Comparative Forecasting Models for Potato Production: Evaluating T-ARMA, ARIMA-ARCH, Weibull, and Score-Driven Approaches in Major Global Producers

Abdullah Mohammad Ghazi Al khatib[1], Bayan Mohamad Alshaib [2], Pradeep Mishra[3*] Walid Emam[4] Yusra Tashkandy[4] and Adelajda Matuka[5]

[1]Department of Finance and Banking, Faculty of Economics, Damascus University, Damascus, Syrian Arab Republic,**abdullah1991.alkhatib@damascusuniversity.edu.sy; genius.275@hotmail.com**. ORCID ID: **https://orcid.org/0000-0002-1352-2348**.

[2]Department of Finance and Banking, Faculty of Economics, Damascus University, Damascus, Syrian Arab Republic, bayan1992.alshaib@damascusuniversity.edu.sy; bayanalshaibm@gmail.com. ORCID ID: https://orcid.org/0000-0002-9876-998X.

[3]College of Agriculture, Jawaharlal Nehru KrishiVishwaVidyalaya (JNKVV), Rewa, India, pradeepjnkvv@gmail.com

[4] Department of Statistics and Operations Research, Faculty of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.

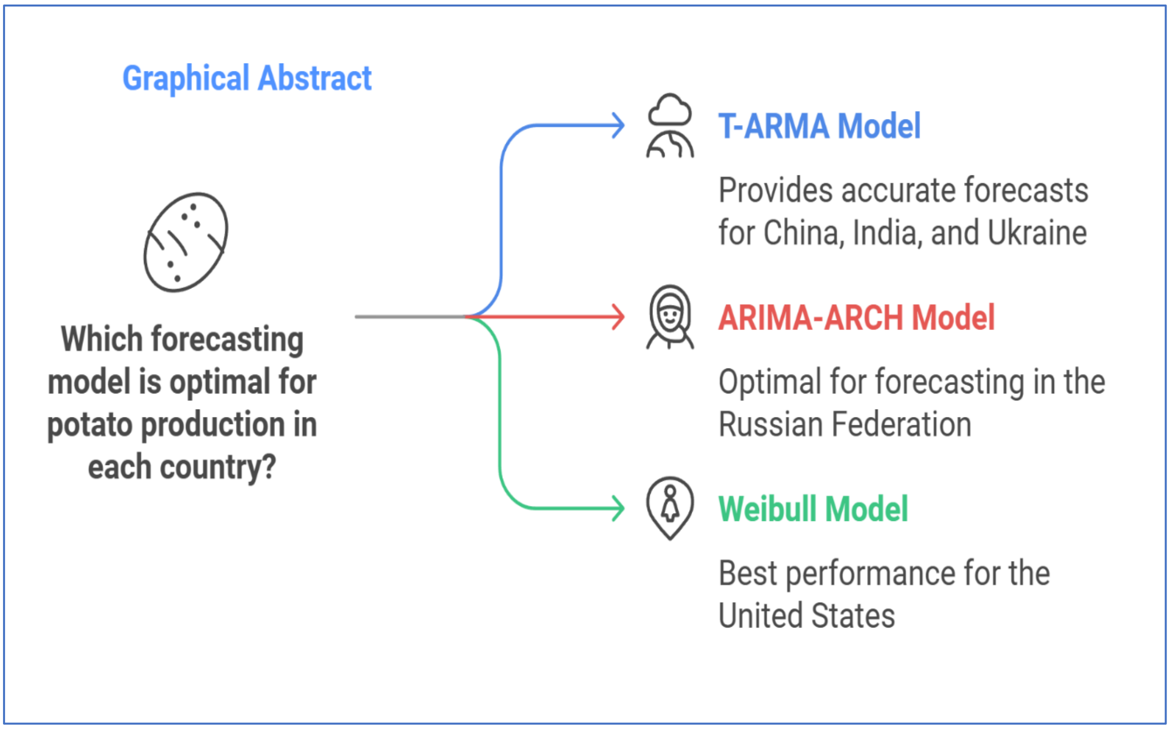
[5]Department of Economics, Faculty of Economics, University of Bologna, Italy, adelajda.matuka@yahoo.com

***Corresponding Author:Pradeep Mishra,College of Agriculture, Jawaharlal Nehru KrishiVishwaVidyalaya (JNKVV), Rewa, India,pradeepjnkvv@gmail.com**

Abstract

This study investigates future trends in potato production across five major global producers (China, India, Russia, Ukraine, and the U.S.) using annual data from 1961 to 2022. We comparatively evaluate T-ARMA, ARIMA-ARCH, Weibull, and Score-Driven models to forecast production for 2023–2030. Results identify the T-ARMA model as optimal for China, India, and Ukraine (validation MAE: 1,668–3,277), while ARIMA-ARCH and Weibull models excel in Russia (MAE: 1,405) and the U.S. (MAE: 547), respectively. Projections reveal a marginal decline in China (96,073 to 96,066 thousand tonnes; -0.007%) and India (55,701 to 55,698 thousand tonnes; -0.005%), contrasted by robust growth in Russia (+31.2%; 20,026 to 26,284 thousand tonnes). Ukraine remains stable (~21,103 thousand tonnes annually), while the U.S. shows a slight decline (18,054 to 18,053 thousand tonnes; -0.005%). These forecasts, derived from regionally tailored models, provide actionable insights for policymakers to enhance food security, optimize supply chains, and prioritize agricultural innovation.

Keywords: Forecasting; Potato Production; T-ARMA Model; ARIMA-ARCH Model; Weibull Model; Score-Driven Models


Graphical Abstract
Which forecasting model is optimal for potato production in each country?
T-ARMA Model
Provides accurate forecasts for China, India, and Ukraine
ARIMA-ARCH Model
Optimal for forecasting in the Russian Federation
Weibull Model
Best performance for the United States

1. Introduction

Potato (*Solanum tuberosum*) is a vital staple crop globally, playing a key role in ensuring food security and supporting economies worldwide. Given the inherent complexities of agricultural production—influenced by factors such as climate variability, technological advancements, and policy interventions—accurate forecasting of potato production is essential for informed decision-making regarding resource allocation, trade policies, and sustainable agricultural planning (FAO, 2023). Mishra et al. (2023a) highlight the importance of potatoes as a staple crop that provides essential nutrients and can play a significant role in food security due to its high caloric content and short growth period. They also note that potatoes are consumed widely in a variety of forms, as well as being a valuable resource for many industries, contributing significantly to economic stability by ensuring market consistency due to their semi perishable nature. These factors emphasize that the accurate forecasting of potato production is essential for both nutritional security and economic stability. Building upon the methodological approach of Abotaleb et al. (2021) who effectively used time series models to forecast rice production in South Asian countries, this study also employs and compares a similar set of approaches to forecast potato production in major global producers. While different from the agricultural context, the study by (Mishra et al. 2020; Mohammed et al. 2021) demonstrated the applicability of time series models for forecasting using complex datasets, which reinforces the potential use of these methods for agricultural forecasts.

This study addresses the need for robust potato production forecasts by applying and comparing multiple advanced time series models to historical data spanning 1961 to 2022 for five major producing countries: China, India, the Russian Federation, Ukraine, and the United States of America. The primary objectives are: (1) to generate accurate potato production forecasts for the period 2023–2030 using four distinct time series models, and (2) to comparatively assess the performance of these four models - T-ARMA, ARIMA-ARCH, Weibull, and Score-Driven – in this specific agricultural context. In line with Mishra et al. (2023b) who employed hybrid models to forecast pulses production, this paper also uses multiple different modeling approaches in order to provide robust production forecasts and emphasizes the need to compare different approaches for accurate forecasting. Moreover, similar to the study of Mishra et al. (2023c) on sugarcane production forecasting in South Asian countries, this study underscores the importance of considering geographical diversity when forecasting agricultural outputs.

This research contributes to the literature by demonstrating the effectiveness of time-varying parameter models and identifying the most suitable forecasting approach for diverse agricultural settings, therefore providing tools and insights that are beneficial to decision makers. While time series models are widely used in agricultural forecasting, a critical gap exists in the literature regarding the comparative performance of advanced models—particularly time-varying parameter models—for multi-country potato production forecasting over extended periods. Traditional approaches often rely on simpler models, which may not fully capture the dynamic complexities of agricultural systems or the diverse factors influencing production across different regions. Therefore, there is a lack of robust evidence regarding the effectiveness of more complex methods like T-ARMA and score-driven models, compared with traditional models such as ARIMA-ARCH and Weibull. This becomes crucial when developing accurate and reliable forecasts to support food security policies in geographically and economically diverse regions.

This study addresses this gap by presenting a comparative analysis of these four time series models – T-ARMA, ARIMA-ARCH, Weibull, and Score-Driven – applied to historical potato production data from five major global producing countries. The novelty of this work lies not only in the breadth of models considered, but also in its focused evaluation of a less utilized method (T-ARMA) in a crucial application setting. The T-ARMA model, with its time-varying parameters and ability to capture dynamic trends and volatility, is specifically evaluated alongside more established approaches to establish its true practical potential for this type of forecasting. The application of these models to five highly relevant and contrasting countries emphasizes the need for techniques that adapt to different agricultural systems.

The story of this research is one of navigating the complexities of global potato production through advanced statistical modeling. It explores how different models capture varying trends across different agricultural systems and their ability to handle the dynamic changes and uncertainties that are inherent in agricultural production. This study tells how an emerging model, the T-ARMA, performed compared to established alternatives, providing new evidence of the suitability of different models for predicting the future of this critical staple crop across varied regional settings. Ultimately, the findings underscore the need for researchers and policy makers to be cognizant of the strengths and limitations of different forecasting techniques.

The primary objectives are: (1) to generate potato production forecasts for 2023–2030 using T-ARMA, ARIMA-ARCH, Weibull, and Score-Driven models, and (2) to identify the most accurate model for each country through comparative validation.

By providing robust empirical evidence and targeted forecasts, this study not only enriches the methodological literature on agricultural forecasting but also offers valuable tools and insights for decision-makers at local and global level.

2. Literature Review

The application of time series models in agricultural forecasting has been extensively explored, with a particular emphasis on methods for handling non-stationarity and volatility. ARIMA models, as detailed by Box and Jenkins (2015), have been foundational, providing a flexible framework for capturing trends and non-seasonal dynamics in agricultural data, while ARCH accounts for heteroskedasticity.

Engle (1982) introduced the ARCH model to address volatility, allowing for the modeling of heteroskedasticity in time series data, and thereby accounting for the increased volatility found in agricultural data.

Recent advancements have seen the development of more sophisticated models, such as time-varying parameter models and non-linear models. Harvey (2013) presents time-varying parameter models, such as the T-ARMA, which effectively capture the dynamic nature of agricultural time series with heavy tails, allowing these models to adapt to changes in the underlying trends. Score-Driven Models, as discussed by Creal et al. (2013), offer an adaptive approach to forecasting by incorporating time-varying parameters based on score functions, which is beneficial for datasets where trends change over time. Additionally, the Weibull distribution, described by Weibull (1951), has been utilized for its suitability in modeling non-linear and skewed data distributions, which allows modelers to accurately portray the underlying non-normal distribution of certain datasets.

Comparative studies by Makridakis et al. (2018) have underscored the importance of selecting appropriate forecasting methods, with hybrid models often showing superior performance, although the use of simpler models has been widely accepted and applied (Box & Jenkins, 1976). Furthermore, as emphasized by Lobell et al. (2011) and Ray et al. (2015), climate change and technological changes significantly impact agricultural yields. This necessitates the integration of these factors and robust forecasting methods into agricultural modelling, particularly when considering future production. Yadav et al. (2024) highlight the necessity for a rigorous, quantifiable assessment of linearity, stationarity, and outliers during the specification of forecasting models. This is essential to ensure that stakeholders and policymakers have access to precise and reliable predictions that they can use to plan future operations effectively. Their work suggests that, beyond just choosing models, a close evaluation of the specific underlying data properties is crucial for robust prediction. Mishra et al. (2023) emphasized that, given the complexities of agricultural production in different regions, no single forecasting model is universally superior, which suggests the need for models that are tailored to specific regions to effectively predict production in those areas.

Despite the breadth of research, a gap remains in the literature regarding the comparative application of these advanced models for multi-country potato production forecasts over extended time horizons, especially when considering time-varying parameter models such as the T-ARMA. This study bridges that gap by applying and comparing a suite of sophisticated models in five major producing countries and aims to provide evidence on the suitability of each modeling approach in various agricultural contexts.

3. Data and Methodology

3.1. Data Description

The dataset comprises annual potato production data, measured in thousands of tonnes (000 tonnes), sourced from the Food and Agriculture Organization (FAO) of the United Nations, specifically from the FAOSTAT database (FAO, 2023). Analyses were conducted using Timeseries Lab software. For five major potato-producing countries: China, India, the Russian Federation, Ukraine, and the United States of America. The dataset span over the period (1961 to 2022). The data were sourced from the Food and Agriculture Organization (FAO) of the United Nations, specifically from the FAOSTAT database (FAOSTAT, 2023), a globally recognized and authoritative source of agricultural statistics. The dataset includes total annual production figures for each country, without distinctions based on specific potato varieties or sub-regions. Prior to analysis, the raw data was inspected for any missing values or inconsistencies; no data imputation or correction techniques were necessary as the dataset was found to be complete and consistent. Summary statistics indicate significant variability in production volumes and trends across the selected countries. China and India typically exhibit the highest annual production, while the Russian Federation and Ukraine show more volatile production patterns, and the United States shows moderate and relatively stable production figures.

3.2. Methodology

This study employs four distinct time series models to forecast potato production:

1. **T-ARMA (Time-Varying ARMA):** The T-ARMA model captures dynamic changes in both the mean and the variance of a time series. It assumes the data follows a Student's t-distribution, which is suitable for heavy-tailed and volatile data often seen in agricultural production. This is achieved by incorporating time-varying parameters into the ARMA (Autoregressive Moving Average) model, allowing it to adapt to changes in the data over time (Harvey, 2013).

2. **ARIMA-ARCH:** This hybrid model combines ARIMA (Autoregressive Integrated Moving Average) with ARCH (Autoregressive Conditional Heteroskedasticity). The

ARIMA component captures trends and seasonality in the time series, while the ARCH component models the volatility and heteroskedasticity of the residuals. This makes it a robust approach for non-stationary data where volatility changes over time (Box & Jenkins, 2015; Engle, 1982).

3. **Weibull Model:** The Weibull distribution is a continuous probability distribution that is used to model time to event data or non-linear datasets. The Weibull Model captures the skewness and non-linearity in the data distribution by modeling the mean of the series based on a Weibull distribution. This model has been proven effective in time series modeling of skewed data, such as is often found in agriculture (Weibull, 1951).

4. **Score-Driven Models:** Score-driven models employ time-varying parameters that adapt based on the score function of the underlying distribution. This means that the model's parameters change over time based on how well the model is fitting to the data; therefore, these models are adaptive to changes in the data dynamics. This adaptive nature enables the model to capture more complex time series patterns, which is useful in non-stationary datasets (Creal et al., 2013).

The dataset was divided into a training set (1961-2016) and a validation set (2017-2022). The models were initially trained using the training set. Each model was optimized to best fit each country's dataset. Then, the optimized models were evaluated using the validation set. Model performance was assessed using metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Log Loss. (Log Loss evaluates probabilistic forecasts by penalizing incorrect predictions based on their confidence. Lower values indicate better-calibrated models) (Hyndman & Koehler, 2006). The model with the lowest errors on the validation set was selected as the optimal model for each country. While the training set (1961–2016) is larger than the validation set (2017–2022), prioritizing validation performance minimizes overfitting risks and ensures generalizability. The model with the lowest validation errors (MAE, RMSE, MAPE, Log Loss) was selected as optimal, as it demonstrates robustness to unseen data. The chosen model was then re-estimated using all available data (1961-2022), and forecasts for the period 2023–2030 were generated using this final model.

4. Results

4.1. Descriptive Statistics

The summary statistics in Table 1 provide a comprehensive overview of potato production trends across five major global producers, highlighting key differences in production scales, variability, and distribution patterns. China emerges as the dominant producer, with a mean annual output of 47,452 thousand tonnes, far exceeding other nations. However, its production exhibits substantial year-to-year fluctuations, as reflected in the high standard deviation of 27,018. Despite this absolute variability, the coefficient of variation (C.V. = 0.57) indicates moderate relative variability, suggesting that production fluctuations are proportional to its large mean output. India follows with a mean production of 20,701 thousand tonnes but displays higher relative variability (C.V. = 0.78), underscoring less consistent production levels compared to China.

The Russian Federation and Ukraine present contrasting profiles. Russia's mean production (28,718 thousand tonnes) is accompanied by moderate variability (Std. Dev. = 5,852) and near-symmetric distribution (skewness = -0.02), reflecting balanced growth and recovery phases, particularly post-2000. Ukraine, with a mean of 19,627 thousand tonnes, shows lower absolute variability (Std. Dev. = 2,694) but a negatively skewed distribution (skewness = -

0.58), indicative of periodic sharp declines, likely tied to socio-economic transitions. The United States, with the most stable production profile, reports a mean of 17,629 thousand tonnes and the lowest variability (Std. Dev. = 3,017), supported by a nearly symmetric distribution (skewness = -0.29) and minimal excess kurtosis (-1.00), highlighting consistent agricultural practices.

Distributional characteristics further differentiate these nations. China and India exhibit positive skewness (0.37 and 0.73, respectively), signaling occasional exceptionally high production years, while Ukraine's negative skewness points to years of significant shortfalls. Excess kurtosis values near zero for most countries suggest distributions close to normal, though China's negative excess kurtosis (-1.39) implies lighter tails and fewer outliers than a typical normal distribution. The interquartile ranges (IQ range), spanning from 45,442 thousand tonnes in China to 5,040 in the U.S., emphasize the vast differences in production scales and central tendencies. For instance, China's 95th percentile (92,473 thousand tonnes) underscores its capacity for peak output, while the U.S. 5th percentile (12,442 thousand tonnes) aligns closely with its stable baseline. These statistical insights contextualize the forecasting challenges and validate the region-specific model selections, as high variability in Asia contrasts sharply with stability in the U.S., necessitating tailored analytical approaches.

4.2. Model Performance

The evaluation metrics for the training and validation datasets are presented in Tables 2 and 3. The T-ARMA model consistently outperformed the other models across most countries, achieving the lowest MAE, RMSE, and MAPE values. For instance, in China, the T-ARMA model achieved an MAE of 2686 and an RMSE of 3597 during the training period, compared to the ARIMA-ARCH model's MAE of 3699 and RMSE of 5005. Similarly, in the validation period, the T-ARMA model demonstrated superior accuracy, with an MAE of 1668 and an RMSE of 2033 for China. The T-ARMA model also showed superior results for India and Ukraine, although for United States and Russian Federation other best performing models were obtained in the validation dataset. The radar charts, displayed in Figure 6 and 7, provide a visual comparison of the performance of the different models across all countries. Specifically, Figure 6 depicts the goodness of fit for each of the models using the training data, showing that T-ARMA model provides the best fit for China and India, while Figure 7 depicts the same information for the validation data and shows that, the best fit models for Russia and the US are ARIMA-ARCH and Weibull, respectively.

4.3. Forecasted Production

The projected potato production for 2023–2030 is presented in **Table 5**. China, expected to remain the largest producer, exhibits a marginal decrease in production from 96,073 thousand tonnes in 2023 to 96,066 thousand tonnes in 2030, a change of approximately -0.007%. India's production is also projected to experience a slight decrease, from 55,701 thousand tonnes in 2023 to 55,698 thousand tonnes in 2030, representing a change of roughly -0.005%. These trends are illustrated in Figures 1 and 2, respectively. The Russian Federation and Ukraine demonstrate moderate growth trends over the forecast period. Specifically, production in the Russian Federation is projected to increase from 20,026 thousand tonnes in 2023 to 26,284 thousand tonnes in 2030, an increase of approximately 31.2%, while Ukraine is projected to produce around 21,103 thousand tonnes per year. These trends can be seen in Figures 3 and 4, respectively. The United States is projected to experience a slight decline, with production decreasing from 18,054 thousand tonnes in 2023 to 18,053 thousand tonnes in 2030, representing a decline of roughly 0.005%.The production trend for the United States can be observed in Figure 5.

5. Discussion

The T-ARMA model's superior performance in forecasting potato production for China, India, and Ukraine, as detailed in Table 4, underscores its effectiveness in capturing the dynamic trends and volatility inherent in agricultural time series data (Harvey, 2013). This aligns with the findings of Makridakis et al. (2018), who emphasized the benefits of time-varying parameter models for handling complex data, particularly when capturing changes over time. In contrast, models based on a Weibull distribution were more accurate in forecasting production in United States and ARIMA-ARCH in Russian Federation. Building on the findings of Al Khatib et al. (2023), nonlinear models offer an advantage over their linear counterparts by revealing dynamic patterns and asymmetries that might remain undetected in a traditional linear analysis. This is highlighting that the most effective methodology is not uniform across different countries. This is also visually demonstrated in Figures 1, 2, 3, 4, and 5, where the fit of the optimal model for each country is shown against the actual production, indicating how different models perform in different contexts. Specifically, the residual diagnostics in Figure 1 indicate no significant issues in model fit for China.

Considering the forecasted production trends presented in Table 5, China and India, despite minor decreases of 0.007% and 0.005% respectively over the forecast period, maintain their dominant positions in global potato production, reflecting the relative stability in these production systems (FAO, 2023). The minor decreases could be explained by the T-ARMA model capturing the slow decline in production in recent years. The Russian Federation, on the other hand, shows the most significant growth, with a notable increase in potato production of approximately 31.2% from 2023 to 2030. This growth pattern is visually clear in Figure 3, highlighting the steeper increase compared to the other countries. Ukraine maintains production levels at around 21,103 thousand tonnes each year. Meanwhile, the United States exhibits a slight decline, with a production decrease of approximately 0.005% over the forecast period, which could be a signal of reorientation in agricultural practices (USDA, 2023). The minor changes in China, India and United States can be visually observed in Figures 1, 2, and 5, respectively, which display consistent but very small variations over the forecast years.

These results have significant implications for policymakers. For countries like China and India, where production shows stability, the focus could be on enhancing supply chain efficiency, storage, and distribution infrastructure to minimize post-harvest losses (FAO, 2023). The notable growth in the Russian Federation, as shown in Figure 3, might necessitate policy interventions to address potential supply gluts and support the development of secondary industries. Conversely, the slight decline in the United States may warrant policies that encourage innovation and investment in potato farming, for example, using new farming techniques that can increase production. The differences in production trends clearly show the importance of tailored approaches for different regions, considering the specific economic, political and agricultural factors and further emphasizing the need for region-specific model selection as highlighted by Mishra et al. (2023) and Yadav et al (2024).

The study of Yadav et al. (2024) showed how different modeling approaches, such as ARIMA, state space, and XGBoost, have their specific advantages and limitations. They highlight how ARIMA models are robust with linear trends, state space models are useful for non-stationarity and outliers, and XGBoost is useful to capture non-linear relationships. By evaluating these different approaches, they were able to capture different patterns in the data and find optimal models for each region, further suggesting the importance of careful model selection. The present research also finds similar support, highlighting how the T-ARMA, Weibull and ARIMA-ARCH models perform differently based on the dataset they are used for. In terms of modeling implications, the T-ARMA model's success across multiple countries suggests its broader applicability in agricultural forecasting. However, the differences in optimal models across countries emphasize the importance of careful model selection and customization based on the specific characteristics of each dataset. Moreover, Al khatib (2023) highlighted that the application of advanced modelling techniques can offer valuable insights to aid policy making in economic and financial spheres. This is particularly relevant when considering that data-driven insights can be used to inform policy making for more relevant and robust approaches for sustainable development.

The residual plots included in Figures 1-5 also support the model choices as they display no major violations of model assumptions. The work of Mishra et al. (2021) highlights that rigorous model evaluation, involving multiple approaches and both training and testing datasets, is paramount for identifying the most suitable models for specific forecasting challenges. Moreover, Al khatib et al. (2021, p. 271) demonstrated that simply using new forecasting methods does not always provide superior results, and highlighted several factors that may influence model accuracy, such as data frequency and complexity, number of observations, seasonality, cyclic and trend variations, stationarity, and the randomness of the data, in addition to the length of the out-of-sample forecasts.

6. Conclusion

This study has provided a comprehensive analysis of potato production forecasting in five major producing countries—China, India, the Russian Federation, Ukraine, and the United States of America—employing a suite of advanced time series models. While the T-ARMA model proved to be the most reliable for predicting potato production trends in China, India, and Ukraine, models based on ARIMA-ARCH and Weibull distribution showed to be more accurate for Russian Federation and the United States, respectively. The forecasted trends, highlighting differing challenges across the selected countries, underscores the need for targeted strategies to support sustainable potato production and ensure food security in the face of global challenges.

The findings of this research emphasize the value of incorporating sophisticated modeling techniques, specifically the T-ARMA model for several countries, into agricultural planning and policy-making. Future research should focus on integrating climate data, market prices, and policy indicators to further improve model accuracy and robustness, particularly for the United States and the Russian Federation. Moreover, future studies should explore the specific policy implementations needed to improve supply chain efficiency in countries like China and India, while considering potential supply gluts and the need to develop secondary industries in countries like the Russian Federation. In addition, incorporating alternative models and more data for the United States, could improve predictions for these countries. By doing so, the agricultural sector can better anticipate and respond to emerging challenges, ensure global food security, and achieve long term economic sustainability. This will ultimately inform more effective policy implementation, leading to better management of potato production and resource allocation, while also contributing towards the development of advanced statistical methodologies for better forecasting.

Highlights:

- **T-ARMA model best forecasts potato production in China, India, Ukraine.**
- **Region-specific time series models ensure accurate potato forecasts.**
- **China, India: slight declines; Russia: 31% growth by 2030.**
- **2023–2030 forecasts guide food security policies via optimal models.**
- **T-ARMA excels in multi-country agricultural forecasting.**

Conflict of Interest:

The authors declare that the research was not conducted under such financial or commercial conditions that could lead to potential conflicts of interest. The authors further declare that they have no financial interests.

Funding:

The study was funded by Researchers Supporting Project number (RSP2025R488), King Saud University, Riyadh, Saudi Arabia.

Acknowledgements: The study was funded by Researchers Supporting Project number (RSP2025R488), King Saud University, Riyadh, Saudi Arabia.

Data availability statement:

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References

Abotaleb M, Ray S, Mishra P, Karakaya K, Shoko C, Al Khatib A, ... and Balloo R (2021) Modelling and forecasting of rice production in South Asian countries. Ama, Agricultural Mechanization in Asia, Africa & Latin America 51: 1611-1627

Al Khatib AMG (2023) The complexity of financial development and economic growth nexus in Syria: a nonlinear modelling approach with artificial neural networks and NARDL model. Heliyon 9(10):1–18. https://doi.org/10.1016/j.heliyon.2023.e20265

Al Khatib AMG, Yonar H, Abotaleb M, Mishra P, Yonar A, Karakaya K, Badr A, Dhaka V (2021) Modeling and forecasting of egg production in India using time series models. Eurasian J Vet Sci 37(4):265–273. https://doi.org/10.15312/EurasianJVetSci.2021.352

Al khatib, A. M. G., Alshaib, B. M., & Kanaan, A. M. (2023) The Interaction Between Financial Development and Economic Growth: A Novel Application of Transfer Entropy and Nonlinear Approach in Algeria. Sage Open, 13(4). https://doi.org/10.1177/21582440231217871

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015) Time series analysis: forecasting and control. John Wiley & Sons

Creal, D., Koopman, S. J., & Lucas, A. (2013) Generalized Autoregressive Score Models with Applications. Journal of Applied Econometrics, 28(5), 777–795

Engle, R. F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica, 50(4), 987–1008

FAO (Food and Agriculture Organization of the United Nations). (2023) FAOSTAT. Retrieved from https://www.fao.org/faostat/

Harvey, A. C. (2013) Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series. Oxford University Press.

Hyndman, R. J., & Koehler, A. B. (2006) Another Look at Measures of Forecast Accuracy. International Journal of Forecasting, 22(4), 679–688

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011) Climate Trends and Global Crop Production Since 1980. Science, 333(6042), 616–620

Makridakis, S., Spiliotis, E., &Assimakopoulos, V. (2018) Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward. International Journal of Forecasting, 34(1), 1–12

Mishra P, Al Khatib AMG, Lal P, Anwar A, Nganvongpanit K, Abotaleb M, ... and Punyapornwithaya V (2023b) An overview of pulses production in India: retrospect and prospects of the future food with an applicationof hybrid models. National Acad Sci Lett 46 (5) 1-8.https://doi.org/10.1007/s40009-023-01267-2

Mishra P, Al Khatib AMG, Sardar I, Mohammed J, Karakaya K, Dash A, Dubey A (2021) Modeling and forecasting of sugarcane production in India. Sugar Tech 23(6):1317–1324. https://doi.org/10.1007/s12355-021-01004-3

Mishra P, Al Khatib AMG, Sardar I, Mohammed J, Ray M, Manish K, ... Rono K (2020) Modelling and forecasting of COVID-19 in India. Journal of Infectious Diseases and Epidemiology 6(5): 1-11

Mishra P, Alakkari KM, Lama A, Ray S, Singh M, Shoko C, ... Karakaya K (2023c) Modeling and forecasting of sugarcane production in South Asian countries. Current Applied Science and Technology 23 (1). 1-15.https://doi.org/10.55003/cast.2022.01.23.002

Mishra, P., Al khatib, A. M. G., Alshaib, B. M., Kuamri, B., Tiwari, S., Singh, A. P., ... &Kuamri, P. (2023a) Forecasting Potato Production in Major South Asian Countries: a Comparative Study of Machine Learning and Time Series Models. Potato Research, 66(4), 479-494. https://doi.org/10.1007/s11540-023-09683-z

Mohammed J, Al Khatib AMG, Mishra P, Adjei P, Singh PK, Priya SK, Das SS (2021) Modeling and forecasting of Covid-19 from the context of Ghana. African Rev Eco Finance 13(2):124–141

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2015) Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLOS ONE, 10(6), e0114702.

USDA (United States Department of Agriculture). (2023) Retrieved from https://www.usda.gov/

Weibull, W. (1951) A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, 18(3), 293–297

Yadav, S., Al khatib, A. M. G., Alshaib, B. M., Ranjan, S., Kumari, B., Alkader, N. A., ... & Kapoor, P. (2024) Decoding Potato Power: A Global Forecast of Production with Machine Learning and State-of-the-Art Techniques. Potato Research, 1-21. https://doi.org/10.1007/s11540-024-09705-4

Table 1: Summary Statistics, using the observations 1961 - 2022

Maximum	Minimum	Median	Mean	Variable
95631.	12010.	34929.	47452.	**China**
56176.	2447.0	16299.	20701.	**India**
39909.	18296.	28953.	28718.	**Russian Federation**
24248.	12723.	19838.	19627.	**Ukraine**
23294.	10941.	18397.	17629.	**United States of America**
Ex. kurtosis	**Skewness**	**C.V.**	**Std. Dev.**	**Variable**
-1.3884	0.37160	0.56938	27018.	**China**
-0.74011	0.73098	0.78086	16165.	**India**
-0.67960	-0.015291	0.20378	5852.1	**Russian Federation**
0.013995	-0.58257	0.13727	2694.2	**Ukraine**
-0.99780	-0.29271	0.17115	3017.3	**United States of America**
IQ range		**95% Perc.**	**5% Perc.**	**Variable**
45442.		92473.	13840.	**China**
23544.		51142.	2815.9	**India**
10218.		38961.	18651.	**Russian Federation**
2946.3		23915.	13927.	**Ukraine**
5040.4		21669.	12442.	**United States of America**

Table 2: Model Performance on Training Dataset (1961-2016)

Country	Model	MAE	RMSE	MAPE	Log Loss
China	ARIMA-ARCH	3699	5005	10.32	10.01
	Weibull Model	3058	3979	7.75	9.58
	Score Driven	2835	3703	7.44	9.64
	T-ARMA Model	**2686**	**3597**	**6.96**	**9.6**
India	ARIMA-ARCH	2264	3451	25.85	9.76
	Weibull Model	2868	4019	18.19	9.27
	Score Driven	1645	2276	11.71	9.15
	T-ARMA Model	**1636**	**2259**	**11.58**	**9.12**
Russia	ARIMA-ARCH	3134	4552	11.09	9.9
	Weibull Model	2774	**3741**	9.85	**9.59**
	Score Driven	2751	3842	9.83	9.67
	T-ARMA Model	**2658**	3897	**9.71**	9.6
Ukraine	ARIMA-ARCH	1906	2495	10.09	9.55
	Weibull Model	1888	2360	10.20	9.23
	Score Driven	-	-	-	**9.15**
	T-ARMA Model	**1692**	**2271**	**9.26**	**9.15**
USA	ARIMA-ARCH	1702	2071	9.79	9.61
	Weibull Model	990	1226	5.72	8.55
	Score Driven	**881**	1119	**5.22**	**8.44**
	T-ARMA Model	883	**1117**	5.24	**8.44**

Table 3: Model Performance on Validation Dataset (2017-2022)

Country	**Model**	**MAE**	**RMSE**	**MAPE**	**Log Loss**
China	ARIMA-ARCH	5476	5667	5.96	10.19
	Weibull Model	4967	5159	5.41	9.93
	Score Driven	2048	2295	2.24	9.33
	T-ARMA Model	**1668**	**2033**	**1.82**	**9.23**
India	ARIMA-ARCH	5517	6223	10.68	10.14
	Weibull Model	9489	10246	18.45	10.87
	Score Driven	3366	3869	6.91	**9.81**
	T-ARMA Model	**3277**	**3674**	**6.28**	10.03
Russia	ARIMA-ARCH	**1405**	**1767**	**7.1**	9.68
	Weibull Model	5180	5405	25.17	10.43
	Score Driven	3256	3590	15.86	**9.61**
	T-ARMA Model	5347	6110	25.41	10.67
Ukraine	ARIMA-ARCH	1161	1211	5.4	9.61
	Weibull Model	772	887	3.57	8.91
	Score Driven	-	-	-	8.73
	T-ARMA Model	**643**	**878**	**3.08**	**8.72**
USA	ARIMA-ARCH	818	997	4.2	9.61
	Weibull Model	**547**	**599**	**2.82**	8.39
	Score Driven	658	778	3.51	**8.18**
	T-ARMA Model	661	781	3.53	**8.18**

Table 4: Optimal Models

Country	Training Dataset	Validation Dataset	Optimal Model
China	T-ARMA Model	T-ARMA Model	T-ARMA Model
India	T-ARMA Model	T-ARMA Model	T-ARMA Model
Russian Federation	Weibull Model	ARIMA-ARCH	ARIMA-ARCH
Ukraine	T-ARMA Model	T-ARMA Model	T-ARMA Model
United States of America	Score Driven models	Weibull model	Weibull Model

Table 5: Projected Potato Production (2023-2030) in 1000 tonnes

Year	China	India	Russian Federation	Ukraine	United States of America
2023	96073.25	55701.46	20026.12	21103.91	18054.23
2024	96072.24	55701.04	21118.16	21103.88	18053.99
2025	96071.23	55700.62	22136.43	21103.86	18053.75
2026	96070.22	55700.20	23085.91	21103.83	18053.51
2027	96069.21	55699.77	23971.25	21103.80	18053.27
2028	96068.20	55699.35	24796.78	21103.78	18053.03
2029	96067.19	55698.93	25566.54	21103.75	18052.79
2030	96066.18	55698.51	26284.31	21103.73	18052.55

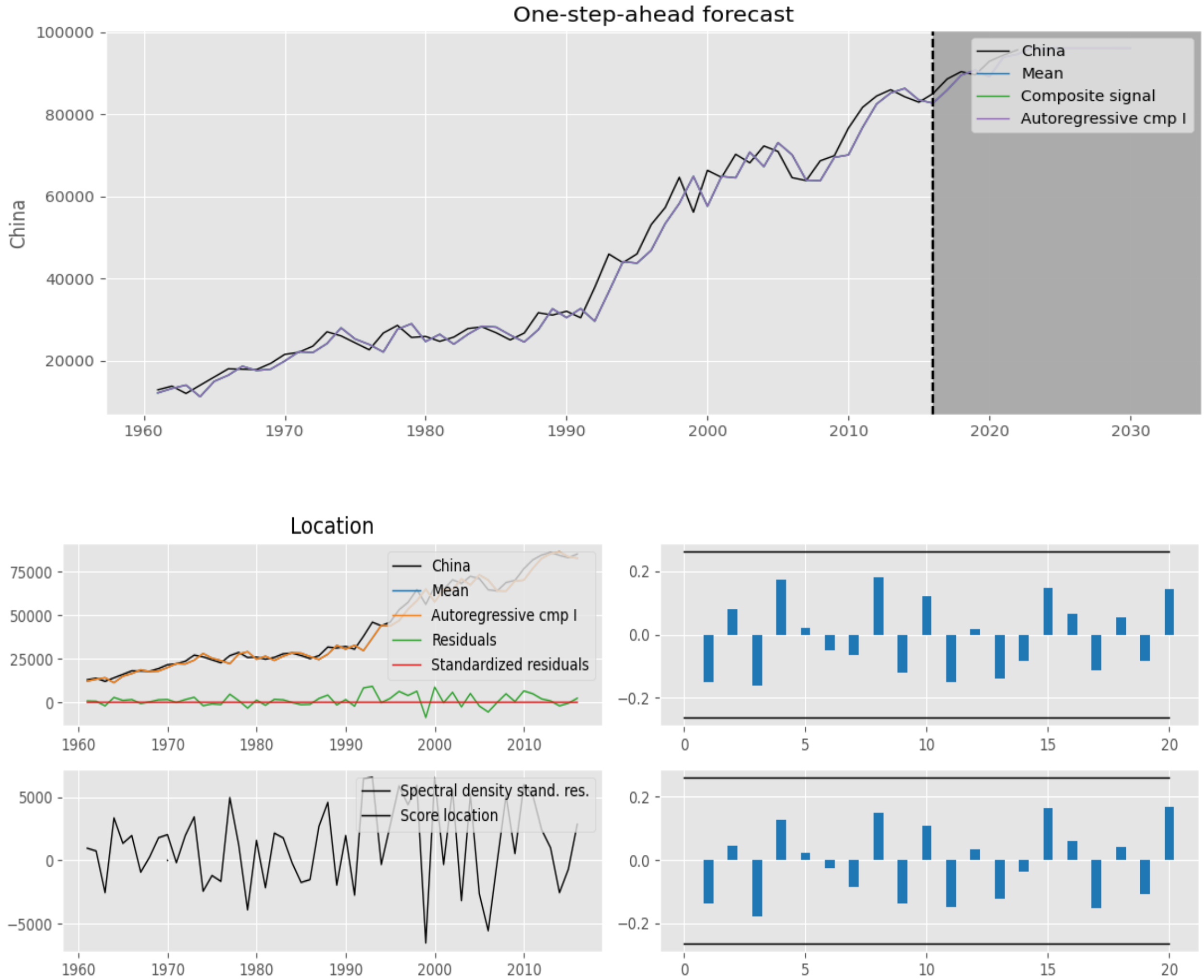


Fig 1: An illustration of the actual and forecasted potato production in China between 1961 and 2030, accompanied by residual diagnostic charts, employing T-ARMA Model

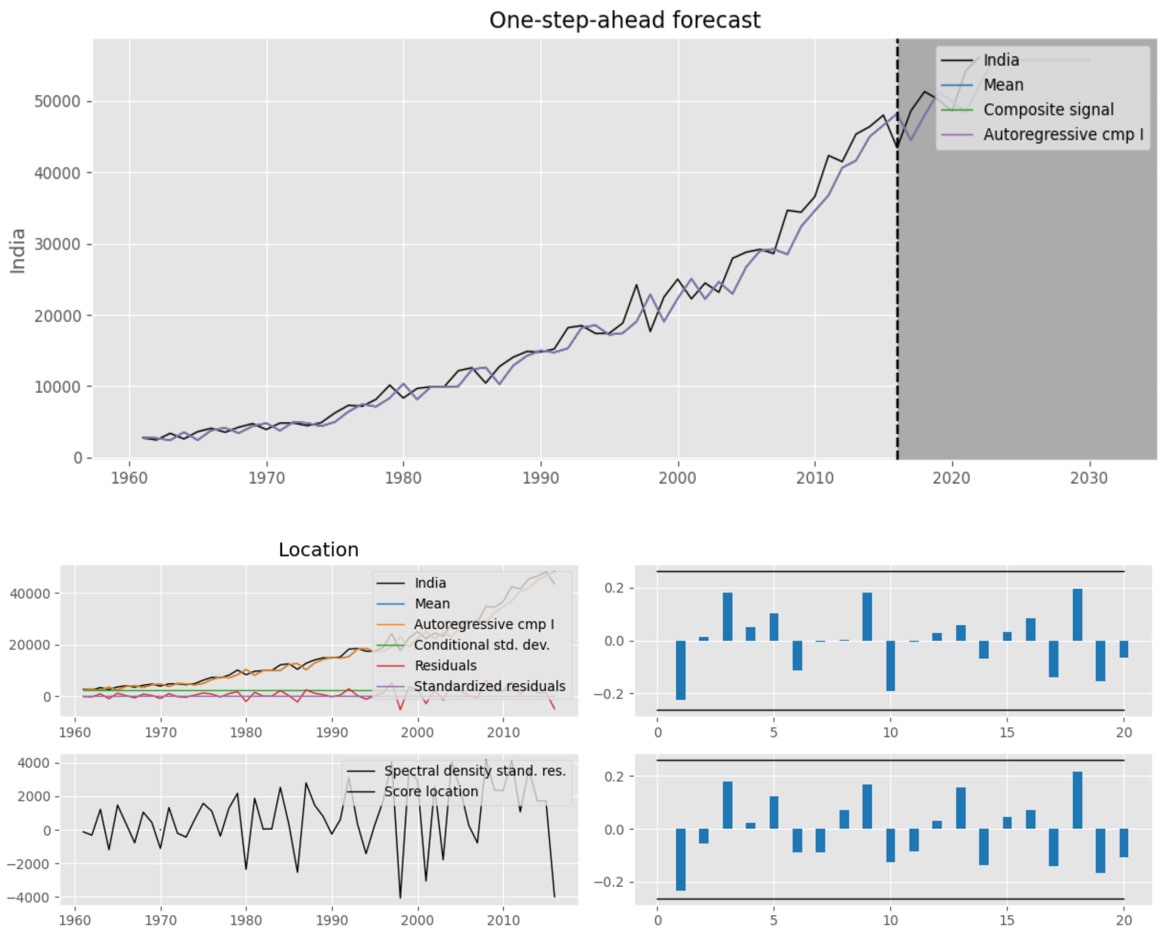


Fig 2: An illustration of the actual and forecasted potato production in India between 1961 and 2030, accompanied by residual diagnostic charts, employing T-ARMA Model

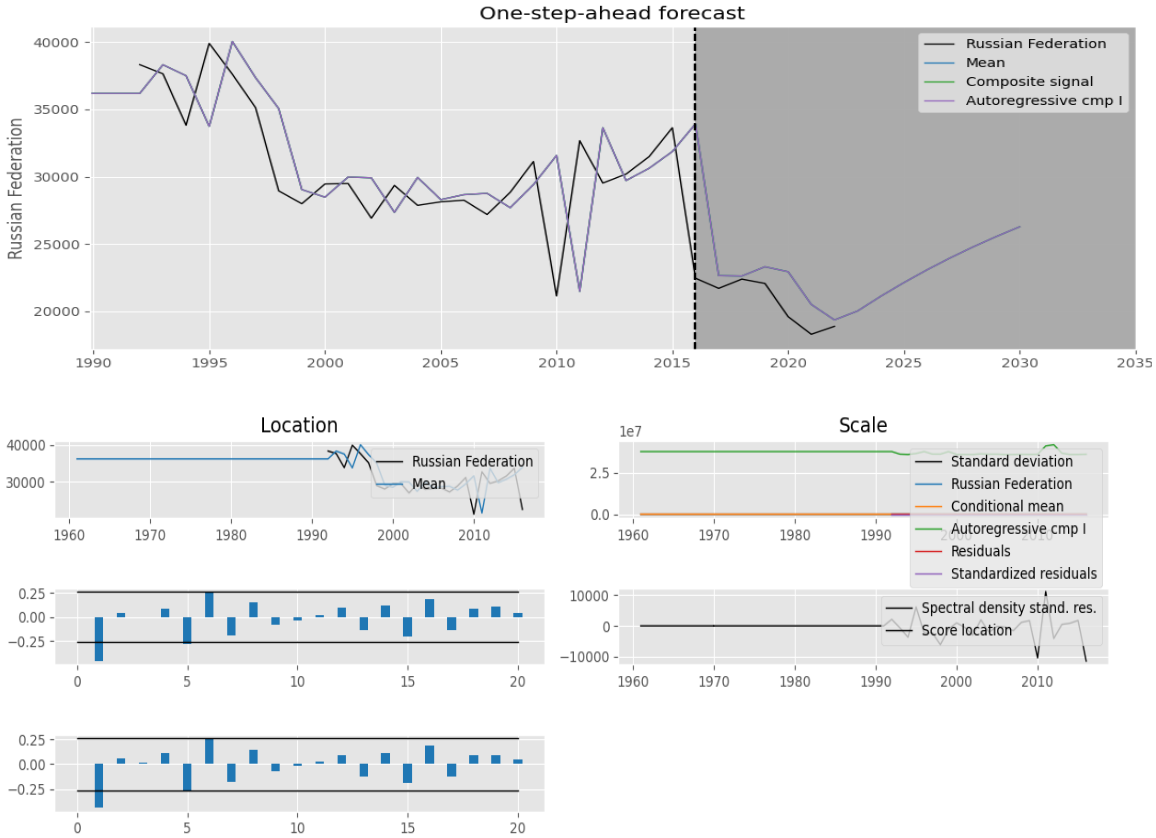


Fig3: An illustration of the actual and forecasted potato production in Russian Federationbetween 1961 and 2030, accompanied by residual diagnostic charts, employing ARIMA-ARCHModel

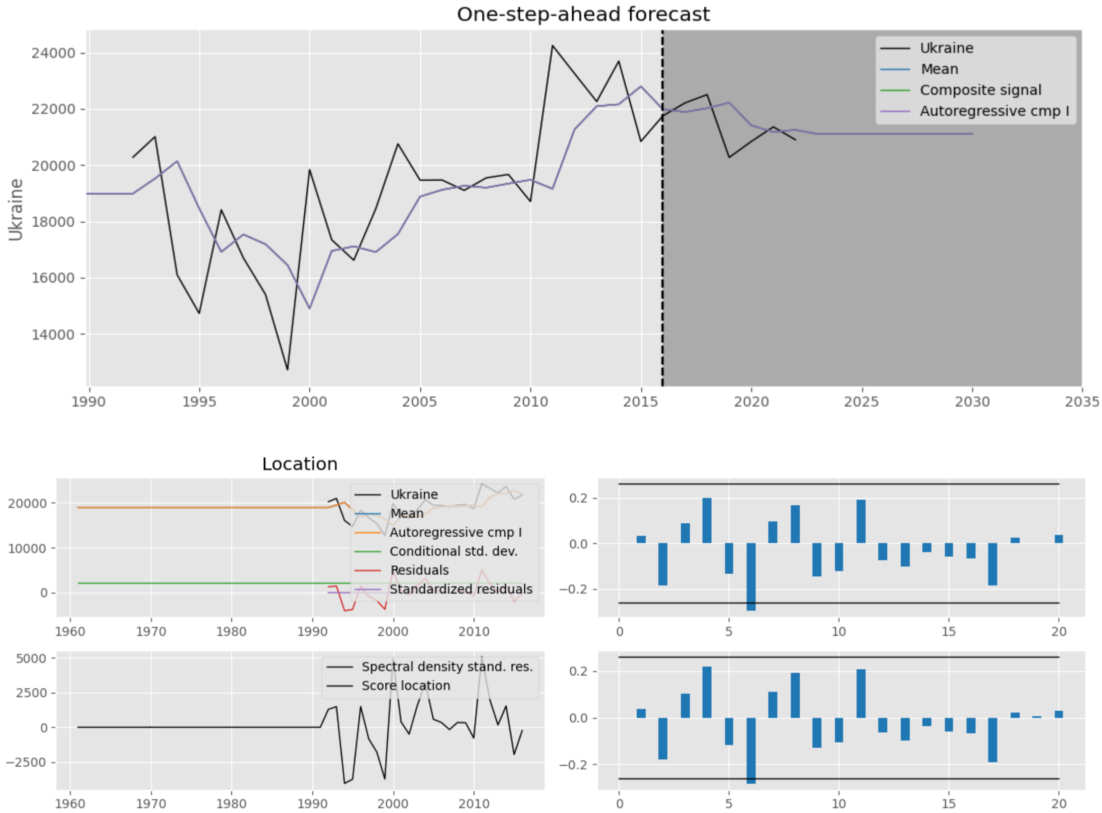


Fig 4: An illustration of the actual and forecasted *potato* production in Ukraine between 1961 and 2030, accompanied by residual diagnostic charts, employing T-ARMA Model

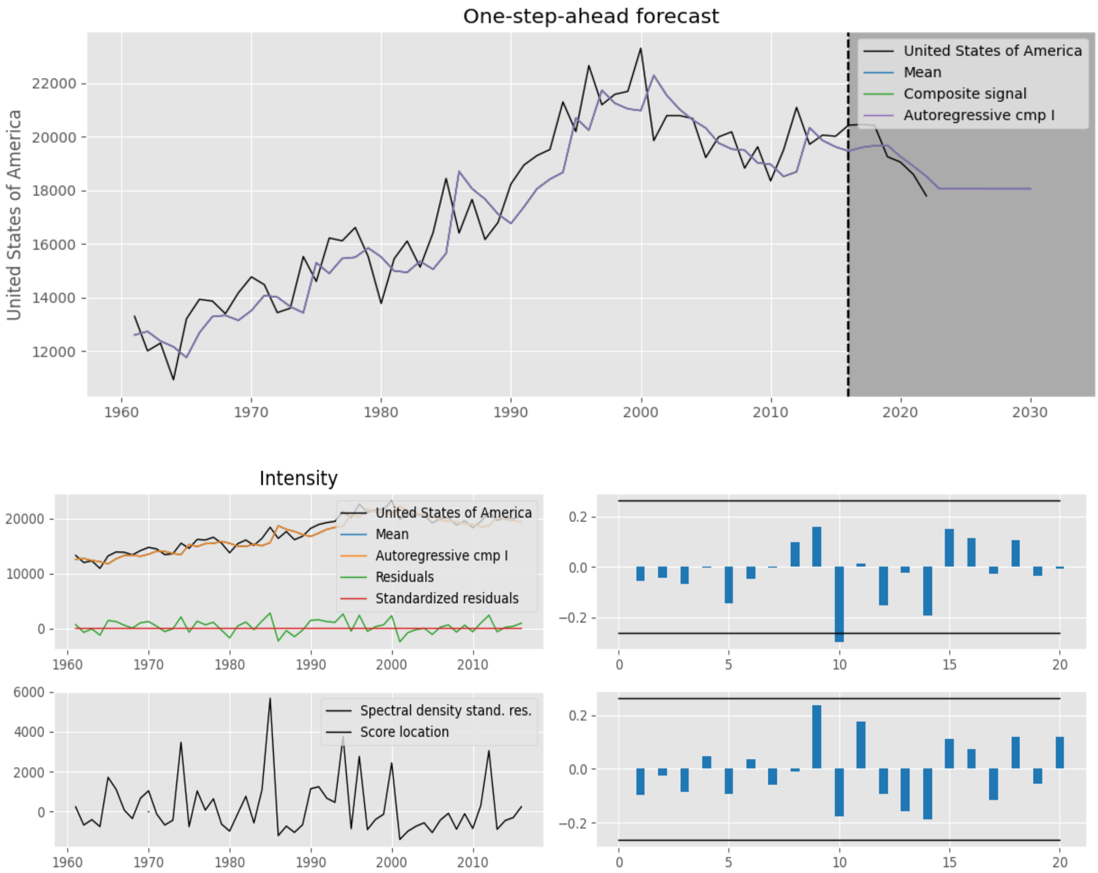


Fig5: An illustration of the actual and forecasted *potato* production in United States of America between 1961 and 2030, accompanied by residual diagnostic charts, employing WeibullModel

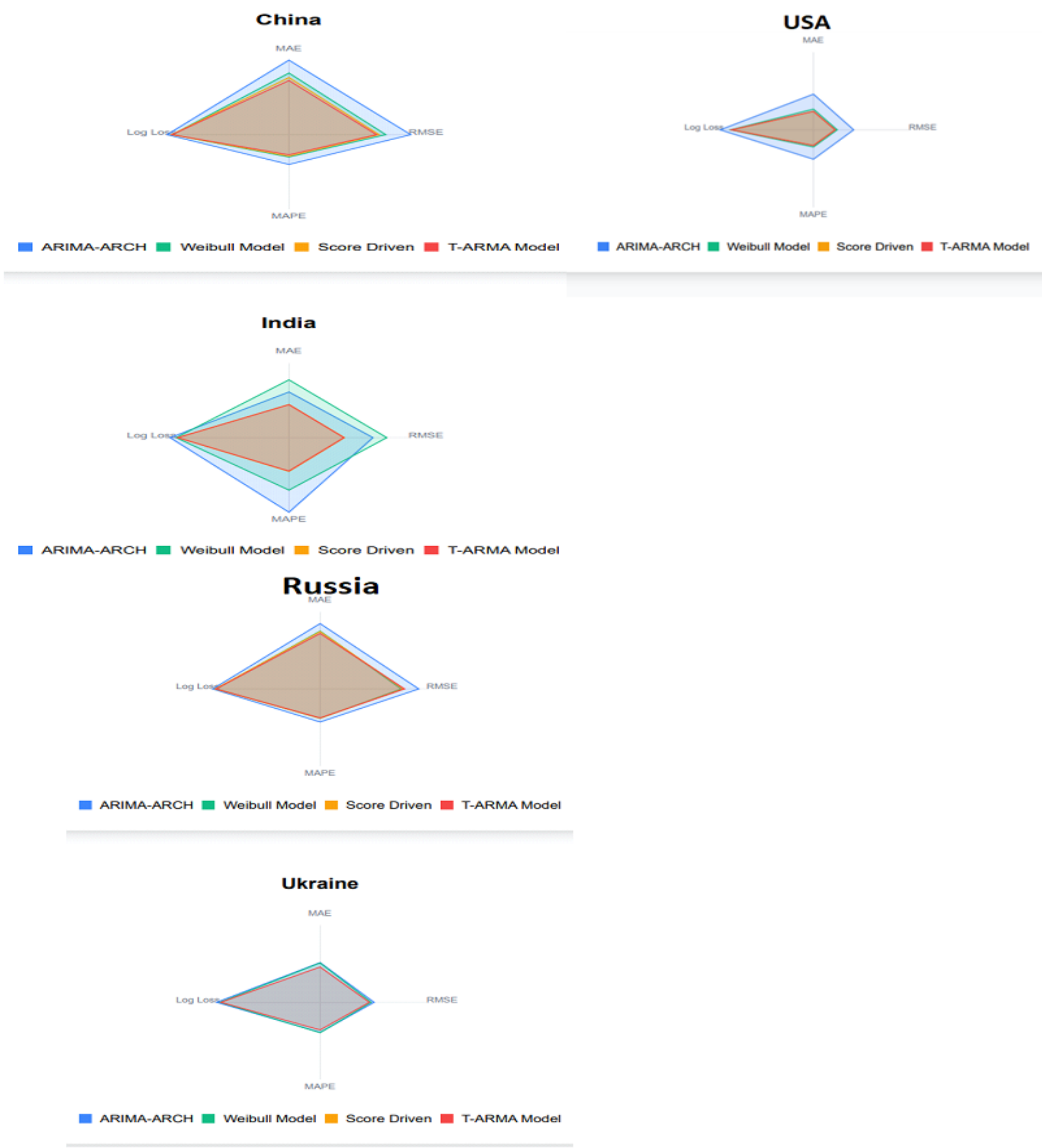


Fig 6: Radar chart comparing the goodness of fit for all models based on the Training Dataset

Fig 6: Radar chart comparing the goodness of fit for all models based on the Validation Dataset