Electric Power Systems Research 202 (2022) 107533

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Automatic classification of special days for short-term load forecasting

Miguel López *, Carlos Sans, Sergio Valero

Miguel Hernandez University of Elche, Universidad Miguel Hernandez de Elche, Ingenieria Mecanica y Energia, Avda de la Universidad s/n, 03202 Elche, Spain

A R T I C L E I N F O

Keywords:
Load forecasting
Power demand
Holidays

A B S T R A C T

Electricity demand presents a repetitive pattern following daily, weekly and seasonal patterns. However, factors like temperature or social events tend to disrupt these patterns introducing outlying data that is difficult to forecast. This paper introduces a new methodology to classify special days without any prior knowledge of the database. Simple classification of special days into two or three categories is insufficient as the consumers' behavior has many shades on these days. However, classifying special days in a wide range of categories required a deep understanding of the consumers' behavior on different days and periods of the year. The methodology proposed describes an algorithm to automate this classification starting from a simple day-of-the-week classification and branching into as many categories as needed to fit a real database. Categories with similar profiles are merged to avoid overfitting and actual outliers are detected to ensure that no false categories are created. The methodology is developed using data from 2010 to 2017 and tested in three different systems. The benchmark used is the classification used by the Transmission System Operator in Spain and the test show that the proposed methodology provides more accurate results without the need of an expert to develop the classification.

1. Introduction

Short-term load forecasting (STLF) is a key task of a Transport System Operator (TSO) in order to ensure the efficiency and technical stability of the system while keeping the costs as low as possible. STLF is an active topic of research and so has been for several decades. Technology advances and availability of large amounts of data along with changing consumer's behavior and distributed generation make the forecasting problem more complex, making STLF research an ongoing task in an ever changing environment.

One aspect of STLF that has been overlooked on most load forecasting research is the effect of special days. On some occasions, the treatment that the proposed forecasting systems apply to special days is not detailed, and their reported errors is not disaggregated to assess the accuracy of their forecast on these special days [1]. However, holidays and special days tend to be responsible for the largest errors in forecasting systems and, therefore, represent significant losses.

STLF research focuses mainly on the design of forecasting engines. Several techniques have been used since its early beginning: linear regressive models [2–4], neural networks [2, 5, 6] and different types of artificial intelligence like evolutionary algorithms [7, 8] or fuzzy logic [5, 7, 9]. More advanced techniques are based on deep learning techniques [10] like long short-term memory (LSTM) [11] or deep convolutional neural networks (DCNN) [12]. The proposed holiday classification method is compatible with the use of any of these techniques.

However, the most relevant factor or innovation in a forecasting model may not be the forecasting engine but rather the input selection and data processing that feeds the engine. The literature review will now focus on input treatment after the superficial review of engines presented before.

Input variables are normally either environmental (weather) or socio-economic. Temperature or humidity are usually considered but there are multiple ways to pre-treat data [2, 4, 13–17]. Socio-economic trends need to be taken into account and can be included in the model by using polynomial functions of time [6] or by using moving functions of previous loads [2]. If the training period is short enough and the model is retrained frequently it can even be ignored [13]. Nevertheless, the most relevant source of daily variations in electricity demand is the type of day. The main differences stem from the days of the week, but there are other factors like national holidays, days adjacent to a holiday, partial holidays, common vacation periods and other calendar factors that cause the load profile of a day to change drastically.

The most common classification used in STLF is simply to use two or three large groups labeled as weekdays, weekends and holiday periods [18–21]. In some cases, in which special-day classification is done based

* Corresponding author.
E-mail address: m.lopezg@umh.es (M. López).

https://doi.org/10.1016/j.epsr.2021.107533
Received 15 October 2020; Received in revised form 2 July 2021; Accepted 16 August 2021
Available online 9 September 2021
0378-7796/

on the demand pattern, more groups (five [22], fifteen [23]) of special days may be used. Generally using a larger number of groups yields a more accurate forecast than using three days because the load profiles of special days may vary in a wide range [22, 24]: Fig 1 shows how both a holiday on a Thursday and the following Friday have a lower load than the same days on adjacent weeks. Nevertheless, this reduction is significantly different for the holiday than for the next day.

The day of the week not only influences regular days (Monday's profile is different than Friday's) but also changes special days [25–27]: a given holiday may alter the load profile differently if it lays on a Monday than if it lays on a Sunday. In addition, adjacency to special days may also cause variations that interact with the day of the week: holidays on Tuesdays or Thursdays cause larger reductions in the previous or following days than if the same holiday lays on a Wednesday because the adjacent day (Monday or Friday) is in between non-working days [22, 25, 26, 28].

In [27], a study of Korean holidays points out that holidays that lay on a Monday or a Saturday incur on the largest forecasting errors. Therefore, it classifies special days into four groups (Tuesdays, Wednesdays, Thursdays and Fridays; Saturdays; Mondays; Sundays). This work only classifies according to the day of the week and, even though it reduces the forecasting error, it appears that using a wider range of factors would yield a more accurate forecast. In [24], the day of the week and the month are used to model the effect of seasonality on electricity consumption. Furthermore, a set of three variables are included to distinguish the effect of holidays, days after a holiday and Easter holidays. A classification based on the profile of the curves is used in [28]. It determines that three categories are needed (weekdays; Saturdays; Sundays and holidays). This work proposes a set of rules that take into account the vicinity of the forecast day to a holiday but it does not distinguish among different types of holidays. Five different categories are used in [22]: weekdays, Saturdays, Sundays, Mondays, holidays. The proposed model uses a separate NN for each category and a fuzzy inference model forecasts the maximum and minimum of the load profile. One problem this configuration may encounter if more categories are defined is that, as the categories become more exclusive, fewer data fit each category and, therefore training data become scarce. Self-Organizing Maps (SOM) are used as a classifier of special days in [23]. A NN is used after classification for forecasting. The number of groups that the SOM assigns vary between 11 and 15. This method has not been considered as it uses prior knowledge regarding the type of special days and the output of the classification is not clear. The classification schemes shown in [25, 29] are somewhat similar in the sense that they both consider whether a special day lay on the same date or the same day of the week. However, the number of rules applied in [25] is four while the number of categories described in [29] and the number of days targeted as special is much larger. In [26], a similar rule-based approach is followed and applied to the French electric system, however it is also limited to seven categories. Seven categories are also used in [30]: four for common holidays and for special national holidays. This literature review includes several examples of unsupervised or rule based classification methods that would not require an extensive knowledge of the database in order to be applied to a new system. However, these classifications are limited to a low number of categories.

In [1], several modeling techniques are applied to the German system. The methods tested use a variety of binary variables representing different types of holidays and it presents the idea of using wider classifications. However, the actual classification of the holidays is taken as an input. An hourly classification is presented in [31], where two parameters are combined to code different degrees of load reduction by the hour. The classification presented in [32] for the holidays in China differentiates between 3-day and 7-day holidays. In addition, it also includes the ordinal of each day within the holiday. These other examples, use slightly more complex classifications and hint that more detailed classifications would outperform simple rule systems. Nevertheless, these expert systems require a thorough analysis of the database, making the design process much more complex. The main objective of this paper is to provide a methodology capable of determining an expert-level classification of special days on a database without any prior knowledge. This classification will be based solely on the date and, therefore, can be applied beforehand and introduced as input to the forecasting engine.

The scope of this paper and its contribution can be summarized as follows:

- To provide a classification of special days that results in a forecast at least as accurate as the one obtained using an expert classification.
- To describe an algorithm that obtains such classification of special days from the hourly load time series data.
- To test the performance of the classification of 3 different databases against expert classifications currently used by the national TSO, Red Eléctrica de España (REE) and prove that it outperforms them with much less design effort.

Section 2 of this paper will describe the forecasting model, the databases and the algorithm. On section 3, the obtained classification is presented along with accuracy results on several databases. Finally, section 4 includes the final analysis and conclusions.

2. Materials and methods

This research work branches off the development of a new

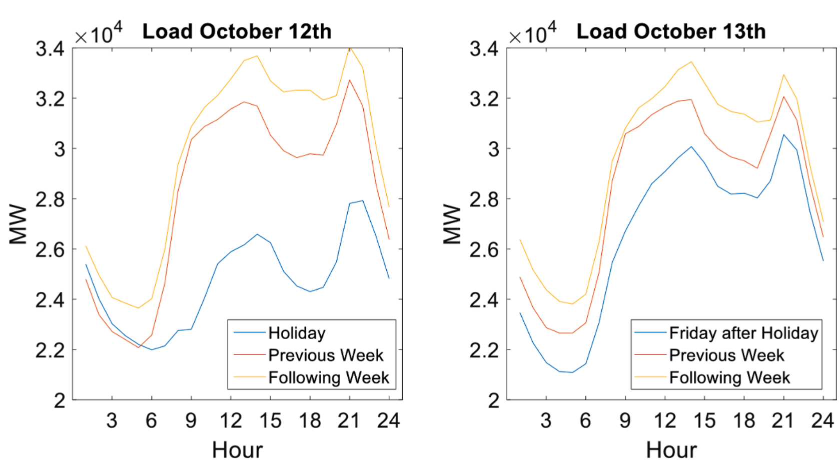

Fig. 1. a) Load profile of national holiday (October, 12th 2017) compared to the same weekday on the previous and following weeks (LEFT). b) Load profile of a Friday after a holiday (October, 13th 2017 compared to the same weekday on the previous and following weeks (RIGHT).

forecasting model for the Spanish TSO (REE) that focused on improving accuracy on special days. The classification developed for that occasion [29] is currently running at the TSO's headquarters and required deep analysis of the load profiles during national and regional holidays, days adjacent to holidays, vacation periods or social events…

That classification will be used as benchmark and the forecasting model will be common to both benchmark and proposed classification. This section starts describing the forecasting model used, the databases and other inputs to the model. Then, it will focus on the process to obtain the classification, the parameters involved and the design choices that arose in the process.

2.1. Model structure

The model comprises two forecasting engines that share the same input and provide hourly forecasts for the current day and the nine following days. In order to provide a more comparable result, this paper will deal with next-day forecasts made at 9am which is found on most works. This paper will focus on the auto-regressive component of the model, as it is easier to analyze its output and the effect of each variable. Nevertheless, the other forecasting engine based on an auto-regressive neural network (NARX) will also be used in the testing phase to understand the effect that the new classification has on the whole model. The following paragraphs show the basic concepts of the AR model and the modifications made for this specific study. For more details on the model the reader can refer to [2].

The forecast of a full 24-hour profile consists on 24 individual forecasts. Therefore, there are 24 sub-models that use the same input but different ouput. During training each model assigns different relevance to each variable as, for example, temperature may cause more significant load variations in the early evening than at mid-night. This permits to obtain a set of 24 coefficients that conform a profile of the effect of any variable along the day.

The auto-regressive model has an exogenous and an auto-regressive (AR) part. The AR component reduces the forecasting error because it captures factors not included in the model but whose effects lag for several samples. This is very useful in forecasting but it may mask the true effect of the predictors used. Therefore, the AR part is removed from the model for the design stage but it is later reinstated to obtain actual forecasts for benchmark:

2.2. Data analysis

2.2.1. Load data

The data for the electricity demand of the Spanish inland system from 2010 to 2018 [33]. This database includes hourly values for the aggregate system and for the different regions that compose it. In this research the inland system will be used to design, train and test the system. Data from Madrid and Catalonia will also be used for testing as another out-of-sample case. It is important to use databases that span long periods because modeling special days requires to have multiple instances of each type of special days, laying on different weekdays. This requires a proper treatment of the long-term trend of the data because relatively old data (7-years old) will be included in the model. The input of the model based on load values has two components: a 52-weeks moving average of previous loads to capture long trends and the last-known value to include recent trends. This last value is not included in the modeling part of this research as it may mask the effect of the predictors.

2.2.2. Temperature data

Temperature data for the whole country is obtained from different weather stations. Maximum and minimum temperature forecasts for these stations for the next 10 days are received every day [34]. The information from all stations is redundant as the series are very highly correlated. The selection of the relevant stations, the linearization of the data and the number of lags to be included in the model is described in [2].

2.2.3. Calendar data

The benchmark model used by REE employs a classification including 53 binary variables to specify the type of day. These variables are thoroughly described in [29] but a brief description is included here as it will serve as the main comparison to assess the performance of the proposed methodology:

- National special days: These are days whose profile is very specific and is not comparable to any other holiday. It includes 11 types of days described by day and month (January, 1st; December, 25th…$^{)}$, and 13 days related to Easter (Good Friday, Easter Monday…).
- National regular holidays: Three categories are used to assign the rest of the holidays found in the Spanish National Gazette. The variables distinguish if the holiday lays on weekend, Monday or Tuesday through Friday.
- Days before and after a holiday: The previous categories affect the adjacent days. In order to address this phenomenon, 4 variables are used to classify days before a holiday (Monday), before a holiday (Tuesday-Friday), after a holiday (Friday) and after a holiday (Monday-Thursday).
- Periods affected by Christmas: The weeks surrounding Christmas are also affected in different ways. Four variables are assigned to days Dec, 20th to Dec, 23rd, one variables for workdays Dec, 27th to Dec, 29th, and another variable for workdays Jan, 2nd to Jan, 5th.
- Summer period: Most Spanish people take some of their vacation during the summer time. To model this behavior, 3 more variables are used to distinguish between the four days of August.
- Regional holidays: Partial holidays are published in the regional gazettes. They are included in the model by a single variable ranging from 0 to 1 representing the fraction of the nation's GDP that the regions on holiday represent.

This classification was created by an expert by analyzing the errors of the model on specific periods and establishing patterns over the years. Other classifications conveying similar information but condensed in less variables have been tested on the neural-network model as this model is expected to deal with more complex relations among variables. However, the most accurate classification for both auto-regressive and neural-network models was the one described here. A classification must consider the following aspects:

- Each category must have sufficient instances.
- All instances within a category must be similar among them.
- All categories must be different among them.

The first two aspects are needed to ensure that abnormalities do not become categories and that only repetitive patterns are treated as such. The third aspect is required to avoid an unnecessarily large number of categories.

This analysis requires deep knowledge and many hours of work and, while it is significantly more accurate than simpler classification it is difficult to justify on smaller systems like regional or city-wide forecasts.

The proposed methodology will provide a classification similar in structure, as it will also use binary variables, but whose application does not require of any prior knowledge of the database nor any external documentation describing holidays.

2.3. Description of the algorithm

The methodology proposed consists on an algorithm that iteratively improves an initial classification by identifying special days not included and configuring the classification to fit their profile. The algorithm is described in Fig. 2, and it can be summarized in three stages: selection of

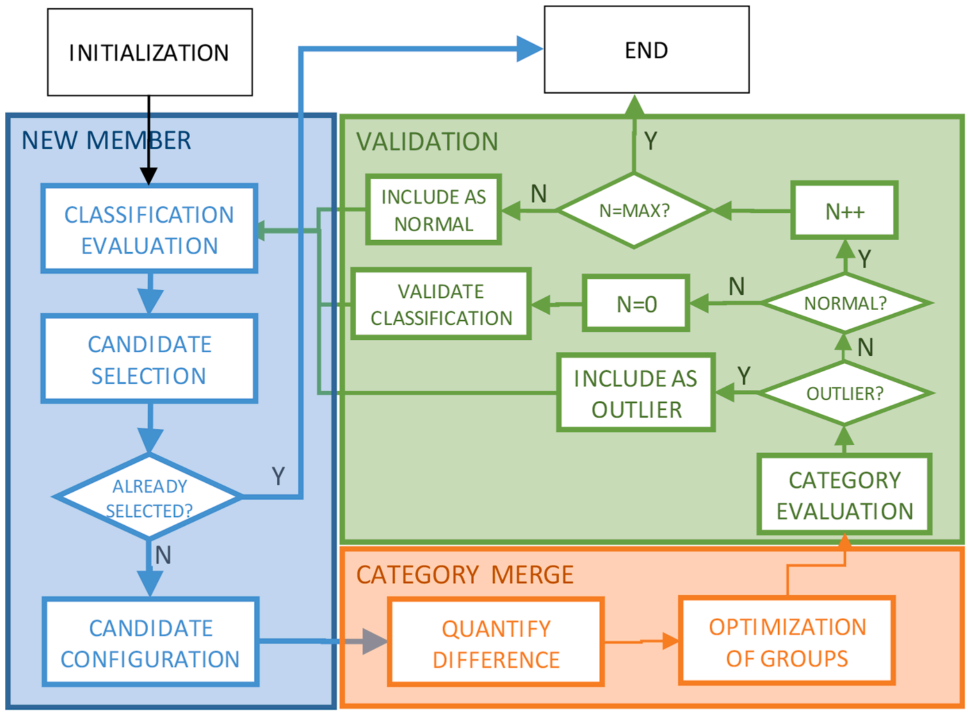

Fig. 2. Flow chart of the proposed algorithm.

new types of special days, merge of similar types of special days and validation of special days. The algorithm uses data from 2010 to 2017 to optimize the classification. Then data from 2018 is used to test the resulting classification in comparison with the benchmark.

2.3.1. *Available categories*

The definition of a special day requires that a certain pattern is followed over the years. Special days caused by strikes or extraordinary events cannot be forecasted accurately because there are no previous examples from which to learn the behavior of the load. Therefore, special days can be defined as either having a specific date (day and month) or as a number related to a fixed event (Good Friday or Monday after DST is implemented). Following these definition, the possible categories that can be considered special days are detailed in Table 1. Other days can be specified as references for the second type to generate new possible categories.

A category may be split into several sub-categories according to the day of week. These sub-categories may or may not be considered as special days. These possibilities are discussed later. All 365 days are considered candidates to type 1 special days. The proposed methodology will determine which ones actually have a specific load profile.

2.3.2. *Initial classification*

The initial classification considers merely the day of the week. The day of the week is a known factor affecting the load profile and, therefore is included by default. The initialization of the algorithm also includes filtering missing or corrupt data, which is done following the rules described in [2].

Table 1
Possible special days categories of types 1 and 2.

Type	Definition		Example	Total
1	Month	Day of month	Jan, 6th; Dec, 25th	365
2	Reference (Good Friday)	# of days to / from reference (-6 to 7)	Good Friday (0); Easter Monday (3)	14

2.3.3. *Evaluation*

Whenever a new classification is proposed, a linear regressive forecasting model is trained by using the established input (long-term, temperature, month and classification variables) from years 2010 to 2017 using (1). Then, the residuals are used to calculate a mean absolute percentage error (MAPE) for each day and each possible category (2 and 3). At this point, categories of type 1 and 2 may have days in common (Good Friday may also be a March, 30th), but this does not cause any difficulty.

$$ln(y_h) = X_{exo}\cdot\beta_h + X_{cl}\cdot\gamma_h + \varepsilon_h \tag{1}$$

$$MAPE_{day} = \frac{\sum_1^{24}|\varepsilon_h/y_h|}{24} \tag{2}$$

$$MAPE_{cat} = \frac{\sum_1^{n} MAPE_{day}}{n} \tag{3}$$

where y_h is the load at hour h, X_{exo} is the matrix of exogenous variables, β_h is the vector of coefficients for hour h, X_{cl} is the matrix of the binary classification of days, γ_h is the vector of coefficients for each category at hour h and ε_h is a vector containing the residuals.

2.3.4. *Candidate selection*

Among all possible categories, the one with the highest MAPE is selected as a candidate to be a special day. Categories that have been optimized already or those marked as containing outliers or marked as normal are not eligible.

2.3.5. *Candidate configuration*

Configuration of the candidate refers to defining how the day of the week should be considered. To this end, two types of categories are defined:

- Priority categories: In these cases, the category may split the days by the day of the week but all days within a group are assigned the same profile.

- Modifying categories: The category may also split by the day of the week. However, the day-of-the-week variable remains active, causing each day of the week within a group to have its own profile.

The configuration of a category entails, on one part, determining whether it is a priority or a modifying category and, on the other part, establishing the grouping of different days of the week.

In order to do this, the category is evaluated using common groupings by days of the week and both priority and modifying types. The result of this process for Christmas Day and Dec, 22nd is shown in Table 2 and Fig. 3.

These two categories exemplify how very specific days like Christmas have one profile (maybe two) regardless of the day of the week and, therefore, should be configured as priority categories. However, another type of days (Dec. 22nd) do not define the profile by themselves but they modify the profile of the day of the week and should be configured as modifiers. In the case of Christmas, all weekdays have the same profile and Sat. and Sun. also share their own profile. However, Dec. 22nd modifies the standard profile of the corresponding days of the week: all weekdays are modified in the same way and Sat. and Sun. are also modified in the same way. The reference for this coefficient is a regular Saturday, which would have a straight line equal to 1.

2.3.6. Merging of categories

Merging similar categories increases the amount of examples they contain which, if the merged categories are truly compatible, can increase accuracy of the forecast. Therefore, in order to find groups of categories suitable to be merged together, the algorithm compares the profile of all sub-categories using a mean absolute percentage difference specified in (4):

$$DIF_{i,j} = \frac{\sum_1^{24} \left|\gamma_{i,h} - \gamma_{j,h}\right| / \gamma_{j,h}}{24} \tag{4}$$

where $\gamma_{i,h}$ is the coefficient for the sub-category i at hour h. Profiles whose difference is lower than a similarity threshold (S_{th}) are paired together as compatible sub-categories. Considering this definition, sub-categories may belong to several pairs which may or may not be compatible among them. The selection of S_{th} is discussed in 2.3.8.

2.3.6.1. Optimization of the groups. The pairs found in the previous section are used to form larger groups of compatible sub-categories. These larger groups must comply the following rules:

- All sub-categories in a group must be compatible between them.
- Any sub-category can be in only one group.

The application of these rules results in multiple possible groups from which the algorithm selects the optimal combination. However, the number of possible groups grows exponentially and it is not possible to evaluate all combinations. To avoid this, an evolutionary heuristic is proposed. This heuristic's flow chart is included in Fig. 4.

The loop is run every time a new category is configured and compatible sub-categories have been detected. A population of N candidates is evaluated on every loop. A candidate is a list of groups, each group containing a number of sub-categories that are compatible among them and that will be merged together:

- INITIALIZATION: The initial population includes the previous winner if the algorithm has been run before
- EVALUATION: Candidates from initialization or previous loops are included in the current population. The population is filled with random candidates until the number of candidates equals n. Each candidate is evaluated by their corresponding MAPE. In order to stimulate the merging of the sub-categories, each merged category discounts a certain amount (C_{tot}) from the reported MAPE. The score of a candidate is described in (5). The selection of C_{tot} is discussed in 2.3.8.

$$SCORE_i = MAPE - C_{tot} \cdot n_{merged} \tag{5}$$

- NEXT GENERATION: The candidates in first quintile (Q_5-1) are passed on to the next generation. The rest of candidates are obtained from the Q_5-1 candidates by randomly adding a new sub-group, eliminating a present sub-group and modifying a present sub-group. This new generation is passed on to the evaluation stage.
- EXIT: The algorithm finishes if the winning candidate repeats 5 times.

The result of the heuristic is a list of several groups of sub-categories to be merged together.

2.3.7. Validation

The final step is to determine whether the final result is actually a properly identified special day, an outlier or just a poorly forecasted regular day:

- Normal day: Error is reduced less than 10% and starting error is lower than 1.5 times the average error for all categories. If considering it a special day does not cause a significant improvement and the forecasting error is not overly inaccurate, then it is considered a normal, poorly forecasted day.
- Outlier: Error on half or more of the days within the category is lower than 1.5 times the average error for all categories and error on the worst day is three times the average error on the same day of the week and category. If most days within the category are not overly inaccurate and the largest error is disproportionately large, then it is considered as an outlier.
- Special day: All other days.

If a category is found to be an outlier or a normal day, it is removed from the possible categories to choose the next candidate from, it is not added to the list of special days and all its sub-categories are removed from the mergers list.

The algorithm stops if three normal days are identified consecutively or if the candidate selected has already been optimized.

2.3.8. Parameter selection

The merging parameter S_{th} determines how similar two profiles must be to be considered compatible. Smaller values will form categories with a low number of instances that are very similar among them while larger number will have the opposite effect. The second merging parameter C_{tot} has a similar effect: smaller values encourage using more variables while larger values acknowledge the advantage of reducing the number of variables in the model. In order to avoid empty categories when the

Table 2
Modelling error for two special days with all possible configurations.

DATE	TYPE	ALL	MON-FRI & SAT-SUN	MON-SAT	MON-FRI & SAT	MON & TUE-FRI	MON-THU & FRI
DEC, 25TH	MODIFIER	7.44%	2.80%	6.73%	5.22%	6.99%	7.46%
	PRIORITY	2.26%	2.18%	5.29%	4.90%	7.01%	7.07%
DEC, 22ND	MODIFIER	2.08%	1.96%	2.00%	1.91%	1.79%	1.91%
	PRIORITY	6.67%	2.74%	4.80%	2.40%	1.89%	2.31%

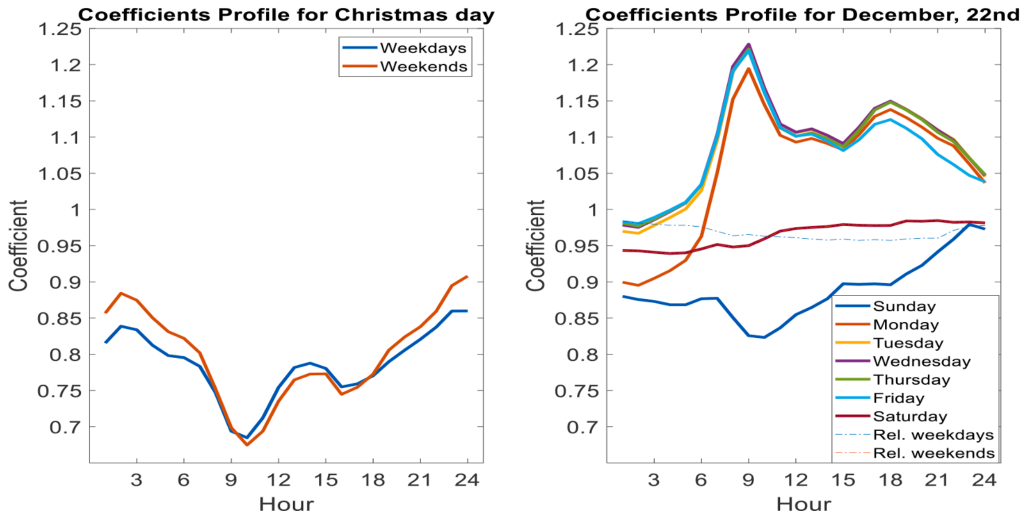

Fig. 3. a) Coefficient profile for Christmas days on weekend and weekdays (LEFT). b) Coefficient profile for Dec. 22nd, on all days of the week. Relative profiles represent the effect of Dec. 22nd, on weekdays and weekends.

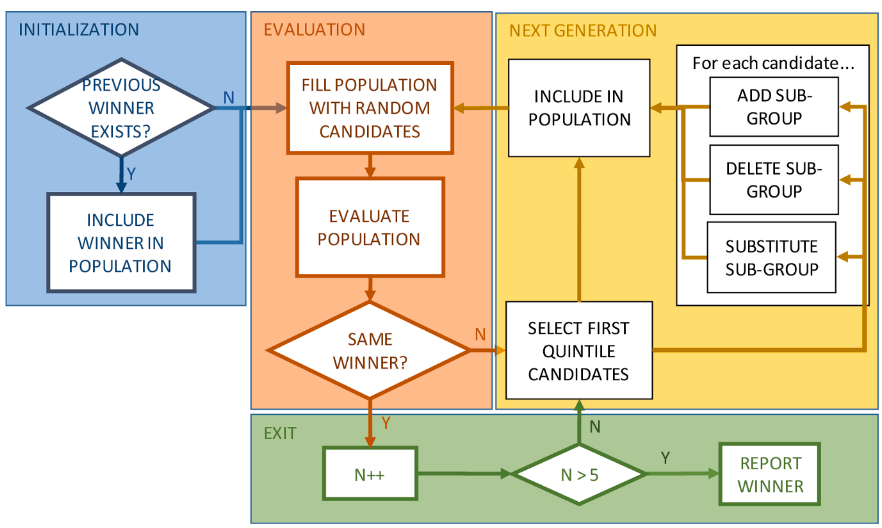

Fig. 4. Flow chart for the grouping optimization heuristic.

training period slides forward both parameter should be selected so that no categories have less than 3 members. In addition, they have been optimized to achieve the lowest MAPE, obtaining values of 0.02 for S_{th} and 0.0005 C_{tot}. The same values were obtained for Madrid and Catalonia data bases. Still, it is possible that some less populated sub-categories remain unmerged after the heuristic. To avoid this issue, each sub-category with less than three members is merged with the sub-category that matches its profile best.

The parameters used in validation were chosen empirically. Non-periodical events like nation-wide strikes or black-outs may be picked-up as candidates to become a category but the conditions for outlier identified properly these cases and ruled them out. In addition, when all special days are identified, the algorithm will attempt to consider normal days as special, however, they will not become a new category if the conditions are not met. The obtained values were valid for all three databases, nevertheless, they could be modified as needed when applied to other databases in order to meet their described goals.

3. Results

The application of the algorithm to the Spanish load produced the classification described in Table 3. The table represents the types of days that belong to each of the 48 merged subcategories and that will be included in the model as binary variables. Some variables include up to 9 sub-categories because their profile is similar (many days in August have similar profiles) while other variables only represent one sub-category because it has a very specific profile (Christmas day). The representation of the classification is cumbersome because of its empiric nature. Nevertheless, the following results show its adequacy for accurate forecasting.

3.1. Training results

The training (2010-2017) results for both benchmark and proposed classifications are shown in Table 4 and Table 5. The first one splits the results according to both classifications, showing how both classifications perform on the days considered as special and as regular by each of them. The second one divides the result by the nature of the holiday.

Table 4 shows that the automatic classification outperforms the expert in both special categories and obtain very similar results in normal days. It comes to the attention that the expert classification considers special more than twice the amount of days that the automatic classification does. This is because the expert uses a variable to model provincial and regional variables. This variable takes the value of the

Table 3
Classification of special days by the proposed algorithm.

#	MEMBERS												
1	1-Nov	4-Jan	10-Aug	9-Aug	22-Aug	11-Aug	11-Sep	23-Aug	24-Aug				
	Sun&Sat	Mon-Thu	Mon-Fri	Mon-Fri	Mon-Fri	Sun&Sat	Tue-Fri	Mon-Fri	Sun&Sat				
2	13-Aug	19-Mar	8-Aug	8-Aug	7-Aug	7-Aug	23-Aug	6-Aug					
	Sun&Sat	Mon-Thu	Mon-Fri	Sun&Sat	Mon-Fri	Sun&Sat	Sun&Sat	Sun&Sat					
3	8-Dec	6-Dec	19-Aug	5-Jan	24-Jun	6-Jan	19-Mar						
	Sun&Sat	Sun&Sat	Mon-Fri	Sun&Sat	Mon-Fri	Sun&Sat	Fri						
4	12-Oct	3-Jan	18-Aug	18-Aug	10-Aug	5-Aug							
	Sun&Sat	Mon-Fri	Mon-Fri	Sun&Sat	Sun&Sat	Mon							
5	11-Aug	6-Aug	24-Aug	21-Dec	5-Aug	8	17-Aug	13-Aug	20-Aug	21-Aug			
	Mon-Fri	Mon-Fri	Mon-Fri	Sun&Sat	Tue-Fri		Mon-Fri	Mon-Fri	Mon-Fri	Sun&Sat			
6	15-Feb	25-Jul	24-Jun	15-Feb		9	16-Aug	9-Dec	2-Nov	4-Jan			
	Mon-Fri	Sun&Sat	Sun&Sat	Sun&Sat			Mon-Fri	Fri	Fri	Fri			
7	01-may-P	01-may-P	14-ago-P	23-Dec		10	14-ago-P	31-oct-P	30-abr-P				
	Mon-Fri	Sun&Sat	Sun&Sat	Mon			Mon-Fri	Mon	Mon				
11	28-Dec	7-Jan	3-Jan	13	2-Nov	9-Dec	13-Oct	15	7-Dec	23-Dec	15-Aug		
	Sun&Sat	Sun&Sat	Sun&Sat		Mon-Thu	Mon-Thu	Mon-Fri		Tue-Fri	Tue-Fri	Sun&Sat		
12	30-Dec	29-Dec	28-Dec	14	01-ene-P	01-ene-P	25-dec-P	16	31-dec-P	24-dec-P			
	Mon-Fri	Mon-Fri	Mon-Fri		Mon-Fri	Sun&Sat	Sun&Sat		Sun&Sat	Sun&Sat			
17	12-Oct	8-Dec	21	27-dec-P	12-ago-P	25	30-abr-P	4-Aug	29	02-may-P	7-Jan		
	Mon-Fri	Mon-Fri		Mon-Fri	Mon		Tue-Fri	Fri		Sun&Sat	Mon-Fri		
18	30-Dec	13-Oct	22	29-Dec	16-Aug	26	20-Aug	11-Sep	30	9-Aug	22-Dec		
	Sun&Sat	Sun&Sat		Sun&Sat	Sun&Sat		Sun&Sat	Mon		Sun&Sat	Sun&Sat		
19	22-Dec	25-Jul	23	17-Aug	22-Aug	27	19-Aug	21-Dec	31	31-jul-P	04-ago-P		
	Mon-Fri	Mon-Fri		Sun&Sat	Sun&Sat		Sun&Sat	Mon-Fri		Mon-Thu	Mon-Thu		
20	12-ago-P	21-ago-P	24	26-dec-P	27-dec-P	28	02-ene-P	GF* +1	32	6-Dec	7-Dec		
	Tue-Fri	Mon-Fri		Sun&Sat	Sun&Sat		Sun&Sat			Mon-Fri	Mon		
33	25-dec-P	36	GF*	39	6-Jan	42	15-Aug	44	26-dec-P	46	GF*-1	48	31-dec-P
	Mon-Fri				Mon-Fri		Mon-Fri		Mon-Fri				Mon-Fri
34	1-Nov	37	24-dec-P	40	GF* +2	43	02-ene-P	45	02-may-P	47	5-Jan		
	Mon-Fri		Mon-Fri				Mon-Fri		Mon-Fri		Mon-Fri		
35	31-oct-P	38	GF* -2	41	GF* +3								
	Tue-Fri												

* GF: Good Friday.

Table 4
Modeling error on special days for expert and automatic classification.

	SPECIAL 1*	NORMAL 1*	ALL	SPECIAL 2**	NORMAL 2**	% SPECIAL DAYS
EXPERT	2.06%	1.95%	1.99%	2.36%	1.92%	34%
AUTOMATIC	2.03%	1.94%	1.97%	2.08%	1.95%	16%
PROPOSED	1.96%	1.93%	1.94%	2.03%	1.93%	31%

* Days considered as special or normal by the expert classification.
** Days considered as special or nomarl by the automatic classification.

Table 5
Modeling error on each type of special days for expert and automatic classification.

	CHRISTMAS	SUMMER	EASTER	NAT'L HOLIDAYS	REG'L HOLIDAY	ADJACENT	NORMAL
EXPERT	2.56%	2.20%	2.08%	2.79%	1.88%	2.17%	1.90%
AUTOMATIC	2.27%	2.07%	2.15%	2.17%	1.95%	2.61%	1.90%
PROPOSED	2.26%	2.05%	2.16%	2.26%	1.82%	2.31%	1.90%
% OF DAYS	5%	17%	2%	2%	18%	2%	58%

percentage of the GNP that the region on holidays represents. Many of the days included in this variable are not picked up by the automatic classification because the impact on national load is negligible. However, this variable can be included in the automatic classification without violating its principle of not requiring an expert as it is readily available information. Therefore, the proposed classification adds this variable which obtains the best results of the three in all categories.

The results are also categorized by the nature of the holiday to better understand its performance. The proposed methodology obtains the best results except in Easter and adjacent days. However, they amount for only 4% of the total of days. It is also worth noticing that the worst performing group drops from 2.79% in the case of the expert (National Holidays) to 2.31% in the case of the proposed method (adjacent days). This shows how maximum forecasting errors are reduced.

3.2. Forecasting results

The proposed methodology is tested with out-of-sample data from year 2018. The results are shown in Table 6. The performance of the benchmark system currently in use at REE is included. This system uses an AR and a NARX models which are combined to provide the definite forecast. The output of the AR, NARX and combined models are included

Table 6
Forecasting error on each type of special days for expert and automatic classification.

		CHRISTMAS	SUMMER	EASTER	NAT'L HOLIDAYS	REG'L HOLIDAY	ADJACENT	ALL SPECIAL	NORMAL
AR	EXP	2.37%	1.69%	1.78%	2.57%	1.43%	1.82%	1.68%	1.51%
	AUT	2.29%	1.35%	1.67%	1.79%	1.33%	1.73%	1.49%	1.43%
NARX	EXP	2.27%	1.62%	2.06%	2.10%	1.34%	1.63%	1.62%	1.44%
	AUT	2.30%	1.68%	2.05%	2.01%	1.35%	1.96%	1.64%	1.45%
COMB	EXP	2.11%	1.50%	1.82%	2.14%	1.23%	1.44%	1.49%	1.32%
	AUT	2.11%	1.35%	1.70%	1.77%	1.21%	1.76%	1.42%	1.31%

using the expert and the proposed automatic classification.

The results show that the proposed classification increases accuracy on special days forecasted by the AR model (1.68% vs. 1.49%) while the result of both NARX methods are equivalent (1.62% vs. 1.64%). In the combined result, which is the actual output, the proposed methodology outperforms the expert classification (1.49% vs. 1.42%). The accuracy improvement is not consistently focused on any time of day but rather evenly distributed. Nevertheless, even though an increased accuracy is very welcome, the most important feature of the proposed methodology is that the classification is obtained without human input, apart from the list of provincial holidays.

3.3. Other databases

The proposed methodology has been tested in regional databases from Madrid and Catalonia. The results coincide with those obtained from the national system: the automatic classification yield a slightly more accurate but statistically equivalent accuracy (2.58% vs 2.55% in Madrid and 3.46% vs. 3.41% in Catalonia) and the number of variable is increased by 10 and 12. Provincial holidays were not needed as a separate input because the algorithm was able to identify them on its own. This is because, at a regional scale, provincial holidays have a larger effect. Nevertheless, if holiday information at a lower level (larger cities or municipalities) was available it could be included in the model in the same way provincial holidays were used in the national model.

4. Conclusions

The problem of modeling special-days load profiles is tackled. The common approach to this problem is to classify each day into foreknown categories. The rules for this classification can be explicit or set by an expert. However, in any case, a deep knowledge of the consumer's behavior is needed to establish the categories.

The methodology proposed establishes a way to classify the days without prior knowledge of the number or nature of the categories. The algorithm is based on the detection of outlying patterns in the load from previous years. Therefore, it is necessary to provide at least 7 years of training data. The algorithm has been thoroughly described by its application to the Spanish inland electric system. The example has been used to describe how to tune the parameters to the database and to carry out each step. Nevertheless, the methodology has been applied to obtain not only a classification for the national system but also for different out-of-sample databases (Madrid and Catalonia).

The classifications have been tested by using the forecasting models currently in use at the TSO's headquarters. A classification made by an expert is currently in use and is considered the benchmark. The comparison of the benchmark classifications (expert) and the proposed (obtained by the new methodology) shows that, while the former one requires more design effort, the second one is more accurate in all three cases. Therefore, the proposed system can be used to effortlessly improve accuracy on special days.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study has been funded by Red Eléctrica de España (REE) as part of its Research and Development activities with Universidad Miguel Hernández (reference REE1.20SW).

References

[1] F. Ziel, Modeling public holidays in load forecasting: a German case study, J. Mod. Power Syst. Clean Energy 6 (2) (2018) 191–207, https://doi.org/10.1007/s40565-018-0385-5. Mar.
[2] M. López, S. Valero, A. Rodriguez, I. Veiras, C. Senabre, New online load forecasting system for the Spanish transport system operator, Electr. Power Syst. Res. 154 (2018) 401–412, https://doi.org/10.1016/j.epsr.2017.09.003.
[3] N. Charlton, C. Singleton, A refined parametric model for short term load forecasting, Int. J. Forecast. 30 (2) (2014) 364–368, https://doi.org/10.1016/j.ijforecast.2013.07.003.
[4] P. Wang, B. Liu, T. Hong, Electric load forecasting with recency effect: a big data approach, Int. J. Forecast. 32 (3) (2016) 585–597, https://doi.org/10.1016/j.ijforecast.2015.09.006.
[5] Z. Yun, Z. Quan, S. Caixin, L. Shaolan, L. Yuming, S. Yang, RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment, IEEE Trans. Power Syst. 23 (3) (2008) 853–858, https://doi.org/10.1109/TPWRS.2008.922249. Aug.
[6] M. López, S. Valero, C. Senabre, J. Aparicio, A. Gabaldon, Application of SOM neural networks to short-term load forecasting: the Spanish electricity market case study, Electr. Power Syst. Res. 91 (2012) 18–27, https://doi.org/10.1016/j.epsr.2012.04.009.
[7] V.H. Hinojosa, A. Hoese, Short-term load forecasting using fuzzy inductive reasoning and evolutionary algorithms, IEEE Trans. Power Syst. 25 (1) (2010) 565–574, https://doi.org/10.1109/TPWRS.2009.2036821. Feb.
[8] J. Wang, S. Jin, S. Qin, H. Jiang, Swarm intelligence-based hybrid models for short-term power load prediction, Math. Probl. Eng. 2014 (2014) 17.
[9] G.-C. Liao, T.-P. Tsao, Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short-term load forecasting, IEEE Trans. Evol. Comput. 10 (3) (2006) 330–340, https://doi.org/10.1109/TEVC.2005.857075. Jun.
[10] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of smart meter data analytics: applications, methodologies, and challenges, IEEE Trans. Smart Grid 10 (3) (2019) 3125–3148, https://doi.org/10.1109/TSG.2018.2818167. May.
[11] B.-S. Kwon, R.-J. Park, K.-B. Song, Short-term load forecasting based on deep neural networks using LSTM layer, J. Electr. Eng. Technol. 15 (4) (2020) 1501–1509, https://doi.org/10.1007/s42835-020-00424-7. Jul.
[12] W. He, Load forecasting via deep neural networks, Procedia Comput. Sci. 122 (2017) 308–314, https://doi.org/10.1016/j.procs.2017.11.374.
[13] J. Moral-Carcedo, J. Pérez-García, Time of day effects of temperature and daylight on short term electricity load, Energy 174 (2019) 169–183, https://doi.org/10.1016/j.energy.2019.02.158. May.
[14] N. Zhang, Z. Li, X. Zou, S.M. Quiring, Comparison of three short-term load forecast models in Southern California, Energy 189 (2019), 116358, https://doi.org/10.1016/j.energy.2019.116358. Dec.
[15] S. Haben, G. Giasemidis, F. Ziel, S. Arora, Short term load forecasting and the effect of temperature at the low voltage level, Int. J. Forecast. 35 (4) (2019) 1469–1484, https://doi.org/10.1016/j.ijforecast.2018.10.007. Oct.
[16] Y. Wang, J.M. Bielicki, Acclimation and the response of hourly electricity loads to meteorological variables, Energy 142 (2018) 473–485, https://doi.org/10.1016/j.energy.2017.10.037. Jan.
[17] M. López, S. Valero, C. Senabre, A. Gabaldón, Analysis of the influence of meteorological variables on real-time Short-Term Load Forecasting in Balearic Islands, in: in 2017 11th IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), 2017, pp. 10–15, https://doi.org/10.1109/CPE.2017.7915137. Apr.
[18] S. Fan, L. Chen, Short-term load forecasting based on an adaptive hybrid method, IEEE Trans. Power Syst. 21 (1) (2006) 392–401, https://doi.org/10.1109/TPWRS.2005.860944. Feb.

[19] K.Y. Lee, Y.T. Cha, J.H. Park, Short-term load forecasting using an artificial neural network, IEEE Trans. Power Syst. 7 (1) (1992) 124–132, https://doi.org/10.1109/59.141695. Feb.
[20] S. Rahman, R. Bhatnagar, An expert system based algorithm for short term load forecast, IEEE Trans. Power Syst. 3 (2) (1988) 392–399, https://doi.org/10.1109/59.192889. May.
[21] W. Christiaanse, Short-term load forecasting using general exponential smoothing, IEEE Trans. Power Appar. Syst., vol. PAS-90 (2) (1971) 900–911, https://doi.org/10.1109/TPAS.1971.293123. Mar.
[22] Kwang-Ho Kim, Hyoung-Sun Youn, Yong-Cheol Kang, Short-term load forecasting for special days in anomalous load conditions using neural networks and fuzzy inference method, IEEE Trans. Power Syst. 15 (2) (2000) 559–565, https://doi.org/10.1109/59.867141. May.
[23] R. Lamedica, A. Prudenzi, M. Sforna, M. Caciotta, V.O. Cencellli, A neural network based technique for short-term forecasting of anomalous load periods, IEEE Trans. Power Syst. 11 (4) (1996) 1749–1756, https://doi.org/10.1109/59.544638. Nov.
[24] A. Pardo, V. Meneu, E. Valor, Temperature and seasonality influences on Spanish electricity load, Energy Econ 24 (1) (2002) 55–70, https://doi.org/10.1016/S0140-9883(01)00082-2. Jan.
[25] S. Arora, J.W. Taylor, Short-term forecasting of anomalous load using rule-based triple seasonal methods, IEEE Trans. Power Syst. 28 (3) (2013) 3235–3242, https://doi.org/10.1109/TPWRS.2013.2252929. Aug.
[26] S. Arora, J.W. Taylor, Rule-based autoregressive moving average models for forecasting load on special days: a case study for France, Eur. J. Oper. Res. 266 (1) (2018) 259–268, https://doi.org/10.1016/j.ejor.2017.08.056. Apr.
[27] K.-B. Song, Y.-S. Baek, D.H. Hong, G. Jang, Short-term load forecasting for the holidays using fuzzy linear regression method, IEEE Trans. Power Syst. 20 (1) (2005) 96–101, https://doi.org/10.1109/TPWRS.2004.835632. Feb.
[28] D. Srinivasan, C.S. Chang, A.C. Liew, Demand forecasting using fuzzy neural computation, with special emphasis on weekend and public holiday forecasting, IEEE Trans. Power Syst. 10 (4) (1995) 1897–1903, https://doi.org/10.1109/59.476055. Nov.
[29] M. López, C. Sans, S. Valero, C. Senabre, Classification of special days in short-term load forecasting: the Spanish case study, Energies 12 (7) (2019) 1253, https://doi.org/10.3390/en12071253. Apr.
[30] J.R. Cancelo, A. Espasa, R. Grafe, Forecasting the electricity load from one day to one week ahead for the Spanish system operator, Int. J. Forecast. 24 (4) (2008) 588–602, https://doi.org/10.1016/j.ijforecast.2008.07.005.
[31] M.S. Kim, Modeling special-day effects for forecasting intraday electricity demand, Eur. J. Oper. Res. 230 (1) (2013) 170–180, https://doi.org/10.1016/j.ejor.2013.03.039. Oct.
[32] K. Zhu, J. Geng, K. Wang, A hybrid prediction model based on pattern sequence-based matching method and extreme gradient boosting for holiday load forecasting, Electr. Power Syst. Res. 190 (2021), 106841, https://doi.org/10.1016/j.epsr.2020.106841. Jan.
[33] "ESIOS REE - Information system for the electric system operator," https://www.esios.ree.es/es, Oct. 28, 2019. https://www.esios.ree.es/es (accessed Oct. 28, 2019).
[34] "AEMET OpenData," http://www.aemet.es/es/datos_abiertos, Sep. 24, 2018. http://www.aemet.es/es/datos_abiertos (accessed Sep. 24, 2018).