Literature Review - Forecasting Food Production and Sales

September 2022

Author(s):
Isabell Galda, Timur Kayhan, University of Duisburg-Essen (Germany),
Chair of Business Administration, especially Service Operations

Contents

List of Figures

List of Tables

1 Introduction

Forecasting is an essential application in supply chain management. Based on past and present data and trends, one can anticipate good's production or sales. There are a lot of mathematical methods available to calculate forecasts. This report will introduce the most prominent ones in relevant papers, define them, and give the reader a better understanding of forecasting in general. Finally, the benefits and the importance of potential applications for ImPUlSe shall be highlighted.

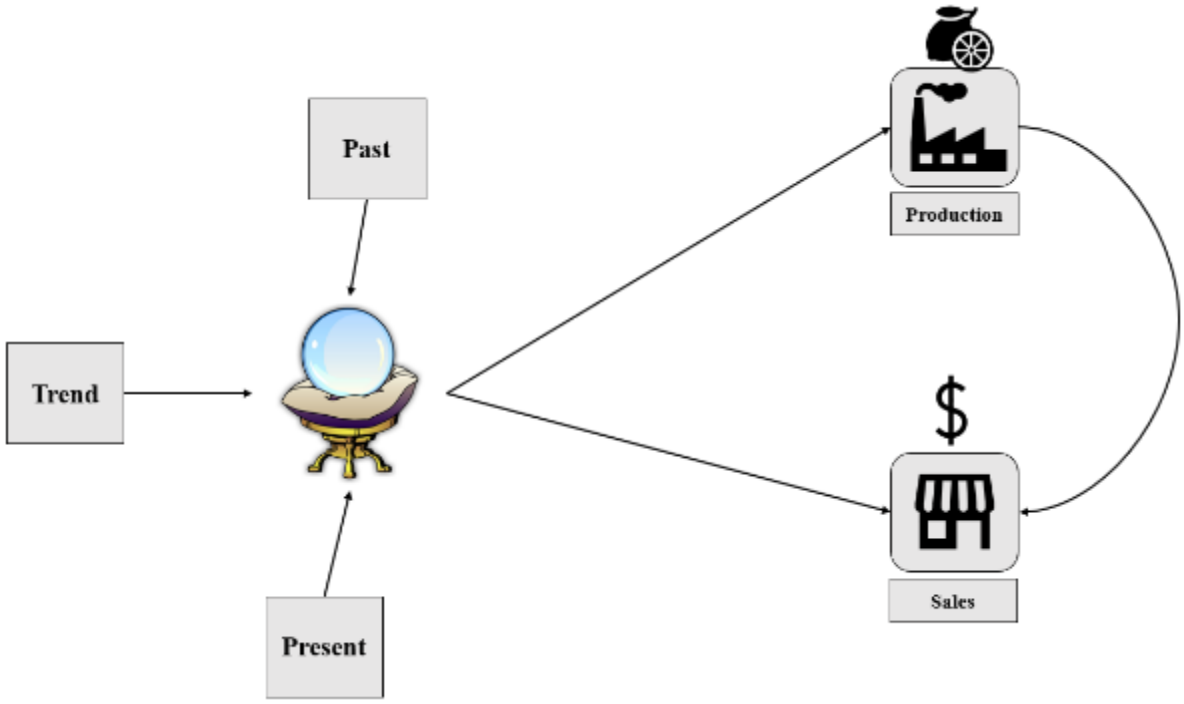

Figure 1: Illustration of the forecasting process for the citrus production (own figure)

Recently, many papers have dealt with the importance of forecasting for various fruits. According to Khan et al. (2020), improving forecasting methods highly benefits developing countries with an increasing population. Yusuf and Salau (2007) state that citrus production is continuously rising due to various health benefits of consuming said fruits. There are also a lot of industrialized countries where fresh orange juice is an integral part of the daily diet. The project ImPUlSe especially concerns countries like Algeria, Egypt, and Tunisia. Farmers can focus on high-demanded crops with low production if predictions are improved. In consonance with Khan et al. (2020), a country's government also has a vital role in the process. By improving said methods, they can support the farmers more efficiently (e.g., by providing loans and new government policies for higher production). In this context, predicting sales and production is imperative to improve the overall supply chain performance.

From the explanations above, it becomes clear that applying forecasting to the prediction of sales and production for perishables highly meets the project targets of ImPUlSe. Forecasting

has the potential to optimize the efficiency of physical supply and consumption patterns. The supply chain can be managed in a more sustainable way and contributes to the reduction of waste.

The following paper is organized as follows: After a brief introduction to the general concept of forecasting and its connection to the ImPUlSe project, section 2 offers a literature review on forecasting with a focus on the FSC. The goal is to offer a comparison of current practices and possible improvements by implementing new pricing strategies.

Section 3 summarizes and concludes the literature review.

2 Literature Review

The literature on forecasting in the fruit production and sales can be categorized into two groups. The first group includes papers that apply regression analysis. This group contains the majority of the papers. The second group consists of papers that apply ARIMA, the auto-regressive integrated moving average.

In the first section of this chapter regression analysis applying papers are focused and the second section of this chapter reviews papers that apply ARIMA.

The papers are analyzed and compared within these sections regarding their research object, applied methods, results and managerial insights.

2.1 Regression analysis

Yusuf and Salau (2007) use regression analysis to forecast the mango and citrus yield and plant area in Nigeria. Three different periods and their data are used for forecasting. 1961 to 2003 contains all periods, 1986 to 2003 is the Structural Adjustment Period and 1991 to 2003 is the Liberalization period. These periods represent the difference in the political framework in Nigeria. The authors assume that normal weather conditions and past trends regarding yield and plant area continue.

The results show that the citrus yield increased between 1988 and 1989. The plant area increased from 30000 hectares in 1961 to 72000 hectares in 1989. The output of citrus production increased by 200 percent. The mango plant area increased from 3200 hectares to 12000 hectares from 1961 to 1998.

Based on these data the coefficients of the regression analysis are estimated. During the Structural Adjustment period, mango production has on average increased by four percent and citrus production by 3.4 percent. During the Liberalization period, the increase is estimated to be 2.1 percent and 2.9 percent for citrus products and mangos, respectively. For all periods, the increase is on average 3.16 percent and 3.18 percent for citrus fruits and mango. The authors conclude that citrus and mango production will continue to increase in the future.

Brinkhoff and Robson (2021) compare different regression models to forecast macadamia nut yield in different regions in Australia. The data are collected from 2012 to 2019. To forecast the yield of the current year, the data from the last two years is used, therefore the forecast years are 2014 to 2019. The authors collect spatiotemporal and meteorological data and apply different regression-solving methods: OLS, Ridge, Lasso, Random Forest, and Support Vector Regression. In addition, the most influential variables are identified; meteorological data are not included. The results show that a Ridge Regression Model is the most suitable model to forecast the macadamia yield. The measurement error is less than 15 percent.

Maskey et al.'s (2019) paper deal with the strawberry yield forecast in California. The authors first investigate the correlation between weather parameters, such as soil moisture and temperature, and yield. The correlated parameters are used to develop three forecast models: Principal Component Regression Analysis and two machine learning models – Neural Network and Random Forest models. The data on yield were collected between February and June 2018. The results show that the temperature-related weather data have the strongest correlation to yield. Regarding the forecast models, the regression analysis forecast was the least exact and the adjusted R^2 was the lowest. Maskey et al. argue that the relationship between the input parameters is not linear. The machine learning algorithms perform better because the adjusted R^2 and predicted yields were the highest and most exact, respectively. Within the two algorithms, the Neural Network is most suitable for forecasting.

Achmakh et al. (2020) use multiple regression analysis to determine olive production in Morocco. The input parameters are the amount of pollen in the air and meteorological data before and during the flowering period, e.g., temperature and precipitation. The data were collected between 2008 and 2018. Based on these data different forecast models were created. The authors proved that there is a correlation between the amount of pollen in the air and olive output. The model with the highest R^2 of 0.98 and the lowest Root Mean Square Error showed that the minimum temperature in July and the rainfall in March have the largest positive impact on olive production. The minimum temperature in June and autumn rainfall have a strong negative influence on the olive output.

2.2 ARIMA

Bayav and Centinbas (2021) use ARIMA to estimate peach production in Turkey for the years 2021 to 2023 based on data from 1971 to 2020. Apart from that, the change in production quantity and production area from 2010 to 2019 are compared for different countries. In Turkey, the production amount increased by 53.98 percent and the production area by 7.41 percent. In some countries, the production amount increased, but the production area decreased. The authors conclude that high-density planting activities and better production techniques are the main cause of this development. It was estimated that the peach production amount in 2023 will be 893.575 tons.

Eyduran et al. (2020) use the 1961 to 2015 harvest area and production amount data for bananas in Turkey to forecast both for the years from 2016 to 2025. The authors conduct a time series analysis. They find a trend in the data and remove it to get stationary data. Then, different ARIMA models and exponential smoothing methods are applied to forecast the production amount and harvest area. The results show that ARIMA models are not suitable in this case. Brown exponential smoothing was found to be the most suitable forecast model. In

2016, the harvest area was measured to be 6175 hectares, in 2025 the area will be 9733 hectares. Concerning the banana production amount, 291,667 tons were calculated for 2016 and 482,093 tons for 2025.

Rueangrit et al. (2020) forecast the production and export amount of durian fruits in the market of Thailand. The harvest of the durian fruit is influenced by seasonal effects, which must be considered when applying forecast models. The underlying time series consists of the monthly data for durian production and export from 2005 to the first half of 2020. The goal is to forecast the production amount and the export amount to both the Chinese market and the rest of the world for the second half of 2020 until December 2021. Several seasonal ARIMA (SARIMA) models are used for forecasting. The production amount is expected to increase by 2.419 percent for the second half of 2020 and by 13.48 percent in the year 2021. The exports to China are expected to increase by 43.398 percent for the second half of 2020 and by 31.299 percent in 2021. The exports to the markets in the rest of the world are expected to increase by 0.542 percent and 6.023 percent for the second half of 2020 and 2021, respectively. Overall, the forecast performance of the models ranges from 78.095 to 88.654 percent.

3 Conclusion

The goal of this report was to review the relevant literature on forecasting in the FSC. Table 1 gives an overview of all considered papers. The reviewed papers were categorized in applying regression analysis and ARIMA. The articles were reviewed in terms of the research object, applied methods, results and managerial insights.

Author	Industry	Region	Products	Forecast	Applied methods
Achmakh, L., Janati, A., Boullayali, A., El Hassani, L., Bouziane, H.	Agriculture	Morocco	Olives	Production	Multiple regression analysis
Bayav, A., Cetinabas, M.	Agriculture	Turkey	Peach	Production	1. Box-Jenkins-method (ARIMA Modell) 2. Competition analysis in comparison with other nations (e.g. Spain, Greece)
Brinkhoff, J., Robson, A. J.	Agriculture	Australia	Macadamia	Yield	Multiple regression analysis and correlation analysis
Eyduran, S. P., Akın, M., Eyduran, E., Çelik, Ş., Ertürk, Y. E., Ercişli, S.	Agriculture	Turkey	Banana	Production	1. Moving Average (autoregressive) 2. Exponential smoothing (Holt) 3. Exponential smoothing (Brown) 4. Exponential smoothing (Damped)
Khan, T., Qiu, J., Ali Qureshi, M. A., Iqbal, M. S., Mehmood, R., Hussain, W.	Agriculture	Pakistan	Diverse - Apples - Bananas - Pears - Citrus fruits - Grapes	Production	1. Levenberg-Marquardt-algorithm (66,5 % accuracy) 2. Scaled Conjugate Gradient Back Propagation (70,2 % accuracy) 3. Bayesian regularization back propagation (76,3 % accuracy)
Maskey, M. L., Pathak, T. B., Dara, S. K.	Agriculture	USA	Strawberry	Yield	1. Regression analysis and correlation analysis 2. Principal component regression 3. Neuronal network 4. Random Forest
Rueangrit, P., Jatuporn, C., Suvanvihok, V., Wanaset, A.	Agriculture	Thailand	Durian fruits	Production	Box-Jenkins-method (ARIMA model)
Ullah, A., Khan, D., Zheng, S.	Agriculture	Pakistan	Peach	Production	Box-Jenkins-method (ARIMA model)
Wen, Q., Mu, W., Sun, L., Hua, S., Zhou, Z.	FMCG	China	Grapes	Sales	Support Vector Machine 1. ε -SVR model 2. LS-SVR model
Yusuf, S. A., Salau, A. S.	Agriculture	Nigeria	Citrus fruits and mangos	Production	Trend analysis via regression analysis -Structural adjustment period (1986 - 2003) - Liberalization period (1991 - 2003)

Table 1: Overview of all considered papers (own figure)

The research objects were different kinds of fruits, like mango, citrus, strawberry, olive, peach, banana, durian fruit and macadamia nuts. All papers forecasted the production or harvest amount, three papers also forecasted the plant area of the fruits.

Concerning the solution methods, regression analysis was applied slightly more often than ARIMA. When deciding about the applied methods, two aspects are important. First, linear regression models may not be suitable if the relationship between the variables is not linear. In this case, other methods are more suitable. Second, when applying ARIMA, the stationarity of the time series data must be ensured.

The results generally indicate that the plant area and production amount of the investigated products will increase in the future due to a growing domestic and international demand for fresh fruits.

This literature review shows the importance and potential application in the ImPUlSe project. One paper was identified that forecasted the citrus plant area and production amount using regression analysis. Furthermore, two papers focused on Turkey as a research object. As part of the project, the harvest area, production, and export amount of citrus fruits could be focused on, also using ARIMA as an alternative method. These analyses could be conducted in the other target countries of ImPUlSe, e.g., Tunisia or Egypt.

References

Achmakh, L., Janati, A., Boullayali, A., El Hassani, L., Bouziane, H. (2020), Forecasting olive (Olea europaea L.) production using aerobiological and meteorological variables in Tétouan (NW Morocco). Aerobiologia, 36(4), pp. 749–759.

Bayav, A., Cetinabas, M. (2021), Peach Production and Foreign Trade of Turkey: Current Situation, Forecasting and Analysis of Competitiveness. ANADOLU Ege Tarımsal Araştırma Enstitüsü Dergisi, 31(2), pp. 212–225.

Brinkhoff, J., Robson, A. J. (2021), Block-level macadamia yield forecasting using spatio-temporal datasets. Agricultural and Forest Meteorology, 303, pp. 108369.

Eyduran, S. P., Akın, M., Eyduran, E., Çelik, Ş., Ertürk, Y. E., Ercişli, S. (2020), Forecasting Banana Harvest Area and Production in Turkey Using Time Series Analysis. Erwerbs-Obstbau, 62(3), pp. 281–291.

Khan, T., Qiu, J., Ali Qureshi, M. A., Iqbal, M. S., Mehmood, R., Hussain, W. (2020), Agricultural Fruit Prediction Using Deep Neural Networks. Procedia Computer Science, 174, pp. 72–78.

Maskey, M. L., Pathak, T. B., Dara, S. K. (2019), Weather Based Strawberry Yield Forecasts at Field Scale Using Statistical and Machine Learning Models. Atmosphere, 10(7), pp. 378.

Rueangrit, P., Jatuporn, C., Suvanvihok, V., Wanaset, A. (2020), Forecasting Production and Export of Thailand's Durian Fruit: An Empirical Study using the Box–Jenkins Approach. Humanities and Social Sciences Letters, 8(4), pp. 430–437.

Yusuf, S. A., Salau, A. S. (2007), Forecasting Mango and Citrus Production in Nigeria: A Trend analysis.

Funding acknowledgements

ImPUlSe is implemented under the umbrella of the Partnership for Research and Innovation in the Mediterranean Area (PRIMA). The PRIMA programme is an Art. 185 initiative supported and funded under Horizon 2020, the European Union's Framework Programme for Research and Innovation.

The PRIMA programme is supported and funded under Horizon 2020, the Framework European Union's Programme for Research and Innovation

The various project partners are each funded by their national funding institution:

SPONSORED BY THE
Federal Ministry of Education and Research

أكاديمية البحث العلمي والتكنولوجي
Academy of Scientific Research & Technology

TÜBİTAK

Directorate-General for Scientific Research and Technological Development ALGERIA
DGRSDT
البحث العلمي في خدمة المواطن

Tunisian Republic
Ministry of Higher Education and Scientific Research

anr agence nationale de la recherche