IMPACT OF CLIMATE CHANGE AND LAND DEGRADATION ON GLOBAL FOOD PRODUCTION

Ramesh Athe[1*], Ameyotosh Michael Roy[2†], Mohammed Abdul Haseeb[2†], Yashu Mittal[2†]

[1]Assistant professor, Data Science, Indian Institute of Information Technology Dharwad, Dharwad, 580009, Karnataka, India.
[2]Undergraduate, Computer science and Engineering, Indian Institute of Information Technology Dharwad, Dharwad, 580009, Karnataka, India.

*Corresponding author(s). E-mail(s): dr.athe9@gmail.com;
Contributing authors: ameyotoshmichael@gmail.com; m.haseeb311@gmail.com; yashumittal2020@gmail.com;
†These authors contributed equally to this work.

Abstract

Using a large dataset from the World Bank, this study explores the complex connections between land degradation, climate change, and world food production. Machine learning techniques and the AutoRegressive Integrated Moving Average (ARIMA) model are used in this work to anticipate agricultural productivity and environmental factors using data from 100 nations and seven important indicators. The study illustrates the DecisionTreeRegressor model39;s improved effectiveness in forecasting future trends in food production through careful data imputation and model evaluation. The study highlights the significance of comprehending and tackling environmental concerns while providing insightful information about the intricate processes influencing global food security. The findings seek to encourage the development of resilient and adaptive methods to lessen the impact of climate change and land degradation on food production by promoting transparency and creativity in policymaking and agricultural planning.

Keywords: Climate change, Land degradation, Food Production Index, Machine learning, DecisionTreeRegressor

1 Introduction

In a time when climate change and land degradation pose enormous difficulties, it is critical to comprehend the complex dynamics of global food production. A wealth of information about many variables relevant to the world's food production and its underlying factors can be found on the World Bank[1]. With the use of this abundance of data, our study sets out to conduct a thorough examination with the goal of identifying the crucial elements affecting food production worldwide within this scenario of environmental sustainability[2].

It is impossible to overestimate how much land degradation and climate change affect the world's food production. Global economic stability, food security, and agricultural productivity faces significant threat by these phenomena[3][4]. Our research aims to provide light on methods to lessen these difficulties and promote sustainable agricultural practices by exploring the intricate relationship between environmental conditions and the dynamics of food production.

Recognizing the difficulty of missing data, we take a judicious approach, impute missing values across the dataset using four important factors: average contributions of agriculture and forestry to the economy, fertilizer usage, GDP of countries, and farmland extent. For each of the seven selected indications, this method results in the formation of 11 unique categories, opening the door for a thorough and thorough investigation.

We use the Autoregressive Integrated Moving Average (ARIMA) model for time series forecasting in addition to machine learning algorithms to investigate how environmental conditions affect agricultural production[5]. The ARIMA model helps us predict future patterns in food production by utilizing trend analysis and historical data, which offers stakeholders, researchers, and policymakers insightful information.

Essentially, we want to provide stakeholders with useful information to help them deal with the obstacles of a world that is changing quickly and promote sustainable agricultural practices by analyzing the intricate relationships that exist between land degradation, climate change, and the dynamics of global food supply..

2 Related Work

Numerous studies have emphasized the substantial impact of climate change and land degradation on global food supply, underlining the importance of understanding and addressing these issues[2][3]. The relationship between environmental changes and agricultural progress has been a major focus of research, with scientists adopting a variety of approaches to uncover the complex interdependencies within these systems.

Previous research has emphasized the significance of using extensive datasets to capture the numerous elements influencing food production. Researchers have underlined the significance of limiting datasets to representative samples to expedite studies and extract valuable insights.

However, issues such as missing data have encouraged the progress of novel approaches to data imputation, which ensure the reliability and validity of research findings[1][6]. In recent years, machine learning approaches have developed as effective tools for assessing the complex interactions between environmental variables and food production outcomes. Models such as DecisionTreeRegressor, LinearRegression, Lasso, and Ridge have helped to identify nuanced relationships, provide useful forecasting capabilities, and improve our knowledge of complex systems[1][7].

While prior research has provided useful insights into the relationship between climate change, land degradation, and global food security, there are still gaps in our understanding of these interactions. The limited exploration of specific approaches linked to data curation and machine learning model selection allows the potential for further research and enhancement in this field.

This paper intends to contribute to the current research by providing a detailed evaluation of the approaches used to study the influence of climate change and land degradation on global food supply. By giving thorough insights into data curation methodologies, machine learning model building, and performance evaluation metrics, this study aims to fill significant knowledge gaps and provide the framework for future research on this vital topic. Through transparent approaches and thorough research, the purpose is to encourage dialogue and innovation in environmental science and agriculture, moving toward resilient and adaptive policies in response to rising environmental concerns[5][7].

3 Methodology

3.1 Data Selection

The first stage was to collect pertinent data about the influence of climate change and land degradation on global food production. The dataset, which originally had information from over 200 countries, was narrowed down to the 100 countries that were most relevant to the study's aims, to streamline and concentrate the analysis.

3.2 Data Selection

After establishing the dataset, the study focused on specific factors to investigate the effects of climate change and land degradation on the global food supply. Seven major variables were selected, including critical components such as arable land, irrigation, precipitation, agricultural land, CO2 emissions, climate-related mortality, and freshwater withdrawals, to provide a thorough picture of the elements pertinent to the research question.

The seven indicators are-

1. Annual freshwater withdrawals, total (% of internal resources)

2. Droughts, floods, extreme temperatures (% of population, average 1990-2009)

3. CO_2 emissions from gaseous fuel consumption (% of total)

4. Average precipitation in depth (mm per year)

5. Agricultural irrigated land (% of total agricultural land)

6. Arable land (% of land area)

7. Agricultural land (% of land area)

3.3 Data Cleaning and Validation

Recognizing the presence of missing values in the dataset, steps were made to guarantee that the analysis remained valid despite these gaps.

3.3.1 Category-based Mean Calculation

To address missing data, we computed the average values for each indicator for each group. This phase was crucial since it allowed us to substitute missing numbers with customized calculations for each nation category. We recognize the importance of correct imputation in maintaining the credibility of our findings. To ensure accurate imputation, we analyzed the impact of agriculture, forestry, and fisheries on food production, fertilizer use, cropland percentage, and GDP.

3.3.2 Imputation Process

We employed a mean technique to replace missing values in the dataset, predicting mean values for key variables within each nation group. Missing data was handled independently for each of the 11 categories and seven indications to guarantee consistency and dataset integrity. We preserved the dataset as separate CSV files to ensure transparency and traceability during the imputation procedure. Validation relied on an iterative technique that tracked progress at each level. Mean imputation was used for both the 'foodproduction.csv' and 'categories.csv' datasets to ensure consistent results throughout the research.

Algorithm used for Imputation:
Step 1: Define Search Function:
Create a search function to check if a string is in a vector. Step 2: Main Function:

- Open'data.csv' file.
- Initialize code and value vectors.
- Read CSV, and store codes and values in vectors.

Step 3: Define Category Vector:

- Define the 2D vector category for country categories.

Step 4: Imputation Loop:

- Iterate over categories, indicators, and values.
- Calculate the mean and replace missing values.

Step 5: Write to New CSV:

- Open 'updated.csv'.
- Write codes and updated values.
- Close the file.

3.4 Machine Learning Model Development

We created a machine-learning model in accordance with completed imputed dataset. This was an exciting stage since it allowed us to apply our theoretical knowledge to real-world situations. The model aimed to investigate how climate change, land degradation, and global food production are interconnected across nation categories, providing valuable insights into the complexities of these relationships. The ML model development process involves the following steps:

3.4.1 Data Integration

The imputed information was combined with the food index dataset to provide full data on climate change, land degradation, and chosen indicators along with insights into world food production.

3.4.2 Machine Learning Model Selection

Our research utilized machine learning algorithms to predict the impact of climate change and land degradation on global food production. We chose four models: DecisionTreeRegressor, LinearRegression, Lasso, and Ridge. These models were chosen for their ability to conduct regression tasks and offer insights into variable relationships.

3.4.3 Model Evaluation Metrics

We assessed each machine learning model's performance using three key metrics: meansquarederror, meanabsoluteerror, and r2score. These metrics evaluated each model's accuracy, precision, and overall fit to the dataset.

After careful examination, the DecisionTreeRegressor model consistently outperformed the other models in the given metrics. Our dataset was best suited to its capacity to forecast complex data linkages accurately. The decision tree model effectively analyzes the influence of climate change and land degradation on global food production.

3.5 ARIMA Model Development

One of the main tools used in time series analysis to predict future trends and patterns in sequential data is the AutoRegressive Integrated Moving Average (ARIMA) model. In the framework of our study, the complex effects of land degradation and climate change are taken into account while predicting future food production indices using the ARIMA model. Understanding ARIMA:
ARIMA comprises three key components: autoregressive (AR), integrated (I), and

moving average (MA).

Component of Autoregressive (AR): An observation and a set of lag observations are compared using the autoregressive component. Put another way, it evaluates the impact of the time series' historical values on current values. 'p', which stands for the number of lag observations in the model, is used to indicate the AR component.

The Integrated (I) Component To achieve stationarity, the integrated component includes differentiating the time series. In time series analysis, stationarity is an essential concept that guarantees statistical attributes like variance and mean don't change over time. The degree of differencing needed to achieve stationarity is indicated by the differencing parameter 'd'.

Component of Moving Average (MA): The link between an observation and a linear mixture of previous error terms is modeled by the moving average component. It records transient oscillations or shocks that last for a while. 'q' stands for the moving average component and represents the size of the moving average window.

3.5.1 How ARIMA Works

Data Preprocessing: The first stage is to prepare the dataset for analysis by handling missing values, making sure the data is coherent, and formatting the data appropriately.

Determining the Model Parameters: The best values for 'p' 'd', and 'q' are found by applying time series analysis techniques like the autocorrelation function (ACF) and partial autocorrelation function (PACF). The setup and forecasting fitness of the ARIMA model are largely dependent on these factors.

Model fitting: After the parameters are established, methods such as maximum likelihood estimation are used to fit the ARIMA model to the dataset. The time series data's underlying patterns and dependencies are captured by the fitted model.

Model Evaluation: Metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared are used to assess the ARIMA model's performance. To ensure the model is reliable and adequate, diagnostic tests and residual analysis are also carried out.

Forecasting: Future values of the time series can be predicted with a well-fitted ARIMA model. To produce forecasts for future time points, the model makes use of previous data and the patterns that have been found. This gives important insights into probable trends and patterns.

In the final analysis, a strong basis for forecasting upcoming patterns in sequential data, such as food production indices, is provided by the ARIMA model. The ARIMA model takes into account the effects of land degradation and climate change while offering insightful information on the intricate dynamics of global food production through the utilization of its components, autoregression, differencing, and moving averages.

3.6 Streamlit App Development

Integrating the ARIMA forecasting tool into the user-friendly Streamlit app transformed accessibility and usability. Users could easily enter certain countries' names and indicator codes into the app's UI. The program dynamically generates graphs in response to user inputs, allowing for quick viewing of ARIMA projections for food production indicators. This interactive tool enabled users to easily examine forecasted trends and patterns, improving their understanding of the predicted effects of climate change and land degradation on global food production. The Streamlit app enabled full investigation and analysis of ARIMA forecasts using a user-friendly interface and real-time visualization capabilities, empowering users to make informed decisions in agricultural planning, environmental protection, and policy creation.

4 Results and Discussion

Our research, based on painstaking data curation and advanced machine learning techniques, provided remarkable insights into the complex dynamics of climate change, land degradation, and global food production. Using the DecisionTreeRegressor model as our primary predictive tool, we handled the complexity of our large dataset with precision. The model consistently outperformed others, demonstrating higher performance across key evaluation criteria.

Evaluation Metrics:
Mean Squared Error (MSE):
The DecisionTreeRegressor model had an MSE of 888.67, showing that it can minimize squared discrepancies between projected and actual values, resulting in accurate forecasts.

Mean Absolute Error (MAE):
With an MAE of 20.92, the model demonstrated its ability to properly estimate the magnitude of errors, resulting in reliable forecasts with low departure from actual values.

R-squared (R^2):
Despite a somewhat negative R-squared score of -0.038, the model showed proficiency in capturing variance in the dataset, with space for improvement.

The findings highlight the vital relevance of understanding the complex interplay between environmental conditions and food production. Our novel categorization algorithms and rigorous evaluation measures ensure the dependability and robustness of our findings, laying the groundwork for future research endeavors.

Furthermore, our analysis provides a critical foundation for dealing with the multiple issues to global food security. Our research encourages policymakers, agricultural stakeholders, and environmental advocates to dialogue and act by fostering transparency, repeatability, and innovation. In summary, our discoveries not only enhance academic debate but also possess wide range ramifications for developing resilient and adaptable measures to offset the effects of climate change and land degradation on global food production.

5 Figures

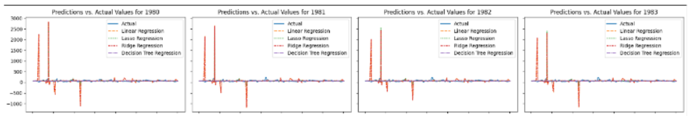

Fig. 1

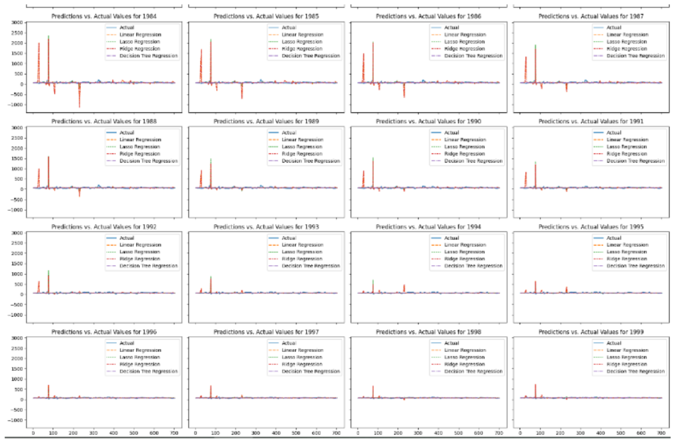

Fig. 2

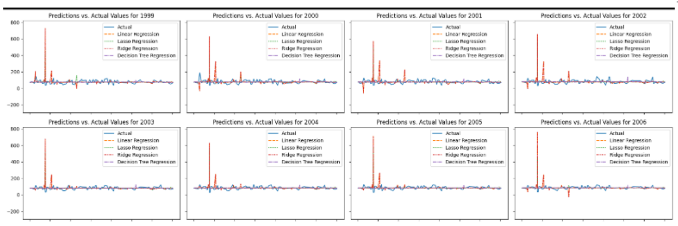

Fig. 3

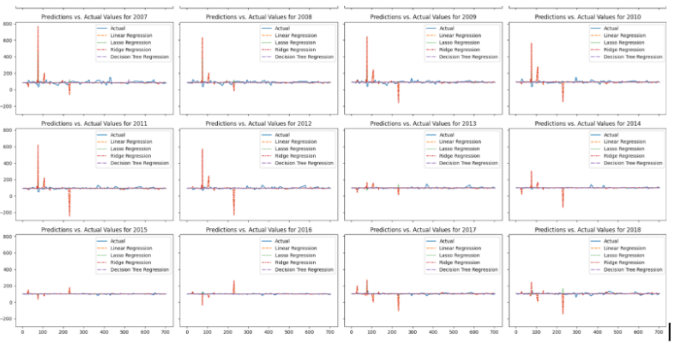
Predictions vs. Actual Values for 2007
Predictions vs. Actual Values for 2008
Predictions vs. Actual Values for 2009
Predictions vs. Actual Values for 2010
Predictions vs. Actual Values for 2011
Predictions vs. Actual Values for 2012
Predictions vs. Actual Values for 2013
Predictions vs. Actual Values for 2014
Predictions vs. Actual Values for 2015
Predictions vs. Actual Values for 2016
Predictions vs. Actual Values for 2017
Predictions vs. Actual Values for 2018
Actual
Linear Regression
Lasso Regression
Ridge Regression
Decision Tree Regression

Fig. 4

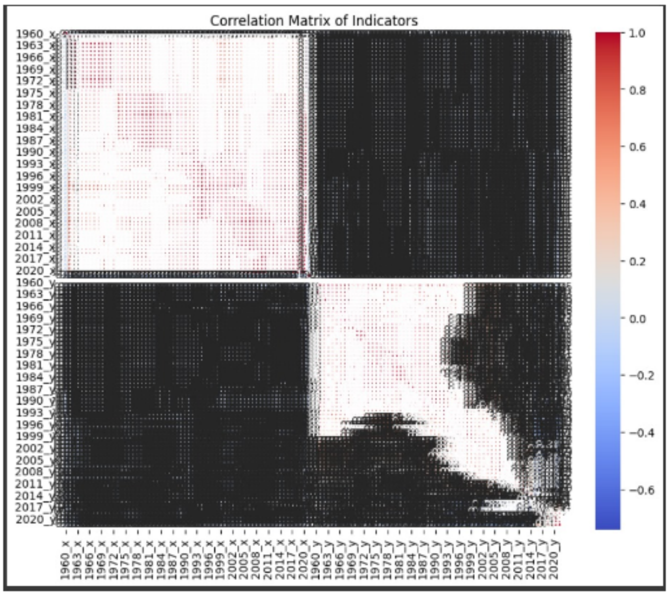
Correlation Matrix of Indicators
1960_x
1963_x
1966_x
1969_x
1972_x
1975_x
1978_x
1981_x
1984_x
1987_x
1990_x
1993_x
1996_x
1999_x
2002_x
2005_x
2008_x
2011_x
2014_x
2017_x
2020_x
1960_y
1963_y
1966_y
1969_y
1972_y
1975_y
1978_y
1981_y
1984_y
1987_y
1990_y
1993_y
1996_y
1999_y
2002_y
2005_y
2008_y
2011_y
2014_y
2017_y
2020_y
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

Fig. 5

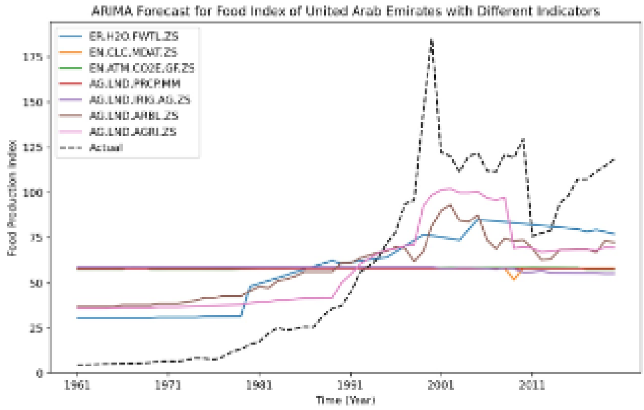
ARIMA Forecast for Food Index of United Arab Emirates with Different Indicators
ER.H2O.FWTL.ZS
EN.CLC.MDAT.ZS
EN.ATM.CO2E.GF.ZS
AG.LND.PRCP.MM
AG.LND.IRIG.AG.ZS
AG.LND.ARBL.ZS
AG.LND.AGRI.ZS
Actual
Food Production Index
Time (Year)
1961
1971
1981
1991
2001
2011
0
25
50
75
100
125
150
175

Fig. 6

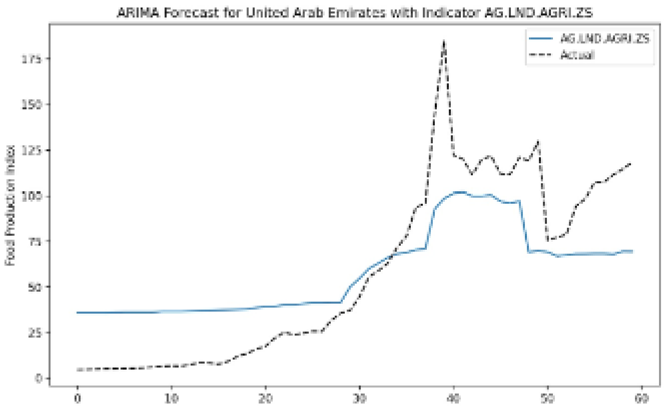
ARIMA Forecast for United Arab Emirates with Indicator AG.LND.AGRI.ZS
AG.LND.AGRI.ZS
Actual
Food Production Index
0
10
20
30
40
50
60
0
25
50
75
100
125
150
175

Fig. 7

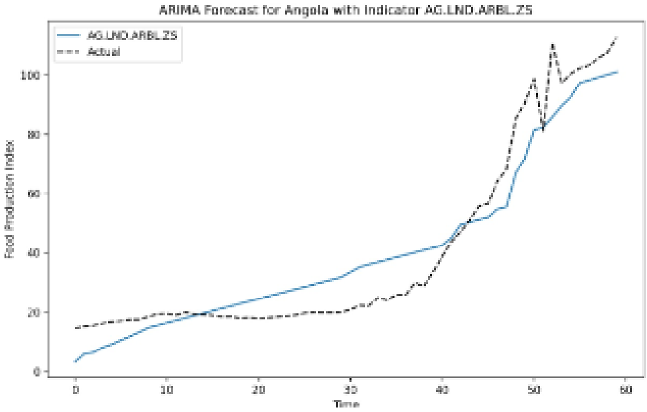
ARIMA Forecast for Angola with Indicator AG.LND.ARBL.ZS
AG.LND.ARBL.ZS
Actual
Food Production Index
Time
0
20
40
60
80
100
0
10
20
30
40
50
60

Fig. 8

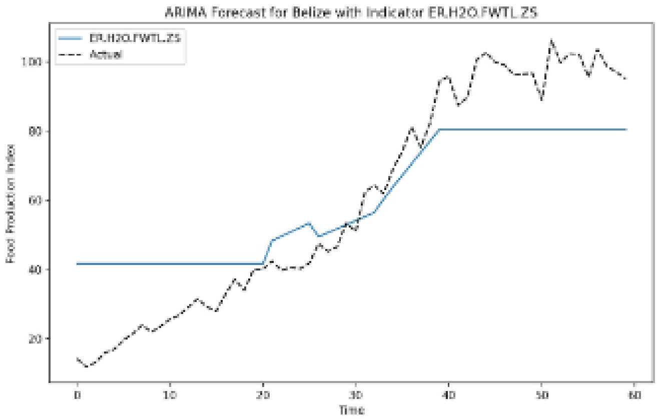
ARIMA Forecast for Belize with Indicator ER.H2O.FWTL.ZS
ER.H2O.FWTL.ZS
Actual
Food Production Index
Time
20
40
60
80
100
0
10
20
30
40
50
60

Fig. 9

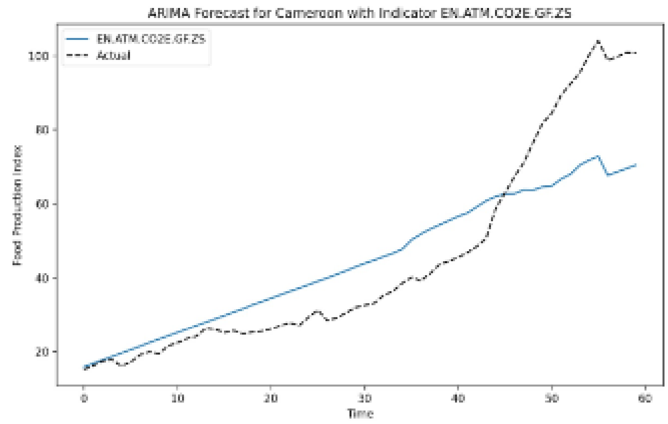
ARIMA Forecast for Cameroon with Indicator EN.ATM.CO2E.GF.ZS
EN.ATM.CO2E.GF.ZS
Actual
Food Production Index
Time

Fig. 10

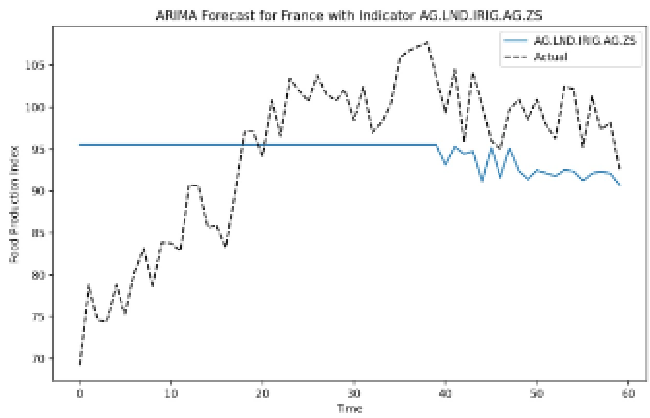
ARIMA Forecast for France with Indicator AG.LND.IRIG.AG.ZS
AG.LND.IRIG.AG.ZS
Actual
Food Production Index
Time

Fig. 11

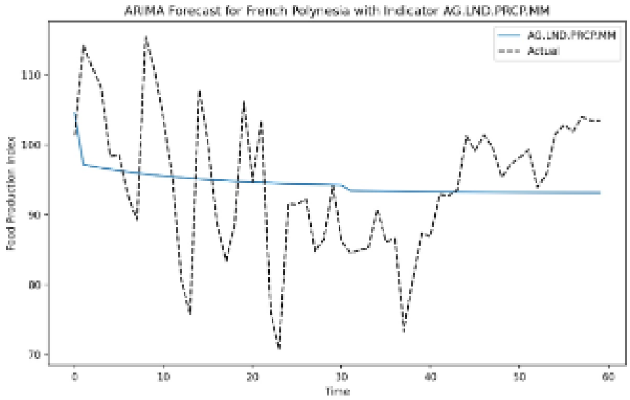

Fig. 12

References

[1] World Bank Open Data - Agriculture, total (% of GDP). https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS Accessed 2024-03-06

[2] Scheffran, J., Mar, A.: Climate change, food security intersectionality. Frontiers in Sustainable Food Systems **5**, 691191 (2021). Accessed 2024-03-06

[3] Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. https://www.ipcc.ch/srccl/chapter/chapter-5/ Accessed 2024-03-06

[4] Porter, J.R., Lobell, D.B., Roberts, M., Schlenker, W., Parry, M.L.C., Howden, M., Burrows, M.B., Willis, I., Dechaussaie, A., Jat, M.A.: Climate change and global food systems: Potential impacts on food security and undernutrition. Global Environmental Change **24**(1), 153–165 (2014). Accessed 2024-03-06

[5] Guo, Z., Li, J., Zhu, Q., Zhang, L., Li, X., He, J., Peng, S.: Food production forecasting with time series and ensemble modeling methods. Sustainability **9**(10), 1749 (2017). Accessed 2024-03-06

[6] Disaster Risk Reduction, U.N.O.: Food Production and Climate Change. https://www.preventionweb.net/files/1090_foodproduction.pdf Accessed 2024-03-06

[7] Food, Organization, A.: The Potential Effects of Climate Change on World Food Production and Security. https://www.fao.org/3/w5183e/w5183e0b.htm Accessed 2024-03-06