A Combined Methodology of Adaptive Neuro-Fuzzy Inference System and Genetic Algorithm for Short-term Energy Forecasting

Konstantinos KAMPOUROPOULOS[1], Fabio ANDRADE[1], Antoni GARCIA[2], Luis ROMERAL[2]
[1]*Fundació CTM Centre Tecnològic, 08242, Spain*
[2]*Electronic Engineering Department, Universitat Politècnica de Catalunya, 08222, Spain*
konstantinos.kampouropoulos@ctm.com.es

***Abstract*—This document presents an energy forecast methodology using Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithms (GA). The GA has been used for the selection of the training inputs of the ANFIS in order to minimize the training result error. The presented algorithm has been installed and it is being operating in an automotive manufacturing plant. It periodically communicates with the plant to obtain new information and update the database in order to improve its training results. Finally the obtained results of the algorithm are used in order to provide a short-term load forecasting for the different modeled consumption processes.**

***Index Terms*—adaptive neuro-fuzzy inference system, energy forecast, genetic algorithm, intelligent energy management systems.**

I. Introduction

With the continuously growing demand of the energy, it is getting more important to develop systems capable to optimize the energy use. Energy management is nowadays a subject of great importance because of the facing problems of the global warming and oil shortage. In the industrial sector, the energy management systems have focused so far on the monitoring and off-line management of energy as it is outlined in [1]. The typical energy management systems are based on the collection of information about the plant's operation using energy meters. Those systems help to monitor the operation of the installations, collect data and generate reports to identify the possible critical points of the consumptions. However, intelligent systems can improve the operation of the energy management systems, offering further functionalities such as predictive maintenance, energy optimization, fault diagnosis and energy forecast.

Different approaches for energy savings and energy prediction have been studied over the past years. Evolutionary algorithms such as Particle Swarms (PSO), Gravitational Search Algorithms (GSA) and Simulated Annealing (SA) have successfully implemented in optimization and control applications [2].

An implementation of a model based on Artificial Neural Network (ANN) was presented in [3-4] and [5] in order to estimate the load forecast in an electrical distribution system while in [6] a comparison between ANN and Fuzzy Logic is made on applications of short-term and medium-term load forecasting. An application of Neuronal Networks (NN) is presented in [7] in which it faces Multi-Input-Multi-Output (MIMO) applications with single input and output (SISO) networks. An ANFIS implementation for energy prediction of regional electrical loads in Taiwan was presented in [8], comparing its performance with other similar techniques (i.e., regression models, ANN-based models, Genetic algorithms and hybrid ellipsoidal fuzzy systems). A cellular multi-grid genetic algorithm is presented in [9] to face balancing problems in assembling lines. Techniques based on cultural algorithms are presented in [10] to resolve complex mechanical design optimization problems in an efficient and effective method.

This document presents the modeling and prediction algorithms that were developed in order to generate customizable mathematical models for different consumptions, as a way to improve the operation of a general energy monitoring system.

The paper is organized as follows: section II describes an overview of the algorithm that has been used for the model's training and the energy forecast while section III outlines a brief explanation about the Genetic Algorithm's operation. In section IV, the proposed methodology is presented explaining the combination of the two algorithms in order to develop a system capable to train the consumption models autonomously. In Section V, the implementation of the system in the pilot plant is explained, presenting the different results that have been obtained during the test and the evaluation of the system. Finally, section VI summarizes the paper and discusses the different conclusions.

II. Energy Prediction Using ANFIS

The training and prediction method that was used in this paper is based on Adaptive Neuro-Fuzzy Inference Systems. This method is based on systems of the type Takagi-Sugeno, combining fuzzy logic and neural networks. The simple forms of the artificial neural networks (ANNs) are not able to give any explicit knowledge or causal relationships for a system and this is their major disadvantage [11-12]. On 1993, Roger Jang [13] developed the ANFIS technique that could overcome the shortcoming of the ANNs and fuzzy systems.

The fuzzy part of the ANFIS is constructed by means of input and output variables, membership functions, fuzzy rules and inference method. The training inputs are also called energy drivers and are variables that can affect the output, such as, in case of the energy consumptions: the daily production, the climatic data, the day of the week, etc

Digital Object Identifier 10.4316/AECE.2014.01002

[14]. The membership functions of the system are the functions that define the fuzzy sets [15-16]. The fuzzy rules have a form of if-then rule and define how the output must be for a specific value of membership of its inputs. In general, the fuzzy systems have different kind of inference methods but ANFIS is based on a particular type of fuzzy system with Takagi-Sugeno rules as inference method [17].

The Fig. 1 shows the architecture of an ANFIS structure with two inputs, four if-else rules and one output. This structure has a maximum of four rules and they are defined in the equations (1)-(4).

$$if\ \ x \in A_1 \wedge B_1 \Rightarrow z_1 = p_1 x + q_1 y + r_1 \quad (1)$$

$$if\ \ x \in A_1 \wedge B_2 \Rightarrow z_2 = p_2 x + q_2 y + r_2 \quad (2)$$

$$if\ \ x \in A_2 \wedge B_1 \Rightarrow z_3 = p_3 x + q_3 y + r_3 \quad (3)$$

$$if\ \ x \in A_2 \wedge B_2 \Rightarrow z_4 = p_4 x + q_4 y + r_4 \quad (4)$$

The first part in the equation (1) is related to the antecedents and the second part of the equation to the consequents. During the training process of the ANFIS structure, these rules are executed in order to be calculated the output of the system. The first layer of the structure (Fig. 1) is called fuzzification. In the second layer, the weight of each rule has to be computed by means of a fuzzy AND operation. In the layer 3, it is made the normalization of the values and in the layer 4 the defuzzification process. Finally in layer 5, the overall output of the system is obtained [18].

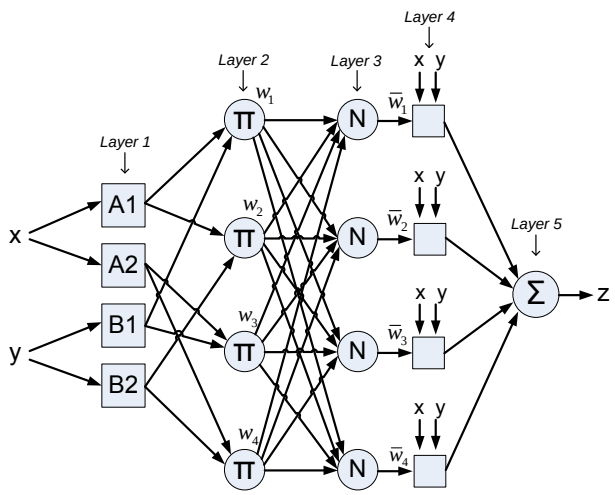

Figure 1. ANFIS architecture with two inputs, four rules and one output.

III. Genetic Algorithms

The Genetic Algorithms is a heuristic method, based on the mechanics of natural selection and natural genetics. They combine parts that present good solutions, in order to improve the final solution [19]. They work with a population of solutions rather than just one and they use selection, crossover and mutation operators to combine parts of parent chromosomes and produce offspring that intends to have the advantages of both parents. The operation of the genetic algorithm works as follows [20]:

a) The initial population is filled with individuals that are generally created at random. Sometimes, the individuals in the initial population are the solutions found by some method determined by the problem domain. In this case, the scope of the genetic algorithm is to obtain more accurate solutions.

b) Each individual in the current population is evaluated using the fitness measure.

c) If the termination criterion is met, the best solution is returned.

d) If it is not met, from the current population individuals are selected based on the previously computed fitness values using the selection function. A new population is formed by applying the genetic operators (crossover, mutation) to these individuals. The selected individuals are called parents and the resulting individuals offspring. Some implementations extend the current population by adding the new individuals and then create the new population by omitting the least fit individuals. Other implementations create a separate population of new individuals by applying the genetic operators. Moreover, there are GAs that do not use generations at all, but continuous replacement.

e) Actions starting from step 2 are repeated until the termination criterion is satisfied. An iteration is called generation [21-22].

In the proposed methodology, the genetic algorithm has a fundamental role in the efficiency of the training results. During the time, the consumption´s behavior may be changed, being dependent on different external parameters (e.g., temperature). Its operation as it is described in more details in the next chapter is to filter the training database of the consumption. Using a binary vector (GA´s chromosome), it evaluates the prediction result of the ANFIS looking for the best training input combination that results the minimum prediction error.

IV. IEMS Structure

This paper presents a short-term load forecasting methodology based on Adaptive Neuro-Fuzzy Inference System (ANFIS) and Genetic Algorithms. The presented methodology has been implemented in the existed energy management system of an automotive manufacturing plant in order to continuously acquire information from the plant. The objective of the system is the generation of the mathematical models of different type of consumptions of the manufactory plant, taking into account different external variables that can affect with a direct or an indirect way the load's energy profile [23] (e.g., climatic data, work calendar, production schedule, etc.).

The system is divided in two main blocks: one of them relates to the part of the training and auto-tuning of the model and the other corresponds to the prediction algorithm.

Training Model: The training part of the methodology is based on a system that acquires periodically information from the database of the plant, in order to analyze the data and detect possible energy pattern behaviors. The period of the tuning of the models depends on the production conditions and the influence of the climatic change in the energy consumption.

The methodology for the training process consists on the use of a genetic algorithm that selects the inputs of the ANFIS system and evaluates its result. The equation of the root mean square error (RMSE), defined on (5) has been selected as the evaluation criterion of the training result. The vector Pi_{real} includes the real values of the checking data set and the vector $Pi_{prediction}$ includes the corresponding prediction results.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_i^{real} - P_i^{prediction}\right)^2} \quad (5)$$

The dataflow of the modeling process of the consumptions is presented on Fig. 2.

Step 1: The first step of the dataflow is the initial configuration of the Genetic Algorithm. In this step, the population size, the crossover, the selection and the mutation functions are configured.

The mutation process is made of a uniform way, where the algorithm selects a fraction of the vector entries of an individual and replaces it with a random number from the entry´s range.

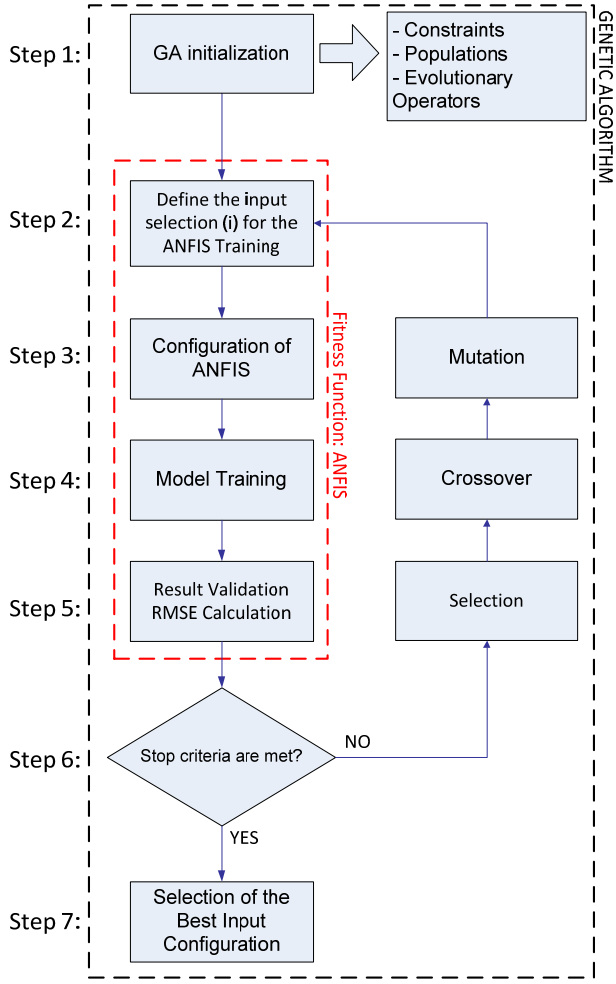

Figure 2. Dataflow of the modeling process.

For the crossover process, a random number of variables are selected from the vector´s entries in order to form the new chromosome.

During the selection process, random chromosomes are being selected in order to continue to the next population.

Step 2: During the second step, the filtering of the data input is made according to the selected chromosome of the population. The vectors that correspond to the 0 values of the chromosome are being filtered and they are not being used during the training of the model.

Step 3: In the step of the ANFIS configuration, the training and checking data are defined. The seventy percent of the input database is used for the model´s training and the rest thirty percent is used for the model validation. Furthermore, numbers of membership functions and rules are defined.

Step 4: During this step, the algorithm applies a combination of the least-squares method and the back-propagation gradient descent method in order to train the FIS membership function and emulate the given training data set.

Step 5: After the finalization of the training process, the checking data are evaluated applying an over fitting model validation and calculating the RMSE of the obtained results. The process between 2 to 5 is being repeated for the whole number of population.

Step 6: In the sixth step, the RMSE of the evaluated chromosome is stored and compared with the rest ones. In case that the optimization criteria have not been reached, the selection, the mutation and the crossover process are applied in order to define the new population that will be evaluated.

In the case that the individual with the smaller RMSE is obtained and the criteria have been met, the algorithm finishes giving as a result the binary vector and its corresponding RMSE value.

Step 7: Finally the de-codification of the chromosome vector is made, defining the optimal data input configuration for the training of the consumption model.

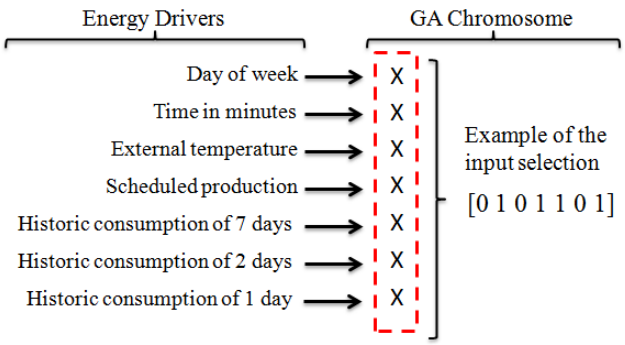

Figure 3. Explanation of the genes that combine the GA chromosome.

During the training process, the Genetic Algorithm forms a binary vector of seven digits with a total of 127 possible combinations. This vector defines the input parameters that are used during the training process of the model and can have binary values between 0000001 and 1111111. The Fig. 3 explains the parameters that correspond to each digit of the filter's vector.

Prediction Algorithm: Once the training process is done and the model inputs have been selected, the mathematical model of each consumption is calculated and stored in the plant's server. For the forecasting calculation, the ANFIS methodology is being used having as inputs the prediction values of the different training parameters that have been selected for each consumption model [24]. This process is being made manually by an operator when a short-term load forecasting is required. The user can carry out the prediction calculation by the use of a graphic user interface that permits to select the desired model and communicates with the server in order to detect the training inputs. After the introduction of the energy carrier's prediction data from the user, the ANFIS calculates the energy forecast for the selected consumption. The block diagram of the iEMS structure is presented in Fig. 4.

V. Implementation and Results

The presented algorithm has been tested during nine months under real operation conditions in the automotive factory plant in order to obtain real data information and analyze the behavior of the energy consumption along the time. In this chapter, the obtained results are presented and the system´s process explained.

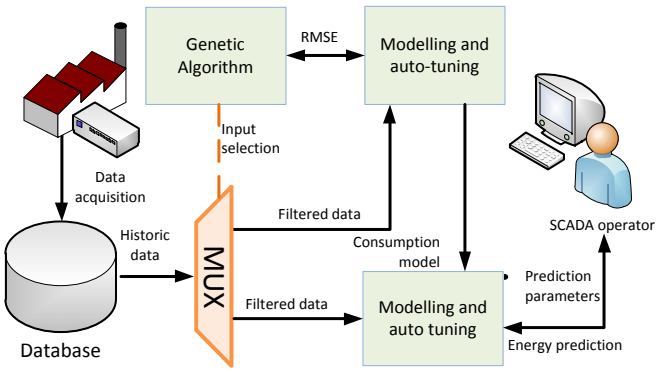

Figure 4. Structure of the proposed intelligent Energy Management System (iEMS).

As it was mentioned before, during the training process of the models, the genetic algorithm can take 127 different combinations of binary values for a vector of seven digits in order to define the dataset that will be used in order to train the model. During the evaluation of the algorithm, the training process has been executed in a total of nine times (one execution every month), completing each time its existing database with the dataset from the last month. The genetic algorithm was configured empirically after a serie of evaluations with a population of 20 individuals per iteration. The selection process uses a stochastic function, while in the mutation process the selection of the individual's fraction is made by the use of a uniform function. Finally the crossover process uses a scattered function to select the genes from the first and the second parent, and combine them to form the child.

Fig. 5 depicts the chromosome´s value of the genetic algorithm, which presents the minimum RMSE during the nine months of operation.

TABLE I. RMSE for the Different Periods of Training

Training Period	RMS error according to the input's chromosome (kW)		
	0111111	***0001101***	***0011100***
1 month	**144,621**	147,945	284,6075
2 months	107,811	**82,713**	93,6317
3 months	733,642	235,065	**197,9428**
4 months	10370	**184,054**	227,2427
5 months	240,632	192,155	188,4783
6 months	310,702	221,579	236,3428
7 months	217,714	195,132	13811
8 months	201,703	166,511	183,7469
9 months	150,113	138,329	173,5739
Training Period	**RMS error according to the input's chromosome (kW)**		
	1001101	***0011101***	***1011011***
1 month	238,710	152,512	396,215
2 months	86,821	152,512	113,967
3 months	270,270	238,347	312,021
4 months	19044	235,693	411,425
5 months	**162,172**	233,749	234,719
6 months	183,711	**150,546**	288,327
7 months	**164,198**	235,421	209,561
8 months	151,029	191,735	**145,732**
9 months	**120,155**	146,781	150,283

In order to evaluate the efficiency of the results of the presented algorithm, each selected chromosome was used to train all the different sets of data (dataset from one month till nine months) in order to calculate the root mean square that results in each case. Table I presents the results of this evaluation. The solution that present the minimum error for each training period has been marked in bold, indicating the optimal input configuration in each case.

From the table´s data, it can be observed that in all the cases the chromosome of the training data is changing during the time. This happens because the consumption´s behavior depends on different variables that vary during the time. Analyzing the genes of the chromosomes that are selected in each case (minimum RMSE), it results that the most important training inputs for the selected consumption is the scheduled production (9 times), the historic consumption of a day (8 times) and the historic consumption of a week (8 times). The following table highlights in detail the number of use of the different input vectors during the nine months of operation. The description and de-codification of the genes is explained in Fig. 3.

TABLE II. Use of the Training Parameters During the Time

Table Head	Genes of the GA´s chromosome						
	1	***2***	***3***	***4***	***5***	***6***	***7***
1 month	0	1	1	1	1	1	1
2 months	0	0	0	1	1	0	1
3 months	0	0	1	1	1	0	0
4 months	0	0	0	1	1	0	1
5 months	1	0	0	1	1	0	1
6 months	0	0	1	1	1	0	1
7 months	1	0	0	1	1	0	1
8 months	1	0	1	1	0	1	1
9 months	1	0	0	1	1	0	1
Times of selection	**4**	**1**	**4**	**9**	**8**	**2**	**8**

Fig. 5 depict the chromosome that was selected in each training period by the GA, while Fig. 6 presents its resulting RMS error.

Comparing the information of the Table I with the selection of the GA, it can be observed that the algorithm detects the changes that appear in the parameter´s correlations during the time and it selects in all the cases the

chromosome that present the minimum error.

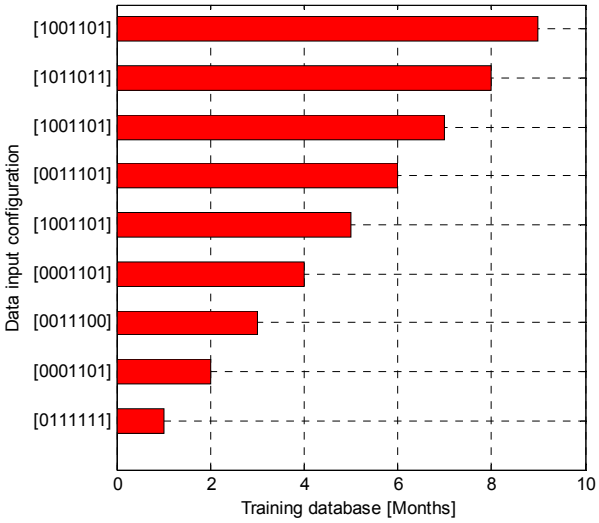

Figure 5. GA's selected chromosome for each training process.

Finally Fig. 6 to Fig. 9 depict a comparison between the real consumption values and the obtained prediction results for different prediction periods.

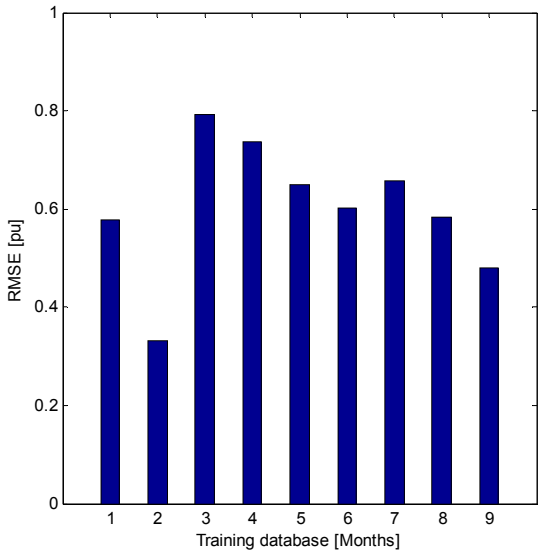

Figure 6. Resulting training RMSE for the different months.

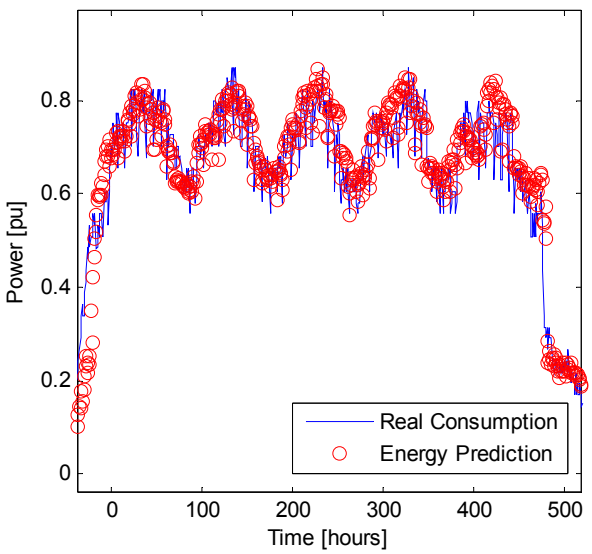

Figure 7. Comparison of the real consumption with the prediction result for a period of 1 month.

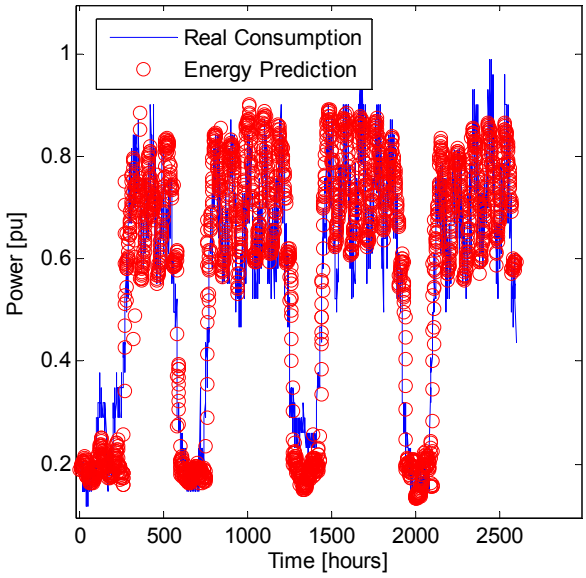

Figure 8. Comparison of the real consumption with the prediction result for a period of 3 months.

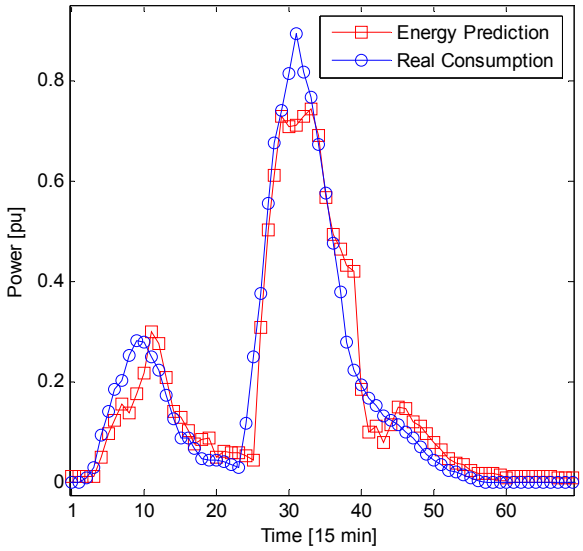

Figure 9. Comparison of the real consumption with the prediction result for a period of 1 day.

VI. Conclusion

This paper proposes a methodology for the modeling and energy prediction for consumptions using a combination of Genetic Algorithms and Adaptive Neuro-Fuzzy Inference Systems. The system is capable to run in an industrial plant, acquire information and detect possible correlations between the energy consumption and different external variables.

The methodology has been evaluated in an automotive factory plant using real consumption data.

The ANFIS is used in order to train the mathematical models and to provide a short-term load forecast while the GA is responsible to analyze the database and the possible correlations with the consumption and evaluate which parameters have to be used during the training and the prediction process.

Analyzing the real data it can be observed that in each training process during a specific period, the energy behavior of the load is changing. The use of the GA to evaluate the training data can offer an improvement in the

prediction results, as it detects modifications in the consumption's behavior and selects always the input dataset that presents the minimum RMSE. Therefore, the system keeps updated with the best possible accuracy, using the new available energy data.

An important advantage of this methodology is that it permits to model different types of consumptions, defining the vectors of the training database and the number of the genes that combine the GA's chromosome.

ACKNOWLEDGMENT

Authors would like to acknowledge the Maintenance department of SEAT for the excellent collaboration and help with the data providing plan. Also, the authors would like to thank their colleagues at Universitat Politècnica de Catalunya and specially to the Motion Control and Industrial Application (MCIA) research group for support and discussions during the preparation of this paper.

REFERENCES

[1] A. Tisot, "Industrial energy management: Doing more with less," Pulp & Paper-Canada, vol. 105, pp. 21-23, Oct 2004.

[2] R. E. Precup, R. C. David, E. M. Petriu, M.B. Radac, S. Preitl, J. Fodor, "Evolutionary optimization-based tuning of low-cost fuzzy controllers for servo systems," Knowledge-Based Systems, vol. 38, pp. 74-84, 2013.

[3] P. J. Santos, A. G. Martins, and A. J. Pires, "Designing the input vector to ANN-based models for short-term load forecast in electricity distribution systems," International Journal of Electrical Power & Energy Systems, vol. 29, May 2007.

[4] Z. Xiao, S. J. Ye, B. Zhong, C. X. Sun, "BP neural network with rough set for short term load forecasting," Expert Systems with Applications, vol. 36, pp. 273–279, 2009.

[5] S. Kouhi, F. Keynia, "A new cascade NN based method to short-term load forecast in deregulated electricity market," Energy Conversion and Management, vol. 71, pp. 76-83, 2013.

[6] A. Badri, Z. Ameli, A. M. Birjandi, "Application of artificial neural networks and fuzzy logic methods for short term load forecasting," 2nd International Conference on Advances in Energy Engineering, vol. 14, 2012.

[7] F. R. Fulginei, A. Laudani, A. Salvini, M. Parodi, "Automatic and parallel optimized learning for neural networks performing MIMO applications," Advances in Electrical and Computer Engineering, vol. 13, no. 1, pp. 3-12, 2013.

[8] L. C. Ying, M. C. Pan, "Using adaptive network based fuzzy inference system to forecast regional electricity loads," Energy Conversion and Management, vol. 49, 2008.

[9] O. Brudaru, D. Popovici, C. Copaceanu, "Cellular genetic algorithm with communicating grids for assembly line balancing problems," Advances in Electrical and Computer Engineering, vol. 10, no. 2, pp. 87-93, 2010.

[10] M. Z. Ali, K. Alkhatib, Y. Tashtoush, "Cultural algorithms: Emerging social structures for the solution of complex optimization problems," International Journal of Artificial Intelligence, vol. 11, no. A13, pp. 20-42, 2013.

[11] D. Wieland, F. Wotawa, G. Wotawa, "From neural networks to qualitative models in environmental engineering," Computer-Aided Civil and Infrastructure Engineering, vol. 17, pp. 104-118, 2002.

[12] A. Mellit, S. Saglam, S. A. Kalogirou, "Artificial neural network-based model for estimating the produced power of a photovoltaic module," Renewable Energy, vol. 60, pp. 71-78, 2013.

[13] J. S. R. Jang, "ANFIS - Adaptive-network-based fuzzy inference system," Ieee Transactions on Systems Man and Cybernetics, vol. 23, 1993.

[14] A. Khotanzad, E. Zhou, H. Elragal, "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment", IEEE Power Engineering Review, vol. 22, pp. 55, 2002.

[15] J. J. Cardenas, A. Garcia, J. L. Romeral, K. Kampouropoulos, "Evolutive ANFIS training for energy load profile forecast for an IEMS in an automated factory," The IEEE 2011 Conference on Emerging Technologies and Factory Automation, 2011.

[16] T. Takagi, M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," Ieee Transactions on Systems Man and Cybernetics, vol. 15, 1985.

[17] D. X. Niu, Y. L. Wang, D. Wu, "Power load forecasting using support vector machine and ant colony optimization," Expert Systems with Applications, vol. 37, pp. 2531–2539, 2010.

[18] K. Kampouropoulos, F. Andrade, J.J. Cárdenas, J.L. Romerar, "A Methodology for Energy Prediction and Optimization of a System based on the Energy Hub Concept using Particle Swarms," The Annual Seminar on Automation, Industrial Electronics and Instrumentation, 2012.

[19] R. Poli, W. B. Langdon, N. F. McPhee and J. R. Koza, "Genetic Programming An Introductory Tutorial and a Survey of Techniques and Applications," CES-475, ISSN: 1744-8050, 2007.

[20] G. Renner, A. Ekart, "Genetic algorithms in computer aided design," Computer-Aided Design, vol. 35, 2003.

[21] D. E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," pp. 60-76, Addison-Wesley Longman, 1989.

[22] E. C. Brown, R. T. Sumichrast, "Evaluating performance advantages of grouping genetic algorithms," Engineering Applications of Artificial Intelligence, vol. 18, 2005.

[23] J. Sheppard, A. Tisot, "Industrial energy management: Doing more with less", Industrial Energy Technology Conference, 2006.

[24] P. A. Gonzalez, J. M. Zamarreno, "Prediction of hourly energy consumption in buildings based on a feedback artificial neural network", Energy and Buildings, vol. 37, pp. 595–601, 2005.