Indian Journal of Science and Technology, *Vol 9(38), DOI: 10.17485/ijst/2016/v9i38/86214, October 2016*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Analysis of Autoregressive Predictive Models and Artificial Neural Networks for Irradiance Estimation

Mauricio Mauledoux[1], Oscar Aviles[1], Edilberto Mejia-Ruda[1] and Oscar I. Caldas[2*]

[1]Mechatronics Engineering Program, Militar Nueva Granada University, Bogota, Colombia; mauricio.mauledoux@unimilitar.edu.co, oscar.aviles@unimilitar.edu.co, edilberto.mejia@unimilitar.edu.co

[2]Industrial Engineering Program, Militar Nueva Granada University, Cajica, Colombia; oscar.caldas@unimilitar.edu.co

Abstract

Objectives: A model predictive controller was designed for a DC microgrid performed in Universidad Militar Nueva Granada at Cajica campus, which requires a 24-hours estimation of solar irradiance. **Methods/Statistical Analysis**: Autoregressive and neural networks based predictive models were designed and tested in order to be used, as well as an Artificial Neural Network (ANN) that was trained to be compared as an alternative solution to the problem. All models were coded and simulated on MATLAB and their performance were verified and mutually compared in order to define the best forecasting approach in the target allocation. **Findings:** The lack of seasons and the stochastically recorded irradiance time series, caused by sudden cloudy moments are the main characteristics of the local weather behavior. Therefore, a 5-years hourly meteorological database was used to estimate and train the ARMAX, NNF, NAR and NARX models, with the main feature of using six time and meteorological variables (air temperature, solar irradiance, atmospheric pressure, day, month and hour of measurements) to estimate a single output of hourly future irradiance. All of them were tested with statistical comparison functions such as square and absolute error criteria, Retrogression coefficient (R) and autocorrelation. **Application/Improvements:** The results let to define the most appropriate model to be used to generate the online data required for MPC designing to assure efficient operation of DC microgrids.

Keywords: Artificial Neural Network, DC Microgrid, Prediction Model, Solar Irradiance

1. Introduction

The current trend of power consumption and generation is the use of DC microgrids[1,2], considering that Distributed Generation units (DG) can either be used as utility-connected systems or stand-alone systems at isolated areas (not connected to the national grid)[2–5]. Regarding the second case, the DG units are usually operated by individuals in a wide range of houses, since small ones (e.g. single households) to large buildings[2,6,7]. The utilities use those DG units in order to enhance system stability and power supply flexibility, quality and expandability as well as to optimize distribution system and costs related with transmission and distribution. Besides, in most cases DG units produce incompatible AC or DC power and as a result, power electronics interfaces must be used to obtain the desired voltage magnitude, frequency and phase angle[8–10] and then connect each DG to the main grid. In order to reduce the corresponding costs, a single power interface could be used for all DG units, which also leads to energy lost reduction[11,12].

The time variability of power loads must be solved by being systematically predictable, due to the fact that an entire day average power consumption could be already known by means of weather forecasting. Therefore, the most appropriate choice of a top-level controller, in order to regulate the entire system dynamics, would be a Model Predictive Controller (MPC)[13–16]. Nevertheless, this kind of controller, as well as others model dependent

**Author for correspondence*

ones, requires a strong mathematical description of the system for ensures accuracy when data is being acquired, along with a significant computing power. Thus, the system model for the particular application, consist in a non-linear equation that defines the dynamics of a photo voltaic modules array, which needs an accurate irradiance forecasting. If the solar irradiance for a previously defined period (say 24 hours) can be predicted, it would be the main tool for setting a better power dispatching plan.

This estimation problem has been studied previously by mean of several approaches, including empirical models[17], analytical models[18], numerical models and statistical[19], as well as ANN[20,21,32,33]. Regarding those based on statistical processes, such as Autoregressive (AR), Autoregressive Moving-Average (ARMA), Moving-Average (MA), Autoregressive-Integrated Moving-Average (ARIMA) have been used widely for modeling the problem mentioned above[21].

ANNs are successful techniques when used in problems that require parameters hard to define and there is a wide group of examples. Besides, a neural network approach does not need to have any information regarding the process that generates data. Instead, it needs a long-term data in order to develop a better model that can be used for its original purpose[21,22].

Hence, this paper is organized as follows: It begins with an explanation of collection method for the time series data and all the parameters and variables related with the process, after the proposed models are introduced according to two types of estimation techniques: Regression and neural networks. All these models are compared by different square and absolute deviation estimators, as well as by the calculation of the retrogression coefficient and autocorrelation error. Finally, correlation charts and plots of predicted and measured data are shown in order to analyze results and conclude.

2. Identification and Modeling

A weather station located at the university campus in Cajica let to collect information hourly to set a 5-years database January 2010-February 2014). The database includes several important variables, such as, atmospheric Pressure (P), Temperature (T), Relative Humidity (RH), time (Month, Day, Hour) and Solar Irradiance (SR). This information allows to notice the main benefit of Colombia's geographic location (over the equator line): Relatively low variability, which causes more concentrated solar energy than the one that can be gotten at the subtropics (trepid or even frigid zones), i.e. sort of steady during a year.

The following model-based techniques were trained with similar parameters. All of them used 30115 samples with 80%, 10% and 10% distribution for training, validation and testing, respectively. Besides, the three neural networks models included a 10-neuron size hidden layer and 10 iterations of delay.

2.1 Auto Regressive Moving Average with Exogenous Excitation

A general Auto Regressive Moving Average with Exogenous excitation (ARMAX) model is defined as follows:

This process model allows to correctly modeling systems behavior considering both control and disturbance inputs, being the basic structure for model predictive control strategies[23]. ARMAX structure is shown in Figure 1[24].

Equation (1) expresses a general form of ARMAX model. *A (q), B (q)* and *C (q)* denote the polynomial of the dynamic process system in Equations (2-4), meanwhile *y (t)* represents the system output, *e (t)* the random prediction error and *x (t)* input variable vector[25,31].

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})e(t) \tag{1}$$

$$A(q^{-1}) = a1 + a_2^{q^{-1}} + a_3^{q^{-2}} + \cdots + a_{n+1}^{q^{-n}} \tag{2}$$

$$B(q^{-1}) = b1 + b_2^{q^{-1}} + b_3^{q^{-2}} + \cdots + b_{n+1}^{q^{-n}} \tag{3}$$

$$C(q^{-1}) = c1 + c_2^{q^{-1}} + c_3^{q^{-2}} + \cdots + c_{n+1}^{q^{-n}} \tag{4}$$

The Equations shown above (*A (q), B (q)* and *C (q)*) can also be expressed as coefficients arrays that are obtained by

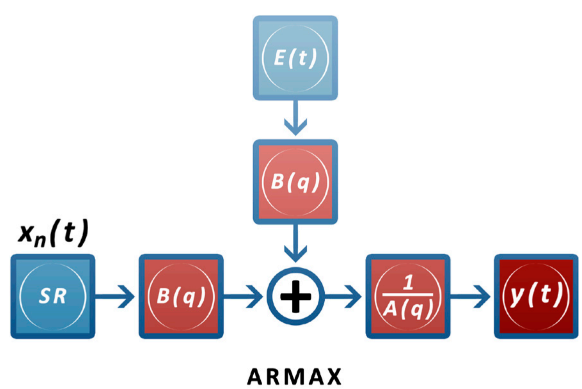


Figure 1. ARMAX model implemented.

fitting the training data with linear curves and q denotes the number of past observed data. The prediction of the ARMAX model is expressed in (5) and developed in (6), (7) and (8)[25].

$$\hat{y}_t = \left[1 - \frac{A(q^{-1})}{C(q^{-1})}\right] y_t + \frac{A(q^{-1})\,B(q^{-1})}{C(q^{-1})\,A(q^{-1})} u_t \qquad (5)$$

$$\hat{y}_t = -c_1 \hat{y}_{t-1} + \cdots + y_t + c_1 y_{t-1} + \cdots - y_t - a_1 y_{t-1} - \square + b_1 u_{t-1} \qquad (6)$$

$$\hat{y}_t = [1 - C][(y]_t - \hat{y}_t) + [1 - A] y_t - B u_t \qquad (7)$$

$$\hat{y}_t = [1 - C] e_t + [1 - A] y_t - B u_t \qquad (8)$$

2.2 Nonlinear Autoregressive ANN

The Nonlinear Autoregressive (NAR) model is a non-linear relation between the past outputs and the predicted process output, which and can be delineated by a high order difference equation[26]. The schematic diagram is shown in Figure 2, where it has ten neurons in the hidden layer. The network is trained with Levenberg-Marquardt back propagation algorithm.

Equation (9) represents a general form of NAR model.

$$y(t) = f\{y(t-1) \ldots y(t-n)\} + e(t) \qquad (9)$$

where $y(t)$ is the output of the model at time t. n is the output order of the closed loop dynamical model and f is a initially unknown nonlinear function, which is approximated by the regression model of (10)[26].

$$y(t) = \sum_{i=1}^{n} a(i) y(t-i) + e(t) \qquad (10)$$

Where $a(i)$ are the coefficients of the linear and nonlinear autoregressive outputs.

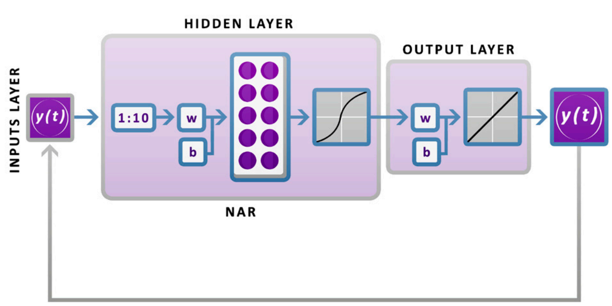


Figure 2. NAR architecture implemented.

2.3 Nonlinear Autoregressive with Exogenous Excitation ANN

The Nonlinear Autoregressive with Exogenous Excitation model (NARX) is characterized by the non-linear relations between the past inputs, past outputs and the predicted process output, which can be obtained by means of the high order difference shown in Equation[26]. The proper blocks diagram is shown in Figure 3 where the network has ten sigmoid hidden neurons in a single layer and a linear output neuron. The network is trained with a Levenberg-Marquardt back propagation algorithm.

Equation (11) shows a general form of the NARX model.

$$y(t) = f\{y(t-1) \ldots y(t-n_y), u(t-1) \ldots u(t-n_u)\} + e(t) \qquad (11)$$

Where $u(t)$ and $y(t)$ represents the input and output of the model at time t, respectively. In addition, n_y and n_u are the input and output orders of the dynamical model, considering that $n_u \geq 0$ and $n_y \geq 1$. The nonlinear function f is unknown and can be approximated by the regression model of (12)[26].

$$y(t) = \sum_{i=0}^{n_u} a(i) u(t-i) + \sum_{j=0}^{n_y} b(j) y(t-j) + \sum_{i=0}^{n_u} \sum_{j=i}^{n_y} a(i,j) u(t-i) u(t-j) + \sum_{i=1}^{n_u} \sum_{j=i}^{n_y} \qquad (12)$$

where $a(i)$ and $a(i,j)$ are the coefficients for the linear and nonlinear original exogenous input; $b(i)$ and $b(i,j)$ are the coefficients of the linear and nonlinear autoregressive variable; $c(i,j)$ are the coefficients of the nonlinear cross variables[26].

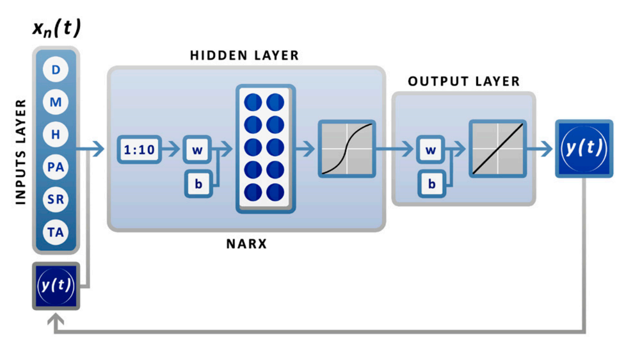


Figure 3. NARX architecture implemented.

2.4 Neural Network Fitting

The Neural Network Fitting (NNF) process uses a two-layer feed-forward network with sigmoid hidden neurons and linear output neurons, that is able to fit problems considerably well, starting from consistent data and enough neurons in its hidden layer. It usually has ten _neurons in the hidden layer as shown in Figure 4.

The network was trained with a Levenberg-Marquardt back propagation algorithm, unless there is not enough memory less. In that case a scaled conjugate gradient back propagation will be used.

Equation (13) describes the mathematical model that the NNF uses.

$$y(t) = f(w_1\, x(t-1) + b_1 \ldots w_n x(t-n) + b_n) \tag{13}$$

where $x(t)$ and $y(t)$ are the inputs and output respectively, f is a nonlinear function, w_n is the polynomial weight coefficient and b_n is the bias for neural network activation.

3. Results and Discussion

All the neural network models benefited of the meteorological data collected mainly time and irradiance. The data set was randomly distributed in an 80/10/10 ratio for training, validation and testing, respectively.

Several validation methods were selected and applied using ANN statistics and including accuracy criteria (level difference) among the current values and the estimations. The techniques applied were such as the Mean Squared Error (MSE), Mean Absolute Error (MAE), Sum of Squared Error (MSE) and Sum of Absolute Error (MAE), according to (14), (15), (16) and (17)[27].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}[y_i - \hat{y}_i]^2 \tag{14}$$

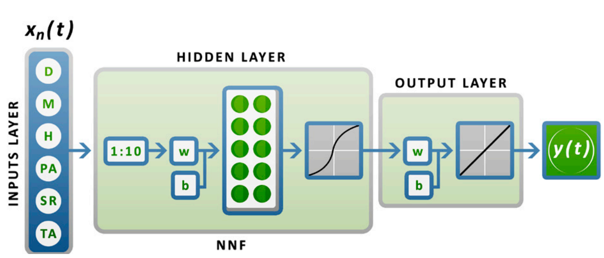


Figure 4. NNF architecture implemented.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{15}$$

$$SSE = \sum_{i=1}^{N}[y_i - \hat{y}_i]^2 \tag{16}$$

$$SAE = \sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{17}$$

Where N is the total number of data, y_i is an array of target values and $\hat{y}_i$ is the corresponding estimated values[28,29]

Another criterion is the squared correlation coefficient (R^2), shown in (14) where $\mathbf{Y_{pt}}$ and $\mathbf{Y_t}$ represent the predicted and measured values of the test set, respectively. $\mathbf{Y_{tr}}$ is the mean value of the training set[30].

$$R^2 = 1 - \left(\frac{\Sigma[y_{pt} - y_t]^2}{\Sigma_j[y_t - \bar{y}_{tr}]^2}\right) \tag{18}$$

The evaluation criteria results can be analyzed in Table 1, performed on all of the applied models.

The neural network fitting model is the one with better results in every evaluation criterion, as can be seen in Table 1, i.e. it has the lower error value on every criterion but R^2, which measures accuracy as nearness to a normalized unit instead of lowest error.

Figure 5 shows the models behavior when using irradiance data from a random selected day of the year.

Table 1. Comparison of MSE and R for dataset

Estimation Method	Error				
	MSE	R^2	SSE	SAE	MAE
NARX	11088.3	0,922	3,197E+08	1,704E+06	56,586
NAR	11031.4	0,923	3,344E+08	1,754E+06	58,245
NNF	10053.4	0,931	3,112E+08	1,635E+00	54,285

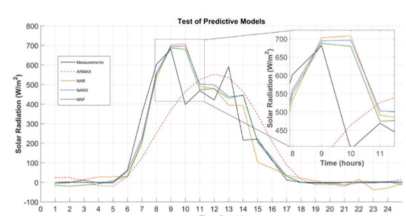
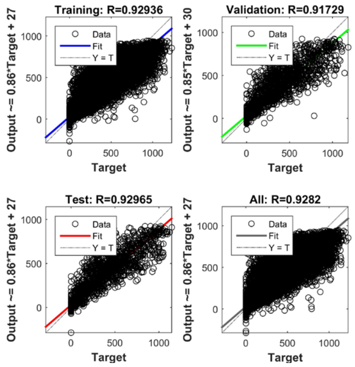
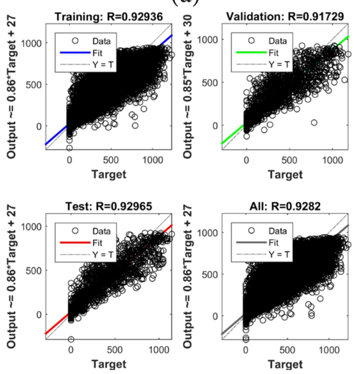


Figure 5. Comparison of scatter plots of the models used for solar irradiance prediction (ARMAX, NNF, NAR y NARX).

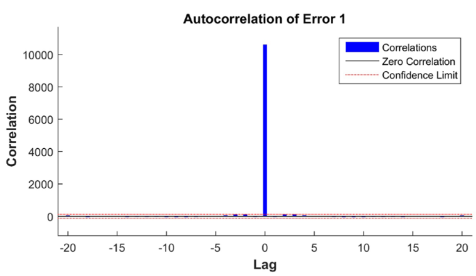


(a)

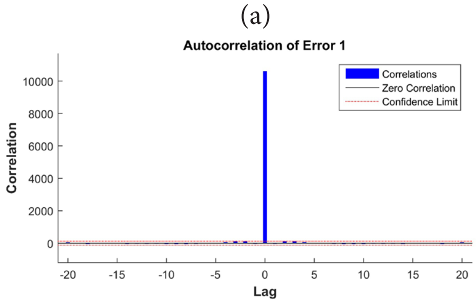


(b)

Figure 6. Comparison of scatter plots of traininig, validation and test of models used for dataset. (A) NAR regression plot, (B) NARX. (C) NNF.



(a)



(b)

Figure 7. Plots of autocorrelation error. (A) NAR, (B) NARX.

This came out as an interesting particular case, due to sudden fall of irradiance just in the noon crest, which is not rare at all and happens due to dense clouds appearing. Most of the models work properly to data noise but ARMAX, which ignores abrupt variations in order to define an average curve fitting, with a corresponding detriment of the prediction accuracy.

Figure 6 shows the regression plots during training, validation and testing.

The autocorrelation plots for NARX and NAR models to assess the training and testing processes were appropriate in base of inputs and outputs. It can be seen on Figure 7.

4. Conclusions

In order to develop a proper hourly irradiance forecasting, two approaches were considered and compared. The first one was de ARMAX model due to its convenience for stochastic time series analysis and the use of both auto regression and moving average polynomials; on the other hand, some Artificial Neural Networks were taken into account as good approaches for functions estimation, usually in machine learning and data mining, but still appropriate for this large number of inputs and data variation. Both prediction methods were employed to create a 24-hour estimation of the UMNG irradiance at Cajica campus. The level of accuracy for each estimation model was measured the following measurements: MSE, MAE, SAE, SSE and both retrogression coefficient and autocorrelation error calculations.

The first conclusion was the early discard of the simple parameter regression technique (ARMAX) because it showed the worst estimation quality, measured as the highest MSE (12680). Nevertheless, all of the ANN models had similar results in terms of MSE (about 11.000), SSE (about 3.25E+11), SAE (about 1.7E+9) and MAE (about 56.3), whi15ch meant a harder selection. Despite of this situation it is important to remark that this also means that the ANN model forecasting could be a feasible approximation of the upcoming hour solar radiation. Therefore, it can be said that the comparative analysis between the estimated and measured data concluded that ANN models have the capability to recognize the relationship between the input and output variables and predict hourly solar radiation with admissible accuracy.

On the other hand, six (6) geographical and meteorological variables were used as inputs for the NARX and ANN models, i.e. air temperature, irradiance, atmospheric pressure, day, month and hour of measurement, which were appropriate and necessary for the hourly solar irradiance as single output. Nonetheless, the NAR model only needed the past time series of irradiance, which implies and additional benefit of this model, due to the fact that the online instrumentation just requires a high precision pyranometer to measure hourly the solar radiation.

Finally, the Nonlinear Autoregressive Artificial Neural Network Model (NAR) showed many advantages in regard to the other considered techniques, remarking the use of less input parameters, which considerably ease the solar irradiance forecasting adoption. Besides, this method can play a remarkable role in the design and performance of a MPC controlled operation in a DC microgrids with the use of renewable resources systems.

5. Acknowledgement

This work was supported by the Universidad Militar Nueva Granada and the department of research under the project titled: Development, automation and control of renewable resource hybrid plant (ING-1577).

6. References

1. Salomonsson D. Modeling, control and protection of low-voltage DC microgrids. Stockholm: Royal Institute of Technology; 2008.
2. Justo JJ, Mwasilu F, Lee J, Jung JW. AC-microgrids versus DC-microgrids with distributed energy resources: A review. Renewable and Sustainable Energy Reviews. 2013; 24:387–405.
3. Laaksonen HJ. Protection principles for future microgrids. IEEE Transactions on Power Electronics. 2010 Dec; 25(12):2910–8.
4. Lund H. Renewable energys trategies for sustainable development. Energy. 2007 Jun; 32(6):912–9.
5. Kanellos FD, Hatziargyriou ND. Control of variable speed wind turbines equipped with synchronous or doubly fed induction generators supplying islanded power systems. IET Renewable Power Generation. 2009 Mar; 3(1):96–108.
6. Noroozian R, Abedi M, Gharehpetian GB, Hosseini SH. Distribution resources and DC distribution system combination for high power quality. International Journal of Electrical Power and Energy System. 2010 Sep; 32(7):769–81.
7. Das A, Balakrishnan V. Sustainable energy future via grid interactive operation of SPV system at isolated remote island. Renewable and Sustainable Energy Reviews. 2012 Sep; 16(7):5430–42.
8. Anand S, Fernandes BG. Modified droop controller for paralleling of DC–DC converters in stand alone DC system. IET Power Electronic. 2012 Sep; 5(6):782–9.
9. Blaabjerg F, Consoli A, Ferreira JA,Van Wyk JD. The future of electronic power processing and conversion. IEEE Transactions on Industrial Applications. 2005 Jan; 41(1):3–8.
10. Li Y, Kao C. An accurate power control strategy for power-electronics interfaced distributed generation units operating in a low-voltage multibus microgrid. IEEE Transactions on Power Electronics. 2009 Dec; 24(12):2977–88.
11. Moreira CL, Resende FO, Peas Lopes JA. Using low voltage microgrids for service restoration. IEEE Transactions on Power Systems. 2007 Feb; 22(1):395–403.
12. Varaiya PPV, Wu FF, Bialek JW. Smart operation of smart grid: Risk-limiting dispatch. Proceedings of the IEEE. 2011 Jan; 99(1):40–57.
13. Heer P. Decentralized model predictive control for smart grid applications. Electrical Engineering; Sweden. 2013.
14. Maciejowski J. Predictive control with constraints. USA: Prentice Hal; 2001 Jun. p. 352.
15. Picasso B, Vito DD, Scattolini R, Colaneri P. An MPC approach to the design of two-layer hierarchical control systems. Automatica. 2010 May; 46(5):823–31.
16. Trangbaek K, Pedersen M, Bendtsen J, Stoustrup J. Predictive smart grid control with exact aggregated power constraints. Smart Power Grids. Berlin Heidelberg: Springer-Verlag; 2011. p. 655–74.
17. Rehman S. Empirical model development and comparison with existing correlations. Applied Energy. 1999 Sep; 64(1-4):369–78.
18. Paulescu M, Fara L,Tulcan-Paulescu E. Models for obtaining daily global solar irradiation from air temperature data. Atmosphere Research. 2006 Mar; 79(3-4):227–40.
19. Mellit A, Kalogirou S, Shaari S, Salhi, Hadj Arab A. Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas: Application for sizing a stand-alone PV system. Renewable Energy. 2008 Jul; 33(7):1570–90.
20. Hontoria L, Aguilera J, Zufiria P. Generation of hourly irradiation synthetic series using the neural network multilayer perceptron. Solar Energy. 2002 May; 72(5):441–6.
21. Mellit A, Eleuch H, Benghanem M, Elaoun C, Pavan AM. An adaptive model for predicting of global, direct and diffuse hourly solar irradiance. Energy Conversion and Management. 2010 Apr; 51(4):771–82.
22. Mellita A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: A review. Progress in Energy and Combustion Science. 2008 Oct; 34(5):574–32.
23. Camacho EF, Alba BC. Model predictive control. London: Springer-Verlag; 2004.

24. Kocur S. Identification and modelling of linear dynamic systems. Advances in Electrical and Electronic Engineering. 2011; 5:40–2.
25. Viana IDE, Palacio JAP, Lemus CG. Estimacion de modelos lineales para el control predictivo de pequenas centrales hidroelectricas. Revista Soluciones de Postgrado EIA. 2012; 9:63–78.
26. Benmouiza K, Cheknane A. Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models. Energy Conversion and Management. 2013 Nov; 75:561–9. ISSN 0196-8904.
27. Braik M, Sheta AF, Arieqat A. A comparison between GAs and PSO in training ANN to model the TE chemical process reactor. Proceedings of the AISB 2008 Symposium on Swarm Intelligence Algorithms and Applications; Aberdeen, Scotland. 2008. p. 1–7.
28. Pelland S, Galanis G, Kallos G. Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model. Progress in photovoltaics: Research and Applications. 2011; 21(3):284–96.
29. Chandel AK, Chandal SS. Artificial Neural Network based prediction of solar radiation for Indian stations. International Journal of Computer Applications. 2012 Jul; 50(9):1–4.
30. Luo H, Liu D, Huang Y. Artificial Neural Network modeling of algal bloom in Xiangxi Bay of Three Gorges Reservoir. 2010 International Conference on Intelligent Control and Information Processing (ICICIP); Dalian. 2010 Aug. p. 645–7.
31. Daneshwar MA, Noh NM. Application of radial basis function neural networks in modeling of nonlinear systems with deadband. Indian Journal of Science and Technology. 2013 Nov; 6(11):5469–73. Print ISSN: 0974-6846. Online ISSN: 0974-5645.
32. Rezvani A, Izadbakhsh M, Gandomkar M, Vafaei S. Investigation of ANN-GA and modified perturb and observe MPPT techniques for photovoltaic system in the grid connected mode. Indian Journal of Science and Technology. 2015 Jan; 8(1):87–95. DOI: 10.17485/ijst/2015/v8i1/54277.
33. Kamali R, Sharifi R, Radmanesh H, Fathi SH. Online voltage estimation for distribution networks in presence of distributed generation. Indian Journal of Science and Technology. 2016 May; 9(18). DOI: 10.17485/ijst/2016/v9i18/71484.