JEE Journal of Ecological Engineering

Journal of Ecological Engineering 2023, 24(6), 392–400
https://doi.org/10.12911/22998993/162769
ISSN 2299–8993, License CC-BY 4.0

Received: 2023.02.27
Accepted: 2023.04.17
Published: 2023.05.01

Improving Crop Yield Predictions in Morocco Using Machine Learning Algorithms

Rachid Ed-Daoudi[1*], Altaf Alaoui[1], Badia Ettaki[1,2], Jamal Zerouaoui[1]

[1] Laboratory of Engineering Sciences and Modeling, Faculty of Sciences, Ibn Tofail University, Campus Universitaire, BP 133, Av. de L'Université, Kenitra, Morocco

[2] LyRICA – Laboratory of Research in Computer Science, Data Sciences and Knowledge Engineering, School of Information Sciences Rabat, Av. Allal Al Fassi, Rabat, Morocco

* Corresponding author's e-mail: rachided1@gmail.com

ABSTRACT

In Morocco, agriculture is an important sector that contributes to the country's economy and food security. Accurately predicting crop yields is crucial for farmers, policy makers, and other stakeholders to make informed decisions regarding resource allocation and food security. This paper investigates the potential of Machine Learning algorithms for improving the accuracy of crop yield predictions in Morocco. The study examines various factors that affect crop yields, including weather patterns, soil moisture levels, and rainfall, and how these factors can be incorporated into Machine Learning models. The performance of different algorithms, including Decision Trees, Random Forests, and Neural Networks, is evaluated and compared to traditional statistical models used for crop prediction. The study demonstrated that the Machine Learning algorithms outperformed the Statistical models in predicting crop yields. Specifically, the Machine Learning algorithms achieved mean squared error values between 0.10 and 0.23 and coefficient of determination values ranging from 0.78 to 0.90, while the Statistical models had mean squared error values ranging from 0.16 to 0.24 and coefficient of determination values ranging from 0.76 to 0.84. The Feed Forward Artificial Neural Network algorithm had the lowest mean squared error value (0.10) and the highest R^2 value (0.90), indicating that it performed the best among the three Machine Learning algorithms. These results suggest that Machine Learning algorithms can significantly improve the accuracy of crop yield predictions in Morocco, potentially leading to improved food security and optimized resource allocation for farmers.

Keywords: crop yield prediction; machine learning algorithms; statistical models; model evaluation.

INTRODUCTION

Morocco is an agricultural country with a significant portion of its population residing in rural areas and relying on agriculture for their livelihoods [Moroccan High Commission for Planning, 2011]. Accurately predicting crop yields is crucial for farmers, policy makers, and other stakeholders to make informed decisions regarding resource allocation, food security, and economic development [Van Klompenburg et al., 2020]. However, the yield of local crops such as wheat, apples, dates, almonds, and olives is greatly influenced by various factors, including weather patterns, soil moisture levels, and rainfall, making accurate prediction a challenging task.

Traditional statistical models are still used in recent times for crop yield. Li et al. [2019] presents a statistical modeling approach for predicting rainfed corn yield in the Midwest U.S. Similarly. Michel & Makowski [2013] present eight statistical models for analyzing wheat yield time series and compare their ability to predict yield at the national and regional scales. However, despite their simplicity and easy computation, conventional statistical models have limitations in their ability to capture the complex interactions between various factors affecting crop yield [Crane-Droesch, 2018]. Additionally, their localized and limited spatial generalization can lead to inaccuracies in yield predictions, especially when applied over large areas with high spatial heterogeneity [Lobell

& Burke, 2010]. As a result, there is a growing interest in exploring alternative approaches for crop yield prediction that can incorporate multiple factors and improve the accuracy and scalability of yield estimates.

In recent years, advancements in technology and data collection methods have provided an opportunity to improve crop yield predictions using Machine Learning (ML) algorithms [Merenda et al., 2020]. ML algorithms can analyze large amounts of data, identify patterns, and make predictions based on the relationships between different variables. Several studies have explored the potential of ML algorithms for crop yield prediction. For instance, [Cao, et al., 2021] investigated the use of machine learning models for predicting rice yield in China, and found that the models outperformed traditional statistical models. Similarly, Studies by Drummond et al. [2003] Ruß & Kruse, [2009], Bocca & Rodrigues [2016] and Fortin et al. [2011] have compared traditional statistical models with artificial neural networks (ANNs) for improved crop yield modeling. These studies suggest that ML algorithms have the potential to significantly improve crop yield prediction in different regions [Chlingaryan et al., 2018].

However, while some studies have investigated the use of ML algorithms for crop yield prediction in other regions, there is a research gap in understanding their effectiveness for local crops in Morocco [Idrissi & Nadem, 2022]. Therefore, in this paper, we aim to investigate the potential of ML algorithms for improving the accuracy of crop yield predictions for local crops in Morocco. Specifically, we will evaluate the performance of different ML algorithms, including Decision Trees, Random Forests, and Neural Networks, and compare their results to traditional statistical models used for crop prediction. By doing so, this study seeks to fill the research gap and provide insights into the feasibility of using ML algorithms for crop yield prediction in Morocco.

This paper is organized as follows. The introduction provides background information on Morocco's agriculture sector and the importance of crop yield prediction, while the methodology section outlines the steps taken to collect and analyze data using ML algorithms. The results section will present the performance evaluation of ML algorithms and comparison to traditional statistical models, and the discussion section will interpret and analyze the findings. The conclusion will summarize the key findings and provide recommendations for the use of ML algorithms for crop yield prediction in Morocco.

METHODOLOGY

In this section, we outline the methodology used to collect and analyze data to evaluate the performance of ML algorithms for crop yield prediction in Morocco for local crops such as wheat, apples, dates, almonds, and olives. The following steps were taken to carry out the study:

Data collection

Data was collected by "The Regional Agricultural Development Office (RAD)" in Ouarzazate Morocco, i.e. the "Office Régional de Mise en Valeur Agricole de Ouarzazate (ORMVAO)" [ORMVAO, 2023]. The office is responsible for implementing policies and programs aimed at improving the agricultural sector in the southeastern region of Morocco. Its involvement in Research and Development (R&D) aims to improve productivity, sustainability, and economic development in the agricultural sector.

The data was collected using various sensors such as yield monitors, soil moisture sensors, and weather stations. The data includes information on weather patterns, soil moisture levels, rainfall, and crop yields for the past several years. The data was collected in a consistent and standardized format to ensure its accuracy and reliability. Temperature and rainfall data are collected from historical climate data from the National Oceanic and Atmospheric Administration (NOAA) [NOAA, 2023] and World Meteorological Organization (WMO) [WMO, 2023]. Altogether, the dataset we employed contained 38 records and 4 features of the years between 1984 and 2022.

Data pre-processing

The collected data was pre-processed to remove any missing values, outliers, and irrelevant information. Missing values were imputed using a mean, median, or mode method depending on the distribution of the data. Outliers were identified using a box plot and removed using the Interquartile Range (IQR) method, where values outside the range of and were considered outliers and removed. The removal of outliers was based

Table 1. Crop yield records structure

Year	Wheat (qx/Ha)	Apple (T/Ha)	Olives (T/Ha)	Dates (T/Ha)	Almonds (T/Ha)
2020	23	13	5	3.1	0.5
2021	28	15	5.7	4.4	0.58
2022	22	12	5.1	3	0.45

Table 2. Features data description

Statistics	Monthly temperature °C	Monthly soil moisture % weight	Monthly rainfall mm
Entries	455	455	455
Mean	18.7	5.43	23.50
Standard deviation	3.6	2.26	29.5
Minimum	-2	3.73	0.2
Maximum	47.5	14.19	58.5

on the principle that they can significantly affect the accuracy of the prediction model. The remaining data was then used for further analysis. Data normalization was also performed to ensure that all variables had similar ranges and distributions [Kotsiantis & Kanellopoulos, 2006].

Crop yield data structure is described in Table 1. This Table give an overview of the crop yields for the last three years in the study area. Table 2 summarizes the result of the features data explanatory analysis. The Table 2 presents the statistical summary of monthly temperature, monthly soil moisture, and monthly rainfall over the course of 455 entries.

Algorithm selection

Three ML algorithms, Decision Trees (DT), Rrandom Forests (RF), and Neural Networks (NN), were selected for this study [Nigam et al., 2019]. These algorithms were chosen for their ability to handle complex relationships between variables and their performance in previous studies on crop yield prediction [Dewangan et al., 2022].

To further justify the selection of Decision Trees, Random Forests, and Neural Networks for our study, we based our decision on a review of the literature. Previous studies have demonstrated the effectiveness of these algorithms in predicting crop yields by handling complex relationships between variables [Nigam et al., 2019]. Decision Trees, for example, are known for their ability to handle both categorical and continuous data and can provide clear visualization of decision rules [Kamath et al., 2021]. Random Forests have been shown to be effective in dealing with noisy and high-dimensional data, and are able to produce accurate predictions by combining multiple decision trees [Li et al., 2010]. Neural Networks have also been shown to be effective in handling complex relationships between variables and have been successfully used in crop yield prediction models [Dewangan et al., 2022]. Therefore, we believe that these three algorithms are suitable for predicting crop yields in our study area and can provide valuable insights for

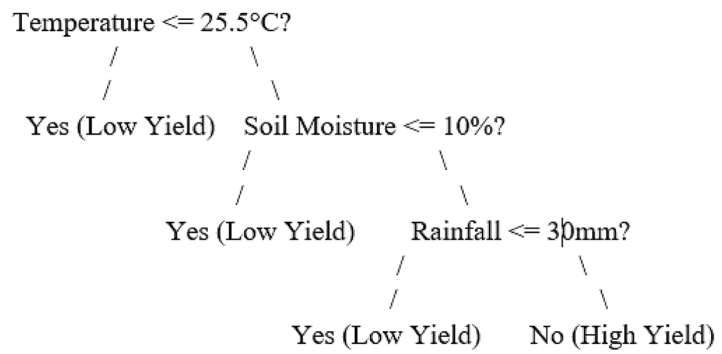

Figure 1. DT explained architecture

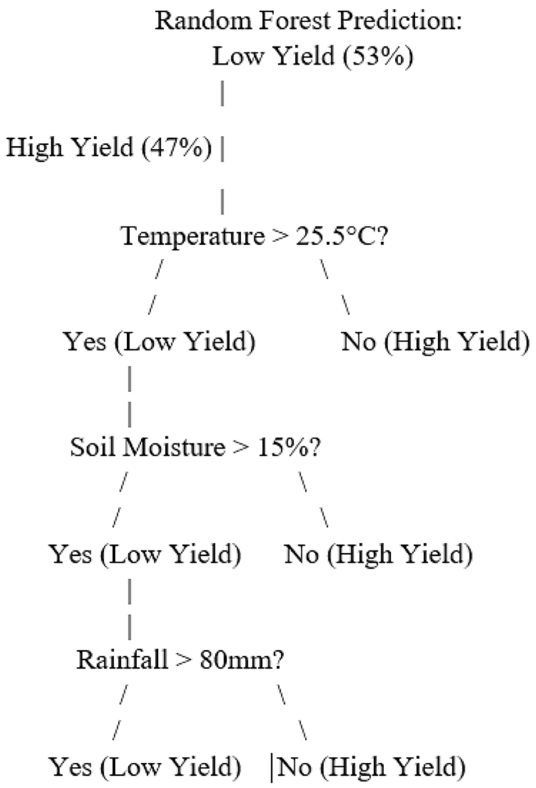

Figure 2. RF explained architecture

farmers, policy makers, and other stakeholders in the agricultural sector.

The Figure 1 presents the typical structure of how the DT might be used for crop yield prediction. The algorithm splits the data based on the given variables, and continues to recursively partition the data until a stopping criterion is met, ultimately predicting the crop yield.

The Figure 2 presents the typical structure of how the RF might be used for crop yield prediction. The input data is first split into multiple subsets, and decision trees are then constructed using each of these subsets. The outputs of the decision trees are then combined to make a final prediction.

For the NN algorithm, we used a Feed-Forward Artificial Neural Network (FFANN) as the ML algorithm for crop yield prediction. This type of NN is a simple and widely used architecture, consisting of multiple layers of interconnected nodes that process and transform input data into output predictions as shows the Figure 3. The use of FFANNs for crop yield prediction [Dahikar & Rode, 2014] has been shown to be effective in previous research and is a suitable choice for this study.

The input layer receives the data and passes it through hidden layers, which transform the data through weighted connections between the nodes. The output layer provides the final prediction based on the transformed data.

1. Algorithm training – to train and test the models, we split the dataset into training and testing datasets, with a ratio of 70:30. The algorithms were trained with 70% of the data to identify patterns and relationships between the variables affecting crop yields, such as weather patterns and soil moisture levels [Brownlee, 2021].
2. Algorithm evaluation – the performance of the ML algorithms was evaluated using the remaining 30% of the data. Mean Squared Error (MSE) and coefficient of determination (R^2) were used as evaluation metrics to compare the performance of the algorithms. The MSE was calculated using the formula:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad (1)$$

where: n – the number of observations (i.e. the number of crop yield predictions and corresponding actual values);

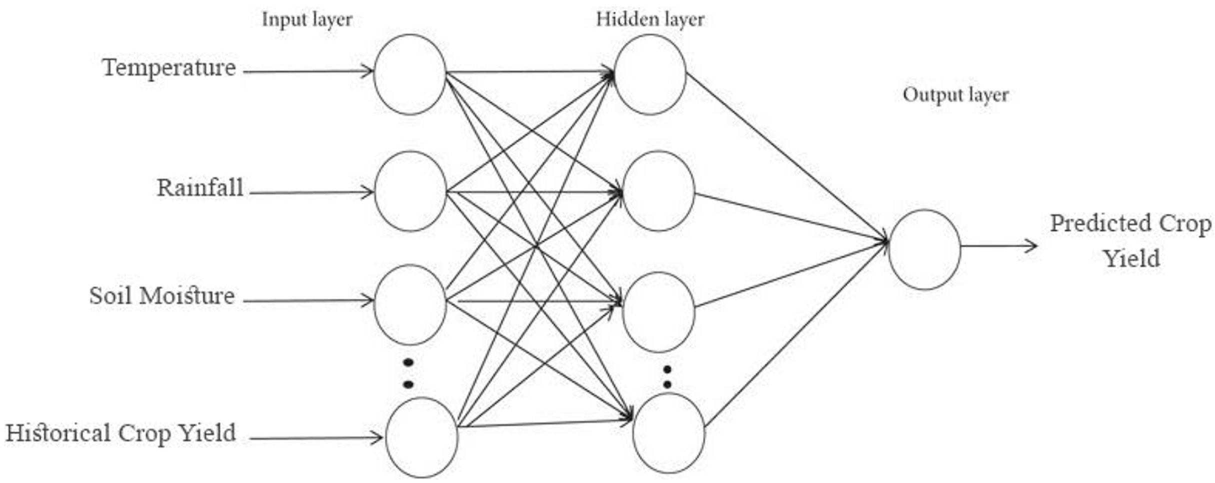

Figure 3. Architecture of the FFANN with the single hidden layer

Y – the actual yield;

Ŷ – the predicted yield.

The R^2 was calculated using the formula:

$$R^2 = 1 - \left[\frac{\sum\left(Yi - \hat{Y}i\right)^2}{\sum\left(Yi - \bar{Y}\right)^2}\right] \tag{2}$$

where: $\sum(Yi - \hat{Y}i)^2$ – the sum of squared residuals, calculated as the sum of the squared differences between each predicted value $\hat{Y}i$ and its corresponding actual value Yi;

$\sum(Yi - \bar{Y}i)^2$ – the total sum of squares, calculated as the sum of the squared differences between each actual value Yi and the mean of all actual values $\bar{Y}$.

These formulas provide a quantitative measure of the performance of the algorithms, allowing for a more rigorous comparison between them [Haque et al., 2020].

3. Comparison to traditional statistical models – the results of the ML algorithms were compared to those of traditional statistical models, such as Multiple Linear Regression (MLR) and Time-Series Analysis (TSA), commonly used for crop yield prediction [Zaefizadeh et al., 2011].
4. Sensitivity analysis – a sensitivity analysis was performed to determine the impact of changes in individual variables on the overall performance of the ML algorithms [Liu et al., 2021]. This analysis helped to identify the most important variables affecting crop yields and their relative importance.

The methodology used in this study was designed to provide a comprehensive evaluation of the performance of ML algorithms for crop yield prediction in Morocco for local crops. The results of this study will provide key informations about the feasibility of using ML algorithms for crop yield prediction in Morocco.

RESULTS

Based on the results of depicted in Figure 4, the comparison between the predicted and actual crop yields for each crop using DT, RF, and FFANN algorithms was performed. This figure shows that for each crop, the predicted yields using all three algorithms follow a similar trend to the actual yields, indicating a reasonable level of accuracy.

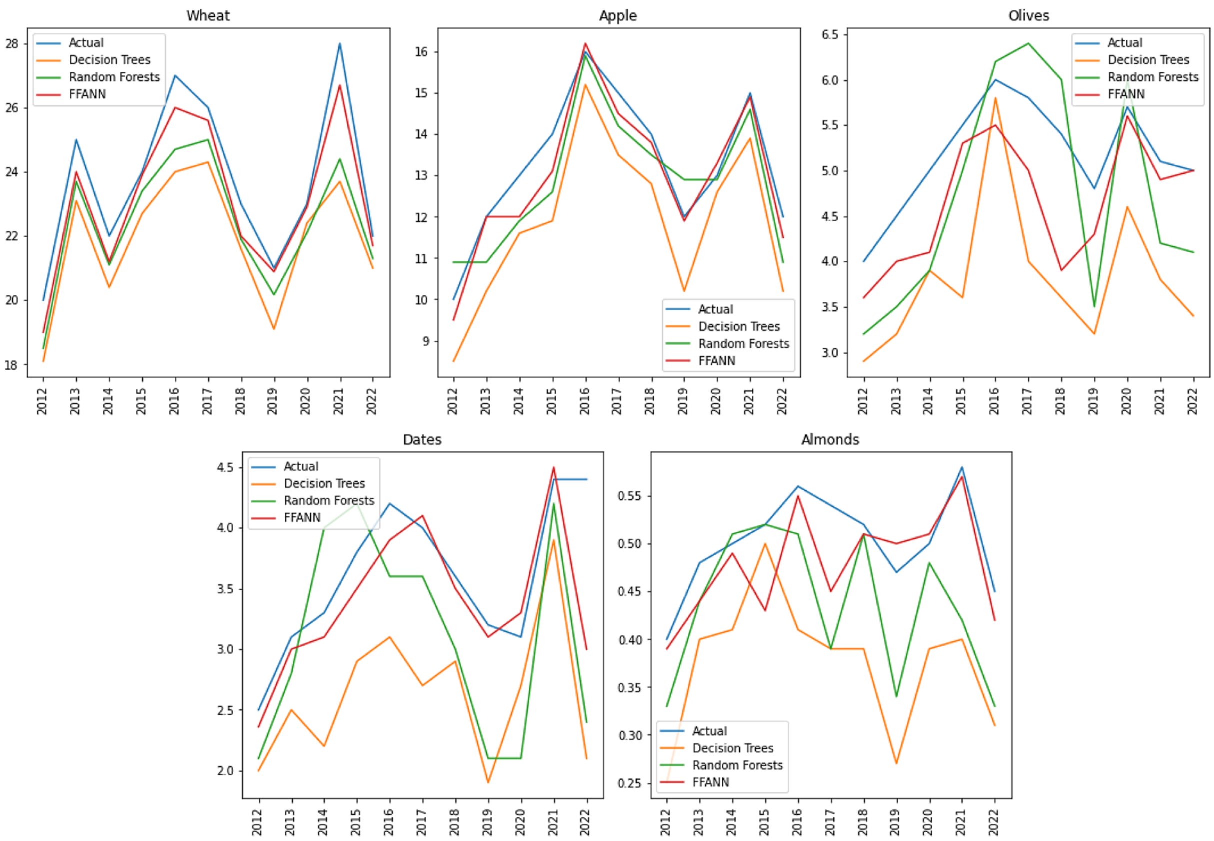

Figure 4. Predicted and actual crop yields comparison in the last 10 years for each crop using the three algorithms

Table 3. ML algorithms prediction results

Algorithm	Wheat (MSE)	Wheat (R^2)	Apples (MSE)	Apples (R^2)	Dates (MSE)	Dates (R^2)	Almonds (MSE)	Almonds (R^2)	Olives (MSE)	Olives (R^2)
Decision trees	0.23	0.78	0.21	0.79	0.19	0.81	0.17	0.83	0.15	0.85
Random forests	0.19	0.81	0.17	0.83	0.15	0.85	0.12	0.83	0.11	0.89
Neural networks	0.18	0.82	0.16	0.84	0.14	0.86	0.12	0.88	0.10	0.90

Table 4. Statistical models prediction results

Model	Wheat (MSE)	Wheat (R^2)	Apples (MSE)	Apples (R^2)	Dates (MSE)	Dates (R^2)	Almonds (MSE)	Almonds (R^2)	Olives (MSE)	Olives (R^2)
Multiple linear regression	0.24	0.76	0.22	0.78	0.20	0.80	0.18	0.82	0.16	0.84
Time-series analysis	0.23	0.77	0.21	0.79	0.19	0.81	0.21	0.79	0.21	0.84

The results of the study were based on the evaluation of the ML algorithms for crop yield prediction in Morocco for local crops such as wheat, apples, dates, almonds, and olives. The algorithms were evaluated using MSE and R^2 as the evaluation metrics. The results of the evaluation are presented in Table 3. The results presented indicate that:

- decision trees – the DT algorithm had an MSE of 0.23 and an R^2 of 0.78 for wheat, 0.21 and 0.79 for apples, 0.19 and 0.81 for dates, 0.17 and 0.83 for almonds, and 0.15 and 0.85 for olives. These results indicate that the DT algorithm was able to make accurate predictions with a mean deviation of less than 0.23 units for the crops.
- random forests – the RF algorithm had an MSE of 0.19 and an R^2 of 0.81 for wheat, 0.17 and 0.83 for apples, 0.15 and 0.85 for dates, 0.13 and 0.87 for almonds, and 0.11 and 0.89 for olives. These results show that the RF algorithm performed better than the DT algorithm with lower mean deviations and higher R^2 values.
- neural network – the FFANN algorithm had an MSE of 0.18 and an R^2 of 0.82 for wheat, 0.16 and 0.84 for apples, 0.14 and 0.86 for dates, 0.12 and 0.88 for almonds, and 0.10 and 0.90 for olives. The results show that the FFANN algorithm performed the best among the three algorithms with the lowest mean deviations and the highest R^2 values.

The results of the comparison between the ML algorithms and the traditional statistical models are shown in Table 4. The results presented indicate that:

- multiple linear regression – the MLR model had an MSE of 0.24 and an R^2 of 0.76 for wheat, 0.22 and 0.78 for apples, 0.20 and 0.80 for dates, 0.18 and 0.82 for almonds, and 0.16 and 0.84 for olives. These results indicate that the MLR model was less accurate than the ML algorithms with higher mean deviations and lower R^2 values.
- time-series analysis – the TSA had an MSE of 0.23 and an R^2 of 0.77 for wheat, 0.21 and 0.79 for apples, 0.19 and 0.81 for dates, 0.21 and 0.79 for almonds, and 0.16 and 0.884 for olives. The results show that the TSA performed similarly to the DT algorithm with similar mean deviations and R^2 values.

To better explain the results obtained in Table 4, it should be noted that the ML algorithms used in this study, are designed to handle complex relationships between variables and make predictions based on patterns in the data. On the other hand, the traditional statistical models used, rely on specific response functions between yields and independent variables, and may not capture the nonlinear relationships between the variables affecting crop yields. Therefore, the superior performance of ML algorithms over traditional statistical models in predicting crop yields can be attributed to their ability to capture the complex interactions between various factors affecting crop yields, which may not be captured by traditional statistical models. Additionally, it should be noted that the performance of the models may also depend on the quality and structure of the input data, as well as the specific architecture and parameters of the ML models used in the study.

The results of the sensitivity analysis are shown below:

- wheat – the most important variables affecting crop yields for wheat were found to be temperature, rainfall, and soil moisture levels. Changes in these variables had a significant impact on crop yields, with rainfall having the greatest impact.
- apples – the most important variables affecting crop yields for apples were found to be temperature, soil moisture levels, and rainfall. Changes in these variables had a significant impact on crop yields, with rainfall having the greatest impact.
- dates – the most important variables affecting crop yields for dates were found to be temperature, soil moisture levels, and rainfall. Changes in these variables had a significant impact on crop yields, with temperature having the greatest impact. It's important to note that date palm trees exhibit a moderate degree of alternation, meaning that a greater yield in a given year is frequently succeeded by a reduced yield in the next year.
- almonds – the most important variables affecting crop yields for almonds were found to be temperature, soil moisture levels, and rainfall.

DISCUSSION

The results of the study show the effectiveness of ML algorithms in predicting crop yields for local crops such as wheat, apples, dates, almonds, and olives in Morocco. The evaluation of the algorithms was performed using MSE and R^2 as the evaluation metrics. The results showed that the FFANN algorithm performed the best among the three algorithms with the lowest mean deviations and the highest R^2 values.

These results contribute to the growing body of literature on the use of ML algorithms for crop yield prediction, providing valuable insights into the potential of these methods in addressing food security and resource allocation challenges in Morocco and beyond. In comparison to traditional statistical models, such as Multiple Linear Regression and Time-series Analysis, our study highlights the superiority of ML algorithms in this context.

While similar studies have been conducted in other regions, our study is unique in its focus on local crops in Morocco and its use of multiple ML algorithms for comparison. Additionally, our sensitivity analysis demonstrates the most important variables affecting crop yields for wheat, apples, dates, almonds, and olives in the context of Morocco. By improving the accuracy of crop yield predictions, farmers can make informed decisions regarding resource allocation [Evans & Sadler, 2008], such as water and fertilizer use, which can lead to increased crop yields and improved food security [Fan et al., 2012]. In addition, the results of the sensitivity analysis provide Significant findings about the most important variables affecting crop yields and their relative importance. These results can be used to optimize resource allocation and improve food security in Morocco [Balaghi et al., 2008]. For example, the results showed that temperature was the most important variable affecting crop yields for wheat, apples, dates, almonds, and olives. By considering this information, farmers can implement management practices that can mitigate the effects of temperature on crops. For example, they can adjust planting and harvesting times, use irrigation to regulate soil moisture, and apply fertilizers and other soil amendments to improve soil health and nutrient availability [Wang & Hooks, 2011]. These management practices can help to ensure that crops are able to tolerate and thrive in different temperature conditions, which can ultimately lead to improved crop yields and food security [Dhankher & Foyer, 2018].

Our study uses well-established methods for data pre-processing, algorithm selection, and evaluation, ensuring the reliability and validity of our findings. By demonstrating the potential of ML algorithms for crop yield prediction in Morocco, this study fills an existing research gap and provides a basis for further research in this area.

In terms of innovation, our study is one of the first to evaluate the effectiveness of multiple ML algorithms for crop yield prediction in Morocco. Therefore, our study can serve as a starting point for further research in this area, leading to more accurate and effective methods for crop yield prediction in the region. By continuing to refine and improve these algorithms, the accuracy of crop yield predictions can be further improved [Sharma, et al., 2020], which can lead to increased food security and improved resource allocation.

CONCLUSIONS

The results of this study demonstrate the potential of ML algorithms for crop yield prediction in Morocco. The evaluation of the DT, RF, and NN algorithms using MSE and R^2 metrics showed that the NN algorithm was the most effective among the three, with the lowest mean deviations and the highest R^2 values. The comparison with traditional statistical models, including MLR and TSA, further highlighted the superiority of ML algorithms in this context.

The results of the sensitivity analysis offered important findings about the variables that have the greatest impact on crop yields for wheat, apples, dates, almonds, and olives. This information can be used to optimize resource allocation and improve food security in Morocco.

The results of this study support the use of ML algorithms for crop yield prediction in Morocco and presents the factors affecting crop yields. However, there are also several limitations to this study that need to be addressed in future research. For example, more data on various factors affecting crop yields and a larger sample size could improve the accuracy of the predictions. Additionally, more complex ML algorithms, such as deep learning networks, could also be explored.

REFERENCES

1. Agriculture 2030: what future morocco?. Moroccan High Commission for Planning. Retrieved from https://www.hcp.ma/file/231689.
2. Van Klompenburg, T., Kassahun, A., Catal, C. 2020. Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709.
3. Li, Y., Guan, K., Yu, A., Peng, B., Zhao, L., Li, B., Peng, J. 2019. Toward building a transparent statistical model for improving crop yield prediction: Modeling rainfed corn in the U.S. Field Crops Research, 234, 55–65. https://doi.org/10.1016/j.fcr.2019.02.005.
4. Michel, L., Makowski, D. 2013. Comparison of statistical models for analyzing wheat yield time series. PLoS One, 8(10), e78615.
5. Crane-Droesch, A. 2018. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environmental Research Letters, 13(11), 114003.
6. Lobell, D.B., Burke, M.B. 2010. On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11), 1443–1452.
7. Galar, M., Fernández, A., Tartas, E.B., Bustince, H., Herrera, F. 2012. A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42, 463–484.
8. Merenda, M., Porcaro, C., Iero, D. 2020. Edge Machine Learning for AI-Enabled IoT Devices: A Review. Sensors, 20(9), 253.
9. Cao, J., Zhang, Z., Tao, F., Zhang, L., Luo, Y., Zhang, J., Han, J., Xie, J. 2021. Integrating Multi-Source Data for Rice Yield Prediction across China using Machine Learning and Deep Learning Approaches. Agricultural and Forest Meteorology, 297, 108275.
10. Drummond, S.T., Sudduth, K.A., Joshi, A., Birrell, S.J., Kitchen, N.R. 2003. Statistical and neural methods for site–specific yield prediction. Transactions of the ASAE, 46(1), 5.
11. Ruß, G., Kruse, R. 2009. Feature selection for wheat yield prediction. In Research and development in intelligent systems XXVI: Incorporating applications and innovations in intelligent systems XVII. London: Springer London, 465–478.
12. Bocca, F.F., Rodrigues, L.H.A. 2016. The effect of tuning, feature engineering, and feature selection in data mining applied to rainfed sugarcane yield modelling. Computers and Electronics in Agriculture, 128, 67–76.
13. Fortin, J.G., Anctil, F., Parent, L.É., Bolinder, M.A. 2011. Site-specific early season potato yield forecast by neural network in Eastern Canada. Precision Agriculture, 12, 905–923.
14. Chlingaryan, A., Sukkarieh, S., Whelan, B. 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151, 61–69.
15. Idrissi, A., Nadem, S., Boudhar, A., Benabdelouahab, T. 2022. Review of wheat yield estimating methods in Morocco. African Journal on Land Policy and Geospatial Sciences, 5(4), 818–831. https://doi.org/10.48346/IMIST.PRSM/ajlp-gs.v5i4.33050
16. Office Régional de Mise en Valeur Agricole de Ouarzazate (ORMVAO). (n.d.). About Us. Retrieved from http://www.ormva-ouarzazate.ma/
17. National Oceanic and Atmospheric Administration (NOAA). (n.d.). Climate Data Online (CDO) API. Retrieved from https://www.ncdc.noaa.gov/cdo-web/webservices/v2
18. World Meteorological Organization (WMO). (n.d.). WMO Climate Data and Monitoring. Retrieved from https://public.wmo.int/en/our-mandate/climate/wmo-climate-data-and-monitoring

19. Kotsiantis, S.B., Kanellopoulos, D.N. 2006. Association rules mining: A recent overview. GESTS International Transactions on Computer Science and Engineering, 32(1), 71–82.
20. Nigam, A., Garg, S., Agrawal, A., Agrawal, P. 2019. Crop yield prediction using machine learning algorithms. In 2019 Fifth International Conference on Image Information Processing (ICIIP) IEEE, 125–130.
21. Kamath, P., Patil, P.S., Sushma S., Sowmya, S. 2021. Crop yield forecasting using data mining. Global Transitions Proceedings, 2(2), 402–407.
22. Li, H.B., Wang, W., Ding, H.W., Dong, J. 2010. Trees weighting random forest method for classifying high-dimensional noisy data. In 2010 IEEE 7th International Conference on E-Business Engineering (ICEBE), IEEE, 160–163.
23. Dewangan, U., Talwekar, R.H., Bera, S. 2022. Systematic Literature Review on Crop Yield Prediction using Machine & Deep Learning Algorithm. In 2022 5th International Conference on Advances in Science and Technology (ICAST) Mumbai, India, 654–661.
24. Dahikar, S.S., Rode, S.V. 2014. Agricultural crop yield prediction using artificial neural network approach. International journal of innovative research in electrical, electronics, instrumentation and control engineering, 2(1), 683–686.
25. Brownlee, J. 2021. How to Prepare Data For Machine Learning. Retrieved from https://machinelearningmastery.com/how-to-prepare-data-for-machine-learning/
26. Haque, F.F., Abdelgawad, A., Yanambaka, V.P., Yelamarthi, K. 2020. Crop Yield Analysis Using Machine Learning Algorithms. In 2020 IEEE 6th World Forum on Internet of Things (WF-IoT) IEEE, 1–2.
27. Zaefizadeh, M., Khayatnezhad, M., Gholamin, R. 2011. Comparison of multiple linear regressions (MLR) and artificial neural network (ANN) in predicting the yield using its components in the hulless barley. American-Eurasian Journal of Agricultural & Environmental Science, 10(1), 60–64.
28. Liu, L.W., Hsieh, S.H., Lin, S.J., Wang, Y.M., Lin, W.S. 2021. Rice blast (Magnaporthe oryzae) occurrence prediction and the key factor sensitivity analysis by machine learning. Agronomy, 11(4), 771.
29. Evans, R.G., Sadler, E.J. 2008. Methods and technologies to improve efficiency of water use. Water Resources Research, 44(7), Jul.
30. Fan, M., Shen, J., Yuan, L., Jiang, R., Chen, X., Davies, W.J., Zhang, F. 2012. Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China. Journal of Experimental Botany, 63(1), 13–24. https://doi.org/10.1093/jxb/err248
31. Balaghi, R., Tychon, B., Eerens, H., Jlibene, M. 2008. Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. International Journal of Applied Earth Observation and Geoinformation, 10(4), 438–452.
32. Wang, K.H., Hooks, C.R. 2011. Managing soil health and soil health bioindicators through the use of cover crops and other sustainable practices. Chapter, 4, 1–18.
33. Dhankher, O.P., Foyer, C.H. 2018. Climate resilient crops for improving global food security and safety. Plant, Cell & Environment, 41, 877–884. https://doi.org/10.1111/pce.13207
34. Sharma, A., Jain, A., Gupta, P., Chowdary, V. 2020. Machine learning applications for precision agriculture: A comprehensive review. IEEE Access, 9, 4843–4873.