United Nations Educational Scientific and Cultural Organization
and
International Atomic Energy Agency

THE ABDUS SALAM INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

**AN ADAPTIVE ARTIFICIAL NEURAL NETWORK MODEL
FOR SIZING STAND-ALONE PHOTOVOLTAIC SYSTEMS:
APPLICATION FOR ISOLATED SITES IN ALGERIA**

A. Mellit
*University Center of Medea, Institute of Engineering Sciences,
Ain Dahab, 26000, Algeria,*

M. Benghanem[1]
*Faculty of Electrical Engineering, University of Sciences and Technology Houari
Boumediene (USTHB), P:32, El-Alia, Bab-Ezzouar, 16111 Algiers, Algeria
and
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy,*

A. Hadj Arab
*Development Center of Renewable Energy (CDER),
P.O. Box 62, Bouzaréah, Algiers 16000, Algeria*

and

A. Guessoum
*Ministry for the Higher Education and Scientific Research, Algiers, Algeria.*

MIRAMARE – TRIESTE

July 2004

[1] Regular Associate of the Abdus Salam ICTP.
Corresponding author: m.benghanem@caramail.com; mbenghan@ictp.trieste.it

**Abstract**

In this paper we investigate, by using an adaptive Artificial Neural Network (ANN), in order to find a suitable model for sizing Stand-Alone Photovoltaic (SAPV) systems, based on a minimum of input data. This model combines Radial Basis Function (RBF) network and Infinite Impulse Response (IIR) filter in order to accelerate the convergence of the network. For the sizing of a photovoltaic (PV) system, we need to determine the optimal sizing coefficients ($K_{PV}$, $K_B$). These coefficients allow us to determine the number of solar panels and storage batteries necessary to satisfy a given consumption, especially in isolated sites where the global solar radiation data is not always available and which are considered the most important parameters for sizing a PV system. Obtained results by classical models (analytical, numerical, analytical-numerical, B-spline function) and new models like feed-forward (MLP), radial basis function (RBF), MLP-IIR and RBF-IIR have been compared with experimental sizing coefficients in order to illustrate the accuracy of the results of the new developed model. This model has been trained by using 200 known optimal sizing coefficients corresponding to 200 locations in Algeria.

In this way, the adaptive model was trained to accept and even handle a number of unusual cases, the unknown validation sizing coefficients set produced very set accurate estimation and a correlation coefficient of 98% was obtained between the calculated and that estimated by the RBF-IIR model. This result indicates that the proposed method can be successfully used for the estimation of optimal sizing coefficients of SAPV systems for any locations in Algeria, but the methodology can be generalized using different locations over the world.

**Nomenclature**

| | |
|---|---|
| $a_i$ | the IIR filter coefficients |
| $bj$ | the recursive filter coefficients |
| ANN | artificial neural network |
| $A_{PV}$ | PV array area (m²) |
| $B_k(t)$ | B-spline function |
| $C_A, C_S$ | sizing coefficients |
| $C_{AOP}, C_{SOP}$ | optimal sizing coefficients for numerical model |
| $C_U$ | useful accumulator capacity (Wh) |
| $E$ | energy function |
| $E_{AUX}$ | auxiliary generator (Wh/day) |
| $f, f_1, f_2$ | sizing parameters for the analytical-numerical model |
| $Fs_1, Fs_2$ | arbitrary factors |
| $Gi$ | energy of linear inflection |
| $h$ | hidden layer |
| $H$ | solar radiation data (Wh/m²/day) |
| IIR | infinite impulse response filter |
| $K_{PV}, K_B$ | optimal sizing coefficients for developed model |
| $L$ | average daily energy consumption (Wh/day) |
| LLP | loss of load probability |
| LMS | linear mean square |
| $m_{ij}$ | vector average of the hidden neurons $i$ |
| $M$ | number of feed-forward delays |
| MLP | multi-layer perceptrons |
| MTM | Markov transition matrices |
| $N_f$ | the number of feedback |
| $N_j$ | the number of days |
| $N_{k,k}$ | basis elements of B-spline function |
| PV | photovoltaic |
| RBF | radial basis function |
| SAPV | stand-alone photovoltaic |
| SOC | state of charge |
| $u$ | the input coefficient to the neural network model |
| $r, r_1, r_2$ | sizing parameters for the analytical-numerical model |
| v | the co-input coefficient to the neural network model |
| $w_i$ | the synaptic weight |
| $y_1, y_2$ | the output parameters for developed model |
| $\gamma_m$ | the free parameter of analytic model |
| $\alpha_1, \alpha_2, \beta, \lambda_i$ | B-spline function coefficients |
| $\eta_{PV}$ | PV array efficiency |
| $\sigma^2$ | the variance from the hidden neuron |

## 1. Introduction

Photovoltaic (PV) applications may offer a promising alternative especially in remote areas as isolated small power generation for essential electric power. All around the world there is a number of small isolated communities, like saharian sites in Algeria and rural villages without access to a large electricity grid. Furthermore, in many places due to the remoteness and due to the cost, it is unlikely that main grid connection will be ever established. However, the need for power still exists. Power systems which can generate and supply electricity to such remote locations are variously termed "decentralized, autonomous or stand-alone". The technology for power production from renewable energy sources are available and reliable so the penetration of the technology depends mainly on the economic feasibility and a proper sizing of the components in order to avoid outages as well as ensuring quality and continuity of supply. Several models have been developed, simulating and sizing PV systems using different operation strategies. Less attention has been given to the simulation of more complex hybrid energy systems incorporating a wider range of components.

The estimation of SAPV system sizing (number of solar cell panels and the size of the storage battery) is very useful to conceive an optimal and economic SAPV system. Several studies have been conducted and were interested in the performance of PV systems.

For instance, Evans *et al* [1] described a method to consider the monthly average output of PV fields; Gupta and Young [2] estimated the excess of energy provided by PV generators using the utilisability method developed by Liu and Jordan [3]. Siegel *et al.* [4] evaluated the monthly average output, the excess of energy and the storage of the batteries. The excess of energy provided by PV systems for an installation having a constant load was also evaluated by Klein [5], Clark et al.[6] using the average utilisability function. All these methods are based on the energy balance of the systems studied to determine their storage capacity and their outputs. Other recent methods estimate the performance of PV systems while being based on the Loss of Load Probability (LLP) technique, defined as the ratio between the energy deficit and the energy demand, both on the load [7].

$$LLP = \frac{\int_t Energy \quad deficit}{\int_t Energy \quad demand} \qquad (1)$$

Thus analytical and numerical methods presenting various algorithms for calculation of the LLP were published in the literature. Among the analytical methods, we can mention those developed by Macomber et al. [8]. Barra et al. [9], Bucciareli [10, 11] and Bartoli et al. [12]. These analytical methods are simple to apply but they are not general. On the other hand, the numerical methods presented by Graham et al. [13], Aguiar et *al*. [14] and Egido and Lorenzo [15], use more complex methods which allow the improvement of the precision of calculation of LLP according to the dimension of the PV-array area and the storage capacity [16, 17].

Virtually, all these methods allow us the estimation of the PV system sizing for a given site, and therefore requires the availability of several parameters such as global solar irradiation data, latitude, longitude, the load, the characteristics of the PV system and the inclination of the panels and, unfortunately, an important computing time for the estimation of the pair $(C_A, C_S)$ for the PV system. The global solar radiation data is considered the most important parameter for an optimal sizing $(K_{PV}, K_B)$ of the PV system, indeed these data are not always available over all areas, specially in isolated sites (remote area, rural zone), these methods are not suitable for the sizing of the PV systems in these sites. In fact, the aim of this study is to present a new technique for the estimation and the modeling of the optimal sizing coefficients particularly in isolated sites. We also investigate the suitability of an adaptive RBF network with IIR filter as a tool for the estimation of the optimal sizing coefficient for SAPV systems in order to improve the results obtained in [18, 19]. These coefficients allow the users of SAPV systems to determine the PV-array area and the storage capacity of the batteries necessary to satisfy a given consumption. The trained network could then be used as a designed tool for estimating the performance of PV systems.

The next section presents a description of SAPV systems, sizing database of optimal sizing coefficients for SAPV systems used in this simulation study and also present different classical models. Section 3 describes an adaptive neural network algorithm implementation. The developed model has been presented in section 4. Section 5 presents a comparison study between classical models and different structures of neural network,

5

also an experimental validation has been presented in this section. An example of PV systems sizing in isolated sites has been presented in the final section.

## 2. Sizing of Stand-Alone Photovoltaic System

### 2.1 Description and sizing of PV system

The size of a SAPV system (Fig.1) is a general concept which includes the sizing of PV-array and the accumulators. A useful definition of such dimensions relates to the load: in daily basis, the PV-array capacity, ($C_A$) is defined as the ratio between average PV array energy production and the average load energy demand. The storage capacity, ($C_S$) is defined as the maximum energy that can be taken out from accumulator divided by the average energy demand [15], so:

$$C_A = \frac{\eta_{PV} A_{PV} H}{L} \quad \text{and} \quad C_S = \frac{C_U}{L} \tag{2}$$

where $A_{PV}$ is the PV array area, $\eta_{PV}$ is the PV array efficiency, $H$ is the average daily irradiation on the PV array, $L$ is the average daily energy consumption, $C_S$ is the storage capacity and $C_U$ is the useful accumulator capacity. Note that $C_A$ depends on the meteorological conditions of the location. That means that the same PV array for the same load can be 'large' in one site and 'small' in another site with lower solar radiation.

The task of sizing a SAPV system consists of finding the better trade-off between cost and reliability. Very often, the reliability is a priori requirement from the user, and the PV engineer problem is formulated as follows: which pair of $C_A$ and $C_S$ values leads to a given LLP value at the minimum cost [15]?

*Empirical models*

The relation between $C_A$, $C_S$ and LLP is exclusively committed to the intuition of the PV sizing. The PV-array dimension is calculated to ensure that the generation during the worst month period exceeds the consumption by a security factor directly established by the PV seizer according to the type of application and his own experience. A similar

6

procedure is used to estimate the battery size. To resume, $C_A = Fs1$ and $C_S = Fs2$, where *Fs1* and *Fs2* are arbitrary factors [15].

## Numerical models

These methods allow the direct calculation of the $(C_A, C_S)$ pair, and in general require simulations of long series of meteorological data in order to obtain the iso-satisfaction curve. The numerical methods suppose that any system is fed by an auxiliary source when the system is missing, the procedure for this method is presented in [20]. Then, the state of the charge of the battery, at the end of the day $j$ is given by

$$SOC_j = \min(SOC_{j-1} + \frac{\eta_{PV} A_{PV} H}{C_U}; 1) \tag{3}$$

The auxiliary generator is managed in such a way that, at the end of the day $j$ it fulfils the battery if the stored energy is lower than the load requirements. Then,

$$SCC_j \geq C_S^{-1} \Rightarrow E_{AUX\,J} = 0 \tag{4}$$

$$SCC_j \leq C_S^{-1} \Rightarrow E_{AUX\,J} = (1 - SOC_j)\frac{L}{C_S} \quad \text{and} \quad SOC_J = 1 \tag{5}$$

where $E_{AUX\,J}$ is the energy supplied by the auxiliary generator in the day j. If the simulation is carried out over a large number of days, N, in order to be statistically meaningful, the LLP value corresponding to the SAPV systems is given by:

$$LLP = \frac{\sum_{j=1}^{N} E_{AUXj}}{N_j L} \tag{6}$$

## Analytical models

The analytical approaches such as those of Macomber *et al.* [9] and Bartoli *et al.* [12] are based on the assumption that the daily solar irradiation is an uncorrelated Gaussian variable having a mean of $H$ and a variance $\sigma$. Both parameters are used as inputs to the model. The main disadvantage of this method is that the use of the error function with an iterative process requires a very long computing time. According to Barra *et al.* [8] the fraction for the energy load converted by the PV system is expressed as:

$$Y = 1 - \frac{E_{AUX}}{L} \tag{7}$$

And they assume that as the monthly average, $Y_m$, relates with the size of the PV system through the formula

$$(C_A - Y_m)(1 - Y_m) = \gamma_m \tag{8}$$

where $\gamma_m$ is the free parameter. This equation represents a hyperbola whose asymptotes are the straight lines

$$Y_m = C_A \quad \text{and} \quad Y_m = 1 \tag{9}$$

In fact, this represents very logical limit conditions. The first means that, for small size PV systems, all the energy produced by the PV array is transferred to the load. The second comes from the consideration that, for very large field areas, the energy supplied by the array is always able to satisfy the load.

Bucciarelli *et al.* [10, 11] presents a model for the LLP derived by approximating the probability density function of the difference between the daily PV array output and the load with two events and by assuming that the daily storage charge/discharge process can be represented as a one-step Markov process.

*Analytic-numeric model*

A more recent study combining at the same time the analytical and numerical methods for the calculation of the LLP gave good results [15]; this combination gives a hybrid method. The relation binding the capacity of PV generator to the capacity of storage is given by the following expression:

$$C_A = f\, C_S^{-r} \tag{10}$$

where, f and r are the parameters which depend on the LLP by a simple regression of the type:

$$f = f_1 + f_2 Log(LLP) \quad \text{and} \quad r = \exp(r_1 + r_2.LLP) \tag{11}$$

where $f_1$, $r_1$, $f_2$ and $r_2$ are the sizing parameters of PV systems.

*B-Spline function model*

The normalized, uniform B-spline function of degree k, denoted as $B_k(t)$, is defined by:

$$B_k(t) = \begin{cases} N_{k,k}(t) & 0 \le t < 1 \\ N_{k-1,k}(t) & 1 \le t < 2 \\ \ldots\ldots & \ldots\ldots \\ N_{0,k}(t) & k \le t < k+1 \\ 0 & t \le 0 \quad or \quad t \ge k+1 \end{cases} \qquad (12)$$

Here the basis elements $N_{j,k}(t)$, $0 \le t \le 1$, are obtained recursively by the following algorithm [21]:

*Algorithm*

Let $N_{0,0}(t) = 1$ and for $i = 1, 2, \ldots, k$

$$\begin{cases} N_{0,j}(t) = \dfrac{1-t}{i} N_{0,i-1}(t) \\ N_{j,i}(t) = \dfrac{1-j+t}{i} N_{j-1,i-1}(t) + \dfrac{1+j-t}{i} N_{j,i-1}(t), \quad j = 1,\ldots,i-1 \\ N_{i,j}(t) = \dfrac{t}{i} N_{i-1,j-1}(t) \end{cases} \qquad (13)$$

Thus, $B_k(t)$ is piece-wise polynomial of degree $k$ with integer knot points and is $k$-$1$ times continuously differentiable. It is noted that $B_k(t)$ for k = 0,1, 2, …..is normalized as follows:

$$\sum_{j=0}^{k} N_{j,k}(t) = 1, \quad 0 \le t \le 1 \qquad (14)$$

And this yields:

$$\int_{-\infty}^{\infty} B_k(t)dt = \int_{0}^{k+1} B_k(t)dt = 1 \qquad (15)$$

For the sake of reference, B-splines of up degree are plotted in Fig. 2.

The principle of this model consists on establishing a spatial interpolation using B-spline function [22, 23], which makes it possible to draw up sizing PV systems maps based on optimal sizing coefficient ($C_{AOP}$, $C_{SOP}$) determined for 60 sites over Algeria and

geographical coordinates corresponding to these sites. In our case, we used as a tool the B-spline function with 2-Degree modeling.

That is to say, a whole of the points $P_i$ of coordinated $(x_i)$ there $i$ pertaining to the whole of realities, one associates to each point $P_i$ a value $G_i$ the energy of linear inflection of a flexible thin section which one forces to pass by the points of cordons $(x_i)$, $G_i$ out of $R^3$ is given by [24]:

$$E(G) = \iint\limits_{R\,R} (\Delta G)^2 \, dxdy \tag{16}$$

The principle of the B-spline with two degrees consists in seeking among the functions $G$, which take the values $G_i$ in points $P_i$ and minimizing the relation (11). According to [24] the only function which minimizes $E(G)$ is given by the following expression:

$$S(x, y) = \sum_{i=1}^{n} \lambda_i k_i(x, y) + \alpha_1 x + \alpha_2 y + \beta \tag{17}$$

where $k_i$, $\lambda_i$, $\alpha_1$, $\alpha_2$ and $\beta$ are the coefficients to be determined. The solution of this equation leads to the determination of the function $S(x, y)$. This makes it possible to estimate the parameter $G_i$ with the point $x_i$

This choice of B-spline function was motivated by the following reasons:

a) The aptitude of the B-spline function to describe complex variations

b) Few data at our disposition

c) Inhomogeneous concentration of these data

Used vectors for interpolation in this study are: the altitude, longitude and the sizing coefficient ($C_{AOP}$, $C_{SOP}$), each vector contains 60 values. The interpolation zone consists of a rectangle of 18° 5 ' with 37° 5 ' of latitude and -9° with 12° 3 ' of longitude, for example Fig.3 shows the sizing map for Algeria.

## 2.2. Constructed database of optimal sizing coefficient for SAPV system sizing

The sizing of the SAPV systems requires the knowledge of one of the components of solar radiation known as daily global irradiation measured by meteorological stations. Unfortunately, these data are available only for few weather stations over Algeria. Therefore, data were collected from a measurement system using a network

configuration [25], also from sunshine duration and temperature using an RBF network model [26] and the Markov Transition Matrices (MTM) approach [14]. As an example, the daily global irradiation for some sites is represented in Fig. 4.

We calculate the various pair ($C_A$, $C_S$) corresponding to 200 sites, using a numerical model [15]. Figure 5 presents the iso- reliability curves for some sites. Next, we calculate the optimal coefficient ($K_{PV}$, $K_B$) based on analytical cost (Table 1), where $K_{PV} = \dfrac{C_{AOP}}{\eta_{PV} H}$ and $K_B = C_{SOP}$. The coefficients $C_{AOP}$ and $C_{SOP}$ are the optimal pair. The obtained optimal sizing coefficient for some sites is given in Table 2, in this case database of optimal sizing coefficients is formed corresponding to 200 sites for standard load (L =1K Wh/day) and LLP=1%. According to this curve (Fig. 5), we note that the coefficient values obtained for the sites located in the north are higher than those obtained in the south. Fig. 6 shows the evolution of optimal sizing coefficients used in this study, from north to south of Algeria, the northern zones are more sky-covered than those in the south.


## 3. Adaptive Artificial Neural Network Algorithm

The concept of Artificial Neural Network (ANN) analysis was discovered nearly 50 years ago [27], but it is only the last 20 years that applications software have been developed to handle practical problems. The history and theory of neural networks have been described in a  large number of published literature and will not be covered in this paper except for a very brief overview of how neural network operates. ANNs are good for some tasks while lacking in some others. Specially, they are good for tasks involving incomplete data sets, fuzzy or incomplete information, and for highly complex and ill-defined problem, where humans usually decide on an intuitional basis. They can learn from examples, and are able to deal with non-linear problems. Furthermore they exhibit robustness and fault tolerance. The tasks that ANNs cannot handle effectively are those of high accuracy and precision as in logic and arithmetic. ANNs have been applied successfully in various fields: mathematics, engineering, medicine, economics, neurology and many others. Some of the most important ones are; in pattern, sound and speech reorganization, in analyzing of electromyography medical signatures, in the identification of military targets and in the identification of explosives in passenger suitcases. They

have also been used in weather forecasting and market trends, in the prediction of mineral exploration sites, in electrical and thermal load prediction, in adaptive and robotic and many others. Neural networks are used for process control because they can build predictive models of the process from multi-dimensional data collected from sensors. ANN models may be used as an alternative method in engineering analysis and prediction. They operate like a black box model, requiring no detailed information about the system. Instead, they learn the relationship between the input parameters and the controlled and uncontrolled variables by studying previously recorded data. ANN can be compared to a multiple regression analysis except that with ANN no assumptions need to be made about the system to be modeled. Neural networks usually perform successfully where other methods do not, and have been applied in solving a wide variety problems, including non-linear problems that are not well suited to classical methods of analysis. Another advantage is their ability to handle large and complex systems with many interrelated parameters. Although several network architectures and training algorithms are available [28], the Multi-Layer feed-forward neural network trained by the Back-Propagation (BP) method is so far one of the most popular. Each type of ANN exhibits its own architecture and learning algorithm. The ANN computation can be divided into two phases: learning phase and testing phase. The learning phase forms an iterative updating of the synoptic weights based upon the error BP algorithm. A schematic diagram of typical multi-layer feed-forward neural network architecture is shown in Fig. 6. The network usually deals with an input and some hidden and output layers. In its simple form, each single neuron is connected to other neurons of a previous layer through adaptable synaptic weights. The number of input and output parameters and the number of cases influence the geometry of the network.

A radial basis function network becomes very popular due to several important advantages over traditional Multi-Layer Perceptrons (MLP) [29, 30, 31, 32]:

a) Locality of radial basis function features extraction in hidden neurons that allows the use of clustering algorithms and independent tuning of RBF network parameters.

b) Sufficiency of one layer of non-linear elements for establishing arbitrary input-output mapping.

c) Solution of clustering problem can be performed independently from the weight in output layers.

d) RBF network output in scarcely trained areas of input space is not random, but depends on the density of the pairs in training data set.

These proprieties lead to potentially quicker learning in comparison to multi-layer perceptrons trained by Back-Propagation (BP). To some extent, RBF networks allow us to actualize a classical idea about training layer by layer. Before beginning a tracking operation using an adaptive neural network model (RBF-IIR), the unknown non-linear plan must be estimated according to the certain model. In this particular estimation process, the model consists of an adaptive neural network topology; the Radial Basis function embedded in the hidden unites. In cascade with the network is a local IIR block structure as shown in Fig. 9. The IIR synopsis network is used to create a double local network architecture that provides a computationally efficient method of training the system, and accordingly results in quick learning, and fast convergence [30]. The approximated signal of the network $\hat{y}(k)$ can be modeled by:

$$\hat{y}(k) = \sum_{i=0}^{M} a_i \ z(k-i)u(k) + \sum_{j=1}^{N_f} b_j \ \hat{y}(k-j)v(k) \tag{18}$$

where $M$ and $a_i$ are the number of feed-forward delays and coefficient of the IIR filter, respectively, $N_f$ and $bj$ are respectively the number of feedback and delays and recursive filter coefficients, respectively. The parameters $u$ and v are the input and the co-input to the model at the example $k$, respectively. The input $v\ (k)$ is usually kept small for the feed back stability purpose, $b$ is the bias value. The RBF network (Fig. 8) has the same structure as the MLP having only one hidden layer, the RBF is applied to the hidden layer [28,32, 33] it is chosen as being Gaussian defined by its average $m$ and its $\sigma^2$ variance, the output layer can be linear or non-linear function. The determination of the network parameters has the same procedure as the MLP; it is also a universal approximator. Therefore, a vector u having $i$ components $u_j$ formed the input layer of the RBF, then a hidden layer contained $h$ neurons and output layer; the expression of the output layer is given by:

$$z = \sum_{i=1}^{h} w_i \exp\left( \frac{\sum_{j=1}^{k} (u_j - m_{ij})^T (u_j - m_{ij})}{2\sigma_i^2} \right) \tag{19}$$

where $m_{ij}$ is the vector average of the hidden neurons $i$, $\sigma_i^2$ is the variance from the hidden neuron $i$ and $w_i$, is the synaptic weights. Determination of the parameters $m_{ij}$, $\sigma_i$, and $w_i$ is done by using the BP algorithm. The neural network parameters $w_i$, $m_{i,j}$, $\sigma_i$, $a_i$ and bi can be optimized in the LMS sense by minimizing the energy function E over the example. Thus: $e(k) = y(k) - \hat{y}(k)$.

The energy function is defined by:

$$E = \frac{1}{2} \sum_{i=1}^{N} e_i(k)^2 \tag{20}$$

To minimize E we may use the method of steepest descent which requires the gradients $\dfrac{\partial E}{\partial w_i}$, $\dfrac{\partial E}{\partial m_{i,j}}$, $\dfrac{\partial E}{\partial \sigma_i}$, $\dfrac{\partial E}{\partial a_i}$ and $\dfrac{\partial E}{\partial b_i}$ for updating the incremental changes to each particular parameter $w_{i,j}$, $a$ and $b$ respectively. Gradients of $E$ are:

$$\frac{\partial E}{\partial w_i} = -\sum_{j=1}^{N} e_j z(\|u_j - m_{i,j}\|^2) \tag{21}$$

$$\frac{\partial E}{\partial m_{i,j}} = -2w_i \sum_{j=1}^{N} e_j z'(\|u_j - m_{i,j}\|^2) \sigma_i (u_j - m_{i,j}) \tag{22}$$

$$\frac{\partial E}{\partial \sigma_i} = -w_i \sum_{j=1}^{N} e_j z'(\|u_j - m_{i,j}\|^2)(u_j - m_{i,j})(u_j - m_{i,j})^T \tag{23}$$

$$\frac{\partial E}{\partial a_i} = -\sum_{j}^{N} u(k)e(k)z(k-j) \tag{24}$$

$$\frac{\partial E}{\partial b_i} = -\sum_{j}^{N} v(k)e(k)\hat{y}(k-j) \tag{25}$$

The incremental changes of each parameter are simply the negative of their gradients, $\Delta w = -\dfrac{\partial E}{\partial w}$, $\Delta m = -\dfrac{\partial E}{\partial m}$, $\Delta \sigma = -\dfrac{\partial E}{\partial \sigma}$, $\Delta a = -\dfrac{\partial E}{\partial a}$ and $\Delta b = -\dfrac{\partial E}{\partial b}$. Thus each coefficient vector $w$, $m$, $\sigma$, $a$ and $b$ of the network is updated in accordance with the rule:

$$w_{i,j}(n+1) = w_{i,j}(n) + \mu_w \Delta w_{i,j}, \qquad\qquad m_{i,j}(n+1) = m_{i,j}(k) + 2\mu_m \Delta m_{i,j}$$

$$\sigma_i(n+1) = \sigma_i(n) + \mu_{\sigma_i} \Delta \sigma_i \qquad\qquad a_i(n+1) = a_i(n) + \mu\, a_i \Delta a_i$$

and $\quad b_i(n+1) = b_i(n) + \mu_b \Delta b_i$

where the subscripted $\mu$ values are fixed learning rate parameters.

## 4. Developed model

A total of 200 patterns has been calculated for optimal sizing coefficient ($K_{PV}$, $K_B$) as described above. From this set 180 patterns were used for the training (learning phase) of the network and 20 were used for testing (testing phase) of the model, these patterns have been randomly selected. The architecture that gave the best result, has two neurons in the input layer and two neurons in the output layer (Fig. 8). However, the number of neurons in the hidden layer must be adjusted during the learning phase, in order to train the network in an efficient manner. A developed model can generate the optimal sizing coefficients from only the geographical coordinates (latitude and longitude). These coefficients allow us to determine the PV-array area ($A_{PV}$) and the useful accumulator capacity ($C_U$) using the following equations:

$$A_{PV} = K_{PV}.\, L \quad \text{and} \quad C_U = K_B.\, L \tag{26}$$

These equations have been obtained from equation (2).

The diagram block of a developed model is represented in Fig. 9. Note that the input/output data are the altitude, the longitude, optimal PV Capacity ($K_{PV}$) and optimal storage capacity ($K_B$) corresponding respectively to $u_1(k)$, $u_2(k)$, $y_1(k)$ and $y_2(k)$.

## 5. Simulation results and application

Once a satisfactory degree of input-output mapping has been reached, the RBF-IIR network training is frozen and the set of completely is an unknown testing optimal sizing coefficient that was applied for validation. After the simulation of many different structures, we found that the best performance is obtained with one hidden layer with 8 neurons. Table 3 displays the confrontation of the statistical features (mean, variance and correlation coefficient) between the measured coefficients and those estimated by our model, it is found that there is no significant difference between the estimated and the

measured coefficient from a statistical point of view. The correlation coefficient obtained for the testing data set is 97.9% for $K_{PV}$ and 98.9% for $K_B$. In this respect, the closest to unity for values are the best estimation accuracy.

Table 4 shows an example of the results obtained after several simulation comparisons in terms of performance among different neural network structures. The performance of the model rises significantly as the number of hidden neurons is increased until 8 neurons. At this point, adding more hidden neurons to the networks results in a slight improvement in performance. The RBF-IIR model presents good results and takes less iteration compared to other neural network structures. Fig. 10 shows clearly that there is almost a complete agreement between the measured and estimated coefficients obtained by our model RBF-IIR, by contribution with the other neural networks.

In order to illustrate the importance of this model, we have made an experimental validation, for tree SAPV systems installed in different locations (Algiers, Tahifet and Ghardaia) in Algeria. Table 5 presents obtained optimal coefficient between classical and neural network models and Table 6 shows the experimental validation between the classical model and the new model. It is noticed that the numerical model has an accuracy value compared to the analytical model which has the simple advantage with to be applied, but the analytical-numerical model presents the more precise results compared to those obtained by the other models. These models require the availability of several parameters for which is applicable, on the other hand the model based on B-spline function gives results close to those obtained by the numerical model, since calculations of the optimal sizing coefficients were made containing this model. With regard to the model based on the neural networks (MLP, RBF and RBF-IIR) it shows an acceptable result and in particular the RBF-IIR model which gives closer results to reality. Although the database has been calculated while itself basing on the numerical model, the advantage of this model makes it possible to generate the PV array area and the useful accumulator capacity for battery for any site in Algeria and requires a minimum of data as an input of the model.

In this section, we presented an application of this model in order to illustrate how we can use it for determining the PV array area and their useful capacity. Firstly, we introduced the geographical coordinates for a given site (Latitude, Longitude) as an input of the model. Then, from the model we obtained the $K_{PV}$, and $K_B$, for a standard

consumption (1 KWh/day). Equation 26 allows us to calculate the $A_{PV}$ and the $C_U$. The number of solar modules and batteries are selected according to the unit dimension of the module and the storage capacity of the battery. Table 7 shows the results obtained for some sites from the north towards the south of Algeria, for sites not appearing in the database (isolated sites); the results have been obtained using a standard load.

For a load different of 1KWh/day we used equation 26, in order to determine the PV array area and the capacity of the batteries corresponding to a given consumption ($A_{PV}$ $=K_{PV} L/1000$ and $C_U =K_B L /1000$). However, a graphic abacus for the sizing of SAPV system has been developed (Fig. 10) corresponding to 10 isolated sites, from this graphic abacus, the users of PV systems can determine the PV array area and the useful accumulator capacity of the battery.

## 6. Conclusion

This article describes how it possible to model and estimate optimal sizing coefficients of SAPV system from a minimum of input data using the RBF-IIR model, once trained, the model estimates these coefficients faster. The validation of the model was performed with an unknown sizing coefficient, which the network has not seen before; the ability of the network to make acceptable estimations even in an unusual day is an advantage of the present method. The estimation with a correlation coefficient of 98 % was obtained. This accuracy is good within the acceptable level used by design engineers.

Classical models of sizing PV systems like empirical, analytical, numerical and analytical-numerical allow the estimation of the sizing of PV system but require the availability of several parameters such as the daily global irradiation data, geographical coordinates (altitude, longitude), the load, the characteristics of stand alone PV system, the inclination of the panels and besides take very much computing time for the estimation of the optimal coefficients. On the other hand, the model that we developed allows the estimation of the PV-array area and the storage capacity from a minimum input data (altitude, longitude) based on the optimal sizing coefficients and does not take very much time simulation. Note that the advantage of this model provides an estimation of the PV-array area and the storage capacity for any site in Algeria particularly in isolated areas, where the global solar radiation data is not always available. Also, this

model gave as good results compared to those given by classical models and other neural network architecture [18, 19].

The results have been obtained for Algerian sites, but the methodology can be applied to any geographical area in the world. Future work will include the investigation of the suitability of the Neuro-Fuzzy model for the sizing of photovoltaic systems in the world.

**Acknowledgments**

**References**

[1]   Evans, D.L. Simplified method for predicting photovoltaic array output. Solar Energy  1980; 27: 555-560.

[2]   Gupta, Y. and Young, S. Method of predicting long-term average performance of photovoltaic systems. Proceedings of systems simulation and economy, Analyze Conference, San Diego, California.1980.

[3]    Liu, B.Y.H., and Jordan, R.C. A rational procedure for predicting the long-term average performance flat-plate solar energy collectors. Solar Energy 1963; 7:53-74.

[4]   Siegel, M.D., Klein, S.A and Beckman, W.A A simplified method for estimating the monthly average of photovoltaic system.  Solar Energy 1981; 22: 413-418.

[5]    Klein, S.A. Calculation of flat-plate collector utilisability.   Solar Energy 1978; 21:325-330.

[6]   Clark, DR., Klein, S.A., and Backman, W.A A method for estimating the performance of photovoltaic system. Solar Energy 1984; 433: 551-555.

[7]   Klein, S.A., and Beckman, W.A. Loss of load probabilities for stand-alone PV. Solar Energy 1987; 39: 499-505.

[8]   Macomber, HL., Ruzek, JB., and Staff of Bird Engineering. Stand-alone photovoltaic system. Preliminary engineering, Design handbook, NAZA, CR165352, NASA Lewis  Research Center 1981.

[9]    Bara, L., Catalakoti, S., Fontana, F., and Lavorane, F.A analytical method to determine the optimal size the photovoltaic plant. Solar Energy 1984;  6: 509-514.

[10] Bucciarelli, L.L.Jr. Solar estimating loss-of-power probabilities of stand-alone photovoltaic solar energy system. Solar Energy 1984; 32; 205-211.

[11] Bucciarelli, L.L.Jr., The effect of day-to-day correlation in solar radiation on the probability of loss-power in stand-alone photovoltaic solar energy system.  Solar Energy 1986; 36:11-18.

[12] Bartoli, B., Cuomo, V., Fontana, F., Serio C., and Silverstrini, V. The design of photovoltaic plants: An optimization procedure. Applied Energy 1984; 18:37-41.

[13] Graham, V.A., Hollands, K.G.T, and Unny, TE., A times series model for Kt with application to global synthetic weather generation . Solar Energy 1988; 40: 83-92.

[14]  Aguiar, R.J., Collares-Pereira and Conde, J.P. Simple procedure for generating sequences of daily solar radiation values using a library of Markov transition Markov. Solar Energy  1988; 40:229-279.

[15] Egido, M., and Lorenzo, E., The sizing of sand-alone PV systems: A Review and a proposed new method. Solar Energy Materials and Solar Cells, 1992; 26: 51-69.

[16] Hadj Arab, A., Ait Driss, B., Amimeur, R., and Lorenzo, E., Photovoltaic systems for  Algeria.  Solar Energy, 1995; 54, 99-104.

[17]  Benghanem, M., and Maafi, A., A simplified method for estimating the performance of stand-alone photovoltaic systems. World Renewable Energy Congress VI, UK, 2000;  3: 1902-1905.

[18] Mellit, A., Benghanem, M., Hadj Arab, A., and Guessoum, A., Modeling of sizing the photovoltaic system parameters using artificial neural network, Proceedings of IEEE,  Conference on Control Application  2003. p. 353-357.

[19] Mellit, A., Benghanem, M., Hadj Arab, A., and Guessoum, A., Sizing the Stand-Alone photovoltaic systems using the neural networks adaptive model. Proceedings of 8th ArabInternational Solar Energy  Conference and the World Renewable Energy Congress, 2004.

[20] Hay, J. E., Mckay, D.C., Estimation solar irradiance on inclined surfaces: a review and assessement of methodologies. International Journal of Solar Energy 1985; 3, 203-240.

[21] Wang,L.-J., et al A Fast efficient computation of cubic-spline interpolation in image code, IEEE Trans. on signal processing , 2004 ; Vol 49, (6): 1189-1197.

[22] Mellit, A., Benghanem, M., Hadj Arab, A., and Guessoum, A., Spatial Modeling of the Optimal Sizing Stand-Alone Photovoltaic System Using Spline Function. Proceedings of 14th International Photovoltaic Science and Engineering Conference,  2004; p.114-115.

[23] Benghanem, M., An optimal sizing method for stand-alone photovoltaic system for Algeria. Proceedings of World Renewable Energy Congress VII, 2002.

[24] Duchon J., Fonctions splines du type plaque mince en dimension deux, (Ed. Université Scientifique et Médicale de Grenoble), N° 231, Séminaire sur l'analyse numérique, Grenoble, France, 1985.

[25] Benghanem, M. and Maafi, A., Data acquisition system for photovoltaic systems performance Monitoring. IEEE Trans. on Instrumentation and Measurement 1998; 1: 30- 33.

[26] Guessoum, A., Boubkeur, S. and. Maafi A., A global irradiation model using radial basis function neural network. World Renewable Energy Congress 1998; 1:332-336.

[27] Kalogirou, S.A., Artificial intelligence in renewable energy systems modeling and prediction, Proceedings of World Renewable Energy Congress VII 2002.

[28] Haykin, S. *Neural Networks*, *A comprehensive foundation*, Ed. Macmillan, 1994; New York.

[29] Kasabov, N., *Foundations of neural network fuzzy systems and knowledge engineering*. MIT. Press, 1994

[30] Leonard, J.A., Kramer, M.A., and Ungar, L.H., Using radial basis function to approximate the function and its error bounds. IEEE Trans. In Neural Networks. 1992; 4: 624-627

[31] Powell, M.J.D., *The theory of radial basis functions approximation, in advance of numerical analysis*, Ox. Charendon. 1992; Press p. 105-210.

[32] Ye X., and Loh N.K., Dynamic system identification using recurrent radial basis function network. Proceedings of American Control Conference, 1993;3, p. 2912

[33] Jordan, M.I, and Jacobs, R.A., Hierarchical mixtures of expert and the EM algorithm. Neural computation, 1994 ; 6 : 181-241.

| PV array | Solarex 50W, 94x50x5 cm |
|---|---|
| Cost | 17 US$ /PW |
| Maintenance cost | 0 US$/year |
| Life time | 10 year |
| Battery | 18Ah, 51x22x22.5 cm |
| Cost | 322 US$/KWh |
| Maintenance cost | 0 US$/year |
| Life cost | 5 year |

Table.1. Data for cost analysis

| Sites | | Optimal sizing coefficients LLP=1%, L=1 KWh/day | |
|---|---|---|---|
| Latitude (°) | Longitude (°) | $K_{PV}$ (Optimal PV array coefficient) | $K_B$ (Optimal battery coefficient) |
| 36 | 2 | 2.202 | 1.95 |
| 35 | 5 | 1.115 | 1.89 |
| 22 | 3 | 0.642 | 0.75 |
| 23 | 4 | 0.631 | 0.76 |

Table.2. Optimal sizing coefficient of stand-alone PV system

| Statistical tests | | | | |
|---|---|---|---|---|
| Optimal sizing Coefficient | Calculated Mean | Estimated Mean | Variance | Correlation coefficient (%) |
| $K_{PV}$ | 1.076 | 1.051 | 0.270 | 97.9 |
| $K_B$ | 1.135 | 2.112 | 0.226 | 98.9 |

Table.3. Comparison between actual and estimated results

| Network structure | | | Mean Square Error (MSE) | # of iterations | MSE for test set |
|---|---|---|---|---|---|
| MLP  Network | | | | | |
| # of input | # of input | # of input | | | |
| 2 | 4    4 | 2 | 0.095 | 640 | 0.087 |
| 2 | 6 | 2 | 0.074 | 935 | 0.060 |
| 2 | 10 | 2 | 0.087 | 1054 | 0.092 |
| RBF Network | | | | | |
| # of input | # of input | # of input | | | |
| 2 | 2 | 2 | 0.051 | 368 | 0.074 |
| 2 | 6 | 2 | 0.043 | 436 | 0.065 |
| 2 | 8 | 2 | 0.035 | 525 | 0.045 |
| MLP-IIR Model | | | | | |
| # of input | # of input | # of input | | | |
| 2 | 6 | 2 | 0.031 | 363 | 0.042 |
| 2 | 8 | 2 | 0.028 | 390 | 0.035 |
| RBF-IIR Model | | | | | |
| # of input | # of input | # of input | | | |
| 2 | 4 | 2 | 0.021 | 163 | 0.032 |
| 2 | 8 | 2 | 0.013 | 250 | 0.022 |

Table.4. Training results from each network structures

| | Site (Alger) | | | | Site (Tahifet) | | | | Site (Ghardaia) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Latitude (°) | Longitude (°) | Radiation (Wh/m²/day) | Load (KWh/day) | Latitude (°) | Longitude (°) | Radiation (Wh/m²/day) | Load (KWh/day) | Latitude (°) | Longitude (°) | Radiation (Wh/m²/day) | Load (KWh/day) |
| | 36.43 | 3.15 | 5.122 | 0.6 | 22 | 6 | 6.936 | 4 | 32.23 | 3.49 | 6.534 | 2 |

Sites

Photovoltaic array area and useful capacity

| | Site (Alger) | | Site (Tahifet) | | Site (Ghardaia) | |
|---|---|---|---|---|---|---|
| Experimental Results | $A_{PV}$ (m²) | $C_U$ (KWh) | $A_{PV}$ (m²) | $C_U$ (KWh) | $A_{PV}$ (m²) | $C_U$ (KWh) |
| | 1.50 | 1.8 | 6 | 3.5 | 4 | 2.5 |
| Simulated results | Optimal sizing coefficient | | | | | |
| Classical model | $K_{PV}$ | $K_B$ | $K_{PV}$ | $K_B$ | $K_{PV}$ | $K_B$ |
| Analytic | 2.453 | 2.781 | 1.435 | 0.812 | 1.893 | 1.196 |
| Numeric | 2.479 | 2.893 | 1.441 | 0.853 | 1.955 | 1.231 |
| Numeric-analytic | 2.543 | 2.931 | 1.453 | 0.871 | 1.958 | 1.239 |
| Spline function | 2.431 | 2.791 | 1.435 | 0.842 | 1.899 | 1.221 |
| NN model | | | | | | |
| MLP | 2.435 | 2.795 | 1.438 | 0.843 | 1.946 | 1.223 |
| RBF | 2.465 | 2.891 | 1.436 | 0.845 | 1.948 | 1.225 |
| MLP-IIR | 2.456 | 2.890 | 1.439 | 0.847 | 1.945 | 1.230 |
| RBF-IIR | 2.473 | 2.894 | 1.440 | 0.851 | 1.951 | 1.234 |

Table.5. Obtained optimal coefficient by classical and neural network model

|  | Classical model | | | | Neural network model | | | |
|---|---|---|---|---|---|---|---|---|
|  | Analytic | Numeric | Numeric-analytic | Spline function | MLP | RBF | MLP-IIR | RBF-IIR |
| $A_{PV}$ Er (%) | 1.9160 | 0.8471 | 1.7200 | 2.8383 | 2.6694 | 1.4199 | 1.7915 | 1.0918 |
| $C_U$ Er(%) | 7.8749 | 3.6986 | 2 .3541 | 7.4884 | 7.3345 | 3.7703 | 3.8062 | 3.6628 |

| Site (Tahifet) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
|  | Classical model | | | | Neural network model | | | |
|  | Analytic | Numeric | Hybrid | Spline function | MLP | RBF | MLP-IIR | RBF-IIR |
| $A_{PV}$ Er (%) | 4.3333 | 3.9333 | 3.1333 | 4.3333 | 4.1333 | 4.2667 | 4.0667 | 4.0000 |
| $C_U$ Er(%) | 7.7586 | 2.5791 | 0.4592 | 3.9192 | 3.7960 | 3.5503 | 3.3058 | 2.8202 |

| Site (Ghardaia) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
|  | Classical model | | | | Neural network model | | | |
|  | Analytic | Numeric | Hybrid | Spline function | MLP | RBF | MLP-IIR | RBF-IIR |
| $A_{PV}$ Er (%) | 5.6524 | 2.3018 | 2.1450 | 5.3186 | 2.7749 | 2.6694 | 2.8278 | 2.5115 |
| $C_U$ Er(%) | 4.5151 | 1.5435 | 0.8878 | 2.3751 | 2.2077 | 2.0408 | 1.6260 | 1.2966 |

Table.6. Comparison between experimental and estimated results by different model

| Sites | | | LLP=1%, L=1 KWh/Day | | | |
|---|---|---|---|---|---|---|
| Latitude (°) | Longitude (°) | $K_{PV}$ | PV-array Area $A_{PV}$ (m²) | $K_B$ | Useful accumulator capacity $C_U$ (KWh) | |
| 36 | 0 | 1.92 | 1.92 00 | 1.74 | 1.7400 | |
| 35 | 8 | 2.59 | 2.5900 | 2.46 | 2.4600 | |
| 34 | 2 | 1.98 | 1.9800 | 1.85 | 1.8500 | |
| 33 | -1 | 1.25 | 1.2500 | 1.52 | 1.5200 | |
| 32 | 9 | 0.96 | 0.9600 | 1.31 | 1.3100 | |
| 31 | -4 | 0.95 | 0.9500 | 1.29 | 1.2900 | |
| 30 | -3 | 0.90 | 0.9000 | 0.97 | 0.9700 | |
| 29 | 5 | 0.87 | 0.8700 | 0.86 | 0.8600 | |
| 27 | 10 | 0.78 | 0.7800 | 0.78 | 0.8300 | |
| 25 | -2 | 0.77 | 0.7700 | 0.76 | 0.7800 | |

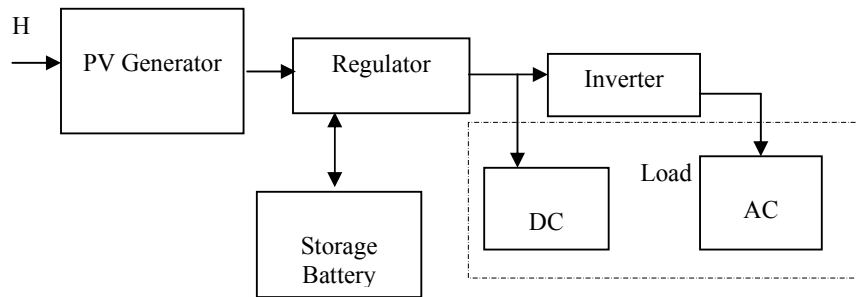Table.7. Sizing example for stand-alone PV system for isolated sites

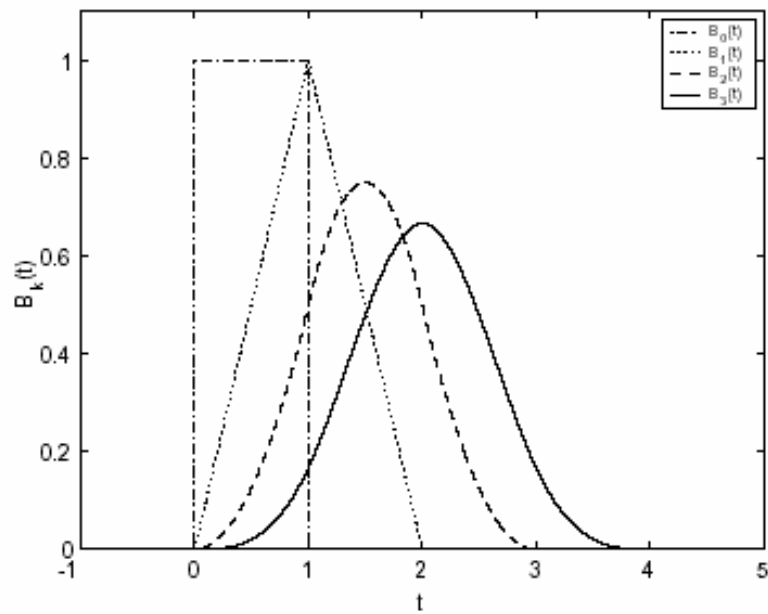Fig.1. Diagram block of simplified stand-alone PV system



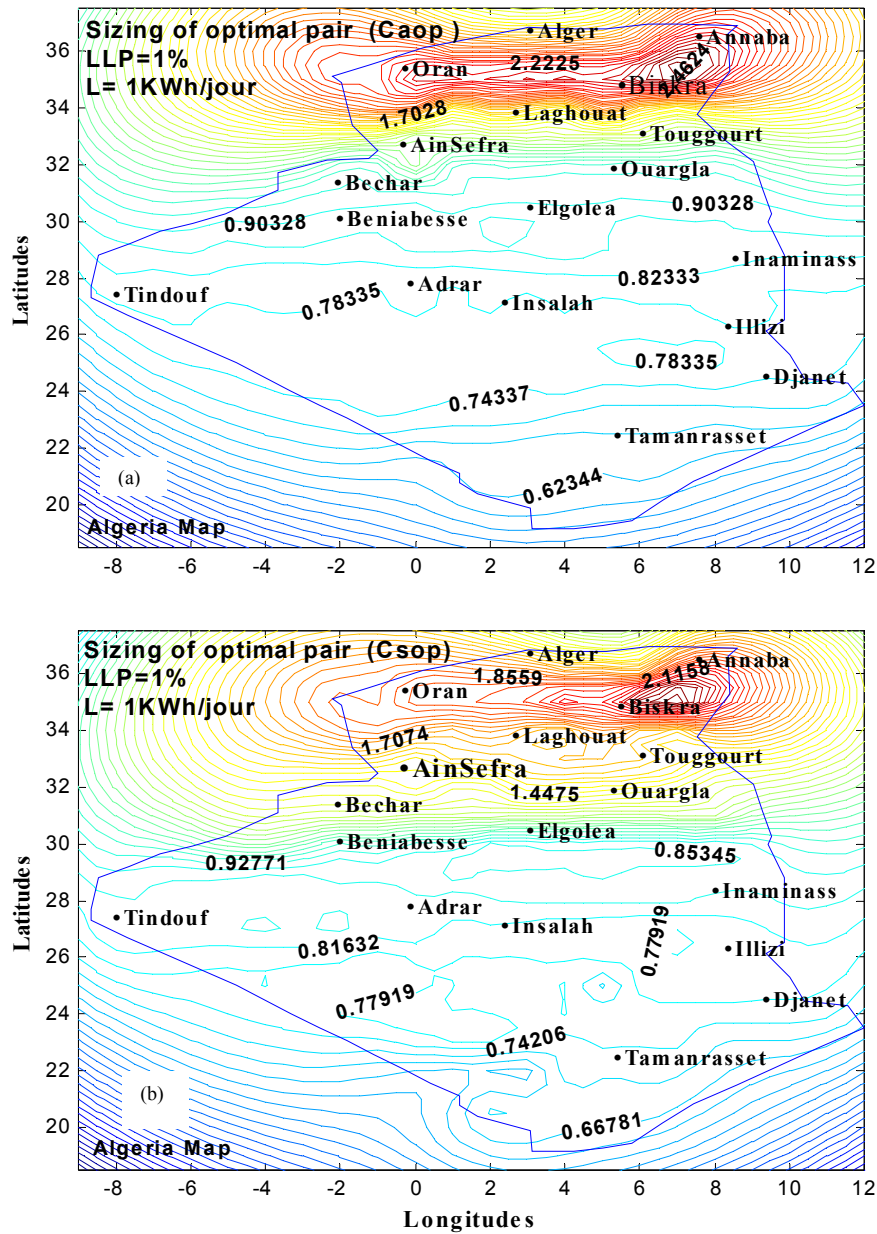Fig.2. B-spline functions $B_k(t)$ for k=0, 1, 2, 3.

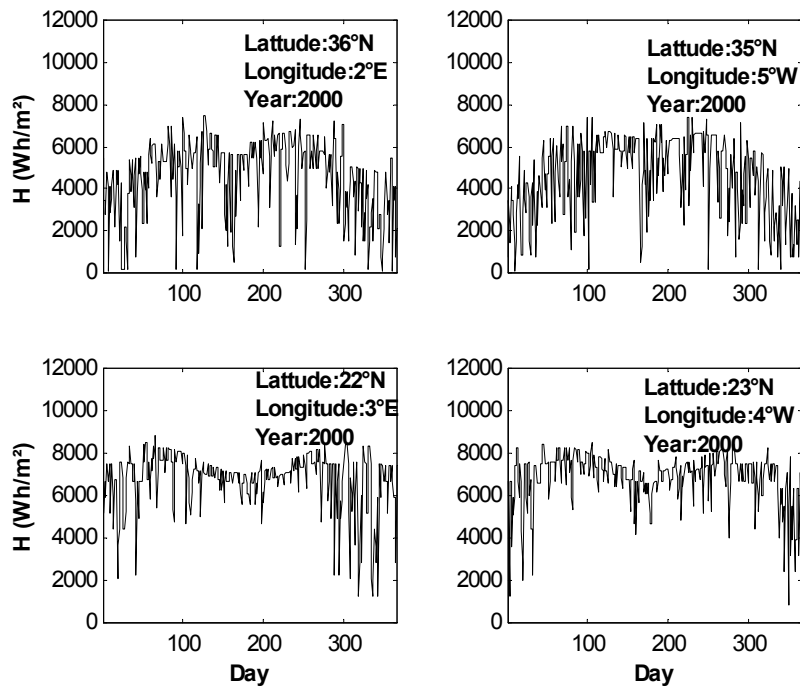Fig.3. Sizing map (a, b) of optimal pair of stand-alone PV system in Algeria

Fig.4. Daily values of inclined global solar radiation data for samples sites
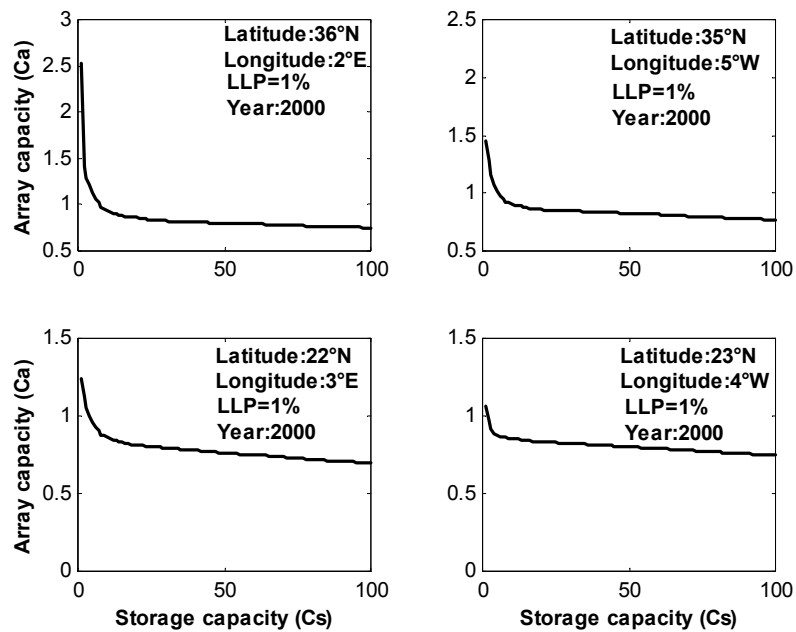


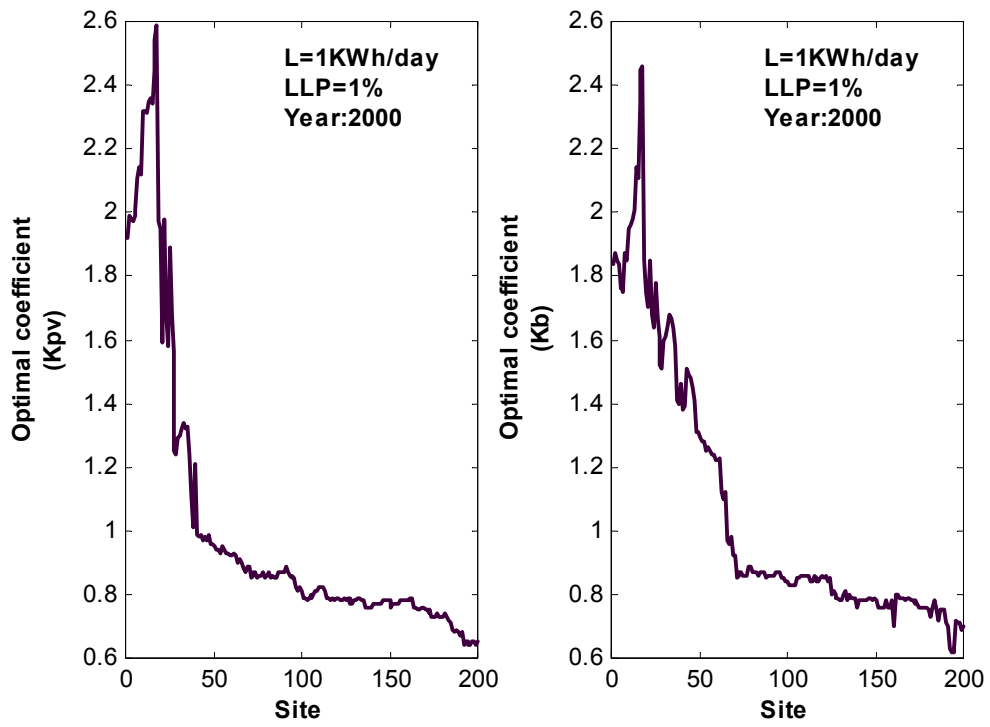Fig.5. Iso-reliability curves for Loss of Load Probability (LLP=1%)

Fig.6. Database of optimal sizing couple of stand-alone PV system for Algeria
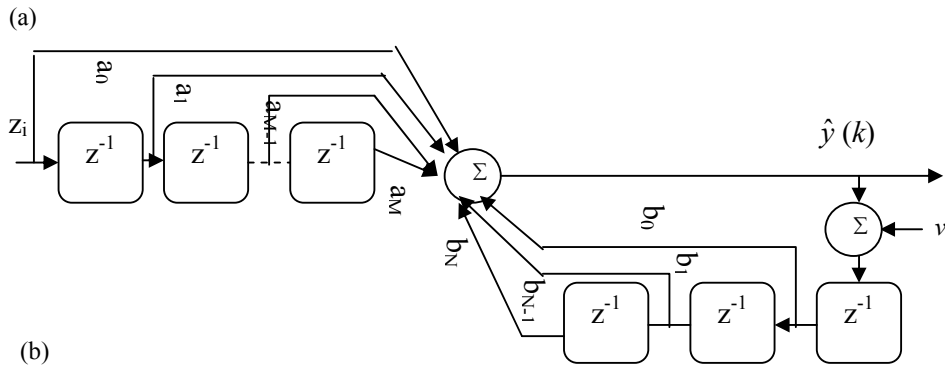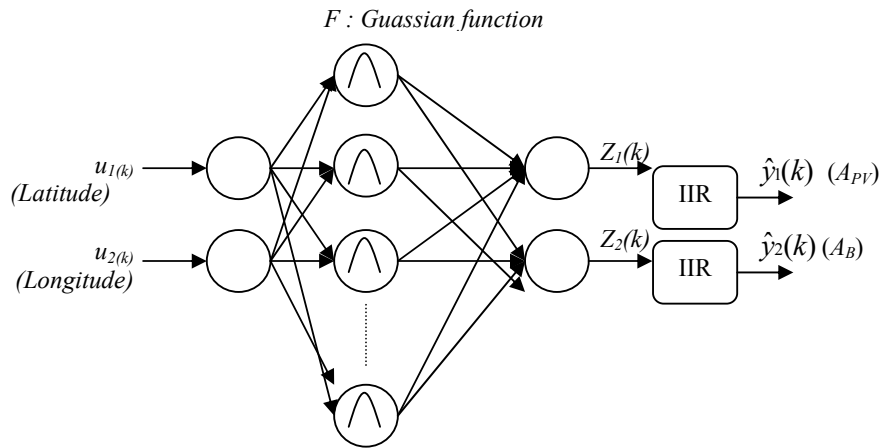
$F$ : Guassian function

$u_{1(k)}$
(Latitude)

$u_{2(k)}$
(Longitude)

$Z_1(k)$

$Z_2(k)$

IIR

IIR

$\hat{y}_1(k)$ $(A_{PV})$

$\hat{y}_2(k)$ $(A_B)$

(a)

$a_0$

$a_1$

$a_{M-1}$

$z_i$

$z^{-1}$

$z^{-1}$

$z^{-1}$

$a_M$

$\Sigma$

$\hat{y}(k)$

$\Sigma$ ← $v$

$b_0$

$b_N$

$b_{N-1}$

$b_1$

$z^{-1}$

$z^{-1}$

$z^{-1}$

(b)

Fig.7. IIR Adaptive RBF network, (a) Neural network architecture, (b) IIR model
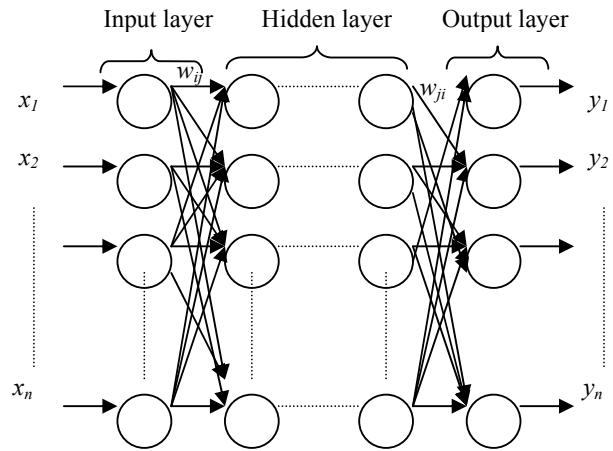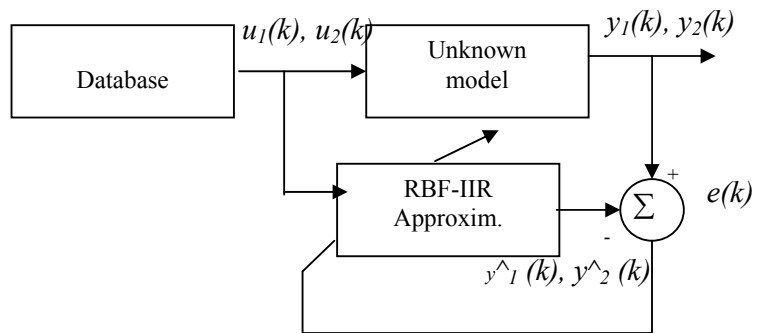
Fig.8. Feed-forward neural network architecture



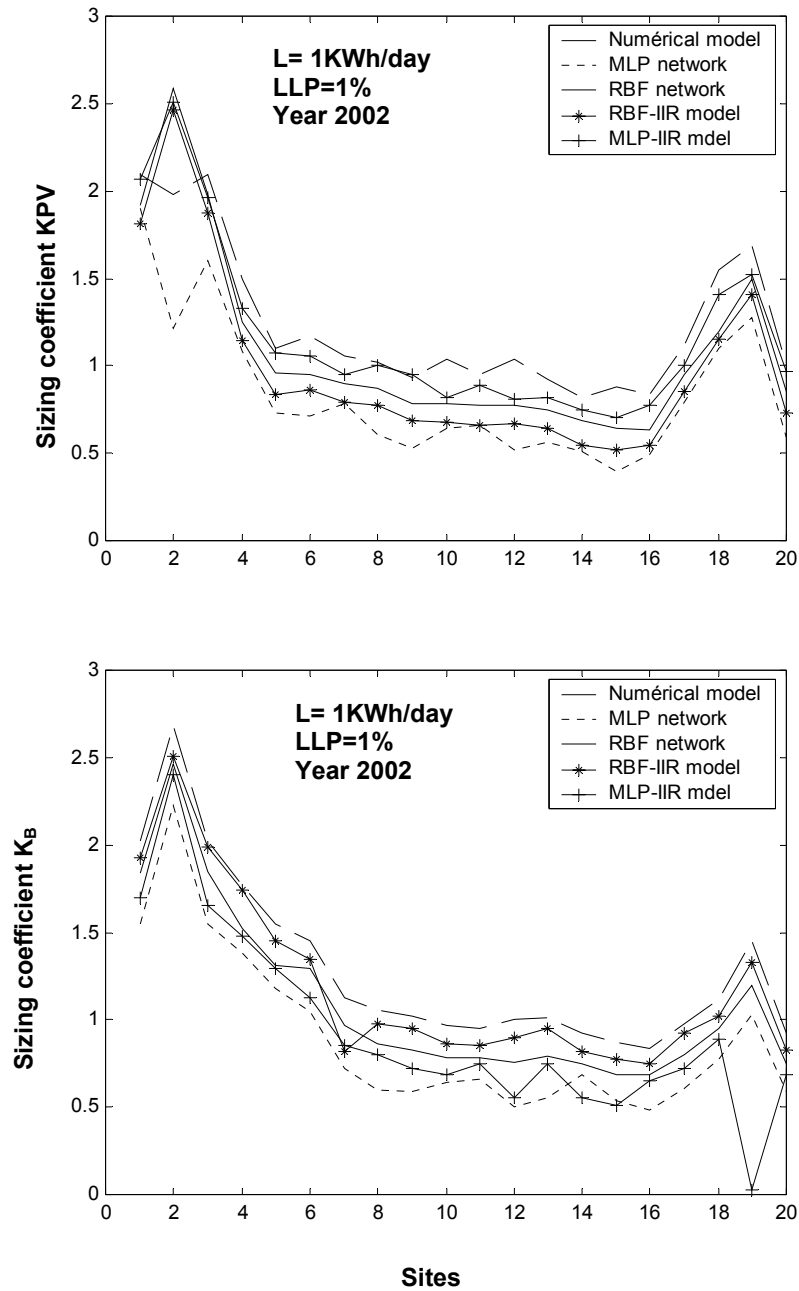Fig.9. Diagram block of developed model

Fig.10. Comparison between measured and estimated data, (a) classical
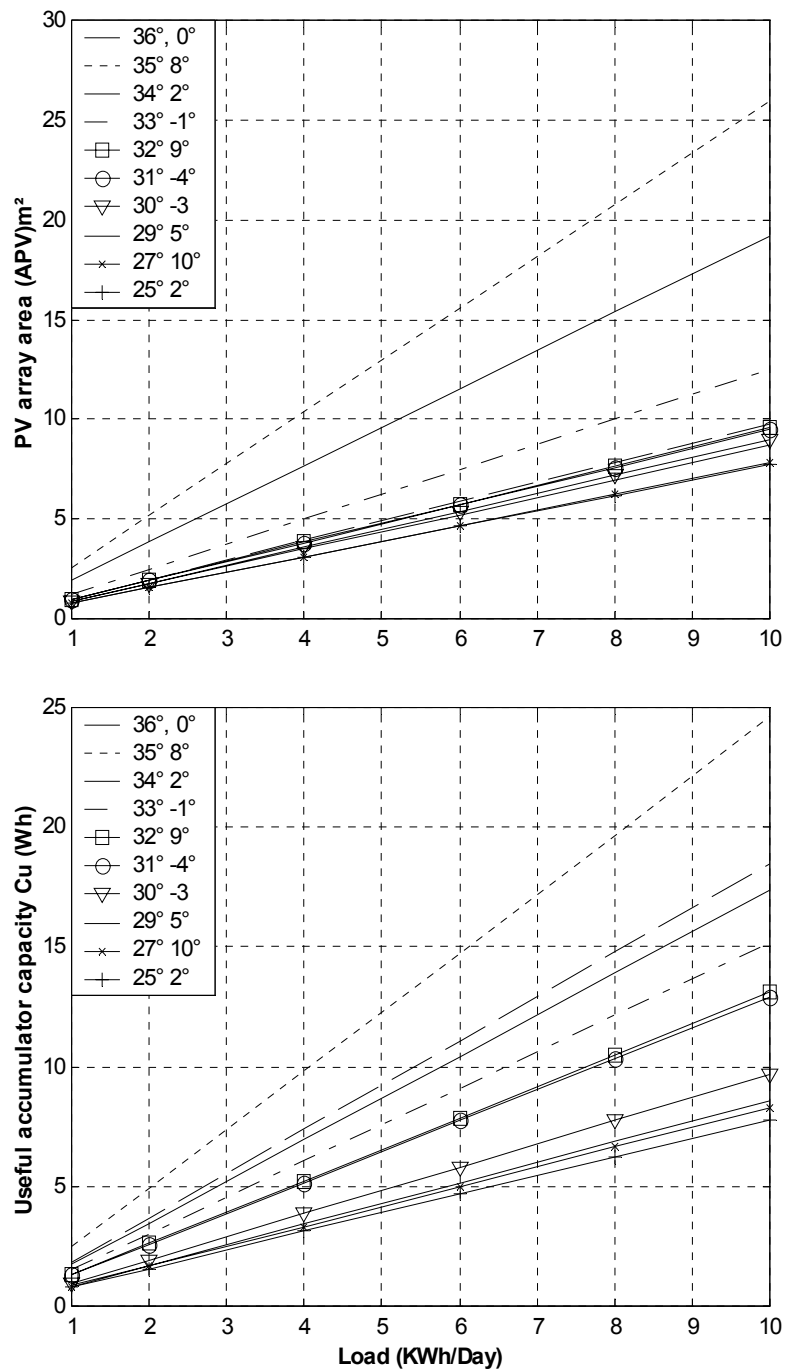models, (b) neural network models

Fig.11 Graphic abacus for sizing of SAPV system (L=1K Wh/Day, 10KWh/day)