ELECTRICITY CONSUMPTION ANALYSIS & ENERGY EFFICIENCY
www.elecon.ipp.pt

Regulatory context of smart grids in Europe and Brazil: current state and trends
Third ELECON Workshop

University of Grenoble Alps– Grenoble Polytechnic Institute – G2ELAB (Grenoble Electrical Engineering Laboratory), Grenoble, France

November 17-18, 2015

Solar Intensity Forecasting using Hybrid Neural Fuzzy Inference System

F. Silva[a], N. Teixeira[a], B. Teixeira[a], T. Pinto[a], I. Praça[a], Z. Vale[a,*], S. L. Abreu[b]

[a]*GECAD – Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, Institute of Engineering – Polytechnic of Porto (ISEP/IPP), Rua Dr. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal*

[b]*General – Alternative Energies Grooup - IFSC – Instituto Federal de Santa Catarina - Campus São José, Rua José Lino Kretzer, 608 - Praia Comprida - CEP 88103-310 - São José, Brazil*

Abstract

This paper presents a method based on Fuzzy Inference Systems (FIS) to forecast solar intensity. FIS are usually used to solve regression tasks in diverse application contexts. The Fuzzy Rule-Based Systems (FRBS) package is a library implemented in R language, which comprises several methods used to solve regression and classification tasks. The Hybrid Neural Fuzzy Inference System (HyFIS) is used in this work, and the forecasting results achieved by using this method are compared to the results obtained from the application of other widely used methods for regression problems, namely Support Vector Machines (SVM) and Artificial Neural Network (ANN). The evaluation of the achieved results using real data from Florianopolis, Brazil, shows that the performance of HyFIS is superior to that of ANN and SVM, thus providing encouraging indicators of the potential of this approach in solving the problem of forecasting of solar intensity.

Keywords: Artificial Neural Networks; Hybrid Neural Fuzzy Inference System; Solar Forecasting; Support Vector Machines

---

1. Introduction

Despite its importance for the existence of life on earth and human beings health, the sun is nowadays a source of clean energy and can contribute to reduce the difficulty in fulfilling the energy demand. Photovoltaic (PV) and solar thermal are the main sources of electricity generation from solar irradiance. In the case of solar thermal energy plants with storage energy system, its management and operation need reliable predictions of solar irradiance with the same temporal resolution as the temporal capacity of the back-up system [1]. The development in the power semiconductor technology has allowed higher efficiencies in the conversion of solar energy into electrical energy trough photovoltaic cells [2] and PV systems have reached the end-user. The spread of PV technology took place and nowadays is being used in several buildings to generate electricity.

The increase on the use of renewable energy sources (RES) affects the behavior of a considerable number of entities from the electricity sector and imposes economical and technical challenges. This is

* Corresponding author.
E-mail address: zav@isep.ipp.pt

mostly due to the distributed nature of RES and variability and unpredictability of generation. Solar energy is clearly the most abundant resource available to modern societies. Usually summer months, such as July and August in the northern hemisphere, have smaller variability. However, even during some sunshine months sudden changes might occur. The variability of the solar resource is mostly due to cloud cover variability and atmosphere conditions [1]. Forecasting renewable resources is, thereby, an important matter from the producers, retailers, aggregators, system operators and market regulators and operators' point of view [3].

Due to its particular characteristics, several approaches are usually used to forecast solar intensity, namely physical models [4], time series analysis [5], and other forecasting algorithms, such as reviewed in [6]. Despite the relevant developments that have been achieved so far, the amount of data that is nowadays available to be used by forecasting algorithms, together with the variability of the associated information, and the necessity for correlating different types of data from different sources, makes the most typically used approaches unable to cope with the current needs. In order to enable a breakthrough in the field, hybrid methodologies that combine the best features of different approaches, are arising as a promising solution.

This paper introduces a methodology based on Hybrid Neuro Fuzzy Inference System (HyFIS) [7] to solve the problem of solar intensity forecasting, by studying the fluctuations in solar intensity in different periods. With this approach it is possible to recognize patterns, which enable strategic support in reducing the inherent risk in the industry. Such tools are very important for entities that are especially dependent on unpredictable resources. To be able to assess the performance of the method, results of the proposed HyFIS methodology are compared with the results of other forecasting methods, namely based on Artificial Neural Networks (ANN) and Support Vector Machines (SVM), which have been presented and evaluated in [8]. The experimental findings consider a case study based on real solar data from a period of ten years, measured in Florianopolis, state of Santa Catarina, Brazil.

After this introductory section. the rest of the paper is organized as follows. In section 2, the used forecasting methodologies are described, namely Fuzzy Inference System (FIS), ANN and SVM. Section 3 introduces the proposed HyFIS methodology for solar intensity forecasting. Section 4 presents some experimental findings that enable evaluating the performance of the proposed HyFIS, using real solar intensity data from Florianopolis, Brazil. Finally Section 5 concludes the paper by providing a discussion on the efficiency of the HyFIS methodology and its comparison to the other considered forecasting approaches: ANN and SVM.

2. Solar Forecasting

Achieving solar forecasts with satisfactory results is a difficult task due to the diversity of climatic factors [6]. Forecasts are crucial to anticipate necessary actions, in order to maximize the energy production resources management. Using this data it is possible to recognize some patterns that allow the extraction of critical information, which are used to predict the solar intensity trends. These patterns are the needful background to support solar dependent producers in their decision making tasks. Solar forecasting can also be a determinant factor for investment in this area, since it can reduce the unpredictability of the environment, for industrial or small producers. Additionally, the investment in clean energy is becoming more relevant to the environmental balance, since the world is facing a resource scarcity problem, with special emphasis on fossil fuels [3].

Given these factors, several studies have been performed using forecasting methods, in order to identify the solar intensity trends throughout the days in different seasons. Among these methods, algorithms based on SVM and ANN achieved results with a satisfactory error rate.

ANNs are artificial networks with the ability of learning, training simulation and predicting data [9]. ANNs are composed by several layers, where the first layer receives input values, the last layer is the output layer which provides the results, and the intermediate layers are called hidden layers. These layers are responsible for detecting features in the input data. This technique works with three types of data: training data to adjust the model parameters; test data to test the model; and validating data to avoid training excess.

SVM [10] is a supervised learning algorithm used for classification and regression problems. SVM is based on decision boundary defined on decision planes for a set of objects of the same membership class. These are separated from others that belong to another set. The separation of planes has the purpose of finding unseen patterns and minimizing the classification error. For example, giving a set of training input x_i and the decision values y_i, between -1 and 1, the objective is finding the best separation of plane, resourcing to the equation $w^t x + b$ for minimizing the distance between the two class memberships.

FIS [11] are systems based on fuzzy logic, which resort to the fuzzy set theory to map input data to output data. In the mapping, the fuzzification process is applied over the input data and a set of IF-THEN fuzzy rules are created. The fuzzification process uses a membership function to convert data inputs into normalized values and create fuzzy sets. Fuzzy sets are groups of linguistic values with an associated numeric interval. Fuzzy IF-THEN rules aim to perform patterns identification and decision support to help solving real world problems. FIS also concider a defuzzification mechanism to convert fuzzy sets into the output values. FIS methods are a good tool in the context of RES forecasting, since they are designed to handle a large number of environment variables.

HyFIS is an improvement of regular FIS, which uses heuristic fuzzy logic rules and input-output fuzzy membership functions that can be optimally tuned from training examples by a hybrid learning scheme comprised of two phases: rule generation phase from data; and rule tuning phase using error backpropagation learning scheme for a neural fuzzy system [12]. HyFIS has already been applied in several works, and it is often used as a basis for developing other methods. In [13] the T2-HyFIS-Yager model is introduced, developed as improvement from HyFIS-Yager-gDIC. This technique combines the original HyFIS network with Gaussian Discrete Incremental clustering (gDIC) to create a more consistent and intuitive framework to emulate human reasoning in a decision-making mechanism. The T2-HyFIS-Yager is a type-2 hybrid neural fuzzy inference system realizing Yagar inference for learning and reasoning with data about corrupted noise. The proposal of T2-HyFIS-Yagar is applied in time-series forecasts to model the signal-to-noise ratio [13]. The Noise Model Creation (NMC) is presented in [14]. This work refers to learning techniques for noise patterns that use HyFIS in a learning engine. The Hybrid ARIMA-HyFIS Model is presented in [15], and considers a hybrid model based on Auto-Regressive Integrated Moving Average (ARIMA) models and HyFIS to tune and prevent univariate time series. The application of HyFIS to forecast solar intensity is presented in section 3.

3. Proposed HyFIS methodology

The HyFIS method has been proposed in [7], and implements an architecture based on the Mamdani model. This is a neuro-fuzzy method composed by a five layers neural network based on fuzzy systems. The knowledge acquisition scheme of HyFIS is composed by two phases named structure learning and parameters learning.

The structure learning uses the Wang and Mendel (WM) technique to generate fuzzy rules from numeric input data, which are read only once. This technique avoids the time spent in the learning process as it happens with the conventional neural networks.

The parameters learning uses a gradient descent based learning algorithm to tune the membership functions' parameters to achieve a good performance level. This phase is composed by a five layer network. Each layer is composed by several nodes which associate a part of the system. Layer 1 nodes receive a vector with crisp values. Layer 2 nodes contain the antecedent part of the IF-THEN fuzzy rules. In this layer, each node receives two crisp values from Layer 1 and uses them as parameters in a Gaussian membership function to convert it into linguistic variables. Layer 3 contains nodes where each represents a fuzzy rule with the t-norm operator AND. Nodes in this layer calculate the firing strength of each fuzzy rule. Each node in Layer 4 represents the consequent part of a fuzzy rule and performs the OR operation. In this layer the output values are represented by the Gaussian function. Layer 5 produces a vector with the output values.

In this work, the FRBS library of R programming language has been used. This library contains an implementation of several methods that apply the concept of fuzzy logic, which have been proposed in [16], to represent, handle and solve real world problems through reasoning representation extracted by human experts in a set of IF-THEN rules. FRBS methods combine various existing approaches, such as heuristic procedures, neuro-fuzzy techniques, clustering methods, generic algorithms and square methods. FRBS is represented by a universal framework named frbsPMML. PMML is a XML based language which provides a standard for description of models produced by data mining and learning algorithms. This facilitates the importation and exportation of data from a FRBS model to frbsPMML.

The arguments used by the implemented HyFIS are presented in Table 1. An explanation on how these arguments are used, and on their contribution to the outcomes of the forecasts are also provided.

Table 1. Arguments list and their description to configure HyFIS method

Arguments	Description
data.train	A matrix (m x n) with crisp values for the training process, where m (rows) is the number of instances of a data series, or a set of parameters. n (columns) is the number of variables inside each instance. The last column contains the result of the combination of the variables of each instance.
num.lables	A matrix (1 x n) where n is the number of linguistic variables and each one has associated a numeric interval with a classification.
max.iter	An integer value representing the maximum number of iterations, i.e. the number of cycles used in the training process. The default value is 10.
type.tnorm	A constant that represents the type of t-norm to be used. The valid values are inside the interval of [1..5]: • 1 or MIN, is the standard t-norm: min: (x1,x2); • 2 or HAMACHER Hamacher product: (x1+x2)/(x1+x2-x1 x2); • 3 or YAGER class Yager: 1-min: (1,((1-x1)+(1-x2))); • 4 or PRODUCT bounded product: (x1+x2-x1*x2); • 5 or BOUNDED delimited product: max (0,x1+x2-1).
type.snorm	A constant which represents the type of s-norm to be used. The valid values are inside the interval of [1..5]: • 1 or MAX standard s-norm: max(x1,x2); • 2 or HAMACHER Hamacher sum: (x1 + x2 - 2x1 * x2)/1 - x1 * x2); • 3 or YAGER Yager class: min(1,(x1 + x2)); • 4 or SUM sum: (x1 + x2 - x1 * x2); • 5 or BOUNDED buonded sum: min(1,x1 + x2).
type.defuz	A constant which represents what defuzzification method to be used: • 1 or WAM weighted average; • 2 or FIRST.MAX first maxima; • 3 or LAST.MAX last maxima; • 4 or MEAN.MAX mean maxima; • 5 or COG means modified center of gravity (COG).
type.implication.func	A value that represent the type of implication function: • DIENES_RESHER(b > 1 - a?b∶ 1 - a); • LUKASIEWICZ (b < a?1 - a + b∶ 1); • ZADEH (a < 0.5‖1 - a > b?1 - a∶ (a < b?a∶ b)); • GOGUEN (a < b?1∶ b/a); • GODEL (a <= b?1∶ b); • SHARP(a <= b?1∶ 0); • MIZUMOTO (1 - a + a * b); • DUBOIS_PRADE (b == 0?1 - a∶ (a == 1?b∶ 1)); • MIN (a < b?a∶ b).

The performance of forecasts is assessed using the Symmetric Mean Absolute Percent Error (SMAPE) error measurement method. SMAPE determines the performance of forecasts by calculating the mean of absolute error in percent. The SMAPE is calculated as shown in the equation (1), where A_t is the real value and F_t is the forecasted value.

$$SMAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|F_t - A_t|}{(A_t + F_t)/2} \tag{1}$$

4. Experimental Findings

This case study aims to verify the performance of HyFIS algorithm in order to ensure that it can overcome the already studied algorithms ANN and SVM. For this study were considered real data from Florianopolis, Santa Catarina, Brazil, for the period between 1990 and 2000. The data includes information about Global, Direct, Diffuse and Extra-terrestrial Irradiance, in W/m^2; temperature in ºC; humidity in %; and wind speed in m/s. More details in the used data can be found in [17].

The first step of this study is the optimization of the HyFIS algorithm through the identification of the most appropriate parameters for the desired type of forecast. For this purpose several forecasts were performed with different combinations of parameters. After sensitivity analysis, it was possible to conclude that the best parameters to use are the parameters presented in Table 2. In can also be concluded that most parameters have a minimal influence in the variation of the results, except for a few parameters, which significantly compromised the quality of results.

Table 2. HyFIS' arguments and respective values

Arguments	Value
num.labels	7
max.iter	10
step.size	0.01
type.tnorm	MIN
type.snorm	MAX
type.defuz	COG
type.implication.func	ZADEH

In the particular case of the max.iter parameter, which relates to the maximum number of iterations, the default value was the one with the best combination of results/runtime. With regard to the step.size value, its decrement did not make a significant impact in the results as it is only necessary to obtain accurate results with up to two decimal places and that is safeguarded using the value 0,01.

After determining the best configuration to be used by HyFIS, the period to simulate the three algorithms has been selected. The hourly solar intensity during a complete week has been selected as experimental target, specifically the week of 27/11/2000 to 03/12/2000.

The target of the forecasts is the value of Global Irradiance (GI). The training data has a training limit of 20 given that this value was responsible for the best forecasts for each algorithm, making it possible to test the algorithms with the same level of knowledge. The training data's input consists of the GI value in the four days preceding the day of output value in the same period. For example, if it is desired to forecast the GI value in the period 12 of day 27/11/2000, the first training case will consist of the GI value in the period 12 of 26/11/2000 as output, and the value of GI in the period 12 of the days between 22/11/2000 and 25/11/2000 as input.

Initially, another approach was tested, in which, the GI value in the last four periods prior to the period intended to be forecasted was considered instead of considering the GI value in the same period of the previous four days. However, this approach resulted in a higher error by the three algorithms.

Finally, after the algorithms optimization phase, the final forecasts have been executed, returning the results presented in Table 3, which shows the comparison of the forecasting error using SMAPE, between the HyFIS, the ANN and SVM; and Fig. 1, which shows the graphical comparison between the forecasting error values achieved by the three methods throughout the test week.

Table 3. SMAPE forecasting error for the HyFIS, ANN and SVM for the complete test week

Algorithm	Average Error
HyFIS	**14,78%**
ANN	**16,62%**
SVM	**18,86%**

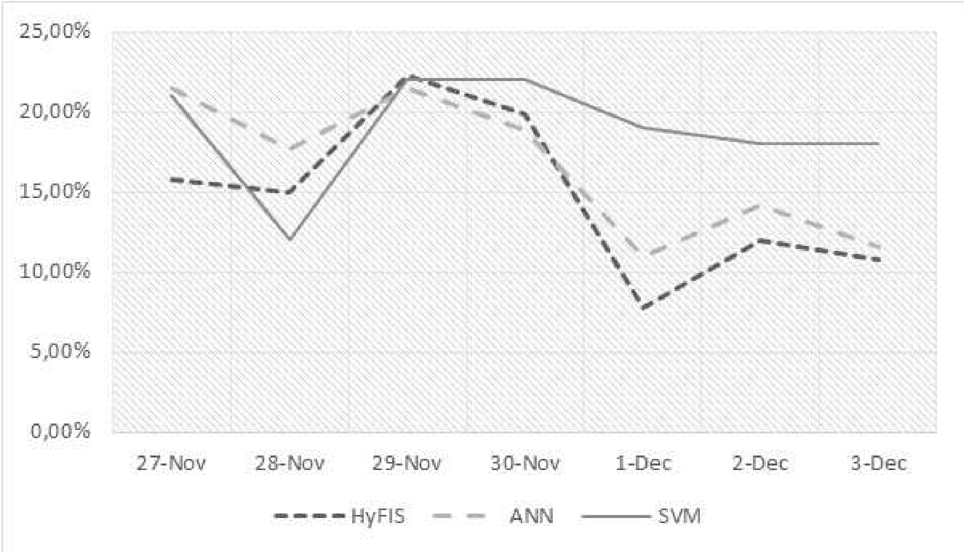

Fig. 1. Comparison between the daily average forecasting error achieved with HyFIS, ANN and SVM

Table 3 shows the average error of each algorithm throughout the test week. These results show that HyFIS is the algorithm with the smallest average percentage error, with 14.78%, followed by ANN with 16.62% (+ 1.84%) and finally SVM with 18.86% (+ 4.08%). The average daily SMAPE error of each algorithm can also be seen in Fig. 1, from which is visible that HyFIS presents lower error values in almost every day of the considered week (4 out of 7 days). The days when the HyFIS did not present the best result (28, 29 and 30 November) coincided with the days when the methods had a smaller difference between their forecasts, being the 29th the day on which the forecasts were closer. It is also possible to verify that the SVM algorithm has its biggest forecasting error, and greater distance from the other algorithms, in the last three days.

Fig. 2 shows the comparison between the real solar intensity values in all periods of day 1 and the values forecasted by the three algorithms. This comparison allows analyzing of how much distanced the different forecasts are from the real values.

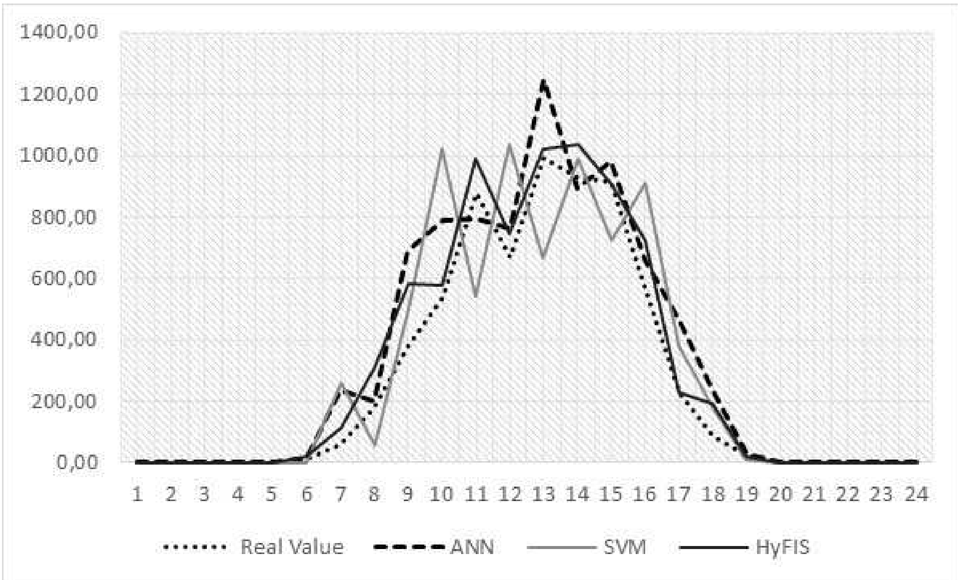

Fig. 2. Algorithms’ forecast results versus real data

As can be seen from Fig. 2, the real GI value in the considered day is greater than 0 between the 6th and 19th periods, which are the periods of sun's exposure. The trend of the GI value is to go up until the period 12nd/13rd and go back down until the 20th period when the sun goes down. By analyzing the figure it is possible to check why the HyFIS had the lowest forecasting error in this day, given that its line is the one that is closest to the real and being extremely close throughout most of the hours of the day.

From this case study, HyFIS has proven that it is possible to overcome the results obtained by ANN and SVM algorithms, whose efficiency has been proven in previous works. The errors for these last two algorithms is superior to HyFIS error, as presented in Table 3.

5. Conclusions

Achieving more efficient forecasting algorithms is a crucial factor for the solar energy production industries development. These algorithms are the main source of decision support systems, which can be used by producer entities to get the best conditions for the business, as well as for management entities, which need reliable forecasts to assure a correct operation.

The FIS methods are particularly effective prediction mechanisms when inserted in complex environments, in which there is a large number of variables, which makes them very useful tools in the solar branch prediction. The study presented in this paper is based on the application of an HyFIS, one of the most promising FIS methods, to the solar intensity forecasting problem.

The presented case study has shown that the difference between the forecasting error of the proposed HyFIS methodology and ANN is approximately 2%; and 4% between HyFIS and SVM, for the considered test week. Therefore, it is possible to identify that the HyFIS algorithm can overcome the ANN and SVM and take another step towards improving the proximity of forecasts to the real values. In addition, the HyFIS leaves a good indication that it can be able to distance itself more of ANN and SVM when performing the analysis with a greater time period in which the error tends to increase and consequently the difference between the algorithms as well.

As future work, other FIS methods can be studied and compared with these results. Namely, adaptive neuro-fuzzy inference system (ANFIS), dynamic evolving neural-fuzzy inference system (DENFIS), Wang and Mendel's fuzzy rule learning method, and genetic fuzzy rule-based systems under the iterative rule learning approach.

Acknowledgements

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under project ELECON - Electricity Consumption Analysis to Promote Energy Efficiency Considering Demand Response and Non-technical Losses, REA grant agreement No 318912 (PIRSES-GA-2012-318912).

References

1. L. Martin, L. Zarzalejo, J. Polo, A. Navarro, R. Marchante, M. Cony, "Prediction of global solar irradiance based on time series analysis: application to solar thermal power plants energy production planning", Solar Energy, 84 (2010), pp. 1772–1781.
2. R. Gupta, G. Gupta, D. Kastwar, A. Hussain, H. Ranjan, "Modeling and design of MPPT controler for a PV model using PSCAD/EMTDC". Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2010 IEEE PES, 11-13 October 2010.
3. F.P. Sioshansi, "Evolution of Global Electricity Markets – New paradigms, new challenges, new approaches", Academic Press, 2013
4. V. Badescu, "Modeling solar radiation at the earth surface", Springer (2008).
5. C. Paolik, C. Voyant, M. Muselli, M. Nivet, "Solar radiation forecasting using ad-hoc time series preprocessing and neural networks". In: Proceeding of the 5th international conference on emerging intelligent computing technology and applications, Ulsan, South Korea; 2009. p. 898–907.
6. H. R. Inman, H. T. C. Pedro, C. R. M. Coimbra, "Solar forecasting methods for renewable energy integration", Progress in Energy and Combustion Science, Elsevier, vol 39, pp. 535-576, 2013.
7. J. Kim, N. Kasabov "Hybrid Neuro-Fuzzy Inference Systems and their Application for On-line Adaptive Learning of

Nonlinear Dynamical Systems” University of Otago New Zeland (1999).

8. T. Pinto, G. Santos, L. Marques, T. M. Sousa, I. Praça, Z. Vale, S. L. Abreu: “Solar Intensity Characterization using Data-Mining to support Solar Forecasting”, Portugal (2015).
9. T. Pinto, et.al., "Dynamic Artificial Neural Network for Electricity Market Prices Forecast", IEEE 16th International Conference on Intelligent Engineering Systems, Portugal, 13-15 June, (2012).
10. R. Pereira, et.al., "Strategic Bidding for Electricity Markets Negotiation Using Support Vector Machines", in Advances in Intelligent Systems and Computing, vol. 293, pp. 9-17, (2014).
11. W. Yaïci, E. Entchev, “Adaptive Neuro-Fuzzy Inference System modelling for performance prediction of solar thermal energy system”, Renewable Energy, Volume 86, February 2016, Pages 302-315
12. J. Kim, N. Kasabov, HyFIS: “Adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems”, Neural Networks, Volume 12, Issue 9, November 1999, Pages 1301-1319
13. S.W. Tung, C. Quek, C. Guan, “T2-HyFIS-yager: Type 2 hybrid neural fuzzy inference system realizing yager inference”, IEEE International Conference on Korea (2009).
14. A. Sakai, M. Mitsuhashi, Y. Kuroda: “Noise Model Creation for Visual Odometry with Neural-Fuzzy Model”, IEEE/RSJ International Conference on Taiwan (2010).
15. S. A. Mangai, K. Subramanian, K. Alagarsamy: “Hybrid ARIMA-HyFIS Model for Forecasting Univariate Time Series”, International Journal of Computer Applications (0975 – 8887) Volume 91 India (2014).
16. L. S. Riza, C. Bergmeir, F. Herrera, J. M. Benítez.: “Learning from Data Using the R Package “frbs””, IEEE International Conference on China (2014).
17. S. L. Abreu, et.al., “Qualificação e Recuperação de Dados de Radiação Solar Medidos em Florianópolis – SC”, 8th Brazilian Congress of Thermal Engineering and Sciences, Porto Alegre, 2000