Systematic modeling of photovoltaic modules based on artificial neural networks

Jose Manuel Lopez-Guede [a,*], *Jose Antonio Ramos-Hernanz* [b], *Ekaitz Zulueta* [a], *Unai Fernadez-Gamiz* [c], *Fernando Oterino* [d]

[a] *University of the Basque Country, University College of Engineering of Vitoria-Gasteiz, Systems Engineering and Automatic Control Department, 01006 Vitoria-Gasteiz, Spain*

[b] *University of the Basque Country, University College of Engineering of Vitoria-Gasteiz, Electrical Engineering Department, 01006 Vitoria-Gasteiz, Spain*

[c] *University of the Basque Country, University College of Engineering of Vitoria-Gasteiz, Nuclear Engineering and Fluid Mechanics Department, 01006 Vitoria-Gasteiz, Spain*

[d] *University of the Basque Country, University College of Engineering of Vitoria-Gasteiz, Nuclear Electronics Department, 01006 Vitoria-Gasteiz, Spain*

ABSTRACT

Using accurate models of photovoltaic modules is of major importance to make realistic simulations of these systems in order to study their elements for a better performance. In this paper we address the problem of the lack of a systematic procedure to obtain accurate artificial neural network based models in an unattended way with large datasets. We face this problem introducing a novel systematic procedure to carry out this task. As proof of concept, we tested the procedure modeling a Mitsubishi Electric PV-TD185MF5 (185 Wp) photovoltaic module placed at the University College of Engineering of Vitoria-Gasteiz (University of the Basque Country, Spain). We have used a dataset of 63,000 samples collected during 18 months (from August 2013 to February 2015). The main findings of the paper are two. The first one is that the systematic procedure works properly because it generated autonomously two very accurate models of I_{PV} (RMSE with unknown test dataset of 0.045 A for a one hidden layer neural network, while 0.042 A for a two hidden layers neural network), i.e., we conclude that the unattended execution of the systematic pro- cedure introduced in this paper has obtained models which have learned the electrical behavior of the photovoltaic module with an accuracy higher than the measurement de- vices precision. We have compared these results with recent relevant papers and we found that the proposed procedure is competitive and improves the state-of-art results. The second finding is that using these models lead

to space savings larger than 99.5% of the original tabular representation of the dataset.

Introduction

Due to the increasing importance of renewable energy, and thereby among them of the photovoltaic energy, there is a great interest in studying its elements for a better perfor- mance. The knowledge of its elements allows more reliable facilities and increased energy production. For example, if accurate models are available, it would be possible to estimate in real time the electricity production from irradiance and temperature data, or even throw alarms if the expected per- formance is not being reached because it could be a symptom of any kind of problem.

When dealing with photovoltaic systems (a cell or a mod- ule), in order to study its behavior it is mandatory to know the value of the supplied voltage and current (V_{PH} and I_{PH}) in different working conditions.

Depending on the requirements, it is possible to use simple models to obtain enough accuracy. The parameters required for its definition can be simply determined from the data sheets provided by the manufacturers. The problem arises when a more complex or detailed model is needed. The embodiment and the parameters of a model of a photovoltaic module are not very different from the ones of a photovoltaic cell. All parameters are the same unless the short circuit voltage V_{OC}, which is divided by the number of cells in series N_S that are in the module.

In the literature there are a number of models of different complexities to explain the electrical behavior of photovoltaic modules. In order to clarify such variety we can make a first division into theoretical and empirical models.

Theoretical models use a characteristic equation [23] (see Section Ideal photovoltaic cell), being among them a number of models which use different degrees of freedom, i.e., some of them are very complete but in other cases authors determine/ approximate them in several practical ways in order to get these models being useful.

The most complete model used in the literature is based on a double diode equivalent circuit which leads to a 7 parame- ters model, i.e., a_1, a_2, R_S, R_{SH}, I_{01} , I_02 and I_{PH}. One of the fist using this model was [14], which adjusted the parameters through LevenbergeMarquardt and NewtoneRaphson algo- rithms. This model is still being used in more recent literature. For example, in Ref. [26] authors introduce a new procedure to calculate the IeV characteristics of only thin-film modules. After describing such procedure (based on an explicit rational form), the model is used to calculate the current of five com- mercial photovoltaic modules, reaching a mean of absolute differences between the calculated and issued current of 0.042 A for five different modules. A more wide field of application can be found in Ref. [11], where the developed model is valid for three types of cells (polycrystalline, amorphous and thin film). In order to estimate the 7 parameters, authors specify an iterative algorithm which uses a number of parameters of the modules provided by the data sheet of the manufacturer. They compare the curves resulting from their algorithm with those of the manufacturer through a plethora of figures showing that they are very close, but no numerical results on their

accuracy are provided. Anyway, the results and the compari- son are entirely theoretical because they are not confronted

with empirical values. The estimation of the 7 parameters can be carried out by means of an evolutionary algorithm as in Ref. [28], where the model is validated with experimental data obtained from a polycrystalline module under 7 different weather conditions with a total of 482 samples (certainly not too many), reaching an average RMSE of 0.0608 A for all these weather conditions.

Another quite complete approach used in the literature is based on a single diode equivalent circuit which leads to a 5 parameters model, i.e., a, R_S, R_{SH}, I_0 and I_{PH}. In Ref. [5] the estimation of the parameters is done again through a genetic algorithm which uses a number of parameters taken from the data sheet of the manufacturer of a concrete monocrystalline module. Authors do not give the resulting RMSE value of their approximation, but the comparison is done with five key points given by the manufacturer, i.e., theoretical values. In Ref. [36] the approach used by authors is to modify the value of R_S and R_{SH} and adjust them by means of an iterative algorithm to fit the theoretical curve to experimental data in the MPP, implementing the complete model under Simulink- SimPowerSystem assisted by Simulink. Authors of [33] use this 5 parameters approach building three versions: from a basic model with two approximations of five magnitudes to a model where they are calculated. Authors compare the results generated by their model with experimental data obtained from a facility by other authors, and study the effect of changing R_S and R_{SH}. In Ref. [7] authors present an interesting use of this kind of model in a hybrid PV and fuel system. Finally, a much more direct approach is used in Ref. [27], where Matlab toolbox components are used and authors generate an application with GUIDE.

A number of authors make some approximations of the characteristic parameters. The most usual of them is to as- sume that R_{SH} ¼ ∞, so I_{SH} ¼ 0 A and the third term of the characteristic equation is discarded leading to a 4 parameters model. In Refs. [13,37] this simplification is carried out when the model of the exactly same module is built following almost the same procedure. These works are quite similar to [30]. However there are several differences: the firsts calculate the temperature coefficient k_0 used to calculate the short cir- cuit current I_{SC}, while the last one uses the value provided by the manufacturer; and the ideality factor of the diode is different between [13,37] and [30]. A model where the ideality factor can be varied to fit it to different PV technologies is proposed in Ref. [34]. As a last 4 parameters model [6], for- mulates a distinct model from the previous ones and develops the model of a microgrid facility in an hybrid PV and fuel system.

The remaining theoretical approaches are even more simple because they consider R_S ¼ 0 U or I_{PH} ¼ I_{SC}, leading to a 3 parameters model. For example, in Ref. [3] almost a toy experiment (due to the small power of the PV module) applied to H_2 extraction is presented, while [4] uses the Rauschenbach model and carries out a comparison with a real farm. A model that defines the magnitude equations using Simulink is pre- sented in Ref. [32]: two models are built, the first one through Simulink tags (from, goto) and the other one by means of the practical realization of the equivalent circuit using electrical components (resistances, diode, etc.), in such a way that instead of using the characteristic equation

they use the

characteristic circuit. In this last model authors consider the approximation that $I_{PH} \approx I_{SC}$.

With regard to empirical models, they are based on learning the model from data. These data, in turn, could be obtained from the characteristic equation when instanced by means of particular characteristic parameters or from direct measurements obtained from real photovoltaic modules, being this second option the most used. The main advantages of using one approach are the drawbacks of the other one and vice versa: using real data it is possible to obtain accurate models, however it is mandatory to have these data and the model is suited only for that photovoltaic module. On the other hand, using data extracted from theoretical models shortens the process and changing characteristic parameters it is possible to obtain models for a number of photovoltaic modules without having them physically.

In Ref. [16] authors use polynomial interpolation and describe a model which allows to obtain only the IV and PV curves and provides further insight into the behavior of the photovoltaic module, but it is suited only for a range of tem- peratures and irradiances.

Other approach to the empirical methods is modeling the behavior of the photovoltaic elements by means of artificial neuronal networks. In Ref. [20] an interesting review of different approaches is presented, while [1,2] developed a practical model, but with very narrow amplitudes for irradi- ance and temperature. There are other authors who are seeking to expand these limits [22], but in spite of expand the range of magnitudes, they still obtain partial models with data captured during quite small time range (two months). In Ref.

[8] authors use RFB neural networks to generate two models. The former is to obtain I_{PV} from irradiance and voltage, and uses 5600 real samples. The second one is devoted to obtain P_{PV}, again from irradiance and voltage using 4600 real sam- ples. Authors reported a relative MSE of 2% and 1% respec- tively, but the test has been done with training data, so in fact they have reported the training accuracy. Two different models based on artificial neural networks are generated in Ref. [25]. Both models have temperature and irradiance as inputs, while power as output. The main difference is that one on them is specialized in cloudy days (S400 $W/m^2/day$) while

them to carry out a sensitivity analysis to study the effect of each one of the inputs of the networks on the generated power.

The main objective of this paper is to address a double problem: the former is to derive systematic procedure to obtain global (not partial) models of photovoltaic modules in an unattended way, i.e., without human intervention. The second one is to generate an accurate model of the Mitsubishi Electric PV-TD185MF5 (185 Wp) photovoltaic module, which will be used as a proof of concept of the systematic procedure. In order to tackle these problems we have used an artificial neural network based approach because they have several advantages: they have shown good learning and generaliza- tion capabilities in different domains, once that they have been trained (i.e., the model has been obtained) the use is very simple, and when they are in use the calculations are per- formed very quickly.

Both previously enumerated problems have been over- comed as shows the obtention of an accurate neural model with a I_{PH} test performance (Root Mean Square Error, RMSE) of 0.0422 A (the 0.52% of I_{SC} ¼ 8.13 A, i.e., better accuracy than the measuring instrumental as will be discussed in Section Experimental design of systematic modeling procedure validation), obtained with a long-time dataset of 18 months and 63,000 samples. Besides, this procedure obtained a model that improves those generated by the state-of-art methods where comparable.

The paper is structured as follows. Section Background recalls some basic concepts on ideal photovoltaic cells and artificial neural networks. The systematic procedure to model the behavior of photovoltaic panels based on artificial neural networks and the formulation of a number of approaches to study are introduced in Section ANN based photovoltaic module modeling systematic procedure. Section Experimental design of systematic modeling procedure validation describes in detail the experimental design to validate the systematic procedure, while the obtained exper- imental results are discussed and compared with previous relevant works in Section Experimental results. Finally, our conclusions are given in Section Conclusions and future work.

the other one in sunny days (>400 $W/m^2/day$). Authors havegathered data during 13 months to build them from a real module, recording 5760 samples as cloudy days, while 2280 samples as sunny days. The training process of each model has been carried out with 70% of samples, while the validation (not test) with the remaining 30%. The reported RMSE for the first model with a test dataset of 4 cloudy consecutive days is 0.115%, while 0.11% for the second one with a test dataset of 4 sunny consecutive days. A more recent work is [35], where authors obtained two neuronal models from experimental data gathered during 4 months, each one for two different modules (monocrystalline and flexible organic modules). The input of the networks are temperature, irradiance and relative humidity, while the output is the power. The 65% of data were used for training (2868 samples), 20% for validation (833 samples) and the remaining 15% for test (662 samples). The correlation coefficient is 0.8046 and 0.8493, while the average error is 0.05664 W and 0.01169 W respectively for each model. Once that the authors have got accurate models, they use

Background

Ideal photovoltaic cell

A first approximation to the ideal photovoltaic cell is to model it as a current source with an anti-parallel diode, where the direct current generated when the cell is exposed to light varies linearly with the solar radiation (sub-circuit enclosed by a dotted line on the left part of Fig. 1). A second and more accurate approximation is shown through overall Fig. 1, where the improvement of the model includes the effects of a shunt resistor and other one in series. The main involved magni- tudes are the following: I_{PH} is the photogenerated current or photocurrent (A), I_D is the current of the diode (A), R_{SH} is the shunt resistance (U) and R_S is the series resistance (U). This equivalent circuit can be used either for an individual cell, a module that consists of several cells or for a matrix that is the union of several

modules.

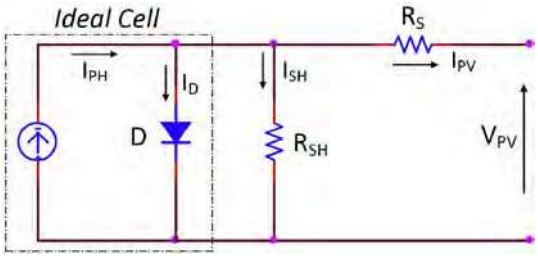


Fig. 1 e Ideal photovoltaic cell.

Based on the ideal equivalent circuit of Fig. 1, we can describe the relationship between the voltage (V_{PV}) and the current (I_{PV}) supplied by the photovoltaic cell by means of the characteristic equation Eq. (1) [14], so that expanding each current term Eq. (2) [23] is obtained:

$$I_{PV} ¼ I_{PH} – I_D – I_{SH}; \quad (1)$$

different situations, the response of a neural network in unseen situations (i.e., with unseen inputs) will probably be acceptable and quite similar to the correct response. So it is said that they have the *generalization property*.

- Real time capabilities: Once they are trained, and due to their parallel internal structure, their response is always very fast. Their internal structure could be more or less complex, but in any case, all the internal operations that must be carried out are a number of multiplications and additions if it is a linear neural network. This fast response is independent of the complexity of the learned models.

There is a number of types of ANN, each of one of them well suited to different problems and applications. Besides, there is one or more learning algorithms for each of those types of ANN, having their own strengths and weaknesses. To learn more about neural networks in general see Refs. [38,9,31,21], and to get deeper insight on identification and control of dynamical systems using ANN see Refs. [29,24].

$$I \quad ¼ I \quad – I \left(e^{\frac{q(VPV + IPVRS)}{aKT}} – 1 \right) – \frac{VPH + IPHRS}{RSH}; \quad (2)$$

I_{PV} I_{PH} 0

ANN based photovoltaic module modeling

where:

I_0 is the saturation current of the diode (A),

q is the charge of the electron, 1.6×10^{-19} (C),

a is the diode ideality factor of the diode,

K is the Boltzmann's constant, 1.38×10^{-23} (j/K),

T is the cell temperature (^{o} C).

In Eq. (2) several structural parameters of the cell are involved, i.e., I_{PH}, I_D, a, R_S and R_{SH}. Since this is a theoretical model, we have to use an estimate of all these parameters, and this circumstance leads necessarily only to approximate values when the study of a particular photovoltaic module is being carried out.

Besides, as one solar cell is only capable of generating very low terminal voltage and output current, for working pur- poses many cells are connected in series to form higher voltage across the terminal and connected in parallel to form a module. For large scale operation of PV generator, modules are connected in series and parallel to form an array. So, due to the large amount of photovoltaic cells used in large real in- stallations, the multiplier effect of these small approximation errors can be relevant.

Artificial neural networks

Using Artificial Neural Networks (ANN) is motivated by its ability to model complex systems [38]. These bio-inspired models have several advantages, and among others, these are the most outstanding to the problem that we are addressing:

- Learning capabilities: If they are properly trained, they can learn complex mathematical models. There are several well known training algorithms and good and tested implementations of them. The main challenge concerning this issue is to choose appropriate inputs and outputs to the black box model and the internal structure.
- Generalization capabilities: Again, if they are properly trained and the training examples cover a variety of

systematic procedure

In this section we describe four models to formulate the approximation to the electrical behavior of a photovoltaic module by ANNs (from subsection Approximation by M_1 ¼ (V_{PH} – I_{PH}) model to Approximation by M_4 ¼ (TGV_{PH} – I_{PH}) model), while subsection General systematic modeling procedure introduces a general system- atic procedure to get the most accurate instance of these models through ANNs.

Building a model or an approximation of the electrical behavior of the photovoltaic module consists in training ANNs to predict the I_{PH} (and in some cases also V_{PH}) given as input specific combinations

of the temperature T and the irradiance G of the surroundings of the module (and in some cases also V_{PH}).

Approximation by $M_1 = (V_{PH} - I_{PH})$ *model*

The main idea that guides the design of this model is the simplicity. In order to build a first model as simple as possible, we have assumed that the temperature T and the irradiance G are quite constant values, and we have discarded them as significant variables. In Fig. 2(a) we can see a schematic representation of this model.

So, following this model, the ANNs have one input neuron because there is only one real valued input (V_{PH}), and there is also one unique output neuron because there is only one real valued target (I_{PH}) associated with each input value. So, both the input and the output layers have only one neuron.

Approximation by $M_2 = (TG - V_{PH}I_{PH})$ *model*

The aim of the design of this model is to characterize the intrinsic behavior of a photovoltaic module, taking into ac- count only exogenous factors. So, we have assumed that only the temperature T and the irradiance G are relevant for this purpose. Fig. 2(b) shows a schematic representation of this model. In this case the ANNs have two input neurons because there are two real valued inputs (T and G), and there are also two output neurons because there are two real valued targets (V_{PH} and I_{PH}) associated with each input combination. In this way, both the input and the output layers have two neurons.

Approximation by $M_3 = (TG - V_{PH}, TG - I_{PH})$ *models*

The purpose of this model is the same that the previous one, i.e., to characterize the intrinsic behavior of a photovoltaic module, but we have considered two different ANN in order to get two smaller and more accurate models than the previous one, at the

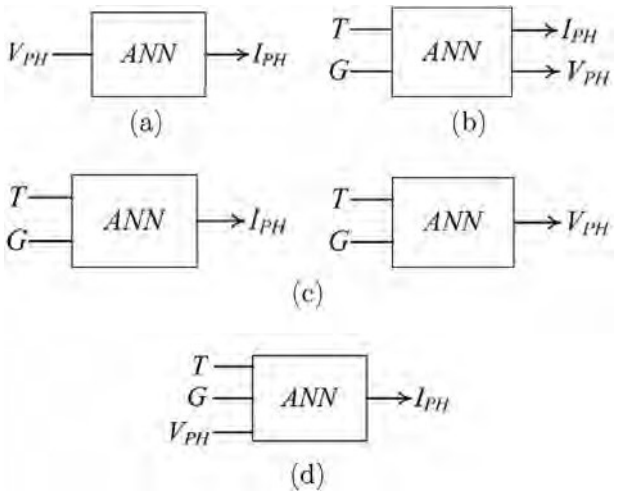


Fig. 2 e Schematic representation of different ANN models.

Fig. 3 e Photovoltaic modules installation (8 pieces), in University College of Engineering of Vitoria-Gasteiz (University of the Basque Country, Spain).

same time that we can use the ANN which models the magni- tude of interest if only one of them is needed. We have consid- ered again that only the temperature T and the irradiance G are relevant for this purpose. Fig. 2(c) shows the double ANN model. Taking into account the these considerations, the ANNs have two input neurons because there are two real valued inputs (T and G), while there is one unique output neuron for each type of ANN because there is only one real valued target (V_{PH} or I_{PH}) associated with each combination of input values.

The complete model of the photovoltaic module is obtained by gathering the responses of the two independent ANNs.

Algorithm 1 General systematic photovoltaic module model learning procedure through ANNs.

1. Obtain raw data of the relevant magnitudes
2. for each ANN learning approximation,
   (a) Adapt the input and output patterns obtained in the step 1 to the description used in the approximation
   (b) Modify the input patterns according to the approximation (normalization and noise addition)
   (c) Partition them into three datasets (60% train, 20% validation and 20% test)
   (d) Train ANNs with the adapted train input and output patterns following the approximation description
   (e) Test the learned approximation with the adapted test dataset in order to validate it

Approximation by $M_4 = (TGV_{PH} - I_{PH})$ *model*

All models pursue to be as accurate as possible, but their design bears in mind other main objectives as simplicity, etc. This last model design, however, has been focused on accu- racy. In order to get this objective the temperature T and the irradiance G have not been considered constant, so they have not been discarded. In Fig. 2(d) we can see a schematic rep- resentation of this complete model.

The ANNs which implement this complete model have three dimensional real valued input patterns (T, G and V_{PH}), i.e., three input neurons, while there is only one real valued target (I_{PH}) associated with each input combination.

General systematic modeling procedure

Once several approaches with different specificities have been introduced in previous subsections, a general specification of the training systematic process for the accurate approxima- tion of the different models by ANNs is given in Algorithm 1. The first step is to obtain raw data of the relevant physical magnitudes to feed all the process. These datasets are adapted to the input/output specification of each approximation, and later a normalization and/or noise addition process is carried out, generating the definitive datasets. They are partitioned into three datasets, i.e., 60% for training ANNs, 20% for validation (overfitting prevention) and the last 20% for test the quality of learned model (prediction capability of the electrical behavior of the photovoltaic module).

Experimental design of systematic modeling procedure validation

This section is devoted to give a detailed description of the experimental design that we have followed to validate our systematic modeling procedure and get accurate photovoltaic module models, giving the specifications of

the photovoltaic module that has been used, the description of the devices used to obtain measurements of all the relevant magnitudes, and much more detailed specification of the systematic procedure that has been followed to train and validate the different ANN based approximations. Subsection Photovoltaic module description describes the main characteristics of the photovoltaic module whose electrical behavior is going to be modeled, while the remaining subsections from Data acquisition procedure to Detailed specification of the systematic modeling procedure are intended as a detailed description of the Algorithm 1.

Photovoltaic module description

In this subsection we introduce the characteristics of the photovoltaic module which is going to be used in the experimental part of the paper to obtain different ANN based approximations. The module has been manufactured by Mitsubishi Electric, the type is PV-TD185MF5 (185 Wp) and8 pieces can be seen in Fig. 3 in the University College of Engineering of Vitoria-Gasteiz (University of the Basque Country, Spain). This module has 50 series connected polycrystalline cells, and the key specifications are shown in Table 1. The performance of solar cells is normally evaluated under the standard test condition (STC), where

Table 1 e Mitsubishi PV-TD185MF5 photovoltaic module characteristics.

Attribute	Value
Manufacturer	Mitsubishi
Model	PV-TD185MF5
Cell type	Polycrystallin Silicon
Size [mm]	156 x 156
Number of cells in series	50
Maximum power [W]	185
Open circuit voltage V_{OC} [V]	30.60

Short circuit current I_{SC} [A]		8.13
Voltage, max power V_{MPP} [V]		24.40
Current, max power I_{MPP} [A]		7.58
Nominal operation cell temp.	[o C]	47.5

an average solar spectrum at AM 1.5 is used, the irradiance is normalized to 1000 W/m^2, and the cell temperature is taken as 25 ^{o}C.

Data acquisition procedure

In this subsection we describe all the details related to the step 1 of the Algorithm 1. The electrical setup and the set of devices needed to obtain the value of all the magnitudes involved in the ANN based approximations design are explained. This procedure has been followed from August 2013 to February 2015, generating nearly 63,000 samples along 18 months, i.e., almost 3500 samples per month. Each sampling period lasted 10 min and the relevant magnitudes were measured while a variable resistance was manually changed.

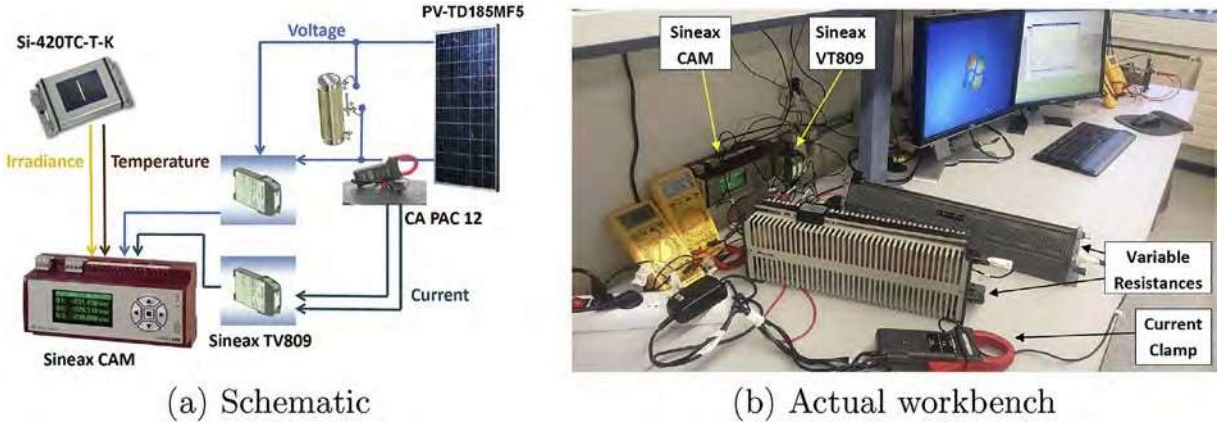
General setup

Fig. 4 shows two versions of the overall setup: on the other hand, Fig. 4(a) shows a schematic diagram with all the devices involved to capture the four relevant magnitudes, i.e., V_{PH}, I_{PH}, temperature T and irradiance G in the surroundings of the photovoltaic module. The voltmeter is placed in parallel with the module to obtain V_{PH} while the amperemeter in series to obtain I_{PH}. Besides, there is a variable resistance to act as a variable load and obtain different pairs of voltage and current with the same irradiance and temperature. The variable resistance value is controlled according to our convenience, but the temperature and the irradiance depends on the climatological conditions. On the other hand, in Fig. 4(b) there can be seen the actual workbench.

Irradiance and temperature sensor Si-420 TC-T-K

This sensor provides the irradiance (W/m^2) and the tempera- ture (^{o}C) of the surroundings of the photovoltaic module along the duration of the data acquisition process, which usually is not long. Once the sensor acquires a stable working temper- ature, both irradiance and temperature values have usually a very small variation, therefore they are assumed as constant by manufacturers when provide the IV and PV characteristic curves of their photovoltaic modules. The sensor provides its output expressed as current.

With regard to the accuracy of the device, it is ±5% when it measures irradiance and ±1.5 ºC when temperature.



(a) Schematic (b) Actual workbench

Fig. 4 e General setup and measurement elements for data logging.

Current clamp Chauvin Arnoux PAC12

The current clamp is used to measure the direct current pro- vided by the photovoltaic module, providing at its output a voltage proportional to the measured current. This clamp has two working scales: one with a narrow input range (0.4e60 A DC) but more accurate readings (precision ±1.5%), and a second one with a broader input range (0.5e600 A DC) but with less accurate readings (precision ±2%). Since the expected value of I_{PV} is less than 10 A, we have chosen the first working mode, which leads to measurements with a maximum error of ±150 mA.

Programmable isolating amplifier Camille Bauer Sineax TV809 The functions of this device are to isolate electrically the input/output signals and amplifying and/or converting the signal level or type (current to voltage, or vice versa) of the input DC signals. In the experiments described in this paper we have used two independent devices. Since the value of both V_{PV} and the output of the current clamps I_{PV} are expressed in volts, and the used datalogger only can read currents at its input, it is mandatory to use two independent devices to transform these magnitudes into current values.

The configuration of the transformations to carry out is done through the TV800plus (V1.11) software, and the main function of both amplifiers is to convert the voltage and cur- rent signals of the photovoltaic module into a 4e20 mA cur- rent signal proportional to the input. The screen to configure the conversion of the voltage supplied by the current clamp into current by one of the amplifiers is shown in Fig. 5.

The accuracy of this device is ±0.2% of the maximum value of the input. In the case of V_{PV} the maximum value is set to 50 V leading to maximum error of ±0.1 V, while the maximum value of I_{PV} is set to 10 A, leading to maximum error of ±20 mA.

Datalogger Camille Bauer datalogger Sineax CAM

This device is devoted to collect and register the four physical magnitudes that later will be used to elaborate the datasets to obtain different approximations as realistic as possible, i.e., irradiance, temperature, V_{PH} and I_{PH}. It is designed for long- term measurements in industrial installations, and it allows continuous measurement and recording of data. In our case the logging frequency was set to 0.5 Hz.

The datalogger is supplied with the CB-Manager software, which integrates several useful functions to configure the I/O interface according to specific requirements and to display the measured values. In Table 2 the used conversions from cur- rents to physical magnitudes are shown, while Fig. 6 shows the configuration screen of the I/O port number 1.

The accuracy of this device is ±0.1% of the maximum value of the input configuration (full scale) which is set to 20 mA, so the measurement error is negligible.

Modification of the datasets

A this point we recall the general systematic procedure described by Algorithm 3.5. After capturing data of the rele- vant magnitudes (step 1 of Algorithm 1) by means of the

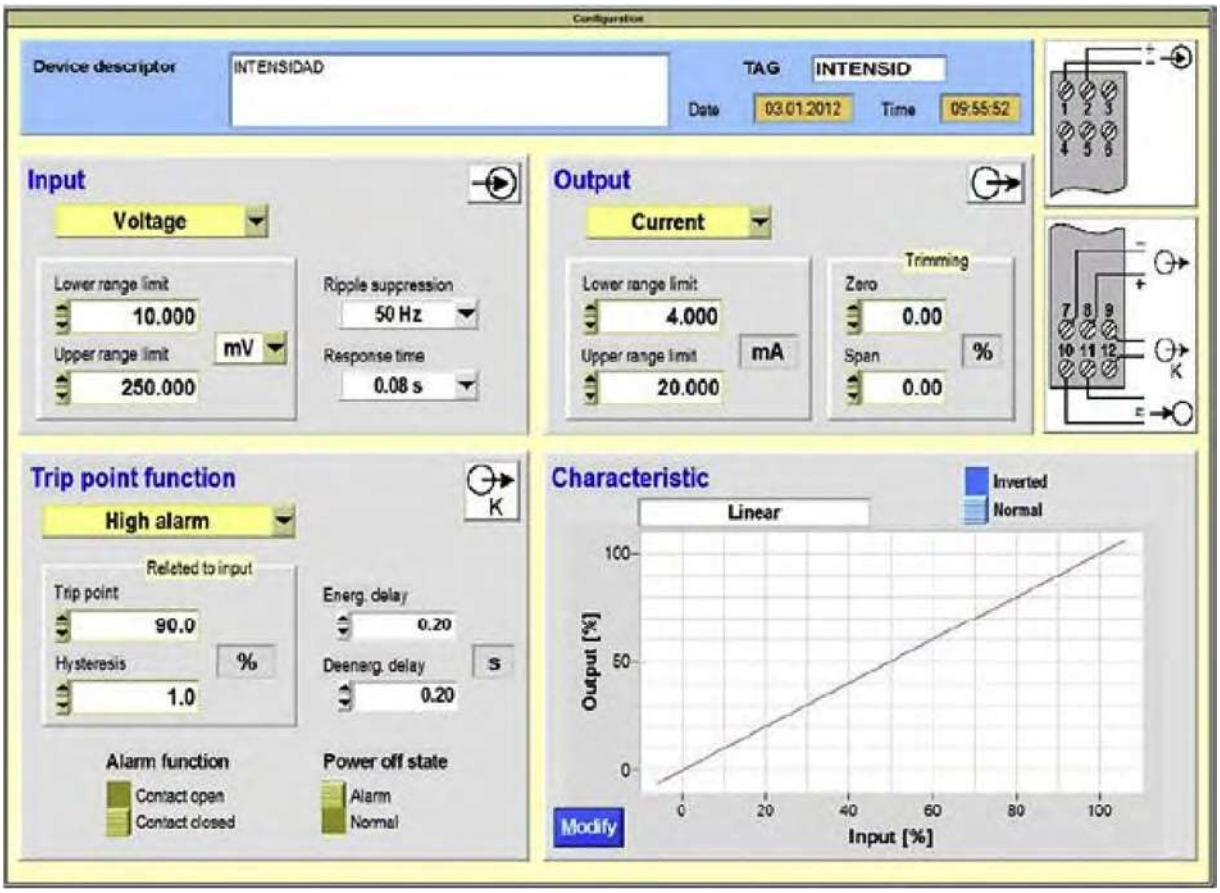


Fig. 5 e VT800Plus software to configure the TV809 Amplifier.

Table 2 ∈ Conversions from current to physical magnitudes.

	Input channel	
Magnitude	4 mA	20 mA
Irradiance G	0 W/m^2	1200 W/m^2
Temperature T	−135.5 °C	76 °C
Voltage V_{PH}	0 V	25 V
Current I_{PH}	0 A	25 A

potentially there are two versions of each training and testing dataset: raw unnormalized and normalized datasets.

- The noise addition process, if applied, we can obtain several datasets versions by the addition of noise to each input attribute based on uniformly distributed pseudo- random numbers in the range $r \in [-1, 1]$ weighted by a parameter $noisew \in [0, 100]$, as indicated in Eq. (3):

$$attribute \leftarrow attribute \cdot \{1 + [noisew \cdot (2 \cdot r - 1)]\} \quad (3)$$

electrical setup described in subsection Data acquisition procedure, a complete dataset is obtained containing an exhaustive sampling of the relevant magnitudes when approximating the electrical behavior of the photovoltaic module. The step 2(a) is particular for each approximation described in Section ANN based photovoltaic module modeling systematic procedure obtaining different datasets suited for each approximation, while in this subsection we are dealing with step 2(b) and step 2(c) of the Algorithm 1.

Regarding step 2(b), there are two modification trans- formations that could be carried out on input patters: a normalization and a noise addition process. We explore the effect of all four possible combinations.

- The normalization of the input attributes of the training and the test datasets sets them into the range $[-1, 1]$. So,

This modification rule is intended for generating additive or subtractive noise of magnitude *noisew* percent of the value of each input attribute. We have used the following values: $noisew \in \{0, 1, 2, 5, 10\}$, i.e., when $noisew = 0$ then the original value of the attributes remain unchanged. No more modifi- cations have been done, i.e., the raw data captured by all the measuring and conversion devices described in subsection Data acquisition procedure have not been filtered looking for outliers nor anomalous values in order to keep the systematic procedure as automatized as possible. This in turns implies that in some cases there are more than one different value of I_{PH} for the same combination of irradiance G, temperature T and V_{PH}; i.e., we are learning a multivalued function.

With regard to step 2(c) of Algorithm 1, we state that the partition of the modified dataset into train, validation and test datasets has been arbitrary. We have partitioned the input

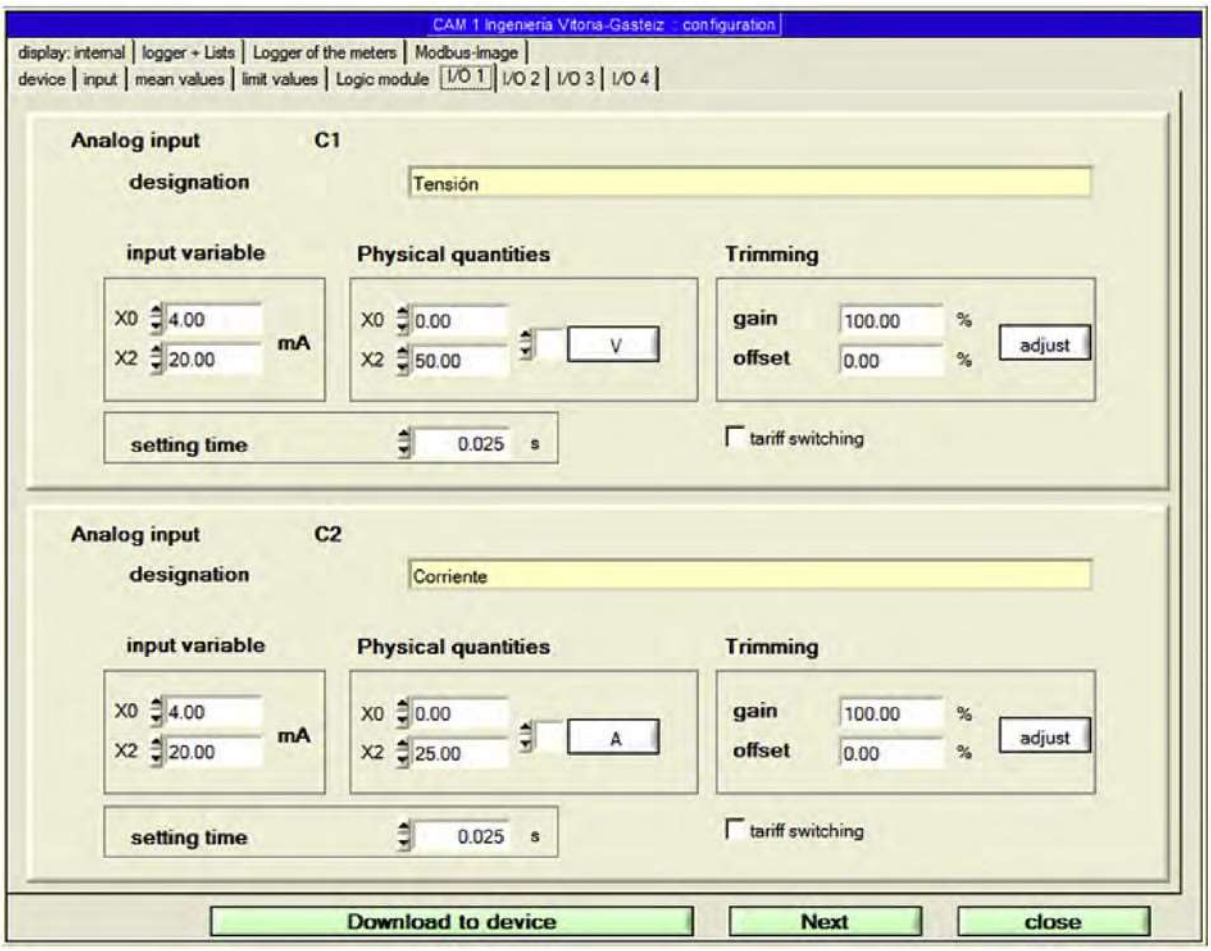


Fig. 6 e CB-Manager software to configure the Sineax CAM datalogger.

and target vectors (accordingly to each approximation defi- nition) using interleaved indices as follows: 60% are used for training, 20% are used to validate that the network is gener- alizing and to stop training before overfitting, and finally, the last 20% are used as a completely independent test of network generalization.

ANN structure and configuration parameters

The first step in the training process to configure an ANN is to fix its structure. We have used feed-forward ANNs with one and two hidden layers, being this one the main structural difference between them.

We complete the configuration of the ANNs with the following description:

- Activation function: We have tested ANNs with different activation functions in the hidden nodes, depending on the number of hidden layers:
  - One hidden layer: tan-sigmoid, log-sigmoid and purelin functions.
  - Two hidden layers: tan-sigmoid, log-sigmoid and pure- lin functions in the first hidden layer, while only purelin in the second one.
- Hidden layers size: We have tested ANNs with different number of hidden nodes in each one of the hidden layers. We have used near to linear fashion spaced natural values depending on the number of hidden layers: Eq. (4) defines the range of the number of nodes when the ANNs have only one hidden layer (200 values), while Eq. (6) shows the range for the first and the second hidden layer when the ANNs have two hidden layers (a maximum of 36 values for each hidden layer):

$$H_1 = \{h_1 \; h_1 \in \mathbb{N} \wedge h_1 \in [1;\ 200]\} \tag{4}$$

$$H = \{h \; h \in \mathbb{N} \wedge h \in [1;\ 20]\} \cup \cup_{p=5}^{20} p^{\infty} \tag{5}$$

$$h_1 \in H \quad \text{if 1st layer} \quad h_2 \in H \ \text{if 2st layer} \wedge h_2 \leq h_1 \tag{6}$$

- Trials: For the resulting combinations of adopting the variants of number of hidden layers, activation function and neurons number previously exposed, we have per- formed 5 trials measuring several performance values in order to analyze the behavior of individual ANNs and also the behavior of all ANNs of a given structure.

Algorithm 2 Detailed specification of the systematic procedure for photovoltaic module modeling

for each formulation of the model $m_j \in M$
  for each normalization possibility $r_q \in R$
    for each noise percentage $n_r \in N$
      Load training and testing datasets with:
        - inputs according to the model m_j
        - outputs according to the model m_j
        - normalization according to r_q
        - noise n_r percent
      for each number of hidden nodes of the 1^{st} hidden layer
        -$h_{1p} \in H_1$ (if only 1 hidden layer), or
        -$h_{1p} \in H_2$ (only if 2 hidden layers)
        for each number of hidden nodes of the 2^{nd} hidden layer
          - $h_{2u} \in H_2$ (only if 2 hidden layers)
          for each activation function $f_t \in F$
            for each initialization $i \in [1, 5]$
              Create an ANN suited to the model m_j with:
                - inputs according to the model m_j
                - outputs according to the model m_j
                - h_{1p} hidden nodes in the 1^{st} hidden layer
                - h_{2u} hidden nodes in the 2^{nd} hidden layer
                  (only if 2 hidden layers)
                - f_t as activation function
              Train the $i-th$ ANN using the training dataset,
                Obtain accuracy individual values
              Test the $i-th$ ANN using the testing dataset
                Obtain accuracy individual values
            Obtain accuracy train/test values for all initializations

Detailed specification of the systematic modeling procedure

The procedure for all model approximations training and validation is specified in Algorithm 2. It is a detailed instance of the general systematic procedure of the Algorithm 1. We define the following complete sets to formalize the algo- rithmic description:

- M ¼ $\{M_1, M_2, M_3, M_4\}$ is the set of the model approximation formulations with which we will address the problem of photovoltaic electric model learning, and M ¼ $\cup_j m_j$, being m_j each model approximation.
- R ¼ $\{true, false\}$ is the set of the possibilities regarding the normalization of the attributes (physical magnitudes of the datasets).
- N ¼ $\{0, 1, 2, 5, 10\}$ is the percentage of noise gain added to each input attribute, and N ¼ $\cup_r n_r$, being n_r the noise per- centage possibilities.
- H_1 defined by Eq. (4) is the set of number of hidden nodes of the hidden layer if the ANN has only one hidden layer, while H_2 defined by Eq. (6) is the set of the of number of hidden nodes of the hidden layers is the ANN has two hidden layers.
- F ¼ $\{tan - sigmoid, log - sigmoid, purelin\}$ is the set of acti- vation functions, and F ¼ $\cup_t f_t$, being f_t the available acti- vation function possibilities for the 1st hidden layer.

Finally, regrading the training algorithm, we have chosen the LevenbergeMarquardt algorithm due to speed reasons, despite of being very memory consuming. In all cases, all the ANN input vectors are presented once per iteration in a batch.

paper, i.e., neither normalization nor noise addition processes were applied to the datasets, nor different ANN structures were tested. In that work two different models were obtained, reporting for both of them a Root Mean Square Error (RMSE) of

0.1 A. It is an accurate result, but both models were required because each of them was suited only for a month, i.e., one of them was trained with a dataset collected in January 2014 while the other one with a dataset collected in July 2014, being this circumstance the main drawback of the reported results of [22] because this approach is appropriate only if the assumption that the temperature and irradiance are con- stants is accepted.

In the first row of Tables 3 and 4 there are the most accu- rate ANNs for this approach with one and two hidden layers respectively. The $Perf_{train}$, $Perf_{val}$ and $Perf_{test}$ columns show the RMSE value with the train, validation and test datasets respectively. The best results were reached normalizing and adding noise to the original datasets, however, as we could expect, the RMSE value for the test dataset $Perf_{test}$ reports a poor accuracy. It is due to the approach assumes that tem- perature and irradiance are quite constants, however, the entire dataset of 18 months has a very wide range of tem- peratures from 3.65 º C to 66.54 º C, and from 33.33 W/m^2 to 1267.2 W/m^2 of irradiances.

Approximation by M_2 ¼ *(*TG – $V_{PH}I_{PH}$*) and*
M_3 ¼ *(*TG – V_{PH}, TG – I_{PH}*) models*

In the second row of Tables 3 and 4 there are the most accu- rate ANNs for the approach M_2 ¼ (TG – $V_{PH}I_{PH}$) with one and two hidden layers respectively, while in the third and the fourth rows of the same tables there are the best results of the

Experimental results

This section is devoted to discuss the experimental results that we have obtained following the systematic procedure and the experimental design of Section Experimental design of systematic modeling procedure validation, and to compare them with previous relevant works of the literature. All the model approaches are discussed, but we pay more attention to the last model M_4 ¼ (TGV_{PH} – I_{PH}) in the last subsection. These experimental results have been obtained after intensive exe- cutions in which 32 workstations (Dell Precision T1700 equipped with Intel(R) Core(TM) i7-4770 CPU@3.40 Ghz and 16 GB RAM) were working during 7 days.

Approximation by M_1 ¼ *(*V_{PH} – I_{PH}*) model*

This model was used with success in Ref. [22] using heuristic methods instead of a systematic procedure as proposed in this

two ANNs which implements the model under the approach M_3 ¼ (TG – V_{PH}, TG – I_{PH}). The results of both approaches are definitely very poor: in general, the RMSE value of the ANNs which implements the approach M_2 is the mean of the two corresponding ANNs which implements the approach M_3. Moreover, these separated ANNs of this last approach do not take advantage of having separated targets. We think that the cause of this problem is that the cartesian product of the input values is not a sufficient domain to learn such complex multivariate functions, but actually we have not proved this hypothesis, and we need to make further work to analyze the poor performance of the trained models under these approaches.

Approximation by M_4 ¼ *(*TGV_{PH} – I_{PH}*) model*

At this point we recall the main idea which drove this approach specification, i.e, not to assume any of the relevant

Table 3 e Best ANN with 1 hidden layer for each approach and their results (RMSE) for the 18 months dataset.

	Norm	Noise	Nodes	Funct	$Perf_{train}$	$Perf_{val}$	$Perf_{test}$
M_1	Yes	5	106	*logsig*	2.121	2.223	1.923
M_2	No	0	196	*logsig*	29.547	33.337	31.609
M_3 *V*	Yes	0	174	*tansig*	60.854	65.200	58.662
M_3 *I*	Yes	0	174	*tansig*	1.998	2.173	2.032
M_4	Yes	0	64	*logsig*	0.047	0.052	0.045

Table 4 e Best ANN with 2 hidden layers for each approach and their results (RMSE) for the 18 months dataset (*Nodes* 1/2 means the nodes of the 1*st* and the 2*nd* hidden layer).

	Norm	Noise	Nodes 1/2	Funct	$Perf_{train}$	$Perf_{val}$	$Perf_{test}$
M_1	Yes	2	20/16	*tansig*	2.176	2.162	1.879
M_2	Yes	1	50/12	*logsig*	36.252	38.245	33.574
M_3 *V*	No	1	45/3	*tansig*	70.808	73.065	65.989
M_3 *I*	No	1	50/8	*tansig*	2.430	2.464	2.262
M_4	No	0	90/6	*tansig*	0.043	0.051	0.042

magnitudes as constant to achieve a model as accuracy as possible.

In the last row of Tables 3 and 4 there are the most accurate ANNs with one and two hidden layers respectively for this last approach when modeling the behavior of the photovoltaic module through data captured along 18 months. The reached accuracy for test datasets $Perf_{test}$ shows a RMSE of 0.045 A and 0.42 A for ANNs with one and two hidden layers respectively. They are very accurate results as expected due to the char- acteristics of this approach. Taking into account that some of measuring equipment has a limited precision (among other devices, for example, the current clamp described in subsec- tion Data acquisition procedure has a maximum error of ±150 mA as stated in that subsection), we can conclude that the unattended execution of the systematic procedure intro- duced in this paper has got an ANN model which has learned the behavior of the

photovoltaic module with an accuracy higher than the measurement precision.

The correlation coefficients (R-value) of the best ANNs are shown in Fig. 7. Both Fig. 7(a) for the best one hidden layer ANN and Fig. 7(b) for the best two hidden layers ANN show a very good correlation coefficient taking into account that R ¼ 1 means perfect correlation between the network response and targets, and the test dataset is composed of more than 12,600 samples (20% of 63,000 samples).

Table 5 and Table 6 give us deeper insight into the learning process of this approach. In the left part of the tables there are the values of the best ANN with a combination regarding the normalization and noise addition processes, while in the right part there are the best mean structures for all the trials with regard to the same combination. Analyzing these tables we can obtain several conclusions, considering that they are ob- tained under the double effect of normalization and noise addition:

- In the case of ANNs with one hidden layer, there are several ANNs with very similar accuracy values. In fact, there are two ANNs with the same RMSE value, and the second one (with normalized dataset) has been selected because it has less nodes in the hidden layer.
- The normalization process in combination with null or low percentage noise addition results in ANNs with two hidden layers with a very similar RMSE value.
- In general, the normalization process carried out with datasets generates smaller ANNs than those generated with unnormalized datasets.
- The normalization process in the case of ANNs with one hidden layer leads to a smaller RMSE; however, in the case of two hidden layers the effect is the opposite, leading to larger RMSE values.

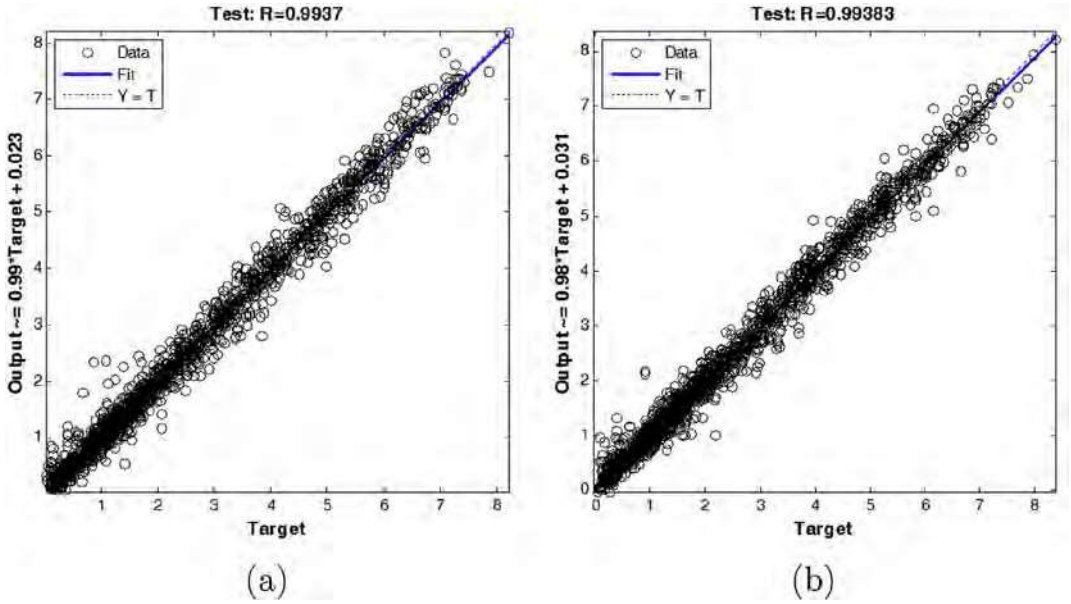


Fig. 7 e Correlation coefficients of the (a) best 1 hidden layer and (b) 2 hidden layers ANNs, for approach M_4 ¼ (TGV_{PH} ¡ I_{PH}) with the 18 months dataset.

Table 5 e Test accuracy (*RMSE*) results for approach M_4 ¼ (TGV_{PH} ¡ I_{PH}), merging the results obtained with all combinations of hidden nodes and activation functions with ANN with 1 hidden layer.

		Best individual			Best mean ANN			
Norm	Noise	Acc.	ANN Nodes	Funct	*Acc* :	$s^{10}-3$	Nodes	Funct
No	0	0.045	129	*logsig*	0.050	6.0	129	*logsig*
	1	0.049	47	*tansig*	0.057	2.8	138	*tansig*
	2	0.055	96	*logsig*	0.061	2.5	66	*logsig*
	5	0.101	148	*logsig*	0.108	6.6	148	*logsig*
	10	0.205	112	*logsig*	0.214	4.2	74	*logsig*
Yes	0	0.045	64	*logsig*	0.051	3.1	94	*logsig*
	1	0.046	117	*logsig*	0.0	3.	15	*tansig*

					52	0	4	
	2	0.049	189	*logsig*	0.0	1.	17	*logsig*
					51	8	4	
	5	0.060	113	*tansig*	0.0	0.	90	*logsig*
					64	6		
	10	0.086	86	*tansig*	0.0	1.	69	*logsig*
					94	6		

Table 6 e Test accuracy (*RMSE*) results for approach M_4 ¼ (TGV_{PH} ¡ I_{PH}), merging the results obtained with all combinations of hidden nodes and activation functions with ANN with 2 hidden layers (*Nodes* 1/2 means the nodes of the 1*st* and the 2*nd* hidden layer).

		Best individual ANN			Best mean ANN			
Norm	Noise	Acc.	Nodes 1/2	Funct	*Acc* :	s$10-3$	Nodes 1/2	Funct
No	0	0.042	90/6	*tansig*	0.0	3.5	70/15	*tansig*
					49			
	1	0.047	95/30	*tansig*	0.0	2.9	95/20	*logsig*
					53			
	2	0.053	65/12	*tansig*	0.0	4.7	70/9	*tansig*
					62			
	5	0.102	100/60	*logsig*	0.1	2.4	100/45	*tansig*
					08			
	10	0.199	65/14	*tansig*	0.2	1.1	18/5	*tansig*
					13			
Yes	0	0.044	55/4	*tansig*	0.0	2.5	100/40	*logsig*
					48			
	1	0.045	80/6	*logsig*	0.0	3.1	55/17	*logsig*
					49			
	2	0.046	65/15	*tansig*	0.0	0.3	80/30	*tansig*
					52			
	5	0.058	70/35	*tansig*	0.0	3.0	35/8	*tansig*
					63			
	10	0.083	80/9	*tansig*	0.0	7.4	80/12	*tansig*
					91			

- As we can see, for both one and two hidden layer ANNs the best results are always obtained without noise addition. This circumstance takes place probably because the measuring devices

introduce noise in the measurements, as stated in the section devoted to their description.

- The addition of noise always generates larger RMSE values.
- The second hidden layer is always much smaller than the first one.
- The best activation functions for this approach are *tansig*

Comparison with previous works

This subsection is devoted to compare the obtained results with previous relevant works of the literature. Table 8 con- tains a summary of the results reported by the most relevant and recent publications of the state-of-art about PV modules modeling. The methodology and the purpose of most of them were recalled in Section Introduction, so only the most

and *logsig*. The remaining activation functions have not been used in hidden neurons to get the best ANN under any combination regarding normalization and noise addition.

As last relevant information with regard to the best one and two hidden layers ANNs trained with the 18 months dataset, Table 7 shows the space requirements needed to store both ANNs. The first one (one hidden layer, Table 7(a)) needs less than 3 KBytes of space to store its configuration, while the second one (two hidden layers, Table 7(b)) uses less than 8 KBytes. Comparing these re- quirements with the space needed to store a tabular representa- tion of the original dataset calculated in Eq. (7), the saving is very relevant, reaching more than 99.84% of space with one hidden layer ANN and more than 99.58% of space with two hidden layers:

$$\underline{values}\,\delta T;\ G;\ \underline{V_{PH}};\ \underline{I_{PH}}\text{Þ}\quad Bytes$$

$$63;000\ samples \times 4$$

Table 7 e Space requirements of the (a) best 1 hidden layer and (b) 2 hidden layers ANNs, for approach $\_M_4 \text{¼} (TGV_{PH}\ \text{¡}\ I_{PH})$ with the 18 months dataset.

	Matrix	Layer	Size	Size (Bytes)
(a)				
1 Hidden layer	Bias	Hidden	64	512
		Output	1	8
	Weight	Input Hidden	64 x 3	15 36
		Hidden	64	512
	TOTAL			25 68
(b)				

2 Hidden layers	Bias	Hidden 1	90	720
		Hidden 2	6	48
		Output	1	8
	Weight	Input Hidden 1	90 x 3	2160
		Hidden 2	6 x 90	4320
		Hidden	6	48
	TOTAL			7304

relevant comments on the results will be given in this sub- section. The meaning of the columns of Table 8 is the following:

- Type: It is the type of PV cells of the modules which have been reported as used in the theoretical or empirical ex- periments of each reference. (poly: polycrystalline, mono: monocrystalline, thin: thin-film, all: the model could be applied to any type of cell because it has an empirical basis and does not depend on any special parameter or charac- teristic of the technology of the cell).
- Model: It indicates whether the model is theoretically or empirically obtained.
- *P*: It is the power of the PV module (watt).
- Mag: It is the magnitude or variable predicted by the model.
- Metric: It indicates the metric used to measure the accu- racy obtained by the model, and it evaluates the difference between the predicted and the expected/real magnitude (RMSE: root mean square error, MSE: mean square error, ME: mean error, R: correlation coefficient, Diff: difference)
- Value: It is the value of the metric when applied to the magnitude.

Before analyzing and comparing the results of previous relevant works, we point out several circumstances related to our general systematic procedure and the experimental setup:

- The PV module that we have used in the real experiments has much more power (185 W) than the PV modules of the previous relevant works (12.4 W, 42 W, 50 W, 55 W, 90 W, 120 W, 125 W, 128 W and 130 W). This makes the absolute value of the metric to be larger while the relative error could be smaller.
- The dataset of our experiments is composed of 63,000 samples gathered during 18 months. It is much more larger than the datasets of the previous relevant works. This makes potentially much more difficult the task of learning a model because there are much more points to learn and to fit.

- The model to be learned by our general procedure is a two inputs model, while some model of the previous works has three inputs. This makes to these works more easy the learning task because they have more information.
- The model generated by our general procedure is unique for the entire dataset, i.e., it is not splitted into two

  submodels as some of the previous works. This makes to these works more easy the learning task because they can specialize each model in a segment of the dataset.

At this point we proceed to compare the models obtained in previous relevant works with the model generated by our general systematic procedure:

- In Ref. [26] a procedure to obtain the model of only thin- film modules is introduced, and apart from its results, the results of the models of a number of different authors are reported. The predicted magnitude is I_{PV}, and accepting roughly the comparison between *Diff* and RMSE, we can see clearly that only if the power of the PV module is much lower than ours then the accuracy is better. Even in this circumstance, in most cases the model generated by our procedure has a better accuracy.
- In Ref. [28] an evolutionary algorithm is used to generate

  humidity, while the output is the power. Again the magnitude predicted by the model is P_{PV} instead of I_{PV} and the metric is ME instead of RMSE, which make that the results are not comparable. However, the R coefficient is comparable and it is clearly worse than ours.

Finally, we recall the work of [22] in order to compare the results obtained in that paper with the new ones from using the systematic procedure introduced in this paper with this modeling approach, but with those two datasets (one with measurements of January 2014 and the other one of July 2014). In Ref. [22] the best test accuracy was obtained with a RMSE of
0.1 A with both datasets. Tables 9 and 10 show that applying the systematic procedure, in general, the RMSE is divided by a factor of 10^3, reaching much better results than those which were already accurate.

the (theoretical) model and the results of a number of

works are reported. In this case the magnitude and the metric used to evaluate the accuracy of all models is the same as ours, and we can conclude clearly that our model is more accurate that all of they, even with a PV module of 185 W instead of only 120 W.

- In Ref. [8] authors do not report the type of used real PV module. They use RFB neural networks to generate two models. The former (ANN_1) is to obtain I_{PV} from irradiance and voltage, while the second one (ANN_2) is devoted to obtain P_{PV} from the same magnitudes. Authors report nu- merical results, but actually the test has been done with training data, so the results are not comparable with

ours because we have reported results with test data. Moreover, the second magnitude nor the metric is the same of ours.

- In Ref. [25] two different models based on artificial neural networks were generated from a real PV module. Both models have temperature and irradiance as inputs while power as output, but they are specialized in cloudy and sunny days respectively. This technique allows to obtain two more specialized models instead of only one, and the predicted magnitude by the model is P_{PV} instead of I_{PV}. So the results are not comparable with ours because we have generated only one global model and the magnitude to predict is different.
- In Ref. [35] two neuronal models from experimental data were built, each one for a different module. The inputs of the networks are temperature, irradiance and relative

Conclusions and future work

In this paper we have discussed one main issue when dealing with photovoltaic module electrical behavior learning to generate accurate global (not partial) models: the lack of a systematic procedure to obtain these accurate models in an unattended way, i.e, with zero human intervention. After giving a background on ideal photovoltaic cells and ANNs, our approach to face this problem has been to propose and describe four models to formulate the approximation to the electrical behavior of a photovoltaic module by ANNs, i.e., M_1 ¼ (V_{PH} – I_{PH}), M_2 ¼ (TG – $V_{PH}I_{PH}$), M_3 ¼ (TG – V_{PH}, TG – I_{PH})
and M_4 ¼ (TGV_{PH} – I_{PH}), and to introduce a general systematic procedure to train and validate the ANN models of these ap- proaches. In order to test such general systematic procedure we have explained in detail the experimental design that we have used, introducing the photovoltaic module (Mitsubishi Electric PV-TD185MF5, 185 Wp) of which an accurate model has been obtained as proof of concept of the general system- atic procedure, the measuring devices of the relevant magnitudes and the extended description of the general systematic procedure. As a result of applying such procedure to a large dataset of 63,000 samples collected during 18 months, two very accurate models ($Perf_{test}$ of one hidden layer ANN with RMSE of 0.045 A, and $Perf_{test}$ of two hidden layers ANN with RMSE of 0.042 A, in both cases less than the measuring devices precision) were obtained. Besides, it has been shown that the

Table 9 e Best 1 hidden layer ANNs for approach M_4 ¼ (TGV_{PH} ï I_{PH}) and their results (RMSE) for each 1 month dataset.

Norm	Noise	1st	Funct	$Perf$ 10−3	$Perf$ 10−3	$Perf$ 10−3

					train	*val*	*test*
January 2014	No	0	9	*tansig*	0.4319	0.3357	0.2359
July 2014	No	0	23	*tansig*	0.3531	1.0929	0.4034

Table 10 e Best 2 hidden layers ANNs for approach M_4 ¼ (TGV_{PH} ¡ I_{PH}) and their results (RMSE) for each 1 month dataset.

	Norm	Noise	1st	2st	Funct	$Perf 10^{-3}$ *train*	$Perf 10^{-3}$ *val*	$Perf 10^{-3}$ *test*
January 2014	No	0	7	4	*tansig*	0.3882	0.6122	0.1971
July 2014	No	0	18	3	*logsig*	0.3731	0.6501	0.3662

proposed procedure is competitive and improves the state-of- art results where comparable.

Future work will consist in analyzing the poor results ob- tained by approaches M_2 and M_3, paying special attention to the presence of outlayers which could condition the smoothness of the multivalued function to learn in both cases.

Acknowledgments

The authors are grateful to the Basque Government (IE101- 279/ETORTEK10/21) by the support of this work through the EUROZONA project (ETORTEK 2010).

references

[1] Almonacid F, Hontoria Leocadio, Aguilera Jorge, Nofuentes G. Una nueva aproximacion para la caracterizacion de modulos fotovoltaicos basada en redes reuronales. Bol del Inst Estud Giennenses 2005;192:287e310.

[2] Almonacid F, Rus C, Hontoria L, Munoz FJ. Characterisation of {PV} {CIS} module by artificial neural networks. a comparative study with other methods. Renew Energy 2010;35(5). ISSN: 0960-1481:973e80. http://dx.doi.org/10.1016/ j.renene.2009.11.018. URL, http://www.sciencedirect.com/science/article/pii/S0960148109004947.

[3] Atlam Ozcan. An experimental and modelling study of a photovoltaic/proton-exchange membrane electrolyser system. Int J Hydrogen Energy 2009;34(16). ISSN: 0360- 3199:6589e95. http://dx.doi.org/10.1016/ j.ijhydene.2009.05.147. URL, http://www.sciencedirect.com/science/article/pii/S0360319909008969. 4th Dubrovnik Conference4th Dubrovnik Conference.

[4] Bandou Farida, Hadj Arab Amar, Said Belkaid Mohammed, Logerais Pierre-Olivier, Riou Olivier, Charki Abderafi. Evaluation performance of photovoltaic modules after a long time operation in saharan environment. Int J Hydrogen Energy 2015;40(39). ISSN: 0360-3199:13839e48. http://dx.doi.org/10.1016/j.ijhydene.2015.04.091. URL, http://www.sciencedirect.com/science/article/pii/S0360319915010034.

[5] Bastidas-Rodriguez JD, Petrone G, Ramos-Paja CA, Spagnuolo G. A genetic algorithm for identifying the single diode model parameters of a photovoltaic panel. Math Comput Simul 2015 ISSN: 0378-4754. http://dx.doi.org/ 10.1016/j.matcom.2015.10.008. URL, http://www.sciencedirect.com/science/article/pii/S0378475415002220.

[6] Bayrak Gokay, Cebeci Mehmet. Grid connected fuel cell and

{PV} hybrid power generating system design with matlab simulink. Int J Hydrogen Energy 2014;39(16). ISSN: 0360- 3199:8803e12. http://dx.doi.org/10.1016/ j.ijhydene.2013.12.029. URL, http://www.sciencedirect.com/ science/article/pii/S0360319913029625.

[7] Bensmail Samia, Rekioua Djamila, Azzi Halim. Study of hybrid photovoltaic/fuel cell system for stand-alone applications. Int J Hydrogen Energy 2015;40(39). ISSN: 0360- 3199:13820e6. http://dx.doi.org/10.1016/ j.ijhydene.2015.04.013. URL, http://www.sciencedirect.com/ science/article/pii/S0360319915008538.

[8] Bonanno F, Capizzi G, Graditi G, Napoli C, Tina GM. A radial basis function neural network based approach for the electrical characteristics estimation of a photovoltaic module. Appl Energy 2012;97. ISSN: 0306-2619:956e61. http://
dx.doi.org/10.1016/j.apenergy.2011.12.085. URL, http://www.
sciencedirect.com/science/article/pii/S0306261911008919. Energy Solutions for a Sustainable World - Proceedings of the Third International Conference on Applied Energy, May 16- 18, 2011- Perugia, Italy.

[9] Chester M. Neural networks. New Jersey: Prentice Hall; 1993.

[10] Das Abhik Kumar. An explicit j-v model of a solar cell using equivalent rational function form for simple estimation of maximum power point voltage. Sol Energy 2013;98. ISSN: 0038-092X:400e3. http://dx.doi.org/10.1016/ j.solener.2013.09.023. URL, http://www.sciencedirect.com/ science/article/pii/S0038092X1300385X. Part C.

[11] Elbaset Adel A, Ali Hamdi, Sattar Montaser Abd-El. Novel seven-parameter model for photovoltaic modules. Sol Energy Mater Sol Cells 2014;130. ISSN: 0927-0248:442e55. http://dx.doi.org/10.1016/j.solmat.2014.07.016. http://www.
sciencedirect.com/science/article/pii/S0927024814003778.

[12] Gong Wenyin, Cai Zhihua. Parameter extraction of solar cell models using repaired adaptive differential evolution. Sol Energy 2013;94. ISSN: 0038-092X:209e20. http://dx.doi.org/
10.1016/j.solener.2013.05.007. URL, http://www.sciencedirect.
com/science/article/pii/S0038092X13001904.

[13] Gonzaez-Longatt F. Model of photovoltaic in matlabtm. In: 2do Congreso Iberoamericano de Estudiantes de Ingenierı´a Electrica. Puerto la CruzeVenezuela: Electronica y Computacion (II CIBELEC 2005); 2006.

[14] Gow JA, Manning CD. Development of a photovoltaic array model for use in power-electronics simulation studies. Electr Power Appl IEE Proc 1999;146(2). ISSN: 1350-2352:193e200. http://dx.doi.org/10.1049/ip-epa:19990116.

[15] Gupta S, Tiwari H, Fozdar M, Chandna V. Development of a two diode model for photovoltaic modules suitable for use in simulation studies. In: Power and energy engineering conference (APPEEC), 2012 Asia-Pacific; March 2012. p. 1e4.
http://dx.doi.org/10.1109/APPEEC.2012.6307201.

[16] Ramos Hernanz JA, Lopez Guede JM, Zamora Belver I, Eguia Lopez P, Zulueta E, Barambones O,

et al. Modelling of a photovoltaic panel based on their actual measurements. Int J Tech Phys Probl Eng (IJTPE) 2014;6(4):37e41.

[17] Ishaque Kashif, Salam Zainal, Taheri Hamed. Simple, fast and accurate two-diode model for photovoltaic modules. Sol Energy Mater Sol Cells 2011;95(2). ISSN: 0927-0248:586e94. http://dx.doi.org/10.1016/j.solmat.2010.09.023. URL, http:// www.sciencedirect.com/science/article/pii/ S0927024810005477.

[18] Ishaque Kashif, Salam Zainal, Mekhilef Saad, Shamsudin Amir. Parameter extraction of solar photovoltaic modules using penalty-based differential evolution. Appl Energy 2012;99. ISSN: 0306-2619:297e308. http://dx.doi.org/ 10.1016/j.apenergy.2012.05.017. URL, http://www. sciencedirect.com/science/article/pii/S0306261912003650.

[19] Jiang Lian Lian, Maskell Douglas L, Patra Jagdish C. Parameter estimation of solar cells and modules using an improved adaptive differential evolution algorithm. Appl Energy 2013;112. ISSN: 0306-2619:185e93. http://dx.doi.org/10.1016/ j.apenergy.2013.06.004. URL, http://www.sciencedirect.com/ science/article/pii/S0306261913005114.

[20] Kalogirou Soteris, Sencan Arzu. Artificial intelligence techniques in solar energy applications. INTECH Open Access Publisher; 2010.

[21] Kriesel D. A brief introduction to neural networks. 2007. URL available at: http://www.dkriesel.com.

[22] Lopez-Guede Jose Manuel, Ramos Hernanz Jose Antonio, Guerrero Ekaitz Zulueta, Fernandez-Gamiz Unai. Towards a systematic neural network based modelization of photovoltaic panels. In: Kurt Erol, editor. The proceedings of third European conference on renewable energy systems-ECRES 2015; 7e10 October 2015.

[23] Luque A, Hegedus S. Handbook of photovoltaic science and engineering. John Wiley & Sons Ltd.; 2003.

[24] Poulsen NK, Norgaard M, Ravn O, Hansen LK. Neural networks for modelling and control of dynamic systems. A practitioners handbook. Number 1 in Advanced Textbooks in Control and Signal Processing. London: Springer-Verlag; 2000.

[25] Mellit A, Saglam S, Kalogirou SA. Artificial neural network- based model for estimating the produced power of a photovoltaic module. Renew Energy 2013;60. ISSN: 0960- 1481:71e8. http://dx.doi.org/10.1016/j.renene.2013.04.011. URL, http://www.sciencedirect.com/science/article/pii/ S0960148113002279.

[26] Miceli Rosario, Orioli Aldo, Gangi Alessandra Di. A procedure to calculate the i-v characteristics of thin-film photovoltaic modules using an explicit rational form. Appl Energy 2015;155. ISSN: 0306-2619:613e28. http://dx.doi.org/10.1016/ j.apenergy.2015.06.037. URL, http://www.sciencedirect.com/ science/article/pii/S0306261915007953.

[27] Molina MG, Espejo EJ. Modeling and simulation of grid- connected photovoltaic energy conversion systems. Int J Hydrogen Energy 2014;39(16). ISSN: 0360-3199:8702e7. http://

dx.doi.org/10.1016/j.ijhydene.2013.12.048. URL, http://www.sciencedirect.com/science/article/pii/S0360319913029820.

[28] Muhsen Dhiaa Halboot, Ghazali Abu Bakar, Khatib Tamer, Abed Issa Ahmed. Parameters extraction of double diode photovoltaic module model based on hybrid evolutionary algorithm. Energy Convers Manag 2015;105. ISSN: 0196- 8904:552e61. http://dx.doi.org/10.1016/j.enconman.2015.08.023. URL, http://www.sciencedirect.com/science/article/pii/S0196890415007694.

[29] Narendra KS, Parthasarathy K. Identification and control of dynamical systems using neural networks. Neural Netw IEEE Trans Mar 1990;1(1). ISSN: 1045-9227:4e27. http://dx.doi.org/10.1109/72.80202.

[30] Oi Akihiro. Design and simulation of photovoltaic water pumping system. PhD thesis. Citeseer; 2005.

[31] Thuijsman F, Braspenning PJ, Weijters AJMM. Artificial neuronal networks. An introduction to ANN theory an practice. volume 931 of Lecture Notes in Computer Science. Berlin Heidelberg: Springer-Verlag; 1995. http://dx.doi.org/ 10.1007/BFb0027019.

[32] Ramos-Hernanz JA, Campayo JJ, Larranaga J, Zulueta E, Barambones O, Motrico J, et al. Two photovoltaic cell simulation models in matlab/simulink. Int J Tech Phys Probl Engineering(IJTPE) 2012;4(1):45e51.

[33] De Soto W, Klein SA, Beckman WA. Improvement and validation of a model for photovoltaic array performance. Sol Energy 2006;80(1). ISSN: 0038-092X:78e88. http://dx.doi.org/10.1016/j.solener.2005.06.010. URL, http://www.sciencedirect.com/science/article/pii/S0038092X05002410.

[34] Tsai Huan-Liang, Tu Ci-Siang, Su Yi-Jie. Development of generalized photovoltaic model using matlab/simulink. In: Proceedings of the world congress on engineering and computer science, vol. 2008. Citeseer; 2008. p. 1e6.

[35] Velilla Esteban, Valencia Jaime, Jaramillo Franklin. Performance evaluation of two solar photovoltaic technologies under atmospheric exposure using artificial neural network models. Sol Energy 2014;107. ISSN: 0038- 092X:260e71. http://dx.doi.org/10.1016/j.solener.2014.04.033. URL, http://www.sciencedirect.com/science/article/pii/ S0038092X14002400.

[36] Villalva MG, Gazoli JR, Filho ER. Modeling and circuit-based simulation of photovoltaic arrays. In: Power electronics conference, 2009; 2009. p. 1244e54. http://dx.doi.org/10.1109/COBEP.2009.5347680. COBEP '09. Brazilian.

[37] Walker GR. Evaluating MPPT converter topologies using a MATLAB PV model. J Electr Electron Eng Aust 2001;21(No.1):49e56.

[38] Widrow B, Lehr MA. 30 years of adaptive neural networks: perceptron, madaline, and back propagation. Proc IEEE Sep 1990;78(9). ISSN: 0018-9219:1415e42. http://dx.doi.org/10.1109/5.58323.