Load forecasting for active distribution networks

Simone Paoletti[†], *Member, IEEE,* Marco Casini, *Member, IEEE,* Antonio Giannitrapani, *Member, IEEE,* Angelo Facchini, *Member, IEEE,* Andrea Garulli, *Senior Member, IEEE,* Antonio Vicino, *Fellow, IEEE*

***Abstract*—This paper addresses the problem of electric load forecasting for distribution networks with Active Demand (AD), a new concept in smart-grids introduced within the EU project ADDRESS. By changing the typical consumption pattern of the consumers, AD adds a new dimension to the problem of load forecasting, and therefore makes currently available load forecasting techniques no more suitable. A new approach to load forecasting in the presence of AD is therefore proposed. The approach is based on a decomposition of the load into its components, namely the base load (representing different seasonal patterns), and a residual term depending both on stochastic fluctuations and AD effects. The performance of the proposed approach is illustrated through a numerical example. Since data sets including AD are not yet available, in the numerical example AD effects are simulated and added to real measurements representing the aggregated load of about 60 consumers from an Italian LV network.**

***Index Terms*—Load forecasting, Active Demand, smart grids.**

I. Introduction

THE advances of information technologies and the increased accessibility of renewable energy resources to end users have triggered new concepts in electricity power distribution and consumption. One of these new concepts is Active Demand (AD), which has been introduced in the context of the European project ADDRESS [1]-[4]. ADDRESS stands for *Active Distribution network with full integration of Demand and distributed energy RESourceS* and its target is to enable the Active Demand in the context of the smart grids of the future. The key idea is that domestic and small commercial consumers will play an active role in the electricity system, adjusting their consumption patterns depending on the dynamics of the electricity markets. AD thus represents a new opportunity for solving network constraints and supporting the development of renewable energy sources through the flexibility it can offer, potentially providing economic benefits to all the participants.

Since individual consumers do not have direct access to the electricity market, a new intermediary function or entity, called the aggregation function or *aggregator*, is needed to coordinate the consumers' behavior with the market. Each aggregator has a pool of subscribers, and is able to send them price-volume signals in order to affect their consumption pattern. These signals essentially consist in specifying a monetary reward (price) if power consumption, during certain hours of the day, is below, above or in-between specified thresholds (volume). Since consumers have a certain degree of flexibility, they might find convenient scheduling or delaying certain tasks (e.g., running an appliance) so that they obtain the reward.

In this way, the aggregator collects a certain amount of energy over specified time intervals, namely the energy shifted by a number of consumers accepting the aggregator's offer. This energy can be used for several purposes. For instance, the Distribution System Operator (DSO) may ask an aggregator to enforce an energy reduction in a given Low Voltage (LV) load area, if an overload is foreseen in that area, in order to counteract possible network unbalancing. Similarly, the DSO could ask for a consumption increase when renewable energy sources are producing. Another reason for the aggregator to collect energy is that he has sold options for providing a certain amount of energy, and the option holder has now decided to exercise the option. Moreover, the aggregator can simply strive to gather energy during the time slots in which the market energy price is higher in order to sell the energy, possibly sharing part of the revenue with its consumers.

To enable AD, the objective of the ADDRESS project is to study, develop and validate technical solutions both at the consumers' premises and at the power system level. This paper focuses on the solutions for load forecasting at the DSO's control centre level by presenting the following contributions:

- analysis of the problem of load forecasting in the presence of AD;
- algorithm for load forecasting in the presence of AD;
- AD simulator to generate data sets for testing the forecasting algorithm;
- testing of the forecasting algorithm, by comparing the proposed approach with classical electric load forecasting approaches not considering explicitly AD.

Since AD is a peculiar concept of the ADDRESS project, all the aforementioned contributions are original.

The paper is organized as follows. Section II motivates the need for load forecasting in active distribution networks, and recalls some classical load forecasting techniques. The problem of load forecasting in the presence of AD is formulated and analyzed in Section III, and a solution to it is presented in Section IV. Section V describes the way AD is simulated in order to generate artificial data sets including AD effects. These data sets are used for the testing of the proposed load forecasting algorithm, whose results are presented in Section VI. Finally, conclusions are drawn in Section VII.

II. Preliminaries

A. Motivations

In the current operation of power networks, load forecasting represents a typical application in Transmission System Operators' control centres, used to aid planning and operational

[†] Corresponding author (e-mail: paoletti@dii.unisi.it).
The authors are with the Dipartimento di Ingegneria dell'Informazione and Centro per lo Studio dei Sistemi Complessi, Università di Siena, Siena 53100, Italy.

decisions at the transmission network level. However, fluctuations in supply and demand consequent, for instance, to more and more widespread distributed generation and use of plug-in electric vehicles, are making load forecasting increasingly important also for the (real-time) operation of distribution networks.

In addition to the motivations described above, in the ADDRESS framework the need for load forecasts at LV network level arises also due to the presence of AD, since this becomes a factor which may change the ”typical” behavior of consumers. The aggregator makes its business by offering load flexibility of its consumers in a given area. The DSO has the role to validate the supply bids sent by aggregators to the market by checking if they are compatible with the network constraints and operation. To do this, the DSO needs new forecasting tools able to take into account the load modifications determined by AD. Indeed, currently available load forecasting techniques do not consider AD among the exogenous inputs.

B. Short review on classical load forecasting

An extended survey and categorization of methods adopted for electric load forecasting is reported in [5]. Provided that the literature offers a wide range of methodologies and models to this purpose, a quite general approach is to model the load based on past data, and then use this model to predict the future load. In this respect, a very rough classification of the most popular techniques for load forecasting consists in distinguishing between linear and nonlinear models. Linear models are based on the assumption that the current load is a linear function of past values of the load itself and, possibly, of exogenous inputs such as temperature, weather forecasts, day type, etc. The most common linear models, borrowed from time series analysis, are the AutoRegressive Moving Average (ARMA) models [6], possibly with eXogenous inputs (ARMAX), or the more general class of Box-Jenkins (BJ) models [7], [8]. However, there are some characteristics of the load that cannot be reproduced by linear models, for instance seasonal characteristics at different time-scales (daily, weekly, yearly). This motivated the adoption of prediction techniques based on nonlinear models such as Nonlinear AutoRegressive models with eXogenus inputs (NARX), Neural Networks (NN) [9], and Support Vector Machines (SVM) [10].

III. LOAD FORECASTING IN THE PRESENCE OF AD

A. Notation

The load consumption in a given area of the distribution network is denoted by y, while the outdoor temperature is denoted by u. The requested AD profile sent by aggregators to consumers is denoted by ad, while the consumers' response to the signal ad is denoted by ad^{true}. Indeed, ad and ad^{true} may differ due to consumers not behaving according to the request.

Let T_s be the sampling time. It is assumed that $n_h = 60/T_s$ is an integer number, representing the number of samples per hour. It follows that $n_d = 24 \cdot n_h$ is the number of samples per day, while $n_w = 7 \cdot n_d$ is the number of samples per week. Typical values in ADDRESS are $T_s = 15$ min, $n_h = 4$, $n_d = 96$ and $n_w = 672$.

Conventionally, some Monday at 00:00 is chosen as the time origin. Hence, if one lets $k = 0, 1, 2, \dots$ be the discrete time index, a new day starts when k is a multiple of n_d, a new week starts when k is a multiple of n_w, and so on. The sample of a variable x at time kT_s from the time origin is denoted by $x(k)$. The forecast of $x(k + h)$, with h a positive integer, computed using only the information available up to time k, is denoted by $\hat{x}(k + h|k)$.

B. Problem formulation and analysis

The next sections focus on electric load forecasting in distribution networks with AD. The considered prediction problem is detailed next.

Problem 1: For fixed prediction horizon $h > 0$, predict the electric load at time $k+h$ based on the following information:

- load observations y up to time k;
- AD signal ad up to time $k + h$;
- temperature observations u up to time k, and predicted temperature $\hat{u}$ from time $k + 1$ to $k + h$.

Problem 1 could be in principle tackled by extending the techniques mentioned in Section II-B, simply considering AD as an additional input. In the so-called *black-box* approach, a mathematical relationship of the following type is estimated using a batch record of data:

$$y(k + h) = f(z(k)) + e(k), \tag{1}$$

where $z(k)$ is a vector of fixed dimension (called *regression vector*) containing (a subset of) the information available at time k, and $e(k)$ is the error process. Concerning the choice of the mapping $f(\cdot)$, it may range from simple linear structures to nonlinear ones (neural networks, kernel methods and support vector machines, etc.). The choice is typically made by considering the mapping that makes the error $e(k)$ “small” not only on estimation data, but also on validation data not used for estimation. The “predictor” is then given by:

$$\hat{y}(k + h|k) = f(z(k)). \tag{2}$$

The approach proposed in this paper to solve Problem 1 can be called *grey-box*, since it tries to exploit the characteristics of the variable to be forecasted (the load) and other available knowledge in order to enhance the prediction accuracy, but also to reduce the computational burden of the estimation algorithm, as is typically expected in model estimation when prior knowledge is used. In the problem under study, two types of prior knowledge are available:

- AD is expressed as load variations with respect to the expected load consumption profile if no AD request were sent. Therefore, assuming that M aggregators operate in the considered area, the load can be decomposed as follows:

$$y = y_b - \sum_{m=1}^{M} ad_m^{true}, \tag{3}$$

  where y is the actual load, y_b is the expected load if no AD requests were sent, and each ad_m^{true} term represents the actual AD profile (including energy payback effects) of the consumers enrolled with the m-th aggregator in response to an AD request ad_m. Indeed, ad_m and ad_m^{true} may differ due to delayed and/or partial response of the consumers with respect to the requested AD profile. Conventionally, positive (negative) values for ad_m and ad_m^{true} mean a decrease (increase) of the power consumption with respect to y_b.

- Classical load series (i.e. not including AD effects) show a strong seasonal behavior. Extracting known periodic (daily, weekly, yearly) patterns from time series helps the model estimation procedure to better capture the stochastic component of the underlying data generation mechanism. Based on this, y_b can in turn be decomposed as follows:
$$y_b = b + r_b, \tag{4}$$
where b is the so-called base load (the periodic pattern) and r_b is the residual due to stochastic fluctuations.

By substituting (4) into (3), one obtains:

$$y = b + r_b - \sum_{m=1}^{M} ad_m^{true} = b + r, \tag{5}$$

where the new residual

$$r = r_b - \sum_{m=1}^{M} ad_m^{true} \tag{6}$$

takes into account all perturbations to the base load b determined by both stochastic fluctuations and AD.

IV. Load Forecasting Algorithm

According to (5), the problem of forecasting y can be decomposed into two subproblems:

1) estimation of b;
2) forecasting of r.

Then, the predicted value of the load can be obtained as the sum of two contributions:

$$\hat{y}(k+h|k) = \hat{b}(k+h) + \hat{r}(k+h|k). \tag{7}$$

The aforementioned subproblems are addressed next.

A. *Estimation of the base load b*

For the sake of simplicity, in the following it is assumed that there are no special days such as holidays. These days have to be treated separately.

An estimate of the base load b can be obtained by applying an exponential smoothing to y_b. The adopted exponential smoothing estimates b as a smoothed version of y_b through the following formula:

$$\hat{b}(k) = \alpha\, y_b(k - n_w) + (1-\alpha)\, \hat{b}(k - n_w), \tag{8}$$

where $\alpha \in (0, 1)$ is called the *smoothing parameter*. Values of α close to one give greater weight to recent changes in the data, while values of α close to zero determine a stronger smoothing of the stochastic fluctuations. Fig. 1 shows an example of extraction of the base load from a load time series not including the effects of AD, by application of the exponential smoothing (8) with $\alpha = 0.1$. In Fig. 1, y_b is the aggregated load of 59 consumers in an Italian LV network. It can be seen in the bottom part of the figure, showing the residuals r_b, that the periodic features of the load are eliminated through the exponential smoothing. Moreover, the base load represents the dominant component of the load profile.

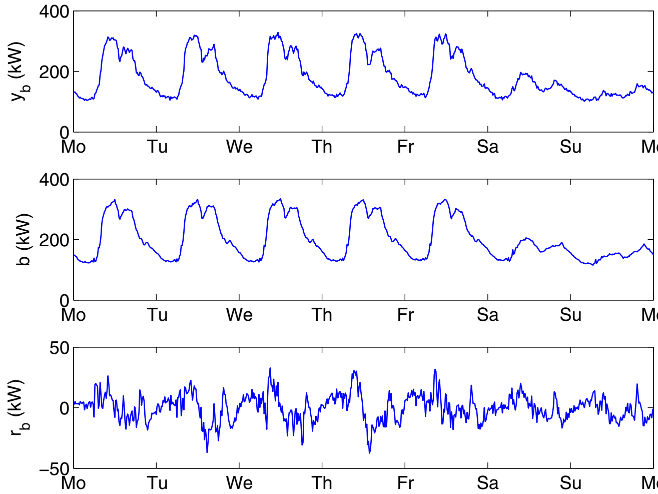

Fig. 1. Application of the exponential smoothing: (top) load time series y_b not including the effects of AD; (middle) base load b estimated via exponential smoothing with $\alpha = 0.1$; (bottom) residuals r_b after the extraction of the base load b from y_b.

Note that, according to (8), the base load is computed on a weekly basis. This may imply that variations of the load level due to, e.g., sudden temperature variations appear in the base load with one week delay. In order to make the estimate more reactive to temperature variations, the right-hand side of (8) can be modified by introducing the additional term $\beta\,(u(k) - u(k - n_w))$, where the coefficient β represents the average load increase per unit of temperature.

When evaluating (8), the most important implementation issue is concerned with the fact that the consumers' response to AD requests, namely the term $\sum_{m=1}^{M} ad_m^{true}$ in (3), may not be known, and therefore y_b cannot be reconstructed exactly from (3). The most straightforward way to overcome this problem is to roughly estimate y_b as follows:

$$\hat{y}_b = y + \sum_{m=1}^{M} ad_m, \tag{9}$$

i.e. by adding the nominal AD profile to the actual load, and then to replace y_b with $\hat{y}_b$ in (8).

B. *Forecasting of the residual r*

From (5), the residuals r are computed through the following relation:

$$r = y - b. \tag{10}$$

The problem of forecasting r is addressed by modelling r on the basis of a record of data previously collected, and then by using the estimated model to predict the future residuals.

Concerning the choice of the model structure, as recalled in Section II-B, one may range from linear structures to nonlinear ones (neural networks, kernel methods and support vector machines, etc.). In principle, no model structure can be a priori preferred to the others, since the available alternatives should be tested on real data in order to figure out which one achieves the best trade-off between prediction accuracy and computational burden. A possible solution based on linear models is proposed next. In spite of their simplicity, linear models have been shown to perform satisfactorily in short-term load forecasting [6].

The following model generating the residuals is assumed:

$$r(k) = \sum_{m=1}^{M} F_m(q) ad_m(k) + \frac{C(q)}{A(q)}\, e(k), \tag{11}$$

where $e(k)$ is a zero-mean white error process, and $A(q)$, $C(q)$ and $F_m(q)$, $m = 1, \dots, M$, are polynomials in the backward shift operator q^{-1}, satisfying $q^{-1}x(k) = x(k-1)$:

$$A(q) = 1 + a_1\, q^{-1} + \dots + a_{n_a}\, q^{-n_a} \tag{12a}$$

$$C(q) = 1 + c_1\, q^{-1} + \dots + c_{n_c}\, q^{-n_c} \tag{12b}$$

$$F_m(q) = f_{m,0} + f_{m,1}\, q^{-1} + \dots + f_{m,n_f}\, q^{-n_f}. \tag{12c}$$

The ARMA term $\frac{C(q)}{A(q)}\, e(k)$ models the stochastic fluctuations r_b, while each Finite Impulse Response (FIR) term $F_m(q)\, ad_m(k)$ models the consumers' response ad_m^{true} to the AD signal ad_m which has been sent by the m-th aggregator (see also Section V). The degrees n_a, n_c and n_f of the polynomials $A(q)$, $C(q)$ and $F_m(q)$ are selected by estimating a model (11) for different combinations of n_a, n_c and n_f, and then choosing the model with the smallest variance of the one-step-ahead prediction errors $\varepsilon(k) = r(k) - \hat{r}(k|k-1)$ computed using validation data, i.e. data not used for estimation. Model (8) can be put in one-step-ahead ($h = 1$) predictor form as follows:

$$\hat{r}(k|k-1) = \sum_{m=1}^{M} \frac{A(q)F_m(q)}{C(q)}\, ad_m(k) + \left[1 - \frac{A(q)}{C(q)}\right] r(k). \tag{13}$$

If $h > 1$, equations for the h-step-ahead predictor can be derived through recursive evaluations of (11).

If the temperature u is available, an additional exogenous term $B(q)r_u(k)$ can be introduced in the right-hand side of (11), where $B(q)$ is a polynomial in the backward shift operator q^{-1} and $r_u = u - b_u$, with b_u obtained by applying an exponential smoothing on a daily basis to u:

$$b_u(k) = \gamma\, u(k - n_d) + (1 - \gamma)\, b_u(k - n_d), \tag{14}$$

and $\gamma \in (0, 1)$. The reason for using r_u instead of u in (11) is that the periodicity of the load due to the daily temperature pattern b_u should be captured by the base load b, while r_u should help to explain the load variations due to intra-day temperature fluctuations.

Remark 1: The main limitation of model (11) could be that the consumers' response to AD requests is assumed to be time-invariant, i.e. the type of response is the same no matter the period of the day when AD is requested. More involved models can be devised to deal with time-varying AD responses. For instance, one could consider piecewise-linear models [11], where the consumers' response is characterized by different parameters $f_{m,i}$ in (12c) depending on the start time of the AD request. On the other hand, the use of linear models allows for a recursive implementation of the parameter estimation procedure. This makes it possible to adapt the model parameters a_i, c_i and $f_{m,i}$ in (12) to the time-varying characteristics of the residuals r.

V. Simulation of Active Demand

Since AD is not yet in practice, there are no real data sets including AD effects which can be used to test the proposed grey-box approach. For this reason, the numerical example presented in the next section is based on a data set where AD effects are simulated and added to a real load time-series. This section explains how AD is simulated.

Actual AD profiles ad_m^{true} are generated by filtering standard AD profiles ad_m (see description below) through a model simulating the response of consumers. Such a model is based on realistic considerations about the consumers' behavior:

- An AD request has a finite duration in time, and therefore it can be assumed that the duration of the consumers' response is finite.
- The consumers may not adhere exactly to the request. This can result in a delayed and/or partial response with respect to the requested AD profile.

If the model of the consumers' response is chosen to be linear, the above considerations imply the use of FIR models of the following type for $m = 1, \dots, M$:

$$ad_m^{true}(k) = \Phi_m(q) ad_m(k) + v_m(k), \tag{15}$$

where $v_m(k)$ is a zero-mean white error process, modelling the random perturbations, and $\Phi_m(q)$ is a polynomial in q^{-1}:

$$\Phi_m(q) = \phi_{m,0} + \phi_{m,1}\, q^{-1} + \dots + \phi_{m,n_\phi}\, q^{-n_\phi}, \tag{16}$$

modelling the consumers' typical behavior through the sum of a finite number of shifted and scaled copies of the requested AD profile.

Examples of AD profiles are shown in Fig. 2. The upper part of the figure shows a standard requested AD profile, composed of two parts. The first part is the actual AD service, whose profile is characterized by the volume V_{ad} and the duration of the service T_{dur}. The second part is the energy payback effect (sometimes called the *rebound* effect), corresponding to the fact that a demand modification requested by an aggregator (the AD service) may be followed by an opposite modification. The lower part of Fig. 2 shows what the actual AD profile could be, obtained by filtering a standard AD profile through model (15).

VI. Numerical example

The numerical example presented in this section is based on a data set of electric load including AD effects, and compares the prediction performance of the model proposed in Section IV with that of a model not taking AD explicitly into account. The aim is to strengthen the motivations for new forecasting tools in active distribution networks with

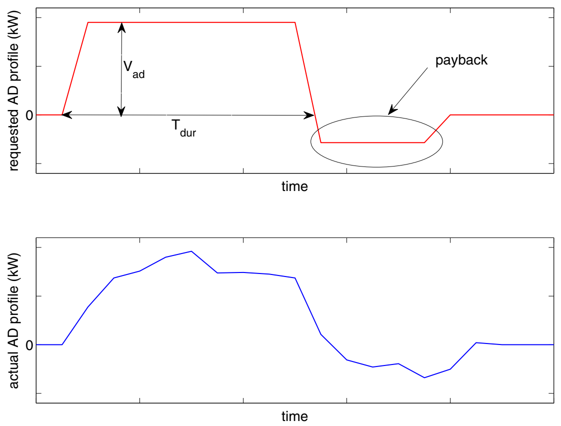

Fig. 2. AD profiles: (top) standard requested AD profile composed by an AD service (with volume V_{ad} and time duration T_{dur}) and a consequent energy payback effect; (bottom) simulated actual AD profile provided by the consumers.

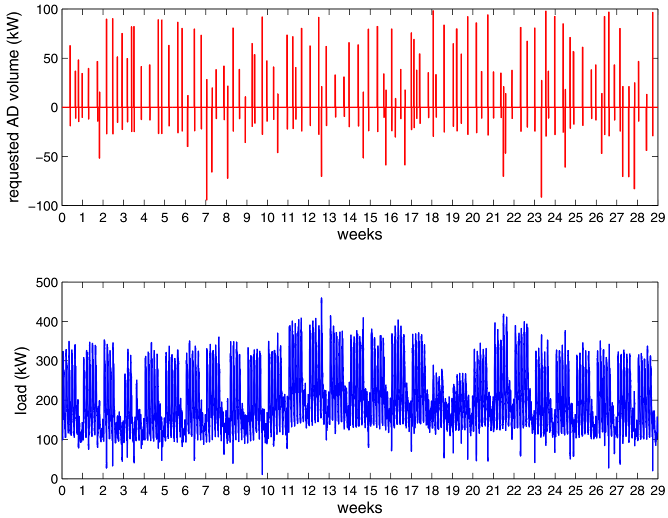

Fig. 3. Data sets used in the numerical example of Section VI: (top) plot of the AD signals sent by the aggregator to its consumers; (bottom) plot of the load time series including the AD effects determined by the consumers' response to the aggregator's AD signals.

AD (see Section II-A) by showing that currently available techniques for load forecasting cannot capture the unusual load modifications determined by AD.

The available data set contains real measurements representing the aggregated electric load of 59 consumers from an Italian LV network. The data set covers 29 weeks from April to October 2008 with sampling time $T_s = 15$ min, for a total of 19488 samples. Temperature records are not available. AD effects are simulated as described in Section V, and added to the load time series. The parameters used in model (15) are reported in Table I, where σ_v is the standard deviation of the error process and the subscript m is omitted because only one aggregator is considered ($M = 1$). Moreover, the requested AD profiles $ad(k)$ are generated by selecting randomly the AD volume V_{ad} in the interval $[30, 100]$ kW and the duration T_{dur} in the interval $[1, 4]$ hours. The energy payback is proportional to V_{ad} with a factor 0.3, and its duration is assumed to be half of T_{dur}. The interval between the end of an AD service and the begin of the next is chosen randomly between 12 hours and 3 days. Fig. 3 shows the resulting data sets, with the requested AD volume in the upper part, and the load including AD effects in the lower part.

A model of the load with the structure proposed in Section IV is estimated from data. This model is referred to as ES+BJ model, because it is composed by the exponential smoothing (8) and by the BJ model (11). For comparison purposes, a model of the load which does not take AD explicitly into account, is estimated. In order to make the comparison as fair as possible, the structure of the latter model is chosen similar to the ES+BJ model, with the only difference that AD is not used as an exogenous input. This implies that the equation of the exponential smoothing (8) is modified as follows, by substituting y_b with the total load y:

$$\hat{b}(k) = \tilde{\alpha}\, y(k - n_w) + (1 - \tilde{\alpha})\, \hat{b}(k - n_w), \qquad (17)$$

while the equation of the BJ model (11) reduces to the ARMA model:

$$r(k) = \frac{\widetilde{C}(q)}{\widetilde{A}(q)}\, e(k), \qquad (18)$$

where, for the sake of the comparison, the orders of the polynomials $\widetilde{A}(q)$ and $\widetilde{C}(q)$ are chosen equal to the orders of $A(q)$ and $C(q)$ in (11), respectively. This model is referred to as ES+ARMA model.

Both the ES-BJ model and the ES+ARMA model are used to predict the load for different prediction horizons h ranging from one sample (15 min ahead) to 96 samples (one day ahead). For fixed prediction horizon h, the performance of each model is evaluated by computing the following index:

$$\text{FIT} = 100 \cdot \left(1 - \frac{\sqrt{\sum_k \left(y(k+h) - \hat{y}(k+h|k) \right)^2}}{\sqrt{\sum_k \left(y(k+h) - \overline{y} \right)^2}} \right) \%, \qquad (19)$$

where $\overline{y}$ is the sample mean of y. The higher the fit, the better the performance. The plots of FIT versus prediction horizon are shown in Fig. 4 for both models. It can be observed that the proposed ES+BJ model exploiting AD performs better than the ES+ARMA model for all the prediction horizons considered. The largest difference between the two models is for $h = 8$ (2 hours ahead), when the ES+BJ model obtains FIT $= 78.3891\%$, while the ES+ARMA model obtains FIT $= 69.0659\%$. The 8-step-ahead prediction errors of the two models are shown in Fig. 5. Visual inspection of the plots confirms that the prediction errors of the ES+BJ model have smaller magnitude on average than the prediction errors of the ES+ARMA model. It can be further observed in Fig. 4 that,

TABLE I
PARAMETERS USED FOR AD SIMULATION

n_ϕ	ϕ_0	ϕ_1	ϕ_2	σ_v
2	0.55	0.30	0.05	10.00

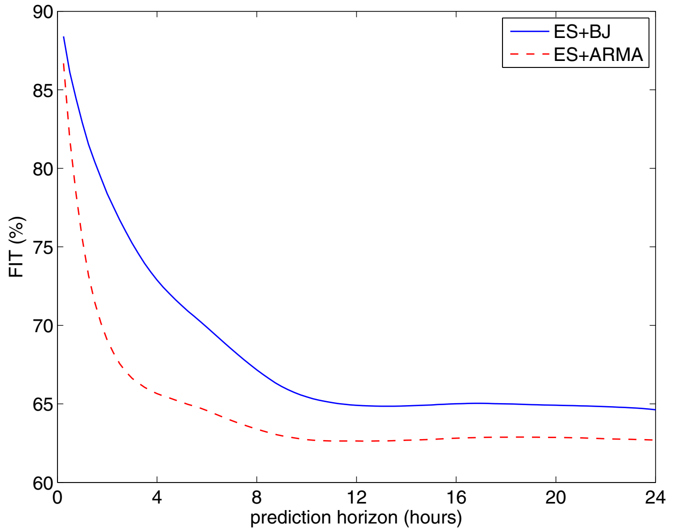

Fig. 4. Plots of FIT versus prediction horizon for models ES+BJ (solid) and ES+ARMA (dashed) considered in the numerical example.

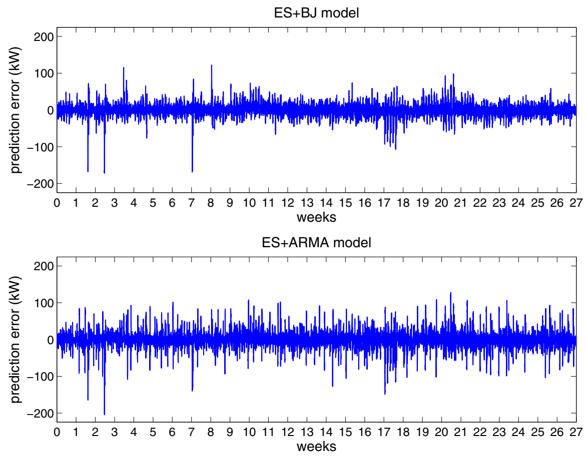

Fig. 5. Plots of the 8-step-ahead prediction errors for models ES+BJ (top) and ES+ARMA (bottom) considered in the numerical example.

as expected, the performance decreases with the prediction horizon for both models. However, the performance of the ES+ARMA model deteriorates more rapidly in the short period than the performance of the ES+BJ model.

It is worthwhile to note that the relative performance between the two models is influenced by several factors, for instance the frequency, the duration and the volume of the AD services. Indeed, if the AD services are very spaced in time, of short duration and of small volume, one may expect that the performances of the two models get close to each other. However, it can be fairly concluded that considering explicitly AD in the models used for load forecasting is always expected to lead to enhanced performance when AD effects of any amount are present in the load time-series to be predicted.

VII. Conclusions

In this paper, an approach to electric load forecasting in active distribution networks with Active Demand has been presented. The approach is based on a decomposition of the load into its components, namely the base load (representing different seasonal patterns), and a residual term depending both on stochastic fluctuations and AD effects. The AD profiles requested by aggregators to consumers are used both in the calculation of the base load, and in the forecasting of the residuals. A numerical example has shown that, when AD will come into practice, it will be appropriate to consider AD as an additional input in load forecasting algorithms, in order to maintain the current standards of accuracy.

It is stressed that, from the technical point of view, to enable and to exploit the benefits of AD, distributed intelligence will be required at different levels of the distribution networks. Concerning load forecasting, this means that measurements of electric load (and temperature, if used) must be available in real-time at the level of the LV load area of interest for AD. Measurements are needed in order to compute the predictions and tune the forecasting models on-line.

Acknowledgements

The authors would like to thank Regine Belhomme (EDF-R&D, France) for her useful comments and suggestions.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 207643.

References

[1] "The ADDRESS project," http://www.addressfp7.org/.

[2] R. Belhomme, R. Cerero Real de Asua, G. Valtorta, A. Paice, F. Bouffard, R. Rooth, and A. Losi, "ADDRESS - Active Demand for the smart grids of the future," in *Proc. 2008 CIRED Seminar: Smart Grids for Distribution*, 2008, paper no. 0080.

[3] E. Peeters, R. Belhomme, C. Batlle, A. Paice, F. Bouffard, S. Karkkainen, D. Six, and M. Hommelberg, "ADDRESS - scenarios and architecture for Active Demand development in the smart grids of the future," in *Proc. 2009 CIRED - 20th Int. Conf. on Electricity Distribution*, 2009, paper no. 0406.

[4] R. Belhomme, R. Cerero Real de Asua, G. Valtorta, and P. Eyrolles, "The ADDRESS project: Developing Active Demand in smart power systems integrating renewables," in *Proc. IEEE/PES General Meeting*, 2011, paper no. PESGM2011-000188.

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

[6] S.-J. Huang and K.-R. Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," *IEEE Trans. on Power Systems*, vol. 18, no. 2, pp. 673–679, 2003.

[7] L. Ljung, *System identification: theory for the user*. Prentice-Hall, Upper Saddle River, NJ, 1999.

[8] G. Box, G. Jenkins, and G. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. Wiley, 2008.

[9] D. Park, M. El-Sharkawi, R. Marks II, L. Atlas, and M. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Trans. on Power Systems*, vol. 6, no. 2, pp. 442–449, 1991.

[10] M. Espinoza, J. Suykens, R. Belmans, and B. De Moor, "Electric load forecasting using kernel-based modeling for nonlinear system identification," *IEEE Control Systems Magazine*, vol. 27, no. 5, pp. 43–57, 2007.

[11] E. D. Sontag, "Nonlinear regulation: The piecewise linear approach," *IEEE Trans. on Automatic Control*, vol. 26, no. 2, pp. 346–358, 1981.