Short-term load forecasting in air-conditioned non-residential Buildings

Yoseba K. Penya, Cruz E. Borges, Denis Agote and Iván Fernández
eNergy Lab
DeustoTech - Deusto Technology Foundation
University of Deusto, Bilbao, Basque Country
{yoseba.penya,cruz.borges,denis.agote,ivan.fernandez}@.deusto.es

***Abstract*—Short-term load forecasting (STLF) has become an essential tool in the electricity sector. It has been classically object of vast research since energy load prediction is known to be non-linear. In a previous work, we focused on non-residential building STLF, an special case of STLF where weather has negligible influence on the load. Now we tackle more modern buildings in which the temperature does alter its energy consumption. This is, we address here fully-HVAC (Heating, Ventilating, and Air Conditioning) ones. Still, in this problem domain, the forecasting method selected must be simple, without tedious trial-and-error configuring or parametrising procedures, work with scarce (or any) training data and be able to predict an evolving demand curve. Following our preceding research, we have avoided the inherent non-linearity by using the work day schedule as day-type classifier. We have evaluated the most popular STLF systems in the literature, namely ARIMA (autoregressive integrated moving average) time series and Neural networks (NN), together with an Autoregressive Model (AR) time series and a Bayesian network (BN), concluding that the autoregressive time series outperforms its counterparts and suffices to fulfil the addressed requirements, even in a 6 day-ahead horizon.**

I. INTRODUCTION

It is common-knowledge that progress affects all aspects of our lives, sometimes for good and sometimes not. This fact has become a painful truth when it comes to the way we consume energy: just think of the new gadgets we have acquired only in the last 5 years. Moreover, the overall change in the regulation has drawn a new scenario in which classical roles have been recast and new have appeared. This process, widely known as *liberalisation* of the energy markets, started in the UK in the 90s and consists of the separation of the electricity generation and retail from the natural monopoly functions of transmission and distribution. In this context, last-mile electricity customers now have the possibility of choosing their retailer (who, in turn, will buy that energy to different producers everyday). Yet the word *possibility* is always twofold, since it comprises both potential profits and loses. For every client, not selecting the most convenient retailer or, directly going for the worst, will definitively make a difference, in this case in the energy bill. Moreover, finding a suitable retailer is not an easy task due to many reasons and this aspect has drawn quite static electricity markets, with few clients switching from their incumbent supplier. First of all, small clients will only get attractive offers if they do form a notable group. On the contrary, huge consumers face the risk of not finding a retailer able to satisfy their needs. At this point, short-term load forecasting tools provide useful information that may help in such challenge: knowing about their consumption (amount, habits, etc.) enables controlling it.

Further, not only consumers profit from these tools; all participants in the electric system do. For instance, since the balance between generation and consumption must be watched out constantly in the power grid, *Transmission System Operators* (TSOs) work with global demand prognoses. Any deviation from them implies an added cost because the consumption is not being managed efficiently. Prediction of the demand from the clients' side may help reduce these deviations, minimising over-costs that the TSO must face. Regarding *retailers*, they always work with client portfolios and being able to foresee their consumption in the short term will enable them to buy more accurately what they need (otherwise, they must sell sparing electricity). The agents in *power derivatives markets* must also tackle a similar situation since this kind of markets that trade energy in the future must adjust the volume of the energy acquired, specially not to buy too much without having then enough consumption on the clients to sell it to (which is one of the phenomenon that has appeared lately due to the demand reduction that the global crisis has caused). Finally, *Distribution Network Operators* (DNOs) require an estimation of the natural growth of the energy demand in their distribution grid, in order to be able to foresee changes or extensions and the subsequent investments. In this way, predicting the demand of their clients may help them achieve this goal.

In a previous work [3], we focused on a very special case of STLF, namely non-residential Buildings STLF. For non-residential buildings we understand schools, universities, public buildings and companies' facilities. They all present a similar consumption curve: stationary, seasonal and regular, coinciding with the times the building is used. Hence, there is no consumption at night (or it is negligible) and, anyway, there exists a notable difference between idle and activity times. Further, many of these buildings are not yet fully-automated. This is, their HVAC system either is manually controlled or it is switched on and off remotely but does not *adapt* to weather changes (e.g. a sudden descent of the temperature due to a storm).

Here, we focus on more advanced, air-conditioned buildings (please note that we included fully-HVAC buildings under the common "air-conditioned" banner), in which weather does show influence on the load. Moreover, there is usually not abundant hourly load historical data and the load profile is sure to vary and evolve over the time. Finally, the solution chosen for this purpose must be simple to tailor to every single case (e.g. there should not be a Neural Network expert in the school to control and periodically tailor the NN that predicts their load profile).

Against this background, we advance the state of the art in two main ways. First, we have successfully validated our methodology to skip the non-linearity of the load in this problem domain. Second, we have tested with real load data the most popular STLF methods, NNs and ARIMA time series, and Bayesian networks (though neither fulfils all requirements put forward), as well as an Autoregressive Model (AM) (which does fulfil all requirements). Our results show that the AR model outperforms the all other methods (including those that do not apply to STLF in air-conditioned non-residential buildings). The remainder of the paper is organized as follows. Section II provides a critical overview on the related work. Section III describes the real-world scenario we took the data from, details the models tested, and empirically evaluates the importance that weather variables really have in the model. Section IV details the experiments carried out and discusses the obtained results. Finally, section V concludes and outlines the avenues of future work.

II. RELATED WORK

As already mentioned, there exists a very large literature on short-term load forecasting (see [1], [2], [4] for a comprehensive survey on STLF) but, comparatively, little on the same topic applied to buildings. In both scenarios, research presents two main branches. The first one includes different types of *statistical* methods, including *univariate time* series, in which the load is modelled according to historical data (e.g. multiplicative autoregressive models [5], dynamic linear [6] or non-linear [7] models, threshold autoregressive models [8], Kalman filtering [9], and Gaussian Process prior [10]), and *causal models*, in which the load is modelled as a function of an exogenous factor(s) (e.g. weather). In this latter group we can place ARMA models (also known as Box-Jenkins [11]), ARMAX models [12], optimization techniques [13], non-parametric regression [14], structural models [15], and diverse curve-fitting procedures [16]. In spite of the large number of alternatives, however, linear regressions [17] have been the most popular election, and, most accurately, ARIMA has been the technique showing the most promising results [18].

In this way, lately the bulk of STLF research has been concentrated on the second group, using several artificial intelligence methods to deal with the non-linearity of the historical load data. In this way, the techniques addressed include fuzzy logic, expert systems, evolutionary algorithms, support-vector machines [19] and, specially, all kinds of neural networks [20]. Though being most promising, NN and SVM must deal with a number of problems. First, either require much more historical data than any of the statistical methods [2]. This data set may also pose a problem to NNs since they fail when it presents random correlations among the inputs and the output because conventional NNs will not set the coefficients for those junk inputs to zero. In this way, irrelevant variables may blur the accuracy of the prediction (for instance, [21] uses Bayesian methods to alleviate this trouble). Moreover, they both relay on a tedious trial-and-error process to tune them up properly. Finally, well-known issues that arise in load forecasting, such as *over-fitting* and *data-ageing*, remain still open.

Yet, as discussed before, STLF in buildings addresses a different problem domain, and there have been a number of interesting initiatives tackling the special features of this scenario. For instance, [22] tried to model the hourly energy use in commercial buildings with Fourier Series. They performed poorer on the weekends due to the fact that they used the whole data series for the modelling. [23] corrected this drawback by distinguishing day types but, still, they focused on hourly modelling rather on load forecasting itself. Regarding artificial intelligence methods, [24] addressed a SVM for predicting the load of a building complex, [21] proposed a NN tuned up by Automatic Relevance Determination in order to optimise the selected input. Moreover, [25] put forward an NN in which the input variables where selected by a version of the Wald's test.

In the same spirit, [20] used the temperature data in a feedback NN with a remarkable MAPE of the 1.945 % (for instance, [26] included inputs about orientation, insulation thickness and transparency ratio without improving that result). As aforementioned, NNs require much historical data (which, in our case may not be available) and, further, a complicated configuration process that yield them unable to be easily adaptable to single small scenarios(say the buildings of a school). Finally, all artificial intelligence methods squander all their efforts on modelling the non-linearity. As we have shown in a previous work, in our problem scenario this can be avoided easily by using the work schedule calender. A similar concept has been applied in [27] (one model for each type of day) and [28] (one model for each hour of the day) to STLF and in [23] to building STLF.

III. SHORT-TERM LOAD FORECASTING

We have recorded the energy consumption data from a new building of the campus that the University of Deusto owns in Donostia-San Sebastián (Basque Country). This building was built 2 years ago and is fully HVAC-equipped. The construction of two new buildings started in January 2009 (our oldest records date from February 2009) and all the electricity is being taken from the same substation we communicate with. This fact converts forecasting into an even more difficult enterprise due to the noise it introduces but simultaneously forces the tested algorithms to demonstrate their ability to successful adapt to evolving data and to sparse training data. Such feature may yield worse results in terms of topping existing STLF solutions but helps us in our goal of finding a good, simple, easy to use, and effortlessly exportable method. Fig. 1 shows the average consumption curve of each day recorded in May 2009.

As it can be seen, the weekly load presents quite a regular profile in working days, with consumption from 7am to 10am (open hours range from 8am to 9pm). On Saturdays, it shows a peak at noon and on Sundays it is almost flat. We have downloaded this data directly from the meter, placed by the Spanish law (54-1997) directly at the transformer, and using the IEC 60870-5-102 standard protocol. The meteorological data was obtained from the Basque Meteorological Agency

(EuskalMet), measured in two points: Zizurkil (20 Km to the South) and Zarautz (20 Km to the West).

A. Models

We will describe here briefly the models we have implemented for the tests.

1) Autoregressive Model: This kind of model is commonly used for modelling univariate time series. We have used one for each different day type (weekday, Saturday and Sunday) and hour:

$$s_t^{h,d} = c + \sum_{i=1}^{q} \varphi_i^{h,d} s_{t-i}^{h,d}, \qquad (1)$$

where c is a constant and $\varphi_i^{h,t}$ are the model parameters. Instead of adjusting these parameters, we have set them to $\varphi_1^{h,d} = 1/3$, $\varphi_2^{h,d} = 1/3$, and $\varphi_3^{h,d} = 1/3$ in order to keep this model as simple as possible and avoid a trial-and-error parametrisation process. Following the methodology introduced in a previous work, we use the work-day schedule to classify the type of day whose load was to be predicted. Therefore, we have adjusted the prediction for each hour applying the AR model depending on the day type and arbitrarily setting $q = 3$. Hence, please note that we have computed the 3 last values of the *same* day type (e.g. from a Tuesday, the previous Monday, Friday, Thursday) and not the three last chronological values (e.g. from a Tuesday, the previous Monday, Sunday, and Saturday).

2) ARIMA model: The Autoregressive Integrated Moving Average is a generalization of an autoregressive moving average (ARMA) model. It joints two steps, one for the data periods showing stationarity and another for the non-stational parts. It is generally referred to as an ARIMA(p,d,q) model where p, d, and q correspond respectively of the order of the autoregressive, integrated, and moving average parts of the model. We tested many combinations and obtained the best results with $p = 3$, $d = 0$, and $q = 3$. An ARIMA(p,d,q) is given by:

$$(1 - \sum_{i=1}^{p} \phi_i L^i)(1 - L)^d X^t = (1 + \sum_{i=1}^{q} \theta_i L_i)\varepsilon_t, \qquad (2)$$

where L is the lag operator, ϕ_i the parameters of the autoregressive part, θ_i the parameters of the moving average part, and ε_t the error terms.

3) Neural network: NNs are non-linear circuits whose *perceptron* (say simple information processors) structure adapts according to the external or internal information that flows through the network during the learning phase. Their output is a linear o non-linear function of the inputs and, therefore, they have been widely used for predicting non-linear data (as in STLF [1], [2], [20], [21], [25], [26]). After many tests, we obtained the best results with a NN design including the temperature-related variables, the value of the previous hour (independently of the day type), and the value of the same hour in the previous same-type day. Moreover, we needed two hidden layers of 35 and 25 perceptrons each.

4) Bayesian model (b): Bayesian Networks (BN) are probabilistic models for multivariate analysis that extend the *Bayes' theorem*. They combine an acyclic directed graph with a probability distribution functions [29]: the graphical model represents the set of probabilistic relationships among the collection of variables modelling the specific problem, whereas the probability function illustrates on each node the strength of these relationships or edges in the graph [30]. The research on BNs has mainly focused on systems with discrete variables, linear Gaussian models or combinations of both since, except for linear models, continuous variables pose a problem for Bayesian networks [31] due to the inherent difficulty of representing a continuous quantity by an estimated magnitude and a range of uncertainty [30]. We have tackled this issue by clustering the values of the load (the variable to be predicted) for each hour (using Agglomerative Hierarchical clustering [32]) and then, by calculating the average load for each of the clusters. In this way, the BN classifies the load into one of those categories and the exact amount predicted is the average load of that class (and the error, the difference between the actual load value and the average load of the assigned cluster). As with the AR model above, we have designed three different BNs, one for each type of day. We have included the weather variables (see next section), type of the day, load value in the previous hour and load value in the same hour of the previous same-type day. Furthermore, for each hour, we re-trained the BN again to include the value of the last hour, and issued the prediction of the next hour. Finally, we have performed the *structural learning* by applying the PC-Algorithm [33], the Expectation-Maximisation algorithm [34] for the *parametrical learning* and the Lauritzen and Spiegelhalter method for conclusion inference over junction trees [29] in order to achieve the *Bayesian inference* (i.e. the actual prediction). Since BN forecasting is out of the scope of this paper, we omit a more detailed description due to the lack of space.

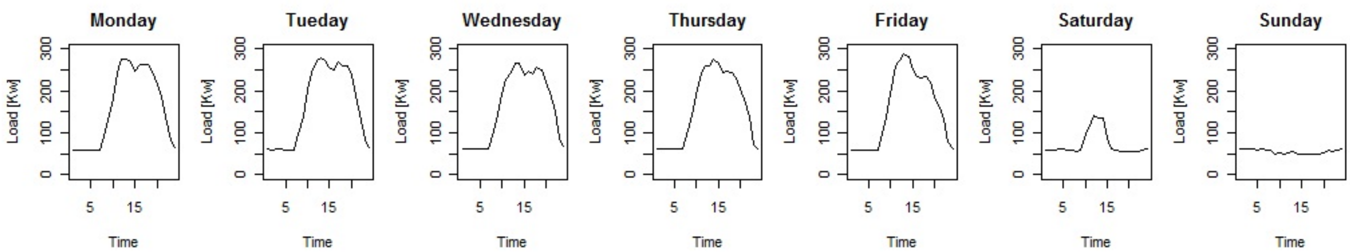

Fig. 1. Average daily load in may 2009.

TABLE I
INFORMATION GAIN ALGORITHM RESULTS OF THE WEATHER VARIABLES.

Variable	**InfoGain** result
dayOfTheWeek	0.667
dayType	0.482
THIZZA	0.216
THIZI	0.216
AirTemperatureZA	0.215
AirTemperatureZI	0.215
WCIZA	0.151
WCIZI	0.151
Rest of weather variables	0

B. Weather influence

We had the following variables to extract the correlations (please note that, we had 2 weather data sources, Zizurkil (ZI) and Zarautz (ZA)): *DayOfTheWeek, TypeOfDay, Season, WindDirectionZA, HumidityZA, PrecipitationZA, SigmaDirectionZA, SigmaSpeedZA, AirTemperatureZA, AverageSpeedZA, WindDirectionZI, HumidityZI, PrecipitationZI, SigmaDirectionZI, SigmaSpeedZI, AirTemperatureZI and AverageSpeedZI*. Moreover, we have calculated two composite variables, namely the Temperature Humidity Index (THI, also known as *discomfort index* or *effective temperature*) and the Wind Chill Index (Wind Chill Index), which are broadly used by utility companies in forecasting [4].

$$THI = T_a - (0.55 - 0.55RH)(T_a - 58), \tag{3}$$

$$WCI = (10\sqrt{v} - V + 10.5)\dot{(}33 - T_a), \tag{4}$$

where T_a is the air temperature in $^\circ F$, RH the relative humidity in percent, and V the wind speed. In order to weight the importance of the weather variables in our model, we have evaluated them with the Information Gain method [35] (widely used for this purposes [36]), which measures, for each variable, the expected reduction in entropy that the presence of that variable causes. Tab. I shows the results obtained; variables that do not appear in the table obtained a result of 0 (i.e. not importance at all). As it can be seen, only temperature and temperature-related variables have a non-negligible weight but, still, not very important if compared with the non-weather variables. Anyway, as we will present in the next section, an statistical method that does no take into account weather variables, outperforms artificial intelligence models that do.

IV. EXPERIMENTS AND DISCUSSION

We carried out the experiments on a Core 2 SU4100 CPU with 4GB RAM and a Gentoo Linux up to date. We used 2 different programs: R 2.12.1 compiled from sources with GCC 4.5.2 (Gentoo), and Weka 3.6.4 compiled from sources with GCC 4.5.2 (Gentoo). We have applied the introduced models to two different datasets representing the hourly load profile of 1 year. The first one (data set 1), showed quite a non-regular profile, mostly on Saturdays, with frequent noisy values due to the construction of the new buildings. The second one (dataset 2), was much regular and homogeneous. We trained the ARIMA with all the existing previous data, starting from the first 12 days of the same type, and forecasted the rest of the values until the end of the year. The AR didn't need training and the NN, contrary to what the literature claims [1], [2],needed surprisingly exiguous training data: only the last 3 same day-type and hour values. We believe that the reason for this phenomenon is that we skipped non-helpful training data by feeding the NN only with same day-type and hour values. Regarding the BN, we used 11 of the months in order to conduct the training (please note that BNs use this process first to build their graph, and then, to obtain the values for each node's probability tables) and the 12th month to issue the forecasting and evaluate it's performance (which is common *usus* in this field). We acknowledge that this methodology may bias the results positively (if that 12th month happens to be a regular average one) or negatively (if heterogeneous and odd). In order to achieve a proper comparison with the other models, the BN should be trained and then tested with predictions over a whole year (we will accomplish this when we have the data). Finally, we re-trained the corresponding BN after each forecast to include the last predicted value and the actual one, in order to keep it constantly up-to-date. In any case, we issued the predictions hour by hour, comparing the predicted results with the real consumption values and computing the Mean Absolute Percentage Error (MAPE). We have selected this error measure to evaluate the performance of the models since it is unit free; this is, it allows comparing the forecasting errors from different measurement units. Moreover, it is the most widely used error measure in forecasting [20]. It is calculated as follows:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\|C_r(i) - C_p(i)\|}{C_r(i)} \times 100, \tag{5}$$

where N is the total number of samples, in this case 24 hours, $C_p(i)$ is the predicted value of the load and $C_r(i)$ the actual one. Tab. II shows the results of the day-ahead forecasting with dataset 1. Please note that we predicted hour by hour a day, computing the corresponding MAPEs, and the one presented in the table is the average value of all the forecasts. Similarly, Tab. III presents the results for the day-ahead forecasting with dataset 2.

TABLE II
MAPE RESULTS IN **DAY-AHEAD** FORECASTING WITH DATASET 1 (%).

DataSet 1	**Sundays**	**Saturdays**	**Weekdays**	**Average**
AR model	12,23	8,53	5,81	8,14
ARIMA	20,21	21,38	17,32	19,13
Neural N	2	5,43	4,45	4,11
Bayesian N	15,99	33,35	18,93	22,75

TABLE III
MAPE RESULTS IN **DAY-AHEAD** FORECASTING WITH DATASET 2 (%).

DataSet 2	**Sundays**	**Saturdays**	**Weekdays**	**Average**
AR model	4,12	5,54	3,52	4,26
ARIMA	10,31	25,81	8,09	13,54
Neural N	1,69	5,22	3,33	3,46
Bayesian N	5,58	9,13	5,90	6,87

As it can be seen, the results with data set 1 are poor everywhere, except for the performance achieved by the AR model in weekdays (5,81%) and the exceptionally good marks of the NN in Sundays (2%). The AR manages to capture the typical consumption curve and only fails in the very odd values (i.e. those 5,81%); the nice results of the NN responds to the following reason: the construction works took place on Sundays only in some months and it succeeded to somehow distinguish which month did have works and which not.

TABLE IV
MAPE RESULTS IN **2-DAYS-AHEAD** FORECASTING WITH DATASET 1 (%).

DataSet 1	Sundays	Saturdays	Weekdays	Average
AR model	19,79	14,10	9,43	13,26
ARIMA	19,24	13,26	9,76	13,05
Neural N	15,88	19,40	11,01	14,46

TABLE V
MAPE RESULTS IN **2-DAYS-AHEAD** FORECASTING WITH DATASET 2 (%).

DataSet 2	Sundays	Saturdays	Weekdays	Average
AR model	8,11	10,29	6,17	7,86
ARIMA	12,04	24,47	9,14	14,62
Neural N	8,28	12,18	7,63	9,12

The ARIMA model shows overall a very poor performance, specially with the noisy data period (dataset 1). The Bayesian network performs acceptably good with a regular dataset but fails when predicting the heterogeneous one. The reason is the clustering of the values to be predicted: data too scarce and different is more difficult to be grouped into a number of homogeneous clusters. The Neural network suffers less from the noisy data or, at least, it shows less difference in the performance with both datasets. Moreover, it obtains on Sundays of dataset 2 a value (1,69%) that stays very close from the best record in the literature (1,51%), showing that it is the best choice when modelling data with slight changes (as it is the case in this dataset). Finally, the AR model shows the best overall performance, with very acceptable marks, specially with *normal* data.

The best record to our knowledge in short-term load forecasting presented a MAPE of the 1.53 % [37]. According to [1], simply the reduction of the 1% in the average forecast error may save hundreds of thousands or even millions of dollars. In our problem domain, non-residential buildings, this possibility is not plausible since an 1% error may be a deviation of some kW. Hence, [20] accomplished a 1,945% (two points above the AR model in dataset 2, only a 1,31% difference if we take just weekdays) using a neural network (as in the case of [37]), with all the problems aforementioned that NN present (for instance, [37] uses a fuzzy set-based classification algorithm to improve the classification ability of their NN). Still, Neural networks (also Bayesian networks) offer a worse trade-off between the difficulty of design, parametrisation, etc., and the performance, in comparison, for instance, with time series.

In the next experiment, we have tested the models when predicting not the next hour but for the next hour of the next day (i.e. day ahead prediction). Tab. IV shows these results with dataset 1 and Tab. V with dataset 2 (since the BN model performed poorly, we skip it's results).

In 2-days-ahead prediction, all maintain their day-ahead behaviour: very bad in the case of ARIMA, and better in ARIMA and AR, with a meritorious 5,75% MAPE in weekdays (it fails in weekends, specially in Sundays). According to these results, it seems that the NN degrades quite quickly (MAPE average with dataset 1 from 6,35 to 14,46%, dataset 2 from 4,86 to 9,12%), where the AR model holds better. In order to further test it's performance, we have further tested it with dataset 2 to predict 3, 4, 5, and 6 days ahead. Tab. VI shows these results.

As we see, the degradation is much worse in Saturdays

TABLE VI
AR MODEL PERFORMANCE (MAPE) IN >2 DAY-AHEAD FORECASTING.

	Sundays	Saturdays	Weekdays
3 days ahead	16,03	17,21	9,53
4 days ahead	17,91	18,83	11,69
5 days ahead	19,32	19,10	11,54
6 days ahead	22,89	19,96	12,32

and Sundays than in weekdays where, even 6 days ahead, remains acceptably accurate (obviously, depending on the actual scenario). Therefore, taking into account the results shown in Tabs. II,III, IV, V, and Tab.VI, we can conclude that the time series model fits the requirements we set out in Section I.

V. CONCLUSION

The old picture in which generation, distribution, and retail of energy (sometimes even transport) was under the auspices of a single company belongs to the past century in many countries. In Europe, the Directive 96/92/EC provided for the legal unbundling of the Transmission Systems and Distribution Systems operations and designed a clear route-map to establish a wholesale electricity market for electricity generation and a retail electricity market for electricity retailing. Allegedly, it would contribute to cheapen the overall cost of energy (the same Directive for the gas sector was approved two years later, 98/30/EC). Moreover, energy cannot be stored or held in stock (as tangible goods can) and the power grid must maintain a balance on real-time between the amount electricity produced and consumed. Otherwise, the risk of a blackout cannot by overseen, both caused by *defect* or *excess* of energy in the grid (i.e. production does not equal consumption). In this way, being the demand of energy stochastic *per se*, all participants in this new scenario must deal with the problem of forecasting at least at the short-run the amount of energy they will have to generate, transport, distribute, and so on.

Research on STLF usually follows two different paths. On the one hand, it handles the demand as a global aspect associated to a territory or country; therefore, this branch is specially suited for generators and Transmission System Operators. On the other hand, it focuses on predicting the demand of minor units (say a company, a village, a building or group of...), which, as we have seen, may help the whole electricity chain, from producers to the client itself. Our work goes in this latter direction, centred in non-residential buildings due to their stable and activity-dependent load profile. We have shown that the activity, which is the stochastic fact that introduces the non-linearity on the load (and which yields it so difficult to forecast), can be best represented by the work day schedule. Further, we have also demonstrated that weather has negligible influence on the load, since a method that does not take it into account outperformed the rest who did.

In our experience, a forecasting method for non-residential buildings should be simple and require no difficult trial-and-error customisation. Moreover, it should be able to work with any or sparse historical data, taking into account that the load data is also susceptible to evolve over time. Finally, it should be as accurate as possible. Under these premises, we have tested several statistical and artificial intelligence methods, namely an Autoregressive model, an ARIMA model, a Neural network, and a Bayesian network. Our experiments have shown that a classification using the work day schedule, and a curve adapting by means of an autoregressive time series suffices to answer the proposed requirements with an acceptable prediction accuracy maintained if we extend the prediction horizon even to 6 days.

Future works will focus first on testing other statistical methods that explicitly includes exogenous variables (to add temperature or THI values) such as ARMAX (Autoregressive moving average model with exogenous inputs model) or NARX (nonlinear autoregressive exogenous model), as well as SVMs (Support Vector Machines), though they usually show poorer results than NNs in this domain [1]. Moreover, we will investigate why the performance of the BN is much worse with noisy data and will use a different clustering technique (most likely a dynamic one) to this end. We will also test the BN with one whole year data, so the results are comparable to the rest. We will also explore the possibility of combining the NN with the AR model, applying one or another depending the day-type and the nature of the load profile. Finally, we plan tackle (normal) short-term load forecasting by applying the same strategy as here to avoid the non-linearity (i.e. use the work-day schedule to predict the day type). In STLF there is no concern for simplicity but new problems arise (such as computing resources required, computing time, etc.).

ACKNOWLEDGEMENTS

This research work has been partially sponsored by the Spanish MEC (ENERGOS TIN2009-00000) and the Provincial Council of Gipuzkoa (EIKE::Energia-Ikerketak 2009).

REFERENCES

[1] H. Alfares and M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23–34, 2002.

[2] V. Hinojosa and A. Hoese, "Short-term load forecasting using fuzzy inductive reasoning and evolutionary algorithms," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 565–574, 2010.

[3] Y. Penya, C. E. Borges, D. Agote, and I. Fernández, "Short-term load forecasting in non-residential buildings," in *Proceedings of the* 10^{th} *IEEE Region 8 AFRICON*, 2011, in press.

[4] E. Feinberg and D. Genethliou, "Load forecasting," in *Applied Mathematics for Power Systems, Chapter 12*, 2005, pp. 269–285.

[5] G. Mbamalu and M. El-Hawary, "Load forecasting via sub optimal seasonal auto regressive models and iteratively reweighted least squares estimation," *IEEE Transactions on Power Systems*, vol. 8, no. 1, pp. 343–348, 1993.

[6] A. Douglas, A. Breipohl, A. Lee, and R. F.N. Adapa, "The impact of temperature forecast uncertainty on bayesian load forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1507–1513, 1998.

[7] R. Sadownik and E. P. Barbosa, "Short-term forecasting of industrial electricity consumption in brazil," *International Journal of Forecasting*, vol. 18, no. 3, pp. 215–224, 1999.

[8] S. Huang, "Short-term load forecasting using threshold auto regressive models," *IEE Proceedings on Generation, Transmission and Distribution*, vol. 144, no. 5, pp. 477–481, 1997.

[9] D. Infield and D. Hill, "Optimal smoothing for trend removal in short term electricity demand forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 1115–1120, 1998.

[10] D. Leith, M. Heidl, and R. Ringwood, "Gaussian process prior models for electrical load forecasting," in *Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems*, 2004, pp. 112–117.

[11] A. Douglas, A. Breipohl, A. Lee, and R. F.N. Adapa, "Practical experiences with modeling and forecast," *Time Series*, 1979.

[12] H. Yang and C. Huang, "A new short-term load forecasting approach using self-organizing fuzzy armax models," *IEEE Transactions on Power Systems*, vol. 13, no. 1, pp. 217–225, 1998.

[13] Z. Yu, "A temperature match based optimization method for daily load prediction considering dlc effect," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 728–733, 1996.

[14] W. Charytoniuk, M. Chen, and P. Van-Olinda, "Non parametric regression based short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 725–730, 1998.

[15] A. Harvey and S. Koopman, "Forecasting hourly electricity demand using time-varying splines," *Journal of American Statistics Assoc.*, vol. 88, no. 424, pp. 1228–1236, 1993.

[16] J. Taylor and S. Majithia, "Using combined forecast switch changing weights for electricity demand profiling," *Journal Operations Research Society*, vol. 51, no. 1, pp. 72–82, 2000.

[17] S. Soliman, S. Persaud, K. El-Nagar, and M. E. El-Hawary, "Application of least absolute value parameter estimation based on linear programming to short-term load forecasting," *Electrical Power and Energy Systems*, vol. 19, no. 3, pp. 209–216, 1997.

[18] M. Hagan and S. Behr, "The time series approach to short term load forecasting," *IEEE Transactions on Power Systems*, vol. 2, no. 3, pp. 785–791, 1987.

[19] A. Jain and B. Satish, "Clustering based short term load forecasting using support vector machines," in *Proceedings of the IEEE Bucharest PowerTech*, 2009, pp. 1–8.

[20] P. González and J. Z. no, "Prediction of hourly energy consumption in buildings based on a feedback artificial neural network," *Energy and Buildings*, vol. 37, no. 6, pp. 595–601, 2005.

[21] D. J. C. MacKay, "Bayesian non-linear modelling for the prediction competition," in *ASHRAE Transactions, V.100, Pt.2*. Atlanta Georgia: ASHRAE, 1994, pp. 1053–1062.

[22] A. Dhar, T. A. Reddy, and D. E. Claridge, "Modeling hourly energy use in commercial buildings with fourier series functional forms," *Journal of Solar Energy Engineering*, vol. 120, no. 3, pp. 217–223, 1998.

[23] ——, "Generalization of the fourier series approach to model hourly energy use in commercial buildings," *Journal of Solar Energy Engineering*, vol. 121, no. 1, pp. 54–62, 1999.

[24] B. Dong, C. Cao, and S. Lee, "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, no. 5, pp. 545–553, 2005.

[25] R. H. Dodier and G. P. Henze, "Statistical analysis of neural networks as applied to building energy prediction," *Journal of Solar Energy Engineering*, vol. 126, no. 1, pp. 592–600, 2004.

[26] B. Ekici and U. Aksoy, "Prediction of building energy consumption by using artificial neural networks," *Advances in Engineering Software*, vol. 40, no. 5, pp. 356 – 362, 2009.

[27] G. Darbellay and M. Slama, "Forecasting the short-term demand for electricity do neural networks stand a better chance?" *International Journal of Forecasting*, vol. 16, pp. 71–83, 2000.

[28] L. Soares and L. Souza, "Forecasting electricity demand using generalized long memory," *International Journal of Forecasting*, vol. 22, no. 1, pp. 17 – 28, 2006.

[29] E. Castillo, J. Gutiérrez, and A. Hadi, *Expert systems and probabilistic network models*. Springer-Verlag, 1997.

[30] J. Pearl, "Graphical models for probabilistic and causal reasoning," in *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Volume 1: Quantified Representation of Uncertainty and Imprecision*, D. M. Gabbay and P. Smets, Eds. Dordrecht: Kluwer Academic Publishers, 1998, pp. 367–389.

[31] R. Hofmann and V. Tresp, "Discovering structure in continuous variables using bayesian networks," in *Advances in Neural Information Processing Systems 8*. MIT Press, 1995, pp. 500–506.

[32] S. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, pp. 241–254, 1967.

[33] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search, 2nd Edition*, 1st ed. The MIT Press, 2001, vol. 1.

[34] U. Kjærulff and A. Madsen, *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*, ser. Information Science and Statistics. Springer, 2008.

[35] D. Roobaert, G. Karakoulas, and N. Chawla, "Information gain, correlation and support vector machines," in *Feature Extraction, Foundations and Applications*, I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh, Eds. Springer, 2006, pp. 463–470.

[36] T. Mitchell, *Machine Learning*. McGraw-Hill, 1997.

[37] M. Daneshdoost, M. Lotfalian, G. Bumroonggit, and J. Ngoy, "Neural network with fuzzy set-based classification for short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1386–1391, 1998.