

Mello, Fernando Dupin da Cunha; Kumar, Prashant; Nascimento, Erick Giovanni Sperandio

## Article

# Advancements in soybean price forecasting: Impact of AI and critical research gaps in global markets

Economies

## Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* Mello, Fernando Dupin da Cunha; Kumar, Prashant; Nascimento, Erick Giovanni Sperandio (2024) : Advancements in soybean price forecasting: Impact of AI and critical research gaps in global markets, *Economies*, ISSN 2227-7099, MDPI, Basel, Vol. 12, Iss. 11, pp. 1-24, <https://doi.org/10.3390/economies12110310>

This Version is available at:

<https://hdl.handle.net/10419/329237>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

Review

# Advancements in Soybean Price Forecasting: Impact of AI and Critical Research Gaps in Global Markets

Fernando Dupin da Cunha Mello <sup>1</sup>, Prashant Kumar <sup>2</sup> and Erick G. Sperandio Nascimento <sup>1,3,\*</sup>

<sup>1</sup> Stricto Sensu Department, SENAI CIMATEC, Av. Orlando Gomes, 1845, Salvador 41650-010, BA, Brazil; fernando.mello@aln.senaicimatec.edu.br

<sup>2</sup> Global Centre for Clean Air Research (GCARE), School of Sustainability, Civil and Environmental Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, UK; p.kumar@surrey.ac.uk

<sup>3</sup> Surrey Institute for People-Centred AI, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, UK

\* Correspondence: erick.sperandio@surrey.ac.uk

**Abstract:** Soybeans, a vital source of protein for animal feed and an essential industrial raw material, are the most traded agricultural commodity worldwide. Accurate price forecasting is crucial for maintaining a resilient global food supply chain and has significant implications for agricultural economics and policymaking. This review examines over 100 soybean price forecast models published in the last decade, evaluating them based on the specific markets they target—futures or spot—while highlighting how differences between these markets influence critical model design decisions. The models are also classified into AI-powered and traditional categories, with an initial aim to conduct a statistical analysis comparing the performance of these two groups. This process unveiled a fundamental gap in best practices, particularly regarding the use of common benchmarks and standardised performance metrics, which limits the ability to make meaningful cross-study comparisons. Finally, this study underscores another important research gap: the lack of models forecasting soybean futures prices in Brazil, the world’s largest producer and exporter. These insights provide valuable guidance for researchers, market participants, and policymakers in agricultural economics.

**Keywords:** soybean; price forecast; artificial intelligence; futures market; spot market



**Citation:** Mello, Fernando Dupin da Cunha, Prashant Kumar, and Erick G. Sperandio Nascimento. 2024.

Advancements in Soybean Price Forecasting: Impact of AI and Critical Research Gaps in Global Markets.

*Economies* 12: 310. <https://doi.org/10.3390/economies12110310>

Academic Editor: Ralf Fendel

Received: 2 September 2024

Revised: 20 October 2024

Accepted: 8 November 2024

Published: 15 November 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

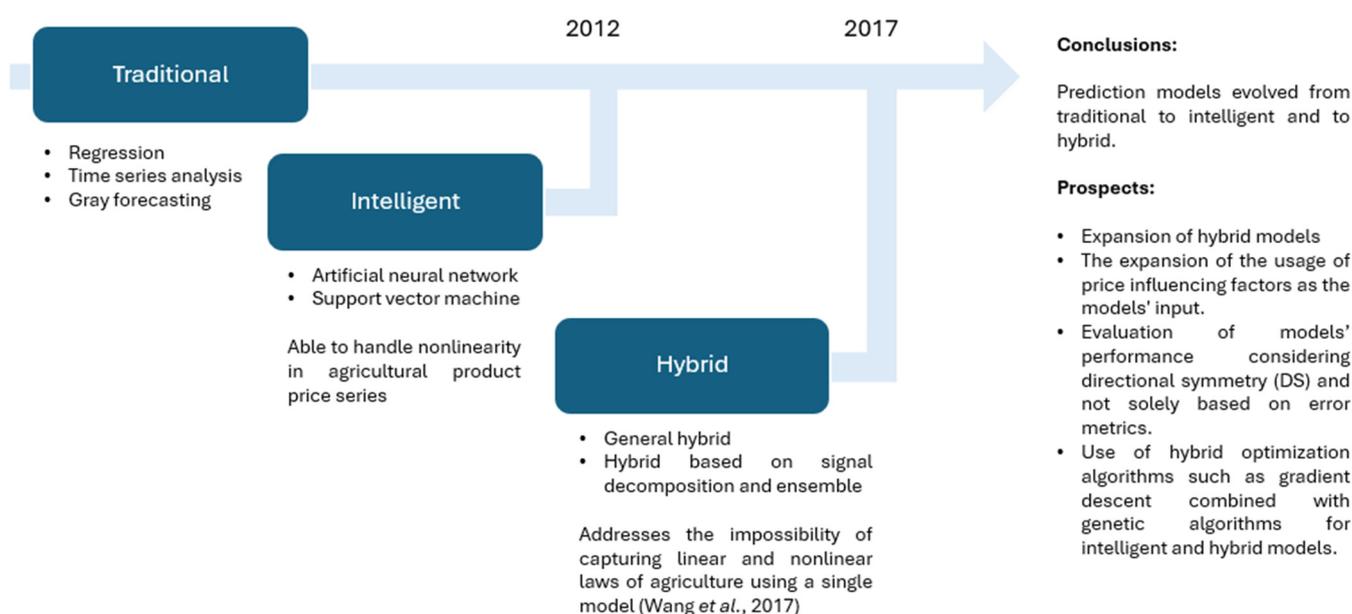
World food security is a pressing global concern, occupying a central position on the agendas of world leaders and international bodies like the Food and Agriculture Organisation of the United Nations (FAO) ([Food and Agriculture Organization of the United Nations 2023](#)). Despite being listed as one of the Sustainable Development Goals (SDG 2), the eradication of hunger remains a formidable challenge. As brought by the FAO report “The State of Food Security and Nutrition in the World 2023”, the intensification of major drivers of food insecurity and malnutrition—conflict, climate extremes, economic slowdowns, and growing inequality—often occurring in combination, has hindered progress towards achieving the SDG2 targets.

Agriculture is the main pillar of food security and therefore stands as a strategic sector, reflecting the collective determination of nations to safeguard the food security of their populations. However, it is also a sector burdened with high risks. These risks stem from the inability to control all variables impacting agricultural outcomes, such as adverse weather conditions, pest infestations, economic downturns, as evidenced by the COVID-19 pandemic, and international conflicts disrupting demand, among others. The repercussions of these risks manifest in fluctuations in agricultural commodity prices. As highlighted by [J. Wang et al. \(2022\)](#), such price volatility serves as a critical factor influencing both economic development and political stability.

Given the volatility in agricultural commodity prices, accurate price forecasting is vital for maintaining stability in global food supply chains and ensuring food security. Precise predictions facilitate optimal resource allocation, risk management, and strategic decision-making for farmers and soybean processing industries. Traders and investors benefit from reliable price forecasts for informed trading strategies. Policymakers use forecasts to shape stable agricultural policies, contributing to global food security. Timely and precise predictions enhance the efficiency of resource allocation, minimising the impact of market fluctuations on food prices (Food and Agriculture Organization of the United Nations 2023). Forecasting in agricultural commodity markets plays a pivotal role in ensuring a resilient and sustainable global food supply chain.

This need for reliable forecasting drives ongoing research into developing and improving price-prediction models. Several research papers (Wang et al. 2020; Wihartiko et al. 2021; Sun et al. 2023) have explored the research landscape and exposed the multitude of methods and their strengths and limitations and pointed to future trends.

L. Wang et al. (2020) present a chronological evolution of forecast models, organising them into three major groups: traditional, intelligent, and hybrid models. They identify the initial papers adopting each group, along with the limiting factors addressed by the introduction of each new group from the prior existing ones. Figure 1 summarises the most relevant outcomes of their review.



**Figure 1.** Summary of the literature review on agricultural product price forecasting methods by L. Wang et al. (2020).

Wihartiko et al. (2021) use a data mining approach to produce an overview based on the appearance of words in the articles. The results showed the methods commonly used in articles published between 2006 and 2020 were AI (30%), data mining (22%), and regression (18%). They recognise the difficulty of predicting prices in a complex supply chain and suggest a future development that would utilise the concept of the multi-objective optimisation problem (MOOP) and neural networks to build a model called the “integrated smart pricing model”. This model would forecast the optimal price and optimal sales time by considering the economic constraints of agricultural business.

Sun et al. (2023) identify the use of hybrid models as a future trend and the adoption of structured and non-structured data for model input as a future trend. They recommend that future proposals should focus on ensuring not only accuracy but also the trend of forecasts.

However, there is a gap in the literature demonstrating how important aspects related to the business environment where the models are created to operate are influencing the

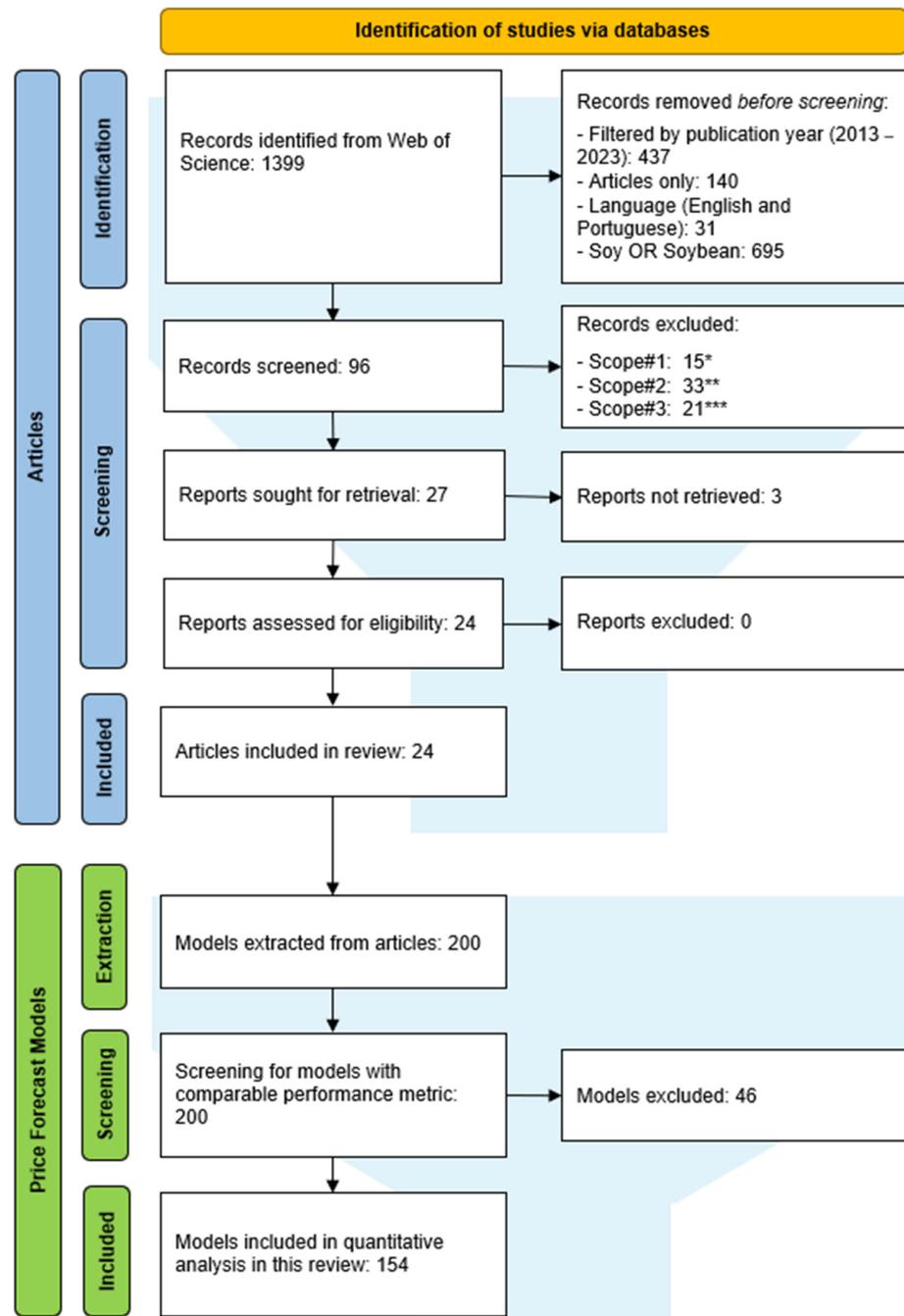
model design and its performance. As an example, the soybean spot market has inherently different dynamics compared to the soybean futures market (Hernandez and Torero 2010). We hypothesise that, instead of reviewing the corpus of published models as a single cluster, as conducted by previous literature reviews, evaluating them through the lens of the specific markets they are designed to serve will yield valuable insights. This approach will help map how the differences between spot and futures markets influence critical model-related design decisions, including the selection of forecast methods, input features, and forecast horizons. Such insights will be relevant in guiding researchers and practitioners in selecting or developing models that enhance decision-making and business operations within their respective market contexts. Additionally, the literature lacks a rigorous quantitative analysis to substantiate the frequent claim that AI improves forecast accuracy in the agricultural commodity price forecasting field. We hypothesise that a comprehensive statistical analysis of published model results will demonstrate that AI-based models consistently outperform traditional models in terms of accuracy. This statistical approach aims to provide empirical evidence supporting the growing affirmation of AI's superiority in this domain, filling a critical gap in the existing literature.

Thus, this review aims to achieve two objectives: (i) to comprehensively analyse the landscape of soybean price forecast models across diverse market types, encompassing the exploration of forecast horizons, prevalence of model categories, complexity of observed features, and representativeness of key producing and consuming regions, and (ii) to produce quantitative evidence to support statistically significant conclusions regarding which type of model has demonstrated the best performance results to date. This work can serve as a reference for researchers and various agents of the soybean supply chain, guiding those interested in conducting research, developing models, or selecting suitable ones to enhance their business operations.

## 2. Scope and Methods

The scope of this literature review article revolves around an in-depth examination of published articles focusing specifically on soybean price forecast models. Notably, this review deliberately excludes other agricultural commodities and soybean by-products such as soybean oil and meal, as their distinct ecophysiological behaviours and market characteristics could potentially confound the comparison of forecast models' performance. Factors such as the number of crops per year, sensitivity to weather variations, volume of production, and importance within the global food supply chain vary significantly among different agricultural commodities. By narrowing the focus solely to soybeans, this review aims to mitigate the influence of these exogenous variables, thus facilitating a more accurate assessment of the performance of various forecast methods.

In conducting this literature review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al. 2021) for identifying, screening, and selecting the articles to be included in this review. Additionally, we incorporated a second stage into the process to guide the extraction, screening, and selection of the forecast models from the chosen articles. The comprehensive process is depicted in Figure 2, outlining the systematic approach undertaken. Detailed descriptions of the steps for selecting articles and models are provided in the following two sections below.



**Figure 2.** Schematic representation of the literature identified in the filtering steps of this systematic review. \* Scope#1 are articles focused on other market-related metrics than price (e.g., volatility, trend, gain); \*\* Scope#2 are articles focused on the market’s nature, not price forecast; \*\*\* Scope#3 are articles not associated with the soybean market or its market price.

### 2.1. Article Selection

The scientific database utilised for this review was the Web of Science, chosen for its comprehensive coverage of the scholarly literature. The search strategy began with a Boolean combination of terms aimed at identifying publications related to the subject of the study with a Boolean combination of terms of agriculture (“agriculture” OR “agricultural”), price (“price” OR “pricing”), and forecast (“forecast” OR “forecasting”), culminating in an initial corpus of 1399 scientific papers. Though the scope of this review is specifically focused on soybean commodities, a broader array of commodities was included in the search process to provide an initial comparative context of soybean representation within

the broader agricultural price forecast domain. A series of filters were then applied to narrow down the publications to the ones relevant to this study. These filters included the following: a filter considering publication year, where publications from the last 11 years were maintained (2013–2023); a filter by document type to restrict the outcome only to research articles; a language filter to retain English and Portuguese research articles; and an inclusion of the term “soy” OR “soybean” in the search query.

The screening phase commenced with 96 articles. Following a thorough examination of abstracts, an additional 69 articles were excluded as they did not align with the specified criteria for soybean price forecasting despite being retrieved through the search query parameters outlined in the preceding paragraph. The reasons for exclusion were categorised into 3 distinct scopes, denoted as Scope#1, Scope#2, and Scope#3, elaborated below.

Scope#1 encompassed articles primarily focused on market-related metrics other than price. Out of the 15 articles classified under this category, 7 were dedicated to predicting price volatility, such as [Marfatia et al. \(2022\)](#), [Chiang et al. \(2019\)](#), [Wang et al. \(2023\)](#), and [Luo et al. \(2022\)](#), 4 to forecasting price trends, e.g., [Xu et al. \(2021\)](#) and [Gu et al. \(2019\)](#), and 4 to projecting price confidence intervals, gains, and variances, such as [Luo et al. \(2023\)](#), [David et al. \(2017\)](#), and [Triantafyllou et al. \(2015\)](#), respectively.

Scope#2 referred to articles excluded from the analysis due to their focus on studying the market’s nature and correlations rather than price forecasting. This category comprised 33 articles. A recurring theme identified was the examination of price correlations between various commodities, including both agricultural and non-agricultural commodities such as oil, as observed in articles by [Chatziantoniou et al. \(2021\)](#), [Wei \(2019\)](#), [David et al. \(2019\)](#), [Uçak et al. \(2023\)](#), [Fiszeder and Orzeszko \(2018\)](#), and [Kumar et al. \(2023\)](#). Additionally, investigations into the impact of the World Agricultural Supply and Demand Estimates (WASDEs) issued by the United States Department of Agriculture (USDA) on soybean market prices were explored in articles by [Isengildina-Massa et al. \(2017\)](#), [Xiao et al. \(2017\)](#), [Goyal and Adjemian \(2023\)](#), [Karali et al. \(2020\)](#), [Karali et al. \(2019\)](#), and [Fernandez-Perez et al. \(2019\)](#). Furthermore, other examples included analyses of how extreme events may affect soybean prices, as evidenced in articles by [Su et al. \(2022\)](#) and [Singvejsakul et al. \(2021\)](#).

Scope#3 comprised articles not associated with the soybean market or its market price. For instance, the articles by [Doi et al. \(2020\)](#), [Iizumi et al. \(2018\)](#), [McNairn et al. \(2014\)](#), [Iizumi et al. \(2013\)](#), [Nejad et al. \(2023\)](#), and [Ju et al. \(2021\)](#) primarily addressed soybean crop yield prediction rather than price, thus necessitating their removal from consideration. This category represented 21 articles.

Three of the twenty-seven full articles intended for retrieval were unavailable in repositories accessible to this research team. Consequently, 24 articles were available for the subsequent phase of data collection, which aimed to catalogue forecast models.

## 2.2. Cataloguing of Soybean Price Forecast Models

The articles were meticulously reviewed to extract the models presented. In addition to extracting the information of the newly proposed models, alternative models developed in the article for comparative analysis were also of interest for this review and therefore were also captured. Thus, it was possible to extract 200 different models from the 24 reviewed articles.

Each model was labelled into one of three major categories. We replicated the same categories adopted by [Wang et al. \(2020\)](#) and repeated by [Sun et al. \(2023\)](#) in their literature reviews concerning agricultural price forecast models. The three categories are traditional models, intelligent models, and hybrid models, as detailed below. Examples of identified articles from each category are provided in Table 1.

Traditional models comprise models based on techniques like regression analysis, time series analysis, including popular methods such as autoregressive integrated moving average (ARIMA) and other variants, and grey forecasting methods, originally introduced by [Deng \(1989\)](#).

**Table 1.** Examples of traditional, intelligent, and hybrid price forecast models included in this review.

Category	Application Example	Reference
Traditional methods	A price forecast model based on seasonal autoregressive integrated moving average (SARIMA) and using price time series as input.	<a href="#">Feliciani et al. (2015)</a>
	A price forecast model based on autoregressive integrated moving average (ARIMA) and using price time series as input.	<a href="#">Li et al. (2022)</a>
Intelligent methods	A model based on the wavelet neural network (WNN) is optimised by the Glowworm Swarm Optimisation algorithm.	<a href="#">Puchalsky et al. (2018)</a>
	A model based on the gradient boosting regression tree (HGBR) utilises price time series and sentiment analysis from data-mined terms as input.	<a href="#">Reis Filho et al. (2022)</a>
Hybrid methods	This study uses dynamic model averaging (DMA) and dynamic model selection (DMS) framework to develop a forecasting model. It has eight predictors, which allows both coefficients and forecasting models to evolve over time.	<a href="#">Xiong and Hu (2021)</a>
	An ensemble model based on extreme gradient boosting machine (XGB).	<a href="#">Ribeiro and Coelho (2020)</a>

Intelligent models comprise models that use AI mechanisms in their forecast process and that gained high relevance in the agricultural price forecasting field due to their ability to capture the nonlinear patterns and to deal with a common lack of normality from input variables affecting the commodity market price ([Sun et al. 2023](#)). Examples of methods found in this field are neural networks in their various forms, such as Long- and Short-Term Memory (LSTM) ([An et al. 2023a, 2023b](#); [Wang et al. 2022](#)), and other methods such as support vector machines ([Jiang et al. 2019](#)) and Bayesian networks ([Drachal 2019](#)).

Hybrid models comprise models that use a combination of two or more forecast models internally and combine their results to achieve a final forecast price number.

A considerable amount of data required to elaborate this review were extracted from each model, and the information is offered as Supplementary Materials in this article. The most relevant information retrieved includes the list of items presented below.

Observed features comprise a list of all variables used by the model as inputs to construct their predictions.

Observed features' sampling frequency refers to the time frequency at which features are observed, such as daily, weekly, or monthly information used.

The forecast horizon is the length of time into the future for which forecasts are to be prepared. This aspect is particularly pertinent because longer forecast horizons often yield lower precision, attributable to heightened uncertainty, the accumulation of errors, and the dynamic nature of systems, among other factors ([Döhrn and Schmidt 2016](#)). Consequently, comparing models with varying forecast horizons might erroneously suggest one model's inferiority over another, when in fact, any disparities in forecast error could stem from the longer forecast horizon of one model relative to its comparator.

Market region refers to the specific geographical area for which the model aims to predict prices (USA, Chicago Board of Trade; India, Akola District Market).

Price type classifies the models into spot or futures prices, reflecting the type of market they intend to forecast. Spot price is the cash price for the immediate purchase and sale of the commodity, whereas the futures price is the price specified in an agreement (futures contract) to deliver a specified quantity of a commodity at a specific future date ([Hernandez and Torero 2010](#)). The futures contracts are important because they represent a financial tool that can be used by any entity to protect its operation against the soybean price fluctuation risk ([Hernandez and Torero 2010](#)).

Evaluation metrics capture all models' performance metrics reported in the articles, such as the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE), among others.

Quantitative results capture the value reported for each evaluation metric.

The screening phase began by identifying a common performance metric that would allow for a standardised comparison of model performance across different articles. After reviewing all the reported performance metrics, the MAPE emerged as the most commonly available metric. Its prevalence across the reviewed papers made it a suitable candidate for establishing a relative comparison metric. As such, the screening process focused on models reporting the MAPE to ensure consistency and comparability in the analysis.

The screening phase started with identifying a performance metric that should be elected for comparing models' performance against each other. The selected metric was the MAPE as this is a dimensionless metric and would accommodate the fact soybean prices in different regions are measured using different units (e.g., USD/bushel or BRL/ton). Moreover, the MAPE was one of the most commonly reported metrics in the population of articles analysed. The second step was to mark among the 200 extracted models those that offered MAPE results. In the end, 154 models providing MAPE results were identified and included in the quantitative analysis.

This review comprises four sections. Section 1 provides a basic overview of the subject, followed by the scope and outline of this paper (Section 2). Section 3 discusses the various findings obtained from the analysis of the published models included in this review, and Section 4 brings a summary of conclusions, recommendations, and areas for future research around the topic covered.

### 3. Findings

#### 3.1. Market Coverage

By analysing the final 24 papers included in this review, it is noticeable that there is a stronger research focus on the markets of the United States and China, with these two regions encompassing more than 70% of the research papers included in this review (Table 2a).

**Table 2.** Market coverage by scientific papers: (a) distribution of articles per market region; (b) distribution of articles per market region and price type.

		(a)			
		Article per Market Region			
		USA	China	Brazil	Others
All articles		10	7	4	3
		42%	29%	17%	13%
		(b)			
		Articles per Market Region			
		USA	China	Brazil	Other
Price Type	Futures Price	6	4	0	0
		60%	40%	0	0
	Spot Price	4	3	4	3
		29%	21%	29%	21%

Going one step further, and categorising these numbers by price type, including futures and spot prices (Table 2b), it becomes evident that the distribution of models for forecasting futures price is responsible for the skewness mentioned in the prior paragraph. The data in Table 2b reveal that all models related to futures price are focused solely on the USA or China. There are no models for futures price covering any other region, including Brazil, despite it being the largest soybean producer and exporter across the world (Dohlman et al. 2001).

A review of the current global soybean market reveals that Brazil is the largest producer and the biggest exporter of this commodity (Dohlman et al. 2001). Notably, Brazil not only

surpassed the USA in 2019 as the biggest producer (Klein and Luna 2021) but also, in 2022, its share of the global soybean exports accounted for an impressive 50% of the total volume traded worldwide (Grand View Research 2024).

As a consequence, all stakeholders in the global soybean supply chain, ranging from countries reliant on imports of this commodity like China, which alone accounts for 70% of global imports (Grand View Research 2024), to industries using soybeans for livestock feed, oil production, biofuels, food, etc., are inevitably impacted by fluctuations in Brazilian production. Accordingly, as stated by the CME Group (Coughlin and Sutton-Vermeulen n.d.), with Brazil's rise as the world's largest producer and exporter of soybeans, financial markets need to offer futures contracts linked to prices in Brazil so that entities with global soy exposure can effectively manage their risks.

Upon observing the data associated with the importance of countries within the soybean market compared to the regions covered by the articles included in this review, the conclusion is that the published academic research lags behind the current market trends and needs. Addressing this gap by developing models focused on forecasting futures prices for Brazil represents a compelling opportunity for future research to enhance the understanding and management of risks within the dynamic soybean market landscape.

### 3.2. Forecast Horizon and Price Type

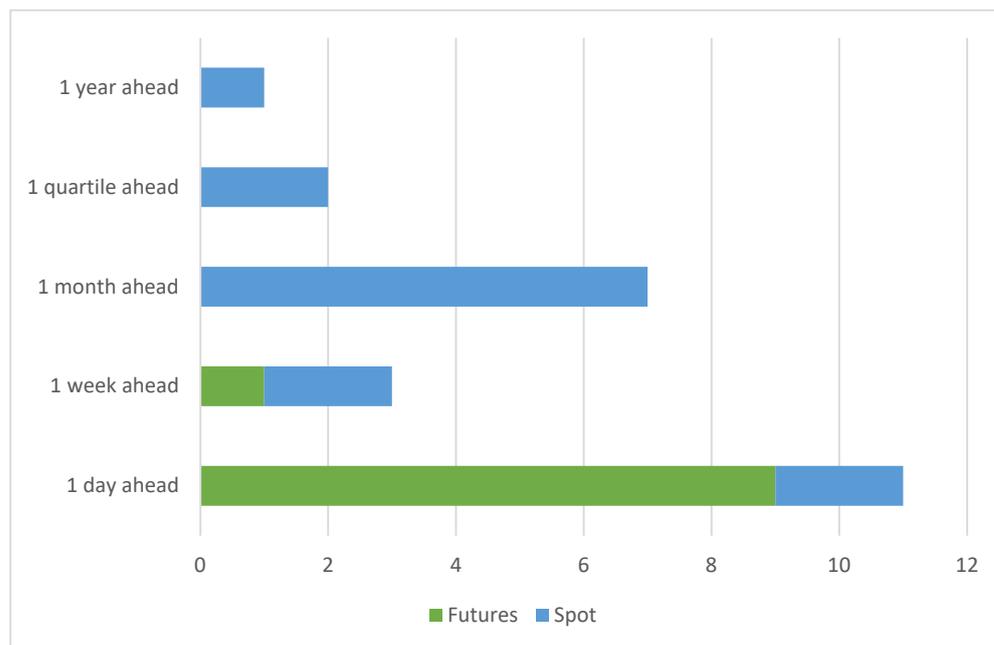
As mentioned in Section 2.2, the forecast horizon is the length of time into the future for which forecasts are to be prepared. The forecast horizon is commonly quantified in steps, with steps referring to an observation interval used in the forecasting process (Hyndman and Athanasopoulos n.d.).

All 24 models reviewed in this article adopt one step ahead as the primary forecast horizon. Additionally, five models (Ahumada and Cornejo 2016b; Crespo Cuaresma et al. 2021; Li et al. 2022; Puchalsky et al. 2018; Jiang et al. 2019) forecast not only one step ahead but multiple steps. For example, Ahumada and Cornejo (2016b) provide three different forecasts: one step ahead, two steps ahead, and four steps ahead. The model presented by Crespo Cuaresma et al. (2021) extends the forecast horizon further, providing forecasts for one, three, six, nine, and twelve steps ahead.

While the adoption of a one-step-ahead forecast horizon by all 24 models initially suggests a significant similarity, we aimed to explore the actual time duration of this step and whether different patterns emerge when comparing the duration of the forecast horizon step between models forecasting the futures market and those forecasting the spot market. To do so, we captured the time length interval of one step for each model and plotted the distribution of all 24 models, also labelling the market type (Figure 3).

The forecast horizon considered by the articles covered in this review ranges from one day to one full year ahead. Upon examining the forecast horizon chosen by the authors, it becomes evident that those creating models to forecast futures prices tend to focus on shorter horizons compared to those forecasting spot prices (Figure 3). This divergence in horizon selection can be explained by the distinct objectives behind these forecasts.

In financial markets, where futures contracts are used, the primary motivation is often related to trading and risk management, particularly hedging price risks for producers and processors (Ouyang and Zhang 2020). These market participants benefit from shorter-term point forecasts that are aligned with trading frequency to optimise profitability and manage price volatility. Speculators in the futures markets also contribute to this focus on shorter horizons, as they react to frequent market fluctuations. Futures prices, driven by both speculative and hedging activities, are influenced by a multitude of volatile factors such as geopolitical events, macroeconomic indicators (e.g., oil prices, stock market indices, exchange rates), and interest rates, which contribute to the necessity for shorter forecast horizons (Ahumada and Cornejo 2016a; Drachal 2019; Xiong and Hu 2021).



**Figure 3.** Distribution of articles by their proposed model's forecast horizon and their split by price type.

On the other hand, forecasting spot prices serves a different economic purpose. Spot prices are more directly tied to domestic supply and demand conditions and reflect longer-term economic trends. Consequently, models designed to forecast spot prices often prioritise longer horizons, focusing on period averages to inform investment decisions, trade terms, and broader macroeconomic factors. As pointed out by [Xu et al. \(2022\)](#), long-term forecasts provide insights into future trends rather than reacting to immediate volatility, making them more relevant for strategic decisions in sectors such as production planning, logistics, and large-scale investments.

Therefore, the selection of forecast horizons aligns closely with the underlying objectives of the forecast: shorter horizons for trading and financial purposes, where managing immediate price volatility is critical, and longer horizons for economic purposes, where understanding trends and period averages provides actionable insights for broader economic activities.

### 3.3. Nominal vs. Real Price Forecast

The distinction between forecasting nominal prices (unadjusted for inflation) and real prices (adjusted for inflation) is crucial for understanding their practical applications. While nominal prices reflect current market prices without adjusting for inflation, real prices provide a clearer view of the true purchasing power of money over time, making them essential for long-term analysis.

This distinction is particularly important for stakeholders making long-term decisions, such as industries planning significant investments in production and infrastructure. For these industries, real prices allow for more accurate forecasts of future costs and returns, providing a solid foundation for decisions involving capital expenditures. In contrast, nominal prices are often of greater interest to market speculators, who focus on short-term trading and seek to capitalise on immediate price fluctuations. Farmers, however, may find value in both perspectives depending on their time horizon. For day-to-day operational decisions, such as selling crops or making purchases, nominal prices are more relevant because they align with short-term cash flow needs and immediate market conditions. On the other hand, for long-term planning—such as making equipment investments or deciding on next season's crop choices—real prices become more important, as they help farmers anticipate the impact of inflation on future purchasing power and profitability.

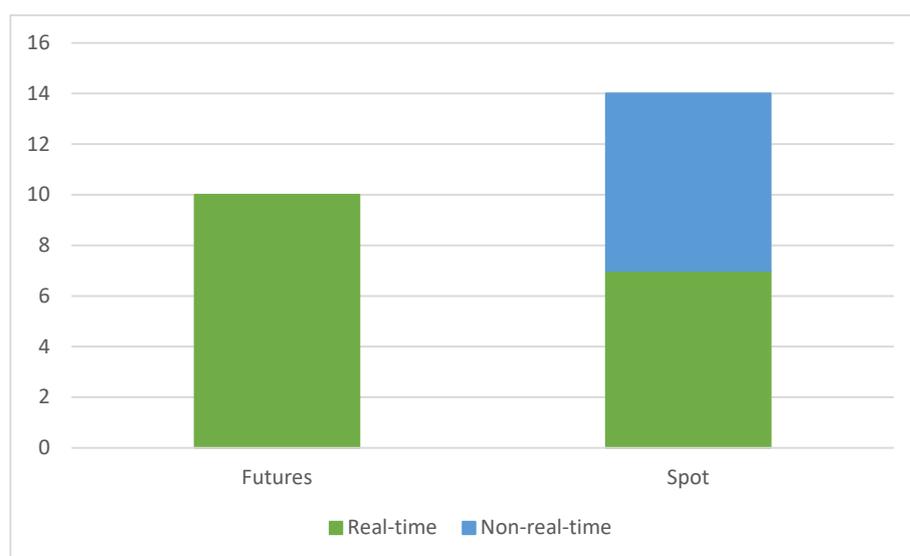
Therefore, while nominal prices are crucial for immediate market decisions, real prices offer a deeper understanding of long-term value and purchasing power.

In this review, we labelled the 24 articles according to whether they forecast real or nominal prices. The results show that all of them focus on nominal prices. While this may be linked to the relatively short-term forecast horizons they use (maximum of 1 year, as shown in Section 3.2), it suggests that not all stakeholders in the soybean supply chain are adequately served by the current research.

### 3.4. Real-Time vs. Non-Real-Time Data Utilisation

A key aspect of evaluating the practical application of soybean price forecast models is determining whether they are based on real-time data or non-real-time data. Real-time data refer to information available at the moment of forecasting, without the benefit of later revisions or updates. In practical applications, especially in volatile markets like soybean futures, forecasts must rely solely on data available in real time. Non-real-time data, on the other hand, often include revisions—such as updates to macroeconomic indicators or trade volumes—that may not be available or updated when the forecast is actually made. Models that do not account for real-time constraints can produce overly optimistic results, as they benefit from hindsight. This can lead to misleading performance metrics in research articles, where models are trained and validated on historical data that include such revisions. In actual production, however, these updates may not be readily available at the time of forecasting, which can result in the model's real-world performance being less accurate than what the reported metrics suggest.

Figure 4 displays the distribution of articles by market type (futures and spot) and their utilisation of real-time and non-real-time data. Notably, all models forecasting prices in the futures market exclusively used real-time data. This finding underscores the critical need for timely and up-to-date information when forecasting futures prices, likely due to the volatility and rapid shifts in futures markets. On the other hand, among the spot market models, there is a more varied approach: out of the fourteen spot models analysed, seven utilised non-real-time data, while seven utilised real-time data. This suggests that for spot price forecasting, non-real-time data, which may include revisions and updates, are less critical. This difference highlights the varying data requirements and flexibility between futures and spot market models, where spot markets may tolerate delayed data revisions, while futures markets demand real-time responsiveness.



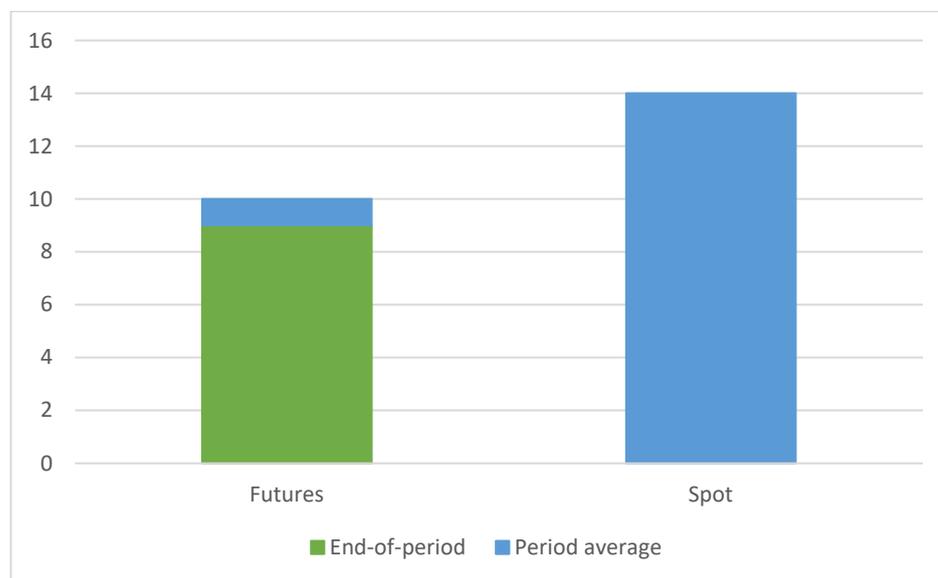
**Figure 4.** Distribution of articles by market type and their split by real-time and non-real-time data utilisation.

### 3.5. End-of-Period vs. Period Average Price Forecasts

End-of-period forecasts predict the price at a specific point in time, such as the last day of a month or quarter, while period average forecasts predict the average price over a set time period (e.g., the daily average across a month). This distinction is significant because each approach serves different practical purposes. End-of-period forecasts are commonly used in financial markets, where traders and investors base decisions on closing prices at specific intervals. In contrast, period average forecasts are more useful for long-term planning and economic analysis, where decisionmakers focus on overall trends rather than short-term fluctuations.

The inclusion of this breakdown is intended to highlight the importance of clarifying the forecast methodology in studies, as different users of these models—such as farmers, policymakers, and traders—require different information based on their respective decision-making horizons. Moreover, inconsistent reporting of whether models forecast end-of-period or period averages complicates cross-study comparisons, potentially leading to the misinterpretation of results.

Figure 5 illustrates the distribution of the 24 articles included in this literature review, categorising them based on whether they forecast end-of-period prices or period average prices and by market type (futures and spot). For futures market models, 90% of the articles forecast end-of-period prices, with only one utilising period averages. This single futures market article executing period average forecasting is [Tang et al.'s \(2023\)](#) article, which proposes a model to forecast weekly futures prices and, to overcome issues like missing daily price data due to holidays, decided to convert daily end-of-period price data into weekly average price data. In contrast, all spot market models exclusively focus on period average prices, with no end-of-period forecasts present in the reviewed literature. This distinction reflects the differing priorities between futures and spot markets—where futures models emphasise short-term precision at specific points in time, and spot market models focus on broader trends over a set period.

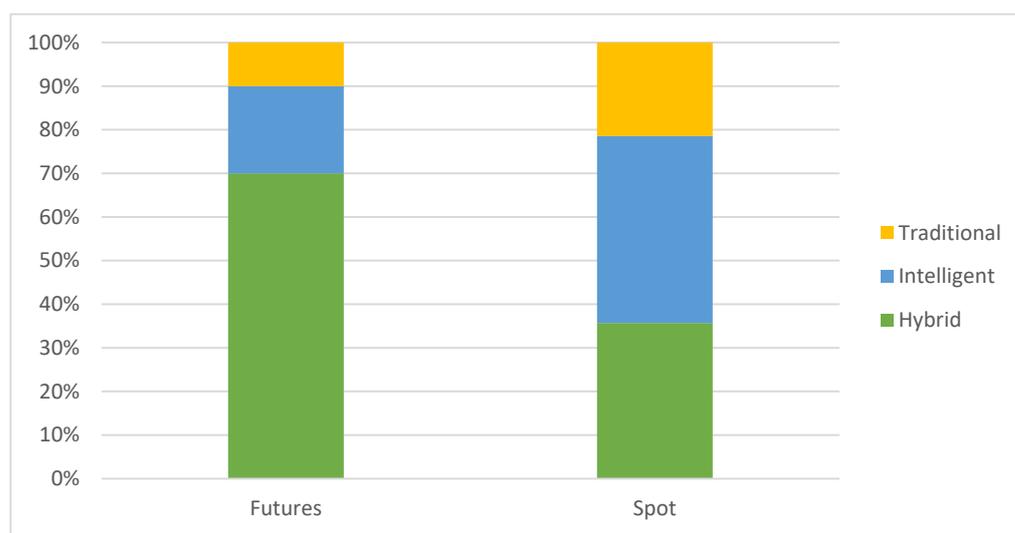


**Figure 5.** Distribution of articles by market type and their split by end-of-period and period average forecast.

### 3.6. Model Category and Price Type

Models designed to forecast futures prices often use hybrid approaches more frequently than those focused on forecasting spot prices (Figure 6). The agricultural commodity futures market, as noted by [Wang et al. \(2022\)](#) and [Huchet-Bourdon \(2011\)](#), is influenced by a multitude of factors such as economic globalisation, financial speculation,

climate change, and oil price fluctuations. This creates a highly complex and nonlinear environment, making accurate forecasts more challenging. Hybrid models have thus become more prominent in the futures market due to their ability to capture these complexities by combining multiple forecast models to improve accuracy. For instance, Wang et al. (2022) employed an Artificial Bee Colony (ABC) algorithm in their hybrid model to forecast soybean futures prices, combining three denoising techniques (Singular Spectral Analysis, Empirical Mode Decomposition, and Variational Mode Decomposition) with different forecasting sub-models to filter noise and enhance prediction.



**Figure 6.** Distribution of articles considering their model's price type and their split by model category.

Traditional models, as used in both futures and spot price forecasting, rely on well-established statistical and econometric methods such as regression analysis and time series models like ARIMA. These models are favoured for their simplicity and robustness when dealing with linear relationships. Traditional models are especially prevalent in the spot market where price behaviour is often seen as being less volatile than futures markets. For example, ARIMA and its variants have long been employed in commodity price forecasting due to their ability to capture and model time-dependent structures within the data (Deng 1989). However, these models may struggle with the nonlinearities that are characteristic of agricultural futures markets, which may explain their lower usage in futures price forecasting. Despite their limitations in handling complex patterns, traditional models remain a foundational approach due to their transparency and interpretability.

Intelligent models leverage AI techniques, including neural networks, support vector machines, and Bayesian networks, to capture nonlinear patterns and deal with the challenges of abnormal data distribution that are often present in agricultural markets (Sun et al. 2023). These models are gaining significant traction in both futures and spot price forecasting due to their capacity to process large datasets and discover hidden relationships that traditional models may miss. For example, LSTM networks (An et al. 2023b) have been applied successfully to forecast prices in markets with complex temporal dependencies. However, while intelligent models excel in dealing with nonlinearities, they often face challenges related to overfitting, requiring careful tuning and validation. Xiong and Hu (2021) propose a dynamic model averaging (DMA) and dynamic model selection (DMS) framework that evolves over time, reflecting the continuous changes in the predictors affecting futures prices.

Hybrid models combine the strengths of multiple forecasting methods, blending traditional, intelligent, and sometimes other forecasting approaches to create more resilient predictions. This combination allows hybrid models to handle both the linear and non-

linear aspects of price movements, making them particularly effective in markets with high volatility and complexity. For instance, dynamic model averaging frameworks and ensemble learning techniques are used to pool the outputs of different sub-models, which can reduce the individual biases of each component model and enhance overall accuracy. In hybrid systems, forecast averaging is often employed, where the final forecast is derived as the weighted or simple average of multiple sub-model predictions.

In summary, the diverse model categories—traditional, intelligent, and hybrid—are applied differently depending on the complexities inherent in the futures and spot markets. The increasing complexity of futures markets has necessitated the rise of hybrid models, while traditional models still hold relevance in simpler spot market forecasting. As the field continues to evolve, intelligent models are emerging as an essential tool in handling nonlinear patterns, though their susceptibility to overfitting and the need for careful calibration remain critical considerations. The combination of different forecasting approaches within hybrid models presents a promising avenue for addressing the diverse challenges of agricultural price forecasting, particularly in futures markets where volatility is high.

### 3.7. Observed Features

The 200 forecast models extracted from the 24 articles analysed in this paper have used a wide variety of inputs (observed features) to produce their forecasts. As a reference, the model proposed by [Xu et al. \(2022\)](#) is fed with 25 variables, including information like soybean planted area, import and export volumes for the commodity, labour cost, land cost, and number of pigs expected by the end of a targeted year, among others. [Ahumada and Cornejo \(2016a\)](#) use 20 different inputs varying from soybean-related data, such as its price and its inventory, to other substitute commodities (e.g., corn) and macro- and microeconomic indexes (e.g., country GDP, currency exchange rate). Some models such as those of [An et al. \(2023a\)](#) and [Reis Filho et al. \(2022\)](#) use data mining techniques and apply sentiment analysis to these data to identify emerging impacting trends that could affect the price.

As an experiment to search for patterns among the available data, we have clustered all the different model inputs using an adaptation of the categories proposed by [Xiong and Hu \(2021\)](#) and bundled the observed features into four categories:

**Price time series:** this category is associated with the utilisation of historical data of the soybean commodity price, which correspond to the most frequent input utilised by the models—in fact, 83% of the 24 articles included in this review use soybean price time series as an input to their models.

**Soybean market fundamentals:** this category includes soybean-related indicators, such as soybean planted area and supply and demand, substitute commodity information (e.g., wheat and corn inventories and prices), and correlated commodity (e.g., annual biofuel production) information.

**Macro- and microeconomic indicators:** this category encompasses observed features like country GDP, currency exchange rate, SP500 index, and oil price.

**Emerging impacting trends:** this last category reflects models containing data mining techniques and sentiment analysis to capture the influence of trending exogenous variables over the price.

Table 3 displays examples of the observed feature types.

The 24 articles were labelled with all applicable input categories. Each article was then assigned a score from 1 to 4 to reflect the number of input categories it uses, constituting a metric we termed observed feature complexity. As an example, the model from [Ribeiro and Coelho \(2020\)](#) was ranked with a score of 2 as it uses soybean price time series at Paranaguá port (price time series category) and exchange rate plus ethanol price (macro and microeconomic indicators category). The model from [An et al. \(2023a\)](#) scored a 4 because of the usage of inputs like soybean price, USD to CNY exchange rate, soybean volume, and trending soybean-related data extracted from the Weibo platform. Consequently, it is considered to have a higher observed feature complexity compared to that in [Ribeiro and Coelho \(2020\)](#). The findings are presented in the next subsections.

**Table 3.** Types of observed features.

Input Type	Application Example	Reference
Price time series	A hybrid model to forecast the USA futures market using soybean futures price history.	D. Y. Wang et al. (2017)
	A hybrid model to forecast Brazil's spot market using soybean spot price history among others.	Ribeiro and Coelho (2020)
Soybean market fundamentals	AI-based model to forecast China's soybean spot price taking as input soybean planted area, total production, and crushing volume, among others.	C. Xu et al. (2022)
	A traditional model to forecast the USA's soybean spot price considering USA, Brazil, and Argentina production volumes, among others.	Crespo Cuaresma et al. (2021)
Macro- and microeconomic indicators	A hybrid model to forecast China's futures market considering USD/CNY exchange rate and crude oil spot price, among others.	Xiong and Hu (2021)
	A hybrid model to forecast USA spot market utilising energy prices, biodiesel production, and S&P 500 index, among others.	Drachal (2019)
Emerging market trends	A hybrid forecast model for China futures market using sentiment from social media text analysis, among others.	An et al. (2023b)
	An intelligent forecast model for USA futures market using domain-specific words acquired via data mining.	Reis Filho et al. (2022)

### 3.7.1. Distribution of Observed Feature Complexity by Model Category

Table 4a displays the distribution of observed feature complexity for each model category. The data reveal that traditional, intelligent, and hybrid model categories have a similar internal distribution of their observed feature complexity, with approximately 50% of the models presenting an observed feature complexity equal to 1. Two outcomes emerge from this scenario:

- Since all three model categories present similar results, there is no evidence suggesting that the observed feature selection made by researchers correlates with a model category choice.
- There appears to be a common explanation for the prevalence of models containing a single observed feature complexity across all model categories, as approximately 50% of models in each category display this characteristic.

In investigating the reason behind the prevalence of models with an observed feature complexity equalling to 1, we delved deeper into the articles encompassing these models (Feliciani et al. 2015; Jiang et al. 2019; Li et al. 2022; Mahto et al. 2021; Manogna and Mishra 2021; Pinheiro and Senna 2017; Puchalsky et al. 2018; Wang et al. 2017; Wang et al. 2022; Xu and Zhang 2022; Yin and Wang 2021). It was noticed that all the models scoring 1 on observed feature complexity share the same input feature: time series. Clues for a plausible hypothesis that could explain the common exclusive use of time series as the observed feature selected by researchers are mentioned in two other articles, Huang et al. (2020) and Yin and Wang (2021). Although the proposed models in these articles are not of complexity 1, they introduce the concept of the Efficient Market Hypothesis (EMH), first formulated by economist Eugene Fama in 1970 (Fama 1970). The concept is that a commodity market could be considered an informationally efficient market, meaning the price history and status fully reflect all available information, and additional input data may not significantly enhance forecast accuracy. It is important to say that the EMH is the subject of ongoing debate and criticism, and further research would be required to draw any firm conclusions.

**Table 4.** Observed feature complexity: (a) distribution of observed feature complexity by model category, (b) distribution of observed feature complexity by price type, and (c) distribution of observed feature complexity by intelligent and non-intelligent models.

(a)					
	Observed Feature Complexity				Weighted avg. Complexity
	1	2	3	4	
Traditional	50%	25%	25%	0%	1.8
Intelligent	50%	25%	25%	0%	1.8
Hybrid	42%	25%	25%	8%	2.0
(b)					
	Observed Feature Complexity				Weighted avg. Complexity
	1	2	3	4	
Spot	36%	21%	36%	7%	2.1
Futures	60%	30%	10%	0%	1.5
(c)					
	Observed Feature Complexity				Weighted avg. Complexity
	1	2	3	4	
Intelligent	53%	29%	12%	6%	1.7
Non-intelligent	29%	14%	57%	0%	2.3

### 3.7.2. Distribution of Observed Feature Complexity by Price Type

Table 4b shows the distribution of observed feature complexity for each price type. The majority of models forecasting futures price use a single category of input features. On the other hand, models forecasting spot prices have a more distributed complexity, showcasing models containing all levels of complexity. To simplify the visualisation of this complexity difference, we calculated a weighted average complexity of the set of models contained in each of the two price types. These values are also shown in Table 4b and convey the set of models used to forecast spot prices have on average (avg. 2.1) a higher degree of complexity regarding the input features than the models created to forecast futures prices (avg. 1.5). These data show that the complexity and multitude of factors affecting the futures price are not driving a more complex selection of input features when compared against models design to forecast spot market prices.

### 3.7.3. Distribution of Observed Feature Complexity by Intelligent and Non-Intelligent Models

During the extraction phase of the models, those classified as hybrids also received a sub-category classification, indicating whether the model was composed of a combination of traditional sub-models, a combination of traditional and intelligent sub-models, or a combination of only intelligent sub-models. This sub-classification allowed us to separate the models into two groups: one called intelligent models, clustering all models that included any level of AI mechanism in their forecast process, and another called non-intelligent models for all the rest. The traditional models were also allocated to the non-intelligent models group, while the models within the intelligent category were allocated to the intelligent models group. We then analysed the complexity of the observed features of these two new categories. The results are shown in Table 4c. Non-intelligent models presented a higher complexity than intelligent models.

### 3.8. Evaluating Benchmarks in Soybean Price Models

Recent advancements in the forecasting literature have raised critical concerns regarding the validity of methods used to evaluate period average price forecasts. Specifically, [Ellwanger and Snudden \(2023\)](#) have shown that many previous studies in the crude oil forecasting literature incorrectly evaluated period average forecasts against a period average no-change benchmark. Their research demonstrates that the appropriate benchmark

for period averages should instead be the end-of-period no-change benchmark, aligning with the random walk (RW) hypothesis used in finance. This adjustment challenges the assumption that period average no-change forecasts provide accurate benchmarks and calls for a reassessment of past results.

In addition, [Farag et al. \(2024\)](#) extend this analysis to futures-based forecasts of commodity prices, including soybeans. They argue that many prior studies used flawed benchmarks when forecasting period average prices by comparing them to a period average no-change benchmark, which has led to misleading conclusions. Their findings emphasise the importance of using the end-of-period no-change benchmark to avoid distortions in forecast evaluation.

In light of these recent findings, we reviewed the 24 papers in our literature analysis to identify whether any of them made the methodological error of comparing period average forecasts against the period average no-change benchmark. According to [Ellwanger and Snudden \(2023\)](#) and [Farag et al. \(2024\)](#), using this incorrect benchmark can lead to invalid conclusions.

To verify this, we started by identifying the articles generating period average forecasts. We then verified whether these articles compare model results to a period average no-change benchmark. The results of this review are presented in Table 5, which highlights the five articles that compare period average forecasts to the period average no-change benchmark. These articles' conclusions regarding the performance of the proposed forecast models against the random walk benchmark are likely questionable, as their methodology does not align with the updated understanding from [Ellwanger and Snudden \(2023\)](#) and [Farag et al. \(2024\)](#).

**Table 5.** Articles comparing period average forecasts against the period average no-change benchmark.

Article Title	Reference
Forecasting food prices: The case of corn, soybeans and wheat	<a href="#">Ahumada and Cornejo (2016a)</a>
Out-of-sample testing price discovery in commodity markets: the case of soybeans	<a href="#">Ahumada and Cornejo (2016b)</a>
Analysis of Agricultural Commodities Prices with New Bayesian Model Combination Schemes	<a href="#">Drachal (2019)</a>
Agricultural commodity price dynamics and their determinants: A comprehensive econometric approach	<a href="#">Crespo Cuaresma et al. (2021)</a>
Agribusiness time series forecasting using wavelet neural networks and metaheuristic optimisation: An analysis of the soybean sack price and perishable products demand	<a href="#">Puchalsky et al. (2018)</a>

### 3.9. Comparison of Forecast Performance: Intelligent vs. Non-Intelligent Models

One of the primary questions this literature review seeks to address is whether soybean price forecast models that apply Artificial Intelligence (AI) techniques outperform those that do not. However, comparing performance across different models is not straightforward due to several inherent challenges. One clear challenge is that models are developed for different markets, each with its own units of measurement (e.g., USD per bushel in the United States vs. BRL per ton in Brazil). To overcome this simple difference, one potential solution is to use dimensionless performance metrics such as the Root Mean Square Percentage Error (RMSPE) or MAPE.

However, while these metrics allow for comparisons across units, they cannot address other challenges and may still lead to distorted results. For example, models that use the same daily data sampled at different frequencies, such as daily versus monthly data, can produce different performance metrics. Daily data capture more granular movements and short-term fluctuations, often leading to smaller percentage errors and more precise short-term predictions. In contrast, monthly data average out daily fluctuations but introduce larger forecast errors due to the increased uncertainty over longer periods. Consequently, models forecasting monthly prices generally produce larger percentage errors than those predicting daily prices. As a result, comparing raw MAPE or RMSPE values across models

with different data sampling frequencies and forecast horizons can lead to misleading conclusions.

The methodology described by Hewamalage et al. (2023) and followed by Filippidis et al. (2024) emphasises the importance of normalising model comparisons by using a common benchmark, such as the RW model, and reporting forecast performance as ratios relative to this benchmark. This approach, exemplified in the equation below using the RMSE, allows for an appropriate comparison of different models:

$$Relative\ RMSE = \frac{RMSE_{model}}{RMSE_{RW}}$$

where the following variables are used:

- $RMSE_{model}$  is the RMSE of the forecast model being evaluated.
- $RMSE_{RW}$  is the RMSE of the RW benchmark.

This approach accounts for differences in forecast volatility, data frequencies, and samples, providing a clearer and more standardised comparison of forecast accuracy. Applying such a benchmark ensures that performance comparisons between AI-based and non-AI models are both fair and accurate, regardless of the inherent differences in market context, data sampling, or forecast horizon.

Based on the above, a sequence of four steps was carried out. The first step was to map the most reported performance metrics used in the 24 articles. The second step involved identifying which of these articles reported the RW benchmark. The third step was to calculate the relative metric, and finally, the samples were divided into two groups—intelligent models and traditional models—in preparation for the statistical comparison of these populations. The results of these screening steps are shown in Table 6.

**Table 6.** Pre-statistical test screening steps: (a) most reported performance metric, (b) RW benchmark availability and accuracy, and (c) split into traditional and intelligent groups.

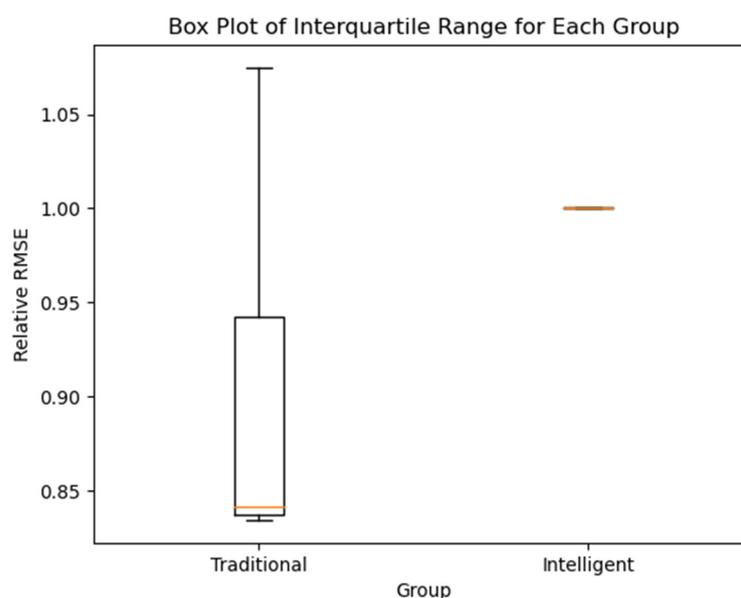
(a)			
Total number of articles: 24			
Most reported metrics			
	Articles	Percentage of total	
MAPE	17	71%	
RMSE	16	67%	
(b)			
Total number of RW articles/(models): 7/(81)			
Proper RW calculation		Inaccurate RW	
Articles	Models	Articles	Models
2	8	5	73
29%	10%	71%	90%
(c)			
Total number of proper RW and RMSE models: 8			
Traditional		Intelligent	
7		1	

The screening of the performance metrics revealed there was not a common candidate across all articles. The most reported metric was the MAPE, reported in 17 out of 24 articles (71% of the total), followed by the RMSE, present in 16 out of 24 articles (67% of the total) (Table 6a).

The second step was to identify which articles reported the performance of the RW forecast as part of their benchmark results. After reviewing the twenty-four articles, we found that only seven articles, encompassing a total of 81 models, reported the performance of the RW benchmark. Of the seven articles, only two (Xiong and Hu 2021; Xu and Zhang

2022) correctly applied the methodology by using end-of-period values to establish the RW benchmark. As discussed in Section 3.8, the remaining five articles inappropriately used a period average no-change benchmark instead of the end-of-period no-change benchmark, which is consistent with the recent findings of Ellwanger and Snudden (2023) and Farag et al. (2024). Consequently, these five articles were excluded from further analysis (Table 6b). This filtering process resulted in only two articles, containing a total of eight models, which met the common benchmark criteria. These two articles (Xiong and Hu 2021; Xu and Zhang 2022) proposed eight models focused on forecasting end-of-period soybean futures price for a 1-day forecast horizon. Xu and Zhang (2022) focus on forecasting Chicago Board of Trade (CBOT) prices, while Xiong and Hu (2021) focus on the Dalian Commodity Exchange (DCR), China.

These eight models reporting RW also shared RMSE as a common performance metric. Using this information, they had their relative performance metric calculated and were separated based on their intelligent or non-intelligent model classification. The traditional group accounted for seven of these models (Table 6c), and the distribution of their RMSE presented a median of 0.84, shown in the box plot of the interquartile range (Figure 7). Only one model belonged to the intelligent group, showing a relative RMSE of 1.00.



**Figure 7.** Box plot of interquartile range for relative MAPE results of intelligent and non-intelligent soybean price forecast models extracted from Xiong and Hu (2021) and Xu and Zhang (2022).

This insufficient number of data points on intelligent models made it unfeasible to carry out the statistical analysis to determine whether intelligent and non-intelligent models exhibit any statistically significant difference in performance. More importantly, this review process highlights a significant opportunity for improvement in the field of soybean price forecasting, particularly in the adoption of common benchmarks, such as the RW model, and in ensuring consistency when reporting performance metrics. Establishing a consensus on benchmark practices and performance metrics would significantly enhance the robustness of model comparisons and facilitate consistent evaluations across different studies.

#### 4. Summary and Conclusions

In light of the extensive analysis conducted on soybean price forecast models published in the last 11 years and covering spot and futures market types, encompassing various forecast horizons, model categories, observed feature complexities, and representations of key producing and consuming regions, this review has pursued two overarching objectives: (i) to comprehensively examine the landscape of soybean price forecast models,

distinguished between those covering futures price forecasts and those covering spot price forecasts, and (ii) to generate quantitative evidence supporting statistically significant conclusions regarding the performance of different model types within each distinct group.

The key conclusions drawn from this work include the following:

- **Market coverage:** Upon observing the data associated with the importance of countries within the soybean market compared to the regions covered by the articles included in this review, the conclusion is that the current state of the art in this field is falling short of the current market trends and needs. Observing the case of Brazil makes it clear that, despite being the largest producer and exporter of soybeans worldwide, the country is not represented in the group of models used to forecast futures prices. This is despite the real business need for stakeholders in the soybean supply chain to hedge their risks against Brazil's soybean price volatility.
- **Forecast horizon and price:** The selection of forecast horizons in soybean price models is driven by the underlying objectives of the forecast. Models focused on financial markets, particularly futures prices, emphasise shorter horizons to manage immediate price volatility and optimise trading strategies. In contrast, models forecasting spot prices prioritise longer horizons, reflecting the strategic needs of economic planning and investment decisions. This duality in horizon selection underscores the importance of aligning model design with the market context in which it operates, offering tailored solutions for distinct market participants.
- **Nominal vs. real price forecast:** The analysis revealed that all models in this review focus on forecasting nominal prices, which may be adequate for short-term decision-making but could overlook the needs of stakeholders requiring long-term planning, such as industries making capital investments.
- **End-of-period vs. period average price forecast:** This review highlights a clear division in the approach taken by futures and spot market models, with futures models predominantly focusing on end-of-period prices and spot market models centring on period averages. This differentiation suggests that futures market participants, including traders and investors, require precise, time-specific forecasts to inform short-term decisions. On the other hand, stakeholders in the spot market, such as policymakers and long-term planners, may prioritise overall price trends and therefore benefit from period average forecasts.
- **Model category and price type:** The analysis reveals that hybrid models dominate in forecasting futures prices due to their ability to account for the high levels of complexity and volatility present in these markets. Traditional models, while simpler and more interpretable, are more commonly used in spot price forecasting, where market conditions are generally less volatile. Intelligent models are applied in both market types. Overall, the findings highlight the importance of selecting the appropriate model type based on the specific dynamics and challenges of the market being forecasted.
- **Observed features:** There is no evidence suggesting that the input features' selection correlates with a model category choice of traditional, intelligent, or hybrid. The higher forecast complexity of the futures market is not driving a more complex selection of input features when compared to models designed to forecast spot market prices. Finally, non-intelligent models presented a higher observed feature complexity than intelligent models.
- **Evaluating benchmarks in soybean price models:** A critical finding emerged from the recent advancements in the literature by [Ellwanger and Snudden \(2023\)](#) and [Frag et al. \(2024\)](#), which raised significant concerns about the use of period average no-change benchmarks in the evaluation of period average forecasts. These studies demonstrated that such benchmarks are inappropriate and often lead to misleading conclusions and that the correct benchmark should be the end-of-period no-change benchmark. Upon reviewing the twenty-four articles included in our analysis, we identified five papers that made this methodological error, making their results questionable under the revised framework proposed by [Ellwanger and Snudden \(2023\)](#) and [Frag et al. \(2024\)](#).

- Comparison of forecast performance of intelligent vs. non-intelligent models:
  - This review identified significant challenges in comparing soybean price forecast models, stemming from the lack of standardised benchmarks and common performance metrics. The inconsistencies in benchmark application and performance metrics across studies limit the ability to robustly evaluate and compare model accuracy.
  - Lack of RW or no-change benchmarks: One notable challenge encountered during this review was the limited number of papers that reported the RW or no-change benchmark. Only seven out of twenty-four articles provided this benchmark, with five of them later being excluded due to methodological inconsistencies. This is a critical limitation because the RW benchmark provides a very relevant reference point for understanding a model's true performance. The absence of this benchmark in many studies inhibits cross-comparisons and makes it difficult for researchers to assess how proposed models perform relative to a simple, well-understood standard.
  - Inconsistent use of performance metrics: Another significant barrier to cross-comparing models is the lack of common performance metrics across studies, which further limits the number of candidate models available for comparison.
  - The findings emphasise the critical need for a unified approach in soybean price forecasting research. Specifically, using a consistent benchmark like the RW model and adopting a common set of metrics would enable more meaningful comparisons and facilitate the tracking of advancements in the field.

The recommendations of this review are as follows:

- Market coverage: Given the underrepresentation of Brazil in models focused on forecasting futures prices, there is a compelling opportunity for future research to address this gap. Pursuing this avenue of investigation could significantly enhance our understanding and management of risks within the dynamic soybean market landscape.
- Nominal vs. real price forecast: Future research should explore the development of models that forecast real prices, providing a more comprehensive tool for stakeholders concerned with long-term profitability and the effects of inflation on the soybean market.
- Real-time vs. non-real-time data utilisation: For those models that utilise non-real-time data, it is recommended that future research examine and report the performance of models in real-world production environments, where access to updated data may be limited. This could involve testing model robustness when real-time data are not available, simulating conditions where data updates are delayed or incomplete.
- End-of-period and period average price forecast: Given the potential for the misinterpretation of forecast results when studies do not explicitly state whether they use end-of-period or period average methods, future research should standardise the reporting of forecast methodologies. Researchers should clearly indicate the type of forecast being produced to facilitate more consistent cross-study comparisons and ensure that the forecast methods align with the needs of the targeted stakeholders. Additionally, exploring hybrid approaches that balance both end-of-period precision and period average trends could provide a more comprehensive view, accommodating a wider range of decision-making needs across both futures and spot markets.
- Observed features: For upcoming research on forecast models in the futures market, the current trend of utilising hybrid approaches as an alternative to dealing with the complexity and nonlinearity of this market should be coupled with an increase in the variety of input factors for those models, exploring AI's ability to identify complex correlations among seemingly unrelated data.
- Evaluating benchmarks in soybean price models: Future studies should incorporate the end-of-period no-change benchmark for evaluating period average forecasts to ensure their results are robust and valid. This adjustment is crucial for forecasting models, as improper benchmarks can significantly distort model performance evaluations.

- Evaluating benchmarks in soybean price models: Researchers should revisit any existing work that used the period average no-change benchmark for period average forecasts. The conclusions from such studies may need to be revised or recalibrated in light of the methodological insights provided by the recent literature.
- Comparison of forecast performance of intelligent vs. non-intelligent models: The lack of common benchmarks and performance metrics across various articles highlights the need for greater standardisation in reporting forecast model performance. This would not only allow for meaningful comparisons among models but also enable the field to track progress more effectively, ultimately fostering the development and evolution of soybean price forecasting research. The following recommendations are made:
  - Authors should prioritise the adoption of a common benchmark, such as the RW hypothesis used in finance, when reporting the performance of soybean price forecast models. This, along with the use of standardised performance metrics, would establish a consistent basis for comparison across different studies.
  - Journals and conferences could promote guidelines for reporting forecasting results, including the recommended benchmarks and performance metrics, to ensure consistency and improve the quality of comparisons within the price forecasting community.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/economies12110310/s1>, Table S1: List of articles included in the screening process; Table S2: List of models and extracted information.

**Author Contributions:** Conceptualization, F.D.d.C.M., P.K., and E.G.S.N.; methodology, P.K., and E.G.S.N.; data curation, F.D.d.C.M.; formal analysis, F.D.d.C.M. and E.G.S.N.; original draft preparation, F.D.d.C.M.; writing—review and editing, F.D.d.C.M., P.K. and E.G.S.N.; supervision, E.G.S.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001 - (Proc. 88881.846220/2023-01), Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) (Proc. 493/2022), and the National Council for Scientific and Technological Development (CNPq, Brazil), through Finance Code 001. Erick G. Sperandio Nascimento is a CNPq technological development fellow (Proc. 308963/2022-9). Prashant Kumar acknowledges the support received through the UKRI (NERC, EPSRC, AHRC) funded RECLAIM Network Plus (EP/W034034/1). The authors also thank the Manufacturing and Technology Integrated Campus—SENAI CIMATEC, the Surrey Institute for People-Centred AI, and the Global Centre for Clean Air Research (GCARE) at the University of Surrey for their valuable support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ahumada, Hildegart, and Magdalena Cornejo. 2016a. Forecasting food prices: The case of corn, soybeans and wheat. *International Journal of Forecasting* 32: 838–48. [[CrossRef](#)]
- Ahumada, Hildegart, and Magdalena Cornejo. 2016b. Out-of-sample testing price discovery in commodity markets: The case of soybeans. *Agricultural Economics* 47: 709–18. [[CrossRef](#)]
- An, Wuyue, Lin Wang, and Dongfeng Zhang. 2023a. Comprehensive commodity price forecasting framework using text mining methods. *Journal of Forecasting* 42: 1865–88. [[CrossRef](#)]
- An, Wuyue, Lin Wang, and Yu-Rong Zeng. 2023b. Text-based soybean futures price forecasting: A two-stage deep learning approach. *Journal of Forecasting* 42: 312–30. [[CrossRef](#)]
- Chatziantoniou, Ioannis, Stavros Degiannakis, George Filis, and Tim Lloyd. 2021. Oil Price Volatility is Effective in Predicting Food Price Volatility. Or is it? *Energy Journal* 42: 25–48. [[CrossRef](#)]
- Chiang, Shu-Mei, Chun-Da Chen, and Chien-Ming Huang. 2019. Analyzing the impacts of foreign exchange and oil price on biofuel commodity futures. *Journal of International Money and Finance* 96: 37–48. [[CrossRef](#)]
- Coughlin, Aison, and Dominic Sutton-Vermeulen. n.d. Relationship Between Global Soybean Prices. Available online: <https://www.cmegroup.com/education/articles-and-reports/relationship-between-global-soybean-prices.html> (accessed on 1 March 2014).
- Crespo Cuaresma, Jesus, Jaroslava Hlouskova, and Michael Obersteiner. 2021. Agricultural commodity price dynamics and their determinants: A comprehensive econometric approach. *Journal of Forecasting* 40: 1245–73. [[CrossRef](#)]

- David, Sergio A., J. A. Machado, Lucas R. Trevisan, Claudio Inacio, Jr., and António M. Lopes. 2017. Dynamics of Commodities Prices: Integer and Fractional Models. *Fundamenta Informaticae* 151: 389–408. [CrossRef]
- David, Sergio Adriani, Claudio M. C. Inácio, Jr., and José A. Tenreiro Machado. 2019. Ethanol Prices and Agricultural Commodities: An Investigation of Their Relationship. *Mathematics* 7: 774. [CrossRef]
- Deng, Julong. 1989. Introduction to grey system. *Journal of Grey System* 1: 1–24.
- Dohlman, Erik, Randall Schnepf, and Chris Bolling. 2001. Soybean Production Costs and Export Competitiveness in the United States, Brazil, and Argentina. *Oil and Oilseed Outlook* 262: 16–24.
- Döhrn, Roland, and Christoph M. Schmidt. 2016. Information or Institution? On the Determinants of Forecast Accuracy. In *Economic Forecasts*. Berlin: De Gruyter Oldenbourg. [CrossRef]
- Doi, Yasuhiro, Yonghee Shin, Wonsik Kim, Jaewon CHoi, and Toshichika Iizumi. 2020. Assessing the subnational-level yield forecast skills of the 2019/20 season NARO-APCC Joint Crop Forecasting Service for Southern Hemisphere countries. *Journal of Agricultural Meteorology* 78: 66–77. [CrossRef]
- Drachal, Krzysztof. 2019. Analysis of Agricultural Commodities Prices with New Bayesian Model Combination Schemes. *Sustainability* 11: 5305. [CrossRef]
- Ellwanger, Reinhard, and Stephen Snudden. 2023. Forecasts of the real price of oil revisited: Do they beat the random walk? *Journal of Banking and Finance* 154: 106962. [CrossRef]
- Fama, Eugene F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance* 25: 383–417. [CrossRef]
- Farag, Markos, Stephen Snudden, and Gregory Upton. 2024. *Can Futures Prices Predict the Real Price of Primary Commodities?* USAEE Working Paper No. 24-629. Available online: <https://ssrn.com/abstract=4906704> (accessed on 1 March 2024).
- Feliciani, Acássio Valente, Adriano Mendonça Souza, and Francisca Mendonça Souza. 2015. Stochastic models of forecasting prices of soybeans in Brazil. *Custos E Agronegocio on Line* 11: 263–80.
- Fernandez-Perez, Adrian, Bart Frijns, Ivan Indriawan, and Alireza Tourani-Rad. 2019. Surprise and dispersion: Informational impact of USDA announcements. *Agricultural Economics* 50: 113–26. [CrossRef]
- Filippidis, Michail, George Filis, and Georgios Magkonis. 2024. Evaluating Oil Price Forecasts A Meta-Analysis. *The Energy Journal* 45: 71–89. [CrossRef]
- Fiszeder, Piotr, and Witold Orzeszko. 2018. Nonlinear Granger causality between grains and livestock. *Agricultural Economics-Zemledska Ekonomika* 64: 328–36. [CrossRef]
- Food and Agriculture Organization of the United Nations. 2023. *Rising Global Food Insecurity: Assessing Policy Response: A Report of Prepared at the Request of the Group of 20 (G20)*. Available online: <https://openknowledge.fao.org/server/api/core/bitstreams/2f7d11db-dad5-442e-934d-836d7cec2723/content> (accessed on 1 March 2024).
- Goyal, Raghav, and Michael K. Adjemian. 2023. Information rigidities in USDA crop production forecasts. *American Journal of Agricultural Economics* 105: 1405–25. [CrossRef]
- Grand View Research. 2024. *Soybean Market Size, Share & Trends Analysis Report By Nature (GMO, Non-GMO), By Form (Raw, Processed), By End-use (Animal Feed, Industrial Use), By Region, And Segment Forecasts, 2024—2030*. Available online: <https://www.grandviewresearch.com/industry-analysis/soybean-market-report> (accessed on 1 March 2024).
- Gu, Quan, Na Lu, and Lin Liu. 2019. A novel recurrent neural network algorithm with long short-term memory model for futures trading. *Journal of Intelligent and Fuzzy Systems* 37: 4477–84. [CrossRef]
- Hernandez, Manuel A., and Maximo Torero. 2010. Examining the dynamic relation between spot and futures prices of agricultural commodities. *Commodity Market Review* 2009–2010: 47–87.
- Hewamalage, Hansika, Klaus Ackermann, and Christoph Bergmeir. 2023. Forecast evaluation for data scientists: Common pitfalls and best practices. *Data Mining and Knowledge Discovery* 37: 788–832. [CrossRef] [PubMed]
- Huang, Joshua, Teresa Serra, and Philip Garcia. 2020. Are futures prices good price forecasts? Underestimation of price reversion in the soybean complex. *European Review of Agricultural Economics* 47: 178–99. [CrossRef]
- Huchet-Bourdon, Marilyne. 2011. Agricultural Commodity Price Volatility: An Overview. *OECD Food, Agriculture and Fisheries Working Papers* 52: 137–85.
- Hyndman, Rob J., and George Athanasopoulos. n.d. *Forecasting Principles and Practice*. Available online: <https://otexts.com/fpp2/> (accessed on 1 March 2024).
- Iizumi, Toshichika, Hirofumi Sakuma, Masayuki Yokozawa, Jing-Jia Luo, Andrew J. Challinor, Molly E. Brown, Gen Sakurai, and Toshio Yamagata. 2013. Prediction of seasonal climate-induced variations in global food production. *Nature Climate Change* 3: 904–08. [CrossRef]
- Iizumi, Toshichika, Yonghee Shin, Wonsik Kim, Moosup Kim, and Jaewon Choi. 2018. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. *Climate Services* 11: 13–23. [CrossRef]
- Isengildina-Massa, Olga, Berna Karali, and Scott H. Irwin. 2017. Do Markets Correct for Smoothing in USDA Crop Production Forecasts? Evidence from Private Analysts and Futures Prices. *Applied Economic Perspectives and Policy* 39: 559–83. [CrossRef]
- Jiang, Feng, Jiaqi He, and Zhigang Zeng. 2019. Pigeon-inspired optimization and extreme learning machine via wavelet packet analysis for predicting bulk commodity futures prices. *Science China Information Sciences* 62: 1–19. [CrossRef]
- Ju, Sungha, Hyoungjoon Lim, Jong Won Ma, Soohyun Kim, Kyungdo Lee, Shuhe Zhao, and Joon Heo. 2021. Optimal county-level crop yield prediction using MODIS-based variables and weather data: A comparative study on machine learning models. *Agricultural and Forest Meteorology* 307: 108530. [CrossRef]

- Karali, Berna, Olga Isengildina-Massa, Scott H. Irwin, Michael K. Adjemian, and Robert Johansson. 2019. Are USDA reports still news to changing crop markets? *Food Policy* 84: 66–76. [CrossRef]
- Karali, Berna, Scott H. Irwin, and Olga Isengildina-Massa. 2020. Supply Fundamentals and Grain Futures Price Movements. *American Journal of Agricultural Economics* 102: 548–68. [CrossRef]
- Klein, Herbert S., and Francisco Vidal Luna. 2021. The growth of the soybean frontier in south america: The case of brazil and argentina. *Revista de Historia Económica—Journal of Iberian and Latin American Economic History* 39: 427–68. [CrossRef]
- Kumar, Pawan, Vipul Kumar Singh, and Sandeep Rao. 2023. Does the substitution effect lead to feedback effect linkage between ethanol, crude oil, and soft agricultural commodities? *Energy Economics* 119: 106574. [CrossRef]
- Li, Jianping, Guowen Li, Mingxi Liu, Xiaoqian Zhu, and Lu Wei. 2022. A novel text-based framework for forecasting agricultural futures using massive online news headlines. *International Journal of Forecasting* 38: 35–50. [CrossRef]
- Luo, Jiawen, Tony Klein, Qiang Ji, and Chenghan Hou. 2022. Forecasting realized volatility of agricultural commodity futures with infinite Hidden Markov HAR models. *International Journal of Forecasting* 38: 51–73. [CrossRef]
- Luo, Rui, Jinpei Liu, Piao Wang, Zhifu Tao, and Huayou Chen. 2023. A multisource data-driven combined forecasting model based on internet search keyword screening method for interval soybean futures price. *Journal of Forecasting* 43: 366–90. [CrossRef]
- Mahto, Anil Kumar, M. Afshar Alam, Ranjit Biswas, Jawed Ahmed, and Shah Imran Alam. 2021. Short-Term Forecasting of Agriculture Commodities in Context of Indian Market for Sustainable Agriculture by Using the Artificial Neural Network. *Journal of Food Quality* 2021: 9939906. [CrossRef]
- Manogna, R. L., and Aswini Kumar Mishra. 2021. Forecasting spot prices of agricultural commodities in India: Application of deep-learning models. *Intelligent Systems in Accounting Finance and Management* 28: 72–83. [CrossRef]
- Marfatia, Hardik A., Qiang Ji, and Jiawen Luo. 2022. Forecasting the volatility of agricultural commodity futures: The role of co-volatility and oil volatility. *Journal of Forecasting* 41: 383–404. [CrossRef]
- McNairn, Heather, Angela Kross, David Lapen, Ron Caves, and Jiali Shang. 2014. Early season monitoring of corn and soybeans with TerraSAR-X and RADARSAT-2. *International Journal of Applied Earth Observation and Geoinformation* 28: 252–59. [CrossRef]
- Nejad, Seyed Mahdi Mirhoseini, Dariush Abbasi-Moghadam, Alireza Sharifi, Nizom Farmonov, Khilola Amankulova, and Mucsi László. 2023. Multispectral Crop Yield Prediction Using 3D-Convolutional Neural Networks and Attention Convolutional LSTM Approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 16: 254–66. [CrossRef]
- Ouyang, Ruolan, and Xuan Zhang. 2020. Financialization of agricultural commodities: Evidence from China. *Economic Modelling* 85: 381–89. [CrossRef]
- Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, and et al. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ* 372: 178–89. [CrossRef]
- Pinheiro, Carlos Alberto Orge, and Valter de Senna. 2017. Forecasting of prices through neural networks and spectral analysis: Evidence for the future market commodity sugar and soy. *Custos E Agronegocio on Line* 13: 103–28.
- Puchalsky, Weslly, Gabriel Trierweiler Ribeiro, Claudimar Pereira da Veiga, Roberto Zanetti Freire, and Leandro dos Santos Coelho. 2018. Agribusiness time series forecasting using Wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack price and perishable products demand. *International Journal of Production Economics* 203: 174–89. [CrossRef]
- Reis Filho, Ivan José, Ricardo Marcondes Marcacini, and Solange Oliveira Rezende. 2022. On the enrichment of time series with textual data for forecasting agricultural commodity prices. *MethodsX* 9: 101758. [CrossRef]
- Ribeiro, Matheus Henrique Dal Molin, and Leandro dos Santos Coelho. 2020. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Applied Soft Computing* 86: 105837. [CrossRef]
- Singvejsakul, Jittima, Chukiat Chaiboonsri, and Songsak Sriboonchitta. 2021. The Optimization of Bayesian Extreme Value: Empirical Evidence for the Agricultural Commodities in the US. *Economics* 9: 30. [CrossRef]
- Su, Yuandong, Chao Liang, Li Zhang, and Qing Zeng. 2022. Uncover the response of the U.S grain commodity market on El Niño-Southern Oscillation. *International Review of Economics and Finance* 81: 98–112. [CrossRef]
- Sun, Feihu, Xianyong Meng, Yan Zhang, Yan Wang, Hongtao Jiang, and Pingzeng Liu. 2023. Agricultural Product Price Forecasting Methods: A Review. *Agriculture* 13: 1671. [CrossRef]
- Tang, Dezhao, Qiqi Cai, Tiandan Nie, Yuanyuan Zhang, and Jinghua Wu. 2023. Agricultural price forecasting based on the spatial and temporal influences factors under spillover effects. *Kybernetes*. Available online: <https://doi.org/10.1108/K-09-2023-1724> (accessed on 1 March 2024).
- Triantafyllou, Athanasios, George Dotsis, and Alexandros H. Sarris. 2015. Volatility Forecasting and Time-varying Variance Risk Premiums in Grains Commodity Markets. *Journal of Agricultural Economics* 66: 329–57. [CrossRef]
- Uçak, Harun, Esin Yelgen, and Yakup Ari. 2023. The volatility connectedness between chicken and selected crops. *Worlds Poultry Science Journal* 80: 55–73. [CrossRef]
- Wang, Deyun, Chenqiang Yue, Shuai Wei, and Jun Lv. 2017. Performance Analysis of Four Decomposition-Ensemble Models for One-Day-Ahead Agricultural Commodity Futures Price Forecasting. *Algorithms* 10: 108. [CrossRef]
- Wang, Jue, Zhen Wang, Xiang Li, and Hao Zhou. 2022. Artificial bee colony-based combination approach to forecasting agricultural commodity prices. *International Journal of Forecasting* 38: 21–34. [CrossRef]
- Wang, Lu, Rui Wu, WeiChun Ma, and Weijiu Xu. 2023. Examining the volatility of soybean market in the MIDAS framework: The importance of bagging-based weather information. *International Review of Financial Analysis* 89: 102720. [CrossRef]

- Wang, Luyao, Jianying Feng, Xiaojie Sui, Xiaoquan Chu, and Weisong Mu. 2020. Agricultural product price forecasting methods: Research advances and trend. *British Food Journal* 122: 2121–38. [[CrossRef](#)]
- Wei, Yanfeng. 2019. The relationship between oil and non-oil commodity prices and China's PPI and CPI: An empirical analysis. *Energy Sources Part B Economics Planning and Policy* 14: 125–46. [[CrossRef](#)]
- Wihartiko, Fajar Delli, Sri Nurdiati, Agus Buono, and Edi Santosa. 2021. Agricultural price prediction models: A systematic literature review. Paper presented at the 11th Annual International Conference on Industrial Engineering and Operations Management, Singapore, March 7–11; pp. 2927–34.
- Xiao, Jinzhi, Chad E. Hart, and Sergio H. Lence. 2017. USDA Forecasts of Crop Ending Stocks: How Well Have They Performed? *Applied Economic Perspectives and Policy* 39: 220–41. [[CrossRef](#)]
- Xiong, Tao, and Zhongyi Hu. 2021. Soybean Futures Price Forecasting Using Dynamic Model Averaging: Do the Predictors Change over Time? *Emerging Markets Finance and Trade* 57: 1198–214. [[CrossRef](#)]
- Xu, Chang, Bo Li, and Lingxian Zhang. 2022. Soybean price forecasting based on Lasso and regularized asymmetric  $v$ -TSVR. *Journal of Intelligent and Fuzzy Systems* 43: 4859–72. [[CrossRef](#)]
- Xu, Xiaojie, and Yun Zhang. 2022. Commodity price forecasting via neural networks for coffee, corn, cotton, oats, soybeans, soybean oil, sugar, and wheat. *Intelligent Systems in Accounting Finance and Management* 29: 169–81. [[CrossRef](#)]
- Xu, Zhiling, Hualing Deng, and Qiufeng Wu. 2021. Prediction of Soybean Price Trend via a Synthesis Method With Multistage Model. *International Journal of Agricultural and Environmental Information Systems* 12: 1–13. [[CrossRef](#)]
- Yin, Tao, and Yiming Wang. 2021. Nonlinear analysis and prediction of soybean futures. *Agricultural Economics-Zemledska Ekonomika* 67: 200–07. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.