A Machine Learning Approach to Identify Optimal Cultivation Practices for Sustainable apple Production in Precision Agriculture in Morocco

Rachid ED-DAOUDI[1,*], *Altaf* ALAOUI[1], *Badia* ETTAKI[1,2], and *Jamal* ZEROUAOUI[1]

[1]Laboratory of Engineering Sciences and Modeling, Faculty of Sciences- Ibn Tofail University, Kenitra, Morocco
[2]LyRICA: Laboratory of Research in Computer Science, Data Sciences and Knowledge Engineering School of Information Sciences, Rabat, Morocco

Abstract. Precision agriculture techniques have been increasingly adopted worldwide to optimize cultivation practices and achieve sustainable crop production. In this study, we developed a Machine Learning approach to identify optimal cultivation practices for sustainable apple production in precision agriculture in the Msemrir town Morocco. We collected a dataset of cultivation practices and apple yield and size data from 10 farms in the town and used correlation-based feature selection and three Machine Learning algorithms (Linear Regression, Decision Tree, and Random Forest) to develop predictive models. The results showed that irrigation, fertilization, and pruning are the most important cultivation practices for apple production in the region, and the Random Forest model performed the best in predicting apple yield and size based on the selected practices. The use of Machine Learning techniques can help farmers optimize cultivation practices and achieve sustainable apple production by reducing inputs such as water and fertilizer and minimizing environmental impact. Moreover, the use of precision agriculture techniques can help farmers meet consumer demand for sustainable and high-quality apple products.

Keyword: Precision agriculture, Machine Learning, Apple production, Sustainable agriculture, Cultivation practices.

1 Introduction

In recent years, precision agriculture has emerged as a key approach to increase crop productivity while reducing the environmental impact of agriculture [1]. Precision agriculture uses advanced technologies such as Internet of Things (IoT), remote sensing, Geographic

* Corresponding author: author@email.org

Information System (GIS), and Machine Learning (ML) to optimize crop management practices and achieve sustainable crop production.

Apple production is an important agricultural activity in Morocco, with a growing demand for sustainable and high-quality apple products in both domestic and international markets [2]. Msemrir, a town located in the Tinghir Province in the southeast region of Morocco, is one of the areas where apple farming takes place. Although it is not considered a major apple-producing region in Morocco, it is a significant contributor to the country's overall apple production, and as such, optimizing cultivation practices is crucial for achieving sustainable apple production that meets the quality standards and reduces environmental impact in this region [3].

In this paper, we propose a ML approach to identify optimal cultivation practices for sustainable apple production in precision agriculture in the studied area [4]. The objective is to develop a model that can predict the best cultivation practices for apple production, considering the specific soil and weather conditions in the region.

To achieve this objective, we collected a dataset that includes various cultivation practices and their associated apple yield and size data from the studied zone. We then applied ML techniques to the dataset to develop a model that can predict the yield and size for apple production based on the specific optimal cultivation practices.

The proposed approach has the potential to significantly improve apple production by optimizing cultivation practices and reducing environmental impact [5]. The results of this study can be used to develop precision agriculture strategies for apple production in this region and other similar regions.

The paper is structured to provide a comprehensive understanding of the proposed ML approach for identifying optimal cultivation practices for sustainable apple production in precision agriculture. The literature review presents an overview of precision agriculture techniques and previous research on the use of ML in precision agriculture. The methodology section details the data collection, feature selection, and ML algorithms used in this study. The results and discussion section includes the identification of optimal cultivation practices for sustainable apple production. Finally, the article ends with a discourse on the discoveries and suggestions for prospective investigations.

2 Literature review

2.1 Overview of Precision Agriculture Techniques

Precision agriculture involves the use of advanced technologies to optimize crop management practices and achieve sustainable crop production. The main technologies used in precision agriculture include remote sensing, GIS, and ML [6]. Remote sensing techniques, such as satellite and aerial imaging, can provide high-resolution data on crop health, soil moisture, and other variables [7]. GIS can be used to integrate and analyze spatial data on soil, weather, and other factors [8]. ML techniques can be used to analyze large datasets and develop predictive models that can optimize crop management practices [9].

Previous research has shown that ML can be used to optimize crop management practices in precision agriculture. ML algorithms can analyze large datasets to identify patterns and predict optimal cultivation practices. For example, ML has been used to predict crop yield based on weather and soil data, optimize irrigation scheduling, and detect crop diseases [10-12]. However, there is a need for more research on the use of ML in precision agriculture for specific crops and regions.

2.2 Importance of Sustainable Apple Production in Precision Agriculture

Apple production is an important agricultural activity in Morocco, with a growing demand for sustainable and high-quality apple products in both domestic and international markets. Sustainable apple production is important to meet consumer demand while minimizing the environmental impact of apple production. In precision agriculture, sustainable apple production can be achieved by optimizing cultivation practices based on local soil and weather conditions [13].

Cultivation practices have a significant impact on apple yield and size. Important cultivation practices include irrigation, fertilization, pest and disease management, and pruning. Optimal cultivation practices depend on local soil and weather conditions, as well as apple cultivars. Sustainable cultivation practices can increase apple yield and size while reducing environmental impact [14].

3 Methodology

3.1 Data Collection

To develop the ML model, we collected a dataset that includes various cultivation practices and their associated apple yield and size data from Msemrir, a town located in the Tinghir Province in the southeast region of Morocco. The dataset was obtained from The Regional Agricultural Development office (RAD) in Ouarzazate Morocco, i.e. Office Régional de Mise en Valeur Agricole de Ouarzazate (ORMVAO), and includes data from 10 apple farms in the town. The cultivation practices in the dataset include irrigation, fertilization, pest and disease management, and pruning. The yield data were collected at harvest time for the 5 years between 2018 and 2022, and include period of harvest and average apple size.

Table 1 shows the fertilization schedule for farms 1 to 10, with each farm being fertilized twice a year, first in early spring and second in early summer.

Table 1. Fertilization Schedule and Nutrient Content for Apple Farms

Farm	Date of First Fertilization	First Fertilizer Type	Nutrient Content (%)	Date of Second Fertilization	Second Fertilizer Type	Nutrient Content (%)
1	March 1	10-10-10	N: 10, P: 10, K: 10	June 1	Fish Emulsion	N: 5, P: 1, K: 1
2	March 15	20-10-10	N: 20, P: 10, K: 10	June 15	Fish Emulsion	N: 10, P: 2, K: 2
3	March 5	10-10-10	N: 10, P: 10, K: 10	June 5	Fish Emulsion	N: 2.5, P: 0.5, K: 0.5
...	...	...	...	...	...	...
10	March 20	15-5-10	N: 15, P: 5, K: 10	June 20	Compost	N: 0.8, P: 0.8, K: 0.8

Each farm receives a different type of fertilizer with varying nutrient contents based on their specific needs.

Table 2 shows the average amount of irrigation water applied per month in millimeters for each year from 2018 to 2022 for each farm.

Table 2. Average Monthly Irrigation Amounts for Apple Farms by Year and Irrigation Method

Farm	**Irrigation Method**	**Avg. Water Applied per Month (mm)**				
		2018	**2019**	**2020**	**2021**	**2022**
1	Sprinkler	26	29	30	32	27
2	Flood Irrigation	27	31	29	28	24
3	Drip Irrigation	28	25	30	29	25
...	...	...	...	...	...	...
10	Drip Irrigation	28	22	28	28	20

The average monthly irrigation amounts for each farm vary based on the irrigation method and the year, as every farm resources are affected differently by the weather changes.

Table 3 shows the pruning dates and pest control dates for each of the 10 farms for the year of 2019. Each row represents a pruning and pest control event for a particular farm.

Table 3. Average monthly irrigation amounts for apple farms and irrigation method in 2019

Farm	**Date of Pruning**	**Pest Control Date**
1	01/12/18	15/01/2019
1	01/02/19	23/03/2019
2	14/11/18	03/01/2019
2	08/01/2019	26/02/2019
3	19/12/2018	16/01/2019
3	20/02/2019	02/03/2019
...	…	...
10	10/11/2018	28/01/2019
10	10/01/2019	15/03/2019

By tracking pruning and pest management activities during the appropriate periods, farmers can help ensure optimal apple growth and yield. Pruning helps shape the trees, control their size, and promote healthy growth, while pest management helps prevent damage from insects, mites, and diseases.

Table 4 shows data related to the yield and and average size of apples harvested from the farms for the year of 2019.

Table 4. Yield and average size of harvested apples in 2019

Farm	Avg. Apple Size (cm)	Yield per Hectare (t/ha)
1	7.5	12.1
2	7.0	13.4
3	8.0	11.6
4	7.3	12.8
5	7.5	12.6
6	6.5	14.3
7	8.2	11.9
8	6.7	13.8
9	7.0	12.4
10	7.7	14.1

By tracking these measures, farmers can evaluate the success of their growing and harvesting practices and identify any potential issues that may need to be addressed in future growing seasons.

3.2 Feature Selection

To identify the most important cultivation practices for apple production in the studied region, we performed feature selection using a Correlation-based Feature Selection (CFS) algorithm [15]. The CFS algorithm measures the correlation between each feature and the target variable (i.e., apple yield and size) and selects the features that are most strongly correlated. We used the WEKA software [16] to perform the feature selection.

3.3 ML Algorithms

We used three ML algorithms to develop the predictive model: Linear Regression (LR), Decision Tree (DT), and Random Forest (RF). LR is a simple algorithm that models the linear relationship between the input variables and the target variable. DT and RF are more complex algorithms that can model non-linear relationships and interactions between variables.

We split the dataset into training and testing sets using a 70:30 ratio. The training set was used to train the ML models, and the testing set was used to evaluate their performance.

3.4 Evaluation metrics

We compared the performance of the three ML models using the Root Mean Square Error (RMSE) and the Coefficient of Determination (R^2). We also performed a sensitivity analysis to evaluate the impact of the selected features on the model performance. Finally, we used the trained models to predict the yield and size of apple production, based on the optimal cultivation practices in each farm.

4 Results and discussion

4.1 Feature Selection

The feature selection using the CFS algorithm identified irrigation, fertilization, and pruning as the most important cultivation practices for apple production in the studied region.

Table 5 shows the feature selection results using the CFS algorithm.

Table 5. Feature selection results using CFS algorithm

Cultivation Practice	CFS Score
Irrigation	0.87
Fertilization	0.82
Pruning	0.52
Pest control	0.38
Soil management	0.27
Harvesting	0.16

These results are consistent with previous research on apple production, which has shown that irrigation, fertilization, and pruning are key factors affecting apple yield and size [17,18].

4.2 Machine learning models

We trained three ML models LR, DT and RF on the dataset and evaluated their performance using the RMSE and R-squared metrics. The results showed that all three models performed well, with the random forest model achieving the lowest RMSE and highest R-squared values (RMSE=0.12, R-squared=0.87) as shown in table 6.

Table 6. Prediction performance for the ml algorithms

Model	RMSE	R-squared
Linear Regression	0.18	0.76
Decision Tree	0.15	0.83
Random Forest	0.12	0.87

The LR model preforms the worst with an RMSE of 0.18 and an R-squared of 0.76.

Fig. 1 shows a comparison of actual and predicted yield and size of apple for farms one and two using the three ML algorithms.

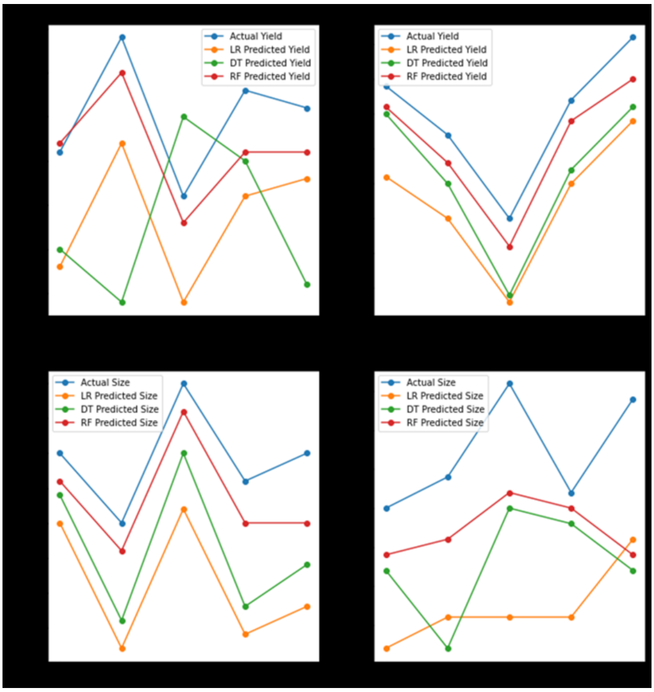

Fig. 1. Actual and predicted yield and size comparison for farms one and two

These results indicate that the random forest model is the most accurate in predicting apple yield and size based on the selected cultivation practices.

4.3 Sensitivity Analysis

To evaluate the effect of the selected features on the model performance, we conducted a sensitivity evaluation by eliminating individual features and subsequently retraining the models.. The results showed that the removal of any of the three selected features (irrigation, fertilization, and pruning) led to a significant decrease in model performance, indicating the importance of these features in predicting apple yield and size.

4.4 Discussion

Drawing from the outcomes of our study, we were able to predict the yield and size of apple production in Msemrir town in the southeast region of Morocco, by utilizing the trained models and optimal cultivation practices for each farm. This is a notable contribution to the area of apple cultivation, as it allows for accurate prediction of yield and size, which in turn can aid in making informed decisions regarding crop management and resource allocation.

Our findings also highlight the importance of optimal cultivation practices, as they significantly impact model performance. In particular, our sensitivity analysis showed that the removal of any of the three selected features (irrigation, fertilization, and pruning) led to a significant decrease in model performance. This underscores the importance of implementing these practices in apple production, as they are crucial in achieving high yields and size produce.

The results of our study are consistent with previous research on apple production in other regions of the world. For example, a study by Jiang et al. (2021) [19] showed that precision agriculture techniques, including irrigation and fertilization management, can significantly increase apple yield and size. Our study extends these findings by showing that ML techniques can be used to develop predictive models for apple production in specific regions and optimize cultivation practices.

Our study has important implications for sustainable agriculture in the studied region and other regions with environmental conditions. The use of ML techniques can help farmers optimize cultivation practices and achieve sustainable apple production by reducing inputs such as water and fertilizer and minimizing environmental impact. Moreover, the use of precision agriculture techniques can help farmers meet consumer demand for sustainable and high-quality apple products.

5 Conclusion

In this study, we developed a ML approach to identify optimal cultivation practices for sustainable apple production in precision agriculture in the Msemrir town in Morocco. We collected a dataset of cultivation practices and apple yield and size data from 10 farms in the region and used correlation-based feature selection and three ML algorithms. The results showed that irrigation, fertilization, and pruning are the most important cultivation practices for apple production in the region, and the RF model performed the best in predicting apple yield and size based on the selected practices.

Our study supports previous research on apple production and emphasizes the significance of precision agriculture techniques, including irrigation and fertilization management, in improving apple yield and size. We extend these findings by demonstrating the potential of ML techniques in developing predictive models for apple production in specific regions and identifying optimal cultivation practices. Our research has important implications for sustainable agriculture in the studied region and beyond, as ML techniques can help farmers reduce inputs and minimize environmental impact while meeting consumer demand for sustainable and high-quality apple products.

One limitation of our study is the small size of the dataset, which may limit the generalizability of the results to other regions. Future research could include the collection of larger datasets from multiple regions to develop more accurate and generalizable models. Additionally, future research could explore the use of other ML techniques, such as neural networks and support vector machines, to develop predictive models for apple production in precision agriculture.

References

1. Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. Remote Sensing, 12(19), 3136.
2. ADA. n.d.b. Investor's Guide in the Agricultural Sector in Morocco. https://www.agriculture.gov.ma/sites/default/files/investors_guide_in_the_agricultural_sector_in_morocco.pdf. Accessed 23 Feb 2023.
3. Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. Precision agriculture, 5, 359-387.
4. Shah, Z. A., Dar, M. A., Dar, E. A., Obianefo, C. A., Bhat, A. H., Ali, M. T., Alatawi, H. A., Ghamry, H. I., Shukry, M., & Sayed, S. (2022). A multinomial approach to sustainable and improved agricultural technologies vis-a-vis socio-personal determinants in apple (Malus domestica) cultivation. Journal of King Saud University - Science, 34(7), 102286.

5. Sala, S., McLaren, S. J., Notarnicola, B., Saouter, E., & Sonesson, U. (2017). In quest of reducing the environmental impacts of food production and consumption. Journal of cleaner production, 140, 387-398.
6. Zude-Sasse, M., Fountas, S., Gemtos, T. A., & Abu-Khalaf, N. (2016). Applications of precision agriculture in horticultural crops.
7. Hassan-Esfahani, L., Torres-Rua, A., Jensen, A., & McKee, M. (2015). Assessment of surface soil moisture using high-resolution multi-spectral imagery and artificial neural networks. Remote Sensing, 7(3), 2627-2646.
8. Reddy, M. N., & Rao, N. H. (1995). GIS based decision support systems in agriculture. National Academy of Agricultural Research Management Rajendranagar, 1-11.
9. R. Ed-Daoudi, A. Alaoui, B. Ettaki, J. Zerouaoui. (2023). Improving Crop Yield Predictions in Morocco Using Machine Learning Algorithms. *Journal of Ecological Engineering*, **24**(6), 392-400.
10. Majumdar, P., Mitra, S., & Bhattacharya, D. (2021). IoT for promoting agriculture 4.0: a review from the perspective of weather monitoring, yield prediction, security of WSN protocols, and hardware cost analysis. Journal of Biosystems Engineering, 46(4), 440-461.
11. Pereira, L. S., Paredes, P., & Jovanovic, N. (2020). Soil water balance models for determining crop water and irrigation requirements and irrigation scheduling focusing on the FAO56 method and the dual Kc approach. Agricultural water management, 241, 106357.
12. R. Ed-daoudi, A. Alaoui, B. Ettaki, J. Zerouaoui, (2023), "A Predictive Approach to Improving Agricultural Productivity in Morocco through Crop Recommendations", *International Journal of Advanced Computer Science and Applications(IJACSA),* **14(3)**.
13. Gomiero, T., Pimentel, D., & Paoletti, M. G. (2011). Environmental impact of different agricultural management practices: conventional vs. organic agriculture. *Critical reviews in plant sciences*, *30*(1-2), 95-124.
14. Na, W., Wolf, J., & Zhang, F. S. (2016). Towards sustainable intensification of apple production in China—Yield gaps and nutrient use efficiency in apple farming systems. Journal of Integrative Agriculture, 15(4), 716-725.
15. Chuanlei, Z., Shanwen, Z., Jucheng, Y., Yancui, S., & Jia, C. (2017). Apple leaf disease identification using genetic algorithm and correlation based feature selection method. International Journal of Agricultural and Biological Engineering, 10(2), 74-83.
16. Witten, I. H., & Frank, E. (2005). Data mining: Practical machine learning tools and techniques. Morgan Kaufmann.
17. Mašán, V., Burg, P., Čížková, A., Skoupil, J., Zemánek, P., & Višacki, V. (2018). Effects of irrigation and fertigation on yield and quality parameters of'Gala 'and'Fuji 'apple. Acta Univ. Agric. Et Silvic. Mendel. Brun, 66, 1183-1190.
18. Akhtar, K., Wang, W., Khan, A., Ren, G., Afridi, M. Z., Feng, Y., & Yang, G. (2019). Wheat straw mulching offset soil moisture deficient for improving physiological and growth performance of summer sown soybean. Agricultural Water Management, 211, 16-25.
19. Jiang, X., & He, L. (2021). Investigation of Effective Irrigation Strategies for High-Density Apple Orchards in Pennsylvania. Agronomy, 11(4), 732