellites
OLI – Operational Land Imager sensor on the Landsat 8 satellite
PCA – The objectives of PCA are to reduce the dimensionality of large datasets by transforming many variables into smaller summary indices which can be used to identify meaningful relationships between underlying variables (Paul et al., 2013)
QA – Quality Assessment
SST – Sea surface temperature
TM – Thematic Mapper sensor on the Landsat 5 satellite
USDA – United States Department of Agriculture
USGS – United States Geological Survey

8. References

Anderson, W., Seager, R., Baethgen, W., & Cane, M. (2017). Crop production variability in North and South America forced by life-cycles of the El Niño Southern Oscillation. *Agricultural and Forest Meteorology*, *239*, 151-165.

Arvor, D., Dubreuil, V., Simões, M., & Bégué, A. (2013). Mapping and spatial analysis of the soybean agricultural frontier in Mato Grosso, Brazil, using remote sensing data. *GeoJournal*, *78*, 833-850.

Arvor, D., Dubreuil, V., Ronchail, J., Simões, M., & Funatsu, B. M. (2014). Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil). *International Journal of Climatology*, *34*(8), 2622-2633.

Avanzi, J. C., Viola, M. R., Mello, C. R. D., Giongo, M. V., & Pontes, L. M. (2019). Modeling of the rainfall and R-Factor for Tocantins state, Brazil. *Revista Brasileira de Ciência do Solo*, *43*, e0190047.

Batistella, M., & Valladares, G. S. (2009). Farming expansion and land degradation in Western Bahia, Brazil. *Biota Neotropica*, *9*, 61-76.

Cattelan, A. J., & Dall'Agnol, A. (2018). The rapid soybean growth in Brazil. *OCL*, *25*(1), D102.

Carvalho, R., Adami, M., Amaral, S., Bezerra, F. G., & de Aguiar, A. P. D. (2019). Changes in secondary vegetation dynamics in a context of decreasing deforestation rates in Pará, Brazilian Amazon. *Applied Geography*, *106*, 40-49.

Cirino, P. H., Féres, J. G., Braga, M. J., & Reis, E. (2015). Assessing the Impacts of ENSO-related Weather Effects on the Brazilian Agriculture. Procedia Economics and Finance, 24, 146–155.

Colussi, J., & Schnitkey, G. (2021). Brazil likely to remain world leader in soybean production. farmdoc daily, 11(105).

Erasmi, S., Schucknecht, A., Barbosa, M., & Matschullat, J. (2014). Vegetation Greenness in Northeastern Brazil and Its Relation to ENSO Warm Events. *Remote Sensing*, *6*(4), 3041–3058.

GLAM-Global Agricultural Monitoring. USDA FAS. Available online: https://glam1.gsfc.nasa.gov (accessed on February 15th, 2024).

Ji, L., & Peters, A. J. (2003). Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. Remote sensing of Environment, 87(1), 85-98.

Júnior, R. D. S. N., & Sentelhas, P. C. (2019). Soybean-maize off-season double crop system in Brazil as affected by El Niño Southern Oscillation phases. *Agricultural Systems*, *173*, 254-267

Martorano, L. G., Siviero, M. A., Monteiro Tourne, D. C., Vieira, S. B., Fitzjarrald, D. R., Vettorazzi, C. A., ... & Lisboa, L. S. S. (2016). Agriculture and forest: A sustainable strategy in the Brazilian Amazon. *Australian Journal of Crop Science*, *10*(8), 1136-1143.

Paul, L. C., Suman, A. A., & Sultan, N. (2013). Methodological analysis of principal component analysis (PCA) method. International Journal of Computational Engineering & Management, 16(2), 32-38.

Reis, L., Santos e Silva, C. M., Bezerra, B., Mutti, P., Spyrides, M. H., Silva, P., ... & Andrade, L. (2020). Influence of climate variability on soybean yield in MATOPIBA, Brazil. Atmosphere, 11(10), 1130.

Richards, P., Pellegrina, H., VanWey, L., & Spera, S. (2015). Soybean development: The impact of a decade of agricultural change on urban and economic growth in Mato Grosso, Brazil. *PLoS one*, *10*(4), e0122510.

Sedgwick, P. (2012). Pearson's correlation coefficient. *Bmj*, *345*.

Sentelhas, P. C., & Pereira, A. B. (2019). El Niño–Southern oscillation and its impacts on local climate and sugarcane yield in Brazil. *Sugar Tech*, *21*(6), 976-985.

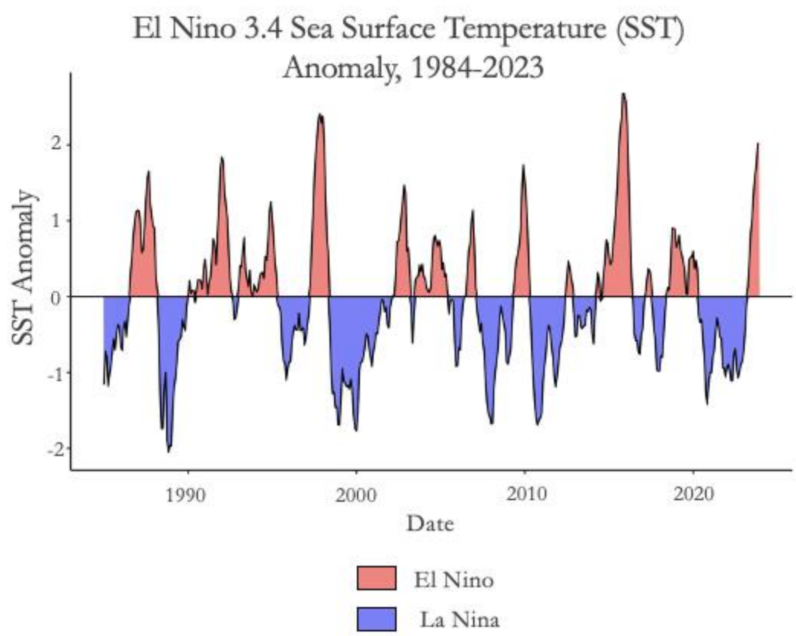
Souza, C. M., Z. Shimbo, J., Rosa, M. R., Parente, L. L., A. Alencar, A., Rudorff, B. F. T., Hasenack, H., Matsumoto, M., G. Ferreira, L., Souza-Filho, P. W. M., De Oliveira, S. W., Rocha, W. F., Fonseca, A. V., Marques, C. B., Diniz, C. G., Costa, D., Monteiro, D., Rosa, E. R., Vélez-Martin, E., … Azevedo, T. (2020). Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sensing*, *12*(17), 2735.

USDA Foreign Agriculture Service. (2024). Brazil Country Summary. Accessed February 14, 2024. https://ipad.fas.usda.gov/countrysummary/Default.aspx?id=BR&crop=Soybean

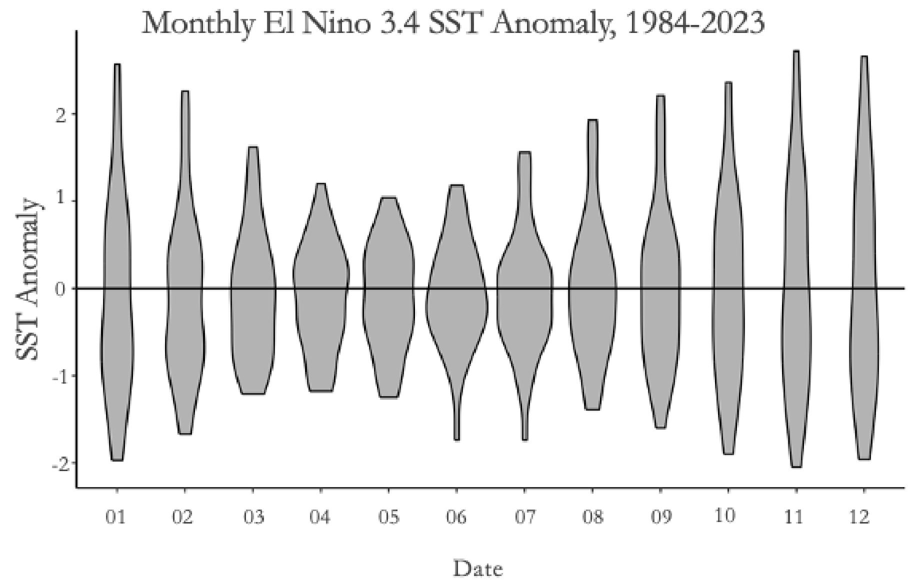
USDA IPAD Foreign Agriculture Service. (2024). Brazil Country Production Summary. Accessed February 27th, 2024. https://ipad.fas.usda.gov/countrysummary/default.aspx?id=BR

Valdes, C. (2022). Brazil's Momentum as a Global Agricultural Supplier Faces Headwinds. United States Department of Agriculture Economic Research Services. https://www.ers.usda.gov/amber-waves/2022/september/brazil-s-momentum-as-a-global-agricultural-supplier-faces-headwinds/

9. Appendices



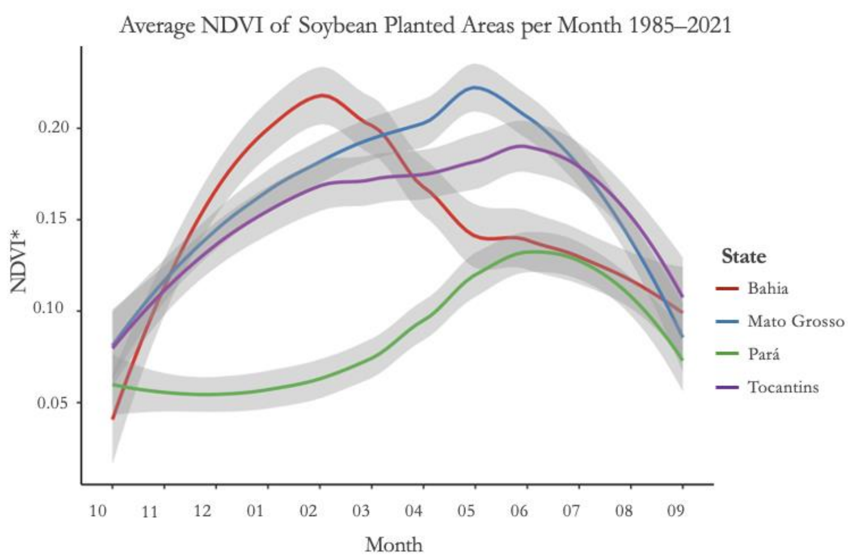
Appendix 1. El Niño sea surface temperature anomaly from 1984 – 2023



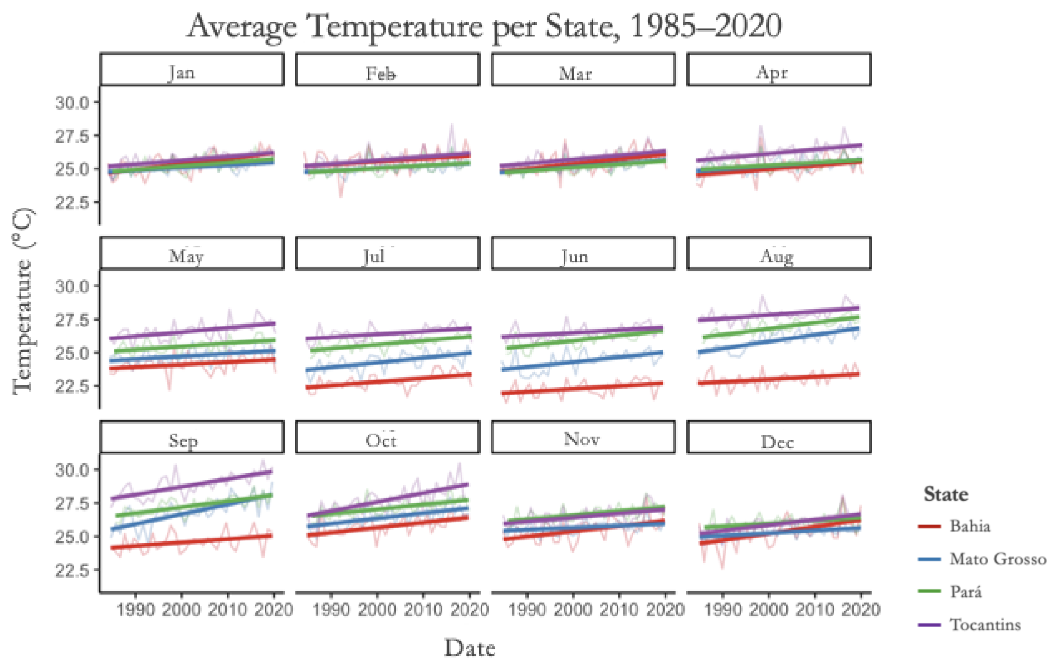
Appendix 2. Monthly El Niño SST anomaly from 1984 – 2023.

State	Growing Season
Pará	Nov–July
Mato Grosso	Oct–June
Tocantins	Oct–June
Bahia	Oct–June

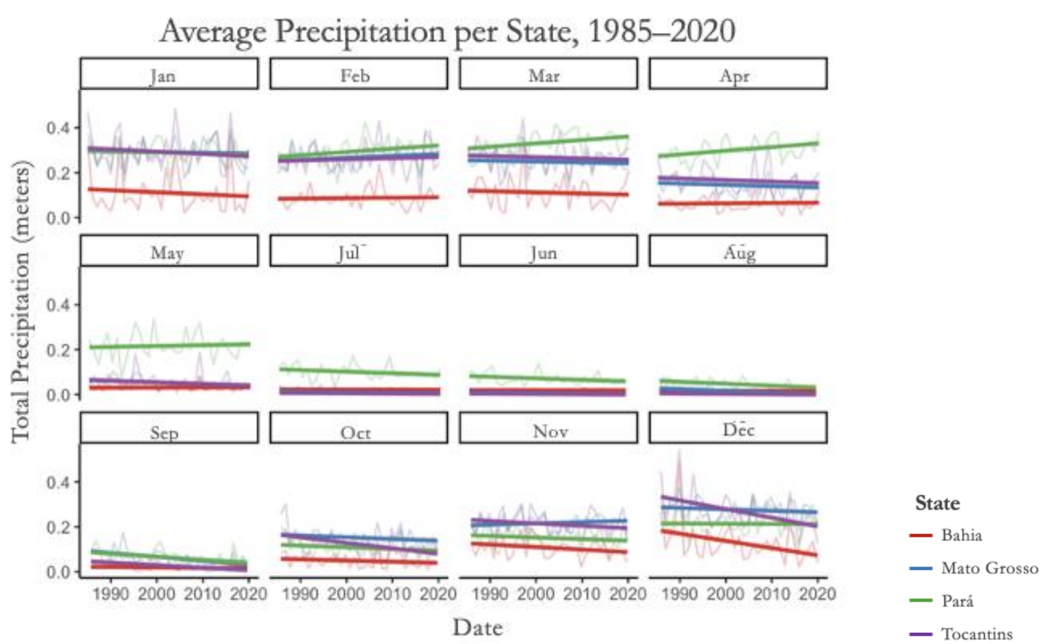
*Scale of -1 to 1, values between 0 and 1 represent vegetation greenness, with higher values corresponding with increased greenness



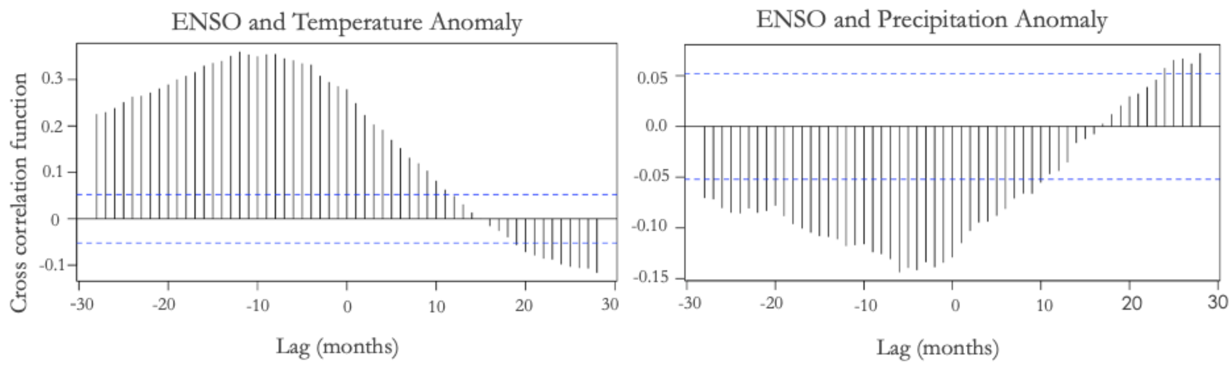
Appendix 3. Average NDVI of soybean planted areas per month between 1985 – 2021. Note the varying growing seasons for each state included in the study.



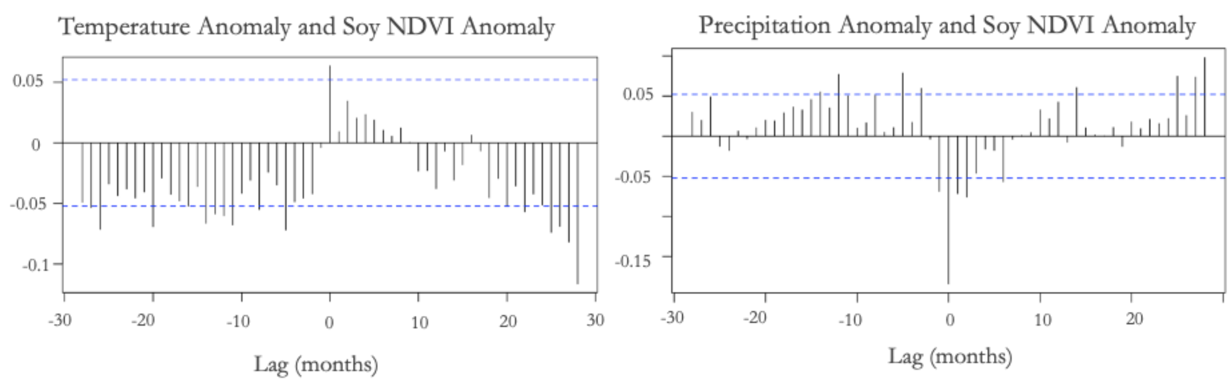
Appendix 4. Average temperature per state from 1985 – 2020.



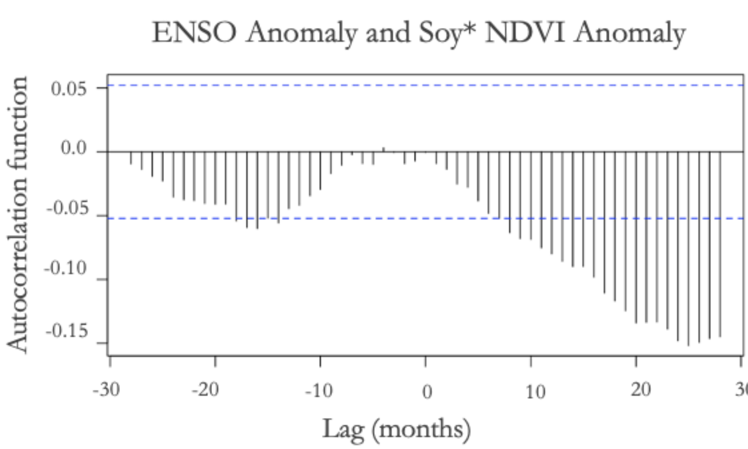
Appendix 5. Average precipitation per state from 1985 – 2020.



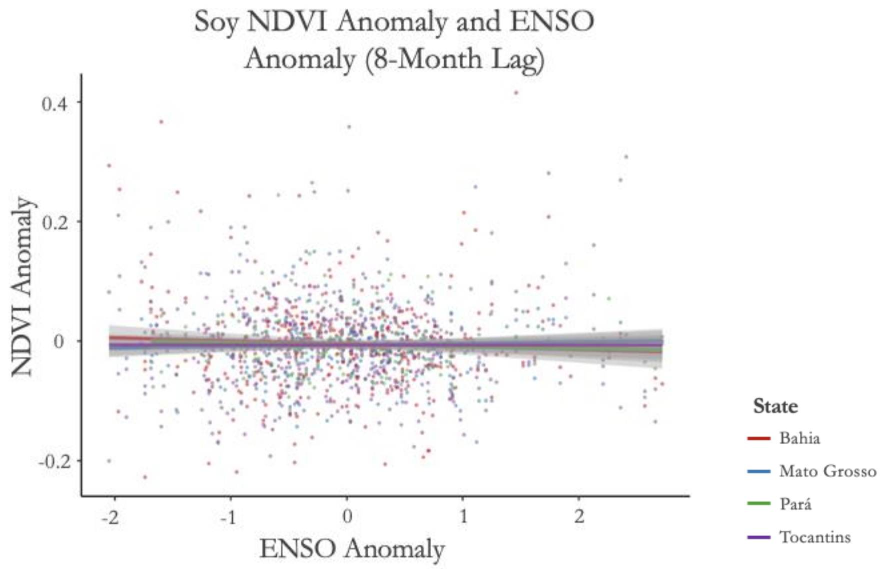
Appendix 6. Comparison of ENSO and temperature and precipitation anomalies. Note the 8-month lag between ENSO anomalies and precipitation and temperature anomalies in the study region.



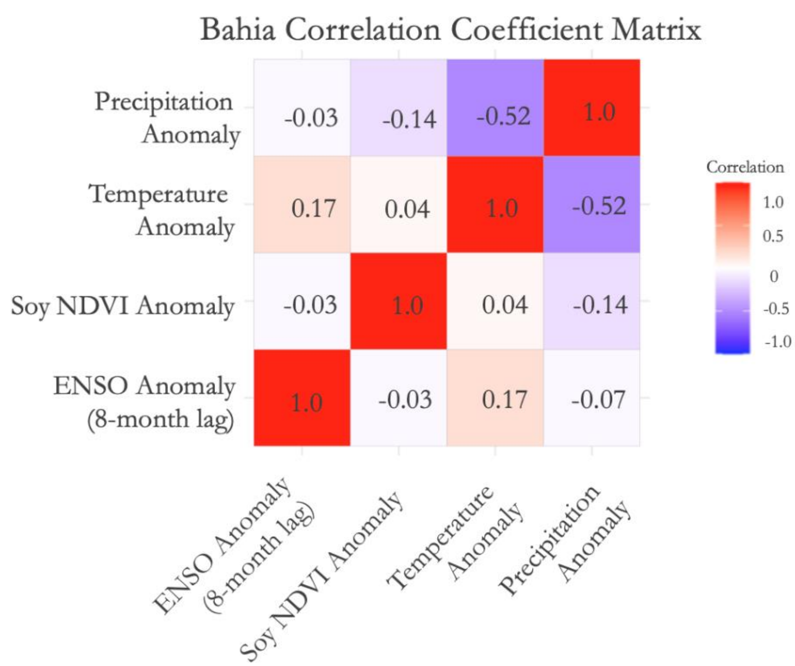
Appendix 7. Comparison of temperature and precipitation and soy NDVI anomalies. Note that there is not a lag observed between NDVI and climatic conditions in the study region. Soy NDVI by MapBiomas.



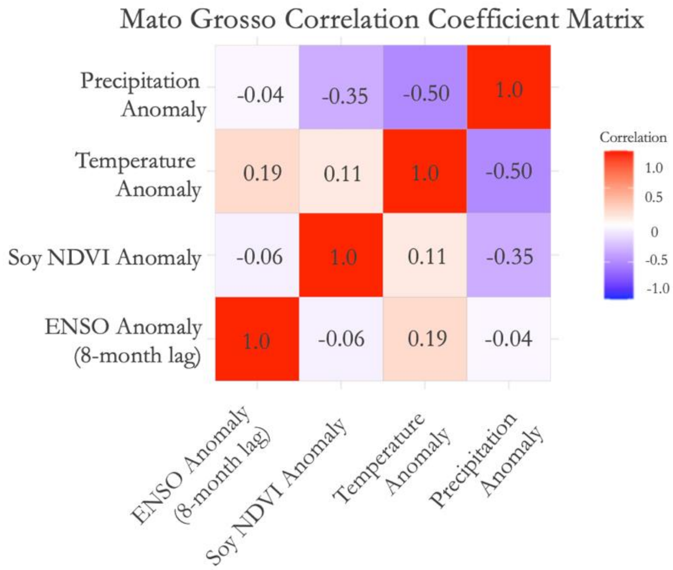
Appendix 8. ENSO anomaly and soybean NDVI anomaly in the study region. Note that there is not a clear lag or statistically significant relationship between ENSO anomaly and NDVI. *Soy crop mask by MapBiomas.



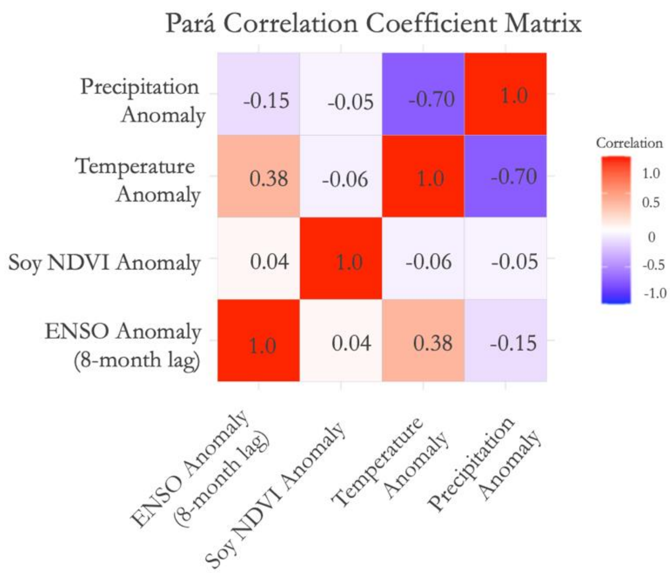
Appendix 9. Soy NDVI anomaly and ENSO anomaly with an 8-month lag.



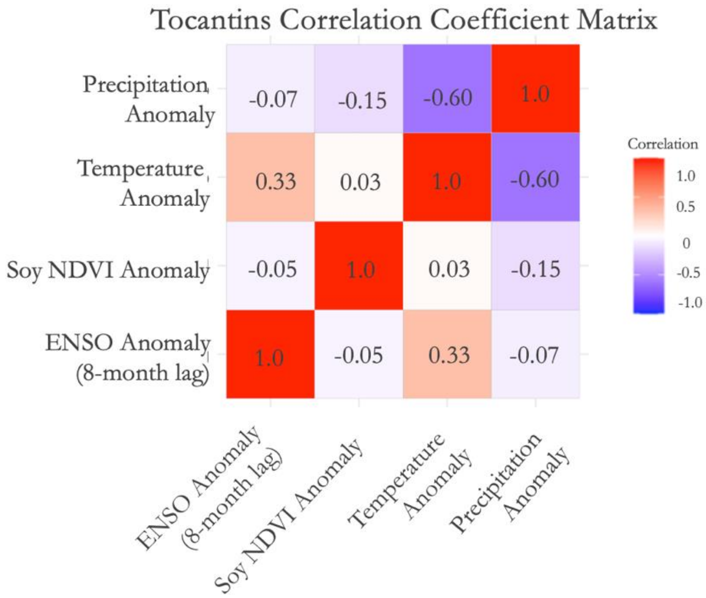
Appendix 10. Correlation Coefficient Matrix depicting relationships among precipitation, temperature, NDVI, and ENSO anomalies for Bahia



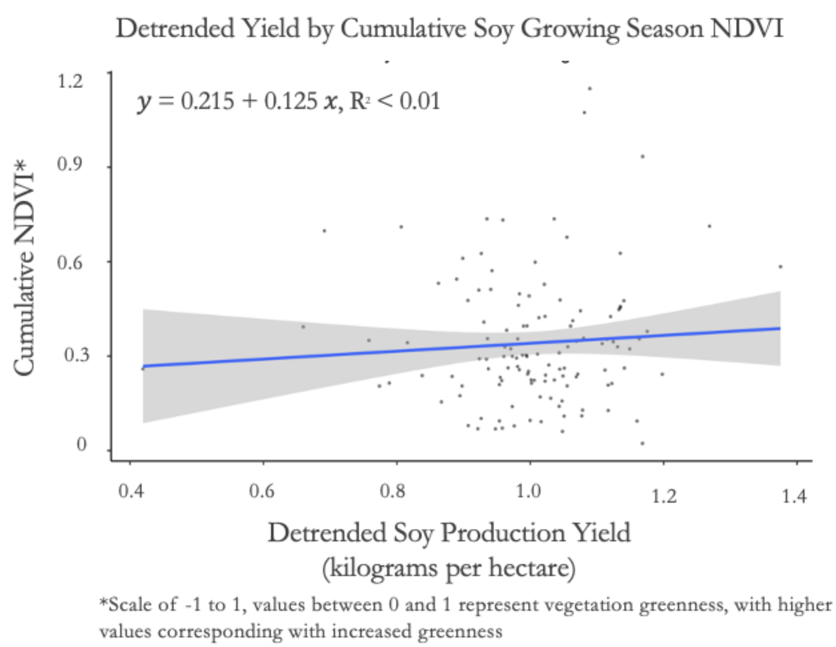
Appendix 11. Correlation Coefficient Matrix depicting relationships among precipitation, temperature, NDVI, and ENSO anomalies for Mato Grosso



Appendix 12. Correlation Coefficient Matrix depicting relationships among precipitation, temperature, NDVI, and ENSO anomalies for Pará



Appendix 13. Correlation Coefficient Matrix depicting relationships among precipitation, temperature, NDVI, and ENSO anomalies for Tocantins



Appendix 14. Detrended yield by cumulative growing season NDVI.