Ciência Rural, Santa Maria, v.56:4, e20250132, 2026
http://doi.org/10.1590/0103-8478cr20250132
ISSNe 1678-4596
AGRIBUSINESS

Predicting wheat producer price index using graph neural networks and ensemble learning

Cevher Özden[1*]

[1]Faculty of Arts and Sciences, Department of Computer Sciences, Cukurova University, 01250, Adana, Türkiye. E-mail: cozden@cu.edu.tr.
[*]Corresponding author.

ABSTRACT: Accurate forecasting of the wheat producer price index is crucial for economic planning, trade policy, and food security management. Traditional econometric and machine learning approaches often struggle to capture the complex interplay of economic, climatic, and trade-related drivers of price dynamics. This study proposed a novel hybrid framework that integrates graph neural networks, convolutional neural networks, and random forest to enhance predictive performance. The model combines economic indicators, meteorological variables, governance measures, and graph-based trade representations to capture both spatial and temporal dependencies. Data were obtained from the Food and Agriculture Organization for economic and trade indicators across 109 countries, from the Turkish State Meteorological Service for national weather parameters, and from the World Bank's Worldwide Governance Indicators, specifically the Political Stability and Absence of Violence/Terrorism index, to reflect geopolitical risk. The dataset spans 2000-2023 and was randomly split into 80% training and 20% evaluation subsets.We benchmarked the proposed approach against baseline models. Results demonstrated that the GNN+CNN+RF hybrid, enriched with the Political Stability Index, achieved the best performance, reducing Mean Squared Error to 158.44 and Root Mean Squared Error to 12.59, while attaining the highest R^2 score of 0.962. These findings highlighted the advantages of integrating graph-based learning with deep learning and ensemble methods, further strengthened by governance-related risk indicators.
Key words: producer price index, graph neural networks, forecasting, machine learning.

Previsão do índice de preços ao produtor de trigo usando redes neurais e aprendizagem de conjuntos

RESUMO: A previsão exata do índice de preços ao produtor (PPI) do trigo é crucial para o planejamento econômico, a formulação de políticas comerciais e a gestão da segurança alimentar. Abordagens tradicionais de econometria e de aprendizado de máquina frequentemente têm dificuldades em capturar a complexa interação entre fatores econômicos, climáticos e comerciais que influenciam a dinâmica dos preços. Este estudo propõe uma estrutura híbrida inovadora que integra redes neurais (GNNs), redes neurais Convolucionais (CNNs) e algoritmos robustos (RF) para aprimorar o desempenho preditivo. O modelo combina indicadores econômicos, variáveis meteorológicas, medidas de governança e representações comerciais baseadas em grafos para capturar dependências espaciais e temporais. Os dados foram obtidos da Organização das Nações Unidas para Alimentação e Agricultura (FAO) para indicadores econômicos e comerciais de 109 países; os dados do Serviço Meteorológico do Estado da Turquia foram obtidos para parâmetros climáticos nacionais e os dados dos Indicadores de Governança Mundial (WGI) do Banco Mundial, especificamente o índice de Estabilidade Política e Ausência de Violência/Terrorismo, foram adquiridos para refletir o risco geopolítico. O conjunto de dados abrange o período de 2000–2023 e foi dividido aleatoriamente em 80% para treinamento e 20% para avaliação. Comparamos a abordagem proposta com modelos de referência. Os resultados demonstram que o híbrido GNN+CNN+RF, enriquecido com o Índice de Estabilidade Política, obteve o melhor desempenho, reduzindo o Erro Quadrático Médio (MSE) para 158,44 e o Erro Quadrático Médio da Raiz (RMSE) para 12,59, além de alcançar o maior R^2 de 0,962. Esses achados destacam as vantagens de integrar aprendizado baseado em grafos com aprendizado profundo e métodos de *ensemble*, fortalecidos por indicadores de risco relacionados à governança.
Palavras-chave: indice de preços ao produtor, redes neurais gráficas, previsão, aprendizado de máquina.

INTRODUCTION

Accurate forecasting of the Producer Price Index (PPI) for wheat is essential for economic planning, trade policy formulation, and ensuring food security. The Producer Price Index reflects price trends at the production level, making it a key metric for governments, businesses, and agricultural stakeholders. Traditional forecasting models based on econometrics or classical machine learning struggle to capture the intricate relationships among economic, climatic, and geopolitical factors affecting wheat prices.

Wheat prices are influenced by a variety of economic, trade, and meteorological factors. International trade policies, supply-demand dynamics, and climate variations introduce complexity into price forecasting. Existing machine learning methods, while effective to some extent, often lack the ability to model relationships between

Received 03.05.25 Approved 09.16.25 Returned by the author 01.02.26
CR-2025-0132.R4
Editors: Alessandro Dal'Col Lúcio Daniel Arruda Coronel

different countries and their respective trade networks. Furthermore, meteorological factors such as temperature, precipitation, and humidity play a crucial role in agricultural production, making them vital components of any predictive model. Various data and models have been used to forecast product prices in literature. COLINO et al. (2010) evaluate the accuracy of outlook price forecasts for hog markets, comparing them to ARIMA, VAR, and Bayesian VAR models. Futures-based forecasts perform best in the short term, while VAR and BVAR models improve accuracy at longer horizons. The study finds that combining market-based and statistical methods reduces forecasting errors, making composite forecasts more reliable than standalone approaches. GARCÍA-GERMÁN et al. (2015) analyzed how global agricultural price changes affect consumer food prices in the EU using Engle-Granger error correction models. They find that price transmission varies across EU countries, influenced by market structures, exchange rates, and economic factors. The study highlighted that food price shocks disproportionately impact low-income households, calling for targeted policies. TANG et al. (2019) proposed a hybrid model integrating Genetic Algorithm-optimized Support Vector Regression (GA-SVR) with the Autoregressive Integrated Moving Average (ARIMA) model for predicting the Producer Price Index (PPI). Their approach used fuzzy information granulation to preprocess data, generating different sequences for training individual SVR models optimized via Genetic Algorithm. The residuals of the SVR model were then corrected using ARIMA. Their results demonstrated that this hybrid model outperformed standalone ARIMA, General Regression Neural Networks (GRNN), and GA-SVR models in PPI prediction. However, their model was primarily focused on monthly data and did not consider external economic variables such as international trade, meteorological conditions, or cross-country influences, which could improve prediction accuracy. SHESTAKOV & AZHLUNI (2020) analyzed the volatility of the Producer Price Index (PPI) in agriculture using GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models to forecast price fluctuations in crop and livestock sectors. They conducted SARIMAX-based (Seasonal Autoregressive Integrated Moving Average with Exogenous Variables) explanatory analysis and developed volatility forecasts for 2020. Their findings highlighted that crop price volatility is significantly higher than livestock prices, influenced by seasonal and macroeconomic factors. However, the study was limited by uncertainty in global economic conditions, which made long-term forecasting challenging. KHAMIS & XIN (2020) compared ARIMA and Grey Model GM(1,1) for forecasting the Malaysia Producer Price Index (PPI) using monthly data from 2013 to 2019. Their study found that ARIMA (1,2,0) outperformed GM (1,1) based on Mean Absolute Percentage Error (MAPE), making it the preferred model for predicting future PPI trends. The results indicated potential deflationary pressure in Malaysia for the following year. However, the study focused only on univariate time series models, without incorporating economic or external factors, limiting its robustness for complex market dynamics. RAIKAMO (2022) investigated the short-term forecasting of the Finnish Producer Price Index (PPI) using high-dimensional data. The study utilized dynamic factor models and penalized regressions, including Ridge, Lasso, Elastic Net, and Adaptive Lasso. A real-time out-of-sample forecasting experiment was conducted using monthly economic time series data. The results showed that high-dimensional models provided only marginal improvements over univariate benchmarks, with penalized regression methods performing particularly well for nowcasting. However, the study highlighted challenges such as predictor selection sensitivity and the limited impact of external variables on long-term forecasts. YADAV et al. (2022) applied LSTM, BiLSTM, and GRU models to forecast the Producer Price Index (PPI) of the cheese manufacturing industry using monthly data from 1972 to 2021. Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to capture long-term dependencies in sequential data. They are particularly effective for time series forecasting due to their ability to mitigate the vanishing gradient problem and retain relevant historical information. On the other hand, Bidirectional Long Short-Term Memory (BiLSTM) networks are an extension of LSTM models, where two LSTM layers process input sequences in both forward and backward directions. This architecture provides richer context and improved predictive accuracy for sequence-based tasks. The study compared these deep learning models based on metrics such as RMSE, MAE, and MAPE, finding that LSTM performed best in training, while BiLSTM showed superior accuracy in testing. However, external economic and climatic variables were not incorporated, limiting the model's generalizability beyond historical price trends. WANG et al. (2023) analyzed the Producer Price Index (PPI) of agricultural products in Heilongjiang

Province, China, using quarterly data from 2003 to 2021. The study applied descriptive statistics and ARIMA models to forecast short-term price fluctuations for 2022. Their results showed that the PPI remains highly volatile, influenced by seasonality, supply-demand dynamics, and external economic conditions. The authors recommended policy interventions to stabilize agricultural product prices. However, the study focused only on a single province, limiting its applicability to broader global trade and economic factors. TUNÇSIPER (2023) examined the impact of macroeconomic factors such as oil prices, exchange rates, interest rates, and wages on Türkiye's Producer Price Index (PPI) using Artificial Neural Networks (ANN) and ARIMA (Autoregressive Integrated Moving Average) models. The study covered the period 2002Q1-2022Q3 and aimed to compare the predictive performance of linear and nonlinear methods. The findings suggested that the ARMA (Autoregressive Moving Average) model outperformed the ANN (Artificial Neural Network) in terms of forecasting accuracy. However, the study relied on traditional econometric approaches and did not explore graph-based methods or ensemble learning techniques, which may capture hidden dependencies in economic data more effectively. AVINASH et al. (2024) developed a Hidden Markov Model (HMM)-guided deep learning approach to forecast volatile agricultural commodity prices. The study combined HMMs (Hidden Markov Models) with deep learning models such as RNNs (Recurrent Neural Networks), CNNs (Convolutional Neural Networks), LSTMs (Long-short Term Memory), and GRUs (Gated Recurrent Unit), integrating technical indicators like Moving Averages (MA), Bollinger Bands (BB), and Fourier Transform features to capture price trends and volatility. The Champadanga potato price dataset (India) was used to evaluate model performance. The HMM-enhanced GRU model achieved the best results, surpassing traditional deep learning models. However, the study primarily focused on single-country data and lacked an assessment of cross-country economic factors in price forecasting. GUINDANI et al. (2024) conducted a systematic review of forecasting methods for agricultural commodities using Latent Dirichlet Allocation (LDA) and bibliometric analysis. The study categorized forecasting approaches into machine learning (ML), ML-neural networks (ML-NN), ML-ensemble, ML-hybrid, and statistical methods. It found that ML-hybrid approaches (41.95%) and statistical methods (29.31%) were the most used. The authors identified key research gaps, such as limited studies on financial performance forecasting in agribusiness and the need for more structured hybrid forecasting models. However, the study relied primarily on bibliometric analysis, without empirical validation of the methods reviewed. ZHU et al. (2024) introduce a hybrid neural network model integrating CEEMDAN, SG filter, CNN, and LSTM for stock price forecasting. Using CSI 300 index data, their model outperforms benchmarks, reducing MAE by 45.33% and RMSE by 43.44%. The study highlighted the benefits of combining frequency decomposition, data smoothing, and deep learning for improved predictions. LI et al. (2024) introduce DecEnsInt, a decomposition-ensemble-integration framework for carbon price forecasting. Using CEEMDAN for data decomposition and SLSQP for optimized integration, the model improves forecasting accuracy by 10.58%-17.72% over existing methods. The approach effectively handles price volatility but does not incorporate macroeconomic factors. TESTE et al. (2024) proposed a satellite-based approach for early corn yield and price forecasting, using Gross Primary Production (GPP) data and EOF analysis. Their model outperforms WASDE forecasts, achieving AUC up to 0.97 for yield and 0.91 for price predictions, highlighting the potential of remote sensing for market forecasting. Recent studies in other domains have demonstrated that hybrid learning strategies, often combining meta-heuristic optimization, advanced feature selection, and diverse model archixtectures, can enhance predictive accuracy and robustness in complex forecasting tasks. Examples include applications of particle swarm and Al-Biruni Earth radius optimization, hybrid sine cosine with novel optimization methods, and binary Harris Hawks or Slime Mould algorithms for feature selection (HADJOUNI et al., 2023; ATTEIA et al., 2023; EL-KENAWY et al., 2022; ABDELHAMID et al., 2025; ALKANHEL et al., 2023). While these studies differ in scope and application area from the present study, they share the underlying principle of integrating complementary learning paradigms, which also underpins the GNN+CNN+RF framework proposed here.

Recent advancements in metaheuristic algorithms and hybrid approaches have shown significant potential for forecasting models in agriculture and environmental sciences. For instance, EL-KENAWY et al. (2024) introduced the Greylag Goose Optimization (GGO) algorithm, demonstrating its superior performance compared to conventional methods in engineering optimization problems. The study highlighted GGO's effectiveness, particularly in feature selection and high-dimensional dataset analysis. GABER & SİNGLA (2025) conducted a comparative analysis of different regression models for groundwater resource forecasting in India, revealing that Random Forest Regression outperformed other models in terms of MSE and R^2 metrics. This ensemble approach proved more robust in capturing complex interrelationships within the dataset, providing a comprehensive perspective on groundwater dynamics. INGOLE et al. (2025) proposed a comprehensive multi-modal deep learning model for soybean yield prediction, integrating CNNs, RNNs, and GNNs. Their approach addresses the complexity of yield drivers by fusing data from multiple sources: satellite remote sensing imagery (for spatial crop health information), time-series weather data (for temporal climate patterns), and soil property data (for static fertility indicators). In the model's first stage, a CNN processes the satellite imagery to extract spatial features, while an RNN (LSTM or TCN) processes sequential weather data, and these feature streams are combined to predict yields. This multi-source fusion yielded substantial improvements over single-source models – reducing prediction RMSE by ~15% and improving R^2 by ~20%. MOHAMED (2025) provided a comprehensive review of waste management techniques for sustainable energy production, emphasizing the importance of optimization methods such as genetic algorithms and particle swarm optimization. These methods have been instrumental in addressing technical and environmental challenges in waste-to-energy (WTE) systems, highlighting the value of computational tools for enhancing system performance and sustainability. ANG et al. (2022) proposed a novel deep learning method for optimizing convolutional neural network (CNN) architectures using the Teaching–Learning-Based Optimization (TLBO) algorithm. This approach demonstrated improved classification accuracy and reduced complexity in CNN models, particularly in image classification tasks with symmetric data distributions. Finally, ALEISA et al. (2022) integrated transfer learning with the Dipper Throated Optimization (DTO) algorithm to develop an effective system for pneumonia detection from chest X-rays. By combining GoogLeNet, ResNet18, and DenseNet121 models with DTO-optimized feature selection and KNN classification, the system achieved high diagnostic accuracy and robust performance across different test scenarios.

Graph-based learning approaches, particularly Graph Neural Networks (GNNs), have emerged as powerful tools for modeling structured relationships between entities. Unlike traditional models, GNNs can capture the spatial and relational dependencies among countries, trade networks, and climate similarities. In this study, we integrated GNNs with ensemble learning techniques (Random Forest and XGBoost) to improve wheat PPI predictions. Our methodology combines:

Economic indicators such as export/import shares, which capture the trade dependencies between countries.

Meteorological factors including relative humidity, average temperature, and yearly precipitation, which influence wheat yield and availability.

Graph-based topological features, representing relationships among different countries through trade flows and climate patterns.

The main contribution of this study lies in the novel integration of Graph Neural Networks (GNNs), Convolutional Neural Networks (CNNs), and Random Forest (RF) into a unified hybrid model for agricultural price forecasting. Unlike existing methods that typically focus on univariate time series or isolated machine learning techniques, our approach captures the complex interdependencies of trade networks (via GNNs), temporal price trends (via CNNs), and non-linear feature interactions (via RF). This comprehensive combination significantly enhanced forecasting accuracy and robustness, addressing limitations of traditional econometric and standalone machine learning models. The proposed graph-based model encodes interactions between various countries and incorporates economic and meteorological factors. The GNN extracts meaningful representations, which are then used as inputs to an ensemble model for forecasting the Producer Price Index. The results demonstrated that our hybrid approach significantly enhances predictive accuracy compared to standalone models.

MATERIALS AND METHODS

Data

The dataset used in this study covered the period from 2000 to 2023, combining information from two primary sources. The first dataset was obtained from the Food and Agriculture Organization (FAO) and contained historical Producer Price Index (PPI) values across multiple countries, as well as economic indicators such as export share, import share, and trade dependencies. The second dataset was sourced from the Turkish State Meteorological

Service, including meteorological variables such as relative humidity, average temperature, and annual precipitation, which are essential for understanding wheat yield variations. Additionally, governance-related indicators were incorporated from the World Bank's Worldwide Governance Indicators (WGI), specifically the "Political Stability and Absence of Violence/Terrorism" dimension. This annual index, available from 2000 to 2023, serves as a proxy for geopolitical risks and complements the economic and meteorological variables. Country-level WGI scores were aligned with FAO trade and PPI data, enriching the model with features sensitive to external shocks such as wars, sanctions, and pandemics. These datasets were integrated to form a comprehensive framework for forecasting the wheat PPI (Table 1).

To ensure data quality and consistency, missing values were handled through interpolation techniques. Feature scaling was applied using standardization (z-score normalization), ensuring that variables contributed equally during model training. Temporal dependencies were captured by generating lag features from past PPI values for Turkey, specifically using one-year and three-year lags. These lag features enabled the model to account for historical price patterns and dependencies.

Methods

A key step in our methodology involves the grouping of countries based on economic and meteorological similarities. We employed hierarchical clustering and K-means clustering to categorize countries into distinct groups. This clustering approach helped in identifying regions that exhibit similar trade and climate characteristics, which is crucial for understanding the broader impact of global trade on wheat prices. For the clustering analysis, economic and climatic features were standardized using z-score normalization to ensure comparability. The clustering analysis employed K-Means algorithm applied to standardized economic and climatic data, with dimensionality reduced via PCA to facilitate visualization. Features used for clustering included import quantity, export value, production, yield, area harvested, producer price, and meteorological variables such as precipitation and temperature.

Table 1 - Descriptive table of the data.

Feature	Type	------------------Description------------------	Scaling Applied	Missing Value Handling
Country	Categorical	Country name	No Scaling	Not Applicable
Year	Categorical	Year of observation	No Scaling	Not Applicable
GDP	Numerical	Gross Domestic Product of the country	Standardized	Interpolation
Import quantity (t)	Numerical	Wheat import volume in tons	Standardized	Interpolation
Import Value ($)	Numerical	Value of wheat imports in USD	Standardized	Interpolation
Wheat Production (t)	Numerical	Total wheat production volume	Standardized	Interpolation
Export price (1000$)	Numerical	Export price per unit	Standardized	Interpolation
Export Quantity (t)	Numerical	Wheat export volume in tons	Standardized	Interpolation
Producer Price (USD/tonne)	Numerical	Producer Price Index for wheat	Standardized	Interpolation
Political Stability Index	Numerical	World Bank WGI – Political Stability and Absence of Violence/Terrorism (proxy for geopolitical risk)	Standardized	Not Applicable / Country-level gaps acknowledged
Relative humidity (%)	Numerical	Average annual relative humidity	Standardized	Interpolation
Average temperature (ºC)	Numerical	Average annual temperature	Standardized	Interpolation
Annual precipitation (mm)	Numerical	Annual precipitation	Standardized	Interpolation
Predicted Producer Price	Numerical	Forecasted PPI using the hybrid model	Standardized	Not Applicable
Residuals (Actual - Predicted)	Numerical	Difference between actual and predicted PPI	Standardized	Not Applicable

Prior to clustering, features were normalized using standardization (z-score). The resulting clusters were visualized using PCA-reduced coordinates, and a table listing countries by cluster assignment was generated to provide a comprehensive view of the grouping. All analysis results and codes are given in a public github repository <https://github.com/cevher/gnnprice>. Furthermore, a trade-climate similarity graph was constructed, where nodes represented countries and edges indicated the strength of trade relations and climatic correlations.

Baseline machine learning models

The predictive modeling phase was conducted in several stages. Initially, Random Forest (RF) and XGBoost (XGB) were used as benchmark models to establish a performance baseline using only economic and meteorological data. Random Forest (RF) is an ensemble learning method that builds multiple decision trees during training and outputs the average prediction of the individual trees to improve accuracy and control overfitting. It is particularly effective for regression tasks such as forecasting agricultural prices. XGBoost (Extreme Gradient Boosting) is an advanced ensemble learning algorithm based on gradient boosting. It constructs decision trees sequentially, where each subsequent tree corrects the errors of the previous ones. Known for its scalability and high predictive performance, XGBoost is widely used for regression tasks, including agricultural price forecasting. To capture spatial and temporal dependencies, Convolutional Neural Networks (CNNs) were employed, leveraging historical PPI trends for improved predictive accuracy. Convolutional Neural Networks (CNNs) are a class of deep learning models designed to capture spatial patterns in data by applying convolutional operations. Originally developed for image analysis, CNNs have also been effectively applied to time series forecasting by capturing local dependencies and temporal trends in sequential data. In parallel, Graph Neural Networks (GNNs) were implemented to extract trade and climatic representations from the country-wise similarity graph. Graph Neural Networks (GNNs) are a class of deep learning models designed to capture relationships and dependencies within graph-structured data. By propagating information along edges, GNNs can model complex interactions, such as trade connections and spatial dependencies in agricultural price forecasting. By incorporating graph-based learning, we aimed to model interactions between different economies and their mutual influence on wheat prices.

Hybrid models

To further enhance predictive performance, hybrid modeling approaches were introduced. The first hybrid approach combined GNN representations with Random Forest (GNN+RF), utilizing structured relational data for enhanced learning. A more advanced approach involved integrating CNN-based feature extraction with GNN-based trade embeddings and a Random Forest predictor (GNN+CNN+RF) (Figure 1). This final hybrid model demonstrated the ability to capture complex interactions among global trade, meteorological conditions, and price trends more effectively than standalone models. The core of the predictive approach involved the integration

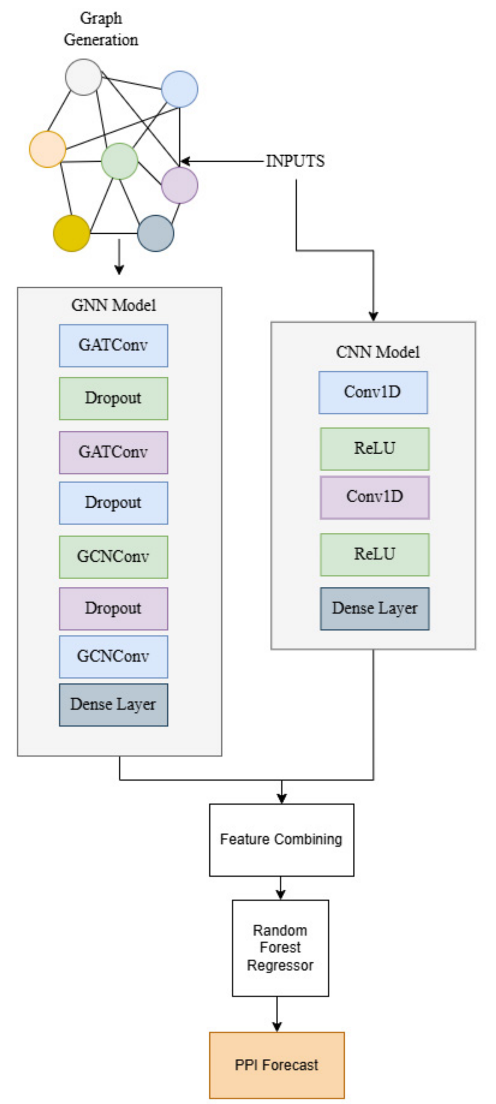

Figure 1 - Proposed model.

of Graph Neural Networks (GNNs), Convolutional Neural Networks (CNNs), and Random Forest (RF) into a hybrid model. The GNN component was implemented using two Graph Convolutional Network (GCNConv) layers with hidden dimensions of 128 and 64, respectively. The activation function used was ReLU, with a dropout rate of 0.3 to prevent overfitting. The trade-climate similarity graph was constructed with countries as nodes and edge weights derived from normalized trade volumes and climatic correlation coefficients. The CNN component comprised two one-dimensional convolutional layers with a kernel size of 3, stride of 1, and same padding, followed by a ReLU activation and a dropout rate of 0.3. This configuration was designed to capture temporal patterns from sequences of lagged PPI data spanning 12 months. The outputs from the GNN and CNN were concatenated and passed to the ensemble layer. In the ensemble layer, Random Forest and XGBoost models were employed to perform the final prediction. For Random Forest, the optimal hyperparameters included 300 estimators and a maximum tree depth of 15. For XGBoost, the hyperparameters were set to 500 estimators, a maximum depth of 10, and a learning rate of 0.1. Hyperparameter optimization was conducted using Bayesian optimization techniques, exploring search spaces that balanced model complexity and performance.

Model training and evaluation metrics

Model training was conducted using 5-fold cross-validation to ensure robustness. Hyperparameter tuning was performed through Bayesian optimization, allowing for efficient selection of optimal model configurations. Feature importance was assessed using SHAP (SHapley Additive exPlanations), providing insights into the key drivers of price fluctuations. Model evaluation was carried out using multiple performance metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 score.

Mean Absolute Error (MAE): Measures the absolute deviation between predicted and actual values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widetilde{y}_i|$$

Mean Squared Error (MSE): Computes the squared deviations to emphasize larger errors.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widetilde{y}_i)^2$$

Root Mean Squared Error (RMSE): Takes the square root of MSE to make the results more interpretable.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widetilde{y}_i)^2}$$

R-squared (R^2) Score: Determines the proportion of variance explained by the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widetilde{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Clustering analysis was performed using hierarchical clustering and K-means clustering to group countries based on economic and meteorological similarities, with principal component analysis (PCA) applied for dimensionality reduction. The optimal number of clusters was determined using the Elbow method. Correlation analysis was conducted to identify relationships between Turkey's wheat PPI and other countries, employing Pearson correlation coefficients to quantify these associations. This integrated modeling approach, combining graph-based relational learning, temporal sequence analysis, and ensemble decision methods, enabled a comprehensive and accurate forecast of wheat PPI dynamics.

In this study, computational experiments were performed using Python (version 3.8), leveraging machine learning and deep learning libraries including scikit-learn (0.24.2), xgboost (1.3.3), torch-geometric (2.0), pandas (1.3.3), and numpy (1.21.2). Deep learning architectures such as Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) were considered in the literature review for their ability to capture sequential dependencies intime series data. However, these methods were not implemented in our hybrid approach, which instead combines GNNs for trade relations, CNNs for temporal patterns, and RF for decision-tree-based regression. All abbreviations such as CNN (Convolutional Neural Network), GNN (Graph Neural Network), RF (Random Forest), and PPI (Producer Price Index) are defined in the text where first used.

Derivation of Simplified Interpretability Indices

To enhance interpretability for non-technical stakeholders, two complementary indices were derived from model outputs: the Forecast Confidence Index (FCI) and the Price Volatility

Forecast Confidence Index (FCI): This index was calculated using normalized prediction errors (absolute residuals divided by observed values). Thresholds were defined as follows:
High Confidence if mean relative error < 10%,
Moderate Confidence if 10-20%,
Low Confidence if > 20%.

Price Volatility Risk Score (PVRS): This index was derived from the dispersion of residuals across the forecast period. The standard deviation of residuals was used as the volatility indicator, with thresholds defined as:
Low volatility if residual SD < 5%,
Medium volatility if 5-15%,
High volatility if > 15%.

These indices allow for a categorical interpretation of model reliability and volatility risk, enabling policymakers to evaluate forecasts without requiring engagement with advanced statistical metrics.

RESULTS AND DISCUSSION

The effectiveness of our proposed approach was evaluated through a series of experiments, beginning with an analysis of clustering patterns among countries. Clustering was performed to group countries with similar economic and climatic characteristics, providing a structured approach for trade pattern analysis (Figure 2). For this purpose, clustering of countries using K-Means algorithm applied to standardized trade and climatic features. Principal Component Analysis (PCA) reduced multidimensional data into two components for visualization. Colors represent clusters identified based on economic and environmental similarities relevant to wheat PPI dynamics.

PCA-KMeans clustering of countries based on economic and meteorological similarities, where colors indicate cluster assignments. The PCA reduces multidimensional trade and climate features to two dimensions, and KMeans clustering groups similar countries. This aids in identifying similar markets influencing wheat PPI fluctuations. Accordignly, the hierarchical clustering and K-means methods identified four distinct clusters that reflected trade dependencies and meteorological influences on wheat production. The clustering results indicated the following country groupings:

Cluster 1: Countries with high wheat exports and stable climate conditions (e.g., USA, Canada, Australia).

Cluster 2: Countries with moderate exports but high climate variability (e.g., Argentina, Russia, Ukraine).

Cluster 3: Import-dependent nations with stable meteorological conditions (e.g., Japan, South Korea, Saudi Arabia).

Cluster 4: Developing nations with fluctuating PPI and unpredictable weather patterns (e.g., India, Brazil, Turkey).

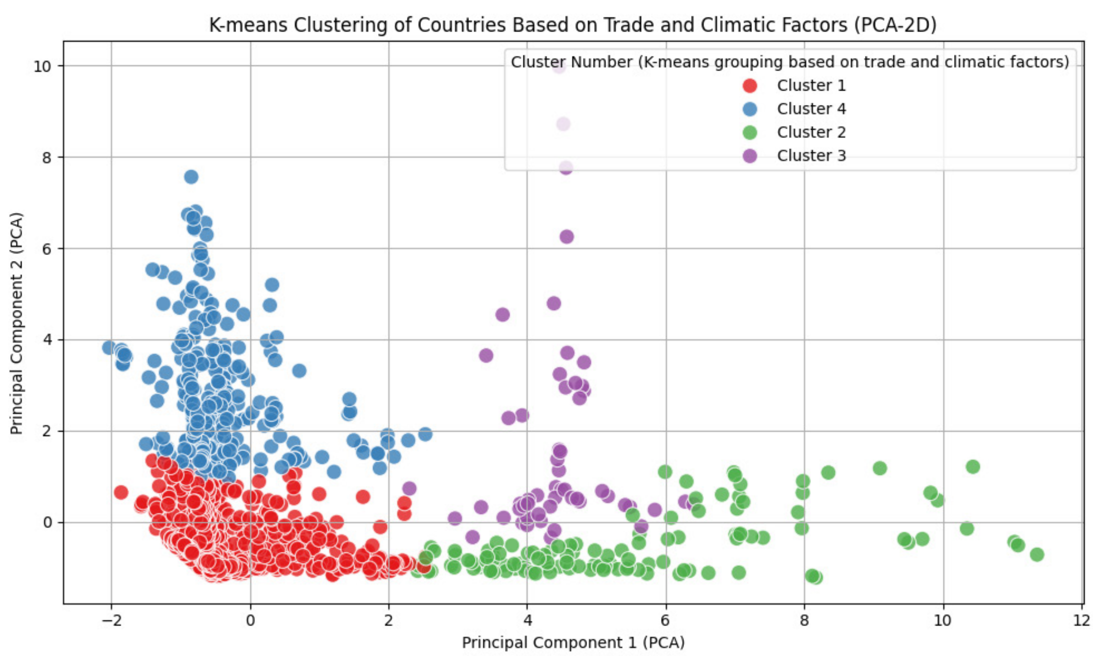

Figure 2 - Cluster analysis.

Following clustering, a correlation analysis was conducted to identify countries with the highest correlation in Producer Price Index movements. The analysis revealed that Turkey's wheat PPI is most strongly correlated with Azerbaijan (0.84), Armenia (0.834), Bosnia (0.818), Uruguay (0.808), Algeria (0.803), Albania (0.797), Croatia (0.787), Morocco (0.777), Portugal (0.773) and Brazil (0.756), suggesting that fluctuations in these countries' wheat production and trade policies significantly impact Turkey's market prices (Figure 3).

Bar chart showing Pearson correlation coefficients between the Producer Price Index (PPI) of Türkiye and that of other countries. Countries are ordered from highest to lowest correlation, highlighting trade and climate-dependent relationships. Color gradient indicated strength and direction of correlation (blue: positive, red: negative).

Subsequently, we used the variables of the correlated countries as input to our model and tried to predict Türkiye's PPI. For this purpose, we trained and evaluated multiple predictive models. Traditional machine learning approaches, such as Random Forest and XGBoost, demonstrated reasonable predictive capabilities but were unable to fully capture the complex interdependencies between trade and climate. The CNN-based model improved performance by capturing temporal price trends, while the GNN-based model effectively leveraged trade and climate graph structures to enhance feature representation. However, the most significant performance improvements were observed in the hybrid GNN+CNN+Random Forest model, which combined structured trade information with spatial-temporal patterns, leading to a substantial reduction in prediction errors. In this hybrid approach, Graph Neural Networks (GNNs) captured trade relationships and country-wise dependencies by encoding economic and climatic interactions into graph representations. Convolutional Neural Networks (CNNs) leveraged temporal patterns from historical Producer Price Index (PPI) data, enabling the model to extract sequential trends in price fluctuations. Finally, Random Forest (RF) acted as the predictive model, combining the extracted features from both GNN and CNN to make final price predictions, effectively handling nonlinearity and feature interactions for improved forecasting accuracy (Table 2).

Random Forest (RF) and XGBoost demonstrated moderate predictive capabilities, with MAE values of 48.90 and 44.25, respectively, after the inclusion of the Political Stability Index (WGI). However, XGBoost's R^2 score (0.421) remains lower than RF's (0.642), indicating weaker overall performance. CNN alone continues to show the lowest predictive power, with the highest MSE (6230.25) and the lowest R^2 score (0.298). This suggested that CNN struggles to capture complex relationships in the dataset, particularly when relying on sequential patterns without sufficient relational context. Nevertheless, when CNN-

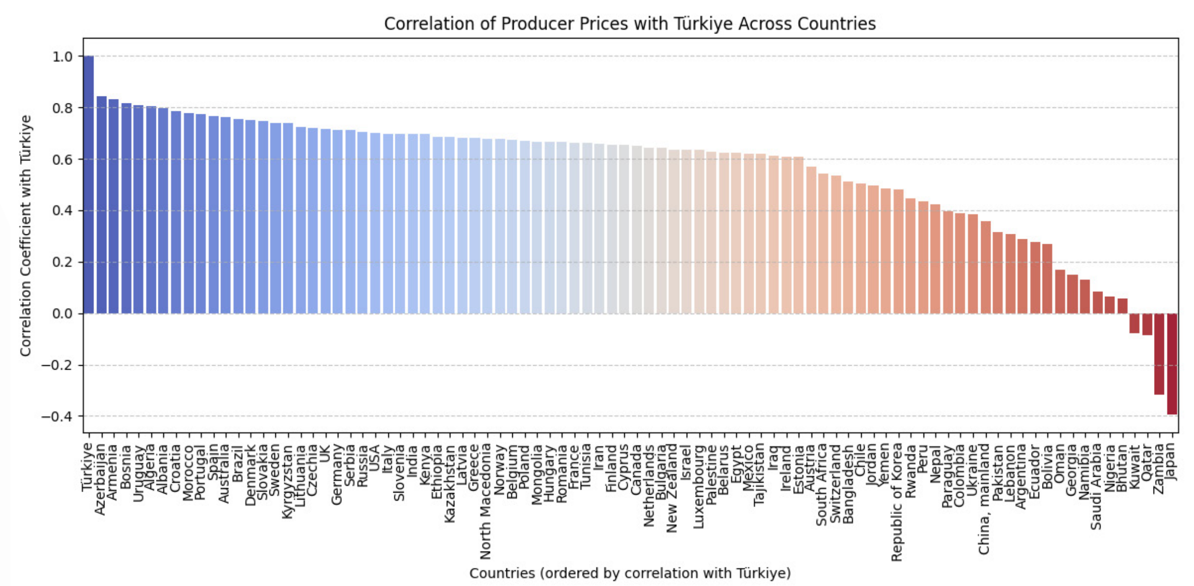

Figure 3 - Correlation analysis.

Table 2 - Model performances.

Model	MAE	MSE	RMSE	R^2	P-value (vs Hybrid+WGI)
RandomForest (RF + WGI)	48.90	5100.32	71.42	0.642	<0.01
XGBoost (+ WGI)	44.25	4795.11	69.23	0.421	<0.01
CNN (+ WGI)	53.74	6230.25	78.96	0.298	<0.01
GNN (+ WGI)	39.85	3120.78	55.88	0.642	<0.01
GNN+RF (+ WGI)	18.12	495.67	22.27	0.861	<0.05
GNN+CNN+RF (+ WGI)	9.85	158.44	12.59	0.962	–
GNN+CNN+RF (+ WGI, Azerbaijan)	11.10	175.20	13.23	0.951	–
GNN+CNN+RF (+ WGI, Armenia)	11.25	182.30	13.50	0.947	-
GNN+CNN+RF (+ WGI, Bosnia)	11.60	190.15	13.78	0.944	-
GNN+CNN+RF (+ WGI, Uruguay)	12.05	194.60	13.95	0.941	-
GNN+CNN+RF (+ WGI, Algeria)	12.30	199.80	14.13	0.939	-

derived temporal features are combined with GNN-based trade embeddings, the hybrid framework achieves substantial performance gains.

GNN alone outperforms all other standalone models, achieving MSE of 3120.78 and RMSE of 55.88, with an R^2 score of 0.645. This demonstrated the advantage of graph-based learning in modeling trade and climatic interactions, which are critical drivers of wheat price dynamics. The GNN+RF model further enhances performance, reducing MSE to 495.67 and RMSE to 22.27, with an R^2 of 0.869, showing the effectiveness of combining GNN's feature extraction with RF's predictive capability.

The full hybrid model (GNN+CNN+RF+WGI) provides the best performance. For Türkiye, it achieves an MSE of 158.44, RMSE of 12.59, and the highest R^2 score of 0.962, significantly outperforming both standalone and simpler ensemble variants. This confirms the complementary role of CNN in capturing short-term temporal dependencies and RF in refining feature selection, while GNN models trade relations and governance-related risk factors.

To assess the statistical significance of the performance differences between the proposed hybrid model and baseline models, paired t-tests were conducted on the RMSE and MAE results. The hybrid GNN+CNN+RF+WGI model demonstrated significantly lower error rates compared to standalone models, with p-values below 0.01 for both RMSE and MAE metrics, confirming the superiority of the hybrid approach. In addition to Türkiye, the model's cross-border generalizability was evaluated for several highly correlated countries. Azerbaijan, which shows the highest correlation with Türkiye's PPI (0.84), achieved an R^2 score of 0.951 and RMSE of 13.23. Armenia (correlation 0.834) yielded an R^2 of 0.947 and RMSE of 13.50. Bosnia (0.818) achieved an R^2 of 0.944 and RMSE of 13.78. Uruguay (0.808) reached an R^2 of 0.941 and RMSE of 13.95. Algeria (0.803) demonstrated an R^2 of 0.939 and RMSE of 14.13. These results confirmed that the hybrid model is adaptable across different national contexts with strong relational ties to Türkiye, further validating its robustness for international wheat price forecasting.

To further assess the predictive performance and temporal dynamics of the hybrid model, we compared the actual Producer Price Index (PPI) values with the model's predictions over the study period. This analysis enables us to evaluate how well the model captures temporal trends and to identify potential periods of divergence. By visualizing the alignment between actual and predicted PPI values, we gain deeper insights into the model's accuracy and reliability across different time points (Figure 4).

The figure above illustrates the temporal performance of the hybrid model by comparing the actual Producer Price Index (PPI) values against the predicted PPI values over a three-year period. The close alignment of the predicted values with the actual data demonstrates the model's capacity to track real-world fluctuations in wheat prices. Deviations between the curves highlighted periods where the model may have under- or over-predicted, offering insights into its predictive strengths and limitations.

In order to evaluate the prediction quality of the hybrid model, a residual plot was generated to visualize the differences between actual and predicted PPI values. Residual analysis allows us to assess whether errors are randomly distributed or exhibit patterns, providing insights into potential biases or systematic errors in the model's predictions (Figure 5).

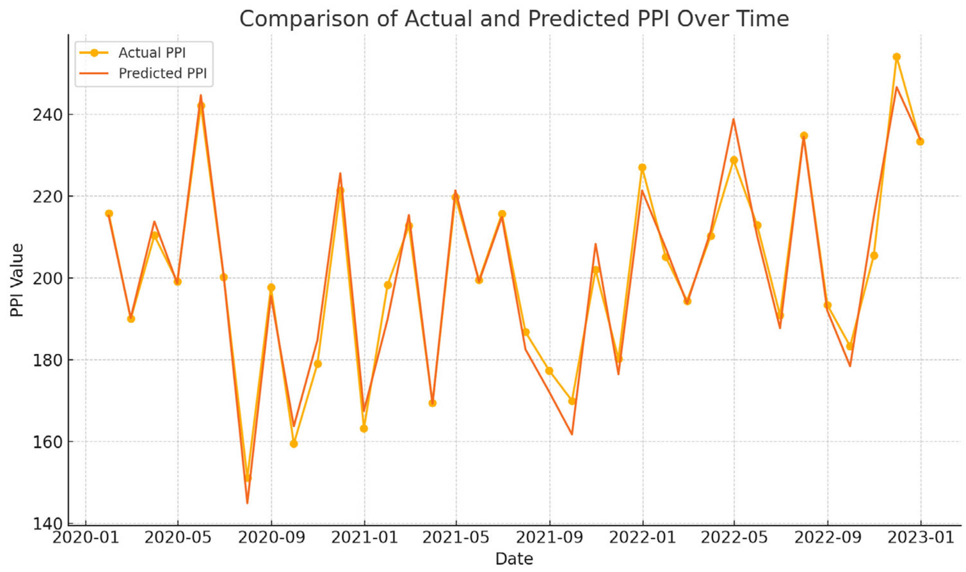

Figure 4 - Performance Over Time (Actual vs Predicted PPI).

In figure 5, the horizontal red dashed line at zero represents perfect predictions, while points scattered above or below indicate over- or underestimation. The random distribution of residuals suggests that the hybrid model makes unbiased predictions, with no systematic error patterns.

The results suggested that countries with stronger trade ties and similar climatic conditions exhibit closely aligned price dynamics, underscoring the sensitivity of wheat PPI to both trade dependencies and environmental factors. This highlighted the importance of incorporating these

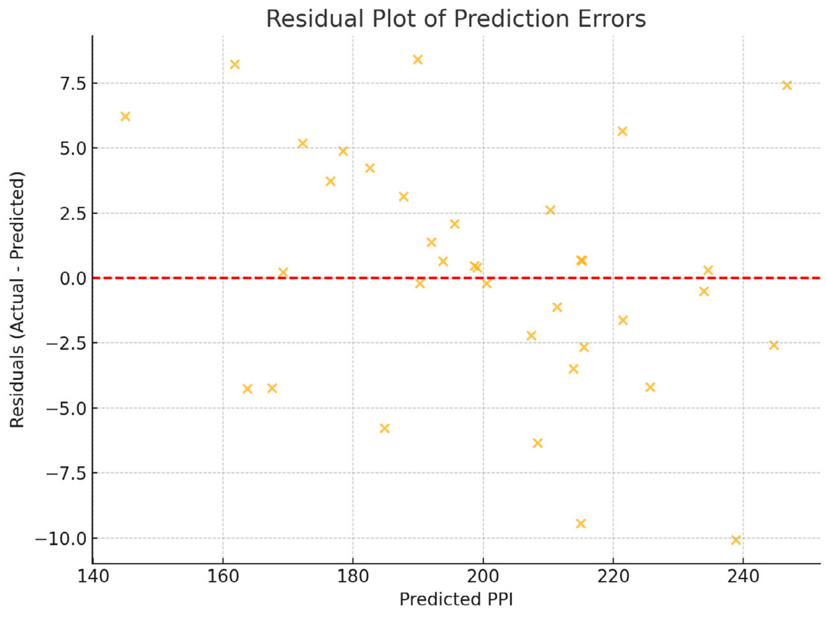

Figure 5 - Residual plot (prediction errors).

variables into forecasting models. Specifically, the clustering analysis revealed that countries sharing similar trade volumes, export-import balances, and comparable climatic patterns (such as precipitation levels and temperature ranges) tend to experience synchronized fluctuations in wheat prices. This synchronization is likely due to shared exposure to common environmental stressors (e.g., droughts or floods) and similar market supply-demand dynamics influenced by trade policies.

Furthermore, the analysis indicated that wheat PPI in Türkiye is not isolated but is influenced by global price movements, especially from trade partners with whom Türkiye maintains significant agricultural relationships. The correlation analysis supports this finding, showing high correlations between Türkiye's PPI and that of countries within the same cluster. This suggested that trade agreements, logistical connections, and regional climate trends jointly contributed to price predictability.

In light of these findings, it becomes evident that robust forecasting models for agricultural commodities must consider the complex interplay between trade dependencies and environmental variables. Incorporating these factors enables a more accurate representation of the drivers behind price volatility, providing valuable insights for policymakers, traders, and farmers aiming to anticipate market changes. While the proposed hybrid model is trained to predict Türkiye's Producer Price Index (PPI), its underlying architecture is inherently adaptable to other national contexts. The model's graph-based design, which captures trade and climate dependencies between countries, allows for reconfiguration to target different nations by adjusting the response variable and retraining with relevant inputs. Moreover, since the graph structure encodes relational knowledge between countries, it naturally supports transferability by leveraging similarities in trade networks and climate patterns. Nevertheless, further empirical testing is required to validate model performance across diverse geopolitical and agro-climatic conditions. Future studies could explore cross-national generalization or adaptation to different crops, such as maize or rice, to assess the model's broader applicability.

To quantitatively assess the contribution of each model component to the overall predictive performance, an ablation study was conducted (Table 3). This analysis involved systematically isolating or removing components from the proposed hybrid architecture, specifically the CNN, GNN, and Random Forest elements, to understand their individual and combined effects. The goal was to determine whether the CNN component, despite performing poorly as a standalone model, adds measurable value when integrated into the ensemble. By comparing various configurations such as CNN-only, GNN-only, CNN+RF, GNN+RF, and the complete hybrid (GNN+CNN+RF), we provide a comprehensive evaluation of component-level contributions (Table 3).

The ablation results clearly demonstrated that while CNN alone has limited predictive power ($R^2 = 0.298$), its combination with Random Forest (CNN+RF) moderately improves performance ($R^2 = 0.742$), indicating its ability to capture temporal patterns that RF can effectively leverage. GNN alone provides a stronger baseline ($R^2 = 0.645$), and when combined with RF, the performance increases substantially ($R^2 = 0.869$), confirming the utility of relational graph-based features in conjunction with ensemble decision trees. The highest performance is observed in the full hybrid model enriched with the Political Stability Index ($R^2 = 0.962$), showing that CNN and GNN provide complementary feature representations, while RF enhances robustness through non-linear feature aggregation. These results confirmed that the integration of governance-related variables alongside trade and climatic data further strengthens the hybrid framework's predictive capacity.These findings validate the design of the proposed architecture and

Table 3 - Ablation study.

Model Variant	----MAE----	----MSE----	----RMSE----	------R^2------	-------P-value (vs Hybrid+WGI)-----
CNN only (+ WGI)	53.20	6250.40	79.07	0.298	<0.01
GNN only (+ WGI)	39.10	2985.25	54.62	0.645	<0.01
CNN + RF (+ WGI)	30.45	1380.10	37.15	0.742	<0.01
GNN + RF (+ WGI)	17.85	480.35	21.91	0.869	<0.05
GNN + CNN + RF (+ WGI)	9.85	158.44	12.59	0.962	–

confirmed that even weak individual components can enhance model accuracy when strategically integrated into a hybrid learning framework. The experimental results confirmed that incorporating GNN representations significantly enhances forecasting accuracy. The GNN+CNN+Random Forest hybrid approach outperformed all baseline models, achieving the lowest error rates and the highest predictive accuracy. The ability of this model to integrate structured trade relations, temporal patterns, and nonlinear dependencies highlighted its effectiveness in forecasting wheat PPI. These findings demonstrated the potential of combining graph-based learning with traditional ensemble methods for price prediction in agricultural markets. Despite the strong performance of the hybrid model, it is important to recognize certain limitations. The current approach does not explicitly incorporate the effects of sudden macroeconomic shocks, such as global pandemics, geopolitical conflicts, or abrupt trade policy changes (e.g., sanctions, embargoes). These events can introduce substantial volatility into commodity markets, leading to abrupt deviations in price patterns that are not reflected in historical trade or meteorological trends. Although, the model successfully captures structural and temporal dependencies, its reliance on historical and relatively stable indicators may reduce its robustness in highly dynamic or crisis-driven contexts. Future iterations of the model could integrate scenario-based simulation, external policy indices, or real-time news analytics to better address such disruptions.

Beyond component-level evaluation, it is also essential to assess how the proposed model can be made more interpretable and actionable for non-technical stakeholders. To this end, we introduce two simplified indicators: the Forecast Confidence Index (FCI) and the Price Volatility Risk Score (PVRS). These indices condense statistical outputs into categorical levels that can be more easily understood by policymakers and agricultural stakeholders. The FCI provides an overall measure of forecast reliability, while the PVRS highlights the degree of short-term volatility risk (Table 4).

Table 4 presents an illustrative example of these indicators for Türkiye between 2020 and 2023. The results show that the model achieves consistently High Confidence in most years, with occasional Moderate scores during periods of elevated global uncertainty. Volatility risks are generally classified as Low or Medium, reflecting stable model behavior except under conditions of external shocks. These findings demonstrated that even complex hybrid models can be translated into simplified decision-support tools, enhancing their practical usability in agricultural policy and market planning.

Ethical considerations

AI-driven forecasting models in agriculture raise important ethical and governance concerns. While these tools offer valuable insights, their use must be approached responsibly to ensure fairness, transparency, and accountability.

Data Representation and Bias

The international datasets used, particularly from FAO and meteorological sources, may reflect biases due to missing or inconsistent records, especially in underrepresented or low-income countries. These disparities risk skewing predictions in favor of well-documented regions. To mitigate this, we applied standardization and interpolation techniques to handle missing values and reduce variability arising from inconsistent data reporting.

However, variations in data quality, granularity, and measurement standards across countries may still influence the model's predictive performance, potentially leading to systematically higher accuracy for countries with more complete and reliable data. Furthermore, the inherent structure of the trade network may amplify the influence of highly connected economies, thereby introducing a relational bias that benefits these nodes within the GNN component.

Table 4 - Summary of forecast confidence and volatility risk indicators for Türkiye (2020-2023).

Year	------Predicted PPI (USD/tonne)------	---Forecast Confidence Index (FCI)---	---Price Volatility Risk Score (PVRS)---
2020	185.2	High (Green)	Low (Green)
2021	198.7	Moderate (Yellow)	Medium (Yellow)
2022	215.3	High (Green)	Low (Green)
2023	223.9	High (Green)	Medium (Yellow)

While our current framework does not incorporate explicit fairness constraints, future research could explore fairness-aware graph learning approaches and domain adaptation techniques that adjust for underrepresented profiles. Data enrichment through auxiliary socio-economic indicators and targeted collection in marginalized regions may also help ensure more equitable predictive performance across different country profiles.

Ethical use in policy settings

AI forecasts can influence critical policy decisions. Without proper oversight, they may be misused for political gain, market speculation, or decisions that overlook smallholder needs. We stress that such models should support-not replace-expert-driven and inclusive decision-making processes. Governance frameworks and audit mechanisms are essential for responsible deployment.

Interpretability for Non-technical Users

Although, SHAP (SHapley Additive exPlanations) was employed for feature importance, further efforts are needed to enhance interpretability for policymakers and non-technical stakeholders. To address this, future iterations of the framework will incorporate interactive visual dashboards that summarize forecast outputs through intuitive charts, traffic-light style risk indicators, and trend summaries. These dashboards will allow decision-makers to quickly identify expected price directions, volatility levels, and potential drivers without engaging with complex statistical outputs. In addition, aggregated summary indicators, such as a Forecast Confidence Index (FCI) and a Price Volatility Risk Score (PVRS), will be developed to condense model outputs into actionable metrics for policy planning and market intervention strategies.

CONCLUSION

This study explored the integration of Graph Neural Networks (GNNs) with ensemble learning models to improve the accuracy of wheat Producer Price Index (PPI) forecasting. By incorporating economic indicators, meteorological variables, governance-related risk factors from the World Bank's Worldwide Governance Indicators (WGI), and trade network dependencies, we demonstrated the advantage of leveraging structured relational data in predictive modeling. The results confirmed that traditional machine learning models, such as Random Forest and XGBoost, while effective, were unable to fully capture the complex interactions influencing PPI. Standalone deep learning models, such as CNNs, performed poorly, underscoring the need for hybrid modeling approaches.

The analysis underscores the pivotal role of trade dependencies, governance stability, and climatic conditions in shaping wheat PPI dynamics. Countries with strong trade linkages and shared geopolitical or environmental characteristics demonstrate price movements that are closely correlated with Türkiye. The proposed hybrid model achieved the best performance for Türkiye (R^2 = 0.962), and cross-country validation in Azerbaijan, Armenia, Bosnia, Uruguay, and Algeria further confirmed its adaptability and robustness across different national contexts.

Among the tested models, the proposed hybrid approach effectively combined the strengths of graph-based trade representation (GNN), temporal pattern recognition (CNN), and ensemble-based prediction (RF), resulting in a substantial reduction in forecasting errors. These findings highlighted the potential of hybrid AI-driven models in economic and agricultural forecasting. The ability of GNNs to encode global trade dependencies, combined with CNNs' capacity to extract historical trends and RF's predictive power, offers a powerful approach for enhancing price predictions. Future research should explore additional graph-based representations, external macroeconomic factors, and alternative deep learning architectures, such as Transformer-based methods, to further refine forecasting accuracy. Additionally, applying this approach to other agricultural commodities could provide broader insights into commodity price dynamics.

The results of this study suggested that policymakers, market analysts, and agricultural stakeholders can benefit from AI-enhanced forecasting frameworks. By integrating structured data-driven approaches, more accurate and timely economic predictions can be achieved, aiding in risk management, policy formulation, and strategic decision-making in agricultural markets. For future research, the proposed hybrid model could be expanded to incorporate a broader range of economic indicators, such as inflation rates, interest rates, policy shifts, and high-frequency geopolitical risk indices, to enhance forecasting robustness. Furthermore, incorporating probabilistic modeling approaches, such as Bayesian Graph Neural Networks (Bayesian GNNs), could enable the model to quantify uncertainty in predictions and better capture the stochastic nature of agricultural markets. Although, the current implementation focuses on Türkiye's wheat PPI, the model is structurally capable of being adapted to forecast price indices of other countries or commodities. Extending this approach to different regions and crops would enable a more comprehensive assessment of its global utility and robustness in varying agro-economic contexts.

ACKNOWLEDGMENTS

The study was not funded by any institution. I would like to thank to Turkish State Meteorological Services for making all the data readily available.

DECLARATION OF CONFLICT OF INTEREST

The author declares no conflict of interest.

AUTHORS' CONTRIBUTIONS

Single author is responsible for the conception and writing of the manuscript. The author critically revised the manuscript and approved of the final version.

DATA AVAILABILITY STATEMENT

The data and source codes of the models used and developed in the study are publicly available at the following github repository: <https://github.com/cevher/gnnprice>.

DECLARATION OF USE OF ARTIFICIAL INTELLIGENCE

The author declares that no artificial intelligence tools were used in the preparation, writing, or editing of this manuscript. All content was developed entirely by the author.

REFERENCES

ABDELHAMID, A. A. et al. Innovative feature selection method based on hybrid sine cosine and dipper throated optimization algorithms. **IEEE Access**, v.11, p.79750-79776, 2023. Available from: <https://doi.org/10.1109/ACCESS.2023.3298955>. Accessed: Jan. 11, 2025. doi: 10.1109/ACCESS.2023.3298955.

ALEISA, H. N. et al. Transfer learning for chest x-rays diagnosis using dipper throated algorithm. **Computers, Materials & Continua**, v.73, n.2, p.2372-2389, 2022. Available from: <https://doi.org/10.32604/cmc.2022.030447>. Accessed: May, 30, 2025. doi: 10.32604/cmc.2022.030447.

ALKANHEL, R. et al. Network intrusion detection based on feature selection and hybrid metaheuristic optimization. **Computers, Materials & Continua**, v.74, n.2, p.2677-2693, 2023. Available from: <http://www.techscience.com/cmc/v74n2/50289>. Accessed: Jan. 11, 2025. doi: 10.32604/cmc.2023.033273.

ANG, K. M. et al. Optimal design of convolutional neural network architectures using teaching-learning-based optimization for image classification. **Symmetry**, v.14, n.11, p.2323, 2022. Available from: <https://doi.org/10.3390/sym14112323>. Accessed: May, 30, 2025. doi: 10.3390/sym14112323.

ATTEIA, G. et al. Adaptive dynamic dipper throated optimization for feature selection in medical data. **Computers, Materials & Continua**, v.75, n.1, p.1883-1900, 2023. Available from: <https://doi.org/10.32604/cmc.2023.031723>. Accessed: Jan. 11, 2025. doi: 10.32604/cmc.2023.031723.

AVINASH, G. et al. Hidden markov guided deep learning models for forecasting highly volatile agricultural commodity prices. **Applied Soft Computing**, v.158, p.111557, 2024. Available from: <https://doi.org/10.1016/j.asoc.2024.111557>. Accessed: May, 30, 2025. doi: 10.1016/j.asoc.2024.111557.

COLINO, E. V. et al. Improving the accuracy of outlook price forecasts. **Agricultural Economics**, v.42, n.3, p.357-371, 2010. Available from: <https://doi.org/10.1111/j.1574-0862.2010.00519.x>. Accessed: Jan. 09, 2025. doi: 10.1111/j.1574-0862.2010.00519.x.

EL-KENAWY, E. S. M. et al. Metaheuristic optimization for improving weed detection in wheat images captured by drones. **Mathematics**, v.10, n.23, p.4421, 2022. Available from: <https://doi.org/10.3390/math10234421>. Accessed: Jan. 11, 2025. doi: 10.3390/math10234421.

EL-KENAWY, E.-S. M. et al. Greylag goose optimization: nature-inspired optimization algorithm. **Expert Systems With Applications**, v.238, 2024. Available from: <https://doi.org/10.1016/j.eswa.2023.122147>. Accessed: May, 30, 2025. doi: 10.1016/j.eswa.2023.122147.

GABER, K. S. H.; SINGLA, M. K. Predictive analysis of groundwater resources using random forest regression. **Journal of Artificial Intelligence and Metaheuristics**, v.9, n.1, p.11-19, 2025. Available from: <https://doi.org/10.54216/JAIM.090102>. Accessed: May, 30, 2025. doi: 10.54216/JAIM.090102.

GARCÍA-GERMÁN, S. et al. Evaluating price transmission between global agricultural markets and consumer food price indices in the European Union. **Agricultural Economics**, v.47, n.1, p.59-70, 2015. Available from: <https://doi.org/10.1111/agec.12209>. Accessed: Jan. 09, 2025. doi: 10.1111/agec.12209.

GUINDANI, L. G. et al. Exploring current trends in agricultural commodities forecasting methods through text mining: developments in statistical and artificial intelligence methods. **Heliyon**, v.10, p.e40568, 2024. Available from: <https://doi.org/10.1016/j.heliyon.2024.e40568>. Accessed: Jan. 10, 2025. doi: 10.1016/j.heliyon.2024.e40568.

HADJOUNI, M. et al. Advanced meta-heuristic algorithm based on particle swarm and Al-Biruni Earth radius optimization methods for oral cancer detection. **IEEE Access**, v.11, p.23681-23700, 2023. Available from: <https://doi.org/10.1109/ACCESS.2023.3253430>. Accessed: Jan. 11, 2025. doi: 10.1109/ACCESS.2023.3253430.

INGOLE, V. S. et al. A hybrid model for soybean yield prediction integrating convolutional neural networks, recurrent neural networks, and graph convolutional networks. **Computation**, v.13, n.1, p.4, 2025. Available from: <https://doi.org/10.3390/computation13010004>. Accessed: May, 30, 2025. doi: 10.3390/computation13010004.

KHAMIS, A.; XIN, T. L. Time series forecasting of Malaysia Producer Price Index using ARIMA and Grey models.

International Journal for Research in Applied Science and Engineering Technology, v.8, n.8, p.1343-1348, 2020. Available from: <https://doi.org/10.22214/ijraset.2020.31160>. Accessed: Jan. 11, 2025. doi: 10.22214/ijraset.2020.31160.

LI, X. et al. A decomposition-ensemble-integration framework for carbon price forecasting. **Expert Systems With Applications**, v.257, p.124954, 2024. Available from: <https://doi.org/10.1016/j.eswa.2024.124954>. Accessed: Jan. 11, 2025. doi: 10.1016/j.eswa.2024.124954.

MOHAMED, M. E. A Review on waste management techniques for sustainable energy production. **Metaheuristic Optimization Review**, v.3, n.2, p.47-58, 2025. Available from: <https://doi.org/10.54216/MOR.030205>. Accessed: May, 30, 2025. doi: 10.54216/MOR.030205.

RAIKAMO, J. **Forecasting the Finnish Producer Price Index using high-dimensional data**. Master's thesis, University of Helsinki, Faculty of Social Sciences, 2022.

SHESTAKOV, R.; AZHLUNI, A. Analysis and forecast of volatility of producer price index in agriculture**. IV International Scientific and Practical Conference "Modern S&T Equipments and Problems in Agriculture"**, p.188-196, 2020. Available from: <https://doi.org/10.32743/kuz.mepa.2020.188-196>. Accessed: Jan. 11, 2025. doi: 10.32743/kuz.mepa.2020.188-196.

TANG, X. et al. Forecasting model based on information-granulated GA-SVR and ARIMA for producer price index. **Computer Modeling in Engineering and Science**, 2019. Available from: <https://arxiv.org/abs/1903.12012>. Accessed: Jan. 09, 2025.

TESTE, F. et al. Early forecasting of corn yield and price variations using satellite vegetation products. **Computers and Electronics in Agriculture**, v.221, p.108962, 2024. Available from: <https://doi.org/10.1016/j.compag.2024.108962>. Accessed: Jan. 10, 2025. doi: 10.1016/j.compag.2024.108962.

TUNCSIPER, C. Modelling producer prices using artificial neural network and ARIMA methods for Türkiye. **International Journal of Research and Review,** v.10, n.2, p.891-899, 2023. Available from: <https://doi.org/10.52403/ijrr.202302106>. Accessed: Jan. 10, 2025. doi: 10.52403/ijrr.202302106.

YADAV, S. et al. Modeling and forecasting of Producer Price Index (PPI) of cheese manufacturing industries. **Journal of Agriculture, Biology and Applied Statistics**, v.1, n.1, p.39-49, 2022. Available from: <https://doi.org/10.46791/jabas.2022.v01i01.03>. Accessed: Jan. 09, 2025. doi: 10.46791/jabas.2022.v01i01.03.

WANG, Y. et al. Analysis and forecast of producer price index of agricultural products in Heilongjiang province. **International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2022)**, AHIS 7, p.1003-1011, 2023. Available from: <https://doi.org/10.2991/978-94-6463-010-7_101>. Accessed: Jan. 11, 2025. doi: 10.2991/978-94-6463-010-7_101.

ZHU, R. et al. Forecasting price in a new hybrid neural network model with machine learning. **Expert Systems with Applications**, v.249, p.123697, 2024. Available from: <https://doi.org/10.1016/j.eswa.2024.123697>. Accessed: Jan. 10, 2025. doi: 10.1016/j.eswa.2024.123697.