Decision support system for photovoltaic fault detection avoiding meteorological conditions

Roberto G. Aragón[a,], M. Eugenia Cornejo[a], Jesús Medina[a], Juan Moreno-García[b], Eloísa Ramírez-Poussa[a]

[a]*Department of Mathematics, University of Cádiz, Cádiz, Spain*
[b]*Escuela de Ingeniería Industrial, Universidad de Castilla La Mancha, Toledo, Spain*

Abstract

A fundamental issue about installation of photovoltaic solar power stations is the optimization of the energy generation and the fault detection, for which different techniques and methodologies have already been developed considering meteorological conditions. This fact implies the use of unstable and difficult predictable variables which may give rise to a possible problem for the plausibility of the proposed techniques and methodologies in particular conditions. In this line, our goal is to provide a decision support system for photovoltaic fault detection avoiding meteorological conditions.

This paper has developed a mathematical mechanism based on fuzzy sets in order to optimize the energy production in the photovoltaic facilities, detecting anomalous behaviors in the energy generated by the facilities over time. Specifically, the incorrect and correct behaviors of the photovoltaic facilities have been modeled through the use of different membership mappings. From these mappings, a decision support system based on OWA operators informs of the performances of the facilities per day, by using natural language. Moreover, a state machine is also designed to determine the stage of each facility based on the stages and the performances from previous days. The main advantage of the designed system is that it solves the problem of "constant loss of energy production", without the consideration of meteorological conditions and being able to be more profitable. Moreover, the system is also scalable and portable, and complements previous works in energy production optimization. Finally, the proposed mechanism has been tested with real data, provided by Grupo Energético de Puerto Real S.A. which is an enterprise in charge of the management of six photovoltaic facilities in Puerto Real, Cádiz, Spain, and good results have been obtained for faulting detection.

Keywords: Photovoltaic Energy Generation, faults detection, OWA operator,

Email addresses: `roberto.aragon@uca.es` (Roberto G. Aragón), `mariaeugenia.cornejo@uca.es` (M. Eugenia Cornejo), `jesus.medina@uca.es` (Jesús Medina), `Juan.Moreno@uclm.es` (Juan Moreno-García), `eloisa.ramirez@uca.es` (Eloísa Ramírez-Poussa)

---

1. Introduction

One important part of the Europe 2020 strategy is focused on reaching 20% of gross final energy consumption from renewable sources by 2020, and increasing this trend to reach at least 32% by 2030. This goal is also pursued by other countries and continents, which also try to fostering the implementation of renewable energy. As a consequence, many research institutions are developing new mechanisms in different research lines, such as in new construction protocols, management, legal aspects, cybersecurity, and the application of artificial intelligence tools [6, 19, 26, 33, 44, 45]. This last line mainly tries to enhance the behavior of the facilities and optimizing the energy generation [38, 39, 51, 53, 55].

In the framework of renewable energy, we have many sources, such as biofuel, geothermal, hydroelectricity, tidal, solar, wave, and wind, being solar energy one of the most important. Many companies around the world are focused on the design of photovoltaic (PV) cells, required accessories, and in the installation of photovoltaic solar power stations. Hence, the innovation of more efficient generation of photovoltaic power is a fundamental goal [4, 48]. In this aspect, fault detection is a critical issue for optimizing the energy generation and encouraging the use of renewable energy. In this research line, many groups are working and different mechanisms have been proposed, as the authors in [28] also highlighted. They for example apply (artificial) neural networks, regression, multi-gene genetic programming, multi-objective and multi-attribute decision-making, and principal component analysis, among others [7, 12, 13, 18, 21, 24, 37, 47, 49, 54]. One important fact, also noted in [7], is that fault diagnosis models are not efficient to be trained and updated, because usually the PV arrays' operating point only contains little information. In the aforementioned paper, the emerging kernel based extreme learning machine (KELM), the Nelder-Mead Simplex (NMS) optimization method and a Simulink based PV modeling approach are combined to present a fault diagnosis model. However, the final result is a complex system which considers variables directly dependent on the operating condition, i.e., the input irradiance and operating temperature. Moreover, they required restrictive boundary conditions, such as all the solar cells in the PV module be homogeneous.

The authors in [13] have proposed a smart algorithm based on support vector machine (SVM) technique for diagnosis short circuit faults in PV systems, which requieres irradiance level and PV module's temperature. [12] also analyzes the theoretical voltage ratio (VR) and power ratio (PR) of the the given grid-connected photovoltaic (GCPV) plant, but irradiance level and PV module's temperature are also required.

In [18], the principal component analysis (PCA) technique is used for extracting and selecting the most relevant multivariate features, and the supervised machine learning (SML) classifiers are applied to faults diagnosis. It also considers outdoor conditions, such as temperature, irradiation and humidity. In

addition, it is only a theoretical method designed on an emulator of a GCPV system.

The method proposed in [21] is based on artificial neural networks (ANN) and considers a radial basis function (RBF), requiring only two parameters as the input to the ANN (solar irradiance and output power). The authors highlight that the fault detection accuracy strongly depends on the varying nature of each PV plant.

The same input variables are also taken into account in [47]. The authors introduce a hybrid methodology which combines the use of a fuzzy system and the use of a machine learning system containing five different trained machine learning models, such as the regression tree, artificial neural networks, multigene genetic programming, Gaussian process, and support vector machines for regression. The Takagi-Sugeno type fuzzy logic inference system is used to select one of the five machine learning techniques that would provide the most accurate prediction, by analyzing the solar irradiation input that is given.

In [24], the mathematical formulation of the different processed carry out in the energy generation is considered, such as the current and voltage characteristics of the solar cell, which also require of the irradiance and temperature of PV module, among others.

[49] highlights the required cost and complex hardware required by the majority of the current fault diagnostic techniques. The authors present in the aforementioned paper a fault diagnosis algorithm based on an artificial intelligent nonlinear autoregressive exogenous (NARX) neural network and Sugeno fuzzy inference. The fuzzy inference requires the actual sensed PV system output power, the predicted PV system output power, and the sensed surrounding conditions. Hence, this method also needs meteorological values and highly monitored plants.

Probability theory, discrete wavelet transform (DWT) and an optimal random forest (ORF) based classifier are combined in [37] to design a procedure to fault detection/classification and zone identification, which is also meteorological data-dependent.

The main problem of the previous mentioned methodologies is that they consider meteorological conditions. Notice that, for example, PV system energy production is non-linear because it is influenced by the random nature of weather conditions. As a consequence, different difficult predictable and unstable variables are taken into account in the different developed methodologies, which shows a possible problem for the plausibility of the introduced methodologies in particular conditions. For example, the strong winds in Cádiz, Spain, remarkably decrease the accuracy meteorological predictions and the meteorological stations must be protected from them, which also decreases their accuracy. Other authors have considered this challenge previously, however, the previous restrictions are very strong to be realistic and portable. For example, the methodology proposed in [54] does not depend on the meteorological conditions, but it needs strong requirements, such as every string must be monitored and the method must be applied to identical strings, which is complicated in usual facilities. Maybe, this method can work in some new installations, but over

3

time different components (investors, photovoltaic cells, etc.) must be replaced by new ones with different features and performance.

On the other hand, the daily maintenance of (medium and small) photovoltaic solar power stations can be very expensive. Therefore, the design of a decision support systems for unsupervised stations is another required challenge. The creation of an automatic intelligence system becomes even more useful when the system provides the outputs in natural language easily understandable and executable by the station manager.

This paper will introduce a mathematical strategy based on fuzzy sets [61] with the purpose of improving the processes involved in the management of renewable energies and energy efficiency, avoiding the consideration of surrounding conditions, unlike other methodologies [12, 13, 47, 49]. We will show that the introduced strategy has good results for faulting detection. Specifically, the company Grupo Energético de Puerto Real S.A. (GEN), located in Puerto Real, Cádiz (Spain), is commended with the management and integral maintenance of the public lighting network of the city. Its main purpose is guaranteeing a service to the citizen of the highest quality, minimizing operating costs and impacts to the environment, through energy efficiency. Hence, one of the activities of GEN is the generation of energy through six heterogeneous photovoltaic facilities. Take into account that these plants have different number and brands of photovoltaic cells, investors, etc. Indeed, they started in different periods.

Therefore, one of the most important and critical aspects related to manage and control this process is the associated degree of imprecision, incompleteness, uncertainty and unpredictability, which highlights the need of using mathematical fuzzy techniques.

Specifically, a decision support system has been presented in this paper to detect anomalous behaviors in the energy generated by the facilities over time. This system has been created by using a mathematical mechanism to compare the energy production among sections (substations) of a big facility or different close facilities, which can be scalable and portable to small, medium and big PV plants. It is clear that a big plant needs to be divided into small subplants for a better monitoring, and also a small plant could also be split. For example, if the plant is installed in a common area of three different neighborhood communities (with different number of people) for self-sufficiency, the installation must be divided into three non-proportional parts, although the PV plant could be relatively small. The main contributions presented in this paper are listed below:

- From the available information about the days in which the facilities have had correct or incorrect performances, we have designed different functions to describe the behaviors of the photovoltaic facilities.

- Taking into account the previously mentioned functions, we have created a decision support system to inform of the performances of the facilities per day, by using natural language.

- We have designed a state machine that allows us to determine the stage

4

of each facility, according to the stages and the performances obtained in the previous days for each facility.

- The designed system has been tested on real photovoltaic stations during 10 months obtaining good results, which have been analyzed in the last section of the paper.

This system is being very useful and it is in production in GEN. Before this system, this company detected the days when any of the stations had an abnormal behavior by making a visual comparison of the daily graphs of the generation of each facility. This caused a waste of time for the person in charge of this task. Since this person was in charge of other functions in the company, many times the data series were not reviewed for several days. This fact caused an accumulation of the error over time. As a result, economic losses occurred since the facilities generate less energy than it is expected. For that reason, the company was interested in the introduced system to receive daily reports related to the status of the facilities, in a completely automatic way and using natural language.

The paper is organized as follows. Section 2 briefly reviews the dataset considered in this study and recalls the definition and some properties of ordered weighted averaging aggregation operators. Section 3 explains the proposed mathematical reasoning to compare the energy production among the photovoltaic facilities, which gives rise to a decision support system to detect anomalous behaviors in the energy generated by the facilities over time. Section 4 is devoted to illustrate and analyze the results obtained by the automatic system described in Section 3. Finally, Section 5 finishes with some conclusions and prospects for future work.

2. Preliminaries

GEN supplies a dataset of a multivalued time series that has the information of the energy generated by each station along with the date and time. Each station has different number of photovoltaic cells and investors, which implies that each station has different peak power[1] and rated power[2]. For example, the peak power ranges from 47,61Kw to 91.125Kw. These data are preprocessed, sorted by date, time and station, in order to be incorporated into the database on which we will carry out the study.

From the collected daily information, a mathematical mechanism to compare the energy production is proposed. This mechanism consists of several phases which will be detailed in the following section. One of these phases makes use of a special kind of mathematical aggregation operator called ordered weighted

---

[1]It is the maximum value of energy that the photovoltaic cells can generate per facility, due to its technical specifications

[2]It is the maximum value of energy that the investors of each facility convert from direct current (DC) to alternating current (AC), due to its technical specifications

averaging (OWA), which were introduced by Yager in the eighties [57] as a solution to the concerned problem of aggregating multi-criteria to provide an overall decision making. On the other hand, in order to analyze the outputs obtained by the decision system we will make use of confusion matrices [5]. Confusion matrices let us measure the accuracy and error obtained in the predictions provided by the decision system in the test done. In our case, they will allow to evaluate the classification of states of each facility. For these reasons, in the following, we will briefly recall some preliminaries about these tools.

2.1. OWA operators

OWA operators have been studied from a theoretical [3, 11, 36, 42, 46, 58, 59, 25, 23, 43] and applied point of view [27, 30]. For example, they have been defined on complete lattices [15, 40, 50], in order to allow the consideration of incomparable elements to be aggregated [11, 36]. OWA operators have been applied to many fields, such as in economy, linguistic, clustering, energy, urban wastewater, etc. For example, a multi-attribute decision-making method was introduced in [17] to deal with linguistic membership and nonmembership degrees. They have also been applied to linguistic decision judgments [8], noise removal in computer vision [22], in fuzzy clustering to handle noise and outliers [52], in order to neutralize the negative effects of possible outlier fuzzy data in the clustering process [14], to provide a decomposition for all the rank-dependent poverty measures in terms of incidence, intensity and inequality [2], to allow experts to express multiple self-confidence levels when providing their preferences [35], etc.

Specifically, an *n-ary ordered weighted averaging* (OWA) operator is a function $F_W \colon [0,1]^n \to [0,1]$ associated with a list of weights $W = \{w_1, \ldots, w_n\}$, such that $w_i \in [0,1]$ and $\sum_{i=1}^n w_i = 1$, and defined as

$$F_W(a_1, \ldots, a_n) = \sum_{i=1}^n w_i a_{\rho(i)}$$

where $\rho \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ is a permutation on the index set satisfying that $a_{\rho(n)} \leq \cdots \leq a_{\rho(1)}$. Hence, the greatest value of the tuple $(a_1, \ldots, a_n)$ is multiplied by the first weight w_1, the second to the second weight and so on. Since two values a_i and a_j, with $i \neq j$, can be equals: $a_i = a_j$, at least another different permutation σ can exist satisfying the ordering among the values: $a_{\sigma(n)} \leq \cdots \leq a_{\sigma(1)}$. However, the function F_W is well defined because, if this happens, the equality $\sum_{i=1}^n w_i a_{\rho(i)} = \sum_{i=1}^n w_i a_{\sigma(i)}$ holds. Since $n = 1$ provides a trivial case, usually $2 \leq n$, as we will consider in the proposed mechanism.

The mathematical formulation of OWA operators allows the user to consider multicriteria, from the least case given by the conjunction "and" and represented by the minimum operator, and the greatest one given by the disjunction "or" and represented by the maximum. Therefore,

$$\min\{a_1, \ldots, a_n\} \leq F_W(a_1, \ldots, a_n) \leq \max\{a_1, \ldots, a_n\}$$

6

for all $(a_1, \ldots, a_n) \in [0,1]^n$. Moreover, two other very interesting properties of the OWA operators are:

- Idempotency: $F_W(a, \ldots, a) = a$, for all $a \in [0,1]$.

- Symmetry: $F_W(a_1, \ldots, a_n) = F_W(a_{\sigma(1)}, \ldots, a_{\sigma(n)})$, for all permutation $\sigma \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ and $(a_1, \ldots, a_n) \in [0,1]^n$.

2.2. Confusion matrix

The confusion matrix is a well-known tool for assessing how good a classification model is [4]. In particular, it serves to show when one class is confused with another one. In the confusion matrix, the rows correspond to predicted conditions and the columns corresponds to actual conditions. Therefore, this kind of matrices are suitable to detect false positives (FP), false negatives (FN), true positives (TP), true negatives (TN), as well as the errors obtained in the predictions. Each of these cases indicates the following: TP for correctly automatic detection alert values, FP for incorrect automatic detection alert values, TN for correctly automatic detection no-alert values, FN for incorrect automatic detection no-alert values. In addition, the errors are computed as Table 1 shows.

	No Alert	Alert	Model error
No Alert	TN	FN	$\frac{FN}{FN+TN}$
Alert	FP	TP	$\frac{FP}{FP+TP}$
Error of use	$\frac{FP}{TN+FP}$	$\frac{FN}{FN+TP}$	$\frac{FN+FP}{TN+FN+FP+TP}$

Table 1: Confusion matrix model.

3. Decision support system to detect and alert of abnormal performances

In this section, we will describe a decision support system to detect anomalous behaviors in the energy generated by the facilities. The main idea in which the proposed automatic detection system is based on is the following: since all the facilities are located in the same city and, as a consequence, they are under the same climatic conditions, it is expected that, when the facilities work in a suitable way, the differences of the energy generation between one facility and another one should be around a constant value. That is, the difference between the production of two facilities must always oscillate in a certain range of values, as long as the facilities are working properly. Consequently, when this difference is out of the expected range of values, the automatic decision system will indicate an anomalous behavior in the production of one of the photovoltaic facilities considered in the comparison. Moreover, this methodology can be exportable and scalable to photovoltaic energy stations with investors, splitting

the collection of energy in different substations and comparing them using the mechanism introduced in this paper.

The mentioned decision support system is supported by a theoretical reasoning which allows us to compare the energy produced by the different photovoltaic facilities in order to detect possible reductions in the performance of the facilities. To explain this theoretical reasoning, we will establish some notation conventions. The company has six different photovoltaic facilities, we will use the notation $\mathcal{I}_i$ in order to refer to the photovoltaic facility i, where $i \in \{1, \ldots, 6\}$. We will use the notation GE_{ij} to refer to the generated energy by the photovoltaic facility $\mathcal{I}_i$ in the hour j, with $j \in \{0, \ldots, 23\}$, these values are measured and provided by the enterprise. Moreover, for testing, we have a learning dataset on days in which the facilities have had a correct performance and days in which they have had an incorrect performance. We will take advantage of the information available regarding the correct and incorrect days of each facility to create the decision system. Specifically, a learning process that will allow us to classify the daily behavior of each facility, based on the information contained in that learning dataset, will be established (the learning procedure will be detailed in Section 3.1). As a result, we will able to detect those days in which a particular facility has an abnormal performance (the procedure of detection will be explained in Section 3.2).

3.1. Learning to classify the performance of the facilities

This section will show the procedure to classify the performance of each facility. In particular, taking into account the GEN dataset of 2019, different parameters will be studied in order to be able to correctly classify the daily performance of the facilities.

Specifically, the decision support system will calculate the difference of the energy production for each pair of photovoltaic facilities basing on the previously mentioned idea. These differences will be evaluated in membership functions that will catalog the energy production of the facilities as suitable or abnormal. This procedure is explained in a detailed way in Algorithm 1 that uses as input the learning dataset. The steps carried out by the algorithm are detailed next.

Algorithm 1 Learning daily performance parameters

1: Calculate the total performance of energy production by each photovoltaic facility.
2: Compute the difference of the total performances of energy production (Daily Performance Matrix DPM).
3: **for** each facility $\mathcal{I}_i$ with $i \in \{1, \ldots, 6\}$ **do**
4:     Divide the daily information in correct days D_C^i and incorrect days D_I^i.
5: **end for**
6: Compute the parameters that define the membership functions.

Line 1: First of all, from the information of the energy production per photovoltaic facility contained in our database, we calculate the total per-

formance of daily energy production. The following formula computes the normalization of the total performance of energy production by the photovoltaic facility $\mathcal{I}_i$, with $i \in \{1, \ldots, 6\}$, at day d:

$$\rho_i(d) = \frac{\sum_{j=0}^{23} GE_{ij}(d)}{PP_i(d)} * 100$$

Notice that, the normalization arises from the ratio of the generated energy $(GE_{ij}(d))$ and peak power $(PP_i(d))$ per day d. It is also important to mention that, in the previous formula, we consider the peak power, instead of the rated power, because the peak power is directly related to the photovoltaic cells that generate the energy. Moreover, the generated energy is always lesser than the peak power.

Line 2: Once we have the total performances of energy productions of the photovoltaic facilities, we can compute the difference between the total performances of energy production provided by two photovoltaic facilities, $\mathcal{I}_i$ and $\mathcal{I}_k$ with $i, k \in \{1, \ldots, 6\}$, at day d, by means of the following formula:

$$\delta_{ik}(d) = \frac{\rho_i(d) - \rho_k(d)}{\max\{\rho_i(d), \rho_k(d)\}} * 100$$

It is important to mention that in the previous formula the difference $\rho_i(d) - \rho_k(d)$ is divided by $\max\{\rho_i(d), \rho_k(d)\}$ in order to obtain a relative difference, which provides an easier interpretation of the result as well as an easier way to establish comparisons between two facilities $\mathcal{I}_i$ and $\mathcal{I}_k$ over time (for a fixed period of days). Moreover, we have that $\delta_{ik} = -\delta_{ki}$. Notice that the sign is important in order to know what station usually generates more energy than the others. All this information is stored in a daily matrix called "daily performance matrix".

Lines 3-5: Furthermore, for each facility $\mathcal{I}_i$ with $i \in \{1, \ldots, 6\}$, we will separate the information contained in the database in two parts attending to the kind of day, correct and incorrect denoted as D_C^i and D_I^i respectively. We can have days with an unclear classification, which will be removed for the test. Anyway, we already have a lot of noise in the data, due to the non-daily manual classification of correct and incorrect days. This is an important problem to take into account and that the use of fuzzy sets will remarkably reduce its impact in the final result.

Line 6: The main idea in which this step is based on is that, from the correct days, we can know what can be considered an acceptable difference and, from the incorrect days, we can detect anomalous differences. In order to simplify the explanation of this step, we will focus the attention on the analysis of a specific photovoltaic facility $\mathcal{I}_i$. For each $\mathcal{I}_k$, with $k \neq i$, we select those days d in which both facilities have had a correct performance and we compute the relative differences $\delta_{ik}(d)$. The least value reached, denoted as b_{ik}, will be the least value indicating a suitable performance of

9

the facility $\mathcal{I}_i$ with respect to $\mathcal{I}_k$, with $k \neq i$. When the number of days in which both facilities have had a correct performance is equal to zero, that is $D_C^i \cap D_C^k = \varnothing$, then the value b_{ik} will be identified from the value a_{ik}, whose computation is explained in the following paragraph.

Considering the information of the days in which the fixed photovoltaic facility $\mathcal{I}_i$ has an incorrect behavior and the rest of facilities $\mathcal{I}_k$, with $k \neq i$, have a correct behavior, we compute the daily relative differences of performance. Under these conditions, the greatest value of $\delta_{ik}(d)$, for each $k \neq i$, denoted as a_{ik}, will determine an interval $[a_{ik}, b_{ik}]$ indicating an anomalous performance of the facility $\mathcal{I}_i$ with respect to $\mathcal{I}_k$ with $k \neq i$. Obviously, any relative difference lower than a_{ik} will indicate a bad performance of the facility $\mathcal{I}_i$ with respect to $\mathcal{I}_k$. Formally, the values a_{ik} and b_{ik}, for each $k \neq i$, will be obtained by means of the following expressions:

$$
\begin{aligned}
a_{ik} &= \max\{\delta_{ik}(d) \mid d \in D_I^i \cap D_C^k\} \\
b_{ik} &= \min\{\delta_{ik}(d) \mid d \in D_C^i \cap D_C^k\}
\end{aligned}
$$

It must be verified that a_{ik} is lesser than b_{ik}. Otherwise, an inconsistency in the data will be detected, which is produced by an inaccurate classification of correct and incorrect days of each facility. In such a case, the roles of the values a_{ik} and b_{ik} will be exchanged, that is:

$$
\begin{aligned}
a_{ik} &= \min\{\delta_{ik}(d) \mid d \in D_C^i \cap D_C^k\} \\
b_{ik} &= \max\{\delta_{ik}(d) \mid d \in D_I^i \cap D_C^k\}
\end{aligned}
$$

For instance, if we focus on the facility $\mathcal{I}_1$ and choose the facility $\mathcal{I}_2$ to compare with it. In the computation of the interval $[a_{12}, b_{12}]$, we obtain the values $a_{12} = -12.550$ and $b_{12} = -21.662$. As we can observe $a_{12} > b_{12}$, hence we have to exchanged these values, obtaining the final interval: $[-21.662, -12.550]$. Table 2 shows a comparative matrix between facilities indicating with 1 that the role of values a_{ik} and b_{ik} have been exchanged and 0, otherwise.

	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{I}_4$	$\mathcal{I}_5$	$\mathcal{I}_6$
$\mathcal{I}_1$	$-$	1	1	1	1	1
$\mathcal{I}_2$	0	$-$	0	0	0	0
$\mathcal{I}_3$	0	1	$-$	1	1	1
$\mathcal{I}_4$	0	0	1	$-$	0	0
$\mathcal{I}_5$	0	1	0	1	$-$	1
$\mathcal{I}_6$	1	1	1	1	1	$-$

Table 2: Table of comparative for each pair of photovoltaic facilities that indicates if an inconsistency in the data has been detected.

We will build a trapezoidal membership function, from the values a_{ik} and b_{ik} for each $k \neq i$, to evaluate the daily relative differences of performance of the

photovoltaic facilities $\mathcal{I}_i$ and $\mathcal{I}_k$. These membership functions $\mu_{ik}\colon \mathbb{R} \to [0,1]$ are defined as Figure 1 shows and they will be used to indicate the degree of suitable performance of the facility $\mathcal{I}_i$ with respect to $\mathcal{I}_k$. Hence, each facility will be associated with five different membership functions.
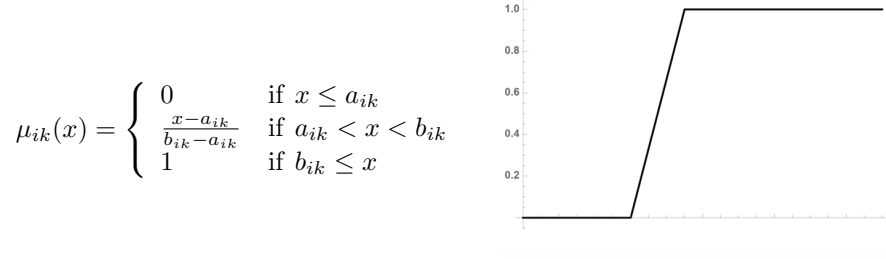
$$\mu_{ik}(x) = \begin{cases} 0 & \text{if } x \leq a_{ik} \\ \frac{x - a_{ik}}{b_{ik} - a_{ik}} & \text{if } a_{ik} < x < b_{ik} \\ 1 & \text{if } b_{ik} \leq x \end{cases}$$

Figure 1: Membership function indicating the degree of suitable performance of the facility $\mathcal{I}_i$ with respect to $\mathcal{I}_k$.

Note that, if $D_I^i \cap D_C^k = \varnothing$, then a_{ik} will be zero and, as a consequence, the system can provide bad classifications. This case arises for example when no information related to incorrect days for station i is available. Therefore, this case must be studied independently considering two cases. (1) If we have information about incorrect days of the facility $\mathcal{I}_k$, specifically $D_I^k \cap D_C^i \neq \varnothing$ holds, then we can determine the value a_{ik} by symmetry, from the values a_{ki}, b_{ki} and b_{ik} by using the formula $a_{ik} = b_{ik} - (b_{ki} - a_{ki})$. With this formula, we make that the distance between the values a_{ik} and b_{ik} coincides with the distance between a_{ki} and b_{ki}, which has sense because the values a_{ki} and b_{ki} are the resulting values obtained from the computation of the opposite differences between the facilities $\mathcal{I}_i$ and $\mathcal{I}_k$. (2) If we do not have information about the incorrect days of $\mathcal{I}_k$ for the correct days of $\mathcal{I}_i$ ($D_I^k \cap D_C^i = \varnothing$), then we will consider that $a_{ik} = b_{ik}$. This fact will cause that any relative difference lesser than the ones registered for the facilities $\mathcal{I}_i$ and $\mathcal{I}_k$ will produce an alert. In this case, the membership function will be defined as in Figure 2. Fortunately, in the considered dataset this second case does not needed to be considered. Table 3 shows with 1 in which combinations a_{ik} has been computed by symmetry (case (1) detailed above).

Repeating this process, we will obtain 5 membership functions to determine 5 degrees of proper performance of the facility $\mathcal{I}_i$ with respect to the rest of facilities.

The methodology presented in this section, including the membership functions that describe the behaviors of the facilities together with the rest of introduced formulas, have been implemented in the automatic system.

3.2. Detecting the abnormal performance of the facilities

This section will detail how the decision support system can detect anomalous performance of the facilities for a particular day d^*, and how the system classifies the performance over time by using a state machine as well.

11

	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{I}_4$	$\mathcal{I}_5$	$\mathcal{I}_6$
$\mathcal{I}_1$	$-$	0	0	0	0	0
$\mathcal{I}_2$	1	$-$	1	1	1	1
$\mathcal{I}_3$	1	0	$-$	0	0	0
$\mathcal{I}_4$	1	0	0	$-$	0	0
$\mathcal{I}_5$	1	0	0	0	$-$	0
$\mathcal{I}_6$	0	0	0	0	0	$-$

Table 3: Table of comparative for each pair of photovoltaic facilities indicating if a_{ik} has been computed by symmetry.

$$\mu_{ik}(x) = \begin{cases} 0 & \text{if } x < b_{ik} \\ 1 & \text{if } x \geq b_{ik} \end{cases}$$

Figure 2: Membership function indicating the degree of suitable performance of the facility $\mathcal{I}_i$ with respect to $\mathcal{I}_k$, when we do not have information about the incorrect days neither $\mathcal{I}_i$ nor $\mathcal{I}_k$.

Algorithm 2 describes the method used to detect the abnormal performance of the facilities. The steps to detect anomalous behaviors for a particular day d^* are the following:

Line 1: We compute the total performance of the relative energy production $\rho_i(d^*)$, for all the photovoltaic facilities $\mathcal{I}_i$, with $i \in \{1, \ldots, 6\}$.

Line 2: We obtain all possible relative differences $\delta_{ik}(d^*) = x_{ik}$, for each pair of facilities $\mathcal{I}_i$ and $\mathcal{I}_k$. These results will be stored in a matrix as the one displayed in Table 4. Notice that the difference of each facility with itself is always equal to zero, as we can see in the main diagonal of the matrix.

Algorithm 2 Detecting abnormal performance

1: Calculate the total performance of energy production by each photovoltaic facility.
2: Compute the difference of the total performances of energy production.
3: Compute the membership function for each difference from the matrix of evaluation of the relative differences.
4: Apply OWA operator to each row.
5: Obtain a linguistic description about the behavior of each facility.

	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{I}_4$	$\mathcal{I}_5$	$\mathcal{I}_6$
$\mathcal{I}_1$	0	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}
$\mathcal{I}_2$	x_{21}	0	x_{23}	x_{24}	x_{25}	x_{26}
$\mathcal{I}_3$	x_{31}	x_{32}	0	x_{34}	x_{35}	x_{36}
$\mathcal{I}_4$	x_{41}	x_{42}	x_{43}	0	x_{45}	x_{46}
$\mathcal{I}_5$	x_{51}	x_{52}	x_{53}	x_{54}	0	x_{56}
$\mathcal{I}_6$	x_{61}	x_{62}	x_{63}	x_{64}	x_{65}	0

Table 4: Matrix of relative differences of energy production for each pair of photovoltaic facilities in a specific day d^*.

Line 3: We will evaluate the values collected in Table 4 in their corresponding membership functions, that is, μ_{ik} will be applied to the value x_{ik}, for all $i, k \in \{1, \ldots, 6\}$ with $i \neq k$ (note that the main diagonal has no value due to it does not make sense to compare the performance of the installations with themselves). Once again we obtain a matrix collecting the resulting evaluations, as it is displayed in Table 5. We recall that the meaning of these values indicate the degree in which the facilities are working properly.

	$\mathcal{I}_1$	$\mathcal{I}_2$	$\mathcal{I}_3$	$\mathcal{I}_4$	$\mathcal{I}_5$	$\mathcal{I}_6$
$\mathcal{I}_1$	$-$	$\mu_{12}(x_{12})$	$\mu_{13}(x_{13})$	$\mu_{14}(x_{14})$	$\mu_{15}(x_{15})$	$\mu_{16}(x_{16})$
$\mathcal{I}_2$	$\mu_{21}(x_{21})$	$-$	$\mu_{23}(x_{23})$	$\mu_{24}(x_{24})$	$\mu_{25}(x_{25})$	$\mu_{26}(x_{26})$
$\mathcal{I}_3$	$\mu_{31}(x_{31})$	$\mu_{32}(x_{32})$	$-$	$\mu_{34}(x_{34})$	$\mu_{35}(x_{35})$	$\mu_{36}(x_{36})$
$\mathcal{I}_4$	$\mu_{41}(x_{41})$	$\mu_{42}(x_{42})$	$\mu_{43}(x_{43})$	$-$	$\mu_{45}(x_{45})$	$\mu_{46}(x_{46})$
$\mathcal{I}_5$	$\mu_{51}(x_{51})$	$\mu_{52}(x_{52})$	$\mu_{53}(x_{53})$	$\mu_{54}(x_{54})$	$-$	$\mu_{56}(x_{56})$
$\mathcal{I}_6$	$\mu_{61}(x_{61})$	$\mu_{62}(x_{62})$	$\mu_{63}(x_{63})$	$\mu_{64}(x_{64})$	$\mu_{65}(x_{65})$	$-$

Table 5: Matrix of evaluation of the relative differences in the corresponding membership function.

Line 4: In this step, we will interpret the results obtained in the matrix of Step 3. To interpret the results obtained for each facility, we will consider the rows of the matrix in Step 3 independently. For instance, if we are interested in analyzing the behavior of the facility $\mathcal{I}_i$, we will consider the i-th row of the matrix. Notice that, the elements in the main diagonal will not be considered to carry out these computations and therefore, each vector is composed of 5 elements.

We will apply an OWA operator to the i-th row to obtain a single value that represents the global degree of proper performance for each facility. Therefore, the values of this i-th row will be sorted from largest to smallest in a vector and they will be weighted according to the values of the vector $W = (0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0)$, in order to assign the value 0 to the greatest

and the least value of the ordered vector, with the goal of avoiding to consider extreme values since they are the most sensitive of being anomalous values. The obtained value y_i will represent the global degree of suitable performance of the facility $\mathcal{I}_i$. We have also tested the weighted vector $W = (0, \frac{1}{2}, \frac{1}{2}, 0, 0)$, but it offers worse results, since less information is taken into account. In the future, more cases of OWA operators will be studied.

Line 5: Finally, the value $y_i \in [0, 1]$ will be empirically identified with a linguistic label following the criteria presented below:

- If $y_i = 1$, then $\mathcal{I}_i$ has a *suitable performance (S)*.
- If $y_i \in [0.75, 1)$, then $\mathcal{I}_i$ has a *lightly anomalous performance (LA)*.
- If $y_i \in [0.45, 0.75)$ then $\mathcal{I}_i$ has an *anomalous performance (A)*.
- If $y_i \in (0, 0.45)$ then $\mathcal{I}_i$ has a *very anomalous performance (VA)*.
- If $y_i = 0$ then $\mathcal{I}_i$ has a *bad performance (B)*.

We will apply this procedure to all rows of the matrix of Step 3, obtaining a linguistic description about the behavior of each facility.

In addition, the mechanism previously presented can be applied to a sequence of consecutive days, which provides a more detailed knowledge about the behavior of the facilities and consequently, a more robust support to make decisions. For example, with the sequential analysis, we can also detect if there is a gradual decrease in the performance of any facility. To achieve this goal, a state machine [20] has been used. State machines allows us to describe the behavior of systems that are composed of a finite number of states. It also allows incorporating events that can occur in the states and that cause the evolution of the machine from a state to another state. For these reasons, they are suitable for simulating the behavior of each facility. A state machine is represented as a directed graph that is formally represented as $M = \{S, \Sigma, T, s_0\}$, where:

- S is a set of finite states, each of which is represented by a circle.

- Σ is a set of input symbols.

- T is the set of transitions. Each transition is represented by an arrow labeled with the input symbols that causes a state change and that begins in one state and ends in another one, where both states can coincide.

- s_0 is the initial state and it is represented by a double circle.

The possible states of each facility are the following:

- OK: The facility works properly.

- NRC: There is no reason to check the facility.

- SBC: The facility should be checked.

	B	VA	A	LA	S
OK	KO	SBC	NRC	NRC	OK
NRC	KO	SBC	SBC	NRC	OK
SBC	KO	KO	SBC	NRC	OK
KO	KO	KO	KO	SBC	NRC

Table 6: Component T of the states machine.

- KO: The facility does not work.

The designed machine is graphically displayed in Figure 3 and it is formally represented as $M = \{S, \Sigma, T, s_0\}$, where $S = \{OK, NRC, SBC, KO\}$, $\Sigma = \{S, LA, A, VA, B\}$ (the linguistic labels defined for the value y_i in page 14), T is given by Table 6 and $s_0 = \{OK\}$.
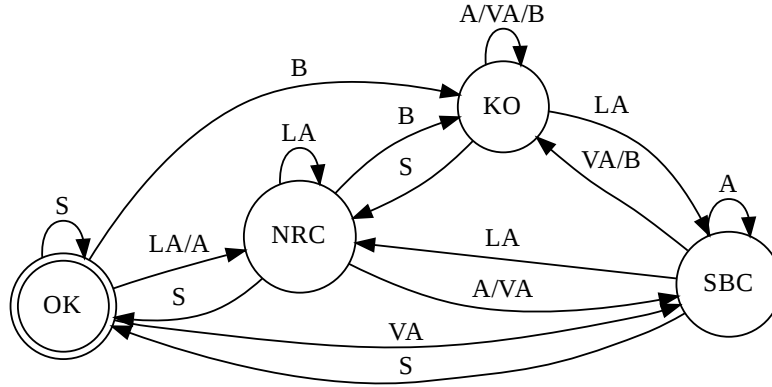

Figure 3: State machine for every facility

The nodes in Figure 3 represent the states of the considered facility, whereas the edges indicate the performances of the considered facility in a specific day. Considering as initial node the state of a particular facility in a specific day, the transition to the next state for the following day will be carried out following the corresponding edge, that is, according to the performance of the facility in the following day. For example, if one given facility is "OK" as initial day (state for Day 1) and next day the facility has a lightly anomalous performance "LA" (performance of Day 2), then the new state of the facility is "NRC"(state for Day 2). Now, the current state is "NRC". If the performance in the next day is "B" (performance of Day 3), then a transition to the state "KO" (state for Day 3) will be. However, if it has an anomalous performance "A" (state for Day 3),

the new state is "SBC" (state for Day 3).

Thus, analyzing the performance in consecutive days, we can take decisions related to the state of the facilities.

4. Illustrative example of the decision support system

In this section, we will show the results obtained by the decision support system described in this paper. We have analyzed and contrasted the results obtained by the presented procedure from January 1st to October 18th 2020, that is, 292 days. Nevertheless, since the automatic system returns the OWA value (y_i) and the state for each facility per each day, we have chosen some consecutive days in order to illustrate how the system may modify the decisions related to the state of the facilities. In particular, we choose the period from April 16th to 19th, 2020 and the facilities $\mathcal{I}_1$ and $\mathcal{I}_3$ to show how the system operates.

The results obtained from the decision system for both facilities are shown in Table 7. In this table, we can see that the facility $\mathcal{I}_1$ on April 16th had the state "NRC" (there is no reason to check the facility). Next day, the system returned 0 as global degree of suitable performance (y_1), which means that the system identifies the value $y_1 = 0$ with the linguistic label *bad performance* (B) and set the state of the facility $\mathcal{I}_1$ as "KO", according to the nodes of Figure 3. In this case, the facility does not work properly and should be inspected. Then, on April 18th the systems returned a global degree of 0.98 and the system identified it as *lightly anomalous performance* (LA) and it set the state to "SBC" (the facility should be checked), by the diagram in Figure 3. At this point, the company examined the facility $\mathcal{I}_1$. On the last day chosen, we obtained the value $y_1 = 1$, that is, *suitable performance* (S) and from the state "SBC" the system changed it to "OK" implying that the possible problem that arose has been already fixed. As we have seen, the facility $\mathcal{I}_1$ went through for all states. This example shows how the decision system automatically detects and alerts a problem (NRC and KO states), and as later, recovers OK once it has been solved, passing through the SBC state. Since the message of the state machine for both detection and fix the problem can take more than one day, the system also return the value of the OWA operator in order to know the behaviour of each station during the day. Hence, the user have enough information for taking a decision.

Day	y_1	Transition	state $\mathcal{I}_1$	y_3	Transition	state $\mathcal{I}_3$
16-04-2020	0.88	LA	NRC	1	S	OK
17-04-2020	0	B	KO	1	S	OK
18-04-2020	0.98	LA	SBC	1	S	OK
19-04-2020	1	S	OK	1	S	OK

Table 7: Results obtained by the automatic system for facilities $\mathcal{I}_1$ and $\mathcal{I}_3$.

On the other hand, we have analyzed all outputs obtained by the decision system by means of a confusion matrix [5] for each facility. We consider that

our system gives an alert when it returns either "SBC" or "KO" states. On the contrary, no-alert corresponds to "NRC" and "OK" states.

Table 8 shows the confusion matrices of each facility corresponding to our decision support system together with the obtained errors of the model, for all data collected in 2020. From the confusion matrices we can say that there are two different scenarios. On the one hand, we can mention that both facilities $\mathcal{I}_2$ and $\mathcal{I}_3$ have worked correctly during all the days of the considered period and our system neither report false positive nor false negative. On the other hand, we can observe that the rest of the facilities have alerts and in this case, our system reports some false negatives. The fact that the system does not detect all the alerts may be given because of the noise in the data. This fact will be improved in the future.

$\mathcal{I}_1$	No Alert	Alert	Model error
No Alert	284	2	0.699%
Alert	0	6	0%
Error of use	0%	25%	0.685%

$\mathcal{I}_2$	No Alert	Alert	Model error
No Alert	292	0	0%
Alert	0	0	0%
Error of use	0%	0%	0%

$\mathcal{I}_3$	No Alert	Alert	Model error
No Alert	292	0	0%
Alert	0	0	0%
Error of use	0%	0%	0%

$\mathcal{I}_4$	No Alert	Alert	Model error
No Alert	183	35	16.05%
Alert	0	74	0%
Error of use	0%	32.11%	11.98%

$\mathcal{I}_5$	No Alert	Alert	Model error
No Alert	68	9	11.69%
Alert	0	215	0%
Error of use	0%	4.02%	3.08%

$\mathcal{I}_6$	No Alert	Alert	Model error
No Alert	184	34	15.60%
Alert	0	74	0%
Error of use	0%	31.48%	11.64%

Table 8: Confusion matrices of each facility of the obtained results in 2020.

In particular, facilities $\mathcal{I}_1$ and $\mathcal{I}_5$ have 2 and 9 few false negatives, respectively, although the operation of $\mathcal{I}_1$ has been better than $\mathcal{I}_5$ taking into account the number of alerts of each one. For these two facilities the system correctly classifies the 99,315% and 96,92% of the considered range of days, and detects 75% and 95.93% of the alerts, respectively. Similarly, for facilities $\mathcal{I}_4$ and $\mathcal{I}_6$ it correctly classifies the 88,02% and 88,36% of the days, and detects the 67.89% and 69.1% of the alerts, respectively. These two facilities have a significant number of alerts, but the nature of these alerts is different. For instance, the facility $\mathcal{I}_4$ had an inverter in bad conditions but the production in comparison with the rest of the facilities is not far enough for the system to detect the failure. However, the facility $\mathcal{I}_6$ had an irregular functioning for most of the time, having low production peaks and recovering the next day. As we previously explained, the system does not detect this type of scenario as an alert until the problem persists for a brief period of time, but the user is alerted by the irregularity of the station. In addition, Figure 4 illustrates these situations where the system does not immediately detect an alert. For example, on day 16th, $\mathcal{I}_1$ has a low production peak but the system does not return an alert until day 17th because its recovery is not fast enough. After that, $\mathcal{I}_1$ returns to normality. Moreover, we can observe on day 20th that $\mathcal{I}_4$ decreased its production and the system alert with NRC. Until June 25th, the system does not alert with SBC, where the differences with the total performance of energy production of each facility are more significant. Since NRC is not classified as alert in Table 8, we are considering these days as FN although the system is sending a warning with the label NRC. Lastly, the irregular performance that $\mathcal{I}_6$ had during these days can be perfectly appreciated in Figure 4.

From the comments above on $\mathcal{I}_4$, we can see that the crisp character of the confusion matrix does not work perfectly to assess the accuracy of our mechanism, transforming four different states into only two values (true or false). For example, it should be noted that most of the false negatives returned by the system correspond to the state "NRC", but this state can be interpreted as a warning of the start of a malfunction, which will be detected if the performance of the installation gets worse the next day. Specifically, all false negatives in $\mathcal{I}_1$, $\mathcal{I}_5$ and $\mathcal{I}_6$ correspond to the state "NRC". Moreover, in the facility $\mathcal{I}_4$, 21 of 35 false negatives also correspond to the state "NRC". Therefore, the system would go from having 80 false negatives to having only 14 false negatives, which corresponds to an error rate of only 4.79%. In the future, we will try to improve the classification of the days (as correct and incorrect performances) and we also try to readjust the values of the intervals for a more optimal classification.

Notice that the system does not report any false positive, therefore all alerts reported are true and it prevents the company from making an effort to check the installation unnecessarily. In addition, according to Table 8, we can ensure that the obtained errors are small and we can assume that our system is reliable.
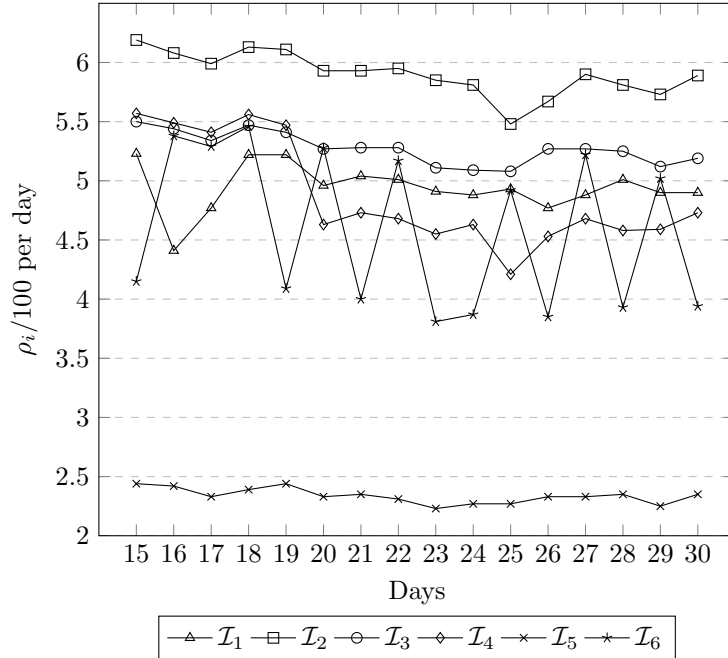
Figure 4: Total performance of energy production of $\mathcal{I}_i$ in June 2020.

5. Conclusions and future work

The main goal addressed in this paper has been to solve the problem of "constant loss of energy production", since when the loss of production is very small, it can be imperceptible from the user and so, it could not be detected during many months. To avoid this situation, a decision support system has been proposed for detecting anomalous behaviors in the energy production of the photovoltaic facilities over time. For the development of this system, different functions which describe the behaviors of the facilities has been designed, as well as a state machine that allows us to determine the state of each facility, taking into account the states and the performances obtained in the previous days for each facility. In addition, an experimental example has been carried out in order to test the proposed automatic decision system obtaining good results.

In the future, we will be focused on different improvements. The first one will be to consider POWA instead of OWA operators as aggregation operators and analyze if the use of these operators improves the obtained results. We also want to develop the system in order to detect if the error in productivity is due to: dirty on the solar cells, investor in poor condition, investor has stopped working, and there are badly functioning photovoltaic cells. In addition, the mechanism will be extended with the consideration of meteorological variables, obtained from weather stations installed in the facilities, in order to obtain a new mechanism to be available to be compared with the current

ones that take into account these variables. Moreover, in line with other failure prognosis approaches [29], we will carry out a predictive maintenance of each installation, anticipating in advance any deficiency that may happen in a particular installation. Furthermore, a power generation predictability module will be designed. In this way, one could know approximately how much energy is going to be produced per installation. This fact can be very interesting if the production is needed to supply an own facility/building. For example, if the installation is going to heat the water of a sport center or school, and we know that not enough energy will be generated, we can take advantage of this information for performing some action in advance to heat the water. These developments will be studied considering different mathematics and artificial intelligence tools [1, 9, 10, 16, 31, 32, 41, 34, 56, 60].

References

[1] L. Antoni, M. E. Cornejo, J. Medina, and E. Ramirez. Attribute classification and reduct computation in multi-adjoint concept lattices. *IEEE Transactions on Fuzzy Systems*, pages 1–1, 2020.

[2] O. Aristondo and M. Ciommi. The decompositions of rank-dependent poverty measures using ordered weighted averaging operators. *International Journal of Approximate Reasoning*, 76:47 – 62, 2016.

[3] G. Beliakov, S. James, and J.-Z. Wu. *Symmetric Fuzzy Measures: OWA*, pages 135–192. Springer International Publishing, Cham, 2020.

[4] R. Ben Ammar, M. Ben Ammar, and A. Oualha. Photovoltaic power forecast using empirical models and artificial intelligence approaches for water pumping systems. *Renewable Energy*, 153:1016–1028, 2020.

[5] R. A. Berk. *Statistical learning from a regression perspective*. Springer Series in Statistics. Springer Science+Business Media LLC, New York, NY, USA, 2008.

[6] K.-H. Chang. A decision support system for planning and coordination of hybrid renewable energy systems. *Decision Support Systems*, 64:4 – 13, 2014.

[7] Z. Chen, L. Wu, S. Cheng, P. Lin, Y. Wu, and W. Lin. Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and i-v characteristics. *Applied Energy*, 204:912 – 931, 2017.

[8] K.-P. Chiao. Multiple criteria decision making for linguistic judgments with importance quantifier guided ordered weighted averaging operator. *Information Sciences*, 474:48 – 74, 2019.

[9] M. E. Cornejo, D. Lobo, and J. Medina. Extended multi-adjoint logic programming. *Fuzzy Sets and Systems*, 388:124–145, 2020. Logic.

[10] C. Cornelis, J. Medina, and N. Verbiest. Multi-adjoint fuzzy rough sets: Definition, properties and attribute selection. *International Journal of Approximate Reasoning*, 55:412–426, 2014.

[11] L. De Miguel, D. Paternain, I. Lizasoain, G. Ochoa, and H. Bustince. Some characterizations of lattice OWA operators. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 25(Suppl. 1):5–17, 2017.

[12] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales. Diagnostic method for photovoltaic systems based on six layer detection algorithm. *Electric Power Systems Research*, 151:26 – 39, 2017.

[13] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, and P. Mather. Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system. *Energy*, 140:276 – 290, 2017.

[14] P. D'Urso and J. M. Leski. Fuzzy clustering of fuzzy data based on robust loss functions and ordered weighted averaging. *Fuzzy Sets and Systems*, 2019.

[15] M. El-Zekey, J. Medina, and R. Mesiar. Lattice-based sums. *Information Sciences*, 223:270 – 284, 2013.

[16] J. M. García, L. Rodriguez-Benitez, L. J. Linares, and G. Triviño. A linguistic extension of petri nets for the description of systems: An application to time series. *IEEE Transactions on Fuzzy Systems*, 27(9):1818–1832, 2019.

[17] H. Garg. Linguistic pythagorean fuzzy sets and its applications in multiattribute decision-making process. *International Journal of Intelligent Systems*, 33(6):1234–1263, 2018.

[18] M. Hajji, M.-F. Harkat, A. Kouadri, K. Abodayeh, M. Mansouri, H. Nounou, and M. Nounou. Multivariate feature extraction based supervised machine learning for fault detection and diagnosis in photovoltaic systems. *European Journal of Control*, 2020. In press. https://doi.org/10.1016/j.ejcon.2020.03.004.

[19] M. A. Heldeweg and S. Saintier. Renewable energy communities as 'socio-legal institutions': A normative frame for energy decentralization? *Renewable and Sustainable Energy Reviews*, 119:109518, 2020.

[20] J. Hopcroft, R. Motwani, and J. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Pearson, 2006.

[21] M. Hussain, M. Dhimish, S. Titarenko, and P. Mather. Artificial neural network based photovoltaic fault detection algorithm integrating two bidirectional input parameters. *Renewable Energy*, 155:1272 – 1292, 2020.

[22] L. G. Jaime, E. E. Kerre, M. Nachtegael, and H. Bustince. Consensus image method for unknown noise removal. *Knowledge-Based Systems*, 70:64 – 77, 2014.

[23] L. Jin and G. Qian. Owa generation function and some adjustment methods for owa operators with application. *IEEE Transactions on Fuzzy Systems*, 24(1):168–178, 2016.

[24] I. U. Khalil, A. Ul-Haq, Y. Mahmoud, M. Jalal, M. Aamir, M. U. Ahsan, and K. Mehmood. Comparative analysis of photovoltaic faults and performance evaluation of its detection techniques. *IEEE Access*, 8:26676–26700, 2020.

[25] A. Kishor, A. K. Singh, S. Sonam, and N. R. Pal. A new family of owa operators featuring constant orness. *IEEE Transactions on Fuzzy Systems*, 28(9):2263–2269, 2020.

[26] P. Kobylinski, M. Wierzbowski, and K. Piotrowski. High-resolution net load forecasting for micro-neighbourhoods with high penetration of renewable energy sources. *International Journal of Electrical Power & Energy Systems*, 117:105635, 2020.

[27] M. Kordestani, A. Chibakhsh, and M. Saif. A control oriented cyber-secure strategy based on multiple sensor fusion. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 1875–1881, 2019.

[28] M. Kordestani, A. Mirzaee, A. A. Safavi, and M. Saif. Maximum power point tracker (mppt) for photovoltaic power systems-a systematic literature review. In *2018 European Control Conference (ECC)*, pages 40–45, 2018.

[29] M. Kordestani, M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani. Failure prognosis and applications: A survey of recent literature. *IEEE Transactions on Reliability*, 70(2):728–748, 2021.

[30] M. Kordestani, M. F. Samadi, M. Saif, and K. Khorasani. A new fault diagnosis of multifunctional spoiler system using integrated artificial neural network and discrete wavelet transform methods. *IEEE Sensors Journal*, 18(12):4990–5001, 2018.

[31] G. Kou, Ö. Olgu Akdeniz, H. Dinçer, and S. Yüksel. Fintech investments in european banks: a hybrid it2 fuzzy multidimensional decision-making approach. *Financial Innovation*, 7:39, 2021.

[32] G. Kou, H. Xiao, M. Cao, and L. H. Lee. Optimal computing budget allocation for the vector evaluated genetic algorithm in multi-objective simulation optimization. *Automatica*, 129:109599, 2021.

[33] R. Leszczyna, T. Wallis, and M. R. Wróbel. Developing novel solutions to realise the european energy – information sharing & analysis centre. *Decision Support Systems*, 122:113067, 2019.

[34] G. Li, G. Kou, and Y. Peng. A group decision making model for integrating heterogeneous information. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 48(6):982–992, 2018.

[35] X. Liu, Y. Xu, R. Montes, Y. Dong, and F. Herrera. Analysis of self-confidence indices-based additive consistency for fuzzy preference relations with self-confidence and its application in group decision making. *International Journal of Intelligent Systems*, 34(5):920–946, 2019.

[36] I. Lizasoain and C. Moreno. OWA operators defined on complete lattices. *Fuzzy Sets and Systems*, 224:36 – 52, 2013. Theme: Aggregation functions and implications.

[37] M. Manohar, E. Koley, and S. Ghosh. Microgrid protection under weather uncertainty using joint probabilistic modeling of solar irradiance and wind speed. *Computers & Electrical Engineering*, 86:106684, 2020.

[38] L. G. Marín, N. Cruz, D. Sáez, M. Sumner, and A. Nuñez. Prediction interval methodology based on fuzzy numbers and its extension to fuzzy systems and neural networks. *Expert Systems with Applications*, 119:128 – 141, 2019.

[39] A. Mattiussi, M. Rosano, and P. Simeoni. A decision support system for sustainable energy supply combining multi-objective and multi-attribute analysis: An australian case study. *Decision Support Systems*, 57:150 – 159, 2014.

[40] J. Medina. Characterizing when an ordinal sum of t-norms is a t-norm on bounded lattices. *Fuzzy Sets and Systems*, 202:75–88, 2012.

[41] J. Medina, J. Moreno-García, E. Ramírez-Poussa, and L. T. Kóczy. Mathematics and computational intelligence synergies for emerging challenges. *International Journal of Computational Intelligence Systems*, 14:818–820, 2021.

[42] J. Medina and R. R. Yager. OWA operators with functional weights. *Fuzzy Sets and Systems*, 414:38–56, 2021. Aggregation Functions.

[43] R. Mesiar, F. Kouchakinejad, A. ˇSipoˇsová, and M. Mashinchi. Owa operators on complete lattices. *IEEE Transactions on Fuzzy Systems*, 26(6):3884–3887, 2018.

[44] M. M. Morato, P. R. Mendes, J. E. Normey-Rico, and C. Bordons. LPV-MPC fault-tolerant energy management strategy for renewable microgrids. *International Journal of Electrical Power & Energy Systems*, 117:105644, 2020.

[45] M. Mousavi, M. Moradi, A. Chaibakhsh, M. Kordestani, and M. Saif. Ensemble-based fault detection and isolation of an industrial gas turbine. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2351–2358, 2020.

[46] X. Pu, L. Jin, R. Mesiar, and R. R. Yager. Continuous parameterized families of rim quantifiers and quasi-preference with some properties. *Information Sciences*, 481:24 – 32, 2019.

[47] S. Rodrigues, G. Mütter, H. G. Ramos, and F. Morgado-Dias. Machine learning photovoltaic string analyzer. *Entropy*, 22(2):205, 2020.

[48] M. I. Said, M. Steiner, G. Siefer, and A. H. Arab. Maximum power output prediction of HCPV FLATCON® module using an ANN approach. *Renewable Energy*, 152:1274 – 1283, 2020.

[49] S. Samara and E. Natsheh. Intelligent pv panels fault diagnosis method based on NARX network and linguistic fuzzy rule-based systems. *Sustainability*, 12(5):2011, 2020.

[50] S. Saminger-Platz, E. Klement, and R. Mesiar. On extension of triangular norms on bounded lattices. *Indagationes Mathematicae*, 19(1):135–150, 2008.

[51] S. R. Sandeep and R. Nandihalli. Optimal sizing in hybrid renewable energy system with the aid of opposition based social spider optimization. *Journal of Electrical Engineering & Technology*, 15:433–440, 2020.

[52] K. Siminski. Fuzzy weighted c-ordered means clustering algorithm. *Fuzzy Sets and Systems*, 318:1 – 33, 2017. Theme: Clustering and Image Processing.

[53] R. Singh and R. C. Bansal. Optimization of an autonomous hybrid renewable energy system using reformed electric system cascade analysis. *IEEE Transactions on Industrial Informatics*, 15(1):399–409, 2019.

[54] T. Takashima, J. Yamaguchi, K. Otani, T. Oozeki, K. Kato, and M. Ishida. Experimental studies of fault location in pv module strings. *Solar Energy Materials and Solar Cells*, 93(6):1079–1082, 2009. 17th International Photovoltaic Science and Engineering Conference.

[55] J. P. Torreglosa, P. García-Triviño, L. M. Fernández-Ramírez, and F. Jurado. Control based on techno-economic optimization of renewable hybrid energy system for stand-alone applications. *Expert Systems with Applications*, 51:59 – 75, 2016.

[56] H. Xiao, Y. Yan, G. Kou, and S. Wu. Optimal inspection policy for a single-unit system considering two failure modes and production wait time. *IEEE Transactions on Reliability*, pages 1–13, 2021.

[57] R. Yager. On ordered weighted averaging aggregation operators in multi-criteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18:183–190, 1988.

[58] R. R. Yager. OWA aggregation of multi-criteria with mixed uncertain satisfactions. *Information Sciences*, 417:88 – 95, 2017.

[59] R. R. Yager and J. Kacprzyk. *The Ordered Weighted Averaging Operators: Theory and Applications*. Kluwer: Norwell, MA, 1997.

[60] J. Yao, J. Medina, Y. Zhang, and D. Ślęzak. Formal concept analysis, rough sets, and three-way decisions. *International Journal of Approximate Reasoning*, 140:1–6, 2022.

[61] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.