

Article

Experimental Analysis of the Input Variables' Relevance to Forecast Next Day's Aggregated Electric Demand Using Neural Networks

Luis Hernández ¹, Carlos Baladrón ², Javier M. Aguiar ^{2,*}, Lorena Calavia ², Belén Carro ², Antonio Sánchez-Esguevillas ², Pablo García ³ and Jaime Lloret ^{4,*}

¹ CIEMAT (Research Centre for Energy, Environment and Technology), Autovía de Navarra A15, salida 56, Lubia 42290, Soria, Spain; E-Mail: luis.hernandez@ciemat.es

² Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad de Valladolid, Campus Miguel Delibes, Paseo de Belén 15, Valladolid 47011, Spain; E-Mails: cbalzor@ribera.tel.uva.es (C.B.); lcaldom@ribera.tel.uva.es (L.C.); belcar@tel.uva.es (B.C.); antsan@tel.uva.es (A.S.-E.)

³ Faculty of Sciences, University of Oviedo, c/Calvo Sotelo s/n, Oviedo 33007, Spain; E-Mail: olbapgg@gmail.com

⁴ Department of Communications, Polytechnic University of Valencia, Camino Vera s/n 46022, Valencia, Spain

* Authors to whom correspondence should be addressed; E-Mails: javagu@tel.uva.es (J.M.A.); jlloret@dcom.upv.es (J.L.); Tel.: +34-983-423-704 (J.M.A.); +34-609-549-043 (J.L.); Fax: +34-983-423-667 (J.M.A.); +34-962-849-313 (J.L.).

Received: 29 March 2013; in revised form: 6 June 2013 / Accepted: 6 June 2013 /

Published: 17 June 2013

Abstract: Thanks to the built in intelligence (deployment of new intelligent devices and sensors in places where historically they were not present), the *Smart Grid* and *Microgrid* paradigms are able to take advantage from aggregated load forecasting, which opens the door for the implementation of new algorithms to seize this information for optimization and advanced planning. Therefore, accuracy in load forecasts will potentially have a big impact on key operation factors for the future Smart Grid/Microgrid-based energy network like user satisfaction and resource saving, and new methods to achieve an efficient prediction in future energy landscapes (very different from the centralized, big area networks studied so far). This paper proposes different improved models to forecast next day's aggregated load using artificial neural networks, taking into account the variables that are most relevant for the aggregated. In particular, seven models based on the multilayer perceptron will be proposed, progressively adding input variables after

analyzing the influence of climate factors on aggregated load. The results section presents the forecast from the proposed models, obtained from real data.

Keywords: artificial neural network; aggregated load; smart grid; microgrid; multilayer perceptron

1. Introduction

A *Smart Grid (SG)* is an “*intelligent*” electrical power distribution grid that integrates information and communication technologies in order to improve the operation of the system, and to optimize the production and distribution of electricity. Other new models based on intelligent environments have been proposed, such as *microgrids*, which can be seen as an aggregation of loads and microgenerators operating as a single system to provide both electrical and thermal power, normally to a small or highly localized geographical region. For any scenario, the deployment of new intelligent devices and sensors in places where historically they were not installed provides an opportunity to come up with new intelligent operations which take advantage of the flexibility, smartness and abundance of information in order to optimize the behavior of the entire system, for instance minimizing resource usage and monetary cost while increasing user satisfaction.

The aggregated load is one of the most interesting parameters to control and forecast in all kinds of energy distribution environments. For example, in [1], Qian Zhang *et al.* proposed an aggregated model using fuzzy theory to obtain further performance improvements to forecast wind farm power output. Therefore, a lot of effort has been put into efficiently and accurately forecasting the aggregated load.

For instance, Hsu *et al.* propose in [2] a model based on *Artificial Neural Networks (ANN)* to forecast the aggregated load in three areas of Taiwan, obtaining results that could be applicable to potential business opportunities. Pilo *et al.* [3] present a *microgrid* central controller aiming to provide business value by aggregating individual and commercial consumers, which uses *ANN* to plan power generation and introduces aggregated load as an input to them, among others. Carpaneto *et al.* [4] characterize the aggregated residential power consumption from a power substation, using the seasonal aggregated load patterns as a critical parameter. Fan *et al.* [5] research the electricity demand and weather data from an electric utility and demonstrate the existence of weather and load diversity within its control area. Then, they develop a multiregion load forecasting system using aggregated load as a factor to analyze load diversity. Pudjianto *et al.* [6] propose a *Virtual Power Plant (VPP)* as the primary vehicle for delivering cost-efficient integration of *Distributed Energy Resources (DER)* into the existing power systems, where the forecast power demand in a primary input of the *Commercial VPP (CVPP)*, so as to interact later with the *Technical VPP (TVPP)*. Ruiz *et al.* [7] provide an optimization algorithm to manage a *VPP* composed of a large number of customers with thermostatically controlled appliances, where the forecast of the aggregated load over a specified period of time is one of the dynamic parameters of the algorithm. *VPPs* pose a challenge for demand forecasting and generation, and a possible approach is to use a management model which takes into account the multiple elements that are part of it, making them cooperate to obtain a demand forecast via *ANN* in disaggregated environments, as shown by Hernández *et al.* [8]. Paoletti *et al.* [9] address the problem of electric load

forecasting for distribution networks with active demand by decomposing the load into two components, the base load (aggregated load representing different seasonal patterns) and a residual term depending on both stochastic fluctuations and active demand effects. Mousavi *et al.* [10] propose an accurate and flexible voltage-dependent load model which is the most promising solution for the aggregated load modeling, and identify aggregated load as a critical parameter in the distribution network. Ipakchi *et al.* [11] present their own view of the transition from the current electricity distribution grid to *SGs*, with a great emphasis on *Demand Response (DR)*, and support of plug-in hybrid electric vehicles as well as *Distributed Generation (DG)* and power storage capabilities. Aggregated load in the control zone is again a key element of this view. Naphade *et al.* [12] point out that the transformation to smarter cities will require innovation in planning, management, and operations, and take utilities as an example of a solution that in the future will combine sophisticated demand (hourly, aggregated, *etc.*) forecasting models based on both historical and real traffic data, which could have an impact on the future demand. Future environments, represented by the umbrella term *Smart World*, will need to be aware of the relation between the local climate variables and their aggregated load in future applications, as shows Hernández *et al.* in [13] and in [14].

The electrical production in *SG/microgrids* is also subject to smart control. Pérez *et al.* [15] present an advanced control approach (Model Predictive Control) designed to manage in real-time the power production of a grid-tied solar photovoltaic power plant with an energy storage system which participates in the market. Ogliari *et al.* [16] introduce an evolutionary algorithm into *ANNs* in order to speed up the convergence when applied to the *ANN* training phase and reduce the overall error in photovoltaic plant production forecasting applications.

In summary, the state of the art suggests that in the future *SG/microgrid* environments the prediction of aggregated load will be a key factor. Planning of generation sources and the available storage will be based on advance knowledge of aggregated load, therefore, forecasting models will be needed in these new environments, which might differ from the traditional solutions applied for national or regional forecasts.

This paper aims at providing a solution for demand forecasting in *microgrid* environments, by proposing different improved models to forecast next day's aggregated load using artificial neural networks, taking into account the variables that are most relevant, which will also be studied a priori to understand their correlation with the aggregated load function to be predicted. In particular, seven models based on the multilayer perceptron will be proposed, progressively adding input variables to analyze the influence of the different factors by comparing their performance. Some of the models will include climate variables that affect demand, and all of them are validated and studied using real data from the real world.

The paper is organized as follows: Section 2 presents the methodologies and real world dataset used in the study. Section 3 shows the correlation analysis of the candidate variables with the aggregated load to be forecasted. Section 4 presents the different forecasting models to forecast based on *ANN*. Section 5 shows the evaluation results of those models and a comparison among them. And finally, Section 6 describes the conclusions and future work.

2. Data, Tools and Methodologies to Be Used in the Study

2.1. Dataset Employed

Electric power demand data to be used along this work was provided by the Spanish electric power supply company *Iberdrola* and gathered from 1 January 2008 to 12 October 2010 from a substation located in Soria (Spain). The data provided includes: day, month, year and aggregated load electric consumptions. The range of consumption varies from 7 to 39 MW, much lower than values typically observed in a large aggregated environment, therefore presenting a *microgrid* size.

The meteorological data used in this study were collected by the Spanish Meteorological Agency *AEMET* from the meteorological station installed in Soria. These data were collected from 1 January 2008 to 12 October 2010. The climate variables considered are: precipitation (mm), air temperature (°C), average wind speed (m/seg), average wind direction (sexagesimal degrees), relative humidity (%), pressure (hPa) and global solar radiation (W/m²).

2.2. Methodology

2.2.1. Analysis of the Variables of Interest

The ultimate aim of this study is to obtain an evaluation of the influence of variables that could be fed into the ANN used by the next day's aggregated load forecasting model, focusing on:

- Past aggregated load: an autocorrelation analysis will be carried out to detect the influence of the previous days' aggregated load on the next day's demand.
- Climate variables: those variables that have a most significant influence on the aggregated load will be detected.

When designing the different ANN models that will be proposed, the strength of the relation between the aggregated load for a certain day and that of the previous days must be assessed, and for that the statistical autocorrelation, defined as the cross-correlation of a signal with itself, will be used as a useful way to find repetitive or periodic patterns in a signal. The statistical autocorrelation of a discrete time series for a given process, X_t , is simply the correlation of the said process with a time-shifted version of the same series. Let X_t represent a second order stationary process with mean μ , then:

$$R(k) = \frac{E[(X_t - \mu)(X_{t-k} - \mu)]}{\sigma^2} \quad (1)$$

where E is the expected value and k is the time-shift. This autocorrelation function varies within the range $[-1, +1]$, where $+1$ denotes total correlation, -1 indicates total anticorrelation and 0 means that there is no correlation. Based on this autocorrelation function, the autocorrelation coefficients for the values of the aggregated load will be obtained.

As far as the climate variables go, this work seeks to examine the relation that exists between them and the aggregated load, so as to determine which can be used in the forecasting models. For this purpose, a correlation analysis between the aggregated load and all the available climate variables will be carried out. In statistics and probability, a correlation measures the strength and direction of the linear relationship between two random variables. Several coefficients measure the correlation degree,

but the most widely used is Pearson's correlation coefficient, which was the one used in this study. It is obtained by dividing the covariance of two variables by the product of their standard deviations. Pearson's correlation can be defined as an index that measures to what extent two quantitative variables are related. Correlation coefficients can range from -1 to $+1$:

- $r = +1$: perfect positive linear correlation.
- $0.0 < |r| < 0.09$: no correlation.
- $0.1 < |r| < 0.25$: small linear correlation.
- $0.26 < |r| < 0.55$: medium linear correlation.
- $0.56 < |r| < 1$: strong linear correlation.
- $r = 0$: both variables are not linearly related.
- $-1 < r < 0$: negative linear correlation.
- $r = -1$: perfect negative linear correlation.

2.2.2. Forecasting Models Based on ANNs

For each of the ANN models developed, a set of indicators must be used to measure the accuracy of the forecasts once the operation stage has been completed. Assessing the electric demand forecasting results is not an easy task, as shown by Hippert *et al.* [17]. These authors also point out that utilities are reluctant to use new non-linear forecasting techniques due to the lack of an evaluation system which can be understood by industry. As a consequence, this study uses three complementary error measures that will be applied to all models:

- *Mean Absolute Percentage Error (MAPE)*: this is the most widely used measure by the industry, and thus allows comparing the results with previous studies. MAPE error is defined as:

$$MAPE = 100 \cdot \frac{\sum_{i=1}^n \left| \frac{A_L(i) - \hat{A}_L(i)}{A_L(i)} \right|}{n} \quad (2)$$

where $A_L(i)$ is the aggregated load corresponding to day i ; $\hat{A}_L(i)$ is the forecast aggregated load for day i ; and n is the size of the sample.

- *Root Mean Square Error (RMSE)*: MAPE error lacks sensitivity for errors which are more than two standard deviations away from the mean. However, this kind of error, though uncommon, is of great importance for utilities. Obviously, the weight of big deviations is greater for a squared function than for the absolute value function, and therefore the former has been selected. RMSE error is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \left(\frac{A_L(i) - \hat{A}_L(i)}{A_L(i)} \right)^2}{n}} \quad (3)$$

- *Maximum Error (ME)*: this measure complements the previous two and evaluates the maximum difference between the forecasts and the real values. A single but very big deviation could have dramatic consequences for a production system. ME error is defined as:

$$ME = \max_i |A_L(i) - \hat{A}_L(i)| \tag{4}$$

3. Analysis of the Relevant Variables to Forecast the Aggregated Load

This section studies the relationship between the aggregated load at any given moment and the aggregated load recorded during previous days, and the relationship between aggregated load and climate variables. This will provide the scientific foundation for the proposal of different forecasting models in Section 4.

3.1. Autocorrelation of Aggregated Load

It is reasonable to think that past values of the load curve are related to current and future values, but in order to measure this relation, the autocorrelation of the aggregated load curve will be studied in this subsection. A maximum of 200 days preceding the day of interest will be considered, which is enough to estimate the optimal number of inputs to be fed to the ANN and allows for enough detail.

Figure 1 shows the analysis of seven specific days chosen randomly for illustrative purposes, each corresponding to a day of the week, including a bank holiday. The legend on the upper-right corner shows the color corresponding to each day of the week. It can be observed that the highest autocorrelation coefficients correspond to the day immediately before the one analyzed, as well as to the same days of the week, up to certain point in the past.

Figure 1. Autocorrelation coefficient for each of the days of the week with respect to all past patterns, arranged by day of the week.

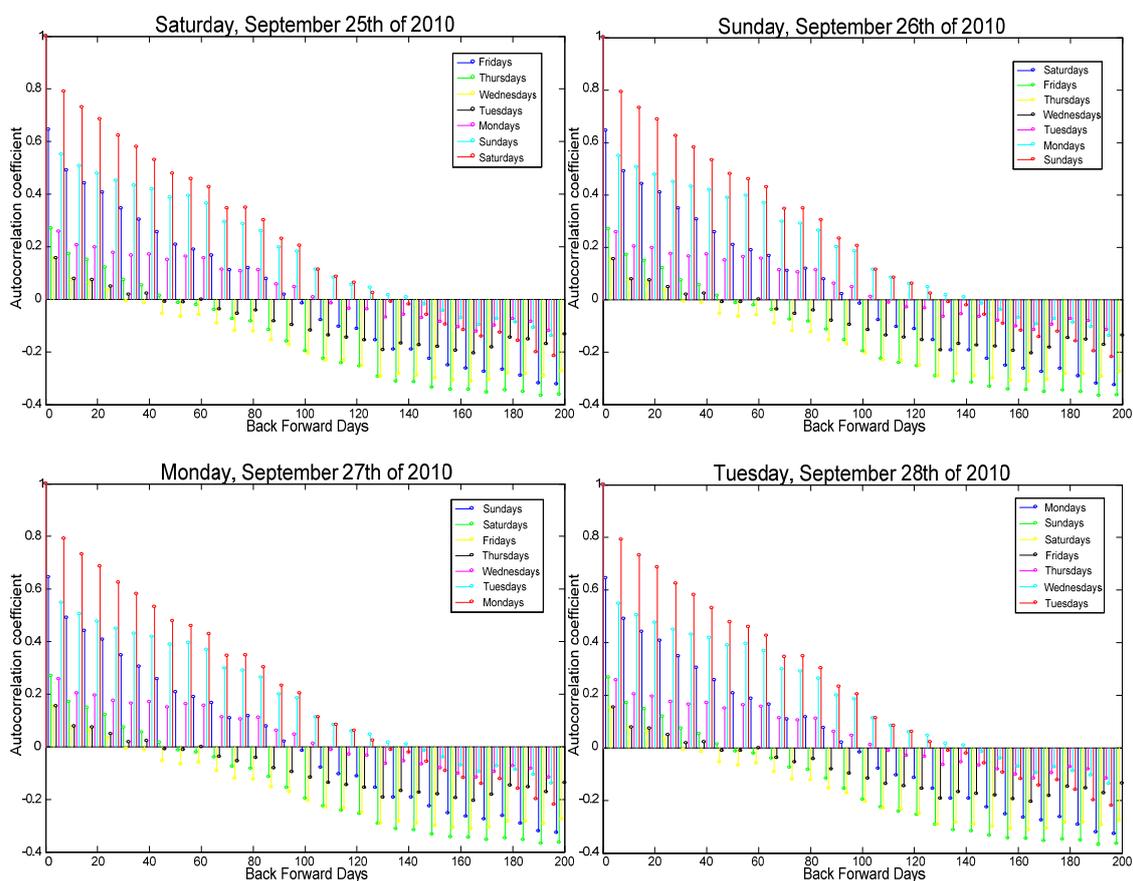
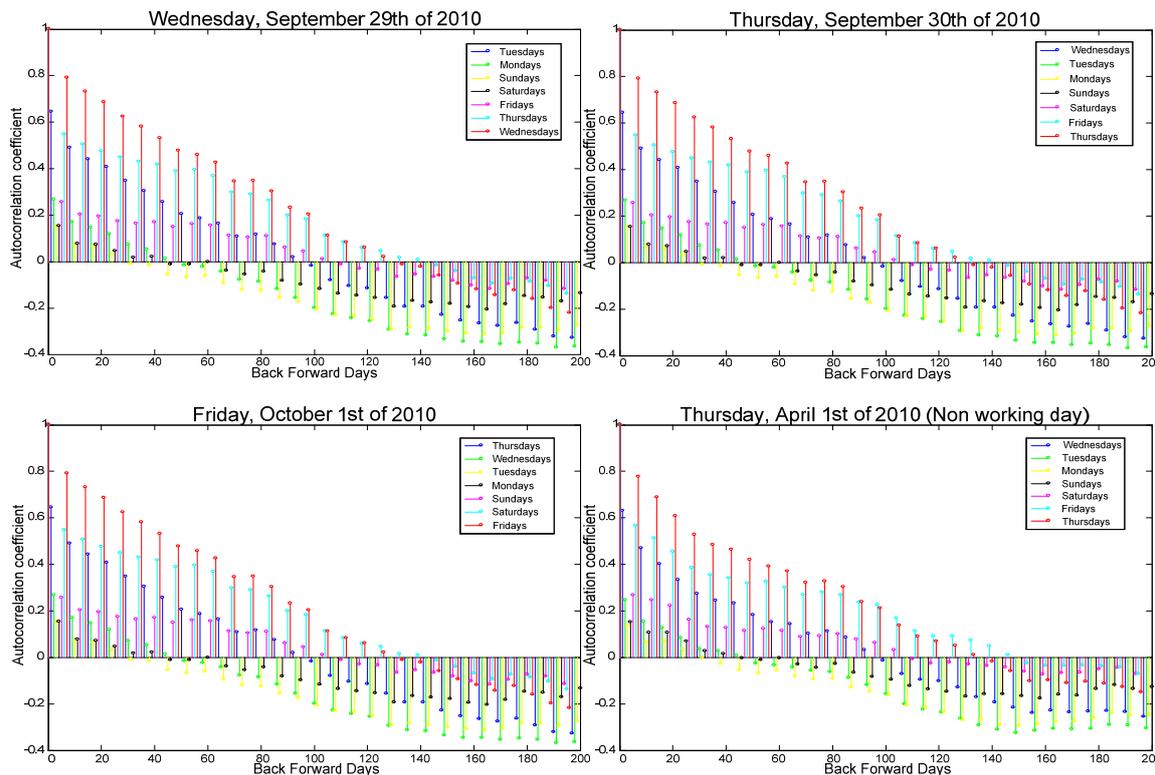


Figure 1. Cont.



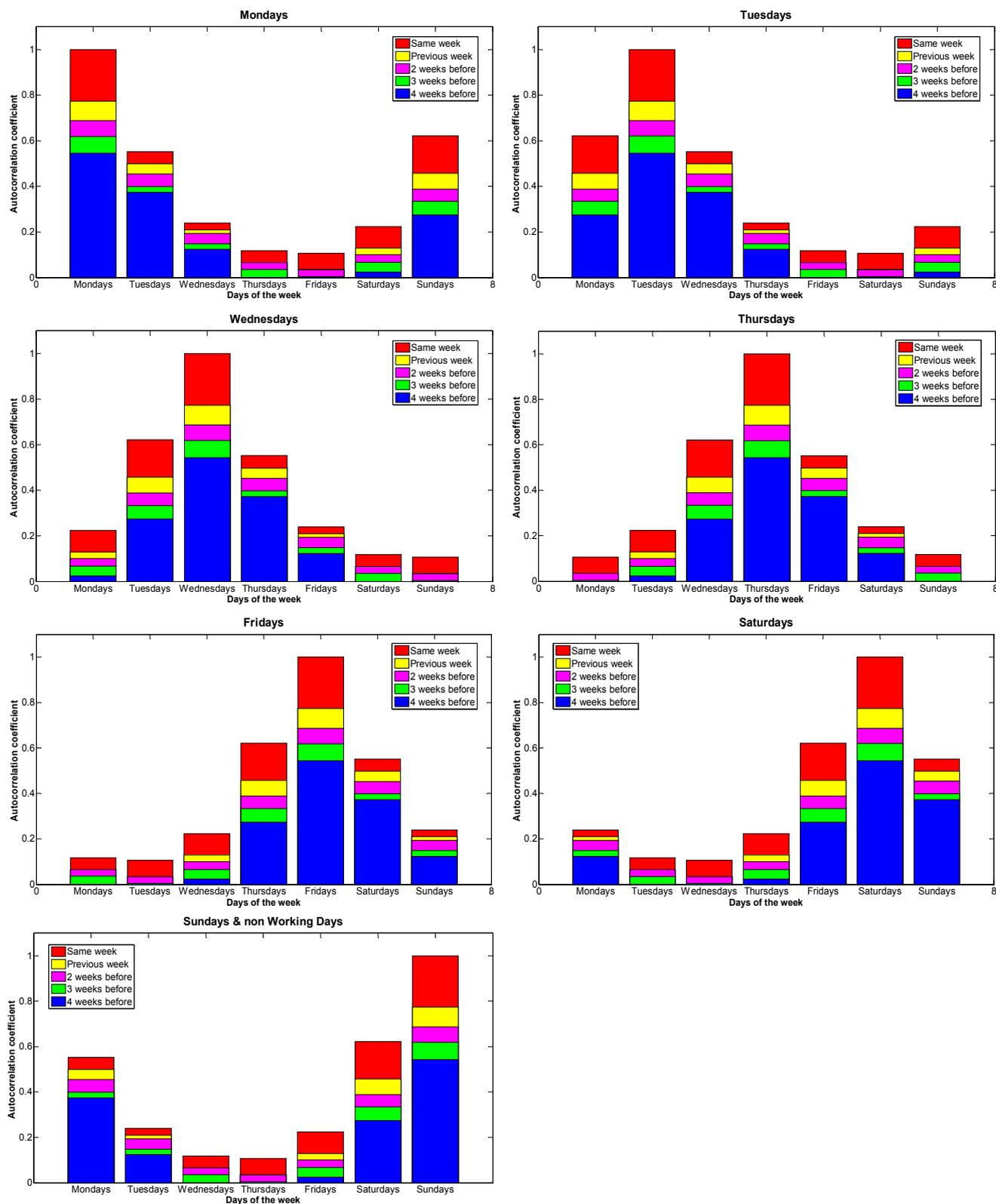
If all correlation values greater than 0.6 are taken as a reference, days can be classified in two groups:

- Working days correlate with the day coming immediately before them and with the same day of the week for the previous three weeks.
- Similarly, non-working days correlate with the day coming immediately before them and with the same day of the week for the previous two weeks.

Figure 1 presents specifically seven example days during the end of September and beginning of October, and an additional example day in April (6 months earlier). The plots show that there is no appreciable significant difference in the autocorrelation coefficients inside a single week, but it is easy to observe that these values change along the year just by looking at the coefficients for the day in April.

A global study of the mean autocorrelation values along the entire database is carried out. For each day, the autocorrelation coefficients with respect to the previous thirty days are obtained, and the average is calculated considering all the values for the same days of the week. This way, an autocorrelation value is obtained between any given day of the week and its thirty previous days, as shown in Figure 2. The obtained results are in accordance with other results reported in the literature, specifically [13]: the days showing a closer relation to the day to be forecasted are those corresponding to the same day of the week from the previous three or four weeks, together with the day immediately before. It is also interesting to see that the least correlation is given when four days have passed (e.g., for Mondays, the least correlated day is Thursday).

Figure 2. Mean autocorrelation coefficients for the aggregated load, arranged by day of the week. (Note: Each bar represents the autocorrelation coefficient of every week day, and each bar portion represents one week delay time.)



3.2. Climate Variables

The aim of this subsection is to obtain a global picture of the relation between the climate variables and the aggregated load, and for that, a correlation analysis has been carried out using all the available patterns.

Table 1 shows the correlation matrix between the aggregated load and the climate variables. It can be seen that mean temperature, mean relative humidity and aggregated solar radiation are the variables that correlate the most to the aggregated load. Mean temperature and aggregated solar radiation across the year show a negative correlation, *i.e.*, the bigger the temperature and the radiation, the smaller the energy consumption and vice versa. A seasonal study shows that mean temperature has a negative correlation with the aggregated load in cold seasons, but this is not true during summer, when the correlation between temperature and load is positive mainly due to air conditioning systems; however, the global annual mean still presents a negative correlation.

Table 1. Correlation between climate variables and aggregated load. Aggregated Load (AL); Mean Temperature (MT); Mean Relative Humidity (MRH); Aggregated Precipitation (AP); Aggregated Solar Radiation (ASR); Mean Wind Direction (MWD); Mean Wind Speed (MWS); and Mean Pressure (MP).

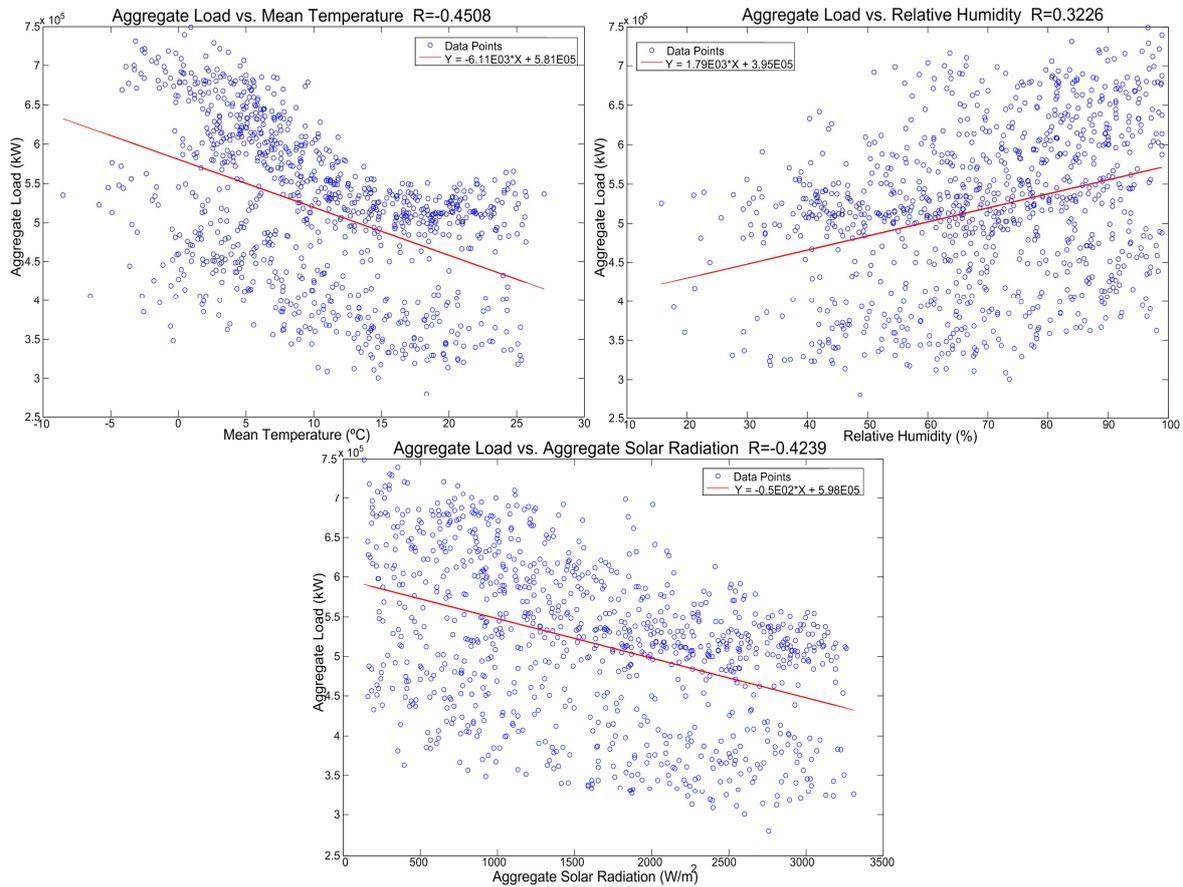
All days	AL	MT	MRH	ASR	MWD	MWS	MP
AL	1.00000	-0.4508	0.3226	-0.4239	0.25437	0.20064	0.02355
MT	-0.4508	1.00000	-0.61656	-0.09186	-0.19514	-0.23928	0.05872
MRH	0.32157	-0.61656	1.00000	0.27902	0.08449	0.24059	-0.06951
ASR	-0.4239	-0.09186	0.27902	1.00000	0.16316	0.01912	0.00309
MWD	0.25437	-0.19514	0.08449	0.16316	1.00000	0.32152	-0.07974
MWS	0.20064	-0.23928	0.24059	0.01912	0.32152	1.00000	-0.08816
MP	0.02355	0.05872	-0.06951	0.00309	-0.07974	-0.08816	1.00000

In contrast, relative humidity shows a mean positive correlation. Figure 3 shows three graphs representing linear regression of the variables that correlate the most to aggregated load. A significant grouping of the points in the cloud is observed, as well as a positive slope for relative humidity and a negative slope for temperature and solar radiation. These results are consistent with those presented in Hernández *et al.* [13].

4. Models Proposed to Forecast the Aggregated Load

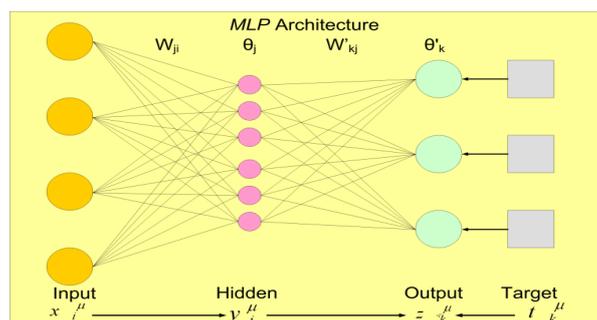
Forecasting models can be divided in time series models (univariate) and causal models, depending on the factors taken into account for the forecast. The former are based on historical values [18–22] whereas the latter focus on social variables and exogenous factors [23–28]. Models based on ANN have received the most attention. Rumelhart [29] developed the *MultiLayer Perceptron (MLP)* algorithm, setting up the basis for a number of architectures; Bromhead and Lowe [30] proposed *Radial Basis Functions* networks (RBF); Elman [31,32] proposed recurrent networks; and Kohonen [33] focused on *Self-Organized Map (SOM)* networks.

Figure 3. Linear regression of the aggregated load on mean temperature, relative humidity and solar radiation. Y-axis shows the aggregated demand (kW) and x-axis, the mean temperature (°C), relative humidity (%) and solar radiation (w/m^2). The correlation coefficient between the aggregated demand and each of the variables is indicated in each graph.



MLP is generated by adding hidden layers to a simple perceptron, which is usually trained using the *backpropagation* algorithm or one of its variants. The *MLP* architecture scheme is shown in Figure 4, where x_i are the inputs; y_i are the outputs from the hidden layer; z_k represent the outputs from the output layer and the global grid; and t_k are target outputs. Additionally, w_{ji} are the weights for the hidden layer; θ_j are their thresholds; and w'_{kj} are the weights for the output layer; while θ_k represent their thresholds.

Figure 4. *MLP* Architecture.



Along this work, all the models proposed have the same output variable, aggregated load. Several forecast models based on *MLP* will be proposed in this section using the results from Section 3 as a basis. The generic network architecture for these models is shown in Figure 4. The subscript of the input variables represents the delay (in days) with respect to the day to be forecast (d) and the subscripts *sin* and *cos* indicate that a periodic variable has not been feed directly, but under the form of its sine or cosine (it has been demonstrated [34,35] that this practice increase the accuracy of the *ANN*).

The models proposed, which differ on the input variables employed, are the following:

- *Forecast with Aggregated Load (F_{AL})*: as shown in Section 3.1, aggregated load is closely related to the previous day's aggregated load, as well as to aggregated loads corresponding to the same day of the week for the previous three weeks, regardless of whether this is a working day or not, and this holds for all days of the week. With this in mind, in the F_{AL} model the chosen inputs are the aggregated load of the previous day, and the same days of the week, of the previous three weeks. This makes 4 inputs and 1 target.
- *Forecast with Aggregated Load and Type of Day [working/non-working](F_{AL_W})*: five new variables are added to the F_{AL} , indicating whether each of the days is a working day or not. They refer to the past days and the day for which the forecast is made. This makes nine inputs and one target.
- *Forecast with Aggregated Load, Type of Day (working/non-working) and Day of the Week (F_{AL_WD})*: 10 new variables are added to the F_{AL_W} , indicating the day of the week (*Sunday* = 0; *Monday* = 1; *Friday* = 5; *Saturday* = 6) in the sine and cosine forms. These refer to the past days and the day to be forecast. This makes 19 inputs and one target. As mentioned, for circular variables (days of the week, *etc.*), the use of two inputs for each variable (in its sine and cosine forms) has been shown to improve performance, since values are uniformly distributed between 0 and 2π , allowing the network to perceive periodic behavior more efficiently and reducing training time, as shown in Ramezani *et al.* [34] and in Razavi and Tolson [35].
- *Forecast with Aggregated Load, Type of Day (working/non-working), Day of the Week and Mean Temperature ($F_{AL_W_DW_MT}$)*: In this model, five new variables are added to the $F_{AL_W_DW}$, representing the mean temperature of the past days and the day for which the forecast is made. This makes 24 inputs and one target.
- *Forecast with Aggregated Load, Type of Day (working/non-working), Day of the Week and Relative Humidity ($F_{AL_W_DW_RH}$)*: In this model, five new variables are added to the $F_{AL_W_DW}$, representing the relative humidity of the past days and the day for which the forecast is made. This makes 24 inputs and one target.
- *Forecast with Aggregated Load, Type of Day (working/non-working), Day of the Week and Solar Radiation ($F_{AL_W_DW_SR}$)*: In this model, five new variables are added to the $F_{AL_W_DW}$, representing the solar radiation of the past days and the day for which the forecast is made. This makes 24 inputs and one target.
- *Forecast with Aggregated Load, Type of Day (working/non-working), Day of the Week and all Weather ($F_{AL_W_DW_allW}$)*: In this model, 15 new variables are added to the $F_{AL_W_DW}$, representing the mean temperature, relative humidity and solar radiation of the past days and the day for which the forecast is made. This makes 34 inputs and one target.

One of the main aspects of the *ANN* is the ability to generalize based on examples. Generalization should be understood as the ability of the network to answer correctly to patterns not used in the training stage. The *ANN* should be able to generalize, as opposed to simply memorizing the answers to known patterns. Ideally, an *ANN* architecture should be trained up to an optimal point where generalization error is minimal. The procedure, called cross-validation, involves a simultaneous training and validation so as to stop in the optimal point, and consists of three stages: training, validation and testing. For all models, the patterns selected to train the *ANN* are subject to cross-validation: 70% training, 15% validation and 15% testing. Of all the data presented in Section 2.1, the training stage used patterns ranging from 1 January 2008 to 11 December 2009, while the operation stage used patterns ranging from 12 December 2009 to 12 October 2010. In the operation stage, a forecast was made for each of the days within the set, after training the models with the corresponding data. For each model and day, the *MAPE* was computed using (2); once the operation stage was over, the *RMSE* and the *ME* were evaluated, via (3) and (4), respectively, in order to compare the models.

5. Results

A script was written for the automation of the creation, training and operation stages. For each of the models, for each of the training functions, and for each hidden layer size (between one and 20 neurons), the script carried out 100 different runs in order to achieve statistically meaningful results which rule out the random factors influencing the *ANN* (such as the initial state).

The training functions considered are: *traingd* is *gradient descent backpropagation*; *traingdm* is *gradient descent with momentum backpropagation*; *traingda* is *gradient descent with adaptive learning rate backpropagation*; *traingdx* is *gradient descent with momentum and adaptive learning rate backpropagation*; *trainrp* is *resilient backpropagation*; *traincgf* is *conjugate gradient backpropagation with Fletcher-Reeves updates*; *traincgp* is *conjugate gradient backpropagation with Fletcher-Ribière updates*; *traincgb* is *conjugate gradient backpropagation with Powell-Beale restarts*; *trainscg* is *scaled conjugated gradient backpropagation*; *trainbfg* is *BFGS quasi-Newton backpropagation*; *trainoss* is *one-step secant backpropagation*; *trainlm* is *Levenberg-Marquardt backpropagation*; and *trainbr* is *Bayesian regulation backpropagation*. *Sse* (*Sum Squared Error*) was used as the network performance function (which ends the training stage).

Table 2 shows the *MAPE* for the operation stage of each of the seven models proposed in Section 4 to forecast aggregated load. For each of the models, all the training functions and the optimal number of neurons are shown, after forecasting the aggregated load for all the days in the operation stage (with 100 iterations in all cases). The optimal values, *i.e.*, the ones that minimize *MAPE* during the operation stage for each of the models, are shown in bold.

Table 2. Comparison of the minimum *MAPE* for all the forecast models proposed, after 100 iterations, and for all training functions. Where: (1) optimal number of neurons for the specific model and training functions; (2) *MAPE* (%) for the optimal number of neurons; and (3) standard deviation (%) for the optimal number of neurons.

	<i>traingd</i>			<i>traingdm</i>			<i>traingda</i>			<i>traingdx</i>			<i>trainrp</i>			<i>traincgf</i>			<i>traincgp</i>		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
<i>F_AL</i>	1	27.84	9.13	1	27.99	8.70	13	8.39	0.96	1	27.07	8.67	7	6.90	0.46	11	6.81	0.37	13	6.98	0.53
<i>F_AL_W</i>	5	40.87	18.85	5	39.26	18.59	6	7.09	1.64	5	42.20	19.12	7	5.34	0.42	15	5.26	0.42	12	5.37	0.51
<i>F_AL_W_DW</i>	1	30.92	9.04	1	28.47	8.67	8	5.31	1.09	1	29.70	10.38	7	3.90	0.44	13	3.61	0.36	13	3.67	0.51
<i>F_AL_W_DW_MT</i>	1	29.90	8.99	1	28.93	10.08	10	5.89	1.05	1	30.39	8.82	8	4.21	0.47	12	3.83	0.39	11	3.99	0.51
<i>F_AL_W_DW_RH</i>	1	27.67	8.01	1	29.52	9.64	9	5.98	1.15	1	29.35	9.19	8	4.16	0.54	15	3.68	0.42	16	3.82	0.41
<i>F_AL_W_DW_SR</i>	1	28.96	9.89	1	29.01	9.24	12	5.89	1.21	1	29.80	8.44	8	4.13	0.60	17	3.66	0.48	15	3.77	0.40
<i>F_AL_W_DW_allW</i>	1	28.14	9.43	1	29.01	9.31	5	6.54	1.68	1	28.42	9.00	7	4.44	0.56	16	3.93	0.49	18	3.99	0.46

Table 2. Cont.

	<i>traincgb</i>			<i>trainscg</i>			<i>trainbfg</i>			<i>trainoss</i>			<i>trainlm</i>			<i>trainbr</i>		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
<i>F_AL</i>	13	6.85	0.36	13	6.96	0.43	10	6.98	0.37	15	7.04	0.45	13	7.06	1.36	11	6.37	0.22
<i>F_AL_W</i>	18	5.31	0.51	15	5.43	0.52	8	5.49	0.35	15	5.54	0.51	19	5.33	0.84	7	4.44	0.15
<i>F_AL_W_DW</i>	16	3.66	0.43	13	3.72	0.57	10	4.06	0.34	12	3.79	0.46	18	3.96	0.85	4	3.13	0.17
<i>F_AL_W_DW_MT</i>	14	3.95	0.48	14	3.96	0.56	8	4.42	0.49	11	4.21	0.68	15	4.53	0.79	3	3.26	0.18
<i>F_AL_W_DW_RH</i>	15	3.85	0.41	15	3.84	0.51	8	4.35	0.45	8	4.01	0.64	14	4.52	1.45	3	3.12	0.17
<i>F_AL_W_DW_SR</i>	16	3.73	0.47	18	3.79	0.58	9	4.43	0.64	8	4.05	0.62	15	4.46	1.64	4	2.98	0.15
<i>F_AL_W_DW_allW</i>	15	4.02	0.40	14	4.05	0.48	10	4.64	0.51	9	4.29	0.64	17	4.94	1.29	3	3.18	0.19

Table 2 shows how the training functions *traingd*, *traingdm* and *traingdx* present minimal operation-stage *MAPE* values which are above 25%. This happens for a low number of neurons, therefore, the *MAPE* will continue to increase as the number of neurons grows. Hence, this section only shows *MAPE* error as a function of the number of neurons for the remaining training functions. Additionally, for all the models proposed, the training function which yields a lower *MAPE* in the operation stage is *trainbr* (which also achieves the lowest standard deviation). Next, the details of the results for each model will be explained.

The models proposed start with a low number of inputs (4), and new variables are gradually added to them. Using *trainbr*, it can be observed how, in general, the addition of new variables decreases *MAPE* error during the operation stage, while decreasing in most cases the neurons needed in the hidden layer.

Figure 5 shows the evolution of *MAPE* during the operation stage depending on the training functions and number of neurons for all the models and with 100 iterations in each case. As it can be seen, *trainbr* produces better results than the other training functions, and some behavioral patterns can be observed for certain training functions in all models. Figure 6 represents the evolution of *MAPE* for each of the days forecast during the operation stage, for all the models proposed. The forecasts with a higher *MAPE* have been labeled, according to the following notation: *Month/Day-Day_of_the_Week Workability*. In addition, *MAPE*, *RMSE* and *ME* values of the operation stage are shown for each model. The *F_AL* simply uses as an input the aggregated load of the days that related the most to the day for which the forecast is made, as explained in Section 3.1. Figure 6 shows that highest *MAPE* values occur when forecasting:

- Non-working days that fall on weekdays, as well as some days that immediately follow them. Since no additional information is available for the different patterns, the network did not take into account whether the day was a working day. The period between 23 June and 28 June, which corresponds to the local festivals. Except for Thursday 24, and Sunday 27, which are non-working days, the rest of the days within the period are working days, but in practice they are similar to non-working days. As previously, this information is not available for the network.
- A high number of Saturdays show an error between 15% and 25%. This can be explained considering that Saturdays are working days (the network forecasts a high aggregated load value), but in fact aggregated load is lower than on the Monday-Friday period. From the previous analysis, the need for the model to know whether the days are working days becomes clear. Therefore, the *F_AL_W* model uses this variable as an input, for the past days and the forecast day. As it can be observed in Figure 6, the model still fails during the local festivals' week, as well as on Saturdays. However, the error for non-working days is lower than the mean in all cases, *i.e.*, adding the new variable (working/non-working) as an input improves the forecast.

Figure 5. MAPE (%) in each of the models for the operation stage, considering all training functions and varying number of neurons. (Note: F_{AL} : Forecast with Aggregated Load; F_{AL_W} : Forecast with Aggregated Load and Workability; $F_{AL_W_DW}$: Forecast with Aggregated Load, Workability and Day of the Week; $F_{AL_W_DW_MT}$: Forecast with Aggregated Load, Workability, Day of the Week and Mean Temperature; $F_{AL_W_DW_RH}$: Forecast with Aggregated Load, Workability, Day of the Week and Relative Humidity; $F_{AL_W_DW_SR}$: Forecast with Aggregated Load, Workability, Day of the Week and Solar Radiation; $F_{AL_W_DW_allW}$: Forecast with Aggregated Load, Workability, Day of the Week and all Weather).

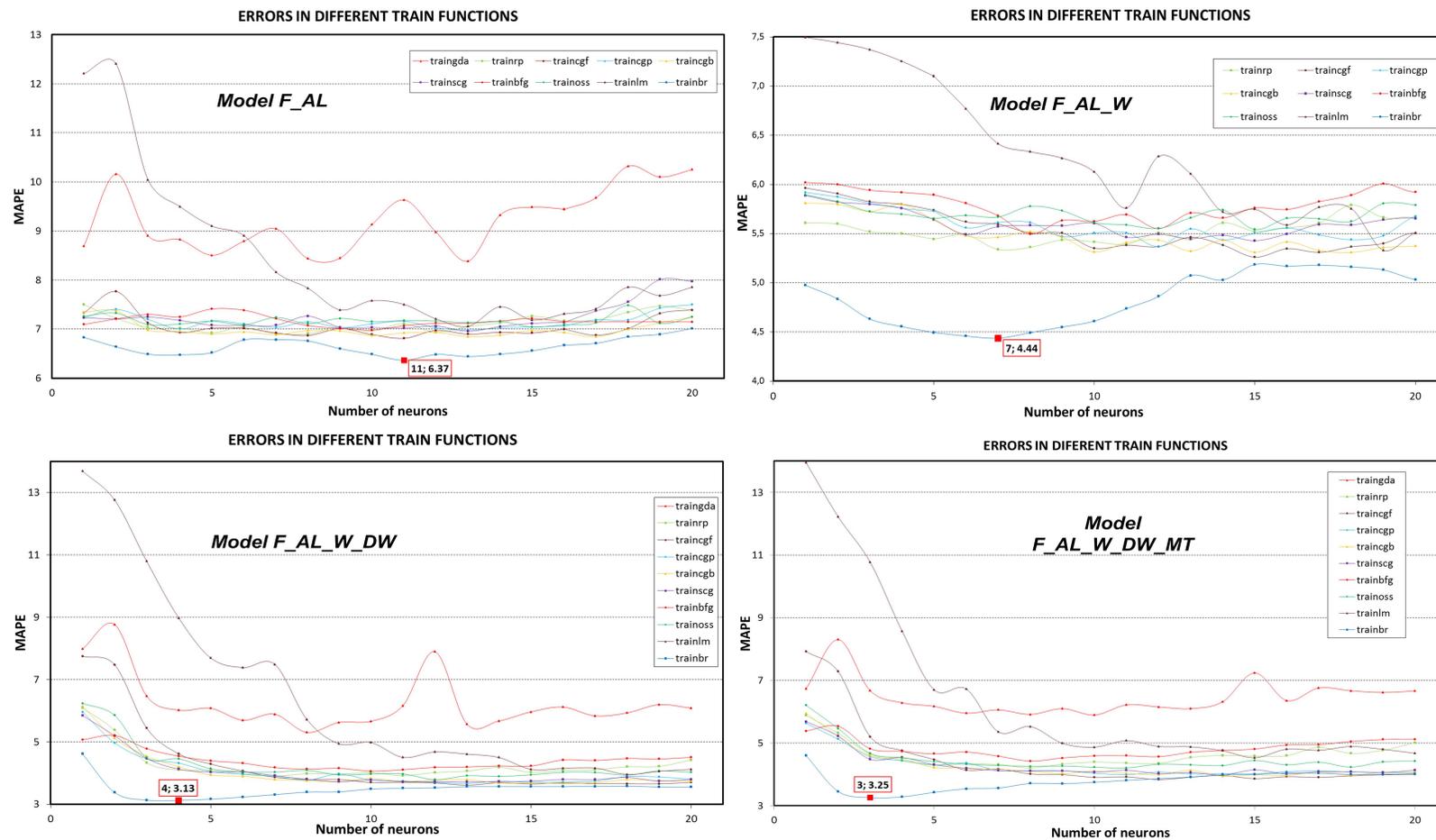


Figure 5. Cont.

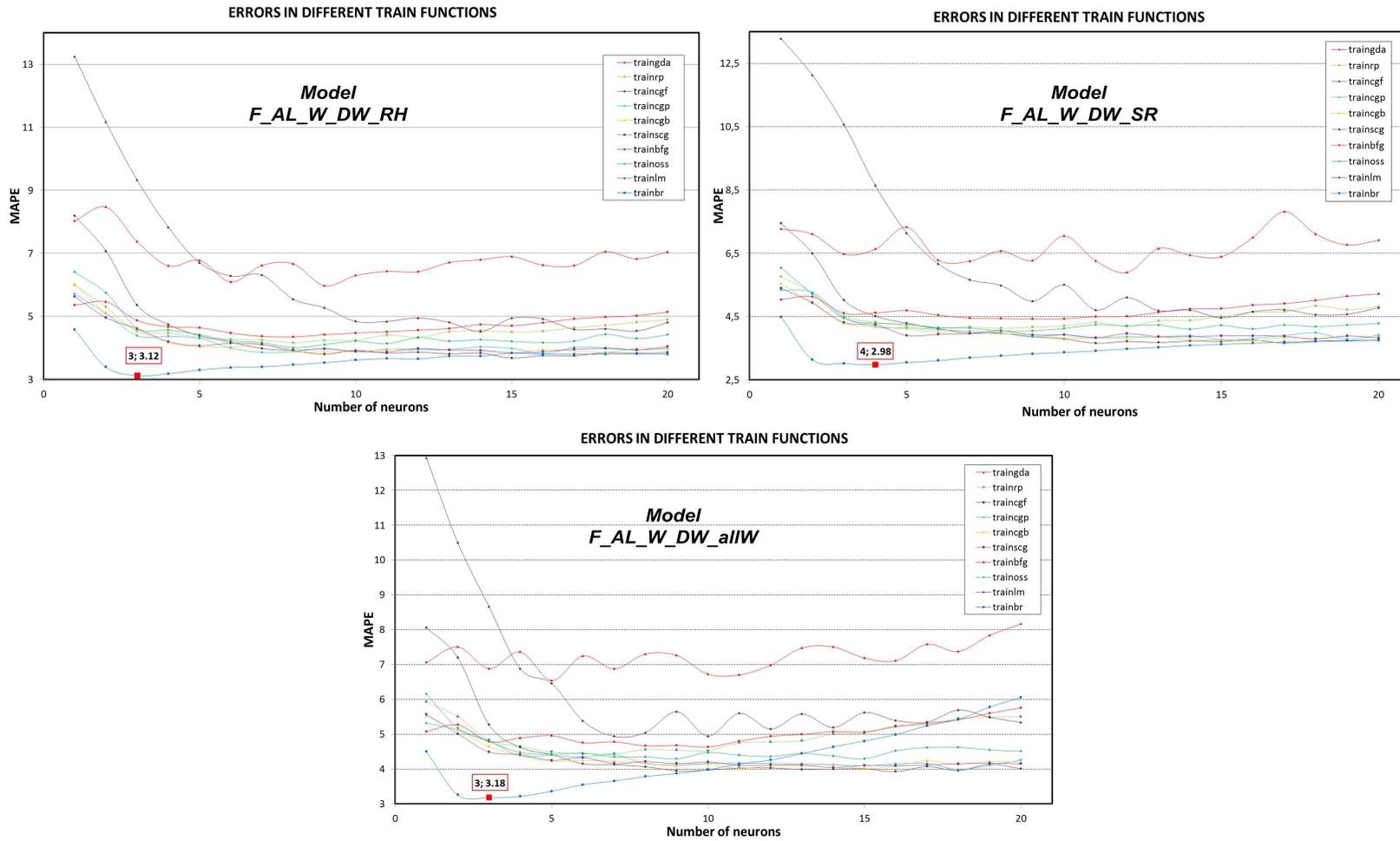
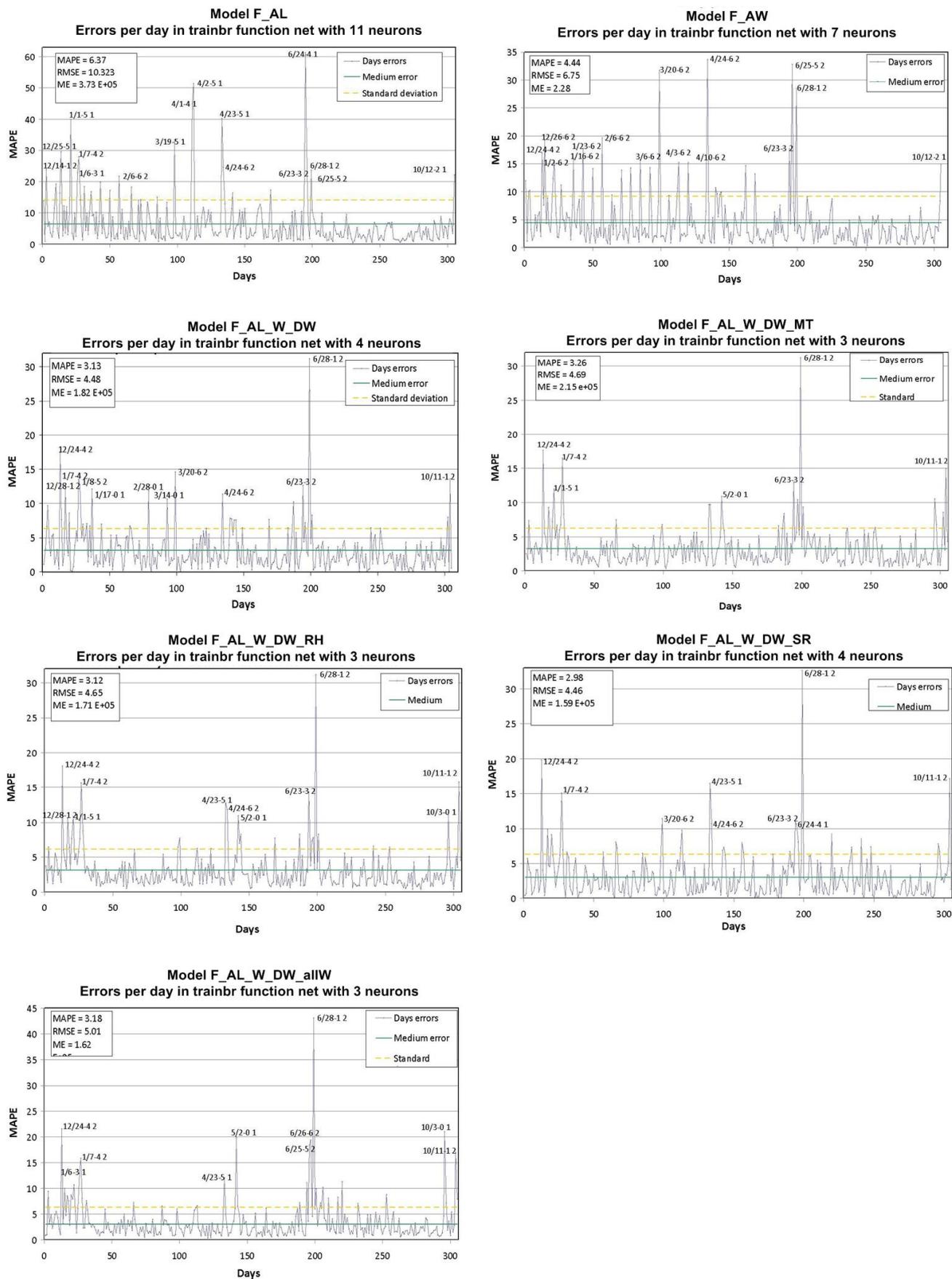


Figure 6. The *MAPE* (%), *RMSE* (dimensionless) and *ME* (W) values of the operation stage are shown for each model with *trainbr*.



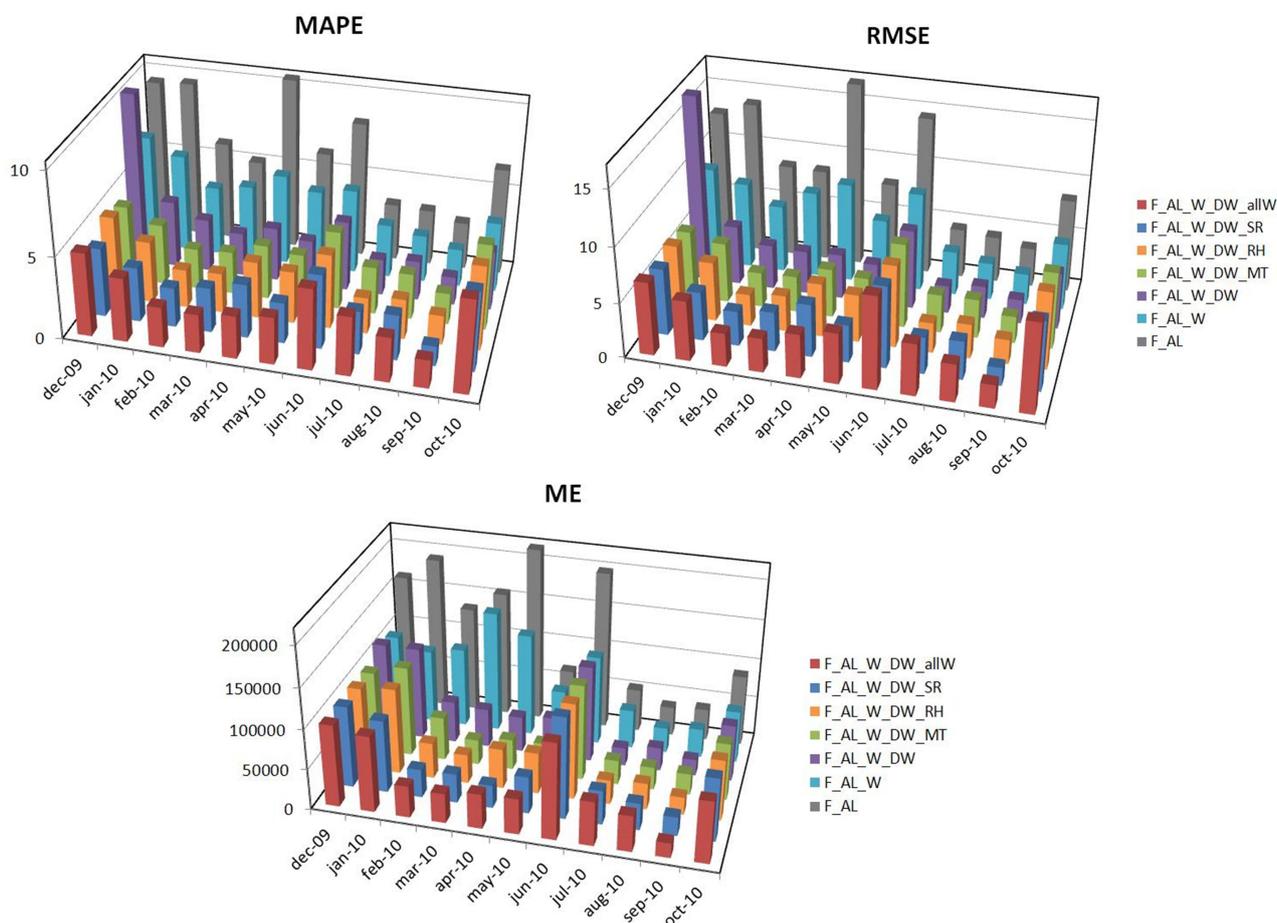
In order to improve the results of the previous model, the days of the week are added as an input, in their sine and cosine form, both for the past days feeding the model and for the day for which the forecast is made. This model is named $F_AL_W_DW$. Figure 6 shows that the model still fails when there are two non-working days in a row, because the patterns are very similar to Saturdays, but no periodicity can be detected for the model. Additionally, the model still fails during the local festivities, but with some improvement. For some Sundays, the forecast seems to go notably wrong; however, the scale of the error-axis has decreased which makes these errors seem high, when in fact they are similar to those produced in the previous models. Also, the forecast in Saturdays has improved notably.

For these three models, when introducing new variables in the ANN the frequency of the days showing a high MAPE decreases; also, Figure 6 shows that the lower bound for the error decreases, from a 20% in F_AL to a 15% in F_AL_W and a 10% in the $F_AL_W_DW$; in the last two, the errors higher than 20% decrease notably (in particular, this happens only once for $F_AL_W_DW$).

The same goes for all the models including climate variables ($F_AL_W_DW_MT$, $F_AL_W_DW_RH$, $F_AL_W_DW_SR$ and $F_AL_W_DW_allW$), whose behavior is very similar to $F_AL_W_DW$. The models present an operation-stage MAPE very similar to that of $F_AL_W_DW$, although the slight improvement of the $F_AL_W_DW_SR$ should be highlighted. However, all the models using climate variables show a slightly higher error when forecasting certain non-working days, as well as days immediately preceding/following those; therefore, it could be possible to use a model that includes climate variables for the forecast in weekdays, and to use the $F_AL_W_DW$ model for the rest. In particular, $F_AL_W_DW_SR$ presents a 2.57% operation-stage MAPE, compared with the 2.79% achieved by the $F_AL_W_DW$; in contrast, the operation-stage MAPE for non-working days and the immediately preceding/following days is 9.87% for $F_AL_W_DW$ and 13.57% for $F_AL_W_DW_SR$.

Finally, Figure 7 shows the operation stage by month, using MAPE, RMSE and ME. MAPE shows, for all models, the high error produced in June compared to the surrounding months, due to the local festivals. For all models, a higher error can be observed in December, January and October compared to the rest of the months, due to the presence of certain holidays that make the forecast harder. The similar performance of the models that include climate variables is confirmed, as well as the relative closeness of these to the $F_AL_W_DW$; however, it is important to note the significant improvement achieved by $F_AL_W_DW_SR$ for December, May and September. Additionally, a relative stability is observed in MAPE, except for June and October. If the RMSE is analyzed, it can be observed that all the models using climate variables show few high errors, while F_AL and F_AL_W present high errors for all months; $F_AL_W_DW$ behaves similarly to the climatic models when it comes to controlling high errors, except in December, which shows that climate variables contribute to a decrease in the quantity of high errors with respect to the average of the month. As far as the ME goes, the obvious resemblance between the climate models and $F_AL_W_DW$ should be noted: all of them show the maximum ME in December, January, June and October, the same months previously mentioned.

Figure 7. Evolution of *MAPE* (%), *RMSE* (dimensionless) and *ME* (W) during all days of the operation stage for each of the models.



6. Conclusions and Future Work

The recent appearance of new operation and working scenarios for electric systems (*SG/VPP/microgrids*) requires an update of the traditional forecast models used in large nation- and region-wide areas, for both the load and the *DG*. In order to work properly in this new scenario, accurate forecasts of different parameters (aggregated load, load curve, *etc.*) and time horizons (next day, next week, *etc.*) will be needed.

So far, demand forecasting models have focused on big regions or even entire countries. With the appearance of the aforementioned energy environments, the chance arises to conduct a more detailed study of the variables affecting electric load at smaller, localized areas, because more disaggregated environments allow for the use of local variables which have a bigger impact, such as climate variables or specific dates/habits, among others. This paper analyzes how aggregated load depends on both climate variables and previous days' aggregated load, so as to have a clearer picture of the variables to use in the forecasting models. This methodology can be used in future studies to identify these relations or even add new variables, so as to propose enhanced forecast models, be it for the aggregated load or other parameters of interest.

Seven demand forecasting models have been presented, each trying to improve the previous, based on the study of the variables of interest. According to the results, the best model using climate

variables has proven to be $F_AL_W_DW_SR$, whereas the best one without climate variables is $F_AL_W_DW$. We proposed to use one or the other depending on the time of year the forecast is made. Irrespective of the model, an *a posteriori* adjustment of the forecast for certain special days (local festivals and holidays) is necessary, based on complementary techniques. Hence, hybrid models will continue to be developed, using *ANN* and *fuzzy logic*.

Acknowledgements

We would like to thank AEMET (the Spanish National Meteorology Agency, Department of Environmental Affairs) for the meteorological information they so kindly provided to us, and Iberdrola, especially to Óscar Villanueva, Head of Business Exploitation in Burgos and Soria (Spain), and Ms. Silvia Herrero, Head of Communication of Iberdrola in Castilla y León (Spain) for the power demand information provided, and Javier Sanjuán, Alvaro González and Alfonso Bordegé from the University of Zaragoza for their collaboration with this study.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Zhang, Q.; Lai, K.K.; Niu, D.; Wang, Q.; Zhang, X. A fuzzy group forecasting model based on least squares support vector machine (LS-SVM) for short-term wind power. *Energies* **2012**, *5*, 3329–3346.
2. Hsu, C.-C.; Chen, C.-Y. Regional load forecasting in Taiwan—Applications of artificial neural networks. *Energy Convers. Manag.* **2003**, *44*, 1941–1949.
3. Pilo, F.; Pisano, G.; Soma, G.G. Neural Implementation of Microgrid Central Controllers. In Proceedings of the 5th IEEE International Conference on Industrial Informatics, Vienna, Austria, 23–27 June 2007; pp. 1177–1182.
4. Carpaneto, E.; Chicco, G. Probabilistic characterisation of the aggregated residencial load patterns. *IET Gener. Transm. Distrib.* **2008**, *2*, 373–382.
5. Fan, S.; Methaprayoon, K.; Lee, W.-L. Multiregion load forecasting for system with large geographical area. *IEEE Trans. Ind. Appl.* **2009**, *45*, 1452–1459.
6. Pudjianto, D.; Ramsay, C.; Strbac, G. Virtual power plant and system integration of distributed energy resources. *IET Renew. Power Gener.* **2007**, *1*, 10–16.
7. Ruiz, N.; Cobelo, I.; Oyarzabal, J. A direct load control model for virtual power plant management. *IEEE Trans. Power Syst.* **2009**, *24*, 959–966.
8. Hernández, L.; Baladrón, C.; Aguiar, J.M.; Carro, B.; Sánchez-Esguevillas, A.; Lloret, J.; Chinarro, D.; Gómez-Sanz, J.J.; Cook, D. A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants. *IEEE Commun. Mag.* **2013**, *51*, 106–113.

9. Paoletti, S.; Casini, M.; Giannitrapani, A.; Facchini, A.; Garulli, A.; Vicino, A. Load Forecasting for Active Distribution Networks. In Proceedings of the 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe), Manchester, UK, 5–7 December 2011; doi: 10.1109/ISGTEurope.2011.6162780.
10. Mousavi, S.M.; Abyaneh, H.A. Effect of load models on probabilistic characterization of aggregated load patterns. *IEEE Trans. Power Syst.* **2011**, *26*, 811–819.
11. Ipakchi, A.; Alboyyeh, F. Grid of the future—Are we ready to transition to a smart grid? *IEEE Power Energy Mag.* **2009**, *7*, 52–62.
12. Naphade, M.; Banavar, G.; Harrison, C.; Paraszczak, J.; Morris, R. Smarter cities and their innovation challenges. *Computer* **2011**, *44*, 32–39.
13. Hernández, L.; Baladrón, C.; Aguiar, J.M.; Calavia, L.; Carro, B.; Sánchez-Esguevillas, A.; Chinarro, D.; Gómez, J.; Cook, D. A Study of the relationship between weather variables and electric power demand inside a smart grid/smart world framework. *Sensors* **2012**, *12*, 11571–11591.
14. Hernandez, L.; Baladrón, C.; Aguiar, J.M.; Carro, B.; Sanchez-Esguevillas, A.J.; Lloret, J. Short-term load forecasting for microgrids based on artificial neural networks. *Energies* **2013**, *6*, 1385–1408.
15. Pérez, E.; Beltrán, H.; Aparicio, N.; Rodríguez, P. Predictive power control for PV plants with energy storage. *IEEE Trans. Sustain. Energy* **2013**, *4*, 482–490.
16. Ogliary, E.; Grimaccia, F.; Leva, S.; Musseta, M. Hybrid predictive models for accurate forecasting in PV systems. *Energies* **2013**, *6*, 1918–1929.
17. Hippert, H.S.; Pedreira, C.E.; Souza, R.C. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Trans. Neural Netw.* **2001**, *16*, 44–45.
18. Douglas, A.P.; Breiphol, A.M.; Lee, F.N.; Adapa, R. The impacts of temperature forecast uncertainty on Bayesian load forecasting. *IEEE Trans. Power Syst.* **1998**, *13*, 1507–1513.
19. Sadownik, R.; Barbosa, E.P. Short-term forecasting of industrial electricity consumption in Brazil. *J. Forecast* **1999**, *18*, 215–224.
20. Huang, S.R. Short-term load forecasting using threshold autoregressive models. *IEE Proc. Gener. Transmi. Distrib.* **1997**, *144*, 477–481.
21. Infield, D.G.; Hill, D.C. Optimal smoothing for trend removal in short term electricity demand forecasting. *IEEE Trans. Power System* **1998**, *13*, 1115–1120.
22. Sargunraj, S.; Sen Gupta, D.P.; Devi, S. Short-term load forecasting for demand side management. *IEE Proc. Gener. Transm. Distrib.* **1997**, *144*, 68–74.
23. Yang, H.T.; Huang, C.M. A new short-term load forecasting approach using self-organizing fuzzy ARMAX models. *IEEE Trans. Power Syst.* **1998**, *13*, 217–225.
24. Yang, H.T.; Huang, C.M.; Huang, C.L. Identification of ARMAX model for short term load forecasting: An evolutionary programming approach. *IEEE Trans. Power Syst.* **1996**, *11*, 403–408.
25. Yu, Z. A temperature match based optimization method for daily load prediction considering DLC effect. *IEEE Trans. Power Syst.* **1996**, *11*, 728–733.
26. Charytoniuk, W.; Chen, M.S.; Van Olinda, P. Nonparametric regression based short-term load forecasting. *IEEE Trans. Power Syst.* **1998**, *13*, 725–730.
27. Taylor, J.W.; Majithia, S. Using combined forecast with changing weights for electricity demand profiling. *J. Oper. Res. Soc.* **2000**, *51*, 77–82.

28. Ramanathan, R.; Engle, E.; Granger, C.W.J.; Vahid-Araghi, F.; Brace, C. Short-run forecasts of electricity load and peaks. *Int. J. Forecast* **1997**, *144*, 161–174.
29. Rumelhart, D.; Hinton, G.; Williams, R. Learning Internal Representations by Error Propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*; Rumelhart, D., McClelland, J.L., Eds.; MIT Press: Cambridge, MA, USA, 1996; Volume 1, pp. 318–362.
30. Broomhead, D.S.; Lowe, D. Multivariable functional interpolation and adaptive networks. *Complex Syst.* **1988**, *2*, 321–355.
31. Elman, J.L. Finding structure in time. *Cognit. Sci.* **1990**, *14*, 179–211.
32. Elman, J.L. Distributed representations, simple recurrent networks, and grammatical structure. *Mach. Learn.* **1991**, *7*, 95–126.
33. Kohonen, T. The self-organizing map. *Proc. IEEE* **1990**, *78*, 1464–1480.
34. Ramezani, M.; Falaghi, H.; Haghifam, M.-R. Short-Term Electric Load Forecasting Using Neural Networks. In Proceedings of the International Conference on Computers as a Tool, Belgrade, Serbia & Montenegro, 21–24 November 2005; Volume 2, pp. 1525–1528.
35. Razavi, S.; Tolson, B.A. A new formulation for feedforward neural networks. *IEEE Trans. Neural Netw.* **2011**, *22*, 1588–1598.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).