Article

Energy Scheduling Using Decision Trees and Emulation: Agriculture Irrigation with Run-of-the-River Hydroelectricity and a PV Case Study

Omid Abrishambaf 1, Pedro Faria 1,*, Zita Vale 2 and Juan M. Corchado 3

1 GECAD—Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, IPP—Polytechnic Institute of Porto, Rua DR. Antonio Bernardino de Almeida, 431, 4200-072 Porto, Portugal; ombaf@isep.ipp.pt
2 IPP—Polytechnic Institute of Porto, Rua DR. Antonio Bernardino de Almeida, 431, 4200-072 Porto, Portugal; zav@isep.ipp.pt
3 Department of Computing Science and Automation, USAL—University of Salamanca, 37008 Salamanca, Spain; corchado@usal.es

* Correspondence: pnf@isep.ipp.pt; Tel.: +351-228-340-511; Fax: +351-228-321-159

Received: 13 August 2019; Accepted: 19 October 2019; Published: 20 October 2019

Abstract: Agriculture is the very backbone of every country. Unfortunately, agricultural sustainability is threatened by the lack of energy-efficient solutions. The threat becomes more evident with the constantly growing world population. The research community must, therefore, focus on resolving the problem of high energy consumption. This paper proposes a model of energy scheduling in agricultural contexts. Greater energy efficiency is achieved by means of PV (photovoltaics) and hydropower, as demonstrated in the conducted case study. The developed model is intended for contexts where the farm is located near a river, so the farmer can use the flowing water to produce energy. Moreover, the model has been emulated using a variety of state-of-the-art laboratory devices. Optimal energy scheduling is performed via a decision tree approach, optimizing the use of energy resources and reducing electricity costs. Finally, a realistic scenario is presented to show the technical features and the practical behaviors of each emulator when adapting the results of the decision tree. The research outcomes demonstrate the importance of the technical validation of each model. In addition, the results of the emulation reveal practical issues that had not been discovered during the theoretical study or during the simulation.

Keywords: agriculture; decision tree; energy scheduling; hydropower; renewables

1. Introduction

Over the last decades, the world population has been incrementing on a daily basis. This, in turn, has led to greater energy and food demands. According to [1], the population is going to continue to soar, reaching somewhere around 10 billion by 2050. It is also expected that by 2035, global energy consumption is going to rise by up to 50% [2], the predicted consumption rise is largely attributed to the agriculture sector and the food supply chain. Food and energy are intrinsic elements of every society. The growing demand for both makes evident the importance of the agriculture sector [3]. In fact, agriculture is the very backbone of every country. At present, it is necessary to invest in the development of energy management systems [4,5]. Today’s agricultural practices still lack energy optimization measures, leading to high energy consumption [6]. It is therefore necessary that the research community address this problem.
The agriculture industry is considered to be the second-largest emitter of greenhouse gases (GHGs) [7], accounting for 21% of the total GHG emissions [8,9]. The reports of the U.S. emission inventory state that methane is the main GHG produced during agricultural practices such as soil management and livestock production. The statistics given in those reports also point to the use of fossil fuels as a significant contributor to GHG emissions in the agriculture sector (between 14% and 30%), where fossil fuels are used in nitrogen-rich fertilizer and to pump water and irrigate crops [10,11]. This means that the agriculture sector can contribute to reducing total GHG emissions by adopting a series of measures, such as exchanging fossil fuels for renewable energy resources (RERs) [8].

More specifically, in the agriculture industry, the energy demand for tactical services, namely, irrigation and water pumping, can be supplied by RERs instead of non-renewable sources [10]. Current agriculture systems should all employ RERs to become environmentally friendly and cost-effective [12,13].

New power system concepts, such as smart grids and microgrids, involve distributed generation (DG) and RERs on the demand side [14]. The agricultural sector is ideal for implementing all types of solutions involving renewable energy resources [15] because farms are located in rural areas where such resources can be easily accessed. Hydro, wind, and solar energy can be used to produce electricity for agricultural purposes, enabling farmers to become largely independent of the utility grid and to reduce electricity costs. The use of smart technologies in the agriculture sector has already attracted a lot of attention. In fact, the use of intelligent systems is not only limited to the electrical grid. Smart solutions can be designed exclusively for the agricultural sector, and this line of research has led to the emergence of a new concept, called smart agriculture or smart farming [6]. This concept also involves the application of the Internet of Things (IoT) in agriculture [16,17]. The ability to combine IoT and other technological paradigms has opened many new possibilities in agriculture, facilitating all types of agricultural practices. An important aspect in agriculture is the weather because crop yields largely depend on it. It is therefore necessary to forecast precipitation probability, air humidity, wind, and solar radiation to enable farmers to take tactical agricultural decisions, namely, field preparation, sowing/planting, irrigation, etc. [18]. For example, forecasting systems have been merged with IoT to gather weather- or field-related data and analyze it to create agriculture-specific classifications and forecast models that are highly accurate [19]. Another important aspect in agriculture is reducing costs, thus, energy production forecasting [20] and energy demand prediction [21] are also essential to minimize energy costs in smart agriculture models. Such models, in other words, enable electricity consumers and prosumers in the agriculture sector to participate in network management scenarios, such as demand response programs [22].

A survey of current literature reveals that much research has been conducted on smart agriculture systems using IoT and wireless sensor networks (WSNs) [23]. Thus, it is an interesting subject undergoing intense study. However, what if there is no internet and mobile network access where the farm is located? Since farms are often situated in remote areas, it is very common for them to lack access to the internet and to high-performance computing machines. Therefore, smart agricultural systems should be capable of operating 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. In this context, the decision tree (DT) approach employs an “if–then” method [24] and can be implemented in any type of controller or programming language [25,26]. However, prior to implementing a model on a massive scale, an emulation phase is required to test and validate the performance of the model in practice. In fact, the emulation phase makes it possible to discover the technical problems experienced by the system, which mostly remain hidden in the simulation and theoretical phases [27].

This paper describes a case study on energy scheduling in agriculture, where PV and hydropower are the available renewable energy resources for energy generation. In the proposed model, it is assumed that the agriculture field is located near a river, so the farmer can benefit from the running water to install a hydropower turbine and produce energy. The model was tested by means of state-of-the-art laboratory devices that emulated energy consumption and energy generation, as well as the typical
agricultural tools such as irrigation equipment and water pumping motors. In addition, several automation mechanisms are presented in this paper. These mechanisms were implemented in the emulators to provide a fully automated model for energy scheduling in agriculture. In this model, energy scheduling is performed by a developed DT that considers a series of factors for the optimal use of resources, including real-time generation and consumption rates, as well as electricity markets. Finally, the presented laboratory model was validated in a realistic scenario, and the performance of the DT in real-time energy scheduling is discussed in detail. The emulated model was run for a short period of time in a laboratory, and real results acquired from real resources were used as inputs to test and validate the performance of each emulator.

There are similar research works that focus on this context. In [28], the authors developed and compared three stand-alone models of an agricultural energy management system employing a hybrid wind-photovoltaic (PV) system. The proposed methodology involved the use an energy storage system that maintained the state-of-the-charge at a maximum level and simultaneously kept the water tank full. The results of this research indicated that, in comparison to other energy management strategies, optimal performance was achieved by the strategy that prioritized battery charging. In [29], a coordination framework was presented to optimize energy usage in an agricultural microgrid. A microgrid is capable of interacting with the energy market and involves the use of energy resources (in this case, renewables were used). The proposed model received forecasts of hourly microgrid consumption, wind power, and market prices to optimally schedule the operation of the irrigation system, pumped-storage unit, and energy trading with the utility grid. The numerical results demonstrated that the optimal operation of the pumped-storage unit can reduce the overall electricity costs of the microgrid. An autonomous approach was presented in [30] to achieve optimal and efficient irrigation scheduling in agriculture. This model gathered data for a series of crop-related parameters, making it possible to calculate evapotranspiration without a lysimeter. The proposed irrigation scheduling was performed using electricity market prices; local RERs irrigated the crops in low-cost periods, at time intervals that prevented the crops from reaching the wilting point.

In [31], the authors proposed a forecasting system with the ability to learn from past cases of forecast errors. The learning feature made it possible for the system to improve its forecasting accuracy over time. The system was then used to forecast agriculture price indices, and specifically, the levels of prices for produce and seasonal differences, demonstrating the benefits of applying this forecasting system in the agriculture industry. An energy production forecasting algorithm was presented in [20] and applied to a PV plant. The method utilized transfer function estimation based on the computation of suitable statistical indicators. The experimental results presented in that work showed the efficiency of the method in terms of forecasting quality. In [21], a microforecasting algorithm was proposed for an energy management system employed in small residential or tertiary industry areas eligible for participation in demand response programs. The algorithm provided daily energy consumption estimations to the connected energy management system. Moreover, it was possible to integrate the energy management system with other methods for the forecast of relevant parameters such as weather conditions. The results of the paper showed the practicality of the developed microforecasting module, which provided appropriate and accurate results.

In [32], a prototype of an IoT-based smart off-grid solar system was developed, with voltage and current sensors implemented to monitor the characteristics of the system. Moreover, the model used a battery for the irrigation systems using fog and sprinkler pumps. The results of the work showed that IoT significantly enhanced the functionalities of the developed prototype, offering an alternative for the use of a green energy resource. In [33], the authors proposed a smart community grid model with several consumers and producers. The presented model utilized a DT approach in the developed optimization algorithm, minimizing the operational costs of the community using demand response programs. The numerical results of this work proved that DT, RER scheduling, and demand response programs benefit both sides of the network.
The main contributions of this paper include:

- A realistic case study in which a farm is located near a river with a hydropower generator;
- The scheduling of the energy resources available in such a scenario;
- The development of a DT, within the context of agriculture, for the optimal management of energy resources in offline mode, with no external server/machine or internet access and no complex optimization computations;
- A laboratory emulation of the developed model for the scheduling/distributed control of PV and hydropower energy resources;
- A validation of the performance of DT for optimal energy scheduling in agriculture, under the practical and technical challenges that emerged during the emulation.

This paper is organized into six sections. Following the introduction, Section 2 describes the developed system model. Section 3 presents the laboratory configurations and implemented model for the agriculture emulation. Section 4 details a case study that was conducted to test and validate the performance of the developed system, and its results are shown and discussed in Section 5. Finally, Section 6 outlines the conclusions drawn from the conducted research.

2. Study Description

This section describes the proposed model of an agricultural energy management system. The main objectives of the study are summarized as follows:

1. Design and implement a suitable architecture;
2. Develop an optimization algorithm for scheduling the energy resources of the system, conduct a theoretical study, and simulate its performance;
3. Use the DT methodology to perform energy scheduling;
4. Conduct a laboratory emulation to test the developed model;
5. Validate the performance of the DT in an emulated scenario and compare the real energy scheduling results obtained in the emulation with the theoretical results obtained from the optimization algorithm.

Furthermore, the proposed energy management model was developed on a series of assumptions. These assumptions include:

- The farm is located near a river (for hydroelectric power generation and water pumping);
- The planted crops must be irrigated twice a day;
- The water used for the irrigation is supplied from a water tank and not the river.

The abovementioned study objectives are addressed separately throughout this research work. Steps 1, 2, and 3 are described in the next section.

2.1. System Architecture

This paper proposes a model of an energy management system that involves the use of RERs and a river turbine. The main purpose of the system is to supply electricity from local energy resources, minimizing the need to purchase electricity from the utility grid, leading to reduced electricity costs. Figure 1 gives an overview of the proposed system model. There are two renewable energy resources in this system to ensure agricultural sustainability: PV panels and a river turbine connected to a synchronous generator. Thus, the model is intended for scenarios where the farm is located next to a river on which a turbine is placed to rotate the shaft of the synchronous generator. Furthermore, there are three electric motors, two of them responsible for irrigation system and the other one responsible for pumping the water from the river to a water tank. Furthermore, there is a level sensor inside of the water tank indicating the amount of water stored in it. Thanks to this sensor, the water pump motor can be controlled accordingly.
To control the system, several distributed programmable logic controllers (PLCs) and energy meters are employed for the decentralized management of the system. There are four distributed PLCs in this model, one for each system player (i.e., PV, electric motors, synchronous generator, and grid connection switchboard). PLCs enable the system to take decisions locally. Moreover, thanks to the TCP/IP communication protocol, all of the players can communicate and achieve the system goals. In fact, system players must continuously send messages that describe their latest status in the network. In this way, the system’s response time to changes is more rapid than in the centralized control approach. Furthermore, distributed control gives greater flexibility and reconfigurability to the system, and it also improves the adaptability of the model.

![Figure 1. The architecture of the presented model for the scheduling of energy resources in agriculture.](image)

There are two layers in the system that are responsible for performing energy scheduling and managing electricity consumption and generation rates:

- Monitoring layer: There are four energy meters connected to distributed PLCs, one for each resource (as shown in Figure 1) to monitor the system’s real-time energy consumption and generation. Furthermore, the level sensor in the water tank indicates whether the tank is full or empty;
- Controlling layer: The system has two tasks in this layer, (1) controlling the status of the water pump motor to turn on or off, depending on whether the tank is empty or full; (2) controlling the depth of the turbine in the river to regulate the speed of the synchronous generator’s shaft.

In fact, the deeper the turbine is submerged in the river, the greater the rotation speed and the production rate of the synchronous generator. This means that the distributed PLC installed on the synchronous generator can adjust the production rate of the turbine by controlling the depth to which it is submerged in the river.

Moreover, the PLC in the grid connection switchboard is responsible for dispatching information about electricity market tariffs among the other players, so the system is aware of current market prices and can perform optimal decision-making in terms of electricity costs.

2.2. Decision-Making Approaches

As mentioned in Section 2.1, the main goal of the system is to supply electricity from local energy resources to minimize electricity costs (ECs). Therefore, to ensure optimal system operation, it is essential to implement a decision algorithm. The system’s objective function for minimizing the EC is shown in Equations (1) and (3), respectively.
shown in Equation (1). In this objective function, $P_{Buy}$ is the power that the system purchases from the utility grid, and $P_{Sell}$ is the surplus of the generation that is injected in the utility grid. These are shown in Equations (2) and (3), respectively.

In Equations (1) to (3), $P_{PV}$ is the PV production, $P_{Sync}$ is the power produced by the synchronous generator, and $P_{Motors}$ stands for the consumption of the electric motors. In addition, $T$ is the time period, and $C$ is the weight of cost for each resource, which is scaled between 0 and 1. The resources whose weight of cost is small (near to 0) are prioritized by the system because they are the low-cost resources that can be used to supply energy.

\[
\text{Minimize} \quad EC = \sum_{t=1}^{T} \left[ \left( P_{PV(t)} \times C_{PV(t)} \right) + \left( P_{Sync(t)} \times C_{Sync(t)} \right) + \left( P_{Buy(t)} \times C_{Buy(t)} \right) - \left( P_{Sell(t)} \times C_{Sell(t)} \right) \right] \tag{1}
\]

\[
P_{Buy(t)} = P_{Motors(t)} - \left( P_{PV(t)} + P_{Sync(t)} \right) \forall t \in \{1, \ldots, T\} \tag{2}
\]

\[
P_{Sell(t)} = \left( P_{PV(t)} + P_{Sync(t)} \right) - P_{Motors(t)} \forall t \in \{1, \ldots, T\} \tag{3}
\]

Equations (4) to (8) show the technical limitations of each resource in terms of its minimum and maximum capacity ($P^{\text{max}}$).

\[
0 \leq P_{PV(t)} \leq P_{PV}^{\text{max}} \forall t \in \{1, \ldots, T\} \tag{4}
\]

\[
0 \leq P_{Sync(t)} \leq P_{Sync}^{\text{max}} \forall t \in \{1, \ldots, T\} \tag{5}
\]

\[
0 \leq P_{Motors(t)} \leq P_{Motors}^{\text{max}} \forall t \in \{1, \ldots, T\} \tag{6}
\]

\[
0 \leq P_{Buy(t)} \leq P_{Buy}^{\text{max}} \forall t \in \{1, \ldots, T\} \tag{7}
\]

\[
0 \leq P_{Sell(t)} \leq P_{Sell}^{\text{max}} \forall t \in \{1, \ldots, T\} \tag{8}
\]

The present mathematical problem was solved as a linear programming optimization problem using the “OMPR” package of RStudio® tools (www.rstudio.com). In this work, a high-performance computer was employed to solve the optimization problem and provide the results. However, it is not advisable to use a computer or a server to implement this decision-making approach because farms are normally located in remote areas that lack access to computers and external servers. Therefore, a methodology should be employed that would enable the system to perform decision making locally in the distributed PLCs.

The DT approach is considered to be a suitable solution since it can easily be implemented in the PLCs or any other controllers. Therefore, a DT was developed to model the proposed objective function using the RPART package of RStudio® tools. RPART is an abbreviation for recursive partitioning, defined as a statistical approach for multivariable analysis [34]. When RPART starts building a DT, it chooses the most effective variable that divides the data into two parts. Afterwards, this method is used separately for every single subsection. This process is continued until the subsections are reduced to a minimum size or no more improvements can be made. The RPART consists of various methods (“anova”, “poisson”, “class”, and “exp”) that can be optionally selected by the user [34]. In this paper, the authors selected the “class” method to build the DT, whose operation is based on the classification of data.

It should be noted that linear programming solvers (e.g., OMPR) are not considered to be a suitable solution for this system since the system is to perform energy scheduling autonomously and in real time using real data. To do this, linear programming solvers would have to be executed every single time a parameter of the system changes. The RPART package provides a set of complementary decisions using “if–then” rules that can be implemented in any type of controller, such as a PLC. Therefore, the system can rely on those predefined rules, implemented in each distributed PLC, to perform energy scheduling autonomously. Moreover, the optimization algorithm may be executed, and
it is not necessary for the system to communicate with a server or external entities during its execution. Then, the optimization solution provided by “OMPR” is compared with the solution provided by DT to demonstrate the superiority of DT and the accuracy of its outputs.

3. Laboratory Implementation

In this section, the developed laboratory model of the agricultural energy management system is illustrated. All of the practical features and automation approaches of real devices and laboratory emulators are described. The automation mechanisms for the monitoring and control of the emulators were developed and implemented by the authors of this paper. Before implementing these automation approaches, the emulators had to be controlled manually by a user. Each emulator in this model is controlled and managed by the distributed PLC installed locally in the emulator. Figure 2 illustrates the two laboratory emulators and their automation mechanisms.

![Figure 2. Automation approaches implemented in laboratory synchronous generator and electric motor emulators for experimenting case studies.](image)

The first emulator used in this model is the synchronous generator, hereinafter syn. gen. emulator (Figure 2A). The second emulator is a group of induction motors, hereinafter motor bench (Figure 2B), to emulate the irrigation system and the water pumping motor.

The syn. gen. emulator includes a 3-kW synchronous generator coupled with a three-phase induction motor. In this model, the induction motor emulates the river turbine described in Section 2.1, which rotates the shaft of the synchronous generator. Thus, controlling the speed of the induction motor in the syn. gen. emulator is a process that corresponds to adjusting the depth to which the turbine is submerged in the river. The speed of the synchronous generator affects its energy generation level. As illustrated in Figure 2A, the PLC installed in the syn. gen. emulator controls the speed of the induction motor as well as the amount of direct current (DC) power provided to the generator. For this purpose, the syn. gen. emulator uses two independent analog output channels. Furthermore, there are three relays in the syn. gen. emulator controlled through the PLC’s three digital output channels: (1) motor relay, which is responsible for turning the induction motor on/off and the related speed controller unit; (2) DC relay, which is accountable for connecting/disconnecting the DC power supply to/from the synchronous generator; (3) grid relay, which is responsible for the islanded-mode or the grid-connected mode of the emulator (in this paper, the emulator is used only in grid-connected mode). In addition, there are two energy meters installed in the syn. gen. emulator, one in the generator side to measure the produced power, and one in the grid side to monitor the utility grid parameters. The PLC acquires the data from these two energy meters using an Ethernet interface, with the MODBUS TCP/IP protocol.
When the system intends to produce energy by means of the syn. gen. emulator, the induction motor begins to rotate the shaft of the generator at a specific speed to achieve a frequency rate (in this case, close to 50 Hz) in the generator. Thereafter, the PLC triggers the DC relay and connects the DC power supply to the rotor of the generator with a nominal value of DC voltage and current specified in the plate of the emulator. This process leads to alternating voltage (AC) in the stator side of the generator. To connect the generator’s stator to the utility grid, the frequency of the AC voltage in the stator should be exactly equal to the frequency of the grid. This frequency synchronization is performed by the PLC since it is aware of the frequency rate in both sides (generator and grid sides). For this purpose, the PLC increases or decreases the frequency of the generator by regulating the speed of the induction motor. When the frequency has been synched in both sides, the difference in AC voltage between the stator of the generator and the utility grid becomes minimal, that is, near to zero. At this point, the PLC triggers the grid relay and connects the utility grid to the stator of the generator. Henceforth, the frequency of the generator’s stator is exactly equal to the frequency of the network (in this study, 50 Hz). While the frequency of the generator is higher than the frequency of the network, the generator produces energy and injects the power to the utility grid with the frequency of the network.

The second emulator utilized in this model is motor bench, as illustrated in Figure 2B). In this emulator, there are two 1.5-kW three-phase induction motors with a constant load and one 3-kW three-phase induction motor with variable load. The motors with constant load always have a constant rate of electricity consumption since a load is already fixed and applied to the shaft of each motor. However, in a 3-kW motor there is an electric brake system applied to the shaft of the motor that enables the system to have a variable rate of loads on its shaft, so the electricity consumed by the motor would vary. The distributed PLC installed in this emulator controls the speed of each motor separately, as well as the electric brake ranging from 0% to 100%, using analog output channels.

In this paper, it is considered that the two induction motors with constant load model the irrigation system of the model, and the induction motor with variable load models the water pump motor. These assumptions have been made considering that the user is able to adjust the flow rate of the water pumped from the river to the tank, so the electricity consumed by the water pump motor varies. It should be noted that these assumptions are only for this specific agricultural energy management system, and that they could vary in different models.

Regarding the RERs used in the developed system, 7.5-kW rooftop PV arrays are employed. This PV system is not an emulator, and in fact, it is already installed in the GECAD research center building in Porto, Portugal, where the developed system model was implemented. To merge this PV system in the model, an energy meter was installed in the AC side of the inverter to measure the real-time PV production and transmit the data to a local distributed PLC assigned to the PV system. Figure 3 shows the methodology implemented for monitoring real-time PV production.

![Figure 3](image_url)
The PV system, as well as the two emulators shown in this section, are all connected together through an internal electricity network as shown in Figure 4. This internal electricity network consists of three power lines, one for each resource of the system. The grid connection switchboard shown in Figure 4 is the connection point of the power lines to the utility grid. In addition, there is a console implemented in the switchboard showing a web-based graphical interface for monitoring the system parameters, such as electricity market tariffs, the status of each emulator, and real-time consumption and generation.

Figure 4. Internal electricity network of the presented system.

In sum, this section provided the laboratory model developed for the agriculture system. Practical features of all the equipment and emulators have been illustrated, and the implemented automation mechanisms have been described.

4. Case Study

This section details the case study that was conducted to validate and survey the performance of the developed model and the energy scheduling algorithms. Figure 5 shows the electricity price data used as input in this case study. All the profiles shown in this case study are for a complete day (24 hours) with a 1-minute time interval (1440 periods).

Figure 5. Electricity market and reference prices considered in the case study.
The profile shown in Figure 5 was created on the basis of the records given in the Portuguese section of the Iberian Electricity Market (www.omie.es). In the proposed agricultural energy management system, the electricity grid is considered as an external supplier, which supports the system whenever the local energy resources cannot supply the electricity demand. In this case study, it is considered that the user can also define reference price (as shown in Figure 5), so that when the electricity market prices are cheaper than the maintenance and technical costs of the local energy resources, the system is allowed to use electricity from the grid instead of using local resources (e.g., synchronous generator).

In addition, Figure 6A shows the total consumption and generation profiles considered in the system, and Figure 6B illustrates the detailed consumption profiles of the irrigation motors as well as the water pump motor. Furthermore, the level of water stored in the tank ranged from 0 to 100%, as shown in Figure 6B. In this case study, the initial tank level was considered to be 50%. The PV production profile shown in Figure 6A is a real profile adapted from the GECAD research center database.

According to the information provided by Food and Agriculture Organization (FAO) [35], the best time to irrigate crops is early in the morning and evening when the transpiration of water on the soil surface is minimal. Therefore, in this case study it was considered that the system irrigates the crops twice a day (as Figure 6B demonstrated), at full capacity in the morning (between 07:00 and 10:00) and at half-capacity in the evening between (17:00 and 20:00). The water used for the irrigation is supplied by the water tank, and the intention of the system is to maintain the tank full. Therefore, the water pump motor is responsible for supplying the tank with water from the river at different water pumping rates defined by the user. Table 1 shows the relation between the electricity consumption of the water pump and the three levels of pumping capacity.

In this case study, the dashed line in Figure 6B indicates the electricity consumption profile of the water pump motor, which depends on the pumping capacity level chosen by the user. Furthermore, the...
energy consumption of all of the irrigation and water pump motors in the system is shown in Figure 6A, and the generation profile is the energy generated by the PV. At some periods in Figure 6A, the energy produced by the PV can supply the demand fully. At other points, surplus energy is generated, which will be injected into the grid. However, there are some periods during which the system consumes more than it generates. In such periods, the use of synchronous generation, as well as the electricity grid, should be optimized and scheduled in order to minimize the electricity costs of the system.

<table>
<thead>
<tr>
<th>Pumping Level</th>
<th>Electricity Consumption</th>
<th>Water Pumping Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 kW</td>
<td>( \frac{1}{3} \times CA )</td>
</tr>
<tr>
<td>2</td>
<td>2 kW</td>
<td>( \frac{2}{3} \times CA )</td>
</tr>
<tr>
<td>3</td>
<td>3 kW</td>
<td>( \frac{3}{3} \times CA ) (full capacity)</td>
</tr>
</tbody>
</table>

In this model, DT is responsible for the optimal scheduling of the energy resources. In fact, all of the profiles shown in Figures 5 and 6 are used as input data during the building of the DT. A large amount of data on the different conditions within the system had to be provided to the “RPART” package of RStudio® as a dataset in order to begin the building of the DT. For this purpose, 5 groups of data were considered. Table 2 shows the dataset for DT.

<table>
<thead>
<tr>
<th>Group 1 (Base)</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Consumption</td>
<td>Figure 6A</td>
<td>Base × 1.2</td>
<td>Base × 1.1</td>
<td>Base × 0.9</td>
</tr>
<tr>
<td>PV Generation</td>
<td>Figure 6A</td>
<td>Base × 1.2</td>
<td>Base × 1.1</td>
<td>Base × 0.9</td>
</tr>
<tr>
<td>Market Prices</td>
<td>Figure 5</td>
<td>Base × 1.2</td>
<td>Base × 1.1</td>
<td>Base × 0.9</td>
</tr>
<tr>
<td>Tank Level</td>
<td>Figure 6B (tank actual level profile)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using the data shown in Table 2, a dataset was created giving numerous possibilities for different conditions within the system. The dataset was used by “RPART” to build the DT. The output of the DT was a set of rules specifying the amount of energy that should be produced by the synchronous generator to supply the electricity demand.

After the DT had been configured, the data shown in Figures 5 and 6 were used to solve the optimization problem discussed previously in Section 2.2. These data were delivered as inputs to the “OMPR” package of RStudio®. In the next section, the optimization solution provided by “OMPR” is compared with the solution provided by DT to demonstrate the superiority of DT and the accuracy of its outputs.

5. Results

This section presents all of the results obtained by the DT and the optimization problem. Furthermore, the functionalities of the system are discussed. Thus, first the DT created by the “RPART” package of RStudio® will be presented, and the accuracy of the tree will be discussed. Then, the DT’s scheduling results will be compared with the optimization results of the “OMPR” package of RStudio®, so that the performance of the DT is validated. In the final stage, the DT itself was implemented in laboratory emulators and energy scheduling was performed for real resources, yielding real results.

5.1. Decision Tree Accuracy

The accuracy of the results of a DT is dependent on a series of factors, such as the nature of the input data, the vastness of the dataset, the calculation methods, etc. In fact, whenever the number of
splits in a DT increases, the DT becomes more complex and precise. Figure 7 illustrates the created DT using the dataset shown in Table 2. The consumption and the PV values represented on the DT are in kW, market prices are in EUR/kWh, and the tank level ranges between 0 and 100%.

Figure 7. Developed DT for energy scheduling of the agriculture system.

The type of DT shown in Figure 7 is a classification tree (CART—classification and regression tree) that is employed to predict a qualitative response. In the classification tree, the training dataset is broken down into smaller subsets until the DT reaches an optimal level with a low error rate. Furthermore, the classification tree provides several decision paths based on the provided training data. The system employs these paths to predict that each piece of data can be placed in the most related category. Each node of the DT shows the utilized variables and thresholds employed for classification. The terminal nodes include the predicted solution in that node. In this system, the final result of the DT (terminal nodes) is the amount of power (in kW) that should be produced by the synchronous generator. The decision rules employed in this DT are presented in Table 3.

All the variables used in the DT have an importance rate, and as shown in Table 3, consumption is the most important variable. More specifically, Figure 8 illustrates the importance of each predictor variable in the DT, scaled to sum to 100%.
The type of DT shown in Figure 7 is a classification tree (CART—classification and regression tree) that is employed to predict a qualitative response. In the classification tree, the training dataset is broken down into smaller subsets until the DT reaches an optimal level with a low error rate. Furthermore, the classification tree provides several decision paths based on the provided training data. The system employs these paths to predict that each piece of data can be placed in the most related category. Each node of the DT shows the utilized variables and thresholds employed for classification. The terminal nodes include the predicted solution in that node. In this system, the final result of the DT (terminal nodes) is the amount of power (in kW) that should be produced by the synchronous generator. The decision rules employed in this DT are presented in Table 3.

All the variables used in the DT have an importance rate, and as shown in Table 3, consumption is the most important variable. More specifically, Figure 8 illustrates the importance of each predictor variable in the DT, scaled to sum to 100%.

Table 3. Employed decision rules to produce the DT.

<table>
<thead>
<tr>
<th>Decision Rules</th>
<th>Sync. Gen. Emulator</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) IF Consumption &lt; 4.3 AND PV &gt;= 1.50</td>
<td>0</td>
</tr>
<tr>
<td>(2) IF Consumption &lt; 4.3 AND PV &lt; 1.50 AND Tank Level &lt; 50</td>
<td>0</td>
</tr>
<tr>
<td>(3) IF Consumption &lt; 1.1 AND PV &lt; 1.50 AND Tank Level &gt;= 50 AND Market Price &lt; 0.15</td>
<td>0</td>
</tr>
<tr>
<td>(4) IF Consumption is 1.2 to 4.3 AND PV &lt; 1.50 AND Tank Level &gt;= 50 AND Market Price &lt; 0.15</td>
<td>0</td>
</tr>
<tr>
<td>(5) IF Consumption &gt;= 4.3 AND Market Price &lt; 0.15</td>
<td>0</td>
</tr>
<tr>
<td>(6) IF Consumption is 1.2 to 2.1 AND PV &lt; 1.50 AND Tank Level &gt;= 79 AND Market Price &gt;= 0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>(7) IF Consumption is 1.2 to 2.1 AND PV &lt; 1.50 AND Tank Level is 72 to 79 AND Market Price &gt;= 0.15</td>
<td>0.5</td>
</tr>
<tr>
<td>(8) IF Consumption is 1.1 to 1.2 AND PV &lt; 1.50 AND Tank Level is 50 to 67</td>
<td>1</td>
</tr>
<tr>
<td>(9) IF Consumption &lt; 1.1 AND PV &lt; 1.50 AND Tank Level &gt;= 50 AND Market Price &gt;= 0.15</td>
<td>1</td>
</tr>
<tr>
<td>(10) IF Consumption is 1.1 to 1.2 AND PV &lt; 1.50 AND Tank Level &gt;= 67</td>
<td>1.25</td>
</tr>
<tr>
<td>(11) IF Consumption is 1.2 to 2.1 AND PV &lt; 1.50 AND Tank Level is 50 to 72 AND Market Price &gt;= 0.15</td>
<td>2</td>
</tr>
<tr>
<td>(12) IF Consumption is 2.2 to 4.3 AND PV is 0.14 to 1.50 AND Tank Level &gt;= 50 AND Market Price &gt;= 0.15</td>
<td>2.25</td>
</tr>
<tr>
<td>(13) IF Consumption is 2.1 to 2.2 AND PV &lt; 1.50 AND Tank Level &gt;= 50 AND Market Price &gt;= 0.15</td>
<td>2.25</td>
</tr>
<tr>
<td>(14) IF Consumption is 2.2 to 4.3 AND PV &lt; 0.14 AND Tank Level &gt;= 50 AND Market Price &gt;= 0.15</td>
<td>2.5</td>
</tr>
<tr>
<td>(15) IF Consumption &gt;= 4.3 AND Market Price &gt;= 0.15</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4 shows the pruning from the “RPART” algorithm for the developed DT. In this table, each row specifies the depth of the tree and the calculations that correspond to that level. In fact, the number of splits in each level is increased until the algorithm reaches an optimal level with a low error rate. Moreover, in Table 4, the relative error indicates the prediction error of the data that was used to make the tree, and the cross-validation error is the amount of error produced by the “RPART” built-in cross-validation. In addition, CP stands for complexity parameter, which is a value in each depth of the tree used to perform divisions in the nodes until the relative error decreases to a desired rate.
In other words, CP controls the size of the tree and chooses the optimal tree size. Moreover, CP $\alpha$ determines how the cost of a tree $R(T)$ is affected by the number of terminal nodes $|T|$, which results in a standardized cost $R_\alpha(T)$ [36]. In Equation (9), a larger amount of $\alpha$ results in small trees and potential underfitting, and small $\alpha$ results in larger trees and potential overfitting. The process of CP calculation is clearly illustrated in Figure 9.

$$R_\alpha(T) = R(T) + \alpha|T|$$ (9)

In fact, Figure 9 shows a summary of the computation process based on the relative error and CP in order to calculate the most optimal size of the tree. The dashed line in Figure 9 is a certain rate of relative error that the algorithm should reach to compute the most optimal results. More simply, the point at which the two lines in Figure 9 have crossed each other is the most optimal size of the tree, which in this case is 15 terminal nodes.

To summarize, the process of building a DT has been outlined, and detailed information has been given regarding its errors. In this section, we explored how it is possible to increase the accuracy of the results obtained by the DT. The output results of the DT (terminal nodes) will be compared with another optimization method to validate its performance.
5.2. Energy Scheduling Results

In this section, the energy scheduling results of the optimization algorithm as well as the developed DT are outlined and compared. The mathematical formulations shown in Section 2.2 were used to present the optimization problem, which was then solved by means of linear programming with the “OMPR” package of RStudio®. The energy scheduling results using the optimization algorithm are shown in Figure 10A. Furthermore, the DT and its decision rules shown in Section 5.1 were also implemented in RStudio®, and the energy scheduling results using DT were acquired, as illustrated in Figure 10B.

By comparing Figure 10A,B, it can be concluded that the developed DT results are very similar to the scheduling results of the optimization algorithm. This can prove the performance of the DT and its decision rules since the system can operate in an optimal way using the developed DT. In both scheduling results shown in Figure 10, PV production is the first source of energy for the supply of the demand, and then the system regulates the output of the synchronous generator to supply the remaining energy needs. In periods of high consumption, the energy supplied by the PV and the synchronous generator is not enough and the system purchases energy from the electricity grid. However, there are several periods in which the PV generation is higher than the consumption rate, so the system injects the surplus energy into the electricity grid. It should be noted that the DT is only applied in the system when the consumption and generation rates are higher than zero since there is no need to apply the scheduling when these two rates are equal to zero.

![Figure 10. Energy scheduling results using (A) an optimization algorithm; (B) DT.](image)

In some periods of the scheduling shown in Figure 10, the system purchases energy from the utility grid instead of using a synchronous generator. This is because in these specific periods, the electricity market prices are cheaper than the reference price defined by the system user. This reference price is calculated according to the service and maintenance costs of the synchronous generator and its river turbine. In other words, if the system recognizes that the market price in a period is cheaper than the maintenance costs of the synchronous generator, it purchases energy from the utility grid and stops the synchronous generator. The same is not true for the PV resource because PV arrays are always able to produce electricity with little or no maintenance cost. To compare the scheduling results in more detail, Figure 11 shows the energy scheduling results for the synchronous generator. In most of
the periods in Figure 11, the differences between the two profiles are not significant. Therefore, it is concluded that the system can rely on the DT to perform the optimal scheduling of the resources.

![Figure 11. Scheduled generation profile for the synchronous generator during the entire case study.](image)

**Figure 11.** Scheduled generation profile for the synchronous generator during the entire case study.

### 5.3. Actual Measurements

This section demonstrates the performance of the agriculture system in the laboratory. All the equipment presented in Section 3 was employed to implement energy scheduling, acquire real results, and test the behavior of each laboratory emulator. In this regard, a 10 minute-cycle was selected between period #300 and #600 of the case study, each period having a duration of 2 seconds in real time. In some experimental results shown in this section, the cycle lasted more than 10 minutes due to the emulator’s technical features.

Figure 12 shows the first experimental results of the syn. gen. emulator. In Figure 12, the setpoints are the scheduled values that the PLC specified to be produced by the synchronous generator to supply the consumption of the system. The real measurement profile consists of the information acquired from the emulator (absolute amount of generated energy), which is measured by the energy meter of the emulator and monitored by the local PLC.

![Figure 12. Generation profile of the syn. gen. emulator.](image)

**Figure 12.** Generation profile of the syn. gen. emulator.

The instabilities of the emulated profile demonstrated in Figure 12 are because of the technical and practical features of the system, such as electrical grid conditions and voltage variations. The PLC is responsible, in this regard, for the control of the rate of generation and for ensuring it is stable and close to the setpoint. Furthermore, the slow response time of the sync. gen. emulator creates a gap...
between the expected and real results. In fact, this is one of the most important aspects of emulation and laboratory experiments; they reveal the practical features and technical issues of each model.

Regarding the consumption emulation of the agriculture system, Figure 13 illustrates the consumption profiles emulated by the motor bench. As shown in Figure 13, the critical moments in this emulation are those where the irrigation motors start operating. In these moments, the motors lead to some instabilities in the system for a short period of time. By comparing Figure 13A,B, it can be concluded that the emulated consumption profile of the two irrigation-related motors is smoother than the water pump motor. This is due to the constant loads applied to the shafts of the irrigation motors in the emulator. In the water pump motor, the PLC should adjust the rate of load applied to the motor shaft, and this causes some variations in consumption rate.

Furthermore, Figure 14 shows the whole consumption profile produced during the emulation of the system as well as the power that was supplied by the utility grid during the emulation. All resources were operated in a grid-connected mode in this laboratory experiment, and the electricity network was considered as an external supplier that supports the system when the local energy resources cannot supply the electricity demand.

**Figure 13.** Emulated consumption profile of the system by the motor bench: (A) irrigation motors; (B) water pumping motor.

**Figure 14.** Power purchased from the utility grid in comparison to total consumption during the laboratory experiment.
As Figure 14 shows, while all the electrical motors of the system (irrigation and water pumping) operate at full capacity, the local energy resources (PV and synchronous generator) are not able to supply the demand, so the rest of consumption is supplied by the electricity network.

5.4. Cost Comparison

The final results involve the comparison of cost. For this purpose, the overall costs of the system were calculated for three different scenarios: (1) without using any RERs and decision making; (2) using RERs and the optimization algorithm for decision making; (3) using RERs and the developed DT. The market prices shown in Figure 5 were considered as the cost of purchasing energy from the utility grid. The cost of selling energy (injecting generation surplus) to the main grid is considered to be 0.0522 EUR/kWh according to the Portuguese regulations—Article 24 of [37,38]. The results of the cost calculations are shown in Table 5.

<table>
<thead>
<tr>
<th>Table 5. Cost calculation during the proposed case study.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenarios</strong></td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>(1)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(2)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(3)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

As can be seen in Table 5, the use of RERs contributes significantly to the reduction of daily electricity costs. The system profits from the surplus energy injected into the utility grid. Thanks to employing RERs on the demand side, electricity consumers take on the role of prosumers who also produce electricity. Furthermore, the use of the DT approach leads to slightly greater cost reductions than those achieved by the optimization algorithm. This is because the DT’s scheduling was more optimal than that of the optimization algorithm.

The profit obtained in scenario (2) comes from the surplus PV energy generation; since the tank was already full, the energy was injected into the grid. In scenario (3), additionally, the DT’s scheduling decisions involve some error. Some energy from the synchronous generator was sent to the grid when this was not optimal.

6. Conclusions and Future Lines of Research

The world’s population is increasing on a daily basis, as a result ensuring agricultural sustainability is more important than before. It is essential to improve energy efficiency in this sector, and we need more automation mechanisms that are compatible with the new concepts of the power system, such as smart grids and demand response programs. In addition, renewable energy resources are easy to access in rural areas and have minimal environmental consequences.

In this paper, an energy-scheduling model was proposed and a case study was conducted to test its viability. Hydropower and renewable energy resources were considered in the model to supply the electricity demand of the farm. Some automation mechanisms were proposed and described in this paper, which were been implemented in real laboratory devices and machines for modelling the consumption and generation of the proposed agriculture system. Furthermore, a decision tree was presented for optimal energy scheduling in the system. The DT was also implemented in a laboratory model to test its performance. In this way, the system was able to minimize its dependence on the utility grid since it could supply the electricity demand with the locally produced energy using a photovoltaic unit and a hydropower turbine.
The system was tested and validated in a realistic scenario, with its energy scheduling performance examined as well as the technical and practical capabilities of the developed model. Real data were used for some of the system aspects; real data on energy resources were used in the development of the decision tree. Furthermore, the precision errors of the system were calculated. In the final stage, the technical behaviors of the emulated agricultural energy scheduling model were discussed.

The results of this paper show that the operation of the decision tree approach is acceptable since it can schedule the use of energy resources in an automated offline mode with no need for an external server or machine for the complex computational algorithm. This is a very important factor in practice since most farms are located in remote areas that lack access to the internet or any high-performance computing machines. Moreover, the practical results of the paper showed a gap between the expected and the real results. This is the main advantage of laboratory experiments—they make it possible to identify practical features and reveal the technical issues of each model, which mostly remain hidden during the phases of theoretical study and simulation. Therefore, laboratory implementation and experiments are essential when validating the performance of any system under practical challenges, such as device response time and electricity network conditions, namely voltage variations and frequency instabilities.

In a future work, it is intended to develop and integrate an autonomous approach to determine optimal irrigation periods using field data, such as soil moisture level, evapotranspiration of crops, precipitation, etc. In this way, the system will not only be able to optimize the use of water in the irrigation process, but it will also be able to optimize and reduce the overall operational costs associated with energy consumption. Forecasting will also be utilized to predict critical parameters (i.e., energy consumption and generation, solar radiation, precipitation, etc.), improving the efficiency of the model. Further lines of research can go beyond model emulations, focusing on the implementation of the model in real scenarios and testing other optimization techniques, such as particle swarm optimization.


Funding: The present research work was conducted and funded within the scope of the following project: Eco Rural IoT project, funded by TETRAMAX-VALUECHAIN-TTX-1, and UID/EEA/00760/2019, funded by FEDER Funds through the COMPETE program and by National Funds through FCT.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).