European Association for the Development of Renewable Energies, Environment and Power Quality (EA4EPQ)

Short-term hourly load forecasting of a hospital using an artificial neural network

D. Morinigo-Sotelo¹, O. Duque-Perez¹, L.A. Garcia-Escudero², M. Fernandez-Temprano², P. Fraile-Llorente¹, M.V. Riesco-Sanz¹ and A.L. Zorita-Lamadrid¹

¹ Department of Electrical Engineering E.I.I. (Sede Paseo del Cauce), University of Valladolid c/ Paseo del Cauce 59, 47011 Valladolid (Spain) Phone number:+0034 983 184407, e-mail: morinigo@eis.uva.es

 ² Department of Statistics and Operational Research, E.I.I. (Sede Paseo del Cauce), University of Valladolid c/ Paseo del Cauce 59, 47011 Valladolid (Spain)
 Phone number:+0034 983 184397, e-mail: lagarcia@eio.uva.es

Abstract. Electricity demand forecasting is important for utilities and for some costumers. It allows to balance energy production and consumption. From the costumer point of view, it is essential for efficient operation, sizing of installation, maintenance scheduling, to name just a few. Load forecasting is a difficult task and there are many tools available to perform it. Among them, Artificial Neural Networks are receiving a lot of attention because it is not needed to know any relationship between the involved variables. But, they are constructed as black boxes, what is one their drawbacks. In this paper, some results of the load demand forecasting of a hospital are shown. It is important the previous statistical analysis of the load curves and how the results are improved adding new information to the training data set, as maximum and minimum daily temperatures.

Key words

Load forecasting, Artificial Neural Network, Control Center of Health Care Facilities, Clustering, Robust Techniques.

1. Introduction

When the jurisdiction of the Health System was transferred to the Regional Government of Castilla y León, the Regional Health Management established a collaboration agreement with the Department of Electrical Engineering and the Department of Energy Engineering and Fluid Mechanics of the University of Valladolid, which is currently serving his ninth stage.

This agreement was created with the aim of establishing a formal channel of cooperation between the Regional Health Management and the University of Valladolid, which permitted the consolidation of a working group in electrical engineering, air conditioning and ventilation guidelines in hospitals and other health care facilities, and the development of an energy management policy for health care facilities. This working group includes staff from both institutions.

This current project involves the following lines: energy tariff optimization, guides and standards for health-care facilities, energy and economic savings, use of renewable energies and energy performance models.

The work presented in this paper is inserted in the last research line previously mentioned, which has the important goal of achieving the best energy efficiency in health care facilities within the Regional Health Management and has been carried out with the participation of the Department of Statistics and Operational Research of the University of Valladolid.

Electricity is one of the main forms of energy that modern life is built upon, used in residential, commercial and industrial sectors. It has two main characteristics that determine the way the electric system is operated, namely [1],[2]:

- 1) Electrical energy cannot be stored in large quantities.
- 2) In most uses, electrical energy cannot be replaced by any other source of energy in a short term.

Due to these characteristics, the system has to maintain a permanent balance between generation and consumption of electricity. Any deviation between supply and demand will generate economic losses to the utility:

- 1) Excess demand will mean compensation to customers, since it has not been possible to supply the required energy.
- 2) Excess supply will imply to operate the electrical system out of the optimum point with energy (and therefore economic) losses (which mean an inefficient operation of the network).

The knowledge of loads at system buses is one of the most important requirements for efficient operation of power distribution systems [1],[3]. Estimation of loads is the basis for the system state estimation and for technical and economic calculations. This makes possible improvement in operation and maintenance of electrical equipment and in planning of network operating configurations [1],[3],[4].

Therefore, it is clear the importance of knowing in advance the demand curve so utility companies are able to plan the most effective way of producing electrical energy.

On the other hand, customers can also be interested in predicting their demand curve since they can obtain some advantages, namely:

- 1) Sizing of the installation taking as reference the forecasted load at each period, for new constructions or whenever a device reaches its life cycle and it is to be replaced.
- 2) Determination of future operational costs.
- 3) Planning of the facility management in order to smooth the load curve and limit peak loads, whenever possible.
- 4) Elaboration of energy efficiency strategies and the evaluation of the introduction of Renewable Energy Sources.

The estimation of the load curves is an active field of research, since slight improvements in the forecasting of the load may result in significant savings. Nevertheless, when considering health care facilities, it must be taken into account that they have certain characteristics that distinguish them from other consumers, since they show a great daily variability and are affected by variables of calendar and also by the seasonality, but not by economic factors [1],[3],[4].

In a hospital it is important to forecast the load curve for the following reasons:

- 1) It makes maintenance planning simpler and easier.
- 2) It allows evaluating the introduction of Renewable Energy Sources to supply the facility.
- 3) It is a key part in integral energy management: economic, energetic and environmental optimization.
- 4) It allows smoothing the load curve by the use of generator sets to supply part of the energy requirements.

With the purpose of encouraging energy efficiency culture to provide an optimal energy management (from an economic and environmental point of view) a pilot project was launched consisting in the installation in some hospitals of a supervisory, control and data acquisition system of the energy supply. Energy saving is the main objective of the plan along with the supervision of the quality of the energy supplied to the hospitals by the utility.

From hospitals point of view, the system should, at least, provide with the following reports:

- 1) Interactive reports in real time of the load curve.
- 2) Warning reports.
- 3) Registered electricity consumption versus electricity supplier's invoice.
- 4) Historical reports.

To achieve objectives 1) and 2) it is necessary a tool that provides a forecast of the standard load curve on a daily basis. This forecasted load could be used to generate warnings about:

- 1) Excess of demanded active power versus contracted power.
- 2) Excess of reactive energy and the necessary corrective measures.

The objective of this paper is to present the results in the development of a short term forecasting tool for the hourly load curve of a hospital in the Castilla y Léon region of Spain.

2. Forecasting Model Construction

Many authors have worked in this field and published results are very different and even sometimes contradictory, revealing how complex this matter is.

The construction of a forecasting model can be summarized in five stages [1]:

- Model formulation. The first assumptions must be specified, as the type of forecast (hourly forecast, peaks forecast [1],[5]), the forecast leadtime (short-, mid- and long-term forecast [1],[4]) and the range or life span of the model (the performance of the model will deteriorate over time and some models will need to be updated regularly [6]). Load forecasting is very influenced by seasonality, day of week and other exogenous variables. Some authors construct a different model for each season or day of the week, establishing a hierarchy of predicting models [7]-[10].
- 2) Model identification. The construction of the empirical model is based on historical data, which can be grouped by their final use in two different sets: identification and validation set. The period covered by the data is also specified, along with the identification of the independent variables to be used in the model, which must be



Fig. 1. Hospital load patterns as a result of the clustering analysis of historical data of several years.

previously arranged and transformed. Authors use normally hourly data in a period covering from several weeks [6],[7] to one or five years [7],[9]-[13].

- 3) Model estimation or construction. In this stage, first the form and type of the model is defined (linear, etc ...) (time series models [11],[13]-[16], artificial intelligent based systems as Artificial Neural Networks (ANN) [1]-[4],[6]-[10],[12] etc). Secondly, the number of free parameters are defined, whose value will be finally estimated by using the identification data set. The choice and identification of the free parameters is very difficult and will have a important influence in the final result [17].
- 4) Model validation. Once the model has been constructed, it must be checked whether it is satisfactory using the validation data set.
- 5) Model use. When the model has passed the validation test, it can be used to make predictions. During the use of the model, new data will be available that can be used to calculate the prediction errors, so it can be decided if the model must be updated

3. Statistical Analysis of Load Curves of a Hospital

The construction of a load-forecasting tool of a hospital is a task not exempt of difficulties since load curves have seasonal and daily variations (especially between winter and summer demands) although curves show similar tendencies in similar days.

The fact that load curves depend strongly on external variables is broadly acknowledged [1],[18]. These variables usually are climatic (temperature, humidity, wind

speed) and economic [1]. On the other hand, the presence of anomalous days must be taken into account, when the load curve behaves in a completely different manner to other more typical days. Examples of anomalous days are bank holidays, special weather conditions and important social events. Therefore, load curves are influenced by many variables and factors that make curve forecasting a very complex task [1],[4].

One important stage in the model construction is the selection and identification of the exogenous variables. In order to identify these variables, a Cluster Analysis, as an unsupervised learning tool, has been applied to the data set (load curves of the hospital of several years). This technique allows us to join together similar load curves without taking into account other sources of information (month, season, week-day...). Moreover, a robust clustering procedure has been applied in order to minimize the harmful effect of "atypical" load curves that unfortunately appear. In fact, following [19], we propose applying the so-called trimmed k-means procedure to the fitted basis coefficients obtained after projecting each load curve onto the space spanned by a B-spline functional base.

This analysis allowed us to classify all electricity demand curves of the hospital in four patterns, as it is shown in Fig. 1. This figure shows the point-wise means of the load curves assigned to each cluster. It was clear the different behaviour between winter and summer seasons and between business days and holidays (including weekends).

The exogenous variables can be of three types: meteorological conditions, economical and day-type [1]. As a result of the clustering analysis, it was determined that weather and day-type variables have a clear influence in the load curves shape and, the economic variable influence can be neglected. However, in a first attempt of the model construction, only day-type (business day or holiday) was considered as an explicit exogenous variable. However, season information was considered indirectly in the training data set considering a long data set covering up to three months. In order to improve the forecast results provided by the ANN constructed, temperature information, as a weather exogenous variable, was included in the training data set in a second attempt, whose results will be explained in the next section. The ANN structure was unchanged but the training data set size was reduced from three months to two months.

4. Artificial Neural Network Construction as a Forecasting Tool

An ANN was chosen as a forecasting tool. This ANN belongs to the Multilayer Perceptron family (MLP) and uses a sigmoid activation function. Many authors are using this type of neural networks as an alternative to statistic models because these require higher computational resources and a previous expertise about the relationship between the variables involved in the model construction [1]. ANN are defined by the number of input neurons, the number of hidden layers, and the number of neurons in the hidden layer [17], which is very difficult to choose. [14] recommends using a short number of hidden layers. Usually, it is enough a unique hidden layer to represent a continuous function. As there is no rule to choose the number of hidden layers and the number of neurons in them, we decided to use only one hidden layer and try with a different number of neurons in it. Some ANNs with different structures were tested, having 5, 10, 15, 20 and 25 neurons in the unique hidden layer. The network with 15 neurons had the lowest average error and less variability in the predictions. One important stage in the model construction is the selection and identification of the exogenous variables. In order to identify these variables, a statistical analysis of the hospital load curves was applied and it is explained in section 3.

As a result of the clustering analysis, economical variables were not included in the first attempt of model construction and only day-type was used as an exogenous variable. However, season information was considered indirectly in the training data set. It is used a large data set covering three previous months to the hour to estimate. So, as a long range of days is used, it can be considered that season data is included indirectly.

Day-type variable was coded as a binary variable [20]. In addition to this exogenous variable, inherent system variables were also included, as the energy demand in previous moments [10],[12]. The following past values of the electricity demand were used as inputs to the neural network:

- 1) Load of the three hours prior to the time to estimate.
- 2) Load at the same time to estimate of the previous day.

3) Load at the same time to estimate of the last week.

When training and using ANNs, input variables should be in the 0-1 ranges, so we have used the maximum and minimum year load demands to normalise input training data.

As our goal is an hourly load demand forecast, along with actual load data, time must be included in the neural network data training set. Regarding time data, it is still important to fit it in the 0-1 range. [18] and [20] proposed to represent time as a 5 digit binary variable, so 5 additional inputs are necessary to inform the ANN about time. Summarising, the first developed ANN had eleven inputs and one output, the estimation of the hospital electricity demand in a specific hour.

Historical data of the three previous months were used to train the ANN, so seasonal influence is considered indirectly.

The used learning algorithm for the MLP-network was the backpropagation algorithm. Error minimization during the training was controlled by the Levenberg-Marquardt algorithm, already implemented in Matlab,



Fig. 2. ANN forecast results for a business day (12th of April, 2010).



Fig. 3. ANN forecast results for a business day (12th of April, 2010).



Fig. 4. ANN forecast results for a business day (4th of August, 2008) without using temperature information.



Fig. 5. ANN forecast results for a business day (4th of August, 2008) considering temperature as additional information for the ANN training.

which was the numerical tool used. The ANN based model only forecasts the electricity demand of hospital in a certain hour.

To forecast load demand in a whole day, previous load estimations are used as inputs to the ANN model to forecast the load demand for the rest of the day.

The load curve provided by the ANN model for a business day (12th of April, 2010) is shown in Fig. 2, where it is also represented the actual load curve for that day. The mean error of the obtained results was inferior to the 3%.

Results for a holiday (6th of December, 2008) are shown in Fig. 3. In this case, the ANN mean error was higher than for business days. The quality of the results was inferior for holydays.

In order to improve these forecasting results, it was decided to train again the constructed ANN adding new information to the training data set regarding climate or weather conditions. The basic structure of the ANN remained unchanged but two additional inputs were added that included information about the maximum and minimum daily temperatures. The training data set was also reduced from three months to two months.



Fig. 6. ANN forecast results for a holiday (2nd of August, 2009) without using temperature information.



Fig. 7. ANN forecast results for a holiday (2nd of August, 2009) considering temperature as additional information for the ANN training.

Some results, including temperature information in the training data set, are shown in Figures 4 to 7. Figures 4 and 6 correspond to two days where the ANN without weather information gave very bad results. The estimated forecasting error was higher than 5% in both cases. These results were highly improved when the ANN was trained using also daily temperature information, as it can be seen in Figures 5 and 7. The estimated error was around 1% in both cases. So, it is clear that weather information is important in order to get good forecasting results.

5. Conclusions

In this conference paper, we present some results of a work that is being carried out and whose goal is to develop a tool to forecast the hourly power demand of a hospital. An ANN was chosen as a forecasting tool among all available tools. All the steps necessary to develop a forecasting tool have been explained. Previous to the tool development, a statistical study has been carried out to identify the variables that have some influence in the model, electricity demand of a hospital.

A Cluster Analysis, as an unsupervised learning tool, has been applied to the data set (load curves of the hospital of several years). A robust clustering procedure has been applied in order to minimize the harmful effect of "atypical" load curves that unfortunately appear. It was observed that season (or weather information) and day-type have an important influence in the respond of the model.

Next, an ANN was constructed to forecast a whole day hospital load curve with a small mean error. Two ANN were constructed. The first ANN considered historical data of the three previous months to the time to forecast and distinguished between business days and holidays, including weekends in this type of days. But, it did not use explicitly weather information as an exogenous variable. However, season information was considered indirectly in the training data set considering a long data set covering up to three months. Although, in some cases, this ANN produced good forecasting results, a second ANN was constructed, using temperature information, as a weather exogenous variable, in the training data set. The structure of the first ANN was unchanged but the training data set size was reduced from three months to two months. In this second attempt, it was observed a significant improvement in the forecasting results.

Acknowledgement

This work has been financially supported by a grant from SACYL (Health System of Castilla y León Region, Spain).

References

[1] K. Warwick, A. Ekwue and R. Aggarwal, Artificial Intelligence Techniques in Power Systems, The Institution of Electrical Engineers, London, (1997).

[2] Y. Al-Rashid and L.D. Paarmann, "Short-term electric load forecasting using neural network models", in IEEE 39th Midwest Symposium on Circuits and Systems, 1996, Vol.3, pp.1436-1439.
[3] T. Senjyu, H. Takara, K. Uezato and T. Funabashi, "One-

hour-ahead load forecasting using neural network", IEEE Transactions on Power Systems, Vol.17, no.1, pp.113-118, Feb 2002.

[4] R. Weron, Modeling and forecasting electricity loads and prices, Wiley, City (2006).

[5] J.-G. Choi, J.-K. Park, K.-H. Kim and J.-C. Kim, "A daily peak load forecasting system using a chaotic time series", in Proceedings of International Conference on Intelligent Systems Applications to Power Systems, ISAP '96, pp.283-287.

[6] I. Drezga and S. Rahman, "Short-term load forecasting with local ANN predictors", IEEE Transactions on Power Systems, Vol.14, no.3, pp.844-850, Aug 1999. [7] K.Y. Lee, Y.T. Cha and J.H. Park, "Short-term load

forecasting using an artificial neural network", IEEE Transactions on Power Systems, Vol.7, no.1, pp.124-132, Feb 1992.

[8] D.C. Park, O. Mohammed, M.A. El-Sharkawi, and R.J. Marks, "An adaptively trainable neural network algorithm and its application to electric load forecasting", in Proceedings of the First International Forum on Applications of Neural Networks to Power Systems, pp.7-11, 23-26 Jul 1991.

[9] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem, J. Farah, and C. Drake, "Practical experiences with an adaptive neural network short-term load forecasting system," IEEE Transactions on Power Systems, Vol.10, no.1, pp.254-265, Feb 1995.

[10] H.O. Sarmiento Maldonado and W.M. Villa Acevedo, "Artificial Intelligence in Forecasting Demands for Electricity: An Application in Optimization of Energy Resources", Revista Colombiana de Tecnologías de Avanzada, 2008.

[11] S. Rahman and R. Bhatnagar, "An expert system based algorithm for short term load forecast", IEEE Transactions on Power Systems, Vol.3, no.2, pp.392-399, May 1988.

[12] C.-N. Lu, H.-T. Wu, and S. Vemuri, "Neural network based short term load forecasting", IEEE Transactions on Power Systems, Vol.8, no.1, pp.336-342, Feb 1993.

[13] A. Muñoz San Roque and T. Czernichow, "Predicción de Series Temporales usando el modelo NARMAX (Aplicación a la predicción de la demanda de energía eléctrica)", Anales de mecánica y Electricidad, Vol.75, 1998.

[14] B. Espinós Pager and C. Pérez Ruiz, Estudio SALCER,

Carbos Software Engineering, 2009.

[15] S. Rahman and I.S. Moghram , "Application of a rulebased technique to weekly load forecast," in Proceedings of IEEE Southeastcon '89, Energy and Information Technologies in the Southeast, pp.380-385.

[16] S. Rahman and O. Hazim, "A generalized knowledgebased short-term load-forecasting technique", IEEE Transactions on Power Systems, Vol.8, no.2, pp.508-514, May 1993.

[17] O. Nelles, "Nonlinear System Identification: From

Classical Approaches to Neural Network and Fuzzy Models", Springer, 2001.

[18] P. Murto, "Neural Network models for short-term load forecasting", Ph.D Thesis, Helsinki University of Technology, 1998.

[19] L.A. García-Escudero and A. Gordaliza, "A proposal for robust curve clustering", Journal of Classification, Vol. 22, 185-201, 2005.

[20] C. Mallo González, "Predicción de la demanda eléctrica horaria mediante redes neuronales artificiales", Departamento de Economía Cuantitativa, Universidad de Oviedo, (http://www.uv.es/asepuma/recta/ordinarios/5/5-1.pdf,

3/1/2011).