Probabilistic Day-Ahead Forecasting of Household Electricity Demand

Alexis Gerossier, Robin Girard, Georges Kariniotakis, Andrea Michiorri

▶ To cite this version:

HAL Id: hal-01518373

https://hal.science/hal-01518373

Submitted on 4 May 2017

PROBABILISTIC DAY-AHEAD FORECASTING OF HOUSEHOLD ELECTRICITY DEMAND

Alexis GEROSSIER
MINES ParisTech, PSL Research University, PERSEE, France, alexis.gerossier@mines-paristech.fr

Robin GIRARD
MINES ParisTech, PSL Research University, PERSEE, France, robin.girard@mines-paristech.fr

George KARINIOTAKIS
MINES ParisTech, PSL Research University, PERSEE, France, georges.kariniotakis@mines-paristech.fr

Andrea MICHIORRI
MINES ParisTech, PSL Research University, PERSEE, France, andrea.michiorri@mines-paristech.fr

ABSTRACT

Forecasting electricity demand at the local level of a building up to a feeder is increasingly necessary in several applications in the smart-grids context. Actors like aggregators and retailers, and tools like home energy management systems, require such forecasts as input. In this paper, a probabilistic day-ahead forecasting model is proposed to predict hourly electrical demand from individual households. This stochastic model uses smart-meter data and temperature predictions to make quantile forecasts. Performance is evaluated using data from a real-life smart grid demonstration site developed in Évora, Portugal as part of the European project SENSIBLE. The proposed model consistently outperforms a persistence model and provides reliable probabilistic forecasts.

INTRODUCTION

Accurate forecasts of electricity demand at regional or national scale are a prerequisite for power system operators to plan electricity production efficiently. Abundant literature exists on large-scale electrical load forecasting. Models in both literature and operational systems are reported to perform to the order of 1%-3%. For instance, Taylor and Snyder report an error of around 2% for day-ahead French and British demand employing advanced univariate methods [1]. With the increasing share of renewable energy sources (RES) in the generation mix, production is more and more decentralized. In addition, storage devices are installed at local level and active demand options are more frequently deployed. Integrating these options calls for the development of smart grids. Forecasting electricity demand at a higher spatial resolution, and not only for a large region, is therefore of interest. This paper deals with the problem of load forecasting at household or building level.

In the current smart grid context, new business models are emerging, such as retail companies. Efficiently predicting household demand is important for these firms to optimise the cost of supply for their customers and anticipate energy purchases. Retailers can offer electricity flexibility to network operators based on their pool of individual clients. They can for example sell a consumption shift for the next day from a peak to an off-peak hour. This type of business usually relies on the day-ahead electricity market (such as the EPEX or Nord Pool spots). For this reason, the focus in this paper is hourly forecasts for the next-day demand.

Although a wealth of literature exists on load forecasting at regional/national scale, few studies examine load forecasting at customer level. The availability of smart-meter data, along with the emergence of applications in the smart-grids context, encourage the development of efficient forecasting approaches. Forecasting household demand is not straightforward [2]. Different households have very different electricity usage depending on the number of inhabitants and their livestyle, and the size of the building. Moreover, consumption in each household is highly volatile from one day to the next, *e.g.* the electrical load profile varies greatly between a Thursday in winter and a Sunday in summer, and throughout a single day due to house occupancy and activities. All of these factors go against the usual single-point predictions. Providing a single value for household consumption for a particular time on the next day is indeed of little relevance because of the high level of inherent uncertainty. That is why in this paper we aim at proposing a model that produces probabilistic forecasts. Therefore, the forecast for one hourly load on the next day is given as a list of quantiles assessing the probability of a result below the quantiles. Uncertainty is then quantified and included in the model instead of being unknown. Interest in probabilistic predictions in the energy sector has grown strongly in recent years, including the organisation of model competitions focusing specifically on this aspect [3].

This paper is organised as follows. In the next section, we introduce the SENSIBLE project for which we developed and evaluated our model. Our statistical method is presented in section 3. Finally, results are reported in section 4.

DATA REQUIREMENTS & QUALITY

Implementing hourly load predictions at household level requires precise, frequently updated measurements of actual electricity consumption in the home. Thanks to the development of smart-meters, measurements are now recorded cost effectively. Such datasets are rarely publicly available due to user privacy issues. In this work we consider data made available as part of the European project SENSIBLE, namely from the demonstrator site in Évora, Portugal, which recorded measurements of a neighbourhood of 226 buildings throughout 2015. All of the buildings are close together, within a circle of 250 m radius, and are mainly residential, with a few restaurants and shops. The neighbourhood forms a pool of customers for a retailer, whose business model is to optimise the cost of electricity for its customers and also sell flexibility to an

electricity market.

In the real world, smart-meters are not perfect: some measurements are missing due to technical problems and a few devices might not work at all. Forecasting models should be robust enough to account for these situations which are applied in real life. For this reason, these problems should be accounted for in the forecasting scheme and should not be removed by selecting a more convenient data subset to develop the model. Table 1 summarises the quality of the measurements for the 226 households.

In addition to smart meter measurements, weather prediction of the next day is necessary to take full advantage of our model.

Table 1 — Overview of measurements quality for the 226 households.

Available measurements	Less than 1 month	From 1 to 11 months	More than 11 months
Number of households	28	34	164

FORECASTING METHOD

In order to build a forecasting model, we initially separate the available data into a training set and a testing set as explained later for the test case considered in this paper. The training set serves for estimating the model parameters.

Whenever there are sufficient data for a specific household, the median consumption for that household for each hour of the week is computed on the training set. This means that the training set for one household contains at least one measure for every hour of the week. If this is not the case, missing hours are either interpolated by a linear regression, or taken from an average of the surrounding households. For every household, the "median week" is obtained by concatenating the 168 (7x24) median hourly loads. This vector is stored and is used as a fallback prediction when the advanced model has problems producing a forecast of adequate quality (i.e. missing input

for a long period). An example of this median consumption during one week can be seen in Figure 1.

The training set is then normalized by these median hourly loads for each household. We thus have to predict the difference between a typical median day and the next day.

Quantile smoothing splines regression

To forecast the difference between the next day and the mean day previously stored, we propose a quantile smoothing spline regression, inspired by the model developed by Gaillard *et al.* [4].

The hourly load y_t at instant t is supposed to be a realisation of an unknown random variable. When making single-point predictions, it is often assumed that this random variable follows a Gaussian distribution and that only its expected value matters, forming the single-point predicted value. In a probabilistic framework, the whole distribution matters, and not only the expected value. In the general case, this distribution cannot be found on the whole infinite semi-axis $\mathbb{R}_+$. Quantile $\tau \in [0,1]$ is therefore used to characterise the distribution. The quantile value for $\tau = 0.2$ means that there is a 20% chance that the actual value will be lower. A list of quantiles can then approximately describe the random variable distribution.

We suppose that each quantile y_t^τ of the random variable, with $\tau \in [0,1]$, depends on n independent variables $x_1, \dots, x_n$ through functions, *i.e.*

$$y_t^\tau = f_{1,h}^\tau(x_1) + f_{2,h}^\tau(x_2) + \cdots + f_{n,h}^\tau(x_n).$$

The functions $f_{i,h}^\tau(.)$ depend on the hour of the day $h \in \{0, \dots, 23\}$ and can be different according to the quantile chosen. It is indeed possible that one variable x_i has practically no influence for low quantiles but strong influence for higher quantiles.

One should be extremely careful when selecting the explanatory variables $x_1, \dots, x_n$ for different reasons. The number n should be low, since the more explanatory variables there are, the more difficult it is for the functions to be trained. The number of training points required increases exponentially with n. By keeping this number low, we also keep it controllable since we can visualise
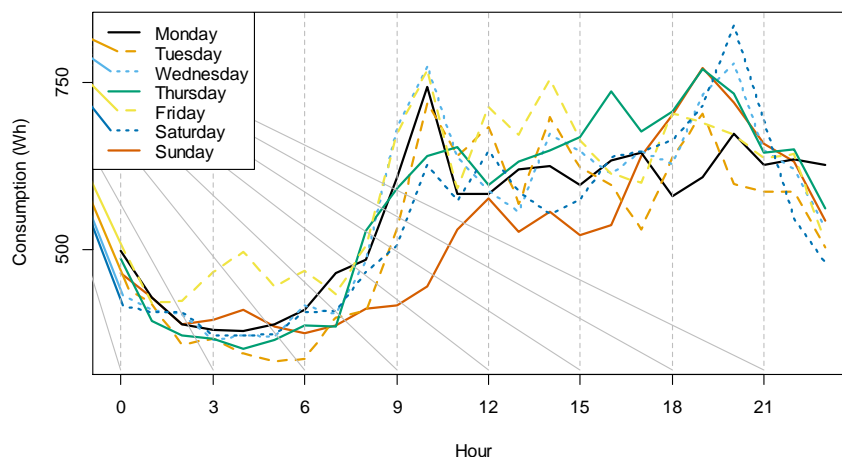

Figure 1 — Median consumption of one household for every day of the week computed over the training set.

what is happening and which factors have strong influences.

For these reasons, we select only three explanatory variables:

- **Consumption of the previous available day at the same hour**. Depending on the time of the day we choose to carry out the prediction, this can be y_{t-24} or y_{t-48} (when the exact horizon is greater than 24 hours, y_{t-24} is still unknown and cannot be used). For reasons of clarity, we always denote this by y_{t-24} in the following.

- **Median consumption of the hourly load during the past week**. For the same reason, this can be either

$$\text{median}(y_{t-24}, y_{t-48}, \dots, y_{t-168})$$

or

$$\text{median}(y_{t-48}, y_{t-72}, \dots, y_{t-192})$$

depending on the exact horizon. Taking the median is more robust than taking the mean and prevents possible missing values weighting too much in the input. This will be referred to as $\overline{y}_t$.

- **Local temperature prediction**. We use the latest available temperature prediction for the neighbourhood obtained from a Numerical Weather Prediction (NWP) model. In the case study of this paper we consider NWPs provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). It is well known that temperature is highly relevant when predicting electrical [5]. Hereafter this variable is denoted as $\hat{T}_t$.

The first two variables where selected by looking at the autocorrelation of the series, in which it is clear that loads at the same hour of previous days are strongly correlated with load at instant t. The temperature variable is a major meteorological factor and is especially important when outside temperature sharply rises or drops between successive days.

By changing the name of the functions, the load at instant t during hour h of the day writes

$$y_t^\tau = a_t^\tau(y_{t-24}) + b_t^\tau(\overline{y}_t) + c_t^\tau(\hat{T}_t). \tag{1}$$

Function estimation

In our application, only 9 quantiles were needed for $\tau \in \{0.1, \dots, 0.9\}$. This totals 648 functions to be estimated (24 hours $\times$ 9 quantiles $\times$ 3 variables). They are estimated with the function rqss developed by Koenker inside the quantreg package [5]. Using all of the values in the training set, we are able to estimate these functions. Once estimated, we can use the functions to predict the future value of the loads by changing inside variables according to the latest known ones.

Figure 2 shows the shape of the three functions for $\tau \in \{0.1, 0.5, 0.9\}$ at different times of the day for household 37.

Use of fallback model

Three kinds of problems can impede the use of equation (1):

- **Not enough data to estimate functions.** This occurs when the training set has too few available data.

- **Unavailable variables.** For example, it is impossible to use equation (1) to predict load on day n at noon if loads at noon on 7 previous days have not been recorded.

- **Extreme situations.** Quantile smoothing splines should not be used to extrapolate. This means that if an extreme situation has never previously occurred (*e.g.* unprecedentedly low temperature in the testing set), equation (1) should not be used.

When one of these three issues arises it is preferable to use the fallback model described earlier, *i.e.*, use as forecasts the median week values computed on the training set or on data from surrounding buildings.

EVALUATION RESULTS

Training and testing sets

As previously described, we have hourly load measurements of the 226 households during one year from 1st January 00:00 to 31st December 23:00. Some
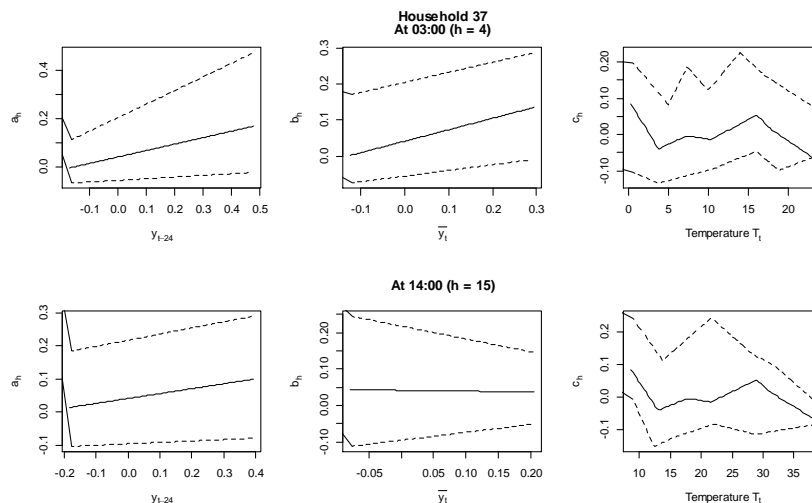

Figure 2 — Estimation of the functions for one household. Top row shows functions at 3:00 during the night, and bottom row shows function at 14:00. Solid lines shows the median ($\tau = 0.5$) and dashed lines show extrem quantiles, for $\tau = 0.1$ and $\tau = 0.9$.

households have missing values as mentioned. Following standard practice, we divide the dataset in two: a training set from 1st January 00:00 to 31st October 23:00 supposed known (7,296 values) and a testing set from 1st November 00:00 to 31st December 23:00 supposed unknown (1,464 values). The training set is used to train the model and estimate its parameters, while the testing set is used to simulate a real and out-of-sample application with an unknown future to predict.

For each household, the aforementioned functions are estimated on the training set. For the testing set, predictions were used in real conditions: each day d at noon, the 24 hourly loads of day $d + 1$ were predicted with 9 quantiles $0.1, 0.2, ..., 0.9$. Figure 3 shows the first three days in November for one household. The actual hourly load is plotted in gold and is to be predicted. Every day at noon (indicated by large ticks on the x-axis), the model issues 24 probabilistic predictions for the next day (as 24 lists of quantiles). The median prediction (for $\tau = 0.5$) is plotted in black. The dark area represents the prediction interval 30%—70% (*i.e.* the actual consumption has a 40% chance of being inside this interval). The light area represents the prediction interval 10%—90% (*i.e.* the actual consumption has an 80% chance of being inside this interval).

To evaluate the quality of the forecasting, several indices are used, two for single-point quality and one for probabilistic quality. We use $\langle y_t \rangle$ to note the mean over index t, and $\widehat{y_t}$ to refer to the estimation of y_t series:

- **MAPE** is the mean absolute percentage error. It is the mean of the error at each instant divided by the mean of the actual consumption. The normalization is useful to compare indices between two households with very different total consumptions. It writes

$$\text{MAPE} = \frac{\langle |y_t - \widehat{y_t}| \rangle}{\langle y_t \rangle}.$$

- **NRMSE** is the normalised root-mean-square error. It is similar to the MAPE except that it penalises large errors more strongly. It writes

$$\text{NRMSE} = \frac{\sqrt{\langle (y_t - \widehat{y_t})^2 \rangle}}{\langle y_t \rangle}.$$

- **CRPS** is the continuous ranked probability score. It is

very useful to evaluate probabilistic predictions as explained in detail by Gneiting *et al.* [7]. Instead of using one single point, it requires the cumulative distribution function $\widehat{F_t}$ found at instant t and compare its shape to the actual value with an integration, *i.e.*

$$\text{CRPS} = \langle \int^{y_t} \widehat{F_t}^2(y)\text{dy} + \int_{y_t} \left(1 - \widehat{F_t}(y)\right)^2 \text{dy} \rangle.$$

With the single-point prediction, the predicted cumulative distribution is Dirac. In this case, the CRPS is similar to the standard MAE (mean absolute error) criterion. That is why CRPS is sometimes mentioned as a generalisation of the MAE.

As a benchmark, a persistence model is considered, where the hourly loads of day d are directly used as a prediction for day $d + 1$. The persistence model is then evaluated with the first two indices but not with the CRPS since it only carries out single-point forecasts.

For each of the 226 households, we compute the above indices. In Table 2, we report only the average values of these indices over the whole set of 226 households and the subset of the 164 households with measurements during more than 11 months (see Table 1).

For both MAPE and RMSE, our model overall improves the persistence model. The improvement is even more significant when only the households with efficient smart meters are considered. This result can be expected because, in this case, the quantile regression model is well trained and carries precise prediction. For the other households (with an important number of missing values), improvement is less striking because the models are poorly trained with few values. Moreover, the evaluation on such households is difficult because just a few values are available to check whether the forecasts are correct or not. When considering the CRPS as a MAE generalisation, we can interpret the CRPS of 0.087 as "similar" to an absolute error of 87 Wh for every hourly load.

Although average indices are useful to globally assess the quality of a forecasting model, the details of error for each household bring additional insights about the quality. Figure 4 represents the MAPE for each one of the 226 households. One can see that, for a few households, errors are greater than 100%, meaning that special care should be
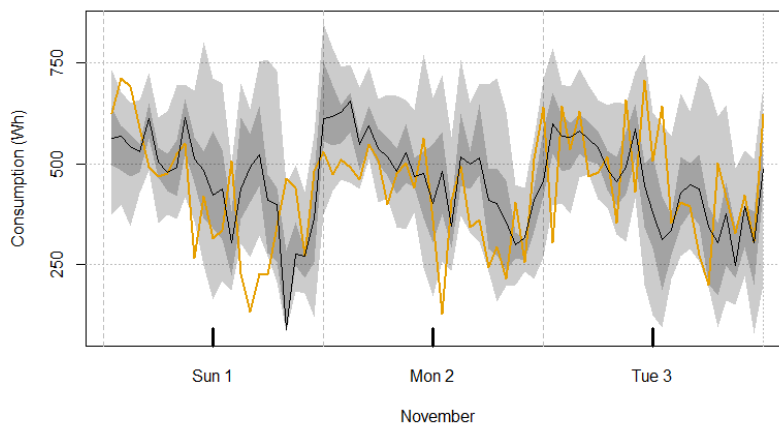

Figure 3 — Example of day-ahead forecasts for one household. The golden line is the actual consumption to predict. The black line is the 50% prediction, the dark shaded area is the 30%—70% interval and the light shaded area is the 10%—90% interval.

devoted to them.

Table 2 — Indices, averaged over all 226 households and over the 164 households with good measurements, evaluated on the testing set with the persistence model and the advanced model proposed here.

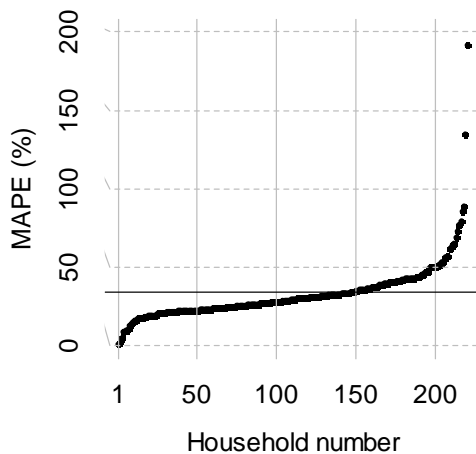|		MAPE	NRMSE	CRPS
226	Persistence	36.2%	49.6%	—
	Advanced	34.2%	43.9%	0.115
164	Persistence	35.6%	48.8%	—
	Advanced	28.9%	38.0%	0.087

Figure 4 — MAPE for each one of the 226 households. The black line represents the mean error.

CONCLUSION

In this paper, we presented an advanced forecasting model for hourly household electricity demand. The model employes a day-ahead horizon and provides probabilistic forecasts by computing a list of quantiles. It uses three inputs: the hourly load of the previous day, the median load of the previous week and the temperature prediction. Each of these three variables is considered to have an effect on the future load through a certain function that is estimated thanks to quantile smoothing spline regression.

The model was evaluated using data from a real-life test case in the city of Évora, Portugal. In a neighbourhood comprising 226 individual buildings, smart-meters recorded the hourly consumption of each building.

For each point of the testing set, we computed the mean of the predicted quantile distribution and used it as a single-point forecast. It was found that this advanced point forecast outperforms persistence by 15% on two different indices. The probabilistic forecasts were evaluated using the CRPS, which is a score assessing the quality of probabilistic models. Finally, fallback models were proposed, making use of surrounding smart-meters to overcome defective ones. This brings robustness to the proposed prediction approach, required to ensuire that reasonable probabilistic forecasts can be maintained. Further work includes implementing this model in an operational platform for on-line evaluation at the Évora demonstation.

REFERENCES

[1] J.W. Taylor, R.D. Snyder, 2012, "Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing", *Omega-International Journal of Management Science,* vol. 40, 748–757.

[2] R.A. Sevlian, R. Rajagopal, 2014, "A model for the effect of aggregation on short term load forecasting", *IEEE Power and Energy Society General*.

[3] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, R.J. Hyndman, 2016, "Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond", *International Journal of Forecasting,* vol. 32(3), 896—913.

[4] P. Gaillard, Y. Goude, R. Nedellec, 2016, "Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting", *International Journal of Forecasting,* vol. 32(3), 1038—1050.

[5] N. Charlton, C. Singleton, 2014 "A refined parametric model for short term load forecasting", *International Journal of Forecasting,* vol. 30(2)., 364—368.

[6] R. Koenker, 2013, "quanreg: Quantile Regression. R package version 5.05", *R Foundation for Statistical Computing,* Vienna, R Foundation.

[7] T. Gneiting, F. Balabdaoui, Raftery A.E., 2007 "Probabilistic forecasts, calibration and sharpness", *Journal of the Royal Statistical Society: Series B,* vol. 69(2)., 243—268.