

# Energy Consumption Forecasting using Genetic fuzzy rule-based systems based on MOGUL Learning Methodology

Aria Jozi<sup>1</sup>, Tiago Pinto<sup>2</sup>, Isabel Praça<sup>1</sup>, Francisco Silva<sup>1</sup>, Brigida Teixeira<sup>1</sup>, Zita Vale<sup>1</sup>

<sup>1</sup>GECAD research group, Polytechnic of Porto (ISEP/IPP), Porto, Portugal  
{arjoz, icp, fspsa, bccta, zav}@isep.ipp.pt

<sup>2</sup>BISITE research group, University of Salamanca, Spain  
tpinto@usal.es

**Abstract**—One of the most challenging tasks for energy domain stakeholders is to have a better preview of the electricity consumption. Having a more trustable expectation of electricity consumption can help minimizing the cost of electricity and also enable a better control on the electricity tariff. This paper presents a study using a Methodology to Obtain Genetic fuzzy rule-based systems Under the iterative rule Learning approach (MOGUL) methodology in order to have a better profile of the electricity consumption of the following hours. The proposed approach uses the electricity consumption of the past hours to forecast the consumption value for the following hours. Results from this study are compared to those of previous approaches, namely two fuzzy based systems: and several different approaches based on artificial neural networks. The comparison of the achieved results with those achieved by the previous approaches shows that this approach can calculate a more reliable value for the electricity consumption in the following hours, as it is able to achieve lower forecasting errors, and a less standard deviation of the forecasting error results.

**Keywords** — Electricity consumption, Forecasting, Fuzzy rule based methods, MOGUL

## I. INTRODUCTION

Electricity consumption forecasting is crucial to enable dealing with the new paradigm of consumers' active participation in the power and energy system [1]. The uncertainty related to the variability of consumption is associated to numerous factors, such as consumers' habits, the environmental temperature, luminosity, etc. Current forecasting methods are not suitable to deal with such a combination of input variables, with often highly variable influence on the outcomes – the actual energy consumption.

Time series forecasting is an attractive domain in the power and energy systems field, as it is essential to enable and adequate energy resources management. With the increase of renewable

generation with an intermittent nature, and the consequent need for the increase in consumers' flexibility, forecasting energy generation and consumption, in addition to other factors, such as market prices, environmental variables, among many others, becomes crucial [2]. In this scope, regression approaches and Artificial Intelligence (AI) based methods, such as Artificial Neural Networks (ANN) are the most common approaches.

Several studies have been published related to energy consumption forecast. An approach based on data mining to forecast electricity consumption of a region, based on the meteorological conditions has been presented in [3]. In this work, the data mining method is implemented using Artificial Neural Networks (ANN), and uses information about several variables such as temperature, humidity and public holiday on which electricity consumption depends, and the daily consumption values as the data.

In [4] the author presents a study about Forecasting electricity consumption using historical electricity consumption in South Africa based on three deferent approaches, namely, Autoregressive Moving Average (ARMA) model, Neural networks and Neuro-fuzzy models. The final results prove that the Neuro-fuzzy models outperformed the neural network and ARMA models in terms of accuracy. And in [5], where the study provides accurate and efficient energy forecasting tool based on Support Vector Machine Regression (SVMR) and the results shows that in this approach the structure of the training set has significant effect to the accuracy of the prediction. Also in [6] the author applies an ANN in order to forecast the consumption of an office building. Most of these works achieve acceptable results. However, the integration of correlated external variables, which can aid the forecast of energy consumption, such as the influence of temperature, brightness, among others, is still insufficiently studied.

This paper proposes the use of a Methodology to Obtain Genetic fuzzy rule-based systems Under the iterative rule Learning approach (MOGUL) methodology (GFS.FR.MOGUL) to forecast energy consumption of an office building. The objective of this study is to forecast a better profile of the energy consumption for the following hours. Results from the proposed approach are compared to those achieved in previous studies, using different techniques, namely two fuzzy based systems: A Hybrid Neural Fuzzy Inference System

---

The present work was done and funded in the scope of the following projects: European Union's Horizon 2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No 703689 (project ADAPT); EUREKA - ITEA2 Project FUSE-IT (ITEA-13023), Project GREEDI (ANIIP2020 17822), and has received funding from FEDER Funds through COMPETE program and from National Funds through FCT under the project UID/EEA/00760/2013

(HyFIS) [7], and the Wang and Mendel's Fuzzy Rule Learning Method (WM) [8]; and also several different approaches based on ANN [6]. The case study is based on data referring to the electricity consumption of a campus building of ISEP/GECAD.

After this introductory section, section II presents the formulation and explanation of the proposed approach, section III presents the achieved results and discusses their comparison to the results achieved by previous methods. Finally, section IV presents the most relevant conclusions and contributions of this work

## II. MATERIAL AND METHODS

This paper implements a genetic fuzzy system for fuzzy rule learning based on the MOGUL methodology (GFS.FR.MOGUL) to forecast the electricity consumption of an office building. The electricity consumption form building N of the Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD) research center located in ISEP/IPP, Porto, Portugal has been chosen to be used in this work. These data is collected and stored through SOICAM (SCADA Office Intelligent Context Awareness Management) [9], a system that is used to manage and simulate the GECAD campus microgrid. GFS.FR.MOGUL approach has been developed based on R programming language. The implementation details and results of this work are discussed and compared in the following sections.

### A. GFS.FR.MOGUL methodology

Genetic fuzzy systems for fuzzy rule learning based on the MOGUL methodology (GFS.FR.MOGUL) is a forecasting method that implements a genetic algorithm determining the structure of the fuzzy IF-THEN rules and the membership function parameters. Two general types of fuzzy IF-THEN rules are considered:

- Descriptive rules;
- Approximate/free semantic approaches.

In the first type the linguistic labels illustrate a real-world semantic and the linguistic labels are uniformly defined for all rules. In contrast, in the approximate approach there is any associated linguistic label.

Modeling a fuzzy IF-THEN rule on a chromosome which consists of the parameter values of the membership function shows that every rule has the own membership function values. A population contains many such generated chromosomes, based on the iterative rule learning approach (IRL). IRL means that the chromosomes will be generated one by one, taking into account the fitness value and covering factor, until there are sufficient chromosomes in the population. After having obtained the population, the genetic algorithm is started, using the genetic operators' selection, mutation, and crossover [10].

Many important statistical properties have to be considered by the Fuzzy Rule Base (FRB) in order to obtain an FRB System FRBS presenting good behavior [11]. In the Generated Fuzzy Rule Bases (GFRB) obtained from MOGUL, there will be consider the satisfaction of two of these statistical properties, which are:

- *completeness*
- *consistency*

As an inductive approach to building GFRBSs is considered, both properties will be based on the existence of a training data set,  $E_p$ , composed of  $p$  numerical input-output problem variable pairs. These examples will present the following structure (1):

$$e_l = (ex_1^l, ex_2^l, \dots, ex_n^l, ey^l), \quad l = 1, \dots, p \quad (1)$$

### 1) Completeness of a Fuzzy Rule Base

As it is explained in [12], it is clear that an FRBS must be always able to infer a proper output for all possible system input. This property is called  $\tau$ -completeness in the field of inductive learning and it may be mathematically formulated using a real value  $t$  by means of the following expressions (2), (3), (4) [13]:

$$C_R(e_l) = \bigcup_{i=1 \dots t} R_i(e_l) \geq \tau, \quad l = 1, \dots, p \quad (2)$$

$$[R_i(e_l) = * (A_i(ex^l), B_i(ey^l))] \quad (3)$$

$$A_i(ex^l) = * (A_{i1}(ex_1^l), \dots, A_{in}(ex_n^l)) \quad (4)$$

Where  $*$  is a  $t$ -norm, and  $R_i(e_l)$  is the *compatibility degree* between the rule  $R_i$  and the example  $e_l$ . Given an FRB composed of  $T$  fuzzy rules  $R_i$ , the *covering value* of an example  $e_l \in E_p$  is defined as (5):

$$CV_R(e_l) = \sum_{i=1}^T R_i(e_l) \quad (5)$$

And the following condition (6) is required:

$$CV_R(e_l) \geq \epsilon, \quad l = 1, \dots, p \quad (6)$$

A good FRB must satisfy both the conditions presented above, to verify the  $\tau$ -completeness property and to achieve an appropriate final covering value.

### 2) Consistency of a Fuzzy Rule Base

A generic set of IF-THEN rules is *consistent* if it does not contain contradictions. It is necessary to relax the consistency property to consider the fuzzy rule bases. And it will be done by means of the positive and negative examples concepts [14]. Equations (7) and (8) are the examples of the positive and negative set for the rule  $R_i$

- Positive:

$$E^+(R_i) = \left\{ e_l \in \frac{E_p}{R_i(e_l)} \geq 0 \right\} \quad (7)$$

- Negative:

$$(8)$$

$$E^-(R_i) = \left\{ e_i \in \frac{E_p}{R_i(e_i)} = 0 \text{ and } A_i(ex^t) > 0 \right\}$$

And by giving the value  $k \in [0,1]$  and equations (9) and (10):

$$n_{R_i}^+ = |E^+(R_i)| \quad (9)$$

$$n_{R_i}^- = |E^-(R_i)| \quad (10)$$

We get that:

$$R_i \text{ is } k\text{-consistent when } n_{R_i}^- \leq k \cdot n_{R_i}^+ \quad (11)$$

So, the way to incorporate the satisfaction of this property in the designed GFRBSs is to encourage the generation of  $k$ -consistent rules. Those rules not verifying this property will be penalized so as not to allow them to be in the FRB finally generated.

### B. Data base description

This study uses the SQL server of GECAD research center located in ISEP/IPP, Porto, Portugal. This server includes several databases with the electricity information of GECAD campus buildings. The database of building N of this research center has been used in this work. This building has five energy meters, and each of them stores the electricity consumption data of one specific part of the building with 10 seconds time interval.

A java based application has been developed in this implementation. Which collect the data from the SQL server and calculate the average of the total electricity consumption of the building N – ISEP/GECAD per each hour. This application also creates a new .csv file in a format that can be used as the input of the forecast method.

## III. RESULTS AND DISCUSSION

The objective of this work is to achieve a better forecast of the energy consumption for the following hour comparing to the results of the previous works. The GFS.FR.MOGUL method has been developed for this forecast. In this process, the GFS.FR.MOGUL method uses the electricity consumption from last 10 days before the hour which is meant to be forecasted to train the method and forecast the final result.

An application based on java programming language has been developed to collect the data from the GECAD SQL server database. This application receives the energy consumption of every energy meter from the database and calculate the total energy consumption of the building as well as the average consumption for every hour.

For every forecasted value the GFS.FR.MOGUL method receive a set of three tables as the input. which are: (i) Testing input, (ii) Training input, and (iii) Training output. The testing input table is the main input of the methods and includes the average of total energy consumption of the building for each hour from last 24 hours before the forecast hour. Table 1 shows a summarized testing input table in case of forecasting the 12hour of 16/3/2016.

Table 1 - summarized testing input table to forecast the 12hour of 16/3/2016

<b>Date</b>	15/3/2016	...	16/3/2016
<b>Hour</b>	12:00	...	11:00
<b>Energy Consumption (w)</b>	1898.15	...	3424.76

The other two tables are used in order to train the GFS.FR.MOGUL method and generate the FUZZY rules. The training input table has the same set of the data as the testing input table but from the last 10 days before the forecasting day. Also, the training output table includes the energy consumption values of the same hour of the target hour from last 10 days. By these two training tables, to forecast every value the GFS.FR.MOGUL method will be trained 10 times by the data from the last 10 days before the target day.

According to equation (1) in this case  $p$  is 10 and every set of data from training input table correspond a set of  $(ex_1^t, ex_2^t, \dots, ex_n^t)$  where every value from training output table is the  $ey^t$ .

In order to test this method, the energy consumption of a total of 12 hours, namely from 12:00 until 23:00 of 16/3/2016, have been forecasted. These hours are chosen because they refer to the hours of greatest activity in the building (higher variation in consumption), thus being the most interesting hours to be forecasted. The MAPE (Mean Absolute Percent Error) error calculation formula is used to compare the forecasted values and the real values of each hour.

Table 2 shows the forecasted value for these 12 hours by the GFS.FR.MOGUL method and the real consumption values of each hour, including the MAPE error for each of the forecasted values.

Table 2 - Results of forecasting the energy consumption of 12 hours of 16/3/2016 by GFS.FR.MOGUL method

<b>Hour</b>	<b>Real Consumption (W)</b>	<b>Forecasted consumption by GFS.FR.MOGUL</b>	<b>MAPE error (%)</b>
<b>12:00</b>	2247.844174	1956.43053	12.96%
<b>13:00</b>	2253.590524	2250.89165	0.12%
<b>14:00</b>	2449.511654	2362.65151	3.55%
<b>15:00</b>	2676.168089	2310.858353	13.65%
<b>16:00</b>	2580.225047	2154.66557	16.49%
<b>17:00</b>	2017.849737	2225.699199	10.30%
<b>18:00</b>	1888.344068	1958.06592	3.69%
<b>19:00</b>	1624.601199	1876.106348	15.48%
<b>20:00</b>	1271.71025	1482.189028	16.55%
<b>21:00</b>	1253.198825	1236.9543	1.30%
<b>22:00</b>	1257.133005	1125.4727	10.47%
<b>23:00</b>	1228.879333	1350.27417	9.88%

During these 12 hours, the largest error belongs to the hour 20 by 16.55% and the lowest error is 0.12% for the hour 13. And also, the average error of these forecasted values is 9.54%.

Many works have been published about forecasting the electricity consumption based on fuzzy rules methods. The work presented in [7] studies about the forecasting of electricity consumption based on the Hybrid Neural Fuzzy Interface System (HyFIS), and in the [8] the electricity consumption has been forecasted by Wang and Mendel's Fuzzy Rule Learning Method (WM). Both of these works forecast the electricity consumption of an office building in a total of 12 hours of an official day. The average MAPE error of these 12 hours for the HyFIS method is 12.42% and this average for the WM method is 18.41%. Which shows that the GFS.FR.MOGUL method by the average of 9.54% has a better performance than the other two fuzzy rule based methods.

Figure 1 shows a comparison between the MAPE errors of the forecasted values by these three methods.

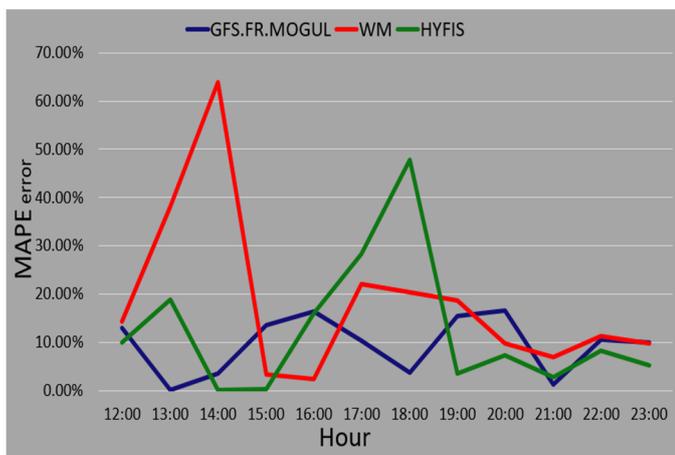


Fig. 1. Comparison between the results of HyFIS [7], WM [8] and the proposed GFS.FR.MOGUL

As Figure 1 shows, the forecasting error of GFS.FR.MOGUL presents less variation than the other two methods. This means that the results of the GFS.FR.MOGUL method during these hours are more reliable than the other results. Table 3 shows the value of the Standard deviation between errors of these three forecast methods.

Table 3 - Standard deviation between errors of HyFIS , WM and GFS.FR.MOGUL

Forecast Method	Forecast Error Standard Deviation (%)
HyFIS	13.90%
WM	17.32%
GFS.FR.MOGUL	5.95%

As Table 3 presents, the GFS.FR.MOGUL method has the lowest Standard deviation between his errors, compare to other two methods. It proves that the possibility of having a large error by the GFS.FR.MOGUL is less than the HyFIS and WM and it makes this method more trustable.

The study presented in [6] addresses the electricity consumption forecast using a set of different Artificial Neural Network (ANN) based approaches, which are summarized in Table 4.

Table 4- Different ANN neurons input / output definition[6]

Id ANN	Input neurons		Output neurons	
	No.	Description	No.	Description
ANN_1	3	- Context data*	1	- Total consumption per hour
ANN_2	3	- Context data	3	- Consumption of HVAC; - Consumption of lights; - Consumption of sockets
ANN_h	4	- Context data; - External environmental data**	1	- Consumption of HVAC
ANN_l	4	- Context data; - External environmental data	1	- Consumption of light
ANN_s	4	- Context data; - External environmental data	1	- Consumption of sockets

\* The context data refer to: hour; day of the week and hour of work (the hour of work is characterized as: time of high, medium and low permanency in the building);

\*\* The external environment data refers to the following variables: temperature, temperature felt, radiation, precipitation and humidity.

Figure 2 shows the electricity consumption forecasting error results achieved by the several ANN approaches (from [6]), the forecasting results achieved by HyFIS and WN (from [7] and [8]) and the results of the GFS.FR.MOGUL method presented in this paper.

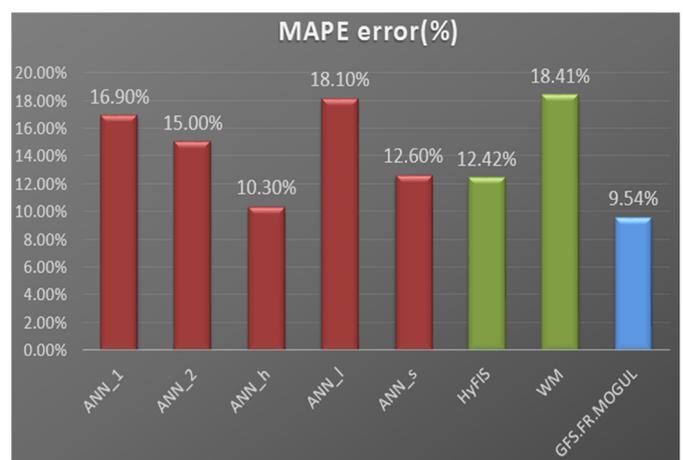


Fig. 2. Average forecasting errors of the HyFIS [7], WM [8], GFS.FR.MOGUL and several ANN approaches [6].

The comparison between the fuzzy rule based methods and several ANN approaches in Figure 2 shows that the GFS.FR.MOGUL method has the lower average forecasting error between these methods and can achieve better forecasting results than the other methods.

#### IV. CONCLUSIONS

This paper addresses the use of a genetic fuzzy systems for fuzzy rule learning based on the MOGUL methodology to forecast the electricity consumption of an office building in the following hours. This method uses the value of the electricity consumption from last hours to preview this value for the next hour.

By comparing the results of the GFS.FR.MOGUL method to the results of some other previous works it is concluded that this method can present a better electricity consumption profile for the following hours. The presented results in this paper in comparison to the results of some other fuzzy rule based methods, namely HyFIS and WM, are more reliable and all calculated errors are closer to the average error. In other methods, there is usually a large difference between some calculated errors and the average error of the method. These results show that the GFS.FR.MOGUL method is more trustable than the other fuzzy rule based methods. In addition, these results have been compared to the results of some different ANN approaches. The average error of the GFS.FR.MOGUL method in comparison to the average errors of the ANN approach is much lower, which makes the GFS.FR.MOGUL method a more trustable method to forecast the electricity consumption.

#### REFERENCES

- [1] Blumsack S, Fernandez A. Ready or not, here comes the smart grid! *Energy*. 2012; 37(1):61-8
- [2] Conejo, A.J., Carrión, M., Morales, J.M. (Eds.), "Decision Making Under Uncertainty in Electricity Markets", *Journal of the Operational Research Society*, Springer, 66, 174, December 2014
- [3] N. Kaur and A. Kaur, "Predictive modelling approach to data mining for forecasting electricity consumption," 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence), Noida, 2016, pp. 331-336.
- [4] L. Marwala and B. Twala, "Forecasting electricity consumption in South Africa: ARMA, neural networks and neuro-fuzzy systems," 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, 2014, pp. 3049-3055.
- [5] N. G. I. S. Samarawickrama, K. T. M. U. Hemapala and A. G. B. P. Jayasekara, "Support Vector Machine Regression for forecasting electricity demand for large commercial buildings by using kernel parameter and storage effect," 2016 Moratuwa Engineering Research Conference (MERCon), Moratuwa, 2016, pp. 162-167.
- [6] E. Vinagre, L. Gomes and Z. Vale, "Electrical Energy Consumption Forecast Using External Facility Data," *Computational Intelligence*, 2015 IEEE Symposium Series on, Cape Town, 2015, pp. 659-664.
- [7] Aria Jozi, Tiago Pinto, Isabel Praça, Francisco Silva, Brigida Teixeira, Zita Vale, "Energy Consumption Forecasting based on Hybrid Neural Fuzzy Inference System"
- [8] Aria Jozi, Tiago Pinto, Isabel Praça, Francisco Silva, Brigida Teixeira, Zita Vale, "Wang and Mendel's Fuzzy Rule Learning Method for Energy Consumption Forecasting considering the Influence of Environmental Temperature"
- [9] Luis Gomes, Pedro Faria, Hugo Morais, Zita Vale, Carlos Ramos, "Distributed, Agent-Based Intelligent System for Demand Response Program Simulation in Smart Grids", *IEEE Intelligent Systems*, vol. 29, pp. 56-65, January/February 2014
- [10] Lala Septem Riza, Christoph Bergmeir, Francisco Herrera, and Jose Manuel Benitez, *Fuzzy Rule-Based Systems for Classification and Regression Tasks*.
- [11] Driankov, D.; Hellendoorn, H.; Reinfrank, M. *An Introduction to Fuzzy Control*; Springer-Verlag: Berlin, 1993.
- [12] O. Cordon, M.J. del Jesus, F. Herrera, and M. Lozano, "MOGUL: A methodology to obtain genetic fuzzy rule-based systems under the iterative rule learning approach", *International Journal of Intelligent Systems*, vol. 14, pp. 1123 - 1153 (1999).
- [13] Herrera, F.; Lozano, M.; Verdegay, J. L. *Fuzzy Sets and Systems* 1998, 100, 143-158.
- [14] Gonzalez, A.; Perez, R. *Fuzzy Sets and Systems* 1998, 96, 37-51.