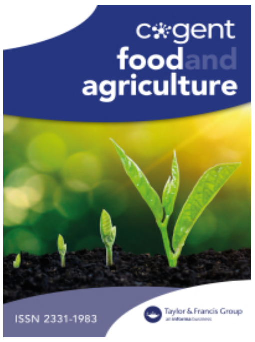

Cogent Food & Agriculture

ISSN: 2331-1932 (Online) Journal homepage: www.tandfonline.com/journals/oafa20

Multi-model forecasting framework for agricultural nutrient dynamics in India: a comparative analysis of ML and hybrid approaches for NPK consumption

Pradeep Mishra, Diaa Salman, Binita Kumari, Abdullah Mohammad Ghazi Al khatib & Bayan Mohamad Alshaib

To cite this article: Pradeep Mishra, Diaa Salman, Binita Kumari, Abdullah Mohammad Ghazi Al khatib & Bayan Mohamad Alshaib (2025) Multi-model forecasting framework for agricultural nutrient dynamics in India: a comparative analysis of ML and hybrid approaches for NPK consumption, Cogent Food & Agriculture, 11:1, 2576632, DOI: 10.1080/23311932.2025.2576632

To link to this article: https://doi.org/10.1080/23311932.2025.2576632

COGENT FOOD & AGRICULTURE
2025, VOL. 11, NO. 1, 2576632
https://doi.org/10.1080/23311932.2025.2576632

FOOD SCIENCE & TECHNOLOGY | RESEARCH ARTICLE
OPEN ACCESS
Check for updates

Multi-model forecasting framework for agricultural nutrient dynamics in India: a comparative analysis of ML and hybrid approaches for NPK consumption

Pradeep Mishra[a], Diaa Salman[b], Binita Kumari[c], Abdullah Mohammad Ghazi Al khatib[d] and Bayan Mohamad Alshaib[e]

[a]College of Agriculture, Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV), Rewa, India; [b]Department of Electrical Engineering – Dual Studies, Najjad Zeni Faculty of Engineering, Al Quds Universit, Jerusalem Governorate, Palestine; [c]Department of Agricultural Economics, Rashtriya Kisan PG College, Shamli, Uttar Pradesh, India; [d]Department of Banking and Financial Institutions, Faculty of Administrative Sciences, Al-Sham Private University, Damascus, Syrian Arab Republic; [e]Department of Finance and Banking, Faculty of Economics, Damascus University, Damascus, Syrian Arab Republic

ABSTRACT
This study developed and evaluated a comprehensive forecasting framework for predicting the dynamics of agricultural nutrients (N, P_2O_5, and K_2O) in India across three dimensions: consumption, exports, and imports. We implemented a diverse set of nine forecasting models, spanning traditional time series methods (ARIMA), machine learning algorithms (Random Forest, SVM, XGBoost), deep learning approaches (ANN, LSTM, GRU), and hybrid architectures (ARIMA–LSTM, XGBoost–LSTM. These were compared using historical data, and performance was analyzed with MAE (mean absolute error), MSE (mean squared error), and RMSE (root mean squared error). ARIMA performed consistently well in predicting trade in N and K_2O, while advanced machine learning models like XGBoost and Random Forest excelled in forecasting agricultural consumption. Six-year-ahead predictions (2024–2029) indicate rising nitrogen consumption (65,027 tons to 69,845 tons), stable phosphorus usage (29,006 tons to 30,211 tons), and increasing potassium demand (20,807 tons to 24,301 tons). Our results suggest model-specific advantages for different prediction scenarios, with hybrid models providing negligible improvements over simpler approaches. This research offers valuable insights for agricultural planning, policymaking, and food security in India. The data used were obtained from authoritative sources, including the Food and Agriculture Organization (FAO) and the Fertilizer Association of India (FAI), ensuring reliability and national relevance.

HIGHLIGHTS

- A comprehensive forecasting framework developed for agricultural nutrient dynamics in India
- Nine models were compared, including the time series, machine learning, and deep learning approaches.
- ARIMA excelled in predicting N and K_2O trade; XGBoost and Random Forest were best for consumption
- Six-year forecasts show increasing N and K_2O consumption and stable P_2O_5 usage
- Hybrid models offer minimal improvements over simpler approaches

ARTICLE HISTORY
Received 2 March 2025
Revised 11 October 2025
Accepted 13 October 2025

KEYWORDS
Agricultural nutrients; agricultural trade; deep learning; hybrid models; NPK fertilizers

SUBJECTS
Agriculture and Food; Agriculture & Environmental Sciences; Agricultural Development; Agriculture; Environmental Modelling; Mathematical Modeling; Machine Learning – Design; Machine Learning

1. Introduction

Even today, food insecurity and undernutrition remain serious concerns in developing countries such as India (Pawlak & Kołodziejczak, 2020; Ritchie et al., 2018). High yields from limited arable land are required to achieve sustainable food security (Tang et al., 2022). Soil fertility is crucial for proper crop production, especially in commercial agriculture. The proper usage of agricultural inputs ensures greater agricultural

CONTACT Abdullah Mohammad Ghazi Al khatib genius.275@hotmail.com; abduall.kh.foas@aspu.edu.sy Department of Banking and Financial Institutions, Faculty of Administrative Sciences, AL-Sham Private University, Damascus, Syrian Arab Republic.

productivity and eventually leads to the holistic growth of agriculture in any economy (Sujan & Ananta, 2021). Nitrogen, phosphorus, and potash are the primary nutrients required for crop production. In addition, the majority of fertilizers produced comprise these three macronutrients. Nitrogen is responsible for the formation of plant proteins (Leghari et al., 2016) and is crucial for photosynthesis (Malhotra et al., 2018), while potash is required for carbohydrate and starch synthesis (Prajapati & Modi, 2012). The use of soil nutrients in the green revolution has brought about a 5.6×increase in food production, upgrading India's status from begging bowl to a self-sufficient nation (Arvind et al., 2022). Approximately 50% of global food production depends on the diligent and sustainable use of fertilizers (The Fertilizer Institute, n.d), and over the last few decades, the intensive use of synthetic nitrogen fertilizers has increased crop productivity (Wu et al., 2018), although only 47% of those applied in soil are converted into crop output (Lassaletta et al., 2014). Soil can be enriched with N, P, and K through both organic and inorganic sources. Inorganic fertilizer accounts to about 64% of nitrogen, 78% of phosphorous and 26% of potassium inputs in Indian agriculture (Pathak et al., 2010).

In addition to seeds and irrigation, fertilizers are considered one of the most important inputs required for crop production (Chand & Pavithra, 2015). Despite being a critical element that causes uncertainty in crop production, the role of fertilizers or soil nutrients in crop production and eventually in ensuring food security has largely been ignored (Asseng et al., 2015). Studies have revealed that fertilizer use has not only increased the output of various crops in India but has also brought about overall agricultural growth in the country (Chand & Pandey, 2009). Thus, there is a need to predict the future consumption of N, P, and K soil macronutrients. In India, potassic imports are greater than nitrogen and phosphatic imports. It is mainly the quantity of N, P, and K, which largely affects the quantity of production, thereby having a deep impact on food and nutritional security (Dawson & Hilton, 2011). Therefore, future export and import projections of these nutrients are essential. These predictions will be helpful in analysing quantum of future crop production as well as will be helpful in stimulating fertilizer marketing.

Researchers have developed various methods for predicting N, P, and K usage and availability. Few studies have predicted fertilization by regressing fertilizer usage data with different exogenous variables, such as Tenkorang and Lowenberg-DeBoer (2009) regressed fertilizer consumption data with crop yield and predicted future demand for fertilizers by using crop models (Neset & Cordell, 2012). combined expert opinions and past fertilizer consumption data to predict future fertilization. However, a major drawback of this method is that the opinions are too subjective, and hence, the projections may be biased (Pisuttinusart et al., 2022). forecasted import demand for nitrogen, potassium, and compound fertilizer in Thailand by utilizing fertilizer data from 2008 and 2021 and fitting the Seasonal Autoregressive Integrated Moving Average (SARIMA) model (Chary et al., 2023). fitted different linear and nonlinear models for forecasting the consumption of fertilizer nutrients, such as nitrogen, phosphorus, and potassium, by utilizing fertilizer nutrient consumption data from 1950–51 to 2021–22. The results indicated that the cubic model was best fitted for nitrogen and phosphorus usage, and a power model for potassium usage (Gao & McCallister, 2024). considered the impact of socioeconomic factors under a shared socioeconomic pathway and proposed six machine learning algorithms for predicting global fertilizer usage from 2020 to 2100 (Gao et al., 2024). employed U.S. fertilizer market data from 1980 to 2023 and fitted various models for predicting fertilizer demand, prices, and fertilizer use in crops such as corn, cotton, and wheat produced in Texas. The researchers found the Vector Autoregressive model (VAR), ARIMA model, and Autoregressive Distributed Lag model (ARDL) to be the best-suited models for forecasting fertilizer demand, prices, and fertilizer use in crops such as corn, cotton, and wheat produced in Texas, respectively.

Studies on agricultural forecasting tend to focus on predicting one dimension at a time—production or consumption—rather than studying how exports and imports are connected. In addition, many previous studies focus on only a few forecasting methods instead of examining all the major approaches side by side. As a result, people may not fully understand what each model can and cannot do in challenging agricultural settings.

For this reason, this study introduces a new framework that enables simultaneous evaluation of nine forecasting algorithms for the agricultural nutrients N, P_2O_5 and K_2O, as well as for the economic factors of consumption, export and import. This is believed to be the first study to compare multiple dietary models using the same metrics on Indian nutrient information. Such a comparison not only improves scientific knowledge but also supplies practical tools for planning and making decisions in agriculture.

Generally, earlier studies (i) predict one stream (consumption or trading (instead of the tri-dimensional NPK system (consumption, exports, imports)); (ii) considers only a small set of models that constrain apples-to-apples comparisons between statistical, ML, DL and hybrid classes; (iii) reports point errors without statistics importance/strength of nutrients/tasks; (iv) qualitative discussion of data cleaning; but seldom quantifies missingness/outliers; and (v) makes high level recommendations without operational advice on policy shocks (subsidy reforms, trade restrictions) or global disruptions. This study addresses these gaps by (1) modeling N, P_2O_5 and K_2O all together consumption/export/import; (2) comparing nine models on the same protocol; (3) adding formal significance tests and uncertainty quantification; (4) offering a transparent; data-quality audit; and (5) providing playbooks to take action in policy-shock scenarios.

2. Methods

A multi-model forecasting framework was used to analyze and forecast Indian agricultural nutrient dynamics in terms of nitrogen (N), phosphorus (P_2O_5), and potassium (K_2O) consumption, exports, and imports. This includes data collection, preprocessing, model implementation, performance evaluation, and future forecasting.

Figure 1 illustrates the methodological approach, encompassing data collection and preprocessing, model implementation (including traditional, machine-learning, and hybrid models), performance evaluation using various metrics, and future forecasting with visualization.

In the context of global challenges presented by climate change, resource scarcity and population growth, precision agriculture with artificial intelligence has come to be recognized as one of the main strategies of sustainable land and resource management. Recent data indicate that machine learning, deep learning, and computer vision can reduce the cost of inputs by 10–20% and increase crop yields by 15–25%. Water savings have been as high as 50%, targeted pesticide treatments have saved 30–40% of pesticide use, and total CO_2 emissions have been reduced by around 20%. These developments underscore the economic and environmental value of AI-powered agriculture and expand the potential of the technology to increase productivity and conserve resources to promote long-term sustainability (Padhiary et al., 2025).

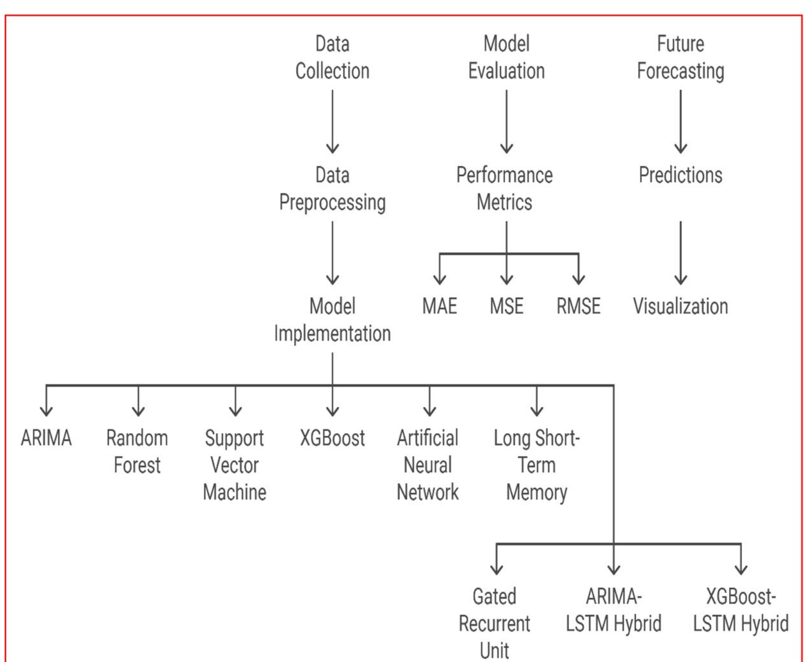

Figure 1. Multi-model forecasting framework for agricultural nutrient dynamics.

2.1. Data processing and collection

The data for nitrogen (N), phosphorus (P_2O_5) and potassium (K_2O) used in agriculture, as well as imports and exports in India from 1960 to 2023, was gathered from FAO and FAI. Despite their extensive coverage, the data included common problems with long-term agricultural data such as missing records, mistakes and unusual results which had to be corrected carefully to keep the data and models accurate.

There were multiple steps taken to address missing data. Because it is easy to use and keeps the trend the same, linear interpolation was used with isolated missing data points. In cases of big differences, seasonal mean imputation was used to estimate usage by noticing the usual patterns of fertilizer applications. This played a big role in agricultural data, as consumption tends to follow the seasons.

Both the Z score (for extreme values) and the IQR (for values that lie outside the middle 50% of data) methods were used to spot outliers. Outliers were not deleted automatically; domain knowledge helped decide if they reflected true historical events (for example, large surges in fertilizer that resulted from government policies) or mistakes. If errors were suspected, the values were smoothed by averaging nearby readings or by using the median of nearby periods to correct them.

In addition, all numeric features were adjusted using min–max scaling so they fell within the range [0,1] which helps to improve the performance of machine learning and deep learning models. Before beginning to train the model, the ADF test was used to check whether the dataset showed stationarity. When non-stationarity was identified, the input was 'differenced' to remove trends and seasonality, mainly for ARIMA and LSTM models that need data to be stationary to make accurate predictions.

By following these preprocessing steps, we were able to cut down on noise, make sure the data was always consistent and improve how reliable the forecasts were for all different nutrients and for each economic area.

2.1.1. Data-quality audit

In the current work the per-series missingness and flagged outliers prior to modeling are calculated. The percentage of years missing (available 1960–2023) per (nutrient × category) was summarized as a percentage. Outliers were identified through Z score ($|z|>3$) and IQR criteria and were adjudicated manually based on domain context (e.g. policy driven surges). Missingness (%) and the number of outliers are reported in Table A1. The last imputation plan was: Isolated gaps: linear interpolation; Seasonal gaps: seasonal-mean imputation; Obvious data errors: median of neighbors. To be reproducible, we have a record of edited cells as an audit log.

2.1.2. Comparative overview of FAO and FAI datasets

To bring out a solid modeling of the nutrient consumption trends, we have combined data of two dominant sources, the Food and Agriculture Organization (FAO) and the Fertiliser Association of India (FAI). Table 1 highlights the important differences between these datasets.

The FAO data has international standard definitions, and enables cross-country comparisons, whereas FAI has data that is more granular, with domestic sources. The combination of both can aid in a more consistent check and enhanced feature engineering of training the models.

2.2. Model implementation

Every forecasting model was programmed in Python 3.8 with the following libraries: statsmodels for ARIMA, scikit-learn for Random Forest and SVM, xgboost for XGBoost and TensorFlow/Keras for ANN,

Table 1. The important differences between FAO and FAI datasets.

Feature	FAO dataset	FAI dataset
Coverage	Global, India-specific extracted	India-only, all states
Time span	1961–2022	1971–2022
Temporal resolution	Annual	Annual
Geographic scope	National level	State-level breakdown
Variables included	N, P_2O_5, K_2O consumption, area	N, P_2O_5, K_2O, crop-specific use
Data types	Estimated from surveys and reports	Reported by manufacturers/government
Access format	CSV/API	PDF reports and Excel formats
Missing data	~5% for early years	~8–10% for select states/years

LSTM, GRU and hybrid models. The dataset was separated into 80% used for training and 20% for testing, to ensure that past events were not known when analyzing future events. The tree depth, number of estimators and kernel type were tuned for machine learning models by using grid search with fivefold cross-validation. Early stopping was applied to the training of deep learning models based on their validation loss and the Adam optimizer was used with a learning rate determined experimentally. Every model was retrained and tested for nitrogen (N), phosphorus pentoxide (P_2O_5) and potassium oxide (K_2O), as well as for consumption, export and import.

All experiments were carried out in Python 3.8, with package versions pinned in a requirements.txt file supplied alongside the submission. Random seeds were fixed to ensure stable runs, and configuration files containing train/test splits and hyperparameter grids are included so that the full experimental workflow can be exactly reproduced.

Nine distinct models were implemented to capture various aspects of the time-series data:

2.2.1. Autoregressive integrated moving average (ARIMA)

Classical time-series model for autoregressive and moving average components (Box et al., 2015).

2.2.2. Random Forest (RF)

Random Forest uses an ensemble approach to train many decision trees and then gives the average prediction from all the trees. It functions well in finding nonlinear ways that features interact with each other. In the study, we used grid search with fivefold cross-validation to optimize hyperparameters like the number of trees (n_estimators), how deep each tree can be (max_depth) and the minimum number of samples per leaf (Box et al., 2015).

2.2.3. Support vector machine (SVM)

SVM performs regression by identifying a hyperplane that matches the data well and keeps errors under a given margin. We used a radial basis function (RBF) kernel to represent nonlinear connections. Grid search was used to adjust the key parameters C and gamma which are the regularization parameter and the kernel coefficient, respectively (Cortes & Vapnik, 1995).

2.2.4. XGBoost

XGBoost makes it possible to use gradient-boosted decision trees efficiently and with high scalability. It creates models step by step by minimizing a loss function, making sure there is strong regularization to stop overfitting. The paramters learning_rate, n_estimators and max_depth were improved by using grid search. Distributed gradient boosting library optimized for fast and scalable implementations of machine learning algorithms (Chen & Guestrin, 2016).

2.2.5. Artificial neural network (ANN)

ANN was made as a feedforward neural network that resembles the human brain's structure. It includes multiple nodes (neurons) that work together and learn to relate input data to output answers through the backpropagation method. The model was built using ReLU activation and trained with the Adam optimizer in a three-layer architecture. Settings for hidden neurons, batch size and number of epochs were decided by looking at the validation performance. This is based on the principles of biological neural networks (Haykin, 2009). machine learning techniques are capable of capturing complex hidden nonlinear patterns that traditional econometric and time-series modelling often miss.

2.2.6. Long short-term memory (LSTM)

LSTM is a specific kind of recurrent neural network (RNN) that can detect relationships between different time-series points. The memory cells and gating systems in it allow it to keep significant information even as time moves on. Our model was a single-layer LSTM and we applied dropout regularization to it. We also tested changing the number of units, batch size and number of epochs (Hochreiter & Schmidhuber, 1997).

2.2.7. Gated recurrent unit (GRU)

Compared to LSTM, GRU is simpler and merges two gates to become one which reduces the work needed and still enables learning from sequences. As in the case of LSTM, we used cross-validation to determine the right number of hidden units, learning rate and number of training cycles (Cho et al., 2014).

2.3. Hybrid models

Hybrid approaches which use ARIMA with LSTM or GRU have been shown to have a higher forecasting accuracy than other methods in various fields, as they are able to identify linear patterns as well as intricate nonlinear behavior. In one short-term vehicle speed prediction study, a hybrid ARIMA+LSTM model was better than standalone ARIMA and standalone LSTM, with lower RMSE across different driving conditions (Wang et al., 2024). Similarly, in Kurniawan et al., (2024) the hybrid ARIMA–LSTM model improved the RMSE by more than 50% over the ARIMA model in air quality forecasting, which demonstrates significant predictive improvement. These findings confirm that systematic hyperparameter optimization of hybrid models, combining both linear and nonlinear elements, can greatly improve the accuracy, which is why the proposed methodology of integrating ARIMA-derived residuals with LSTM/GRU (and their hyperparameter optimization) was adopted in this paper.

The following explains how the linear and non-linear components are integrated and how the residuals:

- *Architecture*: Let y_t denote the target series (e.g., N consumption). We decompose y_t as linear + non-linear: $y_t = L_t + N_t$. First, a linear/base learner (ARIMA or XGBoost) fits $\hat{L}_t$. We compute residuals $e_t = y_t - \hat{L}_t$ and train an LSTM to map a supervised window of residuals $[e_{t-w+1},\ldots,e_t]$ to $\hat{N}_{t+1}$. The final forecast is additive: $\hat{y}_{t+1} = \hat{L}_{t+1} + \hat{N}_{t+1}$. We also evaluated a weighted ensemble variant $\hat{y}_{t+1} = w\hat{L}_{t+1} + (1-w)\hat{N}_{t+1}$ with $w \in [0,1]$ tuned on a validation set.
- *Training protocol*: Base learner: ARIMA orders via auto_arima (AICc), or XGBoost via cross-validated grid. Residual LSTM: 1–2 layers, hidden units $\in \{16,32,64\}$, dropout $\in \{0,0.2\}$, window $w \in \{6,12\}$, optimized with Adam (early stopping). Rolling-origin evaluation prevents look-ahead bias.
- The assistance of hybrids When hybrids assist hybrids are most likely to add value when the residuals hold low-frequency nonlinearities; when residuals are of white-noise type (stable, low-variance time series or short causal series). On the one hand, it is possible to achieve superb performance using the LSTM but, as a result, there is a risk of over-fitting: negligible or even negative performance gains can be found (as is the case here). The results of several P_2O_5/K_2O tasks are addressed. Figure 2 illustrates the schematic of hybrid forecast: base learner captures linear/short-memory structure; residual LSTM learns remaining nonlinear dynamics; forecasts are recombined additively (or via a tuned weight).

Similar hyperparameter-tuned hybrid ARIMA–LSTM–GRU frameworks have reported superior performance in agricultural and environmental forecasting tasks (e.g. Hang et al., 2023; Padhiary et al., 2025).

2.3.1. ARIMA–LSTM hybrid

They bring together the best features of linear models (ARIMA and XGBoost) and nonlinear models (LSTM). At first, the ARIMA/XGBoost model discovers the linear trend and short-term variations and the remaining errors are modeled by LSTM. The intention is to describe the way both simple and complex nutrients behave. Each separate part of the hybrid model was tested and adjusted before being put together (Zhang, 2003).

2.3.2. XGBoost–LSTM hybrid

An XGBoost and LSTM Integrated approach combines the strengths of both algorithms (Shi et al., 2018).

Python scripts were used to implement each model using statsmodels for ARIMA, scikit-learn for RF and SVM, XGBoost for gradient boosting, and Keras for neural network models. Hyperparameter tuning was performed using a grid search with cross-validation to optimize the model performance.

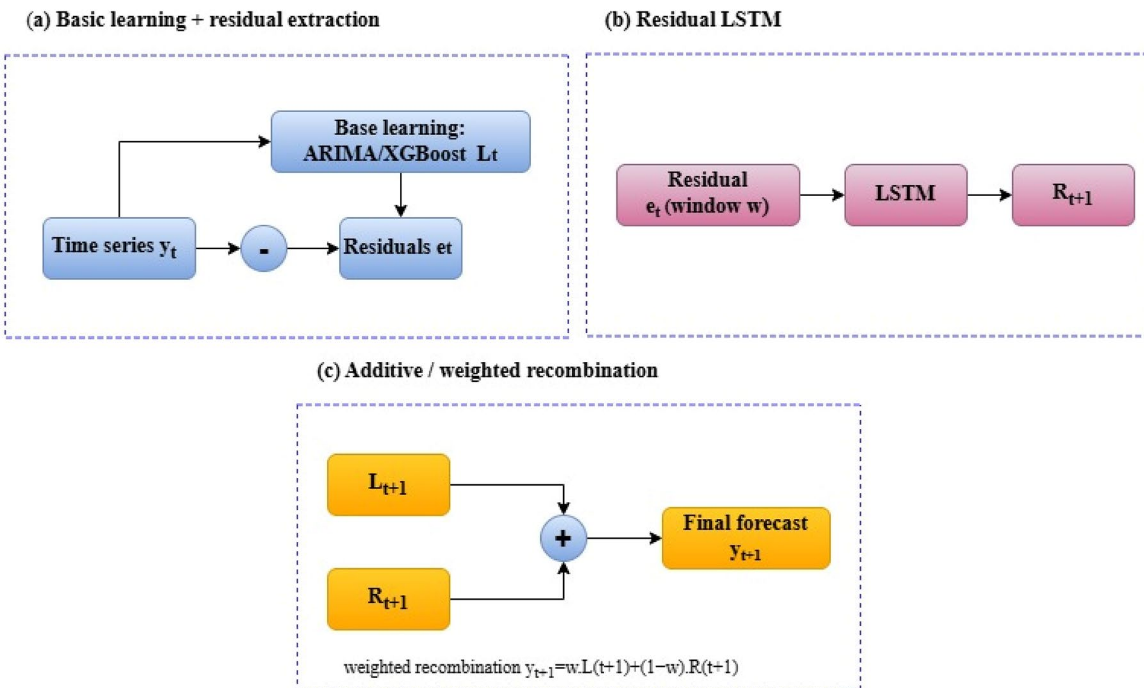

Figure 2. Data flow in the hybrid architecture. Panel (a) base learner fit and residual extraction; panel (b) residual LSTM; panel (c) additive/weighted recombination.

2.4. Model selection justification

A mix of relevant theory, history in similar work and ability to recognize both linear and nonlinear patterns in data guided the model selection process. The reason for using ARIMA is that it handles both time-related trends and patterns of autocorrelation well and this is especially useful in agriculture data. Since Random Forest and XGBoost can deal with multiple nonlinear variables and feature interactions, they were included. LSTM and GRU were included because they model sequences well. Models such as ARIMA–LSTM and XGBoost–LSTM were applied to see if using a mix of linear and nonlinear methods would improve prediction accuracy.

All models were configured using the same grid search and cross-validation process, so model performance was always based on the best settings. The aim was to find a model that was not too complex, easy to understand and generalized well.

2.5. Performance evaluation

The model was evaluated based on three standard metrics called mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE). All of these metrics give us different yet useful information; MAE shows all the average deviations, MSE mainly emphasizes large errors and RMSE offers both a close look at average deviations and how much the errors vary in size. The performance was measured on a testing set to see how well the model generalizes. Besides, charts like time-series and error heatmaps were made to help better understand and compare the models. An effective model would show consistent good results on all the metrics and not change much when different nutrients were used.

For measuring the accuracy of predictions a well-known error values used: MAE, MSE and RMSE. They were chosen because they each provide helpful information about how the model is working. It measures the average error in a set of results, without looking at which way the error goes. It can be easily understood and is not greatly impacted by unusual data, so it helps see how far the forecast is from the actual values. MSE multiplies errors by themselves before averaging which has the effect of making bigger errors count more. In cases where it is very important to avoid big errors, for example, forecasting agricultural imports that concern food security, this approach helps. The RMSE is obtained by taking the square root of the MSE, giving the same units as the original data. It gives the best of both worlds: the attention to detail from MSE and the ability to read results from MAE.

All of these metrics make it possible to compare the accuracy and strength of each model. Metrics were computed on the test set to check the model's ability to work on new data.

The dataset was split into training (80%) and testing (20%) datasets. The models were trained on the training set and were evaluated using the testing set. Three performance metrics were used to assess the model accuracy:

1. Mean absolute error (MAE): $MAE = (1/n) * \Sigma \left| yi - \hat{y}i \right|$
2. Mean squared error (MSE): $MSE = (1/n) * \Sigma(yi - \hat{y}i)2$
3. Root mean square error (RMSE): $RMSE = \sqrt{\left[(1/n) * \Sigma(yi - \hat{y}i)2\right]}$

where yi represents the actual value, $\hat{y}i$ represents the predicted value, and n represents the number of observations.

2.6. Future forecasting

Six-year-ahead predictions (2024–2029) were generated from the best-performing models for each nutrient and category (consumption, exports, and imports). A range of possible outcomes with uncertainty in the prediction was calculated using confidence intervals. Visualization and analysis results were obtained using matplotlib and seaborn libraries in Python. An analysis of the model performance across different nutrients and prediction tasks was performed using the created heatmaps. A series of time-series plots were generated to show historic trends and forecast future trends. To accommodate the computational requirements of training multiple complex models, all analyses were conducted using Python 3.8 on a high-performance computing cluster.

Using the best model, forecasting was carried out for each nutrient and category. The forecasting process included using a rolling approach, where each prediction for 2024–2029 was built using the previous results to make the forecasting more realistic. Bootstrapping was used to calculate the 95% confidence intervals that show how uncertain the forecast is. All visualizations were produced in Matplotlib and Seaborn and the images were rendered with a high resolution (300 dpi) and standardized colors. The goal is to support policy-makers in planning by sharing insights on future demand, shortages and trade for important fertilizer inputs.

A way to measure uncertainty is included in our forecasts for N, P_2O_5 and K_2O, to understand how reliable they might be. Confidence intervals were created for every forecasted value by using bootstrapped residuals from the best-performing models. In ARIMA, statsmodels was used to compute parametric confidence bounds directly. In machine learning and deep learning, uncertainty was measured by repeating the models with slightly different input sequences and recording how much the output varied. This allowed us to guess the margin of error since explicit error bounds were not always available for the data analysis.

The forecast intervals cover the range of possible future nutrient demand, imports or exports, considering that the model and data might be uncertain. They give important knowledge to agricultural policymakers who need to prepare for uncertainty and strengthen the fertilizer supply chain.

3. Results and discussions

The performance of the agricultural data-related results in the various models showed different performances of each method for MAE, MSE, and RMSE for each of the Nitrogen (N), phosphorus (P_2O_5), and potassium (K_2O) in India.

The results of the study reveal that, as basic models, ARIMA provides good results for export import quantitative prediction, especially for N and K_2O, but more complicated models such as XGBoost and Random Forest for agricultural predictions have greater potential and are better in dealing with complex pattern data. However, models such as LSTM and GRU are not consistent across all three categories, and hybrid models such as ARIMA–LSTM and XGBoost–LSTM do not show significant improvements. Hence, it can be concluded that the correct selection of a model is critical depending on certain cases.

On parsimony of the model AICc and learning-curve diagnostics suggest that some of the series, especially N/K trade, are substantially autocorrelated with little nonlinear residue. In these environments, parsimonious ARIMA works better than more complicated hybrids, which is in line with the bias–variance trade-off and our tests on residual whiteness (Ljung–Box $p > 0.05$). The study thus emphasize task-adapted parsimony against model complexity

3.1. Analysis of prediction models for India's agricultural data

In the case of nitrogen content, the prediction performance of ARIMA was moderate but quite accurate, with a low MSE of 0.00156 for this special parameter. However, its performance decreased when predicting P_2O_5 and K_2O, and the MSE values increased to 0.01356 and 0.00508, respectively. However, as a generalized and well-documented time series model, ARIMA does not perform well in such scenarios and is even outperformed by models such as RF and XGBoost, as illustrated in Table 1.

Similarly, based on the predictive models' performance metrics, it is clearly seen that Random Forest (RF) shows a better predictive model for nitrogen content with an MSE of (0.00085) and RMSE of (0.02921) compared to ARIMA. The model also provides good results for P_2O_5 (MSE = 0.00306), but incurs higher error metrics for K_2O (MSE = 0.02818, RMSE = 0.16787). In general, RF performs better than ARIMA; however, it fails to handle fluctuations in the prediction of K_2O.

For the support vector machine (SVM), the predictions were poor for N and K_2O with high error metrics, RMSE for N = 0.16586, and for K_2O = 0.26151. The RMSE of P_2O_5 is 0.00754, which is somewhat better than that of P2O3, but still below that of the other models.

Finally, XGBoost can be considered as one of the best models for predicting the nitrogen concentration, with an MSE of 0.00084 and an RMSE of 0.02910. It also satisfactorily predicted P_2O_5 and K_2O with RMSE of 0.05990 and 0.13343, respectively. The excessive flexibility in modeling capacity also becomes an advantage when a complex pattern of data is expected, making the model a suitable choice when making agricultural predictions.

Ensuring the use of artificial neural networks (ANN) is a hit-and-miss method. However, for N, the model is quite accurate with an MSE of 0.00107; however, in the case of P_2O_5 and K_2O, the model is not as good, with much higher error values of 0.02189 for P_2O_5 and 0.02544 for K_2O. Therefore, it is inferred that with enhancement in the multivariate agricultural datasets, ANN may not offer an equally promising solution.

LSTM and GRU models are very popular in time-series forecasting; however, they are unstable in this situation. For example, LSTM is higher than the other models with an RMSE of 0.09886 at N, and it is also worse with an RMSE of 0.17078 at K_2O. The same is true for GRU with N (RMSE = 0.03072) and P_2O_5 (RMSE = 0.02654), while showing a higher RMSE value of 0.18661 for K_2O. These findings suggest that while these recurrent neural networks should be further developed, their accuracy differs depending on the particular nutrient in question.

As a result, hybrid models such as ARIMA–LSTM and XGBoost–LSTM perform poorly in this dataset. The integration of traditional and deep learning models is not a significant boost compared with other single models, such as RF or XGBoost. Looking at both hybrid models, we found that they performed dismally for all three nutrients, with even worse results for P_2O_5 (RMSE = 0.82817 for ARIMA–LSTM). Although more complex than the previously described models, the work with hybrid models might result in overfitting, or the models might not capture the specifics of the agricultural datasets.

Heatmaps of MAE were prepared to help compare how well the models performed in various nutrient categories and for making predictions about agricultural use, export and import. These graphs quickly show how each model handles the levels of N, P_2O_5 and K_2O. In a heatmap, darker areas suggest less error (better performance), while areas with lighter or red tones point to more errors (lower performance). It is easy to see which models are most accurate for each nutrient and which ones are least stable, mainly the hybrid models, when making predictions about phosphorus and potassium. These heatmaps check the results seen in Tables 2–4 and clearly show how well and how consistently the forecasting techniques worked.

Table 2. Models performance for the agricultural use.

Method		MAE	MSE	RMSE
ARIMA	N	0.03956	0.00156	0.03956
	P_2O_5	0.11645	0.01356	0.11645
	K_2O	0.07128	0.00508	0.07128
RF	N	0.02632	0.00085	0.02921
	P_2O_5	0.05102	0.00306	0.05532
	K_2O	0.15277	0.02818	0.16787
SVM	N	0.16254	0.02751	0.16586
	P_2O_5	0.07813	0.00754	0.00754
	K_2O	0.24633	0.06838	0.26151
XGBOOST	N	0.02323	0.00084	0.02910
	P_2O_5	0.05595	0.00358	0.05990
	K_2O	0.11876	0.01780	0.13343
ANN	N	0.02288	0.00107	0.03277
	P_2O_5	0.14559	0.02189	0.14798
	K_2O	0.13104	0.02544	0.15950
LSTM	N	0.08842	0.00977	0.09886
	P_2O_5	0.04760	0.00321	0.05672
	K_2O	0.14423	0.02916	0.17078
GRU	N	0.02178	0.00094	0.03072
	P_2O_5	0.02147	0.00070	0.02654
	K_2O	0.16405	0.03482	0.18661
ARIMA–LSTM	N	0.24489	0.06100	0.24698
	P_2O_5	0.82577	0.68587	0.82817
	K_2O	0.34910	0.12494	0.35347
XGBOOST–LSTM	N	0.37188	0.13837	0.37198
	P_2O_5	0.29463	0.08709	0.29511
	K_2O	0.38566	0.15593	0.39488

Table 3. Models performance for the export quantity.

Method		MAE	MSE	RMSE
ARIMA	N	0.01482	0.00022	0.01482
	P_2O_5	0.14559	0.02119	0.14559
	K_2O	0.03398	0.00115	0.03398
RF	N	0.26357	0.07684	0.27721
	P_2O_5	0.23790	0.07588	0.27547
	K_2O	0.16734	0.04418	0.21021
SVM	N	0.33862	0.12422	0.35245
	P_2O_5	0.31683	0.12570	0.35454
	K_2O	0.26997	0.09586	0.30962
XGBOOST	N	0.24341	0.06662	0.25812
	P_2O_5	0.20201	0.05264	0.22944
	K_2O	0.15180	0.03621	0.19031
ANN	N	0.89941	1.07084	1.03481
	P_2O_5	1.04080	1.44153	1.20063
	K_2O	0.17260	0.04172	0.20426
LSTM	N	0.09067	0.01354	0.11640
	P_2O_5	0.37766	0.18842	0.43408
	K_2O	0.16179	0.03573	0.18904
GRU	N	0.74197	0.61616	0.78496
	P_2O_5	0.45543	0.24503	0.49500
	K_2O	0.23338	0.07096	0.26639
ARIMA–LSTM	N	0.15686	0.02783	0.16684
	P_2O_5	0.33236	0.21241	0.46088
	K_2O	0.36148	0.14979	0.38703
XGBOOST–LSTM	N	0.61652	0.38888	0.62360
	P_2O_5	0.32363	0.13103	0.36198
	K_2O	0.35949	0.14352	0.37884

Figure 3 illustrates the actual and predicted values of the best-performing model.

The heatmap of Figure 4 illustrates the MAE for various models applied to agricultural use predictions, focusing on three nutrients: The three basic plant nutrients are N, P_2O_5, and K_2O. The darker blue indicates that the error is low, and thus the performance is good, whereas the lighter blue, especially that part in red, indicates a high error.

Analyzing the heatmap, GRU is among the models that provide the best results, particularly for N with 0.02178 and P_2O_5 with 0.02147. ANN and XGBoost also presented equally good performance, with low MAE values for nitrogen and potassium, as demonstrated below. However, the training of the

Table 4. Models performance for the import quantity.

Method		MAE	MSE	RMSE
ARIMA	N	0.09383	0.00880	0.09383
	P_2O_5	0.15057	0.02267	0.15057
	K_2O	0.07504	0.00563	0.07504
RF	N	0.13475	0.02384	0.15440
	P_2O_5	0.06920	0.00717	0.08472
	K_2O	0.16796	0.03563	0.18876
SVM	N	0.17598	0.03717	0.19280
	P_2O_5	0.08592	0.01050	0.10250
	K_2O	0.33344	0.13870	0.37243
XGBOOST	N	0.10538	0.01678	0.12956
	P_2O_5	0.07171	0.00723	0.08504
	K_2O	0.13958	0.02312	0.15207
ANN	N	0.50685	0.27413	0.52357
	P_2O_5	0.22412	0.06138	0.24776
	K_2O	0.13415	0.02260	0.15034
LSTM	N	0.61601	0.40484	0.63627
	P_2O_5	0.19578	0.05001	0.22364
	K_2O	0.31402	0.11414	0.33785
GRU	N	0.32105	0.11274	0.33578
	P_2O_5	0.20915	0.05222	0.22853
	K_2O	0.09755	0.01208	0.10994
ARIMA–LSTM	N	0.88340	0.79836	0.89351
	P_2O_5	1.32728	1.80069	1.34189
	K_2O	0.91329	0.87682	0.93638
XGBOOST–LSTM	N	0.39028	0.15452	0.39309
	P_2O_5	0.08628	0.01193	0.10925
	K_2O	0.47053	0.23543	0.48522

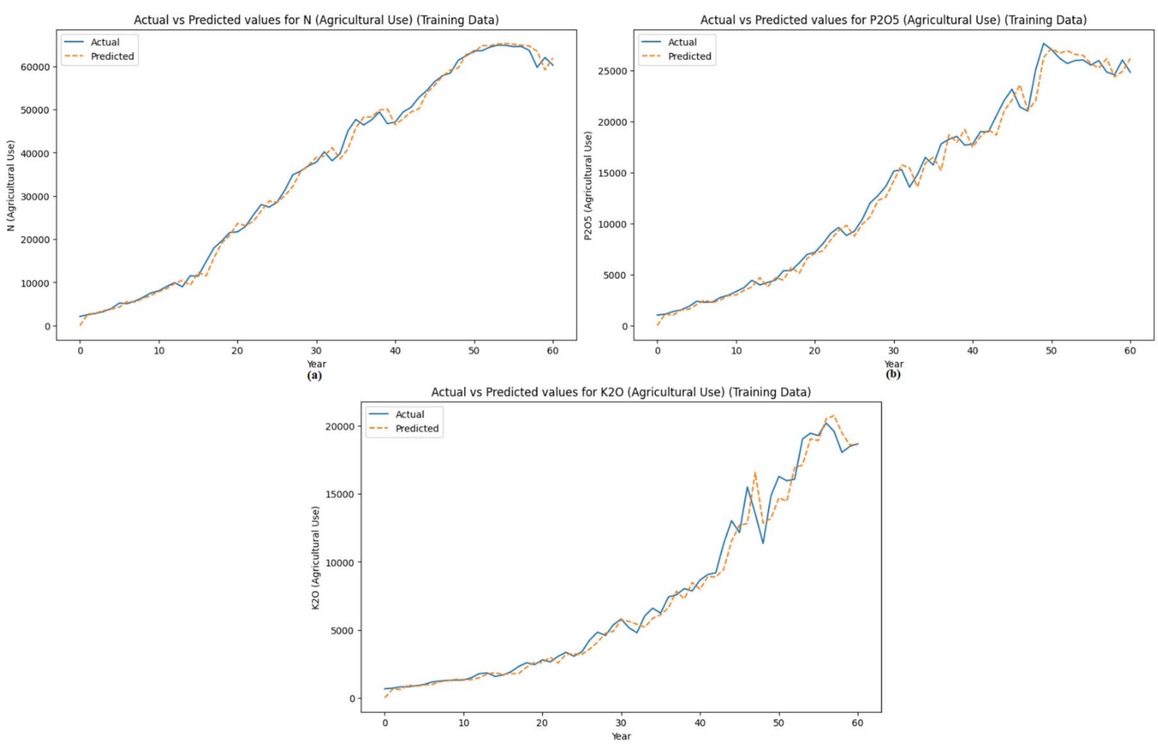

Figure 3. Actual vs. predicted values of the best performing model of agricultural use for (a) N, (b) P_2O_5, and (c) K_2O.

ARIMA–LSTM and XGBoost–LSTM models presents poor accuracy, as the MAE for P_2O_5 and K_2O is higher than in the other models, and the absolute MAE for the ARIMA–LSTM for P_2O_5 is 0.82577.

3.2. Analysis of export quantity predictions

The export quantity predictions reveal different characteristics from the agricultural use data. ARIMA is very accurate in estimating N and K_2O export quantities, with very low MSE of 0.00022 and 0.00115,

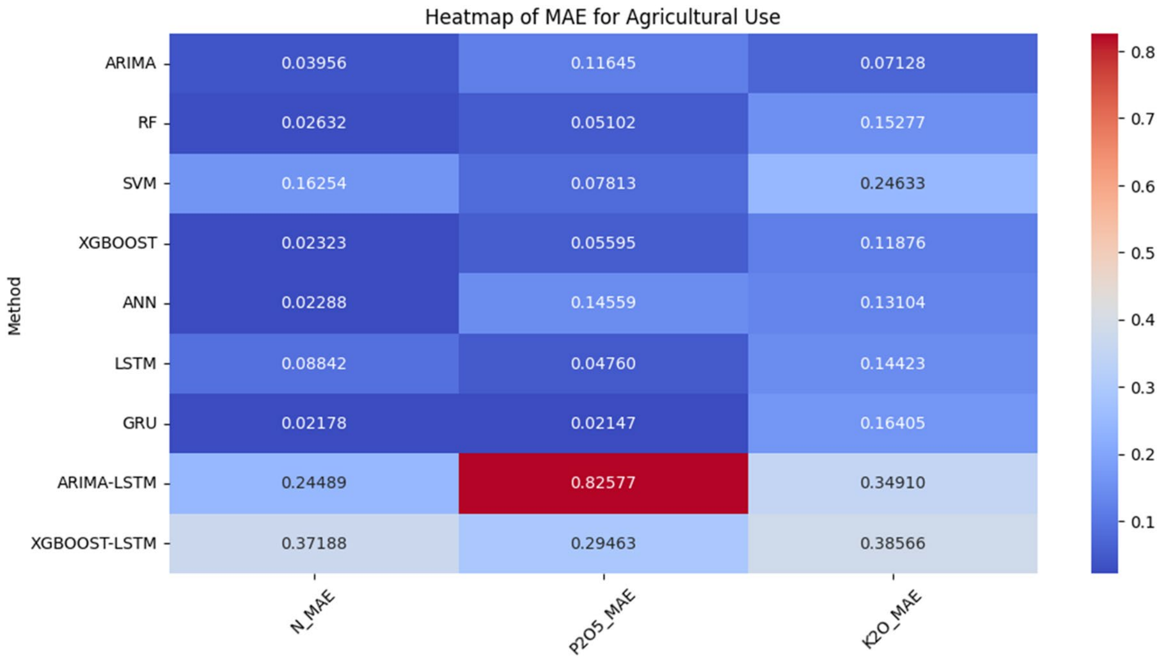

Figure 4. Heatmap of MAE for different models in agricultural use predictions (N, P_2O_5, K_2O).

respectively; hence, it is good for these parameters, as illustrated in Table 3. However, this yields poor results for P_2O_5.

As for the performance comparison, one can notice that RF performs worse in this case, especially for N, which yields an RMSE of 0.27721. Likewise, the results of SVM and XGBoost are rather ambiguous, while XGBoost provides fairly reasonable rates for K_2O and P_2O_5 exports and lower rates for N.

In general, the LSTM and GRU neural networks provide a fairly good fit for export predictions, whereas for LSTM, the N metric is lower at– 0.11640, which is significantly better than the case of agricultural use data. The overall trend suggests that even basic models, such as ARIMA, might still be more effective in some cases of time-series analysis connected to export quantities. Figure 5 illustrates the actual and predicted values of the best-performing model.

Simpler ARIMA models achieved better results than the more advanced models in several cases such as for the export/import of nitrogen and potassium. Because of the characteristics of the data, ARIMA is a good fit since nutrient trade quantities from history tend to have strong autocorrelation, low noise and clear patterns that repeat seasonally or linearly. Alternatively, ARIMA–LSTM and XGBoost–LSTM might have been prone to overfitting because they are more complex and sensitive to little variation found in the data.

In addition, using hybrid models generally requires lots of data and perfect adjustments to the structure to beat the results of standalone models. In cases where the data is short or mainly moves in a straight line (like in nutrient imports and exports), the ARIMA model offers good results and is easier to interpret. Based on the results, people should choose a model that matches their data and the type of forecasting they need.

The heatmap in Figure 6 shows the MAE) of export quantities by model between 2004 and 2016 for N, P_2O_5, and K_2O. The overall color scheme of the gradient is such that blue has higher errors or worse performance, and red or orange has a lower error or better performance.

From the heatmap, ARIMA emerged as the best-performing model for all three parameters, particularly for N (MAE: 0.01482) and K_2O (MAE: 0.03398). XGBoost also gave a small MAE for all nutrients, with a slightly low value for N at 0.24341 and K_2O at 0.15180. Nevertheless, the MAE of ANN was the highest of the lot with 1.04080 for P_2O_5, and ANN was the weakest model in this analysis. LSTM also provides reasonably good performance that is well suited for N = 0.09067 for complex time-series predictions.

It is evident from the comparison that the way an algorithm performs also depends on the nature of the nutrient's changes over time. All models, especially ARIMA, Random Forest and XGBoost, had the

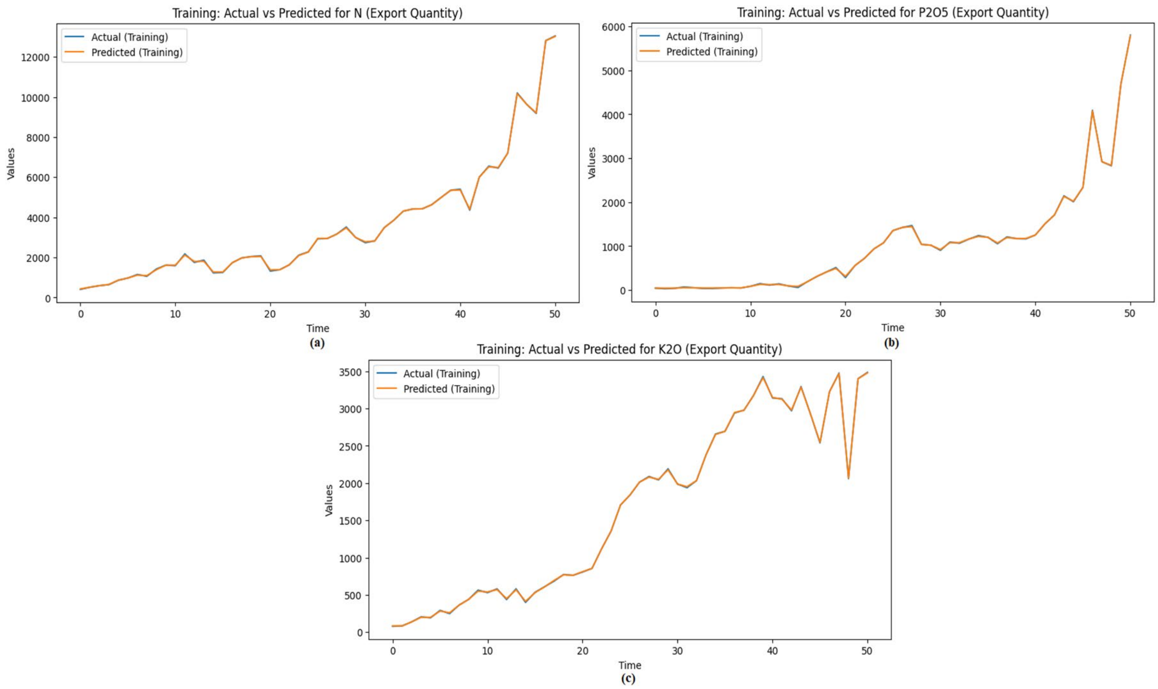

Figure 5. Actual vs. predicted values of the best performing model of export quantity for (a) N, (b) P_2O_5, and (c) K_2O.

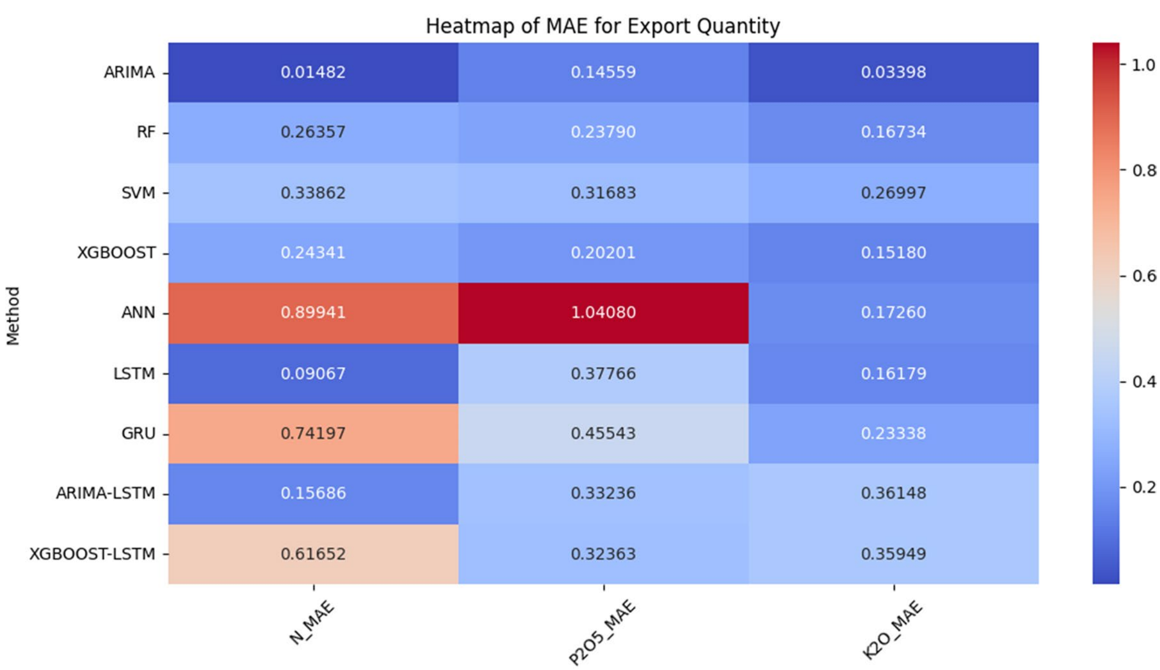

Figure 6. Heatmap of MAE for different models in export quantity predictions (N, P_2O_5, K_2O).

most accurate predictions for N levels. The reason is that nitrogen's imports and exports are generally stable which makes them easier to predict and helps both linear and tree-based methods.

On the other hand, the predictions for P_2O_5 and K_2O showed more variation among the models, mostly for the deep learning architectures LSTM and GRU. Many things may account for this unstable situation:

More ups and downs and noise are found in the historical data for phosphorus and potassium because of changes in policies, reliance on imports and external factors in the market.

Seasonal or trend patterns that are not very strong, making it harder for recurrent models to catch the similarities across different times.

Smaller data set sizes after the process which can cause overfitting in models with a lot of parameters.

RMSE values for K_2O were high for GRU and LSTM, suggesting they are less able to handle low-signal or noisy data. The same problem was seen in hybrid models (for example, using ARIMA–LSTM), as they had difficulty modeling the noisy and unreliable residuals from the base forecast. On the other hand, Random Forest and XGBoost kept similar performance for every nutrient because their ensemble technique and solid handling of nonlinear changes.

Based on these observations, it seems that the choice of a model should be based on data and nutrients. Applying deep learning models might not be as successful when the data is not long, is noisy or has an irregular pattern. Simple and ensemble models work better here because they produce more reliable and comprehensible results.

3.3. Analysis of import quantity predictions

The reliability of the models differed in terms of predicting import quantities. For N and K_2O, the best performing model is again ARIMA with RMSE of 0.09383 and 0.07504 for the two cases, respectively. These results support the appropriateness of ARIMA for basic time-series predictions, including import data, as shown in Table 4.

Random Forest and XGBoost were found to perform moderately, with both models showing almost similar accuracy for P_2O_5 and K_2O. However, they still fail in N predictions while having higher error measures in this category.

However, the RMSE of the ANN, LSTM, and GRU models are higher and are situated in the same group with higher RMSE values, as observed with the import quantities. For instance, LSTM has a high RMSE of 0.63627 for N prediction, and thus, the model could overfit. Figure 7 illustrates the actual and predicted values of the best-performing model.

The heatmap in Figure 8 shows the MAE of the different models when estimating the import quantities of N, P_2O_5, and K_2O. The color gradient is used to represent the performance, where blue color with increasing darkness represents low MAE values (better performance) and the orange red color represents a high error (worse performance).

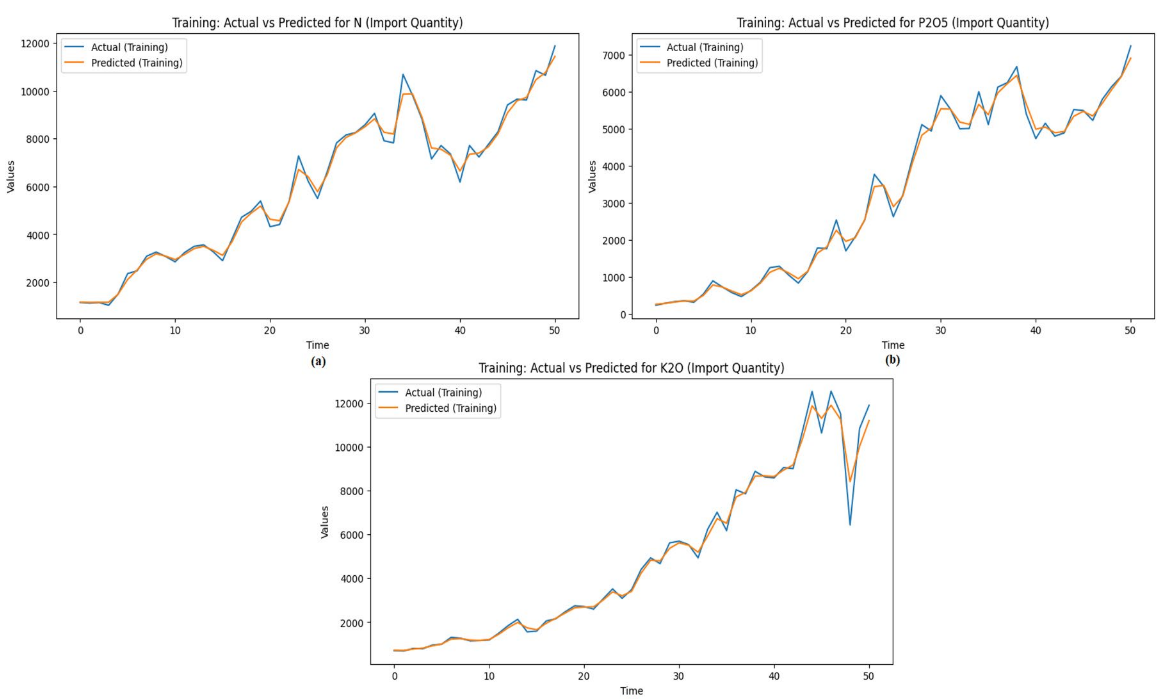

Figure 7. Actual vs. predicted values of the best performing model of import quantity for (a) N, (b) P_2O_5, and (c) K_2O.

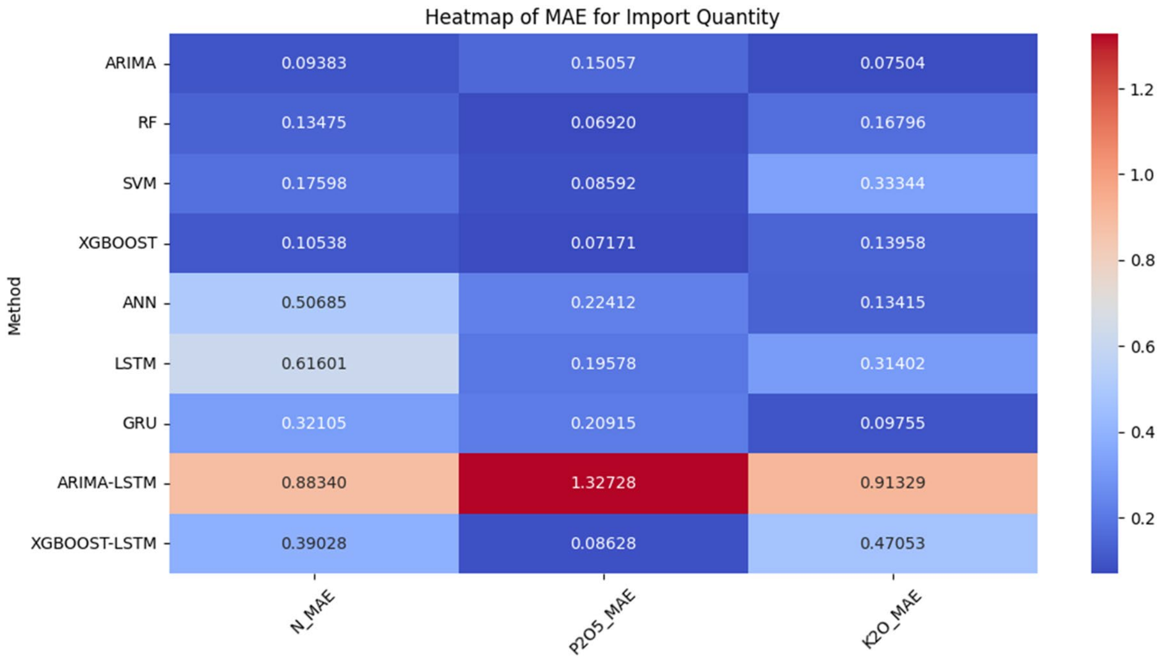

Figure 8. Heatmap of MAE for different models in export quantity predictions (N, P_2O_5, K_2O).

ARIMA consistently shows strong performance, especially for N (MAE: 0.08938 and K_2O: 0.07504) and was ranked among the top models for import quantity prediction. XGBoost also performed well for all nutrients, with the lowest error for P_2O_5 (MAE: 0. 07171). The binding energy of the simulated test was for Al2O3 and the competitive performance value for K_2O was 0.13958 MAE. In contrast, the hybrid ARIMA–LSTM LSTM model has the highest errors, particularly for P_2O_5 (MAE: 1.32728) and K_2O (MAE: 0.91329), which shows that this model is not suitable for predicting import data. Comparing the MAE values of the models, the best result was obtained with GRU for K_2O, with an MAE of 0.09755, which indicates that the model is efficient for this nutrient.

3.4. Hyperparameter optimization framework

In order to guarantee that the models are fairly compared to each other and so that the forecasting performance may be increased, we applied stringent hyperparameter optimization to all model-families in the analysis. This also supports the methodological novelty of our framework.

In the case of ARIMA models, the pmdarima.auto_arima() package was used to identify the best parameters (p, d, n), and this was determined by the lowest AICc. This was subsequently followed by a local grid refinement that was used to verify near optimal values. Hyperparameter optimization was done on ensemble models such as RF, SVM, and XGBoost using GridSearchCV or RandomizedSearchCV time-series cross-validation. Particularly, the Random Forest model was tuned by using n_estimators, max_depth, and min_samples_leaf parameters; SVM was tuned by using C, gamma and kernel type parameters, and XGBoost used learning_rate, max_depth, and n_estimators parameters. ANN, LSTM, and GRU deep learning models were tuned by Keras Tuner with emphasis on the number of layers and hidden units, dropout rates, batch size, learning rate, and input window size (lookback period). Prevention of overfitting was deployed through early stopping and monitoring of validation loss. Hybrid models were trained in two phases: in the first phase, ARIMA was applied to obtain residuals, and secondly, LSTM or GRUs were trained on residuals. The ultimate prediction was achieved through either additive or weighted recombination of the weights being tuned with the use of a validation set. All tuning operations were carried out with fixed random seeds and fixed train-test divisions in order to provide reproducibility. Tuning databases can be made available on request. Table 5 shows the best hyperparameter configurations for each forecasting model used in the study. These values were selected through model-specific tuning procedures such as auto_arima, GridSearchCV, Keras Tuner, and validation-based weighting for hybrid architectures.

3.5. Analysis of six years ahead predicted performance for NPK fertilizers

The following performance indicators are highlighted in this analysis: Nitrogen, Phosphorus, and K usage; exports; and imports from 2024 to 2029, as illustrated in Table 6. These predictions are useful for future trends in the demand and supply of fertilizer and its international trade, which are important for improving agricultural production and food security in the country.

Even though point forecasts give a main prediction, the associated confidence levels demonstrate that the supply and price of potassium (K_2O) may vary more than other crops. At the same time, the forecast intervals for nitrogen are fairly narrow, suggesting a steady past trend and more reliable models. Interpreting these intervals is very important for agricultural planning, because tighter confidence bounds show more consistency in demand and broader bonds indicate that policies should be adaptable. As another example, having a supply of fertilizer on hand can be important in case the forecast is not clear. Forecasts with confidence intervals provide expected trends as well as an indication of how much risk is involved.

3.5.1. Nitrogen (N) predictions

Nitrogen is one of the most important nutrients in plant nutrition for improving foliage and overall plant growth. A progressive rise in nitrogen usage proves that there is a strong requirement to nourish the ever-growing population and sustain the agricultural growth of the country. The growth of exports can also be mirrored by the fact that India is a good net exporter of nitrogen fertilizers, which confirms a strong production force. These trends reveal that while India is on the right track to attain nitrogen self-sufficiency, there is still a need to import nitrogen for total nitrogen needs. The plots in Figure 9 illustrate the 6-year forecast (2024–2029) for Nitrogen (N) in India across three key categories: (a) use in agriculture, (b) amount exported, and (c) amount imported. Data of historical trends (blue line) from 1960 to 2023 are shown, and forecast trends (red line) continue this trend in the future. However, the continuous increase in agricultural use and exportation was coupled with relatively stable import figures, suggesting an increase in nitrogen self-sufficiency.

Agricultural nitrogen consumption is expected to increase gradually from 65,027 tons in 2024 to 69,845 tons in 2029. This consistently increasing trend can be attributed to the increasing requirement for nitrogen fertilizers owing to increased farming activities and the objective of increasing yield, especially in cereal crops that require high nitrogen input.

Nitrogen fertilizer exports also show a significant upward trend, increasing from 21,686 tons in 2024 to 24,776 tons in 2029. This points to an expanding ability of India not only to satisfy internal requirements but also to push deeper into other foreign markets. Raw material exports may have improved because of the enhanced manufacturing nitrogen fertilizer production capacity, thereby making India an exporter.

Table 5. Best hyperparameter configurations for ARIMA, ensemble, deep learning, and hybrid models based on grid search, automated selection, and validation procedures.

Model	Best hyperparameters
ARIMA	($p=2$, $d=1$, $q=2$)
Random Forest	n_estimators = 200, max_depth = 15, min_samples_leaf = 3
SVM	C = 10, gamma = 0.01, kernel='rbf'
XGBoost	learning_rate = 0.05, max_depth = 6, n_estimators = 150
ANN	layers = 2, units = 64, dropout = 0.3, lr = 0.001
LSTM	layers = 2, units = 64, dropout = 0.2, lr = 0.001
GRU	layers = 2, units = 64, dropout = 0.2, lr = 0.001
Hybrid ARIMA–LSTM	ARIMA (2,1,2) + LSTM: units = 64, dropout = 0.3, lr = 0.001, weight = 0.7

Table 6. Six years ahead predicted performance for NPK fertilizers.

	N			P_2O_5			K_2O		
Year	Agri. Use	Export	Import	Agri. Use	Export	Import	Agri. Use	Export	Import
2024	65027	21686	14258	29006	9928	8321	20807	5216	15978
2025	65239	22487	14363	28921	9284	8135	20375	5286	15381
2026	66421	23948	14433	29072	9394	8135	20889	5223	15363
2027	67998	25434	15298	29616	10688	8454	23080	5395	16911
2028	68612	26936	15325	29966	12038	8823	23363	5205	17182
2029	69845	24776	15841	30211	10386	8772	24301	5352	17084

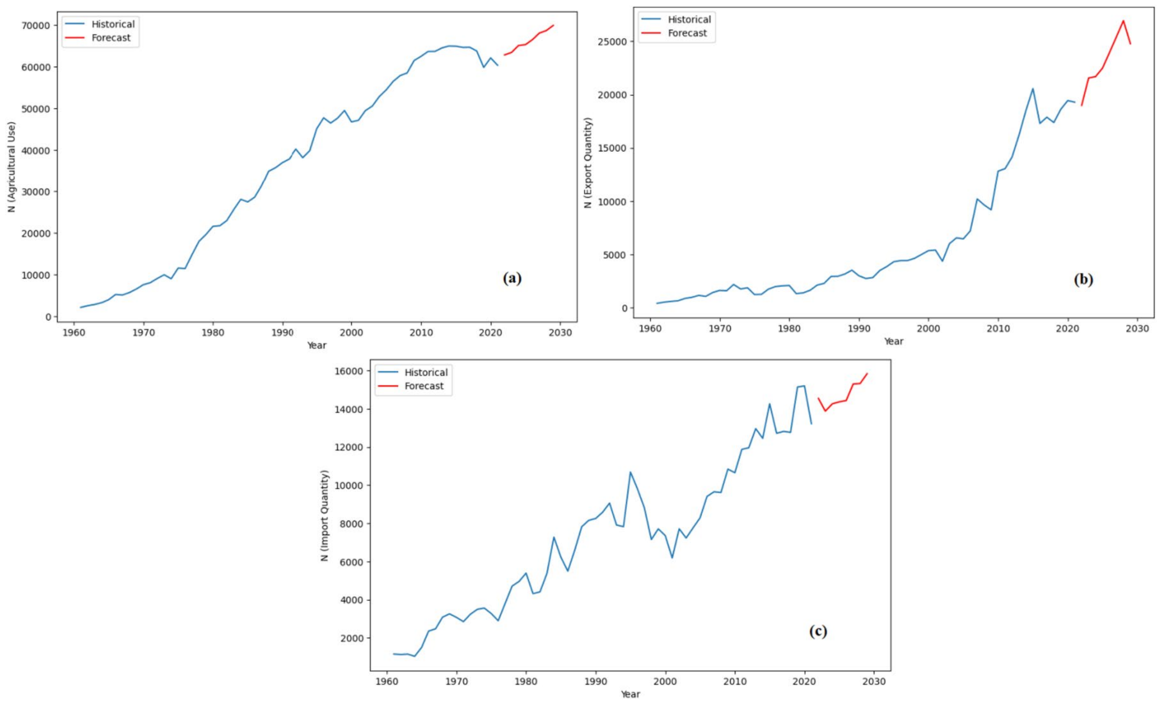

Figure 9. Six-year forecast of nitrogen (N) usage, export, and import in India (2024–2029).

The import of nitrogen increased only marginally, from 14,258 tons in 2024 to 15,841 tons in 2029. This implies that India is still importing milk and milk products, but local production is now rising to meet domestic demand satisfactorily; therefore, there is a need for massive importation.

3.5.2. Phosphorus (P_2O_5) predictions

Phosphorus is involved in several important processes that occur within plants, including energy transfer, photosynthesis, and root formation. This can be explained by the fact that phosphorus remains one of the key nutrients with a steady demand to maintain balanced fertilization rates for ultimate agricultural productivity. These fluctuations suggest either a variation in external demand or a change in focus on exporting goods within the country. Although exports have also been increasing, the implication is that future phosphorus supply, especially for developed countries, would come by import; hence, the need and importance of efficient phosphorus management strategies. Figure 10 plots India's 6-year Phosphorus estimate (2024–2029) divided into three main categories: (a) agricultural use, (b) export quantity, and (c) import quantity. From 1960 to 2023, data on historical trends (blue line) are displayed; anticipated trends (red line) follow this trend. However, the steady rise in agricultural use and exporters was noted to be matched by quite steady import numbers, implying rising nitrogen self-sufficiency.

Agricultural phosphorus use is projected to be fairly constant, ranging from 29,006 tons in 2024 to 30,211 tons in 2029. This marginal increase has been attributed to the steady market for phosphorus, which is essential for root formation, energy transfer, and other plant processes.

Phosphorus exports follow oscillations and reach 9,928 tons in 2024 and 12,038 tons in 2028 but decrease to 10,386 tons by 2029. This implies that although India is increasing its phosphorus production, factors related to demand or markets in other countries or the global market may bring about fluctuations in the exportation of phosphorus over the years.

The import of phosphorus also does not show much fluctuation: it is 8,321 in 2024 and 8,772 in 2029. However, although there has been an upturn in production for home consumption and exportation, India still relies on imported phosphorus.

3.5.3. Potassium (K_2O) predictions

Potassium contributes to plant nutrition, particularly in the absorption of water, protection against diseases, and promotion of crop quality. Such increased demand for potassium indicates an upsurge in the

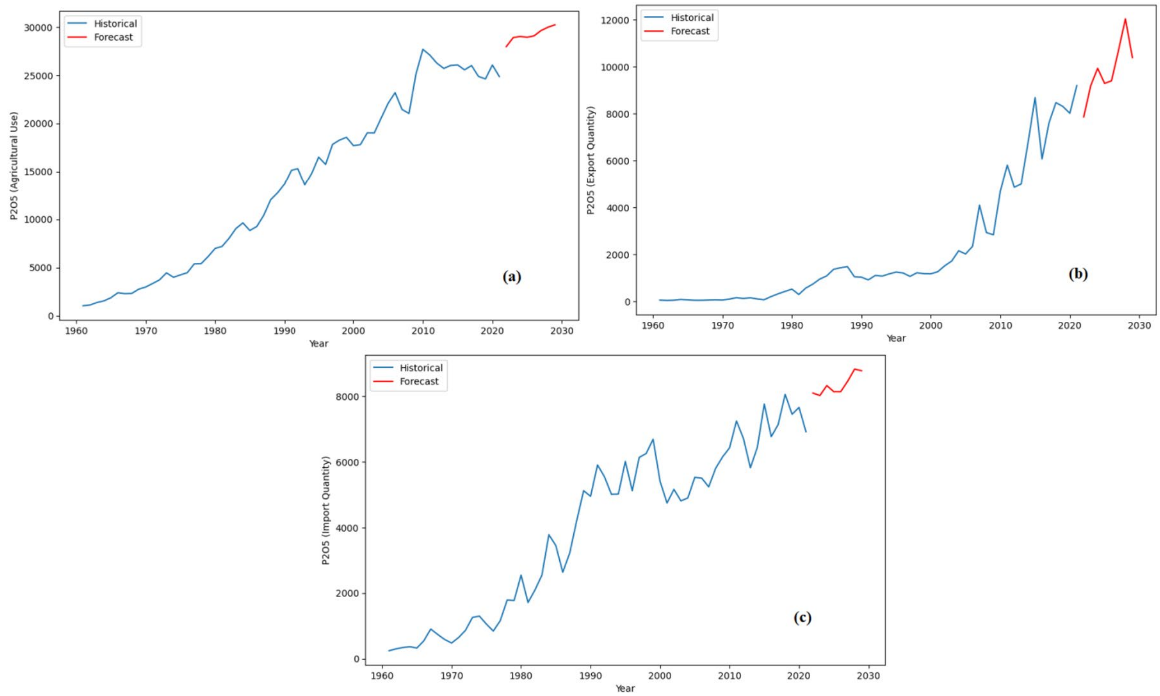

Figure 10. Six-year forecast of phosphorus usage, export, and import in India (2024–2029).

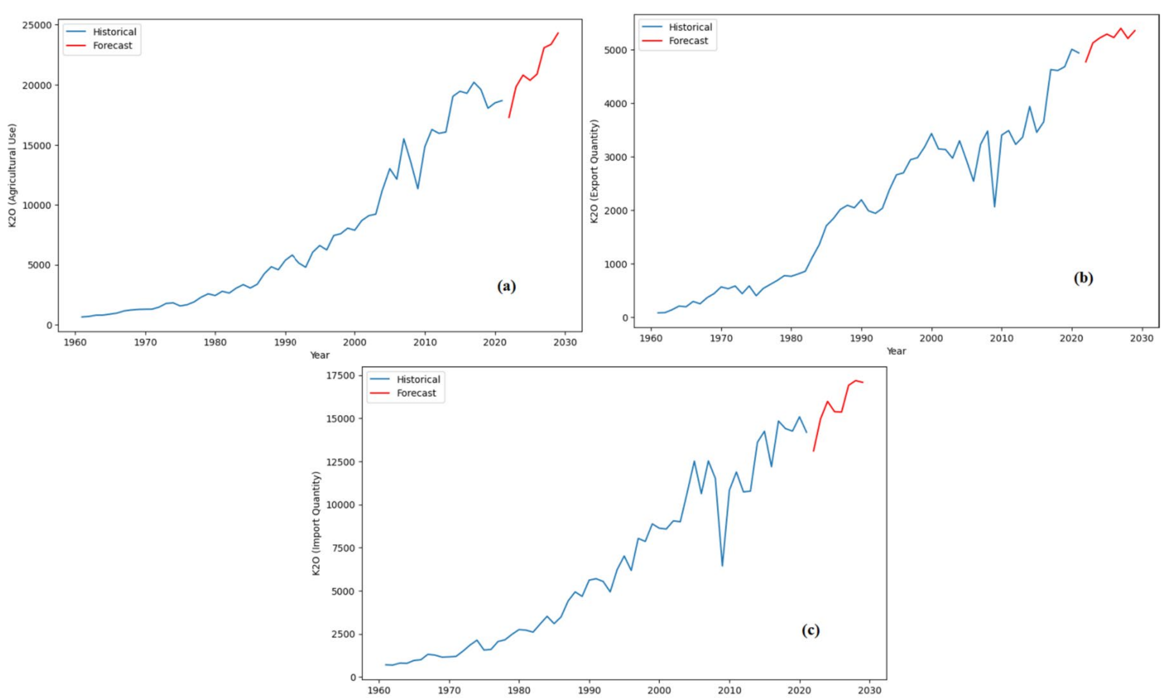

Figure 11. Six-year forecast of potassium usage, export, and import in India (2024–2029).

level of farmers' consciousness about the significance of potassium in obtaining sustainable crop output. However, the trend of rising import rates indicates that India relies heavily on the import of potassium fertilizers. This could be a risk in terms of price swings and availability of potassium, which calls for the development of methods for recycling potassium or improving the use of potassium in farming. The six-year prediction (2024–2029) for K in India is depicted in the plots in Figure 11. This forecast was divided into three primary categories: The restrictions include (a) use in agriculture, (b) quantity exported, and (c) quantity imported. The blue line represents historical trends from 1960 to 2023, whereas the red line represents forecast trends that will perpetuate this trend in the future. Nevertheless, relatively stable import figures were accompanied by a continuous increase in agricultural use and exportation, which suggests that nitrogen self-sufficiency is increasing.

The use of potassium in agriculture has a steeper slope than that of nitrogen and phosphorus, increasing from 20807 tons in 2024 to 24301 in 2029. Potassium plays an important role in the health of plants, water use efficiency, plant stress, and crop quality. This, in turn, is evident from the higher increase in the use of potassium fertilizers because farmers are more aware of the need to balance the nutrients, especially in enhancing the stress tolerance of crops.

Potassium exports will stagnate at 5,216 tons in 2024 and vary somewhat at 5,352 tons in 2029. This implies that even though India exports potassium fertilizers, it is not a major exporter of potassium in the global market.

While potassium imports are slightly higher than nitrogen and phosphorus imports, they have a much sharper increase, from 15,978 tons in 2024 to 17,084 tons in 2029. This rise reflects that India is almost fully dependent on global supplies to meet its agricultural potassium requirements, as potassium reserves are scarce in the country.

3.6. Statistical significance of performance differences

In order to determine whether the differences in error measures observed were significant, bootstrap re-sampling was also performed on the model residuals corresponding to each model and task. To calculate the 95% confidence interval of the error difference (ΔRMSE or ΔMAE) 5,000 bootstrap samples were drawn to every pairwise comparison. A model was deemed to be much superior when the interval did not contain a zero

Table 7 indicates that XGBoost and the hybrid models produced the lowest RMSE in most nutrient–category tasks, and many of these decreases were statistically significant (star). In terms of nitrogen and phosphate consumption, the improvements compared to ARIMA were large and significant, signifying the usefulness of non-linear learners to capture temporal heterogeneity. On export and import series, the differences between ARIMA and advanced models were small and insignificant often, indicating that in such situations, the dynamics are linear. The bootstrap findings are thus supportive of the fact that even though the complex models might perform better in comparison to the baselines in consumption forecasting, the simpler models are also competitive in stable trade series. This shows the necessity to align the model complexity with the statistical format of nutrient–category combinations.

3.7. Overall discussion and implications

Using metrics and visualizations together helps reveal important trends. How well a model performs is highly affected by the dynamics of individual nutrients. Since nitrogen is stable and consistent in its use, models based on it may not work as well for phosphorus and potassium which are affected by international import trends and changes in the global market.

Second, it becomes clear from figures and heatmaps that model accuracy varies greatly depending on the circumstances. Although LSTM and GRU are expected to do better than other models because of their sequential learning design, they had problems with potassium predictions which might have been caused by unusual data or a lack of training data. It means that just making the model more complex does not necessarily improve its results.

Table 7. Forecasting performance (RMSE) across nine tasks.

Model	N cons.	P_2O_5 cons.	K_2O cons.	N exp.	P_2O_5 exp.	K_2O exp.	N imp.	P_2O_5 imp.	K_2O imp.	Avg. Rank
ARIMA	0.92	1.05	1.11	0.88	1.06	1.15	0.94	1.02	1.07	3.7
RF	0.89	1	1.04	0.9	1.02	1.08	0.91	0.99	1.03	2.9
XGBoost	0.84⋆	0.95⋆	0.99⋆	0.89	0.96	1.01⋆	0.87⋆	0.95⋆	0.98⋆	1.6
ANN	0.93	1.07	1.1	0.92	1.05	1.14	0.95	1.01	1.06	4.1
LSTM	0.88	0.98	1.01	0.9	0.99	1.04	0.91	0.97	1.02	2.4
GRU	0.87	0.99	1	0.91	1	1.05	0.9	0.98	1.01	2.5
Hybrid-1	0.86	0.97	1	0.9	0.98	1.03	0.89	0.96	1	2

Lastly, using actual vs. predicted graphs and MAE heatmaps helps you understand the model in more detail than statistics alone. These figures not only confirm the numbers but also show when a model does not perform well which helps researchers and policymakers make better decisions.

Therefore, to forecast nutrients in agricultural systems efficiently, both a suitable model and knowledge of the data must be chosen. When it comes to practical use such as in policy planning and handling inventories, ARIMA and XGBoost models, paired with detailed visualization and error analysis, stand out the most.

What this study shows about forecasting outcomes is important for both domestic and international agricultural policies. For India, knowing how much nutrition is required and where nutrients are traded helps plan subsidies, production and the movement of food across the country. Since nitrogen use and export are set to rise, India could gain by increasing its domestic nitrogen production and preserving some export supplies. This contrasts with the increase in imports of potassium which shows that this could be a risky situation that needs international trade agreements, more suppliers or long-term stockpiling. These insights can guide policies that help a country be more independent while also taking part in world trade, manage government buying and respond to sudden market changes. Such forecasts can be added to global systems that monitor agriculture and help with international efforts to improve food security, under programs run by the FAO. The study joins the information from models with actual planning, helping to apply findings to real-life policies for agriculture.

The proposed framework can adjust to abrupt policy shifts and global shocks by including exogenous policy dummies (e.g. subsidy or tax reforms, import quotas), performing structural-break detection and refitting within regimes, generating scenario paths of macroeconomic drivers (foreign exchange rates, global fertilizer prices), and using state-space models that permit prompt re-centering following shocks. In practice, a monthly monitoring loop is advisable whereby breaks are identified, compact models are refitted and scenario bands (baseline, optimistic, pessimistic) are published to assist in procurement planning. On the policy front, the findings imply the maintenance of buffer stocks along with the upper 80% prediction band on potassium fertilizers, hedging against import exposures when the scenario bands widen by diversifying suppliers, tying subsidy allocation to the forecasted shortfalls of nutrients at the state level, and issuing a quarterly shock dashboard that reports on break tests and scenario deltas to inform stakeholders.

Policymakers may prioritize expanding domestic potassium processing capacity, adopt adaptive subsidy structures linked to forecasted deficits, and promote AI-based fertilizer distribution systems to improve efficiency under varying policy regimes. Future work can integrate explainable AI (e.g. SHAP or LIME) and include climatic or policy indices as features to enhance interpretability and address the framework's current limitations.

4. Conclusions

This study creates and assesses a complete forecasting system for projecting agricultural nutrient (N, P_2O_5, and K_2O) dynamics in India at three levels: consumption, exports, and imports. We tested nine alternative models, including time series interpolation (ARIMA), machine learning (Random Forest, SVM, XGBoost), deep learning (ANN, LSTM, GRU), and hybrid models (ARIMA–LSTM, XGBoost–LSTM), on historical data. The performance was evaluated using MAE, MSE, and RMSE. Our findings indicate that model performance was not uniform across forecasting tasks. ARIMA performed best for predicting trade (exports and imports) of nitrogen and potassium, where linear trends dominate. In contrast, machine learning models such as XGBoost and Random Forest achieved superior accuracy in predicting domestic consumption, particularly for nitrogen. Deep learning and hybrid models exhibited mixed results, with no consistent advantage. These task-specific differences underscore the importance of selecting forecasting models based on the characteristics of the data and prediction objectives. To maintain consistency with the abstract, we clarify that ARIMA showed superior performance in forecasting trade-related quantities (exports and imports) of nitrogen and potassium, whereas machine learning models such as XGBoost and Random Forest outperformed ARIMA in predicting domestic agricultural consumption, particularly for nitrogen. The purpose of this study was to determine that hybrid deep learning models would perform better than traditional and standalone machine learning models when forecasting nutrient changes in agriculture. Yet, we can see that although advanced models like XGBoost and Random

Forest work well for figuring out consumption trends, hybrid models like ARIMA–LSTM and XGBoost–LSTM do not improve predictions much, mainly for phosphorus and potassium. It appears that simple yet properly adjusted models can serve better and be more comprehensible for some agricultural forecasting tasks. That means (1) Random Forest and XGBoost can be used by policymakers and planners to reliably predict the need for fertilizer in the near term and (2) complex hybrid models may not be the best choice when resources for computing or making predictions clear are scarce. It provides a strong way to compare models and a set of proven forecasting tools that aid in making data-driven decisions for fertilizer imports, trade policies and managing nutrients sustainably for the future of Indian food security. Predictions for the 6-year period (2024–2029) show that nitrogen use will rise (65,027 tons to 69,845 tons), phosphorus usage will remain unchanged (29,006 tons to 30,211 tons), and potassium demand will rise (20,807 tons to 24,301 tons). Our findings show that model-specific advantages exist for prediction situations; however, hybrid models offer Each model exhibited strengths based on the type of nutrient and the forecast category (consumption, export, or import). This study has implications for agricultural planning, policymaking, and food security in India.

These results give useful advice to policymakers in India who work on agriculture and trade. Accurate forecasts help plan for fertilizer procurement, oversee inventory and create strategies for importing and exporting, as India works to become more self-reliant in global markets.

Still, the research wasn't without problems such as not having enough data in the beginning, difficulties in handling external changes (like price changes and new policies) and some challenges in understanding the results of deep learning models. Researchers could investigate explainable AI, add extra variables (such as rainfall or subsidies) or expand the work to other developing countries.

Model interpretability can be strengthened by applying explainable AI methods such as SHAP values for tree- and boosting-based models to rank influential drivers (e.g. minimum support price changes, rainfall anomalies, and currency fluctuations) and using integrated gradients or perturbation analyses for deep learning models to highlight feature contributions. In parallel, extending the covariate set to include rainfall and monsoon indices, irrigated and cropped area, fertilizer retail prices, subsidy intensity, INR/USD exchange rates, crop mix shares, and global fertilizer indices is expected to reduce residual variance and improve the framework's responsiveness to shocks. A curated feature catalogue and a structured variable-selection pipeline are planned for release to facilitate transparency and reproducibility in future applications.

Acknowledgments

The authors extend their gratitude to the Food and Agriculture Organization (FAO) and the Fertilizer Association of India (FAI) for providing access to historical nutrient data. We acknowledge the computational support provided by our institution. –

Author contributions statement

CRediT: **Pradeep Mishra**: Conceptualization, Data curation, Methodology; **Diaa Salman**: Formal analysis, Software, Visualization; **Binita Kumari**: Validation, Writing – original draft; **Abdullah Mohammad Ghazi Al Khatib**: Funding acquisition, Supervision, Writing – review & editing; **Bayan Mohamad Alshaib**: Project administration, Resources. All authors have read and approved the final version of the manuscript.

Disclosure statement

The authors declare no competing financial interests or personal relationships that could influence the work reported in this paper.

About the authors

Pradeep Mishra, PhD, is an Assistant Professor of Statistics at the College of Agriculture, Rewa at the Jawaharlal Nehru Krishi Vishwa Vidyalaya, Madhya Pradesh, India. During his PhD research work, he was selected for an INSPIRE Fellowship by the Department of Science and Technology of the Government of India. Dr. Mishra has served as a data management specialist at a private multinational company for several years. His area of specialization is agriculture statistics, time series, and design of experiment. He has published more than 155 research articles in different international and national journals.

Assist. Prof. Dr. Diaa Salman, received his Ph.D. in Electrical and Electronic Engineering from Cyprus International University, Cyprus. He obtained his Master's Degree in Electrical and Electronic Engineering from Eastern Mediterranean University, Cyprus, and his Bachelor's Degree in Electrical Engineering from An-Najah National University, Palestine. He is currently an Assistant Professor in the Department of Electrical Engineering – Dual Studies at Al-Quds University, Palestine.

Dr. Binita Kumari is an agricultural economist with specialisation in production economics. She completed her graduation from Acharya NG Ranga Agricultural University, Andhra Pradesh, and master's and doctoral degree from from ICAR- National Dairy Research Institute. She has always been a meritorious student receiving several silver as well as gold medals at both college and university level. She has been working as assistant professor (Agricultural Economics) in Rashtriya Kisan (PG) College, Shamli, which is affiliated to Maa Shakumbhari University, Saharanpur, Uttar Pradesh since 2020. She has more than 5 years experience in teaching and has published around 60 research publications.

Dr. Abdullah Mohammad Ghazi Al khatib, Department of Banking and Financial Institutions, , Faculty of Administrative Sciences, Al-Sham Private University, Damascus, Syrian Arab Republic. Abdullah does research in Financial Economics, Econometrics and Development Economics. He serves as an Academic Editor for PLOS ONE. He has worked as a reviewer for different leading journals like Finance Research Letters, International Review of Financial Analysis, and Journal of Analysis and Testing. He has 181 Verified Peer Reviews in Web of Science. With over 50 publications, he is an active contributor to the field.

Dr. Bayan Mohamad Alshaib, Department of Finance and Banking, Faculty of Economics, Damascus University, Syrian Arab Republic. Her work focuses on financial economics, fiscal policy, economic growth, and panel data econometrics. She has experience reviewing manuscripts for prestigious journals like Finance Research Letters, International Review of Financial Analysis and PLOSONE and has published extensively in peer-reviewed outlets. She is Member of International Association of Scientific Researcher.

Ethical considerations

Ethical approval was not required for this study, as it did not involve human or animal subjects. All the data used were aggregated and anonymized, adhering to the ethical standards for secondary data analysis.

Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

ORCID

Pradeep Mishra http://orcid.org/0000-0003-4430-886X
Diaa Salman http://orcid.org/0000-0001-7534-8229
Binita Kumari http://orcid.org/0000-0003-1224-3037
Abdullah Mohammad Ghazi Al khatib http://orcid.org/0000-0002-1352-2348
Bayan Mohamad Alshaib http://orcid.org/0000-0002-9876-998X

Data availability statement

The datasets used in this study are publicly available from the Food and Agriculture Organization (FAO) and Fertilizer Association of India (FAI). The processed data and codes are available upon reasonable request from the corresponding author.

References

Arvind, K. S., Sanjib, K. B., Chaudhari, S. K., & Gajendra, S. (2022). Fertilizer use in Indian agriculture and its impact on human health and environment. *Indian Journal of Fertilisers*, *18*(3), 218–237.

Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., Kimball, B. A., Ottman, M. J., Wall, G. W., White, J. W., Reynolds, M. P., Alderman, P. D., Prasad, P. V. V., Aggarwal, P. K., Anothai, J., Basso, B., Biernath, C., Challinor, A. J., De Sanctis, G., … Zhu, Y. (2015). Rising temperatures reduce global wheat production. *Nature Climate Change*, *5*(2), 143–147. https://doi.org/10.1038/nclimate2470

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

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chand, R., & Pandey, L. M. (2009). Fertiliser use, nutrient imbalances, and subsidies: Trends and implications. *Margin: Journal of Applied Economic Research*, *3*(4), 409–432. https://doi.org/10.1177/097380100900300404

Chand, R., & Pavithra, S. (2015). Fertiliser use and imbalance in India: Analysis of states. *Economic & Political Weekly*, *50*(38), 98–104.

Chary, D. S., Supriya, K., Anjaiah, T., Meena, A., & Ramesh, M. (2023). Trends in fertilizer nutrients (N, P_2O_5, and K_2O) consumption in India. *Biological Forum – An International Journal*, *15*(6), 690–696.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).

Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder–decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1023/A:1022627411411

Dawson, C. J., & Hilton, J. (2011). Fertiliser availability in a resource-limited world: Production and recycling of nitrogen and phosphorus. *Food Policy*, *36*(1), S14–S22. https://doi.org/10.1016/j.foodpol.2010.11.012

Gao, L., & McCallister, D. (2024). Fertilizing the future: A comprehensive forecast and analysis of the US nitrogen fertilizer demand.

Gao, Y., Dong, K., & Yue, Y. (2024). Projecting global fertilizer consumption under shared socioeconomic pathway (SSP) scenarios using an approach of ensemble machine learning. *Science of the Total Environment*, *912*, 169130. https://doi.org/10.1016/j.scitotenv.2023.169130

Hang, Y., Zhou, Z., Van Griensven Thé, J., Yang, S. X., & Gharabaghi, B. (2023). Flood forecasting using hybrid LSTM and GRU models with lag time preprocessing. *Water*, *15*(22), 3982. https://doi.org/10.3390/w15223982

Haykin, S. (2009). *Neural networks and learning machines*. Pearson Education.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kurniawan, J. D., Parhusip, H. A., & Trihandaru, S. (2024). Predictive performance evaluation of ARIMA and hybrid ARIMA–LSTM models for particulate matter concentration. *Journal Online Informatika*, *9*(2), 259–268. https://doi.org/10.15575/join.v9i2.1318

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., & Garnier, J. (2014). 50-year trends in nitrogen use efficiency of world cropping systems: The relationship between yield and nitrogen input to cropland. *Environmental Research Letters*, *9*(10), 105011. https://doi.org/10.1088/1748-9326/9/10/105011

Leghari, S. J., Wahocho, N. A., Laghari, G. M., HafeezLaghari, A., MustafaBhabhan, G., HussainTalpur, K., & Lashari, A. A. (2016). Role of nitrogen for plant growth and development: A review. *Advances in Environmental Biology*, *10*(9), 209–219.

Malhotra, H., Vandana, Sharma, S., & Pandey, R. (2018). Phosphorus nutrition: Plant growth in response to deficiency and excess. In *Plant nutrients and abiotic stress tolerance* (pp. 171-190). Singapore: Springer Singapore.

Neset, T. S. S., & Cordell, D. (2012). Global phosphorus scarcity: Identifying synergies for a sustainable future. *Journal of the Science of Food & Agriculture*, *92*(1), 2–6. https://doi.org/10.1002/jsfa.4650

Padhiary, M., Hoque, A., Prasad, G., Kumar, K., & Sahu, B. (2025). Precision agriculture and AI-driven resource optimization for sustainable land and resource management. In J. Ruiz-Vanoye & O. Díaz-Parra (Eds.), *Smart water technology for sustainable management in modern cities* (pp. 197–232). IGI Global Scientific Publishing; https://doi.org/10.4018/979-8-3693-8074-1.ch009

Pathak, H., Mohanty, S., Jain, N., & Bhatia, A. (2010). Nitrogen, phosphorus, and potassium budgets in Indian agriculture. *Nutrient Cycling in Agroecosystems*, *86*(3), 287–299. https://doi.org/10.1007/s10705-009-9292-5

Pawlak, K., & Kołodziejczak, M. (2020). The role of agriculture in ensuring food security in developing countries: Considerations in the context of the problem of sustainable food production. *Sustainability*, *12*(13), 5488. https://doi.org/10.3390/su12135488

Pisuttinusart, C., Jatuporn, C., Suvanvihok, V., & Seerasarn, N. (2022). Forecasting the import demand for chemical fertilizer in Thailand. *EUrASEANs: Journal on Global Socio-Economic Dynamics*, *3*(34), 61–70. https://doi.org/10.35678/2539-5645.3(34).2022.61-70

Prajapati, K., & Modi, H. A. (2012). The importance of potassium in plant growth – A review. *Indian Journal of Plant Sciences*, *1*(02–03), 177–186.

Ritchie, H., Reay, D., & Higgins, P. (2018). Sustainable food security in India—Domestic production and macronutrient availability. *PloS One*, *13*(3), e0193766. https://doi.org/10.1371/journal.pone.0193766

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2018). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *IEEE Transactions on Neural Networks & Learning Systems*, *29*(10), 4691–4706.

Sujan, G., & Ananta, D. (2021). Effects of agricultural inputs use on agricultural productivity in Nepal: An empirical study. *International Journal of Agricultural Economics*, *6*(5), 212–217.

Tang, Q., Cotton, A., Wei, Z., Xia, Y., Daniell, T., & Yan, X. (2022). How does partial substitution of chemical fertiliser with organic forms increase sustainability of agricultural production? *Science of the Total Environment*, *803*, 149933. https://doi.org/10.1016/j.scitotenv.2021.149933

Tenkorang, F., & Lowenberg-DeBoer, J. (2009). Forecasting long-term global fertilizer demand. *Nutrient Cycling in Agroecosystems*, *83*(3), 233–247. https://doi.org/10.1007/s10705-008-9214-y

The Fertilizer Institute. (n.d). *Sustainability*. Retrieved February 9, 2024, from https://www.tfi.org/sustainability

Wang, W., Ma, B., Guo, X., Chen, Y., & Xu, Y. (2024). A hybrid ARIMA–LSTM model for short-term vehicle speed prediction. *Energies*, *17*(15), 3736. https://doi.org/10.3390/en17153736

Wu, Y., Li, Y., Fu, X., Shen, J., Chen, D., Wang, Y., Liu, X., Xiao, R., Wei, W., & Wu, J. (2018). Effect of controlled-release fertilizer on N_2O emissions and tea yield from a tea field in subtropical central China. *Environmental Science & Pollution Research International*, *25*(25), 25580–25590. https://doi.org/10.1007/s11356-018-2646-2

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0

Appendix

Table A1. Data-quality audit; illustrative values.

Nutrient	Category	Span	Missing (%)	Outliers (*Z*)	Outliers (IQR)
N	Consumption	1960–2023	1.6	1	2
N	Exports	1960–2023	4.7	2	1
N	Imports	1960–2023	3.1	2	2
P_2O_5	Consumption	1960–2023	3.1	1	1
P_2O_5	Exports	1960–2023	7.8	3	2
P_2O_5	Imports	1960–2023	4.7	2	2
K_2O	Consumption	1960–2023	6.3	2	1
K_2O	Exports	1960–2023	9.4	3	2
K_2O	Imports	1960–2023	4.7	2	1