Energy Resource Scheduling in an Agriculture System Using a Decision Tree Approach

Omid Abrishambaf, Pedro Faria, Zita Vale
Polytechnic of Porto (IPP), Porto, Portugal
ombaf@isep.ipp.pt, pnf@isep.ipp.pt, zav@isep.ipp.pt

***Abstract*—Agriculture sector is backbone of each country. Nowadays the energy efficiency in this sector is at a very low level, which shows the necessity of more investments in this regard. By appearance of smart grid technologies, some new concepts were also appeared in the agriculture sector, such as smart farm, and smart agriculture. This paper provides an energy management system for an agriculture field equipped with renewable energy resources and a river turbine. A decision tree is developed in this paper to schedule and optimize the use of energy resources for reducing the electricity costs. Decision tree method enables the system to obtain optimal scheduling of energy resources in offline mode, without using any external server/machine or internet access. A case study validates the performance of developed decision tree, and the errors and accuracy of all gained results are discussed.**

***Index Terms*— Agriculture system, Decision tree, Energy scheduling, Renewable resources, River turbine.**

I. Introduction

Energy demand has been increasing in the last decades, together with food demand. These two elements are the bases of each society, which the increment of their demands brings several challenges. It is estimated that the demands of energy and foods will be increased by more 40% and 50% respectively till 2030 [1]. This makes the role of the agriculture system more tangible than before. In fact, agriculture is a fundamental sector of each country that nowadays requires more investment and development to have a contribution to the energy management systems [2]. Studies show energy consumption efficiency in agriculture systems is still at a low level. This means the challenges of energy consumption and its efficiency should be tackled [3]. Furthermore, studies indicate the agriculture systems are in charge of 15%-25% greenhouse gas emission. Therefore, traditional agriculture system should be equipped with Renewable Energy Resources (RERs) to become environmentally friendly and cost-effective [4].

Since the new concepts of power system, such as smart grids and microgrids, were implemented, the use of smart technologies in the agriculture sector attracted a lot of attention. In fact, the use of intelligent systems is not only limited to the smart grids, and the agriculture systems can also benefit from these smart technologies, which leads to creating more new concepts, namely smart agriculture or smart farm [3]. In this way, all electricity consumers and prosumers in the agriculture sector would be intelligent enough in order to participate in the network management scenarios, such as demand response programs [5][6].

By a simple look on the current literature, it can be found that there is a huge range of research works focused on smart agriculture systems using the Internet of Things (IoT) and Wireless Sensor Network (WSN) [7]. This is an interesting and hot topic in the current advancement of smart grid technologies. Now, what happens when there is no internet access in the field? This is very common in the agriculture sector, since they may be located in distant areas. Therefore, the smart agriculture system should also be capable to perform decision making in offline mode, without using any external server/machine or internet access. This shows the need for developing an offline agriculture system that can optimally use the available energy resources. Decision Tree (DT) approach is a solution in this context, which employs "if-then" method [8] and can be implemented in any types of controllers or programming languages [9][10].

This paper presents an energy management system for an agriculture field, which perform energy scheduling using DT approach. The system is equipped with RERs and a river turbine enabling the system to supply the electricity demand from the local resources in order to minimize the electricity costs.

There are several similar works in this context. In [11], the author presented a framework to optimize energy management of an agricultural microgrid equipped with RERs and a pumped-storage unit. In [12] is proposed to improve the efficiency of the irrigation based on the water requirement. In [13], the authors provided a DT approach for energy resource management of a community of consumers to minimize the operational costs. However, the focus of this paper is to develop a DT for the agriculture system that can manage the energy resources in offline mode, without using any external server/machine or internet access.

After this introductory section, the proposed agriculture system is shown in Section II. Section III presents a case study to validate the performance of the model, and its results are shown in Section IV. Conclusions are presented in Section V.

The present work was done and funded in the scope of project: Eco Rural IoT (TETRAMAX-VALUECHAIN-TTX-1), CEECIND/02887/2017, and UID/EEA/00760/2019 funded by FEDER Funds through COMPETE and by National Funds through FCT.

II. MODEL DESCRIPTION

This section focuses on the developed method for energy resource scheduling in the agriculture field. The system utilizes a DT methodology to perform scheduling of the resources in order to have the most optimal and economical solution. For this purpose, firstly the agriculture system considered in this paper is demonstrated, and then, the employed DT approach will be described.

Fig. 1 illustrates an overall view of the agriculture system presented in this paper. Since the system is equipped with renewable resources, the most important intention of the model is to use the own produced energy for the local electricity demand. There are several models regarding energy scheduling in the agriculture system. As an example, the model described on [11], used several RERs to supply a local agriculture microgrid. In the same work, an energy management system has been developed to optimally schedule the irrigation based on several input data, such market price and real-time generation rate.

However, in this paper (as Fig. 1 shows) the energy resources are Photovoltaic (PV) Panels and a synchronous generator. The main focus of this system is given to the synchronous generator. In this regard, it is considered that the agriculture field is located next to a river, where a turbine is placed on the river rotating the shaft of the synchronous generator.

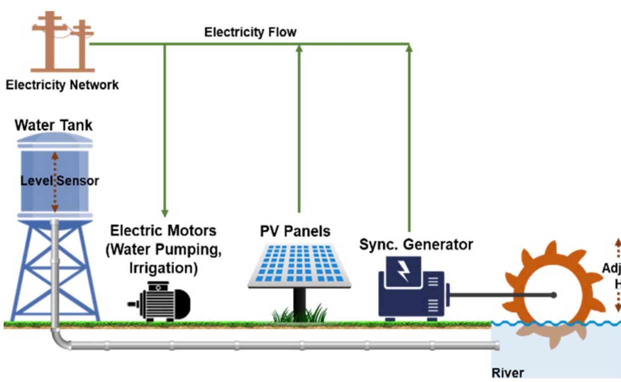

Fig. 1. The overall perspective of the proposed agriculture system.

Electric motors are responsible for irrigating, and for pumping water to a water tank. Also, there is a level sensor inside of the water tank that enables the system to turn the pumping motor off when the tank is full.

In the main core of the system, it is considered that there is a Programmable Logic Controller (PLC), and several energy meters for managing the proposed system. In fact, the PLC manages the system based on the monitoring of consumption and generation of the resources. Also, the level sensor in the water tank indicates that the tank is empty or full. Furthermore, the PLC controls the plunging level of river turbine to regulate the speed of the synchronous generator´s shaft.

Also, the PLC controls the status of the water pump to turn on/off. In fact, while the river turbine is plunged in the deeper height of the river, the rotation speed is increased, and therefore, the generation rate will be higher. Another feature of the proposed agriculture system is that the electricity tariffs are pre-defined in the PLC as some decision rules. This means, the system is aware of the electricity market prices, and therefore, it can take the best optimal decision for the system for using the energy resources.

As it was mentioned, the main intention of the system is to supply the electricity demand from the local energy resources (PV and synchronous generator). Therefore, it is necessary to have a decision algorithm in the PLC to perform the optimal operation. For this purpose, the DT methodology has been selected, since it can easily be implemented in the PLC or any other controller. Therefore, it is avoided the communicating with a server as all the decisions are made locally by the PLC. Fig. 2 demonstrates the set of the rules used by DT for supplying the load demand.

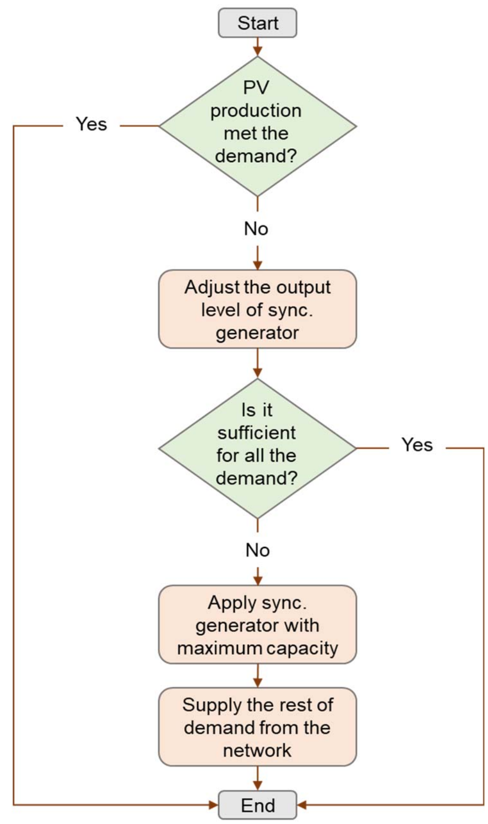

Fig. 2. The defined sequence of supplying the electricity demand.

The first rule defined for the DT is based on the amount of PV production. Since the energy produced by the PV panels is not controllable, it will be the priority of the system to supply the loads from that produced energy. After then, if the PV production is not sufficient for the load, the PLC adjusts the energy production of the synchronous generator by regulating the plunge height of river turbine. In the final stage, if both PV and synchronous generator are not enough for the consumption, the system utilizes electricity network to supply the rest of consumption. In other words, the output of the DT would be a set of rules specifying the amount of required energy to be produced by the synchronous generator.

In this paper, the DT is built by RStudio tools (www.rstudio.com) using the RPART package. In fact, RPART stands for Recursive PARTitioning, which is a statistical approach for multivariable analysis. The RPART includes various methods that can be optionally chosen by the user. If no method is specified in the model, the routine attempts to intelligently select the most suitable method. As an example, since in this paper the variables are as factors, the routine chooses "class" method for creating the DT, which operates based on the classification of the data.

III. Case Study

In this part, a case study is explained for validating the functionalities of the developed model for the energy resource scheduling in the agriculture field using DT. Therefore, various conditions of the system with a significant number of data should be provided to the DT as a dataset to gain the most optimal approach.

A maximum capacity rate is specified for each resource of the system. In this regard, it is considered that the PV has a maximum production capacity of 7 kW, the synchronous generator has a nominal generation capacity of 3 kW, and the electric motors consume 6 kW in the maximum consumption capacity. These specified values are only assumption for this case study, and they may be varied in the real infrastructures.

For creating the dataset for DT, 5 different conditions of the system including 5 groups of data are considered. Table I shows the dataset configuration of DT. The PV production profiles (A to E) in Table I, are shown in Fig. 3. All information of this case study is for 96 periods of 15 minutes, presenting 24 hours a day.

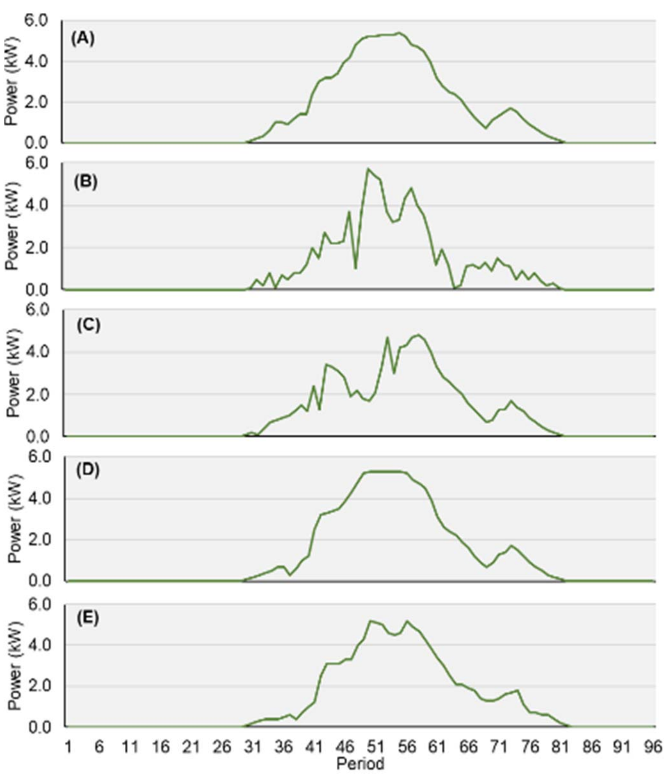

Fig. 3. Considered PV profiles.

The profiles demonstrated in Fig. 3 are real data for 5 different days adapted from PV system of GECAD research center, in Porto, Portugal. Moreover, Fig. 4 shows the electricity prices (as it also shown on Table I), which are based on Time Of Use (TOU) tariffs of incumbent Portuguese electricity retailer (EDP Commercial – www.edp.pt).

TABLE I. DATASET CONFIGURATION FOR DT.

	Group 1 (Base)	**Group 2**	**Group 3**	**Group 4**	**Group 5**
Market Price	0.15/0.25 EUR	Base × 0.8	Base × 0.9	Base × 1.1	Base × 1.2
Electric Motors	5 kW	Base × 0.8	Base × 0.9	Base × 1.1	Base × 1.2
PV	Profile (A)	Profile (B)	Profile (C)	Profile (D)	Profile (E)
Tank Level	Empty (0)	Empty (0)	Full (1)	Full (1)	Full (1)

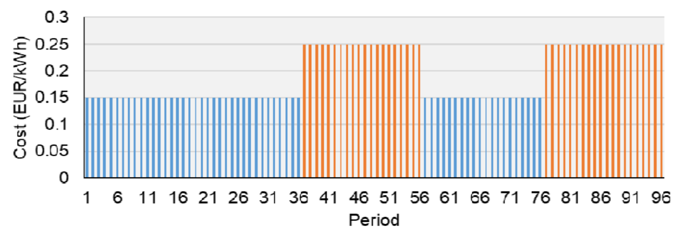

Fig. 4. Electricity market prices.

From the combination of data shown on Table I, a dataset including 60000 possibilities are created and provided to the DT as input. The output of the DT would be a set of rules specifying the amount of required energy to be produced by the synchronous generator. In order to validate and compare the results of the DT, two scenarios are proposed. In the first scenario, it is considered that the system is able to control the Tank Level on two levels: empty or full (as shown in Table I). In the second scenario, it is assumed that the system can control the Tank Level on five levels: 0%, 25%, 50%, 75%, and 100% respectively in Group 1 to 5. Therefore, there would be more complexity in the second scenario. The rest of the data in both scenarios are considered as same as each other, as demonstrated in Table I.

IV. Results

In this section, firstly, the errors of DT in each scenario has been calculated separately and compared to each other, as Table II shows. Then, the focus is given to the DT with more precise results and less amount of errors.

TABLE II. ERRORS CALCULATION IN TWO PROPOSED SCENARIOS.

	Results Error	**No. Splits**	**Rel error**	**xerror**	**xstd**
Scenario 1	0.0208	22	0.2761	0.2777	0.0044
Scenario 2	0.0255	18	0.3428	0.3467	0.0049

In Table II, Rel error is a relative error, xerror is the cross-validation error, and the xstd is standard deviation error. As Table II shows the DT of scenario 1 provides more accurate results with less amount of errors. The number of splitting nodes in scenario 1 is higher than scenario 2, which improves the performance of DT with fewer errors. Fig. 5 illustrates the entire DT of scenario 1. All the values shown in Fig. 5 are in kW, except market price, which is EUR/kWh.

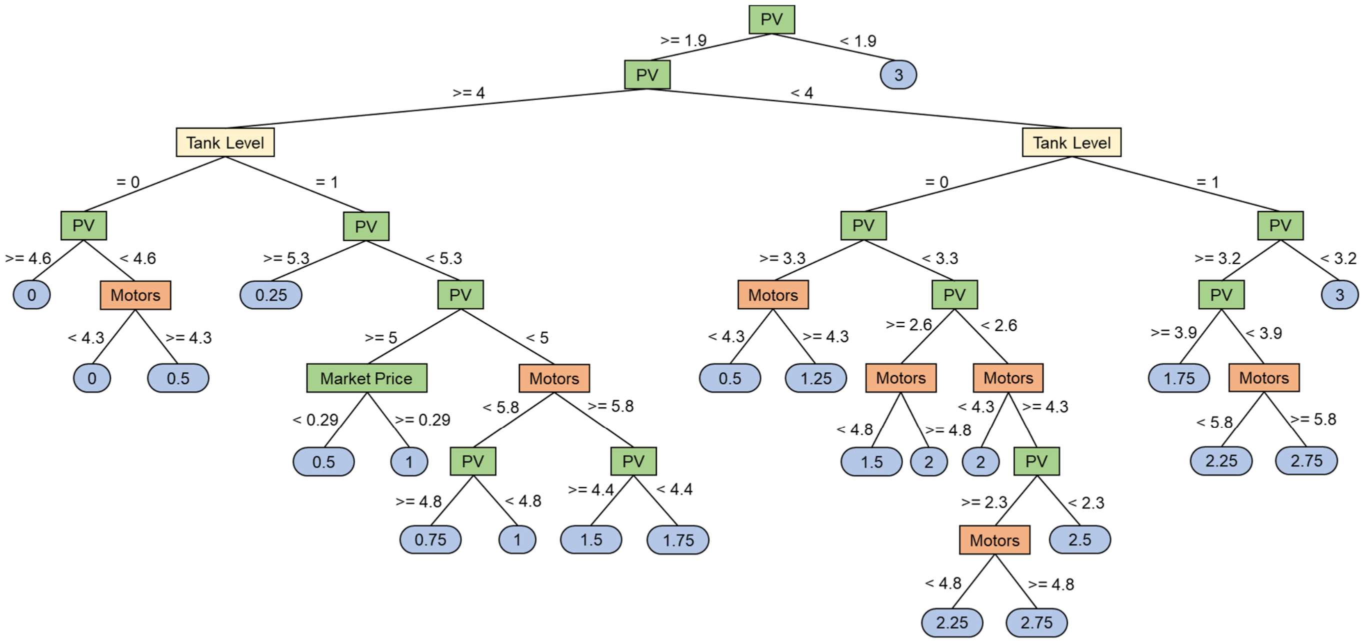

Fig. 5 – Illustration of the DT in scenario 1 for the energy resource scheduling of the proposed model.

In fact, Fig. 5 shows all the nodes, branches, and leaves of the developed DT. The information demonstrated in each node is the employed variables and thresholds that have been utilized for data classification. The terminal nodes (blue nodes in Fig. 5) demonstrate the predicted solution on that node, which is computed based on the average solution of the training data that were provided to the tree as input. In practice, the output of the DT (colored by blue in Fig. 5) is the amount of energy targeted for the generator. The decision rules used in this DT are:

x1—PV;
x2—Electric Motors;
x3—Market Price;
x4—Tank Level;
Y—Synchronous Generator (output).

Decision Rules for Energy Resource Scheduling.

1)	when **x1 >= 4.6 & x4 is 0**:	Y = 0
2)	when **x1 is 4.0 to 4.6 & x4 is 0 & x2 < 4.3**:	Y = 0
3)	when **x1 >= 5.3 & x4 is 1**:	Y = 0.25
4)	when **x1 is 4.0 to 4.6 & x4 is 0 & x2 >= 4.3**:	Y = 0.5
5)	when **x1 is 3.3 to 4.0 & x4 is 0 & x2 < 4.3**:	Y = 0.5
6)	when **x1 is 5.0 to 5.3 & x4 is 1 & x3 < 0.29**:	Y = 0.5
7)	when **x1 is 4.8 to 5.0 & x4 is 1 & x2 < 5.8**:	Y = 0.75
8)	when **x1 is 4.0 to 4.8 & x4 is 1 & x2 < 5.8**:	Y = 1
9)	when **x1 is 5.0 to 5.3 & x4 is 1 & x3 >= 0.29**:	Y = 1
10)	when **x1 is 3.3 to 4.0 & x4 is 0 & x2 >= 4.3**:	Y = 1.25
11)	when **x1 is 2.6 to 3.3 & x4 is 0 & x2 < 4.8**:	Y = 1.5
12)	when **x1 is 4.4 to 5.0 & x4 is 1 & x2 >= 5.8**:	Y = 1.5
13)	when **x1 is 4.0 to 4.4 & x4 is 1 & x2 >= 5.8**:	Y = 1.75
14)	when **x1 is 3.9 to 4.0 & x4 is 1**:	Y = 1.75
15)	when **x1 is 1.9 to 2.6 & x4 is 0 & x2 < 4.3**:	Y = 2
16)	when **x1 is 2.6 to 3.3 & x4 is 0 & x2 >= 4.8**:	Y = 2
17)	when **x1 is 3.2 to 3.9 & x4 is 1 & x2 < 5.8**:	Y = 2.25
18)	when **x1 is 2.3 to 2.6 & x4 is 0 & x2 is 4.3 to 4.8**:	Y = 2.25
19)	when **x1 is 1.9 to 2.3 & x4 is 0 & x2 >= 4.3**:	Y = 2.5
20)	when **x1 is 2.3 to 2.6 & x4 is 0 & x2 >= 4.8**:	Y = 2.75
21)	when **x1 is 3.2 to 3.9 & x4 is 1 & x2 >= 5.8**:	Y = 2.75
22)	when **x1 is 1.9 to 3.2 & x4 is 1**:	Y = 3
23)	when **x1 < 1.9**:	Y = 3

As it was expected and shown on Fig. 5, PV and Electrical Motors are the two most effective variables of the DT. More specifically, while all predictive variables have been scaled to sum to 100%, the importance of PV variable in the DT is 63%, the importance of Electric Motors is 17%, Market Price is 13%, and Tank Level is 7%.

Regarding the errors of the model, Table III demonstrates the pruning from RPART algorithm for DT. As it was explained previously, Rel error is a relative error, and xerror is the cross-validation error. In fact, the Rel error indicates the prediction error of the data that was utilized to make the tree, and the xerror is the amount error produced by the RPART built-in cross-validation. In Table III, each row indicates the depth of the tree and the corresponding calculations performed on that level. In the second column of the same table, CP stands for Complexity Parameter, which is a value in each depth of the tree to perform divisions in the nodes until the relative error (Rel error in the fourth column) goes below a desired rate.

TABLE III. PRUNING TABLE OF THE PROPOSED DT IN SCENARIO 1.

Depth	CP	No. splits	Rel error	xerror	xstd
1	0.0857	0	1.0000	1.0000	0.0077
2	0.0666	2	0.8285	0.8285	0.0071
3	0.0571	3	0.7619	0.7619	0.0069
4	0.0476	5	0.6476	0.6476	0.0065
5	0.0380	6	0.6000	0.6000	0.0063
6	0.0285	7	0.5619	0.5619	0.0061
7	0.0222	8	0.5333	0.5352	0.0060
8	0.0190	11	0.4666	0.4685	0.0056
9	0.0142	18	0.3333	0.3348	0.0048
10	0.0100	22	0.2761	0.2777	0.0044

Fig. 6 shows this process more clearly for both proposed scenarios. Also, as it is clear in Table III, the number of splits in each level is increased until the algorithm reaches the optimal level with less error.

Actually, Fig. 6 illustrates a summary of the calculation process based on the rel. error and CP for indicating the most optimal size of the tree in scenario 1 and 2. The dashed line in Fig. 6 is a certain rate of rel. error that the algorithm should reach for the most optimal results. In other words, the point that the two lines in Fig. 6 have crossed each other, is the most optimal size of the tree (23 terminal nodes in scenario 1).

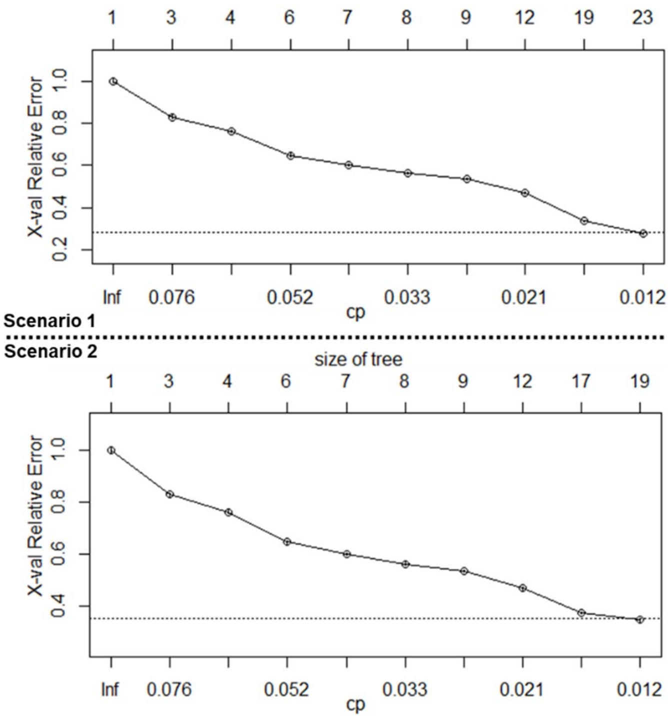

Fig. 6. Pruning plot of the DT in both scenarios.

V. CONCLUSIONS

In this paper, an energy management system was proposed for an agriculture field. A decision tree was presented for the system to provide the optimal scheduling of energy resources for minimizing the electricity costs. The agriculture system was equipped with renewable energy resources and a synchronous generator connected to a river turbine. In fact, the output of the decision tree was the amount of required energy to be produced by the synchronous generator. In this way, the system was able to prevent purchasing energy from the utility grid as much as possible, since it can supply the electricity demand from the local produced energy.

The case study validated the functionalities of the proposed model. Several conditions with real groups of data adapted from the real resources have been provided to the developed decision tree, and its performance was surveyed. Furthermore, the errors of the results and the precision of them were calculated and demonstrated. The decision tree approach can perform optimization and resource scheduling in offline mode regarding using external server. This is a very important factor in practice, since most of the agriculture fields are in distant areas without access to the internet.

For the future work, it is intended to consider an energy storage system in the related agriculture model, to provide more flexibility.

REFERENCES

[1] M. Li, Q. Fu, V. P. Singh, D. Liu, and T. Li, "Stochastic multi-objective modeling for optimization of water-food-energy nexus of irrigated agriculture," *Adv. Water Resour.*, vol. 127, pp. 209–224, May 2019.

[2] A. O. Avcıoğlu, M. A. Dayıoğlu, and U. Türker, "Assessment of the energy potential of agricultural biomass residues in Turkey," *Renew. Energy*, vol. 138, pp. 610–619, Aug. 2019.

[3] M. Mohammadi, Y. Noorollahi, B. Mohammadi-ivatloo, M. Hosseinzadeh, H. Yousefi, and S. T. Khorasani, "Optimal management of energy hubs and smart energy hubs – A review," *Renew. Sustain. Energy Rev.*, vol. 89, pp. 33–50, Jun. 2018.

[4] J. Farfan, A. Lohrmann, and C. Breyer, "Integration of greenhouse agriculture to the energy infrastructure as an alimentary solution," *Renew. Sustain. Energy Rev.*, vol. 110, pp. 368–377, Aug. 2019.

[5] O. Abrishambaf, P. Faria, and Z. Vale, "SCADA Office Building Implementation in the Context of an Aggregator," in *2018 IEEE 16th International Conference on Industrial Informatics (INDIN)*, 2018, pp. 984–989.

[6] O. Abrishambaf, P. Faria, Z. Vale, and J. M. Corchado, "Real-Time Simulation of a Curtailment Service Provider for Demand Response Participation," in *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T&D)*, 2018, pp. 1–9.

[7] Đ. Banđur, B. Jakšić, M. Banđur, and S. Jović, "An analysis of energy efficiency in Wireless Sensor Networks (WSNs) applied in smart agriculture," *Comput. Electron. Agric.*, vol. 156, pp. 500–507, Jan. 2019.

[8] M. Amayri and S. Ploix, "Decision tree and Parametrized classifier for Estimating occupancy in energy management," in *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2018, pp. 397–402.

[9] N. Putjaika, S. Phusae, A. Chen-Im, P. Phunchongharn, and K. Akkarajitsakul, "A control system in an intelligent farming by using arduino technology," in *2016 Fifth ICT International Student Project Conference (ICT-ISPC)*, 2016, pp. 53–56.

[10] L. He, Z. Xiaoxian, C. Liying, and L. Dexin, "A study on evaluation of farmland fertility levels based on optimization of the decision tree algorithm," in *Proceedings of 2012 2nd International Conference on Computer Science and Network Technology*, 2012, pp. 158–161.

[11] A. Ghasemi, "Coordination of pumped-storage unit and irrigation system with intermittent wind generation for intelligent energy management of an agricultural microgrid," *Energy*, vol. 142, pp. 1–13, Jan. 2018.

[12] O. Abrishambaf, P. Faria, L. Gomes, and Z. Vale, "Agricultural Irrigation Scheduling For a Crop Management System Considering Water and Energy Use Optimization," in *ICEER2019 - 6th International Conference on Energy and Environment Research: "Energy and environment: challenges towards circular economy*, 2019.

[13] O. Abrishambaf, P. Faria, and Z. Vale, "Participation of a Smart Community of Consumers in Demand Response Programs," in *2018 Clemson University Power Systems Conference (PSC)*, 2018, pp. 1–5.