IET Generation, Transmission & Distribution

Research Article

Adjustment of discrete load changes in feeder databases for improving medium-term demand forecasting

ISSN 1751-8687
Received on 1st February 2017
Revised 25th May 2017
Accepted on 7th July 2017
E-First on 21st August 2017
doi: 10.1049/iet-gtd.2017.0129
www.ietdl.org

Gonzalo Paredes[1] ✉, Luis Vargas[1]
[1]*Department of Electrical Engineering, Universidad de Chile, Av. Tupper 2007, Santiago, Chile*
✉ *E-mail: goparede@ing.uchile.cl*

Abstract: A discrete load change (DLC) event may be described as an abrupt change in feeder demand. These events are due to network reconfigurations, or the connection/disconnection of large consumers to the grid. This phenomenon affects the performance of load forecasting methods and, in general, it may worsen any planning or operational application that uses feeder demand records as input. This study proposes four load adjustment (LA) methods to correct this type of distortion from distribution system demand database. The methods are tested by using real demand values, encompassing six years of hourly data registered in 169 feeders, of a distribution company. To test the effectiveness of the LA methods, in medium-term load forecasting, a comparative study using different forecasting techniques is performed. Results show that demand forecasting, with *DLC* adjustment, improve their average performance over 33% compared to the case were this phenomena is not considered.

1 Introduction

Today more than ever the incorporation of new technologies to the distribution grids is transforming the paradigms of how the operation and planning of the distribution systems is done [1]. Conventional networks are evolving into more sophisticated grids, which require information flow between generation and consumption, a higher participation from the end users, and a growing need for a more flexible operation of the system [2–5].

Modern smart grids require high accuracy in the knowledge of consumer behaviour [6], where demand forecasting has been identified as a key process with links to many areas of system planning and operation [7–10]. This process requires the availability of a reliable and timely demand database, which is a cornerstone in distribution companies [11–12].

The reconfiguration of feeders in distribution networks, i.e. the load transfer between primary feeders, is a common operation to relieve overloading and reduce system losses. This switching operation can achieve load balance among distribution feeders and is registered as a discrete demand change in the primary substations data [13]. Another source of discrete changes is the connection (or disconnection) of large consumers. In both cases, the effect on the load data set is a distinctive abrupt change in the demand level that remains in time. In this paper, this phenomenon is called discrete load change (*DLC*).

DLC events produce distortion in the readings of monitoring equipment at a feeder level [14]. In this reference it is shown that data shifts, due to feeders switching, contaminates registers and reduces forecasting accuracy in the medium and long term. In addition, feeders' reconfiguration has been recognised as an important element for estimating future demand values in the short-term (24 h ahead) [15]. In that reference the authors conclude that switching operation in primary networks has a substantial impact on load profiles, hence, producing errors in future demand forecasting.

In [16] the effect of reconfiguration events over an artificial neural network (ANN) model is evaluated. In that work the authors show that reconfiguration events degrade the one-day ahead forecasting. This degradation is reduced when new samples (after the *DLC* event) are used to retrain the ANN.

In [17] the *DLC* effects are referred as load transfer coupling, and they are studied in the context of several months and several years, but not in daily operation such as several hours or days. In [18] in order to improve forecasting accuracy, abnormal data (load re-allocation, feeder reconfiguration, or faults) are eliminated. In [13] reconfiguration events at a feeder level are identified as abnormal changes in demand and are treated as special cases to train an ANN for load forecasting.

Most of the work on medium and long-term demand forecasting has been focused on aggregated data, where few variables are projected (e.g. maximum demand, average demand and so on) [19–22]. However, modern technologies, such as electric vehicles, renewable-based generation in distribution grids and energy storage devices, usually require hourly (or shorter) demand estimations to study their performance in future scenarios [23–26]. Therefore, long and medium-term load forecasting with hourly resolution has emerged as an important problem in modern energy systems [27–28].

In [29] the authors propose an additive semi-parametric model to perform load forecasting in short and medium-term for more than 2200 feeders of the French distribution network. They show that *DLC* events degrade the accuracy of models and, to get rid of the problem, they eliminate the time series that present *DLC* events.

According to the literature, the impact of *DLC* events depends on the timeframe of the application. On the one hand, in short-term load forecasting the impact is low as *DLC* is sporadic events (sometimes with one or two events per year) [16]. Also, in this timeframe, less data is required to achieve a good forecast (e.g. two weeks data is needed to perform one-day ahead prediction [16]), so a *DLC* can be treated as an especial case [13]. On the other hand, in medium and long-term applications, eliminating data containing a *DLC* (as proposed in [18]) could reduce significantly the amount of data available. Then, with reduced datasets the training process of methods like ANN or support vector regression (SVR) could become difficult or even infeasible [30].

As *DLC* events will become more common in future networks, their importance on medium-term load forecasting will increase accordingly, as it will have a direct impact on investment in new grid components or in the upgrade of the existing ones [31]. This work is motivated by the fact that *DLC* events on feeders' data can deteriorate significantly the performance of medium-term demand forecasting with hourly information, and it is focused on the characterisation of the phenomenon and the data pre-processing to adjust demand with *DLC* events, regardless of the forecasting techniques used afterwards.

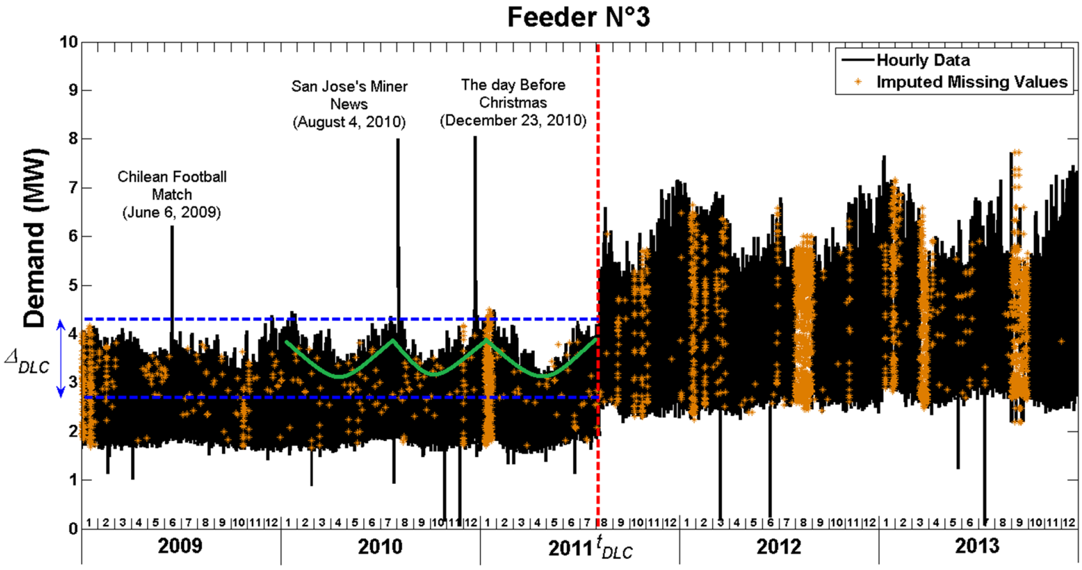

Fig. 1 *Example of a DLC in a feeder*

This paper presents efficient methods to adjust *DLC* events in load databases. Specifically, the methods allow to pre-process data registered previous to the occurrence of a *DLC* event, in order to have sufficient and reliable feeders demand databases to train models and, then, perform the forecasting. Evaluations of the proposed adjustment methods in the context of medium-term load forecasting with hourly granularity for 20 feeders of Santiago's distribution network are presented. Also, a standard forecasting *Naive Benchmark* is used to test the proposed *Load Adjustment* (*LA*) methods over 169 feeders at the distribution company of Santiago, Chile. This *Naive Benchmark* approach includes a trending forecasting of feeders' demand, which was done with the current spatial forecasting model used by the distribution company.

The remaining of this paper is organised as follows. In Section 2, DLCs definition, characterisation and its detection are presented. Section 3 describes four LA methods to correct *DLC* events. In Section 4, the effect of *DLC* events on medium-term demand forecasting is illustrated by using a *Naive Benchmark* and *Artificial Intelligence* (AI) methods, which are applied to a database of 20 feeders located in the distribution system of the city of Santiago, Chile. In Section 5, the effect of *DLC* events in one-year ahead demand forecasting is evaluated for the *Naive Benchmark* approach by using an hourly demand (HD) database of 169 feeders. Finally, Section 6 summarises the main findings of this work.

2 DLCs characterisation

In this section, a characterisation of the *DLC* phenomenon is developed, where the main features for the detection of *DLC* events are shown. Throughout the paper we use $HD^f(h, j)$ to represent the real demand at hour h in the year j, for feeder f.

2.1 Definition of DLCs

A *DLC* event could be described as an abrupt increase (or decrease) in demand $HD^f(h, j)$, which persists over time (more than one month). In order to define the main parameters to characterise a *DLC* in distribution feeders, in this subsection a real case from the distribution company of the city of Santiago is analysed.

In real databases there are missing values and outliers, which has to be pre-process before any forecasting is made. In Fig. 1, data demand $HD^f(h, j)$, in a feeder f, in the period 2010/2013 is shown. In this figure, missing data is represented with yellow dots (see middle of January 2011 and August 2012 data). Spikes in Fig. 1 represent special events such as a football match, a peak demand during the news of the San Jose's Miners rescue and a daily peak the day before Christmas.

In Fig. 1, there is a DLC at the beginning of August 2011, which is characterised by a clear increase in the demand (marked with the dotted-red line), which looks like a shift in load data of nearly 1.5 MW (shown between blue lines). It is important to keep in mind that this change does not correspond to consumption trends, as it only occurs once and remains for the rest of the period.

In this work, two parameters are used to characterise *DLC*. The first parameter measures the abrupt change in load, referred as the size of the *DLC*, and is portrayed as the Δ_{DLC} variable in Fig. 1. The other parameter is the day of occurrence, which is defined as t_{DLC} in Fig. 1.

Notice also that in Fig. 1 there is a seasonal pattern of the load. It corresponds to an increase followed by a decrease in load with a period of 6 months approximately. For illustrative purposes, it is highlighted with a green line starting in 2010.

The *DLC* phenomenon has two possible explanations. On the one hand, this could be a reconfiguration of the feeder, i.e. the transference of load among feeders. Or, it could be the connection/disconnection of a large customer (e.g. a large commercial building). In any case, it is a discrete load increase, whose effect is clearly different from the typical demand growth of existing customers.

2.2 Characterisation of DLC events

In the demand database, a *DLC* effect may be modelled as a discrete amount added to the load (see Δ_{DLC} in Fig. 1), so the demand of the remainder of year 2011 is increased. In this case, from beginning of August 2011 to the end of 2013 a fix amount, close to 1.5 MW in Fig. 1, is added to the demand.

Notice that *DLC* starts as a sudden load increase (or decrease) which has a lasting effect rather than a seasonal effect. In order to illustrate this feature, Fig. 2 shows a demand panel with four feeders on a time span of 6 years (which was extracted from the same database used in Fig. 1). In Fig. 2, *DLC* events are highlighted with a red-dotted line. These *DLC* events produce load changes that last from a few months to years, and they are not repeated on a yearly basis, i.e. they do not correspond to seasonal changes. These events were analysed and confirmed by the distribution company's personnel working on planning. Furthermore, the database containing the demand of 169 feeders was carefully examined by the company's expert, who identified *DLC* events from 2008 to 2013.

2.3 DLC detection in feeders databases

A distinctive characteristic of *DLC* is that it produces permanent changes in a demand feeder. Therefore, for its proper detection it is necessary to confirm that the resulting demand change persists over

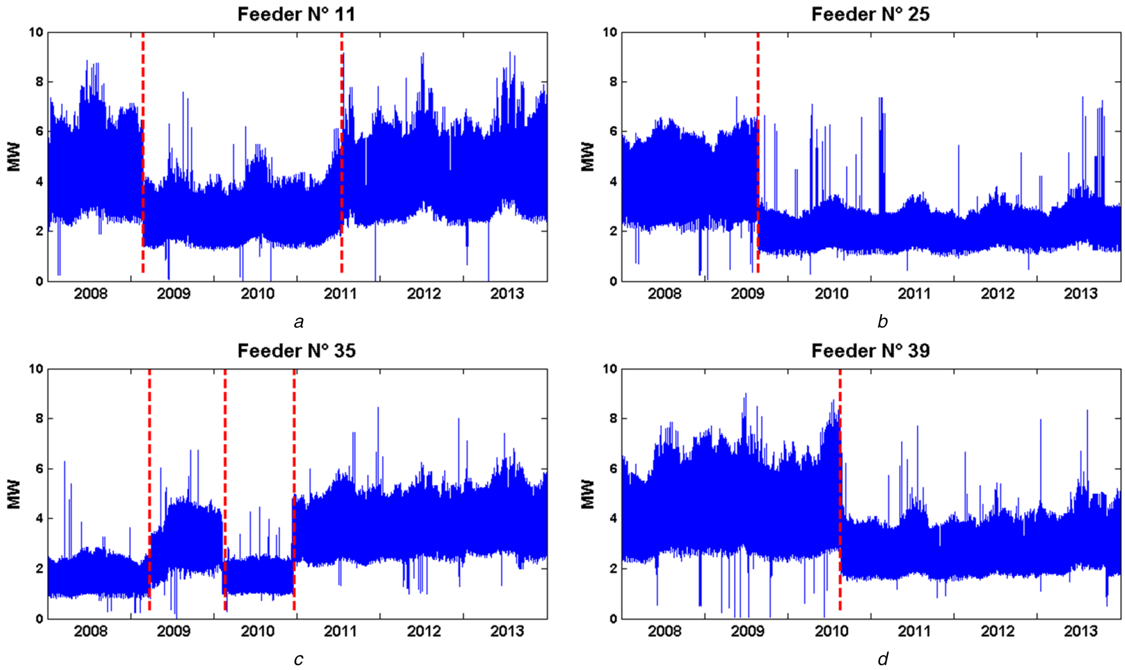

Fig. 2 *Example of a DLC in feeders*

time. In addition, a *DLC* event could be confused with daily variation, weekly patterns or even with a seasonal behaviour. To illustrate the above mentioned problems, Fig. 1 shows one example in a period of 5 years in a real feeder.

In Fig. 1, a distribution planning engineer could identify a *DLC* event in 2011 (shown with the dashed red line). In this example, at the beginning of year 2011 there is a continuous decrease in the load, which is followed by a period of constant increase (green line highlighted in Fig. 1). This pattern could be preliminary attributed to a *DLC* in 2011, but as it is also found in years 2010 and 2009, so it actually corresponds to a seasonal load pattern.

As this work is focused on the adjustment of abrupt changes in load database and not on detecting this phenomenon, an expert identification of the DLC events is considered. In other words, the expert identification provides what we call *Ground Truth Identification* (*GTI*), which later on is required as an input for the adjustment methods. Note that this approach allows eliminating the error that could be added by an automatic DLC detection method from the final results of this work.

The identification of abrupt changes in data has been studied previously [32]. The application and development of algorithms to detect these abrupt changes can be seen in a wide branch of areas, namely time series image processing [33], climate change applications [34], fault detection [35], IP network anomalies detection [36], land cover changes [37], and medical treatments [38]. In addition, this problem has been addressed in time domain and also in frequency domain [39], where the methodologies and algorithms must be tailored to the particularities of each specific application.

A preliminary *DLC* detection, made by human inspection or by using computers, should assesses that the size of the *DLC* is different for the maximum and minimum values of the hourly feeder's demand, which depends on the consumption patterns of end users. Fig. 3 shows aggregated daily values for the same load data of Fig. 1, where the black line is the daily maximum, the green line is the daily average and the purple line is the daily minimum. In this figure, *DLC* event is also highlighted in dotted-red line. The horizontal dotted and dash lines correspond to the average demand considering all data after and before the DLC event (highlighted with the vertical dotted-red line).

Notice that the difference between the average daily statistics before and after the *DLC* event is different for each curve. This is clear from Fig. 3, where Δ_{DLC} considering the daily minimum is $\Delta_{\text{DLC}}(\min) = 2.61–1.75 = 0.86$ MW, whereas the difference for the daily average is $\Delta_{\text{DLC}}(\text{avg}) = 4.29–2.70 = 1.59$ MW and for the daily maximum it is $\Delta_{\text{DLC}}(\max) = 5.69–3.51 = 2.18$ MW.

According to the field experience, not always the biggest Δ_{DLC} comes from the maximum demand statistic, and it is necessary to consider the three statistics to detect a *DLC* event in the general case. Regarding the correction of the *DLC*, different sizes of *DLC* must be tested for different demand levels in the time series.

3 LA methodology for the correction of DLC

The purpose of the *LA* methodology is to eliminate the effect of the *DLC* on the load database. In Fig. 4, the HD [HD is the entire time series $\text{HD}^f(h, j)$] and all the variables used for the proposed *LA* are illustrated. In this figure, a typical discrete load increase of 1.5 MW (Δ_{DLC}) occurs at the beginning of August 2011, in the 947th day of the time series ($t_{\text{DLC}} = 947$).

The variables used for *LA* are

- U_b is the average of daily maximum ($D_{\max}$) for the period before the *DLC* occurrence in t_{DLC} (pink line).
- M_b is the average of daily average demand (D_{avg}) for the period before the *DLC* occurrence in t_{DLC} (green line).
- L_b is the average of daily minimum ($D_{\min}$) for the period before the *DLC* occurrence in t_{DLC} (orange dash line).
- U_a, M_a, L_a are the corresponding values for the period after the *DLC* event in t_{DLC}.

The proposed *LA* method assumes that if the *DLC* occurs in a given year, its effect may be captured by adding (or subtracting) the constant change Δ_{DLC} to the data before the *DLC* occurrence. Thus, the proposed strategy for *LA* consists of adjusting the registers before the *DLC* (from January 2009 to the beginning of August 2011 in Fig. 4), whereas the remaining values (after the *DLC*) are kept unaltered (from August 2011 to December 2013 in Fig. 5).

In this work, four *LA* strategies are tested and evaluated, which differ mainly in the way they compute the size Δ_{DLC}:

- *LA-A*. Here two Δ_{DLC} are tested. In the first place, it is computed as the difference $\Delta_M = M_a - M_b$ which is added to all demand records with values above the mean M_b. In the second place, Δ_{DLC} is calculated as $\Delta_L = L_a - L_b$ is subtracted to all records below or equal to M_b.

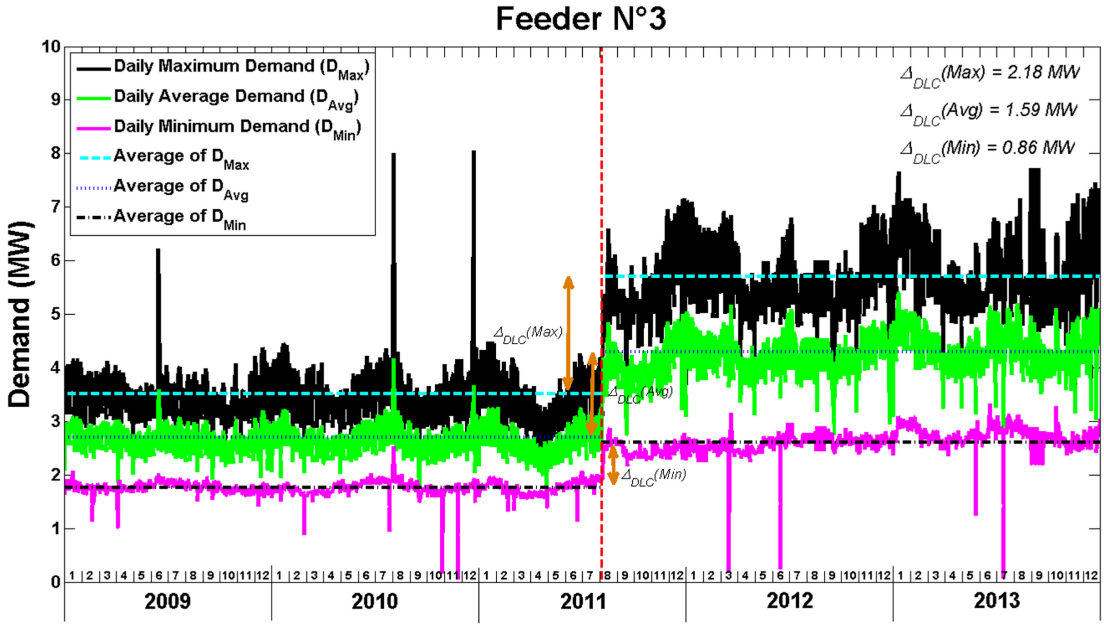

Fig. 3 *Daily statistics of feeders electric demand*

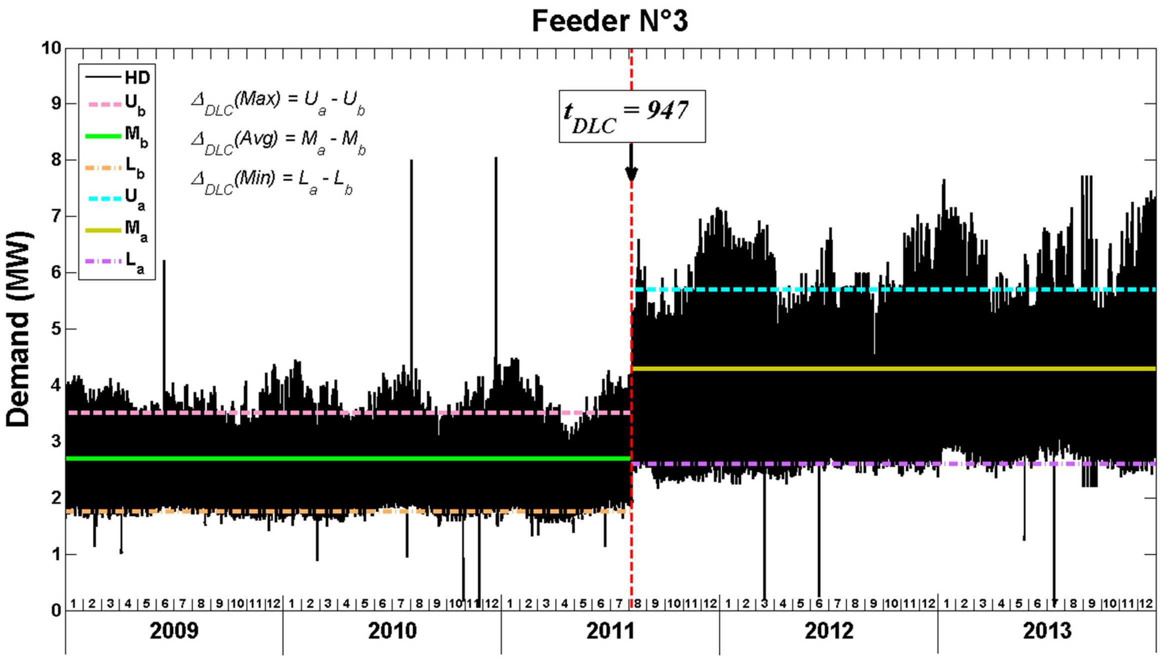

Fig. 4 *Variables used in the proposed LA methods*

- *LA-B*. Here also two *DLC* sizes are used. The first size Δ_{DLC} is estimated as $\Delta_U = U_a - U_b$, which is added to the records above M_b; whereas the second uses Δ_L, and it is subtracted to all records below or equal to M_b.
- *LA-C*. Here $\Delta_{\mathrm{DLC}} = \Delta_M$ is used, which is added to all records.
- *LA-D*. Here seven *DLC* sizes are used, one for each day of the week. This is done by subtracting a representative HD matrix before the *DLC*, named $\boldsymbol{R}_b(h, t_d)$, from a representative HD matrix after the *DLC*, named $\boldsymbol{R}_a(h, t_d)$. These matrices have one row for each hour h of the day (24 rows) and one column for each type of day t_d (seven types of days, from Monday to Sundays). The elements of matrices $\boldsymbol{R}_b(h, t_d)$ and $\boldsymbol{R}_a(h, t_d)$ are calculated as the average of all days before ($r^b_{h,td}$) and after ($r^a_{h,td}$) the *DLC*.

Then, the difference $\Delta_R(h, t_d) = \boldsymbol{R}_a(h, t_d) - \boldsymbol{R}_b(h, t_d)$, is added to adjust the load before the *DLC* for each type of day.

Notice that methods *LA*-A to *LA*-C perform LA by increasing or reducing the demand before a *DLC* event. The fourth *LA*-D method follows the same strategy but the adjustment incorporates the changes in the daily patterns.

4 Effect of DLCs on medium-term forecasting

In order to illustrate the effect of *DLC* events on load forecasting, in this section, two exercises are presented. In the first exercise, an example of the *Naive Benchmark* approach, on one feeder for 3 years, is presented, whereas in the second the effect of using *DLC* adjustment on a small data set of 20 feeders, by using two popular *AI* forecasting techniques, is shown.

4.1 Effects of DLC event on MTLF: Naive Benchmark example

A comparison of the real load, using the same feeder of Fig. 1, with two load forecasting exercises for the period 2012–2013, is presented in Fig. 5.

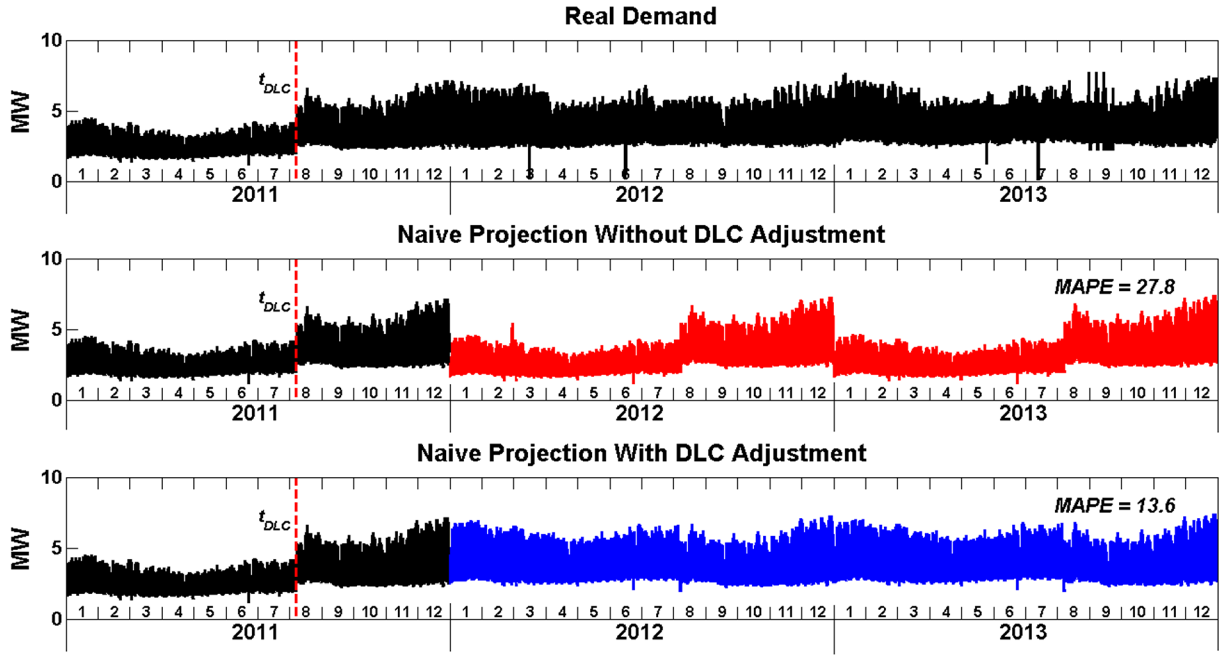

Fig. 5 *Historic values and load forecasting*

Table 1 Forecasting errors with and without LA

Forecast	MAPE, %		MAE, MW		RC, %
	Without LA	With LA	Without LA	With LA	
2012–2013	27.8	13.6	1.56	0.76	104

Table 2 MAPE indicator with and without DLC adjustment

Method	Without LA	With LA	RC
NAR, %	28.7	11.6	147.4
e-SVR, %	26.8	10.8	148.1

In Fig. 5, the black time series show the real behaviour of the demand (historic values of $\mathrm{HD}^f(h, j)$) for years 2011–2013.

The *Naive Benchmark* load forecasting used in this exercise is obtained from the product of the electricity load of the previous year and the growth rate ($\mathrm{GR}_{j,f}$) estimated by the distribution company, for each feeder f and each year j, based on where the feeder is located [29]. This approach was defined by Hyndman and Athanasopoulos [40] as a naive approach with drift.

The distribution company uses a spatial regression model to estimate the growth of different areas of the city based on historical demand and economic variables. The projected growth rate estimated in 2010 for the feeder used in the example was 1.6% in 2012 and 2% in 2013. The red time series in Fig. 5 correspond to a straightforward forecasting by using those constant annual growth rates applied to the year 2011. Note that from January to August (months 1–8) for years 2012 and 2013 there is a clear difference between the real demand (black lines) and the projected values (red lines). The reason explaining this behaviour is the blind repetition of the load pattern of year 2011, which includes the *DLC* in the month of August 2011.

The adjustment consists of adding Δ_{DLC} to all records previous to the day of the event t_{DLC}. Let us call $\overline{\mathrm{HD}}^f(h, j)$ the adjusted demand after applying the *LA-D* method, in hour h, year j and feeder f. Then, the *Naive Benchmark* forecast $\widetilde{\mathrm{HD}}^f(h, j+1)$, for year $j+1$, hour h and feeder f is calculated as follows:

$$\widetilde{\mathrm{HD}}^f(h, j+1) = \overline{\mathrm{HD}}^f(h, j)(1 + \mathrm{GR}_{j,f})$$

This procedure has been applied to the 2011 database by using the same growth rates (1.6% in 2012 and 2% in 2013) in order to obtain the forecasts for 2012 and 2013. Results with the *DLC* adjustment LA-D are shown in Fig. 5 with the blue lines. It is clear from that figure that the best forecast is given by the blue line (is the closest to the real data).

In order to measure the performance of the forecasting, the mean average error (MAE), mean average percentage error (MAPE) and a relative change (RC) indicator are used. MAPE and MAE are common measures of errors, whereas RC in this work is defined as

$$\mathrm{RC} = \frac{100 * (I_{WA} - I_A)}{I_A}$$

where I_{WA} is the forecasting indicator (MAE or MAPE) without adjustment of the database and I_A is the corresponding indicator with the adjustment.

Results for MAPE and MAE of the real and forecasted demand (same data as shown in Fig. 5) for the period 2012–2013 are shown in Table 1.

Results in Table 1 show that the adjustment of *DLC* reduces dramatically the MAE and MAPE errors, where average improvements (RC) are over 104% when compared to the case without *DLC* correction.

4.2 Measuring the effects of DLC events using AI forecasting techniques

In order to show the effect of using the *DLC* adjustment technique with more sophisticated load forecasting methods, the proposed *DLC* LA method LA-D is used to feed two selected forecasting techniques: non-linear autoregressive neural network (NAR) [41] and SVRs (ε-SVR) [42]. By following the same idea of the previous subsection, the exercise consists of a comparison of the forecasting results in two cases. In the first case, data is adjusted according to the *DLC* proposed techniques, whereas in the second case NAR and SVRs are applied directly to the untreated data. In these tests, data from 2008 to 2011 was used for training and setting parameters, whereas data from 2012 to 2013 was used for validation purposes. Results for MAPE indicators are shown in Table 2.

Table 3 LA methods performance

Year	Without LA, %	Average MAPE of next year with LA, %				Average RC, %			
		LA-A	LA-B	LA-C	LA-D	LA-A	LA-B	LA-C	LA-D
2009	31.3	22.8	26.6	23.4	17.5	37	18	34	79
2010	30.5	18.0	23.4	18.3	10.9	69	30	66	179
2011	29.7	18.9	24.0	21.5	14.7	58	24	38	102
2012	26.2	18.1	22.7	19.4	13.6	45	16	35	93
2013	27.1	20.5	23.5	16.6	15.8	33	15	63	71

Results in Table 2 show a consistent improvement when the proposed LA procedure is used with more sophisticated forecasting methods, such as NAR and ε-SVR. For both univariate methods, the input data was processed in order to consider intraday, intraweek and intrayear seasonal cycles [43]. By considering that data are on an hourly basis, the length of intraday (s_1), intraweek (s_2) and intrayear cycles (s_3) are: $s_1 = 24$, $s_2 = 24 \times 7$, $s_3 = 24 \times 7 \times 52$. With these definitions and d_t as a demand register in time t, the input variables used for all forecasting evaluations are the following:

$$d_1, d_2, d_3,\ d_{s_1}, d_{s_1+1}, d_{s_1+2},$$
$$d_{s_2}, d_{s_2+1}, d_{s_2+2},\ d_{2s_2}, d_{2s_2+1}, d_{2s_2+2},$$
$$d_{3s_2}, d_{3s_2+1}, d_{3s_2+2}, d_{s_3}, d_{s_3+1},\ d_{s_3+2}\,.$$

In all evaluations, the best parameters for each model are found based on a greedy search approach.

For the NAR method a search of the best number of delays (from 1 to 10) and the best number on neurons in the hidden layer (from 5 to 15) was implemented. A linear transfer function was used in the input layer and a log-sigmoid transfer function was used in the output layer, as load forecasting must be always positive. Regarding data, the training set (2008–2011) was divided into 70% to train, 15% for testing and 15% for validation. By using this validation results, the best 10 models configurations out of 110 models tested, for each feeder, were selected to produce the final forecasting evaluation presented in Table 2.

In the case of ε-SVR the same previous approach was followed. Linear and non-linear kernels were tested and a search for the best ten models was performed. For the non-linear tests, the Gaussian radial basis function (RBF) was used. As previous work shows that ε-SVR is less sensitive to the ε parameter [44], in this application ε was fixed at 0.1. In these cases, the search for the best hyperparameters considered σ^2 and C with values from 2^{-9} to 2^9. In the linear tests, the search for the best hyperparameters considered ε between 0.1 and 1, and C with values from 2^{-9} to 2^9. Again the best ten models were selected to calculate an average perform of this technique with and without *LA* as shown in Table 2.

In the literature, it is reported that the overall MAPE errors with DLC events is above 20% [29], which are larger than those shown in Table 2. In summary, significant improvement for medium-term load forecasting accuracy is achieved when proper detection and correction of *DLC* events is performed on real field data.

5 Evaluation of LA methods

In this section, the performance of the LA methods is evaluated. The methods are tested on an HD database of 169 feeders located in the city of Santiago with registers from 2008 to 2013. These data represent real scenarios in different conditions, e.g. feeders having significant random fluctuations, feeders with and without trending, feeders with and without seasonal variations and so on.

In order to compare the performance of the proposed LA methods, it is necessary to build the true identification of DLC occurrences. This is done by an extensive work, where the database containing the demand at each feeder was carefully examined by an expert, who identified *DLC* events from 2008 to 2013. The result is an expert-built indicator $I_{\mathrm{DLC}}^{\exp}(t)$ that provides the *GTI*, which is required as an input for the *LA* methods.

As the focus of these tests is on the performance of *LA* methods, rather than in the forecasting technique, only the *Naive Benchmark* approach is used. Thus, forecasting is calculated simply by multiplying the demand of the previous year with the growth rate ($\mathrm{GR}_{j,f}$) projected by the distribution company with the current spatial forecasting model, for each feeder f and each year j (same procedure shown in Section 4.1).

Furthermore, to simplify the comparison, a special testing database is built, which is constructed with feeders that have a *DLC* in the first year and, simultaneously, they do not have a *DLC* in the next year. It is found that 71 demand time series fulfil these conditions. The next step is to apply the *DLC*-*LA* methods to adjust the demand at each year j.

In order to measure the performance of *LA* methods, the error between the projected demand in year $j+1$, $\widetilde{\mathrm{HD}}^f(h, j+1)$, and the actual register of that year, $\mathrm{HD}^f(h, j+1)$, is calculated. Table 3 shows the performance for the four *LA* methods, which is measured with the MAPE and RC indicators. In addition, for reference purposes, in Table 3 the second column indicates the corresponding MAPE when no *LA* is applied.

From Table 3, columns 3–6 show that all *LA* methods achieve a reduction in MAPE as compared with the reference case (second column). Last four columns show the MAPE improvement as a percentage with respect to the reference case (column 2). It can be seen that for the *LA-D* method, improvements over 71% on average are achieved in all years. In fact, from the total of 71 demand series evaluated from 2009 to 2013, it is found that *LA-D* method is able to reduce the forecasting errors in 61 cases.

6 Conclusion

In this work, the DLC phenomenon, resulting from network reconfigurations or the incorporation of large consumers in distribution feeders, is characterised, identified and adjusted.

Tests are carried out by using field data from a distribution company, covering 169 feeders and a time span of 6 years.

Results show a notorious improvement in mid-term load forecasting when the adjustments methods (*LA*) are used. The best method, which uses seven *DLC* sizes, one for each day of the week, is able to improve load forecasting over 71% on average for all years, when compared to the case where no *DLC* is adjusted.

Future work is focused on applications of these methods to medium-term load forecasting with hourly granularity and incorporating weather variables.

7 Acknowledgments

This research was funded by the Complex Engineering Systems Institute, ISCI (CONICYT: FB0816). Special thanks to Mauricio Riveros, from the planning area of the distribution company in the city of Santiago, Chile, for the authors support in developing this work.

8 References

[1] Energetics Incorporated: 'The future of the grid: evolving to meet America's needs – final report', Department of Energy United States of America, Tech. Rep. GS-10F-0103J, 2014

[2] National Institute of Standards and Technology: 'Nist framework and roadmap for smart grid interoperability standards, release 1.0', U.S. Department of Commerce, Tech. Rep. 1108, 2010

[3] Gharavi, H., Ghafurian, R.: 'Smart grid: the electric energy system of the future', *Proc. IEEE*, 2011, **99**, p. 6

[4] Sánchez-Jiménez, M.: 'European smartgrids technology platform. Vision and strategy for Europe's electricity networks of the future', European Commission, Tech. Rep. KI-NA-22040-EN-C, 2006
[5] Farhangi, H.: 'The path of the smart grid', *IEEE Power Energy Mag.*, 2010, **8**, (1), pp. 18–28
[6] Ferreira, H., Lampe, L., Newbury, J.: '*Power line communications. Theory and applications for narrowband and broadband communications over power lines*' (John Wiley & Sons, 2010)
[7] Chan, S.-C., Tsui, K.M., Wu, H.*, et al.*: 'Load/price forecasting and managing demand response for smart grids: methodologies and challenges', *IEEE Signal Process. Mag.*, 2012, **29**, (5), pp. 68–85
[8] Vardakas, J.S., Zorba, N., Verikoukis, C.V.: 'A survey on demand response programs in smart grids: Pricing methods and optimization algorithms', *IEEE Commun. Surv. Tutor.*, 2015, **17**, (1), pp. 152–178
[9] Rahimi, F., Ipakchi, A.: 'Demand response as a market resource under the smart grid paradigm', *IEEE Trans. Smart Grid*, 2010, **1**, (1), pp. 82–88
[10] Hong, T.: 'Energy forecasting: past, present, and future', *Foresight: Int. J. Appl. Forecast.*, **2014**, (32), pp. 43–48
[11] Wang, D.T.C., Ochoa, L.F., Harrison, G.P.: 'Modified GA and data envelopment analysis for multistage distribution network expansion planning under uncertainty', *IEEE Trans. Power Syst.*, 2011, **26**, (2), pp. 897–904
[12] Zou, K., Agalgaonkar, A.P., Muttaqi, K.M.*, et al.*: 'Distribution system planning with incorporating DG reactive capability and system uncertainties', *IEEE Trans. Sustain. Energy*, 2012, **3**, (1), pp. 112–123
[13] Sun, X., Luh, P.B., Cheung, K.W.*, et al.*: 'An efficient approach to short-term load forecasting at the distribution level', *IEEE Trans. Power Syst.*, 2016, **31**, (4), pp. 2526–2537
[14] Willis, H., Powell, R., Wall, D.: 'Load transfer coupling regression curve fitting for distribution load forecasting', *IEEE Trans. Power Appar. Syst.*, May 1984, **PAS-103**, (5), pp. 1070–1076
[15] Yasuoka, J., Brittes, J.L.P., Schmidt, H.P.*, et al.*: 'Artificial neural network-based distribution substation and feeder load forecast', 16th Int. Conf. and Exhibition on Electricity Distribution, 2001, Part 1: Contributions CIRED (IEE Conf. Publ No. 482), Amsterdam, vol. **5**, p. 5
[16] Fidalgo, J.N., Lopes, J.A.P.: 'Load forecasting performance enhancement when facing anomalous events', *IEEE Trans. Power Syst.*, 2005, **20**, (1), pp. 408–415
[17] Bai, X., Gang, M., Ping, L.: 'A method of spatial load forecasting based on feeder', Third Int. Conf. Electric Utility Deregulation and Restructuring and Power Technologies, 2008, DRPT 2008, 2008, Nanjuing, pp. 1548–1553
[18] He, D., Habetler, T., Mousavi, M.*, et al.*: 'A ZIP model-based feeder load modeling and forecasting method', 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, 2013, pp. 1–5
[19] Weron, R.: '*Modeling and forecasting electricity loads and prices: a statistical approach*' (John Wiley & Sons, 2007)
[20] Kandil, M.S., El-Debeiky, S.M., Hasanien, N.E.: 'Long-term load forecasting for fast developing utility using a knowledge-based expert system', *IEEE Trans. Power Syst.*, 2002, **17**, (2), pp. 491–496
[21] Gonzalez-Romera, E., Jaramillo-Moran, M.A., Carmona-Fernandez, D.: 'Monthly electric energy demand forecasting based on trend extraction', *IEEE Trans. Power Syst.*, November 2006, **21**, (4), pp. 1946–1953
[22] Hyndman, R.J., Fan, S.: 'Density forecasting for long-term peak electricity demand', *IEEE Trans. Power Syst.*, May 2010, **25**, (2), pp. 1142–1153
[23] Al-Hamadi, H., Soliman, S.: 'Long-term/mid-term electric load forecasting based on short-term correlation and annual growth', *Electr. Power Syst. Res.*, 2005, **74**, (3), pp. 353–361
[24] Asber, D., Lefebvre, S., Saad, M.*, et al.*: 'Modeling of distribution loads for short and medium-term load forecasting', Power Engineering Society General Meeting, 2007, June 2007, pp. 1–5
[25] Filik, U., Gerek, O., Kurban, M.: 'Hourly forecasting of long term electric energy demand using a novel modeling approach', Fourth Int. Conf. Innovative Computing, Information and Control (ICICIC), 2009, December 2009, pp. 115–118
[26] Hong, T., Wilson, J., Xie, J.: 'Long term probabilistic load forecasting and normalization with hourly information', *IEEE Trans. Smart Grid*, 2014, **5**, (1), pp. 456–462
[27] Filik, Ü.B., Gerek, Ö.N., Kurban, M.: 'A novel modeling approach for hourly forecasting of long-term electric energy demand', *Energy Convers. Manage.*, 2011, **52**, (1), pp. 199–211
[28] Xia, C., Wang, J., McMenemy, K.: 'Short, medium and long term load forecasting model and virtual load forecaster based on radial basis function neural networks', *Int. J. Electr. Power Energy Syst.*, 2010, **32**, (7), pp. 743–750
[29] Goude, Y., Nedellec, R., Kong, N.: 'Local short and middle term electricity load forecasting with semi-parametric additive models', *IEEE Trans. Smart Grid*, 2014, **5**, (1), pp. 440–446
[30] Zhang, G.P.: 'Avoiding pitfalls in neural network research', *IEEE Trans. Syst. Man Cybern. C Appl. Rev.*, 2007, **37**, (1), pp. 3–16
[31] Khuntia, S.R., Tuinema, B.W., Rueda, J.L.*, et al.*: 'Time-horizons in the planning and operation of transmission networks: an overview', *IET. Gener. Transm. Distrib.*, 2016, **10**, (4), pp. 841–848
[32] Zurbenko, I., Porter, P.S., Gui, R.*, et al.*: 'Detecting discontinuities in time series of upper-air data: development and demonstration of an adaptive filter technique', *J. Clim.*, 1996, **9**, (12), pp. 3548–3560
[33] Verbesselt, J., Hyndman, R., Zeileis, A.*, et al.*: 'Phenological change detection while accounting for abrupt and gradual trends in satellite image time series', *Remote Sens. Environ.*, 2010, **114**, (12), pp. 2970–2980
[34] Faghmous, J.H., Kumar, V.: 'Spatio-temporal data mining for climate data: advances, challenges, and opportunities', 'In *Data mining and knowledge discovery for big data',* (Springer Berlin Heidelberg, 2014), pp. 83–116
[35] Gustafsson, F., Gustafsson, F.: '*Adaptive filtering and change detection*', (Wiley, New York, 2000, vol. **1**)
[36] Thottan, M., Ji, C.: 'Anomaly detection in IP networks', *IEEE Trans. Signal Process.*, 2003, **51**, (8), pp. 2191–2204
[37] Verbesselt, J., Hyndman, R., Newnham, G.*, et al.*: 'Detecting trend and seasonal changes in satellite image time series', *Remote Sens. Environ.*, 2010, **114**, (1), pp. 106–115
[38] Wagner, A.K., Soumerai, S. B., Zhang, F.*, et al.*: 'Segmented regression analysis of interrupted time series studies in medication use research', *J. Clin. Pharm. Ther.*, 2002, **27**, (4), pp. 299–309
[39] Antoniadis, A., Gijbels, I.: 'Detecting abrupt changes by wavelet methods', *J. Nonparametric Stat.*, 2002, **14**, (1-2), pp. 7–29
[40] Hyndman, R.J., Athanasopoulos, G.: '*Forecasting: principles and practice*' (OTexts, 2014)
[41] Hudson, M., Hagan, M., Demuth, H.: 'Neural network toolboxTM reference', The MathWorks, 2015
[42] Chang, C., Lin, C.: 'LIBSVM: a library for support vector machines', ACM Transactions on Intelligent Systems and Technology, 2011
[43] Taylor, J.W.: 'Triple seasonal methods for short-term electricity demand forecasting', *Eur. J. Oper. Res.*, 2010, **204**, (1), pp. 139–152
[44] Smets, K., Verdonk, B., Jordaan, E. M.: 'Evaluation of performance measures for SVR hyperparameter selection', IEEE Int. Joint Conf. Neural Networks 2007, IJCNN 2007, 2007, pp. 637–642