A SEASONAL AND MONTHLY APPROACH FOR PREDICTING THE DELIVERED ENERGY QUANTITY IN A PHOTOVOLTAIC POWER PLANT IN ROMANIA

Prof. George Căruțașu Ph. D
The Romanian-American University
Faculty of Computer Science for Business Management
Bucharest, Romania
Lect. Alexandru Pîrjan Ph. D
The Romanian-American University
Faculty of Computer Science for Business Management
Bucharest, Romania

Abstract: In this paper, we present solutions that facilitate the forecasting of the delivered energy quantity in a photovoltaic power plant using the data measured from the solar panels' sensors: solar irradiation level, present module temperature, environmental temperature, atmospheric pressure and humidity. We have developed and analyzed a series of Artificial Neural Networks (ANNs) based on the Levenberg-Marquardt algorithm, using seasonal and monthly approaches. We have also integrated our developed Artificial Neural Networks into callable functions that we have compiled using the Matlab Compiler SDK. Thus, our solution can be accessed by developers through multiple Application Programming Interfaces when programming software that predicts the photovoltaic renewable energy production considering the seasonal particularities of the Romanian weather patterns.

JEL classification: C53, L86, O13, Q42, Q47.

Key words: Artificial Neural Networks, Levenberg-Marquardt, renewable energy, overfitting, performance indicators

1. INTRODUCTION

In the recent years, one of the most important issues at global level is the sustainable development, targeting the evolution of the humankind as to sustain the natural resources on which the society and economy rely. One of the most important factors of this sustainable development is the supplying of energy at affordable costs, in secure ways, valuing the nature, taking into account the economic and social needs of the society.

Energy is an essential factor in improving human welfare, raising living standards or even in eradicating poverty. However, most current methods of supplying and using energy are, unfortunately, unsustainable. In many areas of the world the energy supplying is not safe and reliable and consequently their economic development is limited. In other cases, the sustainable development is impeded due to the environmental degradation caused by the production and use of energy [1].

One of the most important factors that have influenced the evolution of many countries from an agricultural-based economy to a modern one, based on industry and services, is the quality of the energy related processes, that should be affordable and adequate to the necessities of the society. Thus, the energy influences all the aspects of the social and economic life, the industry, the commerce, the living standards, the prosperity and, moreover, the development of a sustainable economy and a clean, healthy environment.

Consequently, the governments should pay special attention to develop the plans, programs, policies and strategies suitable for their countries, to understand the local and global effects, impacts and implications of their decisions. In order to obtain data that describes the quality and quantity of energy and to develop a proper policy, one should correctly compute and analyze specific indicators. These indicators represent an efficient way of highlighting statistical data, of analyzing the factors that could influence the energy and many aspects related to the social life, the environment, the overall progress. Such indicators also provide information about the way in which all of the above-mentioned factors could be influenced and improved over time. Choosing suitable indicators helps to analyze the policies and to take adequate decisions in order to develop the best strategy, leading to the desired sustainable development.

In this context, due to their undeniable advantages, the renewable energy resources usage must be encouraged as these resources offer secure, practically unlimited energy at affordable prices, preserving traditional resources and reducing their imports, stimulating the progress and development at all the levels, creating new jobs, reducing the pollution and gas emissions [2,3]. This kind of energy is generated from renewable, naturally replenished resources (geothermal heat, tides, rain, wind, sunlight).

Many renewable energy resources are related to the solar radiation. Practically, in each hour the amount of solar radiation received by the Earth represents a greater amount of energy than the worldwide demand for an entire year, while in three days the received solar radiation is comparable to the amount of energy that could be generated by all the fossil energy resources that the humanity knows in the present [4]. The sun energy represents a free, inexhaustible resource that is harnessed and becomes useable through the solar energy technology.

In order to monitor, evaluate and analyze the renewable energy, the management of the renewable resources should be sustained by a decision support system (DSS). After analyzing various approaches that have been used in the literature for developing such systems [5-8], due to their multiple advantages, we have decided to present in this paper a series of Artificial Neural Networks (ANNs), that we have developed, trained, validated and tested. These solutions can be used to design and develop a DSS having as a main purpose the forecasting, monitoring and analyzing the specific performance indicators related to renewable energies [9].

In the first stage of our research we have analyzed and compared the main properties of three of the most important algorithms on which could be based the development of such neural networks: the Scaled Conjugate Gradient (SCG), the Bayesian Regularization (BR) and the Levenberg-Marquardt (LM) algorithms. Afterwards, based on these algorithms, we have developed, trained, validated and tested a series of Artificial Neural Networks, using the Neural Network Toolbox from the development environment MatlabR2015a [10]. In this paper, we present the results that we have obtained when we have developed the Artificial Neural Networks based on the Levenberg-Marquardt algorithm (as the results obtained using this algorithm were slightly improved compared to the results provided by the other two algorithms).

The Levenberg-Marquardt (LM) algorithm is useful for training Artificial Neural Networks and has many applications in mathematics and computer science. It is also known as the Damped Least-Squares (DLS) algorithm, being based on the least squares method. The LM algorithm aims to build a mathematical function or curve (that may be subject to restrictions), corresponding to a data set. The LM is used in a wide range of applications related to the curves adjustment problems and as well as other adjustment algorithms, it provides a local minimum, not a global one [11].

The experimental data that corresponds to the input and output parameters consist in a number of 8413 samples (after excluding the irrelevant samples), obtained through hourly measurements over one-year period (from December 2013 to November 2014), in a photovoltaic power plant (PPVP) located in the Izvoru village, in the Giurgiu County, Romania. This power plant with an installed capacity of 9.6 MW harnesses a number of 40026 panels, with a capacity of 240 Wp per panel. The PPVP is connected to an existing 110/20kV substation.

In contrast to a previous work [12] that analyses the forecasting accuracy obtained by developing a global neural network based on a whole two-year data set, in this paper we fine tune the research, by developing seasonal and monthly approaches, thus obtaining photovoltaic renewable energy prediction solutions that take into account the particularities of each season and month, accounting for the specific of the Romanian weather patterns.

In the following, we present and study the most important results obtained after having developed, trained, validated and tested the seasonal and monthly Artificial Neural Networks that can be afterwards used in the purpose of forecasting, monitoring and evaluating the performance indicators specific to the renewable energy field in Romania.

2. THE ARTIFICIAL NEURAL NETWORKS USED IN FORECASTING, MONITORING AND EVALUATING THE PERFORMANCE INDICATORS SPECIFIC TO PHOTOVOLTAIC POWER PLANTS

Each of the developed forecasting solutions takes as input data the humidity (in percentages), the atmospheric pressure (in hPa), the solar irradiation level (in W/m^2), the environmental temperature (in Celsius degrees), the present module temperature (in Celsius degrees) and forecasts the quantity of delivered energy (in kWh). We have developed and tested a series of ANNs, having various architectures, in order to identify the one that provides the best results (in terms of forecasting accuracy). The best results were obtained when the Input data layer had 5 neurons (corresponding to the input parameters), the Hidden layer had a size of 14 neurons, while each of the Output layer and Output data contained 1 neuron (Figure no. 1).

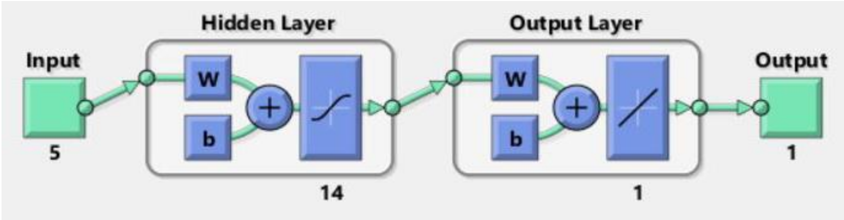
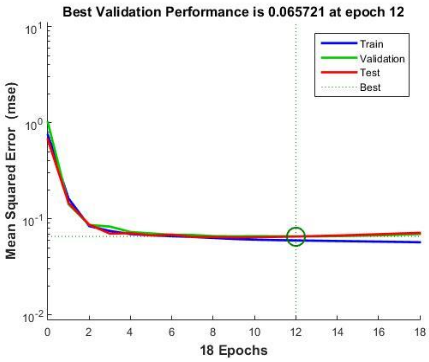


Figure no. 1. The Architecture used for developing the Artificial Neural Networks

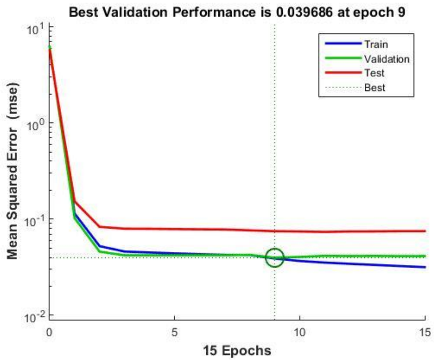
After a series of tests, we have allocated the subsets of the whole data set to the different stages of developing ANNs (training, validating, testing), choosing for each of these stages the amount of data that has provided the best forecasting accuracy: 60% of the samples were used in the training phase, while the remaining percentage of samples was equally allocated for the validation and testing phases. In all of the above-mentioned phases, the samples were randomly allocated. Based on the Levenberg-Marquardt (LM) algorithm, we have developed 4 seasonal and 12 monthly Artificial Neural Networks and we have named these networks in accordance to the period that was taken into account (e.g. NNSpring, NNMarch, etc).

In the following, we analyze the performance and the forecasting accuracy of the monthly Artificial Neural Networks developed for the months of June, July and August, based on the Levenberg-Marquardt algorithm, respectively NNJune, NNJuly and NNAugust. In order to further improve the obtained results, we also analyze the performance of the ANN developed for the Summer season, NNSummer.

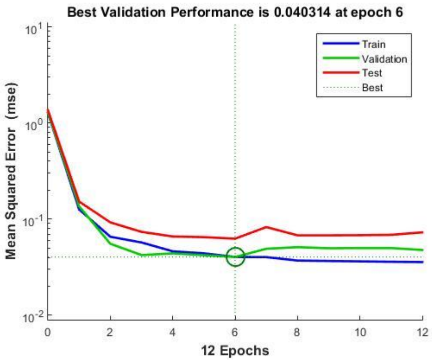
In this purpose, we have generated and analyzed a series of plots: the performance analysis, highlighting the Mean Squared Error MSE and its minimum value (Figure no. 2); the error histograms (we have highlighted the range of errors); the regressions (we have highlighted the correlation coefficient R and its minimum value) for all the above mentioned Artificial Neural Networks.



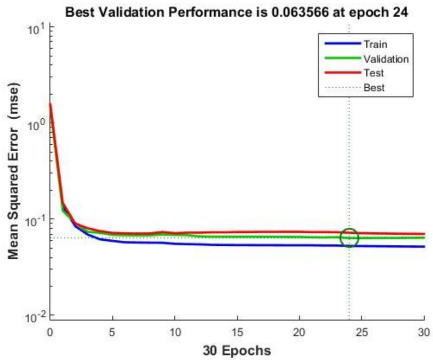
a) The validation performance of the NNJune Artificial Neural Network



b) The validation performance of the NNJuly Artificial Neural Network



c) The validation performance of the NNAugust Artificial Neural Network



d) The validation performance of the NNSummer Artificial Neural Network

Figure no. 2. The validation performance for the analyzed Artificial Neural Networks

The validation performance plot of the NNJune, NNJuly, NNAugust and NNSummer Artificial Neural Networks is described through the curves corresponding to the training, validation and testing processes. Analyzing the plot, we conclude that the best validation performance has been obtained at the 12th, 9th, 6th, 24th epoch, for the NNJune, NNJuly, NNAugust, respectively for the NNSummer ANNs. In these cases, the values of the Mean Squared Error were 0.065721 for NNJune, 0.039686 for NNJuly, 0.040314 for NNAugust and 0.063566 for NNSummer.

We have paid particular attention when studying the obtained results to the comparison of the test and validation curves. Thus, we have noticed that in the monthly approach cases of the July and August ANNs, the test curves increase significantly before the validation ones. This indicates that the monthly approach poses the risk of overfitting the data. The overfitting phenomenon may cause large prediction errors when one wants to predict new outputs using another input data set [10]. The forecasting is prone to be erroneous even if the neural network has been successfully trained, as the network is able to predict correctly the training set while producing significant errors when forecasting based on new input data samples.

As in the seasonal approach this phenomenon does not occur, we can conclude the seasonal approach offers the most reliable results. The results are justified by the fact that the number of samples available in the monthly approach for training the ANNs is considerably smaller when compared to the seasonal approach.

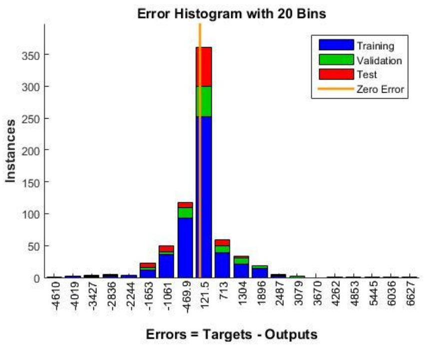
Subsequently, we have represented and analyzed the error histograms, through which we have highlighted the range of errors in the cases of the NNJune, NNJuly, NNAugust and NNSummer Artificial Neural Networks, developed using the Levenberg-Marquardt algorithm.

In the histogram plots, we have represented the testing data with red bars, the validation data with green bars and the training data with blue bars. In the case of the NNJune Artificial Neural Network, we notice that the interval in which most of the errors falls is [-1653,1896], in the case of the NNJuly ANN errors range between -1316 and 1192, in the case of the NNAugust ANN most of the errors falls in the interval [-1146,1290], while in the case of the NNSummer ANN errors range between -724.9 and 1311.

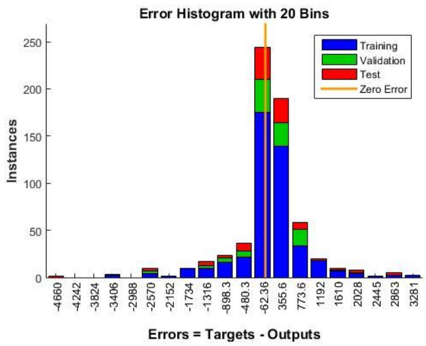
The smallest interval is the one obtained in the seasonal approach. One can observe that in the case of the NNJune ANN we have obtained a number of 350 out of

687 samples with errors around 0 (51%), in the case of the NNJuly ANN a number of 250 out of 639 samples with errors around 0 (39%), in the case of the NNAugust ANN a number of 450 out of 732 samples with errors around 0 (61%), while for the NNSummer Artificial Neural Network, 1200 out of the total of 2058 samples have errors around 0 (58%). From this point of view, we consider that the monthly approach is better for June and July than the seasonal approach, but due to the fact that it has a smaller range of errors, the seasonal approach is more desirable.

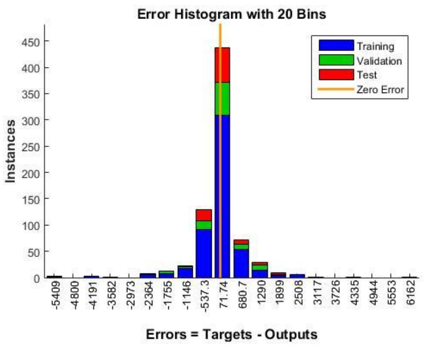
In both the monthly and seasonal approach we have noticed that, even if most of the errors fall in a certain interval (as we have discussed above), there remain some training points whose errors are outside this range. We have used a series of methods (rechecking data, retraining the networks) for reducing the number of such points, called outliers, samples for which the fitting is weaker than for the majority. Even if after applying these methods the outliers keep appearing, we have managed to reduce significantly their number (Figure no. 3).



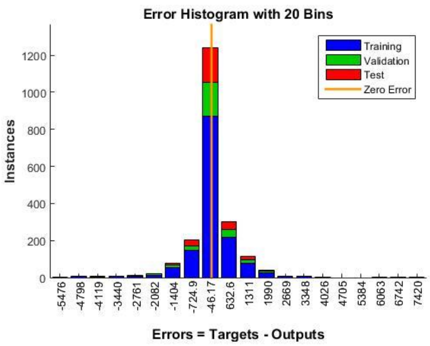
a) The error histogram of the NNJune Artificial Neural Network



b) The error histogram of the NNJuly Artificial Neural Network



c) The error histogram of the NNAugust Artificial Neural Network



d) The error histogram of the NNSummer Artificial Neural Network

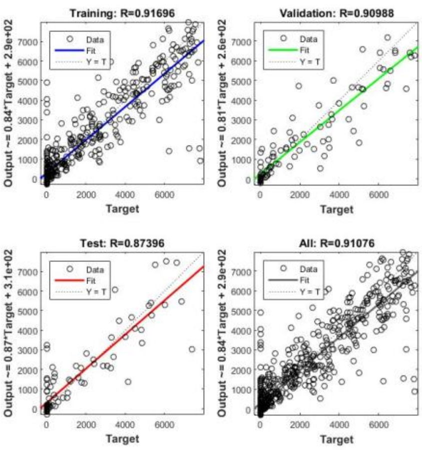
Figure no. 3. The error histogram for the analyzed Artificial Neural Networks

We have also represented the regression plots between the targets and network outputs, for all the developed ANNs. These plots are useful in validating the networks.

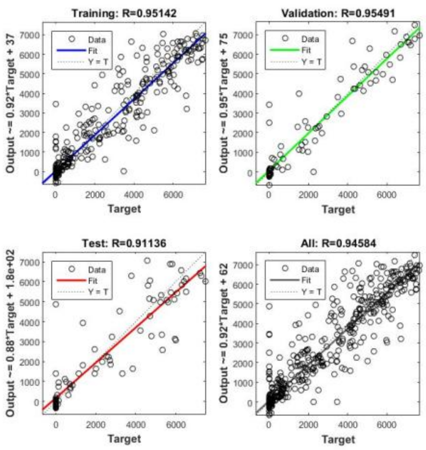
Through the regression plots, we have compared for each ANN the outputs and targets which, in an ideal case, should be the same. The correlation coefficient R gives information about the relationship between the outputs and targets and its degree of linearity. Thus, if R=0 the targets and outputs are not linked through a linear relationship, while the case when R=1 corresponds to the ideal case, when the targets and outputs coincide. For each of the four ANNs, we have represented the regression plots that correspond to the training phase, the validation step, the testing as well as to all the data samples. In these plots, we have represented the best-fit regression between the ANN targets and outputs with a solid line and the perfect, ideal fit, with a dashed line.

In all the cases, the values obtained for the correlation coefficient reflect a good fit. Thus, in the case of the NNJune Artificial Neural Network all the values of R are greater than or equal to 0.87396, in the case of the NNJuly Artificial Neural Network all the values of the correlation coefficient are greater than or equal to 0.91136, in the case of the NNAugust Artificial Neural Network all the values of R are greater than or equal to 0.93383, while in the case of the seasonal ANN, NNSummer, all the values of the correlation coefficient are greater than or equal to 0.90435.

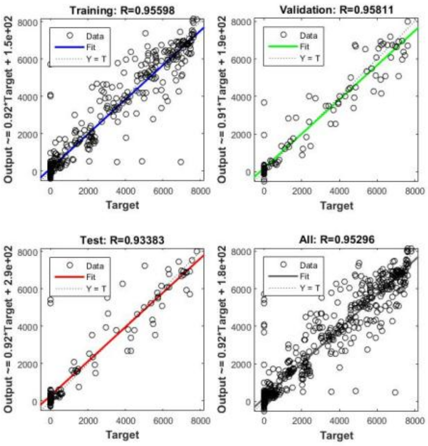
Regarding the values of the correlation coefficient R, the results provided by the July and August ANNs are improved compared to the seasonal approach. But, due to the fact that, as we have mentioned before, in the case of these months the overfitting process occurs, the best results remain those provided by the seasonal approach (Figure no. 4).



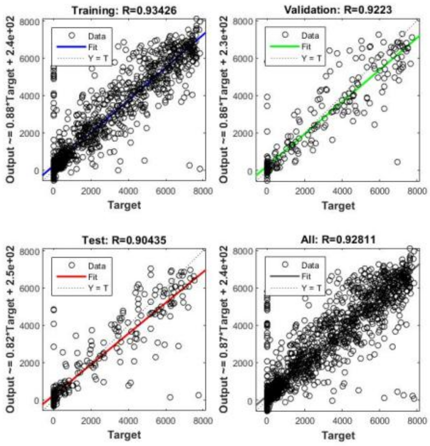
a) The regressions for the NNJune Artificial Neural Network



b) The regressions for the NNJuly Artificial Neural Network



c) The regressions for the NNAugust Artificial Neural Network



d) The regressions for the NNSummer Artificial Neural Network

Figure no. 4. The regressions between the network targets and network outputs for the analyzed Artificial Neural Networks

After analyzing the performance plots, we have synthesized the performance and the forecasting accuracy provided by all the monthly and seasonal ANNs (Table 1).

Table no. 1. A comparison analysis of the results provided by all the 16 monthly and seasonal ANNs

No.	The developed ANN	MSE	Range of errors	R is greater than or equal to	Remarks (risk of overfitting)
1	NNSpring	0.03258	[-680.2, 969.7]	0.95286	NNMarch
2	NNMarch	0.01762	[-950.6, 1256]	0.9546	
3	NNApril	0.025939	[-1116, 1280]	0.95779	
4	NNMay	0.069623	[-1734, 1591]	0.9077	
5	NNSummer	0.063566	[-724.9, 1311]	0.90435	NNJuly NNAugust
6	NNJune	0.065721	[-1653,1896]	0.87396	
7	NNJuly	0.039686	[-1316, 1192]	0.91136	
8	NNAugust	0.040314	[-1146,1290]	0.93383	
9	NNAutumn	0.044292	[-703.9, 1901]	0.90808	NNOctober
10	NNSeptember	0.050563	[-856.6, 2267]	0.92079	
11	NNOctober	0.04675	[-1549, 1541]	0.88734	
12	NNNovember	0.064865	[-941.7, 1973]	0.92290	
13	NNWinter	0.016752	[-404.6, 497.5]	0.91291	NNJanuary NNFebruary
14	NNDecember	0.037217	[-687.6, 686.1]	0.90965	
15	NNJanuary	0.036589	[-461.9, 913.6]	0.78025	
16	NNFebruary	0.03436	[-737, 429.3]	0.86860	

Analysing the results, we remark that the accuracy of the forecast is generally good in all the cases. Even if in some situations a few of the parameters registered by the monthly ANNs are slightly improved compared to the seasonal ones, we consider that the best results are those provided by the seasonal approach, due to the fact that in this case the overfitting process does not occur. In the case of the monthly approach, the process occurs for the ANNs mentioned in the last column of **Table 1**, thus in every season there is at least one month in which for the corresponding ANN the overfitting process occurs.

Comparing the results provided by the seasonal ANNs, we notice that the best value of the MSE coefficient and the narrowest range of errors is obtained for the NNWinter network, while the best value of the correlation coefficient R is registered by the NNSpring network. When comparing the recorded results corresponding to the monthly approach, we remark that the best value of the MSE coefficient is provided by the NNMarch network, the narrowest range of errors is provided by the NNFebruary network, while the value of the correlation coefficient R that is nearest to 1 is registered by the NNApril network. The ANN that offers the best overall forecasting results is the NNWinter network.

In order to facilitate the development of custom tailored solutions for a wide range of software applications that forecast the photovoltaic renewable energy production, we have integrated our developed Artificial Neural Networks into callable functions that we have compiled using the Matlab Compiler SDK, thus obtaining C/C++ shared libraries, .NET assemblies, Excel COM Add-ins and also JAVA packages. Therefore, our solution can be accessed by developers through multiple APIs (Application Programming Interfaces) when programming software that predicts the photovoltaic renewable energy production taking into account the seasonal particularities of the Romanian weather patterns.

3. Conclusions

The Artificial Neural Networks prove to be a versatile tool in predicting the photovoltaic renewable energy production, successfully managing the ever-changing nature of the weather patterns. We have developed and analyzed a monthly and seasonal approach in the purpose of forecasting the quantity of delivered energy when having access to the data measured from the panels' sensors: solar irradiation level, present module temperature, environmental temperature, atmospheric pressure and humidity.

Based on the obtained results, in the case of the Izvoru photovoltaic power plant we have noticed that the overall performance is better when considering the seasonal approach than in the case of the monthly one. In some cases, the monthly approach has the drawback of overfitting the data.

In Romania, it is mandatory for the renewable energy producers to provide to the ANRE (Romanian National Energy Regulatory Authority) daily estimations, as accurate as possible, regarding the quantity and parameters of the produced energy. The developed solutions can be a useful tool in this purpose and also for evaluating the potential of producing renewable photovoltaic energy in a specific geographic area, taking into account the weather patterns of Romania.

Acknowledgments

This paper presents a series of results obtained within the SIPAMER research project ("Sistem Inteligent pentru Predicţia, Analiza și Monitorizarea Indicatorilor de Performanţă a Proceselor Tehnologice şi de Afaceri în Domeniul Energiilor

Regenerabile"), PNII – "Parteneriate în domeniile prioritare", PCCA 2013, code 0996, no. 49/2014, financed by the National Authority for Scientific Research (NASR).

REFERENCES

1. Vera, I. A., Langlois, L. M., Rogner, H. H., Jalal, A. I., Toth, F. L. Indicators for sustainable energy development: An initiative by the International Atomic Energy Agency. Natural Resources Forum, 2005
2. Freris, L. L., Infield, D. G. Renewable energy in power systems, Chichester, U.K.: John Wiley & Sons, 2008
3. Luo, F. L., Ye, H. Renewable energy systems: Advanced conversion technologies and applications, CRC Press, 2012
4. Nassehzadeh, Sh.T., Behboodi, E., Aliyev, F.Q. Towards Renewability by Applying Solar Energy Technologies For Improved Life Cycle. International Journal on Technical and Physical Problems of Engineering (IJTPE), International Organization of IOTPE, 2012
5. Abuella, M., Chowdhury, B. Solar power forecasting using artificial neural networks. North American Power Symposium (NAPS), 2015
6. Yan, X., Abbes, D., & Francois, B. Solar radiation forecasting using artificial neural network for local power reserve. International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), 2014
7. Qazi, A., Fayaz, H., Wadi, A., Raj, R. G., Rahim, N., & Khan, W. A. The artificial neural network for solar radiation prediction and designing solar systems: A systematic literature review. Journal of Cleaner Production, 2015
8. Rana, M., Koprinska, I., Agelidis, V. G. Forecasting solar power generated by grid connected PV systems using ensembles of neural networks. International Joint Conference on Neural Networks (IJCNN), 2015
9. Lungu, I., Bâra, A., Caruțașu, G., Pîrjan, A., Oprea, S.V. Prediction Intelligent System in ohe Field of Renewable Energies Through Neural Networks. Economic Computation and Economic Cybernetics Studies and Research, 2016
10. *** https://www.mathworks.com/help/index.html
11. Nocedal, J., Wright, S. J. Numerical optimization, New York: Springer, 2006
12. Oprea, S.V., Pîrjan A., Lungu I., Fodor A. Forecasting solutions for photovoltaic power plants in Romania. Proceedings of The 15th International Conference on Informatics in Economy, 2016