Research on World Agricultural Economy
https://journals.nasspublishing.com/index.php/rwae

RESEARCH ARTICLE

Empirical Comparison of Facebook Prophet and Traditional Models for Tomato Price Forecasting in Greece

Eirini Kostaridou [1], Nikolaos Siatis [2], George Lampiris [2], Eleni Zafeiriou [1*]

[1] *Department of Agricultural Development, School of Agricultural Science and Forestry, Democritus University of Thrace, 68200 Orestiada, Greece*

[2] *Independent Researcher, 16344 Ilioupoli, Greece*

ABSTRACT

Agricultural product prices are crucial to the income and livelihoods of millions worldwide, making accurate forecasting essential for market participants. The European Union regularly reports agricultural product prices through an online databank, ensuring up-to-date price information is widely available. This study evaluates the effectiveness of Facebook Prophet, a modern forecasting tool, in predicting tomato prices in Greece from 2013 to 2024.The results show that Facebook Prophet outperforms traditional forecasting models, providing more accurate predictions with lower Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) values. Notably, the model exhibited superior performance in forecasting tomato prices over a six-month horizon compared to conventional seasonally adjusted models. This demonstrates Facebook Prophet's potential to significantly improve decision-making in agricultural markets, offering reliable price forecasts to stakeholders such as farmers, traders, and policymakers. Furthermore, the study incorporated feedback from market participants, which provided valuable insights into market practices and conditions. This integration of practical knowledge with advanced forecasting techniques enhanced the interpretation of the results, making them more applicable to real-world scenarios. Overall, the findings suggest that Facebook Prophet holds considerable promise for future agricultural price forecasting, with potential applications across various commodities. These insights pave the way for more precise

*CORRESPONDING AUTHOR:
Eleni Zafeiriou, Department of Agricultural Development, School of Agricultural Science and Forestry, Democritus University of Thrace, 68200 Orestiada, Greece; Email: ezafeir@agro.duth.gr

ARTICLE INFO
Received: 30 July 2024 | Revised: 26 August 2024 | Accepted: 30 August 2024 | Published Online: 9 December 2024
DOI: https://doi.org/10.36956/rwae.v5i4.1295

CITATION
Kostaridou, E., Siatis, N., Lampiris, G., et al., 2024. Empirical Comparison of Facebook Prophet and Traditional Models for Tomato Price Forecasting in Greece. Research on World Agricultural Economy. 5(4): 594–607. DOI: https://doi.org/10.36956/rwae.v5i4.1295

agricultural forecasts, benefiting all market participants by supporting more informed and timely decision-making.

Keywords: Tomato Price; Greece; Prophet; Market; Time Series; Forecast Models; SARIMA

1. Introduction

Tomatoes are a cornerstone of Greek agriculture, both economically and culturally, serving as a staple in the Mediterranean diet and a key ingredient in numerous traditional dishes. Greece's climate and geography make it an ideal location for tomato cultivation, which significantly contributes to both domestic consumption and export markets. This study investigates the price fluctuations and market trends of tomato production in Greece, informed by empirical data collected through interviews with producers and traders during the second quarter of 2024 [1].

The tomato holds a central place in Greek cuisine, forming the basis of dishes ranging from the iconic Greek salad to a variety of sauces and stews. Recognized for its health benefits, the Mediterranean diet emphasizes fresh vegetables, with tomatoes being a daily staple in Greek households. Consequently, fluctuations in tomato production or pricing can have wide-reaching effects on consumers, the food industry, and the agricultural economy at large [2].

Tomato cultivation in Greece is structured around two primary planting seasons, which align with the country's climatic conditions and agricultural practices. The first planting season occurs in October, predominantly in Crete, where traditional greenhouses are utilized. Despite being less technologically advanced, these greenhouses benefit from Crete's mild winter climate, allowing for a prolonged growing season that produces high-quality tomatoes during the off-peak months.

The second planting season begins in April, involving both open-field cultivation and the use of advanced, climate-controlled greenhouses. These greenhouses are primarily located in southern continental Greece, with significant operations also present in central and northern regions. This seasonal planting schedule closely mirrors those of major European Union (EU) and non-EU tomato producers, including Belgium, the Netherlands, Spain, and Italy. The synchronization of planting schedules across these regions is driven by the tomato plant's life cycle, local climatic conditions, and market-driven factors such as holidays and vacation periods, which create predictable peaks and troughs in supply and demand [3].

In Greece, tomato production is concentrated in two major regions: the Peloponnese peninsula in southern Greece (around 37° N latitude) and the island of Crete, just north of 36° N. The Peloponnese is characterized by a mix of traditional greenhouses and open-field cultivation, whereas Crete is renowned for its extensive greenhouse production. Central and northern Greece, while hosting fewer producers, are notable for their modern, technologically advanced greenhouses that significantly contribute to the country's overall tomato output [3].

As depicted in **Table 1**, the area of land dedicated to open-field tomato cultivation varies significantly over time. This variability is largely attributed to the relatively low investment required to switch between different types of vegetable crops in open fields. In contrast, greenhouse cultivation, which is capital-intensive, shows little variation in the area over time. However, switching between different crops within greenhouses is feasible and contributes to variations in cultivated areas. The decision to grow tomatoes in open fields is heavily influenced by the "expected" income from sales, which, in turn, affects the availability of the product during periods when open-field cultivation is possible due to favorable weather conditions. This trend is evident in the productivity data presented in **Table 1**.

Market participants, including producers and traders, indicate that the surplus of open-field tomatoes during the spring period (April to June), when most open-field production reaches the market, exerts significant downward pressure on prices. This seasonal oversupply contrasts with the more stable output from greenhouse production, which benefits from controlled growing conditions [2, 3]. Unfortunately, detailed monthly production data by cultivation method (open field vs. green-

Table 1. Evolution of open field cultivation versus greenhouse production for tomatoes in Greece.

Year	Open Field Tomatoes				Greenhouse Tomatoes			
	Area	Volume	Productivity	Overall Volume	Area	Volume	Productivity	Overall Volume
	(ha)	(tn)	tn/ha	%	(ha)	(tn)	tn/ha	%
2000	22.449	931.857	41.5	83%	3.742	186.371	49.8	17%
2001	21.840	969.205	44.4	78%	4.368	271.377	62.1	22%
2002	22.365	941.250	42,1	78%	3.550	263.550	74.2	22%
2003	21.120	895.000	42.4	83%	3.520	179.000	50.9	17%
2004	21.420	940.000	43.9	83%	3.570	188.000	52.7	17%
2005	18.000	835.000	46.4	83%	3.000	167.000	55.7	17%
2006	17.400	775.000	44.5	83%	2.900	155.000	53.4	17%
2007	16.080	711.950	44.3	83%	2.691	142.960	53.1	17%
2008	15.000	669.300	44.6	83%	2.520	134.931	53.5	17%
2009	16.320	822.550	50.4	83%	2.720	164.510	60.5	17%
2010	14.440	699.039	48.4	71%	2.568	288.191	112.2	29%
2011	9.759	349.882	35.9	54%	2.783	293.164	105.4	46%
2012	8.819	319.633	36.2	50%	2.804	325.762	116.2	50%
2013	9.019	325.924	36.1	47%	2.981	365.079	122.5	53%
2014	8.880	279.114	31.4	42%	3.064	384.190	125.4	58%
2015	6.425	236.075	36.7	41%	2.819	340.768	120.9	59%
2016	6.264	243.026	38.8	43%	2.591	321.721	124.2	57%
2017	6.597	232.942	35.3	43%	2.666	309.661	116.2	57%
2018	8.466	191.317	22.6	37%	2.690	327.072	121.6	63%
2019	7.896	169.923	21.5	37%	2.610	293.281	112.3	63%
2020	8.625	195.033	22.6	32%	2.716	405.988	149.5	68%
2021	4.974	184.512	37.1	32%	3.109	397.892	128.0	68%

* Source: Ministry of Agricultural development and Food.

house) are not available, which limits a more granular analysis of these effects.

The Greek tomato market is characterized by a high degree of fragmentation, with production scattered among numerous individual producers. Unlike in some other agricultural sectors, few tomato producers in Greece are unionized or engage in collective bargaining. This decentralized structure means that most significant growers market their products independently, often under their own brand names and through their own sales networks. This independence can lead to wide variations in pricing and marketing strategies, depending on the producer's capacity and market reach [4].

On the demand side, major retailers in Greece maintain their procurement processes, sourcing tomatoes either directly from large-scale producers or through intermediaries and traders. Despite the lack of a formalized, regulated market, tomato trading is monitored by government authorities, primarily through two major wholesale markets located in Athens and Thessaloniki. These markets play a crucial role in price discovery and distribution, with market prices being reported weekly. This pricing information is collected through a European-wide network and made available on an online platform, providing transparency and helping to stabilize the market.

Given the structure of the market, Greece's tomato industry does not operate in isolation. The country's production and pricing are influenced by trends in neighboring countries that are either major producers and exporters of tomatoes or those that require imports to meet domestic demand. Within the EU, Greece's primary trade partner in the tomato sector is Turkey. Although Greece is an EU member state, its proximity to Turkey and the latter's significant agricultural output makes it a key source of imports, especially during periods of short supply in Greece.

In terms of intra-EU trade, data availability is limited, but interviews with market participants suggest that a substantial percentage of tomato imports into Greece originate from the Netherlands and Belgium. These countries are known for their high-tech greenhouse production, allowing them to supply tomatoes year-round, often filling gaps in the Greek market during off-peak seasons or when domestic production falls short. The overall trade status from 2013 until 2024 is presented in **Table 2** and is indicative of the effect that seasonal production has on the overall performance of the cultivation of tomato in a macroeconomic level. The country is exporting large quantities of tomato but simultaneously imports significant quantities maintaining a trade surplus in terms of volume. At the same time a trade deficit is recorded for several years in values of goods traded. This may be attributed to the perishable nature of the product and limitations in bargaining among sellers and buyers of the product in the region.

Current research on agricultural prices is revied extensively in [5] where most available methods for forecasting are presented. A large portion of literature discusses the price asymmetry [6–8] observed between consumer and producer prices, mainly utilizing the macroeconomic indicators of PPI and CPI (producer price index and consumer price index) [4]. Empirical analysis is also utilized in some instances to capture the momentum of local markets [9, 10]. Recent studies have explored the use of seasonal time series and non-linear models for forecasting prices of similar crops, including perishable vegetables [11], as well as commodities that exhibit more stable price behaviors [12]. Understanding the roles of different players along the value chain in shaping the final price is essential when selecting the appropriate forecasting model and conducting price dynamics analysis. The influence of various stakeholders—such as traders, retailers, and consumers—on price formation varies depending on the characteristics of the crop. For goods that can be stored for extended periods, traders tend to have greater control over pricing, as they can regulate supply by adjusting inventory levels. Conversely, for highly perishable goods with shorter shelf lives, consumers and retailers play a more significant role in price determination, as immediate demand and quick turnover are critical. These differences in price formation dynamics also affect the choice of forecasting methodology, as noted in prior research [13].

The aim of this paper is to investigate the forecasting power of econometric models, well established in literature, and also evaluate the novel algorithms from Facebook Prophet as a quick and open-source alternative to complicated prediction software. The forecasts' performance against six monthly values actualized during the study period will also be discussed and the empirical aspects will be presented.

The conclusions of this research can be valuable for various market participants in assessing and refining their risk management practices. For farmers, while they may not directly utilize the findings due to limited access to or understanding of advanced analytical tools, they can benefit through their cooperatives or collective organizations, which can provide them with up-to-date price forecasts. This collective approach enables farmers to make more informed decisions about production and sales. For traders and retailers, the study's outcomes can support more effective procurement planning and better-informed contractual agreements, both upstream (with suppliers) and downstream (with retailers and consumers). Accurate price forecasts allow for improved inventory management, reducing the risk of overstocking or shortages. Finally, policymakers should carefully examine these forecasts, recognizing that macroeconomic issues such as trade deficits, if left unaddressed, can have long-term negative impacts on the market. Effective policies are needed to prevent market stagnation resulting from poor planning. These policies could include investments in or initiatives to expand controlled-environment agriculture (CEA), such as greenhouse farming, to reduce reliance on open-field cultivation. Such measures would help stabilize product availability, reduce seasonal price volatility, and ensure a more consistent supply to the market.

2. Data and Methodology

2.1. Data and Software

Data used in the present study were downloaded from the official European Union database for fruit and vegetables price data modified to a monthly time-series from January of 2013 to December of 2023, accessed in August 2024. Additionally, trade data were obtained by the same source's special page for tomato trade and include traded volumes and values as well as price per unit of traded goods [14]. Timeseries were downloaded and manipulated with standard spreadsheet software. All tomato price and trade data refer to "round tomatoes" as this is the most popular variety in Greece.

To analyze the complex interactions, the study will utilize the time-series data and apply econometric anal-

Table 2. Greek trade balance in volume and value.

	Import		Export		Trade Balance	
Year	**Volume (tns)**	**Value (€,000s)**	**Volume (tns)**	**Value (€,000s)**	**Volume (tns)**	**Value (€,000s)**
2013	8.125	4.558	20.667	11.594	12.542	7.036
2014	18.491	12.946	78.122	33.718	59.631	20.772
2015	12.350	7.918	35.704	13.384	23.354	5.466
2016	18.226	13.044	41.366	13.908	23.140	864
2017	18.057	13.280	36.779	12.619	18.722	−661
2018	29.774	22.603	37.691	12.744	7.917	−9.859
2019	17.897	14.989	35.548	10.844	17.651	−4.145
2020	13.625	10.793	36.371	10.884	22.746	91
2021	19.670	19.192	38.823	12.337	19.153	−6.855
2022	15.750	16.407	32.049	14.307	16.299	−2.100
2023	26.077	27.772	36.847	19.303	10.771	−8.469
2024	8.151	8.928	31.315	15.274	23.164	6.346

* Data supplied on request from the Greek Statistical Authority (www.statistics.gr). Year 2024 data are temporary.

ysis using EViews 12 software suite. Standard statistical tools will also be employed to identify patterns and correlations between various factors influencing the tomato market.

Additionally, for the forecasting of tomato price, Facebook Prophet was utilized to offer comparison among different forecasting methods.

2.2. Applied Methods

Prior to performing fitting of any of the forecasting methods discussed in detail below, the tomato price time series as published on agridata.eu databank, is evaluated. The descriptive statistics of the series are shown below in subsequent **Tables 3** and **4.**

The time series was found to be stationary on all tests performed. The stationarity tests for the time series are as follows:

Based on the analysis, stationarity for all time series is confirmed using the ADF and PP tests, whereas the KPSS test indicates stationarity only after taking the first differences of the original series and its log-transformed version. Additionally, a review of the correlogram for the tomato price (TP) time series suggests that the data exhibit seasonality.

2.2.1. SARIMA – Seasonal Autoregression Integrated Moving Average Model Forecasting

The ARIMA model can be extended to include seasonality, leading to the Seasonal ARIMA (SARIMA) or ARIMA(p,d,q)(P,D,Q)$_s$ model, where s represents the seasonality period, and P, D, Q are the seasonal counterparts to the autoregressive (AR), differencing (I), and moving average (MA) terms. The SARIMA model incorporates both non-seasonal and seasonal components [15, 16]. The notation ARIMA(p,d,q)(P,D,Q)$_s$ refers to the following:

p: the number of non-seasonal autoregressive terms
d: the number of non-seasonal differences to make the series stationary
q: the number of non-seasonal moving average terms
P: the number of seasonal autoregressive terms
D: the number of seasonal differences
Q: the number of seasonal moving average terms
s: the seasonal period (for example, s = 12s = 12s = 12 for monthly data with annual seasonality)

The non-seasonal part of the SARIMA model is the standard ARIMA model, which is expressed as:

$$\Delta^d y_t = c + \sum_{i+1}^{p} \phi_i \Delta^d y_{t-i} + e_t + \sum_{j=1}^{q} \theta_j e_{t-j} \quad (1)$$

Where:

ϕ_i are the coefficients of the autoregressive (AR) terms.
θ_j are the coefficients of the moving average (MA) terms.
d is the number of non-seasonal differencing terms to make the time series stationary.
e_t is white noise.

The seasonal part introduces additional AR, differencing, and MA components, but applied at the seasonal level. For a time-series with seasonality of period s (e.g., s = 12s = 12s = 12 for monthly data with annual season-

Table 3. Descriptive statistics of tomato price in Greece time series.

Timeseries	Mean	Median	Maximum	Minimum	St.Dev.	Skewness	Kyrtosis	Jarque-Bera	Probability
Tomato Price Greece	0.78	0.75	1.34	0.37	0.20	0.61	2.82	7.73	0.021

Table 4. Unit root test results.

t-Test Values	Augmented Dickey Fuler (ADF)	Philips Perron (PP)	Kwiatkowski–Phillips–Schmidt–Shin (KPSS)
TP	−5.596 (−2.884)	−5.625 (−2.884)	0.919 (0.463)
log(TP)	−5.786 (−2.884)	−5.828 (−2.884)	0.969 (0.463)
D(TP)	−8.026 (−2.884)	−31.524 (−2.884)	0.378 (0.463)
D(log(TP))	−7.878 (−2.884)	−32.516 (−2.884)	0.301 (0.463)

* numbers in parenthesis indicate 5% acceptance level.

ality), the seasonal ARIMA model is written as:

$$\Delta_s^D y_t = \sum_{k=1}^{P} \Phi_k y_{t-k\cdot s} + e_t + \sum_{l=1}^{Q} \Theta_l e_{t-l\cdot s} \quad (2)$$

Where:

- Φ_k are the seasonal autoregressive (SAR) coefficients.
- Θ_l are the seasonal moving average (SMA) coefficients.
- D is the number of seasonal differences to achieve stationarity at the seasonal level.
- s is the length of the seasonal cycle.

The seasonal differencing operator Δ_s^D is defined as: $\Delta_s^D y_t = y_t - y_{t-s}$.

If higher-order seasonal differencing is required (D > 1) $\Delta_s^D y_t = \Delta_s^{D-1} y_t - \Delta_s^{D-1} y_{t-s}$.

The complete SARIMA model combines both non-seasonal and seasonal components. It is written as:

$$\phi(B)\, \Phi(B^s)\, (1-B)^d (1-B^s)^D y_t = \theta(B)\, \Theta(B^s)\, e_t \quad (3)$$

Where:

- $\phi(B) = 1 - \phi_1 B - \ldots.. - \phi_p B^p$ is the non-seasonal AR part.
- $\Phi(B^s) = 1 - \Phi_1 B^s - \ldots.. - \Phi_p B^{sP}$ is the seasonal AR part.
- $\theta(B) = 1 + \theta_1 B + \ldots.. + \theta_q B^q$ is the non-seasonal MA part.
- $\Theta(B^s) = 1 + \Theta_1 B^s + \ldots.. + \Theta_: B^{s:}$ is the seasonal MA part.
- $(1-B)^d$ is the non-seasonal differencing operator.
- $(1-B^s)^D$ is the seasonal differencing operator.
- B is the backshift operator: $By_t = y_{t-1}, B^2 y_t = y_{t-2} \ldots\ldots B^s y_t = y_{t-s}$

Thus, SARIMA incorporates both short-term (non-seasonal) and long-term (seasonal) patterns by combining AR, MA, and differencing for both types of components [17, 18].

ARIMA model estimation of TP time series was performed with EViews 12 software suite. The built-in automatic ARIMA estimation was used to specify the AR and MA orders of the model, including seasonal components. The time series' correlogram was examined to assess stationarity and seasonality of the data in levels. Stationarity was achieved in 1st differences, but seasonality remained, according to the relevant correlogram. The hyperparameters were input to the software with maximum values of four for the AR and MA components, two for the respective seasonality factors with periodicity of 12 and maximum differencing of two for the original and logged values of the time-series. Selection criterion was KPSS with 5% threshold, with hyperparameters selection using the maximum likelihood estimation method. The total estimated models were 225 and the proposed model was SARIMA (0,1,2) × $(1,1,1)_{12}$, with the following estimation equation:

$$\left(1 - 0.981B^{12}\right) \Delta y_t = 0.003 + (1 - 0.498B - 0.502B^2)(1 + 0.882B^{12})e_t \quad (4)$$

This estimation was used to forecast the next 12 periods of TP (12 months of 2024), and the output and relative graph are shown in the results section.

2.2.2. Ratio to Moving Average Multiplicative Adjustment Method – Moving Average Model

The Moving Average model is a simple and widely used statistical approach for forecasting time series data which focuses on modeling the error terms (shocks or innovations) of a time series rather than the values themselves [19, 20]. It accounts for multiplicative seasonality where the seasonal variation changes with the trend, as appears to be the case of the original time-series. Data are decomposed into trend cycle, seasonal and shock or innovation components. The centered moving average (CMA) is calculated from the following equation and isolates the trend cycle component.

$$CMA\,(t) = \frac{y_{t-5} + y_{t-4} + \cdots + y_{t+6}}{12} \quad (5)$$

The trend cycle is removed by calculating the ratio of the reported tomato prices to the CMA and the seasonal components are isolated as follows:

$$Ratio\,(t) = \frac{y_t}{CMA\,(t)} \quad (6)$$

The software then automatically performs grouping and averaging of the seasonal indices and relative normalization for a monthly dataset, to construct the deseasonalized time-series. In our case, the original time-series was seasonally adjusted with simple moving average, as described above, then the first differences of the logged values were subjected to the process described in Section 2.2.1, but the selection method was now performed according to the least squares estimation. The estimated SARIMA model order is (2,1,4) × $(2,1,1)_{12}$, and the respective estimation equation is the following:

$$\begin{gathered}\left(1 - 0.748B^{12} + 0.087B^{24}\right)\\ \left(1 - 1.323B + 0.956B^2\right)\left(1 - B^{12}\right)\\ \left(1 - B\right)\Delta log y_t = 0.003 +\\ (1 + 1.818B - 1.355B^2 + 0.108B^3 +\\ 0.326B^4)(1 + B^{12})e_t\end{gathered} \quad (7)$$

The output and relative graph is shown in the results section.

2.2.3. TRAMO/SEATS

The TRAMO/SEATS model is a widely used method for time series decomposition and seasonal adjustment. It was developed by the Bank of Spain and stands for "Time Series Regression with ARIMA Noise, Missing Observations, and Outliers" (TRAMO) and "Signal Extraction in ARIMA Time Series" (SEATS). In the first step (TRAMO) the original time-series is modeled with an ARIMA process and then corrected for outliers, missing or seasonal data. In the SEATS process, the time-series produced on the TRAMO step, is decomposed into four components, trend, cycle, seasonality and noise,

$$\mathrm{Y}_t = T_t + C_t + S_t + I_t \quad (8)$$

Where:

T_t: is the trend component
C_t: is the cycle component
S_t: is the seasonal component
I_t: is the irregular component

Then using the state-space representation of the ARIMA model the components are estimated through filters. The equations used are:

$$Z_t = FZ_{t-1} + Ge_t \quad (9)$$

and

$$X_t = HZ_t + u_t \quad (10)$$

Where:

Z_t: is the state vector
F: is the state transition matrix
G: is the noise operator vector
H: is the observation matrix
e_t: is the noise from the ARIMA model
u_t: is the observation noise

Then the seasonality is removes by applying the filter to the seasonal component $\widehat{\mathrm{Y}}_t = \mathrm{Y}_t - S_t$ where $\widehat{\mathrm{Y}}_t$ is the final seasonal adjusted series.

With EViews 12, the performance of TRAMO/SEATS forecast is an automated process, which was implemented in our research.

In our case, tomato price was manipulated with TRAMO/SEATS followed by the estimation of a SARIMA model of the first differences of the logged values of the outcome time-series, with order (1,1,3) × $(1,1,1)_{12}$, with the maximum likelihood method and other parameters as described in Sections 2.2.1 and 2.2.2, and the respective estimation equation is:

$$\begin{gathered}\left(1 - 0.0559B^{12}\right)\left(1 + 0.965B\right)\left(1 - B^{12}\right)\\ \left(1 - B\right)\Delta log y_t = -0.003 + (1 - 0.563B -\\ 0.763B^2 + 0.430B^3)(1 + B^{12})e_t\end{gathered}$$

The next 12 monthly prices were forecasted for 2024, and the relevant values and graph are shown in the results section.

2.2.4. STL (Seasonal-Trend Decomposition Using Loess)

The STL model is a method for decomposing a time series into three components: seasonal, trend, and residuals [22–24] similarly to the above-mentioned principles. In this instance, the Loess process is implemented to assist on the decomposition of the time-series in its trend, seasonal and residual components.

$$\mathrm{Y}_t = T_t + S_t + R_t \tag{11}$$

Where:

T_t: is the trend component
S_t: is the seasonal component
R_t: is the residual component

Loess is an acronym for Locally Estimated Scatterplot Smoothing and algorithmically estimates a smoothed version of the time-series without the assumption of a specific model, but instead by focusing on the individual datapoints and their linear regression [15]. The datapoints are included in the estimation window for seasonality smoothing. The parameters inserted in EViews 12 dialog box were 35 for season, 19 for trend and 13 for the robustness parameter. Since STL cannot be used for forecasting, similarly to the processes demonstrated above, ARIMA [25] estimation is performed on the final, seasonally adjusted time-series with EViews 12 and the order of the model is (4,1,2), deriving from the process of model selection with maximum likelihood method estimation described previously for uniformity. The estimation equation is as follows:

$$\begin{gathered}\Delta Y_t = 0.003 + 1.161\Delta Y_{t-1} - \\ 0.634\Delta Y_{t-2} + 0.081\Delta Y_{t-3} - 0.215\Delta Y_{t-4} - \\ 1.650e_{t-1} + 0.864e_{t-2} + e_t\end{gathered} \tag{12}$$

The forecasted values and graph are shown in the results section.

2.2.5. Facebook Prophet

Facebook Prophet forecasts were performed in Python environment, with the multiplicative forecast option [26]. The software utilizes trend, seasonal effect and holiday data to scale up or down the effects of holidays depending on the underlying trend. This translates to seasonality or holidays having a proportional effect on the overall trend. For example, if the trend in tomato prices rises, the seasonal increase (such as a summer spike in prices) also grows proportionally larger. Conversely, when the trend is lower, the seasonal impact will diminish accordingly. The estimation is in the form of three time series combined with the following equation:

$$\mathrm{y}_{(\mathrm{t})} = \mathrm{T}_{(\mathrm{t})}\cdot(1 + \mathrm{S}_{(\mathrm{t})} + \mathrm{H}_{(\mathrm{t})}) + \epsilon_{\mathrm{t}} \tag{13}$$

The software estimates upper and lower boundaries for the forecast and plots the output as shown in the results section.

What makes Facebook Prophet different from the traditional time-series modeling – fitting and forecasting processes, is the presence of the Logistic Growth Model. According to the original white paper introducing the algorithm to the broader scientific community, this is a feature required from the original scope of the algorithm, which was to predict the number of users and other time-series related to the Facebook social network. In nature these numbers are usually limited by physical factors such as the total people living in a given area, or people with internet access etc.

$$g\left(t\right) = \frac{C}{1 + exp(-k(t-m)} \tag{14}$$

Where:

g(t) is the predicted value (growth) at time t,
C is the carrying capacity, which represents the upper limit or maximum possible value that the series can grow to,
k is the growth rate, which controls how quickly the growth happens,
t is time,
m is the offset parameter, which shifts the time at which the maximum growth rate occurs,
exp(−k(t−m)) is the exponential decay term.

In our study, the carrying capacity C can be interpreted as either the maximum "allowed" price of tomato or the minimum possible price of tomato, given that when the system reaches this capacity, the market forces explained previously will act to affect the direction and value of the growth rate of the time-series in question.

2.2.6. Forecasting Averaging

Two simple but universally accepted methods for averaging forecasts are employed: simple mean and trimmed mean [27]. The simple mean method calculates the arithmetic mean of all forecasts, without giving priority or weight to any of them. It is a simple yet effective method for combining forecasts when all forecasts are considered equally reliable.

$$\hat{y}_{mean} = \frac{1}{n}\sum\nolimits_{i=1}^{n}\hat{y}_i \tag{15}$$

$\hat{y}_i$ is the *i*-th forecast,
n is the number of forecasts.

The trimmed mean is a variation of the simple mean that aims to neutralize the effect of outliers that could influence the accuracy of the combined forecasts. This technique removes a certain percentage of extreme forecasts from both the high and low ends before calculating the mean.

$$\hat{y}_{trim} = \frac{1}{n-2k}\sum\nolimits_{i=k+1}^{n-k}\hat{y}_i \tag{16}$$

$\hat{y}_i$ is the *i*-th forecast (after sorting the forecasts)
n is the total number of forecasts
k is the number of forecasts trimmed from both the top and bottom, based on the trimming percentage [28].

In the present research, only the highest and lowest values forecasted for the TP time-series were excluded for the estimation of the trimmed mean value.

2.2.7. Forecast Evaluation

Three simple methods for the evaluation of the forecast models were employed. During the study in the beginning of 2024, six monthly prices of tomato in Greece materialized and they are used to evaluate the performance of the tested models. First, the Root Mean Square Error (RMSE) is calculated between actual and forecasted values of each forecasting method.

$$RMSE = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(y_1 - \hat{y}_i)^2} \tag{17}$$

This simple calculation offers an easy to compare measure for each forecast method employed. Secondly, Mean Absolute Error (MAE) is demonstrated as the mean of the absolute differences between the actual values and the predicted values according to the formula

$$MAE = \frac{1}{n}\sum\nolimits_{i=1}^{n}|y_i - \hat{y}_{i|} \tag{18}$$

Finally, the Mean Absolute Percentage Errors (MAPE) method is applied by applying the following formula to all forecast model outputs.

$$MAPE = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \cdot 100 \tag{19}$$

The model with the lowest among the calculated values is considered the best performing and, the respective evaluation results with all above methods and for each forecasting process are shown in **Table 5**.

Table 5. Values of adjusted R^2 and Akaike Information Criteria for the developed models.

	Adjusted R^2	**AIC**
SARIMA	0.383	−0.934
MOVING AVERAGE	0.369	−0.741
TRAMO/SEATS	0.362	−0.844
STL	0.229	−1.229
PROPHET	0.533	

3. Results and Discussion

3.1. Results

The results of the EViews 12 forecasting process are shown below for the price of tomato in Greece. Following the steps described above in the methodology section, the forecasted values and graphs are presented in **Figures 1–5**.

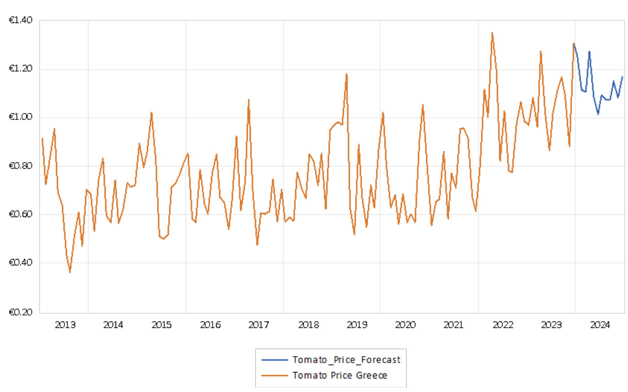

Figure 1. SARIMA model forecast graph.

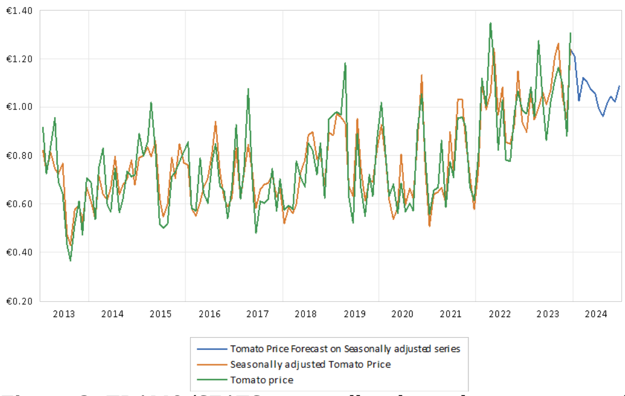

Figure 2. TRAMO/SEATS seasonally adjusted time series and forecast graph.

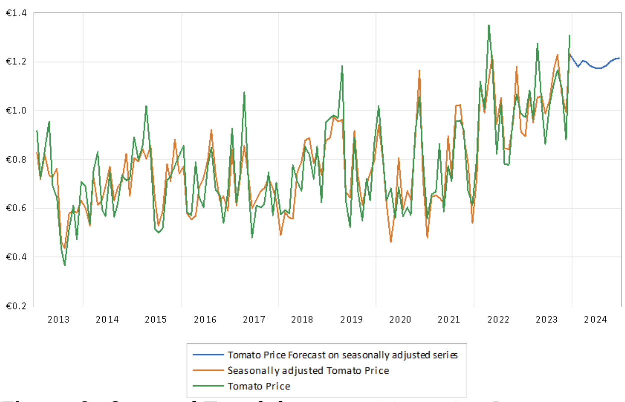

Figure 3. Seasonal-Trend decomposition using Loess seasonally adjusted time series and forecast graph.

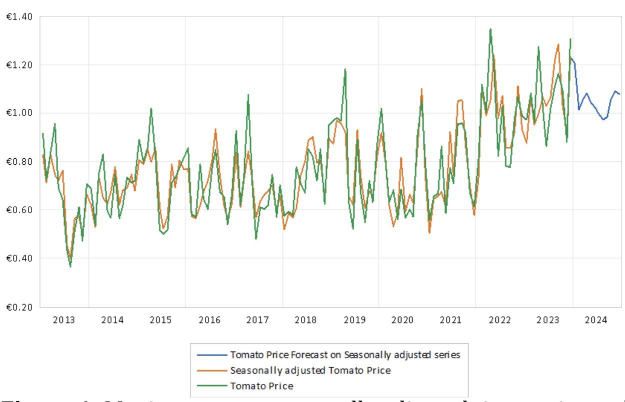

Figure 4. Moving average seasonally adjusted time series and forecast graph.

The initial evaluation of the models is done by comparing the resulting adjusted R^2 values for each individual model. The models fitted with EViews 12 the values are automatically reported after model fitting, but for FBPROPHET the value was calculated after the report using standard spreadsheet applications. The reported values indicate a better performance for FBPROPHET which presents the highest adjusted R^2 value, as shown in **Table 5**. An indication of model performance is also evident by the values of Akaike Information Criteria shown for the ARIMA based models. For FBPROPHET this criterion is non-applicable.

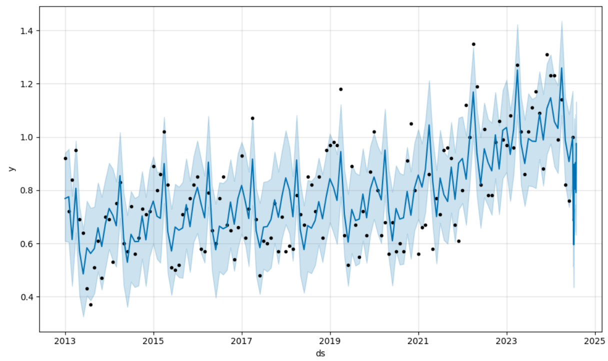

Figure 5. Facebook Prophet forecast of time series values.

The output of the employed forecasting models is shown in **Table 6**. The table is populated with the tomato prices made available for the first semester of 2024, along with the estimated prices from various forecasting models.

Although informative, the table above does not provide clear information on the performance of various models, therefore **Table 7** is constructed to demonstrate the results of the comparison methods based on squared differences and absolute variation between forecasted and actual tomato prices.

3.2. Discussion

The employed forecasting models' evaluation demonstrates prophet algorithm superiority for the prediction of tomato price in Greece, even against combined forecasts with simple and trimmed means in which the prophet predictions participate. This observation can be explained by the built-in information regarding the behavior of the time series regarding seasonality and trend along with the multiplicative principle.

The reliability of the mathematical forecasting as demonstrated is limited by the information included in the time series. However, the availability of market data affecting the price is limited, however estimated to be very high. As per the interviews with market participants, the area of tomato grown in open fields is deter-

Table 6. Forecast models' outputs and actualized values.

	TRAMO	STL	Moving Average Model	ARIMA	FBPROPHET	Simple Mean	Trimmed Mean	6M2024 Actuals
Jan-24	1.21 €	1.20 €	1.21 €	1.26 €	1.15 €	1.21 €	1.21 €	1.23 €
Feb-24	1.03 €	1.18 €	1.02 €	1.11 €	1.06 €	1.08 €	1.07 €	1.23 €
Mar-24	1.12 €	1.20 €	1.06 €	1.11 €	1.03 €	1.10 €	1.10 €	0.99 €
Apr-24	1.11 €	1.20 €	1.08 €	1.27 €	1.26 €	1.18 €	1.19 €	1.14 €
May-24	1.07 €	1.18 €	1.04 €	1.09 €	0.99 €	1.07 €	1.05 €	0.82 €
Jun-24	1.05 €	1.17 €	1.02 €	1.01 €	0.91 €	1.03 €	1.03 €	0.76 €
Jul-24	0.99 €	1.17 €	1.00 €	1.09 €	1.00 €	1.05 €	1.04 €	
Aug-24	0.96 €	1.18 €	0.97 €	1.07 €	0.97 €	1.03 €	1.02 €	
Sep-24	1.02 €	1.19 €	0.98 €	1.07 €	0.86 €	1.02 €	1.02 €	
Oct-24	1.04 €	1.20 €	1.05 €	1.15 €	0.75 €	1.04 €	1.08 €	
Nov-24	1.02 €	1.21 €	1.09 €	1.08 €	0.69 €	1.02 €	1.06 €	
Dec-24	1.08 €	1.21 €	1.08 €	1.17 €	0.68 €	1.05 €	1.12 €	

Table 7. Performance comparison of forecast models.

Comparison Method	TRAMO	STL	Moving Average Model	ARIMA	FBPROPHET	Simple Mean	Trimmed Mean
RMSE	0.18	0.24	0.17	0.17	0.13	0.18	0.17
Absolute differences	0.02	0.03	0.02	0.03	0.08	0.03	0.02
	0.20	0.06	0.22	0.12	0.18	0.16	0.17
	0.13	0.21	0.07	0.12	0.04	0.11	0.10
	0.03	0.05	0.06	0.13	0.12	0.04	0.05
	0.25	0.36	0.22	0.27	0.16	0.25	0.22
	0.29	0.42	0.26	0.25	0.15	0.27	0.27
MAE	0.15	0.19	0.14	0.15	0.12	0.15	0.14
MAPE	17%	22%	15%	17%	12%	15%	14%

mined by individual growers and cannot be monitored in real time but is only reported on a yearly basis by the authorities. In other words, the availability of product in. i.e., May or June, is determined when the decision to grow tomatoes in open fields by thousands of growers in March. It is worth mentioning that the elasticity on cultivated area of open field tomatoes is higher than this of greenhouse tomatoes as shown in the data of **Table 1**. Having not included this information to a forecast model limits the possibility of a reliable forecast. On the other hand, since growers take into consideration the previous prices observed, in the form of Price Expectations, the lags or previous prices are expected to affect their decision, and by including these prices in a forecasting model inserts this information to the model. Similarly, greenhouse growers can switch among some options, i.e., tomato varieties, peppers or cucumbers depending on their Price expectations.

Another parameter that can limit the effect of open field tomatoes on the overall price is of course climate. Even if the decision to grow tomatoes in open fields is taken by a significant number of growers bounding area for tomato cultivation, the produce of this area is very influenced by the climate conditions in the coming months. Hence, if the weather is favorable, the produce will be higher and the supply of product in the market will be increased pushing the prices to the threshold. If the weather conditions prove unfavorable, the produce will be limited and thus the supply will be limited pushing the price to the ceiling.

The threshold and ceiling prices are dictated by several factors as well as the other factors discussed. The threshold price is estimated to be the minimum price for which the product will leave the producer covering harvest and transportation cost. This price is affected by the overall economic environment, labor costs and availability, energy prices etc. The ceiling price is the maximum price a trader or retailer is willing to pay the producer before they seek alternatives in EU or non-EU suppliers. These two prices are not known or forecasted as they relate to other markets and macroeconomic factors, thus cannot be included in the forecast. Again, by utilizing time series, some of this information is included in the previous prices and are included in the trends of the time series of the price of tomato, but their effect cannot be decomposed and studied individually.

The graph below (**Figure 6**), provides a valuable illustration of the relationship between import volumes, import prices, and the cumulative price of tomatoes in Greece from 2013 to mid-2024. It reveals that during periods of high import volumes, the import price tends to act as a ceiling for the overall price, often crossing above it. Another notable observation is that import volumes increase when import prices are low, while they drop to nearly zero when import prices exceed domestic prices. This pattern highlights the strong interconnection between domestic tomato prices and those in external markets, suggesting that price movements in connected markets significantly influence local pricing dynamics.

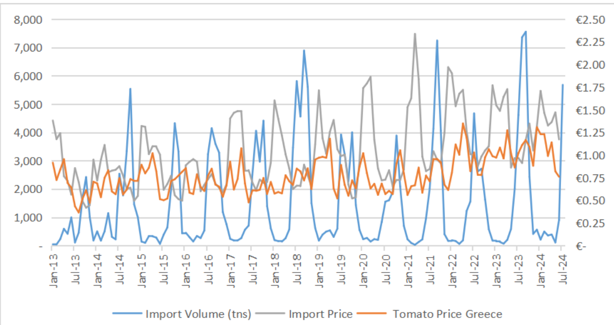

Figure 6. Tomato volume imports and prices.

* Source: Statistical authority of Greece, data on demand (www.statistics.gr).

Finally, there is a discrimination between competitive products inflows from different origins. As per the market participants, the substitute of open field tomatoes produce are mainly neighboring open field products mainly from Turkey, whereas for the greenhouse products the substitutes are mainly EU producers of greenhouse products. This is in part because of the established procurement networks of retailers buying mainly greenhouse products with access to alternative EU suppliers. An overview of the market and climate forces impact on tomato price based on empirical data is shown in **Table 8.**

The Facebook Prophet algorithm's ability to incorporate the Logistic Growth Model enhances its forecasting process by better capturing dynamic factors that are often qualitative in nature and difficult to quantify. These factors, such as market saturation or shifts in growth rates, are critical in real-world scenarios but are often excluded from traditional models due to their complexity. Prophet addresses this by allowing flexible growth modeling through parameters like carrying capacity, growth rate, and changepoints, enabling the model to reflect non-linear growth patterns more accurately. The interconnected nature of the markets and characteristics of the product (such as perishability and dispersed production locations) shift price determination factors toward the consumer and retailer side. This makes the forecasting model more aligned with the target market for which the Facebook Prophet algorithm was originally designed. While this novel algorithm may not be expected to outperform traditional models in forecasting other types of agricultural products, such as commodities, further research is needed to explore its potential in this area. However, the inclusion of these qualitative factors into forecasting models poses significant challenges. These aspects, often derived from empirical knowledge, are not easily measured or integrated into a formal model. Future studies should focus on refining methods for incorporating such parameters, potentially leading to more accurate forecasts. By doing so, stakeholders in various markets—such as business analysts, economists, and policymakers—could benefit from improved forecasting precision, enabling better decision-making and resource allocation.

Table 8. Forces influence on production and price of tomato in Greece.

	Open Field Production	**Greenhouse Production**
Climate	High	Medium
Price expectation	High	Medium
Intra EU trade price	Low	High
Extra EU price	High	Medium
Macroeconomic factors	High	High

Author Contributions

Conceptualization, E.K., N.S.; methodology, E.K., N.S.; software, N.S.; validation, E.Z., G.L.; formal analysis, E.K., N.S., E.Z., G.L.; investigation, E.Z.; resources, E.Z., E.K. and N.S.; data curation, E.K., N.S.; writing—original draft preparation, E.K., N.S., G.L.; writing—review and editing all authors supervision, E.K., N.S.; project administration,

all authors. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Ethical review and approval were waived for this study as not applicable for studies not involving humans or animals.

Informed Consent Statement

Not applicable.

Data Availability Statement

Authors should state where data supporting the results reported in a published article can be found, and under what conditions the data can be accessed. They also include links (where applicable) to the data set.

Acknowledgments

The authors wish to express their gratitude to the market participants who voluntarily shared their insights on the formulation of tomato prices in Greece. The participants have chosen to remain anonymous by their own preference.

Conflicts of Interest

All authors disclosed no conflict of interest.

References

[1] Alamanos, E., Bourlakis, M., Tzimitra-Kalogianni, I., 2013. Segmenting Greek tomato consumers: Policy and marketing insights towards a healthy diet. British Food Journal, 115(4), 488–507. DOI: https://doi.org/10.1108/00070701311317801

[2] Michalis, E., Giatra, C.E., Skordos, D., et al., 2023. Assessing the Different Economic Feasibility Scenarios of a Hydroponic Tomato Greenhouse Farm: A Case Study from Western Greece. Sustainability, 15(19), 14233. DOI: https://doi.org/10.3390/su151914233

[3] Dimokas, G., Tchamitchian, M., Kittas, C., 2009. Calibration and validation of a biological model to simulate the development and production of tomatoes in Mediterranean greenhouses during winter period. Biosystems Engineering, 103(2), 217–227. DOI: https://doi.org/10.1016/j.biosystemseng.2009.01.004

[4] Rezitis, A.N., Pachis, D.N., 2016. Investigating the Price Transmission Mechanisms of Greek Fresh Potatoes, Tomatoes and Cucumbers Markets. Journal of Agricultural and Food Industrial Organization, 14(1), 91–108. DOI: https://doi.org/10.1515/jafio-2015-0004

[5] Sun, F., Meng, X., Zhang, Y., et al., 2023. Agricultural Product Price Forecasting Methods: A Review. Agriculture, 13(9), 1671. DOI: https://doi.org/10.3390/agriculture13091671

[6] Girapunthong, N., Vansickle, J.J., Renwick, A., 2023. Price Asymmetry in the United States Fresh Tomato Market. Journal of Food Distribution Research, 34(3), 51–59.

[7] Anastasiadis, F., Apostolidou, I., Michailidis, A., 2020. Mapping Sustainable Tomato Supply Chain in Greece: A Framework for Research. Foods, 9(5), 539. DOI: https://doi.org/10.3390/foods9050539

[8] Puria, P., Singh, K.M., Parihar, L., 2023. Decomposition and forecasting of tomato price in Bhopal district, Madhya Pradesh. Agricultural Economics Research Review, 36, 111–122. DOI: https://doi.org/10.5958/0974-0279.2023.00020.4

[9] Epperson, J.E., Fletcher, S.M., Collins, M.F., 1987. Price forecasting and trigger price probability estimation. International Journal of Forecasting, 3(2), 281–287.

[10] De Gooijer, J.G., Hyndman, R.J., 2006. 25 years of time series forecasting. International Journal of Forecasting, 22(3), 443–473. DOI: https://doi.org/10.1016/j.ijforecast.2006.01.001

[11] Giri, A., Giri, V.R., 2024. Forecasting cauliflower prices in Nepal: A comparative analysis using seasonal time series and nonlinear models. Cogent Food and Agriculture, 10(1), 2340155. DOI: https://doi.org/10.1080/23311932.2024.2340155

[12] Majhi, S.K., Bano, R., Srichandan, S.K., et al., 2023. Food price index prediction using time series models: A study of cereals, millets, and pulses. DOI: https://doi.org/10.21203/rs.3.rs-2999898/v1

[13] Jayapal, S., Muntean, C.H., 2023. Food Demand Prediction using Statistical and Machine Learning Models MSc Research Project Data Analytics. Available from: https://norma.ncirl.ie/6594/1/sasiku

marjayapal.pdf
[14] European Commission, 2024. Tomato Trade Dashboard. Available from: https://agridata.ec.europa.eu/extensions/DashboardTomato/TomatoTrade.html
[15] Falatouri, T., Darbanian, F., Brandtner, P., et al., 2022. Predictive analytics for demand forecasting – A comparison of SARIMA and LSTM in retail SCM. Procedia Computer Science, 200, 993–1003. DOI: https://doi.org/10.1016/j.procs.2022.01.298
[16] Chatfield, C., 2003. The Analysis of Time Series. Chapman and Hall/CRC: London, UK. DOI: https://doi.org/10.4324/9780203491683
[17] Damrongkulkamjorn, P., Churueang, P., 2023. Monthly energy forecasting using decomposition method with application of seasonal ARIMA. In Proceedings of the 2005 International Power Engineering Conference; Singapore, 29 November–2 December 2005.
[18] Adanacioglu, H., Yercan, M., 2016. An analysis of tomato prices at wholesale level in Turkey: An application of SARIMA model. Available from: www.custoseagronegocioonline.com.br
[19] Johnston, F.R., Boyland, J.E., Meadows, M., Shale, E., 1999. Some properties of a simple moving average when applied to forecasting a time series. Journal of the Operational Research Society, 50(12), 1267–1271. DOI: https://doi.org/10.1057/palgrave.jors.2600823
[20] Akpinar, M., Yumusak, N., 2016. Year ahead demand forecast of city natural gas using seasonal time series methods. Energies, 9(9), 727. DOI: https://doi.org/10.3390/en9090727
[21] Gómez, V., Maravall, A., 1997. Programs TRAMO and SEATS: instructions for the user. Banco de España.
[22] Theodosiou, M., 2011. Forecasting monthly and quarterly time series using STL decomposition. International Journal of Forecasting, 27(4), 1178–1195. DOI: https://doi.org/10.1016/j.ijforecast.2010.11.002
[23] Ouyang, Z., Ravier, P., Jabloun, M., 2021. STL Decomposition of Time Series Can Benefit Forecasting Done by Statistical Methods but Not by Machine Learning Ones. Engineering Proceedings. 5(1), 42. DOI: https://doi.org/10.3390/engproc2021005042
[24] Cleveland, W.S., 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Association, 74(368), 829–836. DOI: https://doi.org/10.1080/01621459.1979.10481038
[25] Hyndman, R.J., Athanasopoulos, G., 2021. Forecasting: Principles & Practice. 3rd ed. Otexts: Melbourne, Australia.
[26] Taylor, S.J., Letham, B., 2023. Forecasting at Scale. DOI: https://doi.org/10.7287/peerj.preprints.3190v2
[27] Blanc, S.M., Setzer, T., 2016. When to choose the simple average in forecast combination. Journal of Business Research, 69(10), 3951–3962. DOI: https://doi.org/10.1016/j.jbusres.2016.05.013
[28] Wichern, D.W., Makridakis, S., Wheelwright, S.C., 1979. Forecasting: Methods and Applications. Journal of the American Statistical Association, 74(367), 733. DOI: https://doi.org/10.2307/2287014