PROPOSAL OF MODEL FOR PREDICTION OF GRAPE PROCESSING AND SPRAYING TIME BY USING IOT SMART AGRICULTURE SENSOR DATA

Jakup Fondaj (1) *, Mentor Hamiti (1), Samedin Krrabaj (2), Xhemal Zenuni(1), Jaumin Ajdari(1)

(1) South East European University, Tetovo, North Macedonia;
(2) University “Ukshin Hoti”, Prizren, Kosovo

* Corresponding Author, e-mail: jf13459@seeu.edu.mk

Abstract: The grape industry's impact on agriculture and the economy requires precise forecasting for processing and spraying schedules to optimize production. This article introduces an innovative IoT-based model for predicting optimal timings in grape processing and spraying. By integrating real-time environmental and viticultural data, the model improves decision-making, enhancing product quality, reducing energy consumption, and increasing operational efficiency. Crucially, SARIMA predictive algorithms forecast parameters like temperature, humidity, wind speed, and air pressure. This comprehensive model transforms the grape industry, offering advanced decision support and promoting sustainable, resource-efficient production. The research signals a potential shift to precision agriculture, balancing economic viability with environmental stewardship in grapes.

Keywords: SARIMA, Internet of Things (IoT), smart agriculture sensor data, artificial intelligence, prediction.

1. INTRODUCTION

Grape production is a significant agricultural sector, with grapes being used for various purposes, including winemaking, juice production, and table grape consumption. The processing and spraying of grapes play a crucial role in the production of grape-based food and beverage products, such as juices, jams, and wines. The quality and taste of these products depend greatly on the timing and duration of processing and spraying. Optimizing these processes is essential for maintaining product consistency, meeting consumer demands, and minimizing resource wastage.

Traditional approaches to determine grape processing and spraying times often rely on fixed time intervals or manual sampling methods. These methods can lead to inconsistent product quality and energy inefficiencies. Moreover, they do not account for the dynamic environmental and viticultural factors that influence grape maturity and composition. Therefore, there is a need for a model that integrates IoT smart agriculture
sensor data to provide accurate and timely predictions of grape processing and spraying times.

The primary objective of this research is to propose a model that uses IoT smart agriculture sensor data to predict the optimal processing and spraying times for grapes. By considering real-time environmental variables, such as temperature, humidity, and viticultural parameters, including grape quantity and quality from previous years, the model aims to provide food and beverage producers with reliable predictions for efficient and high-quality grape processing.

Based on our previous research [17, 18, 19], we have chosen to use the SARIMA [22] model for further prediction in this model because it has shown better performance than other algorithms we tested [19].

Next, we present related work in this field, followed by an explanation of the research methodology in section three. In section four, we demonstrate the prediction using SARIMA for four evaluation parameters. Section five presents the proposed model for predicting grape processing and spraying times using IoT smart agriculture sensor data. Finally, in section six, are the conclusions and future work.

2. RELATED WORK

The review article [1] provides an overview of the application of IoT in agriculture. It discusses the potential benefits of IoT in improving crop management and highlights the importance of real-time monitoring and data-driven decision-making.

The study [2] explores the use of remote sensing techniques, including thermal imaging and hyperspectral imaging, for crop stress detection. The integration of such sensors in IoT platforms can provide valuable data for monitoring grapevine health and phenological stages, contributing to the prediction of grape spraying time.

The study [3] proposes a model for predicting grape maturity using IoT and machine learning techniques. It utilizes environmental data, such as temperature and humidity, collected by IoT sensors to train a machine-learning model. The model demonstrates promising results in accurately predicting grape spraying time.

This research paper [4] presents a grape ripening prediction model that incorporates IoT sensors. The model considers environmental parameters, including temperature, humidity, and solar radiation, along with vine phenology data. The integration of IoT sensors enhances the accuracy of grape ripening prediction, enabling precise determination of spraying time.

The study [5] focuses on grapevine yield prediction using IoT and machine learning techniques. By analyzing data from IoT sensors, including climate, soil moisture, and vegetation indices, the model accurately predicts grapevine yield. This approach can also be extended to predict grape spraying time, considering the correlation between yield and grape maturity.

The research paper [6] explores the automated detection of crop phenology in grapevines using IoT sensors. By analyzing data from temperature and humidity sensors, the model detects phenological stages such as budburst, flowering, version, and ripening. The accurate detection of grape ripening stage contributes to predicting optimal grape spraying time.
The review paper [7] provides an overview of the application of IoT in agriculture, including smart agriculture and precision viticulture. It discusses the potential benefits of IoT in optimizing grape processing and spraying by enabling real-time monitoring and data-driven decision-making.

The systematic literature review [8] focuses on predictive models in precision viticulture. It discusses the use of IoT sensors and data analytics for monitoring grapevine phenology, environmental conditions, and grape composition. The review highlights the potential for utilizing IoT data to predict optimal grape processing and spraying time.

The study [9] presents an IoT-based grapevine intelligent management system using wireless sensor networks. It discusses the integration of various sensors to monitor environmental conditions, soil moisture, and grapevine growth. The data collected by the sensors can be utilized to predict grape processing and sterilization time.

The research paper [10] proposes a model for predicting grape berry quality using IoT sensors and machine learning algorithms. It highlights the importance of real-time data from IoT sensors, such as temperature, humidity, and sugar content, in accurately predicting grape processing and sterilization time.

The study [11] focuses on the development of a precision viticulture system based on IoT and cloud computing for vineyard management. It discusses the integration of IoT sensors for monitoring environmental conditions, grapevine phenology, and grape composition. The collected data can be utilized to predict grape processing and sterilization time.

The review article [12] provides an overview of smart precision viticulture systems. It discusses the use of IoT technologies, including sensors, actuators, and data analytics, for monitoring vineyard conditions, grapevine growth, and grape composition. The review emphasizes the potential for utilizing IoT sensor data in predicting grape processing and sterilization time.

The study [13] focuses on grape spraying quality estimation using IoT and machine learning techniques. It discusses the integration of IoT sensors for monitoring environmental conditions, grape maturity, and grape quality parameters. The study highlights the potential for utilizing IoT data to predict optimal grape processing and sterilization time.

The paper [14] focuses on the development of a process for spray forecasting in agriculture. It proposes a methodology that utilizes data analysis techniques to predict crop yields and improve decision-making for farmers. The authors highlight the importance of accurate forecasting to optimize agricultural practices and ensure better outcomes.

The paper [15] focuses on the use of machine learning algorithms for crop selection and yield prediction in agriculture. It proposes an approach that leverages various machine-learning techniques to analyze agricultural data and provide recommendations for crop selection. The authors highlight the potential of these algorithms to improve decision-making and enhance crop yield in agricultural practices.

The paper [16] addresses the challenge of crop yield prediction in Punjab state by proposing an interactive model. It develops a classification-based approach that utilizes machine learning techniques to predict the yields of different crops. The authors
highlight the importance of accurate yield prediction in agricultural planning and decision-making for farmers.

3. RESEARCH METHODOLOGY

To conduct this research, we have done several research methods as data collection for IoT smart agriculture sensor data, then we have done data preprocessing, feature extraction, model development, and model evaluation.

3.1. Data Collection: The proposed model hinges on gathering extensive data sourced from IoT smart agriculture sensors strategically positioned in grape vineyards. In this subsection, we introduce the dataset utilized for the predictions, acquired from the Kosovo Hydrometeorological Institute. This dataset comprises crucial parameters such as temperature, humidity, wind speed, and air pressure. The selection of these parameters for prediction is rooted in the research, which underscores their pivotal role in determining grape quality and quantity. We possess a dataset spanning 5 to 6 years, and we will illustrate it through visualization in the subsequent figures.

The temperature forecasting data we utilized covers a six-year timeframe, starting from early 2017 and concluding at the end of 2022 (refer to Figure 1). We sourced this dataset from IoT sensors, which continuously recorded temperature values over this period. At the end of 2019 and the beginning of 2020, there are constant values because the sensors not working.

Fig. 1. Temperature dataset
In the winter season, temperatures typically remain close to 0 degrees Celsius, occasionally dropping to negative values. During this period, the temperature readings are consistently measured and closely aligned with each other. As spring emerges, temperature fluctuations become more noticeable, influenced by shifting weather patterns, which include sunny days, precipitation, and cloudy conditions. Transitioning into summer, temperatures regain stability, mirroring the warm and sunny climate prevalent in Kosovo, albeit with occasional minor fluctuations. In Figure 2 we present the humidity dataset. At the end of 2019 and the beginning of 2020, there are constant values because the sensors not working.

Shifting the focus to humidity predictions, the examination exposes discernible trends. The period from winter to early spring exhibits elevated humidity levels, which gradually diminish from April to mid-May. Throughout the summer season, humidity tends to remain relatively low. As autumn draws near, humidity levels ascend once more, peaking during the winter months. It's worth noting that even in air humidity measurements, we encounter anomalies that can impact the forecasting process. In figure 3 we present the wind speed dataset dataset.
An additional crucial parameter incorporated into the model involves the analysis and prediction of wind speed. We have accessed data from the Kosovo Hydrometeorological Institute, covering the Prizren region during the period spanning from 2019 to 2022. In Figure 4 we present the air pressure dataset. In the last part of 2020, the data were missed because the sensors didn’t work and we don’t have data for that period.

The fourth parameter under scrutiny within our proposed model is air pressure. We have acquired the dataset from the Kosovo Hydrometeorological Institute, spanning five years from 2018 to 2022. The air pressure dataset is four-year data from 2019 to 2022.
It's noteworthy that in the previous year, there was a notable increase in air pressure, a factor that bears significance for grape cultivation, as elucidated in the results of the prediction algorithms, which we will present in the following subchapters.

### 3.2. Data Preprocessing
Preprocessing encompasses various essential procedures, such as data cleansing, normalization, and outlier identification. These measures serve to uphold the integrity and quality of the collected sensor data, making it ready for subsequent analysis.

### 3.3. Feature Extraction
Feature extraction methods are utilized to pinpoint pertinent attributes within the gathered sensor data. We employ various techniques, including statistical measures, time-series analysis, and feature selection algorithms based on machine learning, to extract valuable features for the prediction model.

### 3.4. Model Development
Various machine learning algorithms, including regression, decision trees, and neural networks, undergo assessment and comparison to determine their appropriateness for predicting grape processing and spraying times. We elucidate the chosen algorithm and provide the rationale behind its selection.

### 3.5. Model Training and Evaluation
The gathered sensor data is partitioned into training and testing sets to facilitate the training and evaluation of the prediction model. We utilize evaluation metrics such as mean absolute error (MAE) [26], root mean square error (RMSE) [27] and mean squared error (MSE) [27] to gauge the model's accuracy and performance. The results of this evaluation are detailed in our previous research [19]. Based on the previous research [19] we have selected SARIMA as the best predictive algorithm in the case and based on it we have evaluated the proposed model by comparing manually the five-year data and the predicted data with SARIMA.

In Tables 1 and 2 below we have presented a part of these data and how we analyze the existing data and prediction. Based on it, we conclude that temperature is higher and also when we have evaluated [19] the algorithms with MAE, MSE, and RMSE the best algorithm is SARIMA. The tables below show that the temperature is higher and the humidity is lower. Based on this information, we propose the right time to do grape spraying for farmers as we present in the chapters below.

#### Table 1. Temperature comparison with the prediction for 2023 by SARIMA

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5/1/2023 13:00</td>
<td>17.407</td>
<td>26.9</td>
<td>30</td>
<td>18.3</td>
<td>24.5</td>
<td>13.885</td>
<td>29.29365</td>
</tr>
<tr>
<td>5/2/2023 13:00</td>
<td>23.916</td>
<td>27.3</td>
<td>30</td>
<td>20.7</td>
<td>25.9</td>
<td>11.633</td>
<td>29.86727</td>
</tr>
<tr>
<td>5/3/2023 13:00</td>
<td>18.272</td>
<td>19.9</td>
<td>30</td>
<td>12.3</td>
<td>24.1</td>
<td>17.662</td>
<td>27.26964</td>
</tr>
<tr>
<td>5/5/2023 13:00</td>
<td>22.152</td>
<td>24.4</td>
<td>30</td>
<td>17.8</td>
<td>21.7</td>
<td>20.574</td>
<td>22.85682</td>
</tr>
</tbody>
</table>

#### Table 2. Humidity Comparison with a SARIMA Prediction for 2023

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023-Sarima</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/1/2017 13:00</td>
<td>68.532</td>
<td>36</td>
<td>50</td>
<td>51</td>
<td>47</td>
<td>37.64664</td>
<td>50.05725</td>
</tr>
<tr>
<td>5/2/2017 13:00</td>
<td>46.05</td>
<td>42</td>
<td>50</td>
<td>46</td>
<td>35</td>
<td>30.60458</td>
<td>43.01347</td>
</tr>
<tr>
<td>5/3/2017 13:00</td>
<td>68.516</td>
<td>68</td>
<td>50</td>
<td>90</td>
<td>32</td>
<td>31.75366</td>
<td>44.16205</td>
</tr>
<tr>
<td>5/4/2017 13:00</td>
<td>46.412</td>
<td>55</td>
<td>50</td>
<td>48</td>
<td>39</td>
<td>36.77398</td>
<td>49.18262</td>
</tr>
<tr>
<td>5/5/2017 13:00</td>
<td>37.517</td>
<td>61</td>
<td>50</td>
<td>38</td>
<td>38</td>
<td>45.54947</td>
<td>57.95816</td>
</tr>
<tr>
<td>5/6/2017 13:00</td>
<td>40.15</td>
<td>80</td>
<td>50</td>
<td>38</td>
<td>39</td>
<td>37.5315</td>
<td>49.94033</td>
</tr>
</tbody>
</table>
4. PROPOSAL OF MODEL FOR PREDICTION OF GRAPE PROCESSING AND SPRAYING TIME BY USING IOT SMART AGRICULTURE SENSOR DATA

This section introduces the structure of the proposed model, delineating the data flow, sensor integration, and prediction components. It elucidates how the model harnesses IoT smart agriculture sensor data to produce precise forecasts of grape processing and sterilization timing.

The model comprises six phases:

1. Data Collection: The initial phase involves gathering data from IoT smart agriculture sensors.

2. Data Preprocessing: In the second phase, data mining techniques are employed to address missing data issues. These gaps are filled by substituting them with data from nearby time periods.

3. Prediction of Weather Conditions: The third phase centers on predicting weather conditions. In this stage, we apply four well-established algorithms for weather data prediction: Random Forest Regression [23], Seasonal ARIMA (SARIMA) [22], NeuralProphet[24], and ANN from KERAS [25]. We then evaluate the results using metrics like mean absolute error (MAE) [26], root mean square error (RMSE) [27] and mean squared error (MSE) [27] to gauge the model's accuracy and performance. The results are presented in our previous research [19], with SARIMA identified as the most effective algorithm.

4. Visualization of Results: The fourth phase focuses on visualizing the results. We utilize the Python Plotly library to create dynamic graphics, facilitating in-depth analysis of the obtained outcomes.

5. Analysis Phase: The main phase is the fifth one, where we analyze the results derived from the prediction phase. The analysis is done manually for the data obtained from the Vineyard Institute in Kosovo.

6. Recommendation Phase: The last phase recommends the optimal grape spraying time for farmers based on the analyses conducted in phase five. This recommendation provides insight into how the suggested timing affects grape quality and quantity.

In Figure 5 below, we present a visual representation of the proposed model.

![Fig. 5. Proposed model for prediction of the time period for grape processing](image)

To ascertain the accuracy and dependability of our proposed model, we undertake a validation process. This section delves into the validation methodologies utilized,
including cross-validation and comparison with ground truth data. These techniques are instrumental in verifying the model's efficacy in predicting grape processing and spraying times. In essence, we suggested the spraying timeframe for 2023 and are currently monitoring the outcomes, which, thus far, have proven favorable in comparison to other farmers who opted for different spraying schedules.

6. RESULT DISCUSSION

Based on the model in Figure 5, its results are discussed. We concentrate on four key parameters: temperature, humidity, air pressure, and wind speed—critical factors influencing grape quality and quantity. We scrutinize data from the Kosovo Vineyard Institute spanning the last five years, encompassing grape quantity and quality. Based on this data and the weather condition predictions, we project conditions for 2023. For instance, if 2022 experienced high temperatures during the grape cultivation period (April to October) with lower grape yields than in 2021, when temperatures were lower, and our 2023 prediction indicates higher temperatures than 2022, we advise farmers to change grape spraying period by extending it, specifically from May 15th to July 05th, 2023, instead of the usual 10 May to 28 June. This means we recommend increasing the number of sprays from five to six and the spraying period to be longer because the temperature has increased and it is necessary to spray the grape to avoid grape illnesses. The spraying effect starts after two days and its effect is obtained after 15 days, this means it needs to be repeated on time to avoid grape illnesses. These findings underscore the importance of pinpointing the correct grape spraying time to bolster both quality and quantity.

7. CONCLUSION AND FUTURE WORKS

This section offers a concise overview of the principal findings and contributions of the proposed model. Our model, designed to determine the optimal grape spraying time in the Kosovo region through predictive analysis, stands as the pioneering initiative in this specific area. Prior research has presented similar models, but they predominantly focused on European regions like Italy, Spain, etc., and none have addressed the specific region we have studied.

Currently, we are in the process of evaluating the proposed model this year, and it is evident that it has had a positive impact on grape quality. The evaluation phase will conclude upon the completion of the grape cultivation season, and our next research endeavor will involve presenting the results from the ongoing case study evaluation.

It is worth emphasizing that our proposed model demonstrates remarkable accuracy, reliability, and practical applicability in the realm of predicting grape processing and spray timing. This model is poised to directly benefit the economic well-being of grape-cultivating families, as it aids in enhancing both the quality and quantity of grapes by providing precise predictions for grape processing and spraying times.

Within the proposed model, we have integrated supplementary sensor data encompassing temperature, wind speed, air pressure, and humidity. Additionally, we have incorporated advanced machine-learning techniques, including Random Forest Regression [23], Seasonal ARIMA (SARIMA)[22], NeuralProphet[24], and ANN from
KERAS [25]. Furthermore, we have taken into account external factors such as seasonal variations and disease prevalence, thereby enriching the model's comprehensive approach.

In future work, we are going to analyze the quality and quantity of grapes we are going to get from farmers which we monitor and use the recommendation, and compare with other cases that didn’t use the recommendation to evaluate the correctness of our model.

REFERENCES


Information about the authors:

Jakup Fondaj – is Ph.D. candidate at the Faculty of Contemporary Sciences and Technologies in the field of Computer Science at South East European University in Tetovo Macedonia. His research is related to IoT and Data sciences. He has a large number of papers in the field of computer science.

Mentor Hamiti – is a full-time professor at the Faculty of Contemporary Sciences and Technologies at South East European University in Tetovo Macedonia. He is a Doctor of Computer Sciences at the Faculty of Contemporary Sciences and Technologies, South East European University, Ilindenska 335, 1200 Tetovo, Macedonia. Specialty: Algorithms, Natural Language Processing, Program Languages and Technologies, Professional Ethics. Thesis: "Text-to-Speech Conversion: Examination, Experimentation, and Application in Albanian". He has a large number of papers in this field.

Samedin Krrabaj – is an associate professor at the Faculty of Computer Science at the University “Ukshin Hoti” Prizren, Kosovo. He is a doctor of technical sciences at the Faculty of Engineering at the University of Pristina, Kosovo. He has a large number of papers in this field.

Xhemal Zenuni – is a full-time professor at the Faculty of Contemporary Sciences and Technologies at South East European University in Tetovo Macedonia. He is a Doctor of Computer Sciences at the Technical University of Sofia, Sofia, R. of Bulgaria. Specialty: Computer Systems, Complexes and Networks. Thesis: "QoS Aware Semantic Service Composition via AND/OR Graphs". He has a large number of papers in this field.

Jaumin Ajdari – is a full-time professor at the Faculty of Contemporary Sciences and Technologies at South East European University in Tetovo Macedonia. He is a Doctor of Mathematical Sciences at the University of Tirana, Faculty of Natural Sciences, Tirana. Specialty: Parallel processing.

Manuscript received on 29 November 2023