GA-ANN Short-Term Electricity Load Forecasting

Joaquim L. Viegas, Susana M. Vieira, Rui Melício, Victor F. Mendes, João C. Sousa

GA-ANN Short-Term Electricity Load Forecasting

Joaquim L. Viegas[1], Susana M. Vieira[1], Rui Melício[1,2],
Victor M.F. Mendes[2,3], João M.C. Sousa[1]

[1]IDMEC, LAETA, Instituto Superior Técnico, Universidade de Lisboa, Portugal
[2]Dep. Física, Escola de Ciências e Tecnologia, Universidade de Évora, Portugal
[3]Instituto Superior de Engenharia de Lisboa

joaquim.viegas@tecnico.ulisboa.pt; ruimelicio@gmail.com

Abstract. This paper presents a methodology for short-term load forecasting based on genetic algorithm feature selection and artificial neural network modeling. A feedforward artificial neural network is used to model the 24-hour ahead load based on past consumption, weather and stock index data. A genetic algorithm is used in order to find the best subset of variables for modeling. Three datasets of different geographical locations, encompassing areas of different dimensions with distinct load profiles are used in order to evaluate the methodology. The developed approach was found to generate models achieving a minimum mean average percentage error under 2%. The feature selection algorithm was able to significantly reduce the number of used features and increase the accuracy of the models.

Keywords: load forecasting, genetic algorithm, feature selection, artificial neural networks

1 Introduction

Energy is of great importance for the functioning and development of all activities in any nation. If the current pattern of energy consumption continues, at a global level, the total world demand will increase by more than 50% before 2030 [1]. Meanwhile, the high demand for energy and its production causes degradation of the environment as most energy resources are non-renewable [2]. Allied to the fact that big changes are currently happening to the in the utility industry due to deregulation and an increase in competition, policy makers and utilities are continuously seeking to identify ways to increase energy efficiency and alternate energy sources [3].

The decision making in the energy sector has to be based in accurate forecasts of energy demand, making forecasts one of the most important tools for utilities and decision makers in the energy sector. Forecasts of different time horizons and different scopes are essential for operation of plants and of the whole power system, "system response follows closely the load requirement" [4,5].

Short-term time horizon load forecasting is usually used for the one-day ahead forecasting and has a strong influence on the operation of electricity utilities. Many decisions depend on this type of forecast, namely scheduling of fuel purchases, scheduling of power generation, planning of energy transactions and assessment of system safety [6].

The load forecast can be related in complex and non-linear ways with various variables such as the past consumption pattern, the season of the year, climatic

conditions and others. Several methods to model these relationships have been applied in the past such as regression, econometric, time series, decomposition, co-integration, ARIMA, artificial intelligence, fuzzy and support vector models [2,7]. In the work of Lin et al. [8,9], a stock exchange index (TAIEX) was used in order to better the performance of short-term load forecasting (STLF) at times of global economic downturn.

As intuition tells us that many different variables are correlated to the load patterns intended to be predicted, feature selection (FS) becomes increasingly important in order to improve the forecasting performance, provide faster, more scalable models and provide an understanding of the influence of the variables on the future load [10]. Genetic algorithm (GA) is a search metaheuristic inspired in the process of natural evolution [11] that can be used for the selection of the best feature subset to be used for forecasting [12].

This paper proposes the use of artificial neural networks (ANN) together with GA feature selection in order to model future load using the best possible subset of available variables. Stock index variables are used in the study, together with weather and past load data, with the intent of finding a relationship between financial markets behavior and electricity demand. The rest of this paper is organized as follows: Section 2 presents a relationship of the paper to cyber-physical systems. Section 3 deals with the applied load forecast methodology. Section 4 presents the case studies for experimental evaluation. Finally, concluding remarks are given in Section 5.

2   Relationship to Cyber-Physical Systems

The smart grid is a concept with the purpose of intelligently integrating the generation, transmission and consumption of electricity through technological means [13,14]. Cyber-physical systems (CPS) are smart systems with physical and computational components, seamlessly integrated and interacting to sense the changing state of the world [15]. Meters, transformers, switchgear, lines, production plants are incorporating a growing number of automatic control components with the need to effectively sense the changes in the environment. Predicting the future state of the power grid, with the help of the proposed forecasting method, can potentially increase the effectiveness of CPS systems and reduce the complexity of the smart grid challenge by enabling prescriptive capabilities in order to give to the systems the ability of acting on change before it happens.

3   GA-ANN Load Forecasting Methodology

3.1   Feedforward Neural Network

ANN models were initially inspired by studies on brain modeling. This type of model generates a non-linear mapping from R^I to R^K; I and K are the dimension of the input and target space [11].

In this work, a feedforward neural network (FFNN) is used, consisting of three layers: an input layer, a hidden layer and an output layer. The input layer is composed of a number of neurons equal to the number of inputs and each of the artificial neurons has a unique input. The hidden layer neurons are fully connected to both the input and output layer and the output layer is composed by a number of neurons equal to the number of outputs. The activation function used for the input and hidden layers is the *tansig* function and, the output layer makes use a linear or function. FFNN have been widely used for load forecasting with success [2,16,17] due their ease of application with inputs from different sources and good performance.

In the proposed methodology the parameters of the model, such as the number neurons of the hidden layer, are optimized using grid search.

3.2 Feature Selection

The objective of FS is to find a subset of the available features by eliminating features with seemingly little or no information useful for prediction and also redundant features that are correlated [18]. FS techniques are usually divided into filter, embedded and wrapper methods. Wrapper and embedded are usually referred to as model-based methods and filter techniques as model-free methods.

Filter techniques assess the features by looking only at the intrinsic properties of the data [12]. Wrapper methods include the classification model within the feature subset optimization. The selected set of features results from training and testing a specific model, rendering this approach tailored to a specific modeling method.

3.3 Genetic Algorithm

This works makes use of GA for wrapper feature selection. This evolution inspired optimization algorithm starts with a random population of individuals (solutions). The main driving operators of a GA are selection (survival of the fittest) and recombination through crossover (reproduction), mutation is also used in order to escape local minima [11]. The typical flow of GA process is shown in Fig. 1.
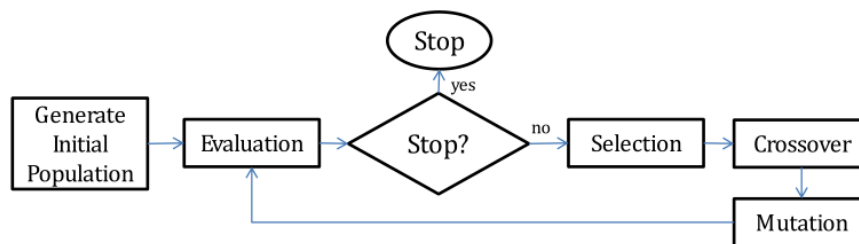
Fig. 1. GA optimization process.

The algorithm starts by generating a random population of individuals, binary encoded with dimension equal to the initial number of features, each chromosome representing if a feature is used in the subset represented by the individual [19]. The

individuals are evaluated using the ANN modeling method and depending on the performance of the individuals, they are selected for the creation of the next generation. The selected individuals are then used to generate new solutions using the crossover operator and mutation is used in order to escape local minima. The selection, crossover and mutation operators used in this work are presented in
Section 4.4.

4   Experimental Evaluation

4.1   Datasets

Three sets with different load patterns and available variables are used in order to better evaluate the proposed methodology. The first dataset is relative to Portugal, the second one to New York West and the third to the city of Rio de Janeiro. The data used is the following:

- **Portugal (PT)**: The historical load data is obtained from the European network of transmission system operators for electricity (ENTSOE) [20], the hourly load data is obtained between and including the days 2010-01-01 and 2013-12-31. The weather data is obtained from the Lisbon Airport Weather Station. The stock data used is from the Portuguese Stock Index 20 (PSI20) which tracks the prices of the twenty listings with the largest market capitalization and share turnover in Portugal [21].
- **New York West (NYW)**: The historical load data was obtained from NYISO [22], the hourly load data was obtained for the same period as the data from Portugal. The weather data was obtained from the Buffalo Weather Station. The stock data used in this case is from Standard & Poor's 500 (S&P 500), which is based on the capitalizations of the 500 largest companies having stock listed on the NYSE or NASDAQ [21].
- **Rio de Janeiro (RIO)**: The historical load and temperature data is the same used in the article of Hippert and Taylor (2010) [16] for the years of 1996 and 1997. No stock data was used with this dataset.

The pre-processing consisted in filling the missing values of the various time series by extending the last existing value. The data was processed for modeling and the input and output vector generated in the way presented by Table 1 and

Table 2. The use of the 24 prior hours and 7 days temperatures was inspired by the work of Sheikhan and Mohammadi [17] and the use of the mean of stock variables follows the work of Lin et al. [8,9]. The stock index i days mean is the average value of the stock index value during the i days before the day for which the loads are predicted.

To validate the proposed methodology, datasets encompassing different geographies with significantly different electricity consumption dynamics were selected. The datasets present the necessary characteristics to test the modelling approach and have not been extensively explored in the literature.

Table 1. List of inputs and outputs for the Portugal and New York West datasets

Outputs
$h + i$ hour's load ($i = 0, \ldots, 23$)
Inputs
$h - i$ hour's history load ($i = 1, \ldots, 24$)
$d - i$ day's maximum temperature ($i = 0, \ldots, 6$)
$d - i$ day's minimum temperature ($i = 0, \ldots, 6$)
Month
Season
Holiday Indicator
Weekday Indicator
Stock index i days mean ($i = 7, 15, 30, 60, 90$)
Day's precipitation

Table 2. List of inputs and outputs for the Rio de Janeiro dataset

Outputs
$h + i$ hour's load ($i = 0, \ldots, 23$)
Inputs
$h - i$ hour's history load ($i = 1, \ldots, 24$)
$d - i$ day's mean temperature ($i = 0, \ldots, 6$)
Month
Season
Holiday Indicator
Weekday Indicator

4.2   Training and Test Data

The data was divided into training and test sets in the following way: for the PT and NYW sets the first two years are used for training and the third for testing; for the RIO dataset, due to lack of data, one year is used for training and 266 days are used for testing.

4.3   Performance Evaluation

This work uses the mean absolute percentage error (MAPE) and normalized root-mean-square error (NRMSE) as performance evaluation metrics. The NRMSE is normalized using the maximum and minimum load values. The MAPE is the percentage based index of indicating forecasting accuracy level and is widely used in the field of electricity load forecasting and is calculated by:

$$MAPE\% = \frac{1}{n}\sum_{i=1}^{n}\frac{ol_i - pl_i}{ol_i} \times 100 \qquad (1)$$

where n is the total number of hours predicted by system, ol_i is the observed load and pl_i the predicted load.

The parameters used for the ANN and GA, determined through grid search, are presented in Table 3 and Table 4.

The operators used by the applied genetic algorithm are the following:

- **Selection**: n three random individuals tournaments are run, where n is the size of the population;
- **Crossover**: Classic three-point crossover is used for selected individuals pairs with probability p_{co};
- **Mutation**: One-bit flip mutation is done on the selected individuals with probability p_m.

n_g represents the number of generations for which the genetic algorithm runs.

Table 3. ANN parameters

Parameter	PT	NYW	RIO
Hidden layer neurons		20	
Output layer neurons		24	
Training MSE goal	0.0035	0.006	0.00375
Relaxed FS training MSE goal	0.0045	0.007	0.00475

Table 4. GA parameters

n_{pop}	p_{co}	p_m	n_g
200	0.9	0.05	80

4.3   Results

Table 5 and

Table 6 present the final results for 24-hour ahead load forecasting using the feature subsets selected by the genetic algorithm. 35 models were trained using the chosen feature subset for each one of the datasets. The minimum, mean and standard deviation of the MAPE is presented and compared with the performance of models trained

without feature selection. The features used for the training of the models without feature selection are the ones used for the ANN parameter tuning with the economic variables removed. This was done in order to have a fairer comparison. For all the datasets the use of feature selection resulted in significantly better results, an improvement of, respectively, 10%, 3.6% and 11% was obtained using the evolutionary approach for best performing models.

Regarding the features chosen by the selection algorithm, Table 7 presents the best feature subset selected by the algorithm for each one of the datasets. For the Portuguese dataset, 16 features are selected in total, consisting on 12 load values of the day before the one predicted, the maximum temperature 6 days before and the indicators for the season, holiday and weekday.

For the New York West dataset the number of loads used is equal to 16, the minimum temperature for the predicted day and 3 historic temperatures, prediction day's precipitation and weekday indicator are used. According to this, a stronger relationship between the load exists between the temperature and load for this dataset, in comparison to the Portuguese data. It is also interesting to note the lack of use of the season indicator. Maybe in this case the load follows closely the temperature in comparison to the Portuguese dataset, where the season is used for correction related to weather changes.

Regarding the Rio de Janeiro dataset, 13 features are selected. 10 historic loads, the prediction days and days before mean temperature and the weekday indicator are used.

It can be noted that the stock variables are not chosen for either one of the datasets. This way it is believed that these variables are not significant for the load forecasting in the Portuguese and New York West cases.

Table 5. Performance evaluation results – MAPE%

Dataset	Feature Selection	Nr. of features	Min. MAPE%	μ(MAPE%)	σ(MAPE%)
PT	GA	16	**2.053**	2.148	0.059
	None	32	2.284	2.386	0.052
NW	GA	22	**2.549**	2.610	0.049
	None	33	2.645	2.752	0.052
RIO	GA	13	**1.935**	2.066	0.071
	None	30	2.178	2.353	0.076

Table 6. Performance evaluation results – NRMSE%

Dataset	Feature Selection	Nr. of features	Min. NRMSE%	μ(NRMSE%)	σ(NRMSE%)
PT	GA	16	**3.698**	3.783	0.051
	None	32	3.795	3.932	0.064
NW	GA	22	**4.595**	4.666	0.033
	None	33	4.655	4.820	0.074
RIO	GA	13	**4.051**	4.252	0.104
	None	30	4.396	4.656	0.122

5 Conclusions

According to the results presented in this report it is concluded that GA feature selection can provide a performance increase in load modeling, more specifically, short-term load forecasting using ANN. Three distinct datasets with very different load patterns were used in order to study the viability of the applied methods, the presented methodology achieved good performance for all of them.

The use of stock variables was found to not be suitable for this modeling approach. The genetic algorithm never selected these features and no improvements were found when forcing their use.

Table 7. Feature selection results

Dataset	Nr. of features	Features selected by feature selection
PT	16	$h-i$ hour's historic loads ($i = 1,3,4,5,6,7,10,15,17,18,21,23$); $d-6$ day's maximum temperature; Season; Holiday; Weekday
NW	22	$h-i$ hour's historic loads ($i = 1,4,5,6,7,8,9,10,13,16,17,18,20,22,23,24$); $d-i$ day's minimum temperature ($i = 4,3,2,0$); Weekday, Prediction day's precipitation
RIO	13	$h-i$ hour's historic loads ($i = 2,3,4,6,7,9,10,12,17,22$); $d-i$ day's mean temperature ($i = 1,0$); Weekday Indicator

Acknowledgements

References

1. CSIRO and The Natural Edge Project, "Energy transformed: sustainable energy solutions for climate change mitigation." (2007)
2. L. Suganthi and A. A. Samuel, "Energy models for demand forecasting: A review," *Renewable and Sustainable Energy Reviews*, 16:2, 1223–1240 (2012)
3. D. Niu, Y. Wang, and D. D. Wu, "Power load forecasting using support vector machine and ant colony optimization," *Expert Systems with Applications*, 37:3, 2531–2539 (2010)
4. E. Kyriakides and M. Polycarpou, "Short term electric load forecasting: A tutorial," in *Trends in Neural Computation*, Springer, 391–418 (2007)
5. H. Hahn, S. Meyer-Nieberg, and S. Pickl, "Electric load forecasting methods: Tools for decision making," *European Journal of Operational Research*, 199:3, 902–907 (2009)
6. J. P. Ignizio, *Introduction to Expert Systems*. McGraw-Hill (1991)
7. K. Metaxiotis, A. Kagiannas, D. Askounis, and J. Psarras, "Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher," *Energy conversion and Management*, 44:9, 1525–1534 (2003)
8. C.-T. Lin and L.-D. Chou, "A novel economy reflecting short-term load forecasting approach," *Energy Conversion and Management*, 65, 331–342 (2013)
9. C.-T. Lin, L.-D. Chou, Y.-M. Chen, and L.-M. Tseng, "A hybrid economic indices based short-term load forecasting system," *International Journal of Electrical Power & Energy Systems*, 54, 293–305 (2014)
10. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research*, 3, 1157–1182 (2003)
11. A. Engelbrecth, "Computational Intelligence: An Introduction." (2007)
12. Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, 23:19, 2507–2517 (2007)
13. M. Welsch, M. Howells, M. Bazilian, J. F. DeCarolis, S. Hermann, and H. H. Rogner, "Modelling elements of smart grids: enhancing the OSeMOSYS (open source energy modelling system) code," *Energy*, 46:1, 337–350 (2012)
14. International Energy Agency, "Technology Roadmap: smart grids," (2011)
15. Energetics Incorporated, "Foundations for Innovation in Cyber-Physical Systems," in *Workshop Report. National Institute of Standards and Technology* (2013)
16. H. S. Hippert and J. W. Taylor, "An evaluation of Bayesian techniques for controlling model complexity and selecting inputs in a neural network for short-term load forecasting," *Neural Networks*, 23:3, 386–395 (2010)
17. M. Sheikhan and N. Mohammadi, "Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection," *Neural Computing and Applications*, 21:8, 1961–1970 (2011)

18.   S. M. Vieira, J. M. C. Sousa, and T. a. Runkler, "Two cooperative ant colonies for feature selection using fuzzy models," *Expert Systems with Applications*, 37:4, 2714–2723 (2010)
19.   C.-L. Huang and C.-J. Wang, "A GA-based feature selection and parameters optimizationfor support vector machines," *Expert Systems with Applications*, 31:2, 231–240 (2006)
20.   https://www.entsoe.eu/, "European network of transmission system operators for electricity," (2015)
21.   http://www.google.com/finance, "Google Finance." (2015)
22.   http://www.nyiso.com/, "The New York Independent System Operator." (2015)