Short-term load forecasting based on a semi-parametric additive model

Shu Fan, *Senior Member, IEEE,* and Rob J Hyndman

***Abstract*—Short-term load forecasting is an essential instrument in power system planning, operation and control. Many operating decisions are based on load forecasts, such as dispatch scheduling of generating capacity, reliability analysis, and maintenance planning for the generators. Overestimation of electricity demand will cause a conservative operation, which leads to the start-up of too many units or excessive energy purchase, thereby supplying an unnecessary level of reserve. On the contrary, underestimation may result in a risky operation, with insufficient preparation of spinning reserve, causing the system to operate in a vulnerable region to the disturbance.**

In this paper, semi-parametric additive models are proposed to estimate the relationships between demand and the driver variables. Specifically, the inputs for these models are calendar variables, lagged actual demand observations and historical and forecast temperature traces for one or more sites in the target power system. In addition to point forecasts, prediction intervals are also estimated using a modified bootstrap method suitable for the complex seasonality seen in electricity demand data. The proposed methodology has been used to forecast the half-hourly electricity demand for up to seven days ahead for power systems in the Australian National Electricity Market. The performance of the methodology is validated via out-of-sample experiments with real data from the power system, as well as through on-site implementation by the system operator.

***Index Terms*—short-term load forecasting, additive model, time series, forecast distribution.**

I. Introduction

LOAD forecasting is a key task for the effective operation and planning of power systems. The forecasting accuracy has significant impact on electric utilities and regulators. Overestimation of electricity demand will cause a conservative operation, which leads to the startup of too many units supplying an unnecessary level of reserve or excessive energy purchase, as well as substantial wasted investment in the construction of excess power facilities. On the contrary, underestimation may result in a risky operation and unmet demand, persuading insufficient preparation of spinning reserve and causes the system to operate in a vulnerable region to the disturbance.

Load forecasting is usually concerned with the prediction of hourly, daily, weekly, and annual values of the system demand and peak demand. Such forecasts are sometimes categorized as short-term, medium-term and long-term forecasts, depending on the time horizon. In terms of forecasting outputs, load forecasts can also be categorized as point forecasts (i.e., forecasts of the mean or median of the future demand distribution), and density forecasts (providing estimates of the full probability distributions of the possible future values of the demand).

Various techniques have been developed for electricity demand forecasting during the past few years [1]. Statistical models are widely adopted for the load forecasting problem, which include linear regression models, stochastic process models, exponential smoothing and ARIMA models [2]–[7]. To incorporate the nonlinearity of electricity demand series, artificial neural networks (ANNs) have also received substantial attentions in load forecasting with good performance reported [1], [8]–[14]. Neural networks have been shown to have the ability not only to learn the load series but also to model an unspecified nonlinear relationship between load and weather variables. Recently, machine learning techniques and fuzzy logic approaches have also been used for load forecasting or classification and achieved relatively good performances [15]–[18]. Although substantial attention has been paid to short-term load forecasting, a few researchers have proposed interval and probabilistic approaches for long-term forecast horizons; notably [19], [17] and [6].

The electricity demands nowadays are nonlinear and volatile, and are subject to a wide variety of exogenous variables, including prevailing weather conditions, calendar effect, demographic and economic variables, as well as the general randomness inherent in individual usage. How to effectively integrate the various factors into the forecasting model and provide accurate load forecasts is always a challenge for modern power industries.

The purpose of this study is to develop short-term load forecasting (STLF) models for regions in the National Electricity Market (NEM) of Australia. This paper follows the regression methodology, but focus on the non-linear relationships between load and various driving variables. This study aims to allow nonlinear and nonparametric terms within the regression framework. In particular, semi-parametric additive models are proposed to estimate the relationships between load and the exogenous variables, including calendar variables, lagged load observations and historical and forecast temperature traces for one or more sites in the target power systems. In addition to point forecasts, forecasting distributions are also estimated based on a bootstrap method.

The proposed methodology has been used to forecast the half-hourly electricity demand for up to seven days ahead for power systems in the NEM. The performance of the methodology is validated via out-of-sample comparisons using real data from the power systems. The proposed forecasting software has been used by Australian Energy Market Operator (AEMO) to

S. Fan and R. J. Hyndman are with the Business and Economic Forecasting Unit, Monash University, Clayton, VIC 3800, Australia. E-mail: (Shu.Fan@monash.edu; Rob.Hyndman@monash.edu)

Manuscript received August 2010.

forecast half-hourly electricity demand in Victoria and South Australia for system scheduling and planning, with satisfactory accuracy.

II. Methodology

A. Model establishment

The proposed semi-parametric additive model is in the regression framework but with some non-linear relationships and with serially correlated errors. In particular, the proposed models allow nonlinear and nonparametric terms using the framework of additive models [20].

Specific features of the models are summarized below:

- previous demand observations are used as predictors;
- temperatures from two sites are considered;
- temperature effects are modelled using regression splines;
- temperatures from the last three hours and the same period from the last six days are considered;
- loads from the last three hours and the same period from the last six days are considered;
- errors are serially correlated.

A separate model for each half-hourly period was fitted. Since the demand patterns change throughout the day, better estimates can be obtained if each half-hourly period is treated separately. This procedure of using individual models for each time of the day has also been applied by other researchers [16], [21], [22]

The model for each half-hour period can be written as

$$\log(y_{t,p}) = h_p(t) + f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t}) + a_p(\mathbf{y}_{t-,p}) + n_t \qquad (1)$$

where

- $y_{t,p}$ denotes the demand at time t (measured in half-hourly intervals) during period p ($p = 1, \ldots, 48$);
- $h_p(t)$ models all calendar effects;
- $f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t})$ models all temperature effects where $\mathbf{w}_{1,t}$ is a vector of recent temperatures at Richmond and $\mathbf{w}_{2,t}$ is a vector of recent temperatures at Melbourne airport;
- $a_p(\mathbf{y}_{t-,p})$ models the effects of recent demands;
- n_t denotes the model error at time t.

The logarithmic demand, rather than the raw demand, is modeled. A variety of transformations of demand from the Box-Cox (1964) class were tried and it was found that the logarithm resulted in the best fit to the available data. Natural logarithms have been used in all calculations.

1) Calendar effects: $h_p(t)$ includes annual, weekly and daily seasonal patterns as well as public holidays:

$$h_p(t) = \alpha_{t,p} + \beta_{t,p} + \ell_p(t), \qquad (2)$$

where

- $\alpha_{t,p}$ takes a different value for each day of the week (the "day of the week" effect);
- $\beta_{t,p}$ takes value zero on a non-work day, some non-zero value on the day before a non-work day and a different value on the day after a non-work day (the "holiday" effect);
- $\ell_p(t)$ is a smooth function that repeats each year (the "time of year" effect).

The smooth function $\ell(t)$ was estimated using a cubic regression spline. A regression spline consists of several polynomial curves that are joined at points known as "knots". Six knots are chosen at equally spaced times throughout the year for smooth function. The choice of knots is made automatically based on the forecasting performance.

2) Temperature effects: The function $f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t})$ models the effects of recent temperatures on the aggregate demand. Because the temperatures at the two locations are highly correlated, these were not used directly. Instead, the average temperature across the two sites,

$$x_t = (w_{1,t} + w_{2,t})/2,$$

and the difference in temperatures between the two sites,

$$d_t = (w_{2,t} - w_{1,t}),$$

were both used. These will be almost uncorrelated with each other making it easier to use in statistical modelling [24]. Then the temperature effects were included using the following terms:

$$\begin{aligned} f_p(\mathbf{w}_{1,t}, \mathbf{w}_{2,t}) = & \sum_{k=0}^{n} \left[f_{k,p}(x_{t-k}) + g_{k,p}(d_{t-k}) \right] \\ & + \sum_{j=1}^{m} \left[F_{j,p}(x_{t-48j}) + G_{j,p}(d_{t-48j}) \right] \\ & + q_p(x_t^+) + r_p(x_t^-) + s_p(\bar{x}_t), \end{aligned} \qquad (3)$$

where

- x_t^+ is the maximum of the x_t values in the past 24 hours;
- x_t^- is the minimum of the x_t values in the past 24 hours;
- $\bar{x}_t$ is the average temperature in the past seven days.
- m is the maximum number of lagged days considered in the model.
- n is the maximum number of lagged half-hourly periods considered in the model.

Each of the functions ($f_{k,p}$, $g_{j,p}$, $F_{k,p}$, $G_{j,p}$, q_p, r_p and s_p) was assumed to be smooth and was estimated using a cubic regression spline. In selecting the knots for the cubic regression splines, different knot positions and different numbers of knots were tried. Since it is not feasible to try all the possible knot positions, we considered knots equally spaced in the possible sample ranges, based on the finding that a minor change of knot position has little effect on the model performance. In this way, the computational burden can be effectively reduced without sacrificing the model's forecasting capacity. The knot positions that provided the best forecasting performance were selected in the model. In this paper, splines with knots at 9, 22 and 29°C were used for f, F and q. For the functions g and G, knots at -2.2 and -0.7°C were used.

3) Lagged demand effects: We incorporate recent demand values into the model. By doing this, some of the serial correlations within the demand time series can be captured within the model, and the variations of demand level throughout the time can be embedded into the model as well.

Demand effects, $a_p(\mathbf{y}_{t-,p})$, model the effects of recent demands using the following terms:

$$a_p(\mathbf{y}_{t-,p}) = \sum_{k=1}^{n} b_{k,p}(y_{t-k,p}) + \sum_{j=1}^{m} B_{j,p}(y_{t-48j}) + Q_p(y_t^+) + R_p(y_t^-) + S_p(\bar{y}_t), \quad (4)$$

where

- y_t^+ is the maximum of the y_t values in the past 24 hours;
- y_t^- is the minimum of the y_t values in the past 24 hours;
- $\bar{y}_t$ is the average demand in the past seven days.
- m is the maximum number of lagged days considered in the model.
- n is the maximum number of lagged half-hourly periods considered in the model.

Each of the functions ($b_{k,p}$, $B_{j,p}$, Q_p, R_p and S_p) is assumed to be smooth and is estimated using a cubic regression spline.

4) Error term: The error term n_t will be serially uncorrelated within each model, as the lagged demand terms remove serial correlations that might otherwise exist. Note that there will still be a small amount of correlations between residuals from the different half-hourly models.

B. Forecasting distribution estimation

Forecasting distributions convey more meaningful information than predicted point values, and are valuable in evaluating and hedging the financial risk accrued by demand variability and forecasting uncertainty. Forecasting distributions (and the prediction intervals that are obtained from them) provide an indication of the forecast accuracy, and give useful information for system scheduling. They also enable planning and operating decisions to be made in a stochastic as opposed to a deterministic context. For instance, the upper bounds of prediction intervals can be used for developing conservative electricity generation plan and schedules, while the use of the lower bounds of prediction intervals reflects an optimistic attitude in scheduling with more attention to over-supply avoidance.

Generally, there are two kinds of methods used in constructing forecasting distributions and prediction intervals: parametric methods and nonparametric methods. Selection of the appropriate method usually depends on the model, computational burden, number of available data, etc.

Typically, the parametric method assumes the disturbance is an i.i.d. normal variate with zero mean and finite variance. Sometimes, another distribution can be assumed. In this paper, the normality of the model residuals are tested using the Anderson-Darling test and Pearson chi-square test [25]. The p-values obtained from the two tests are extremely small, indicating the model residuals are not normal.

Another practical problem when using the parametric method is that the forecasts of the dependent variable usually rely on forecasts of the explanatory variables. For instance, the temperatures used for load forecasting in this study are forecasts provided by meteorological service. In such cases, the exact distribution of the load forecasts depend on the distribution of the explanatory variables and so is extremely difficult to determine [26].

A widely used nonparametric approach is the bootstrap. As a distribution-free method, bootstrapping is robust against violations of the normality assumption and offers a promising method of computing a forecast distribution. The bootstrap approach to the construction of forecasting distributions can be summarized as follows:

- Estimate the forecasting model using historical data.
- Create N artificial sample sets by sequentially substituting the randomly resampled residuals into the estimated model;
- Re-estimate N models based on the artificial sample sets, and obtain N *simulated forecasts* for each forecast horizon by substituting the original data into the re-estimated models;
- Resample another set of residuals, and substitute these into the original estimated model to obtain the *simulated actuals*;
- The differences between the *simulated actuals* and the *simulated forecasts* are the simulated forecast errors, which can be used to construct empirical forecasting distributions, by centering the simulated forecast errors around the point forecasts of the original model.

For the distribution forecasts in this work, the above bootstrap implementations will not be directly applicable due to some practical problems. First, there are 48 different models estimated for each half-hourly period within the day, and up to seven days ahead load forecasts are calculated with multi-period ahead forecasts in an iterative manner. By using the above bootstrap approach, the computational burden will be extremely heavy, making it impractical for an industrial application like STLF, which requires timely updates in a real-time operation. Second, the load forecasting models use temperature forecasts of up to seven days ahead from a meteorological service, while the information regarding the variances of the temperature forecasts are not provided, making it difficult to incorporate this additional source of randomness.

To address the above issues, a modified bootstrap method is proposed in this paper. Instead of bootstrapping residuals in the model learning procedure, we focus on the forecasting residuals in the real application with forecasted temperatures, which contains the uncertainties from both the models and forecasts of exogenous variables. Specifically, the modified bootstrap method can be summarized as follows:

- Estimate the forecasting model using historical data;
- Calculate load forecasts using half-hourly temperature forecasts of up to seven days ahead, and calculate the forecasting residuals between the forecasts and the actual demands;
- Accumulate the forecasting residuals for a contiguous period;
- Bootstrap the forecasting residuals to obtain simulated forecast errors. Block bootstrapping is used since there are correlations between the forecasting residuals from different half-hourly models, and growing variances also result from multi-step ahead forecasts iteratively derived;
- Construct empirical forecasting distributions by centering the simulated forecast errors around the original point

forecasts;
- Continue accumulating the historical temperature forecasts and repeat the above procedures every time the model has been updated.

The revised bootstrap implementation is fast and straightforward, and incorporates the randomness from the model and the exogenous variables. A potential limitation of this bootstrap method is that it requires the accumulation of a substantial number of historical forecasting residuals before reaching a reasonable estimation of the forecasting distribution. In addition, the randomness in the estimated coefficients is not incorporated.

III. Model implementation

A. System and data description

The forecasts presented in this paper are for half-hourly "native demand" in Victoria, Australia, being the demand met by both scheduled and non-scheduled generators supplying the Victorian region of the NEM. Each day is divided into 48 half-hourly periods which correspond with NEM settlement periods.

Victoria is the second most populous state in Australia. Geographically the smallest mainland state, Victoria is the most densely populated state, and has a highly concentrated population of about 5.4 millon in Melbourne, the state capital, and in nearby areas. Half-hourly demand and temperature data were obtained from the AEMO from 1997 to 2009.

Time plots of the half-hourly demand data are illustrated in Figures 1–2, which clearly show the intra-day pattern, the weekly seasonality and the annual seasonality.

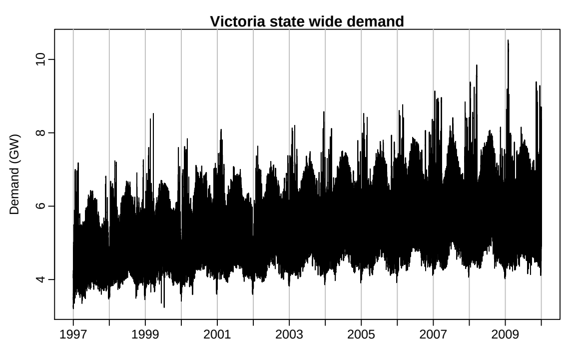

Fig. 1. Half-hourly demand data for Victoria from 1997 to 2009.

Half-hourly temperature data for two different locations in high demand areas (Richmond and Melbourne Airport) have been considered in this study. The relationship between demand and average temperature is shown in Fig. 3, where a non-linear relationship between load and temperature can be observed.

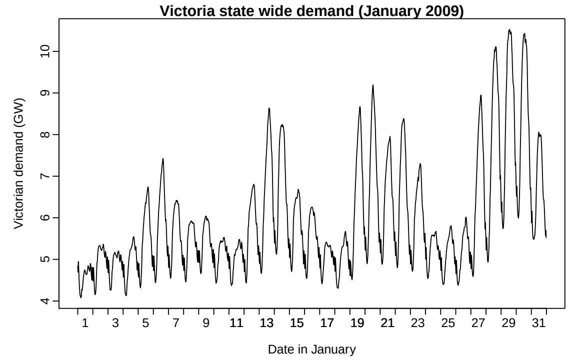

Fig. 2. Half-hourly demand data for Victoria, January 2009.

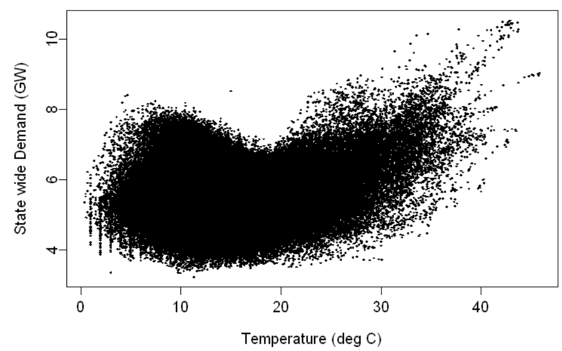

Fig. 3. Half-hourly Victoria electricity demand (excluding major industrial demand) plotted against temperature (degrees Celsius).

B. Variable selection

A highly significant model term does not necessarily translate into good forecasts. Instead, we need to find the best combination of the input variables for producing accurate demand forecasts.

A separate model of the equation (1) for each half-hourly period has been estimated, i.e., 48 half-hourly demand models. For each model, the lagged demands, lagged and future temperatures and calendar variables were selected through a cross-validation procedure [27]. That is, the data have been separated into training and validation sets, and then the input variables are selected by minimizing the accumulated prediction errors for the validation data set. Here the Mean Absolute Percentage Error (MAPE) is used as the selection criterion.

To select the input variables for the half-hourly demand model, we began with the full model including all demand, temperature and calendar variables. The predictive capacity of each variable in the model was tested independently by dropping each term from the model while retaining all other terms. Omitted variables that led to a decrease in MAPE were left out of the model in subsequent tests. Thus, a step-wise variable selection procedure was carried out to identify the best model based on out-of-sample predictive accuracy.

For instance, we first dropped from the model temperature differentials from the same period in the previous six days, one at a time, and located the model with the lowest MAPE as the best model to this stage. Then in the next stage, we froze the terms representing temperature differentials from the same period in previous days, and dropped temperature differentials from the last six hours, one at a time. A similar process was repeated until every group of input variables had been explored. The best model for this application contains the following variables:

- demands around the same time period for the last 2 days;
- the maximum demand in the last 24 hours;
- the minimum demand in the last 24 hours;
- the average demand in the last seven days;
- the current temperature and temperatures from the last half-hour period;
- the current temperature differential;
- the maximum temperature in the last 24 hours;
- the minimum temperature in the last 24 hours;
- the average temperature in the last seven days;
- the day of the week
- the holiday effect
- the day of year effect

IV. FORECASTING RESULTS

According to the Annual Electricity Market Performance Review 2008, published by Australian Energy Market Commission (AEMC) [28], the historical demand forecasting performances in January, during which the highest temperatures of the year are usually observed, are usually the worst of the year due to the extremely high temperatures, indicating a hard case for load forecasting. Therefore, January 2009 has been selected to evaluate the performance of the proposed model. The data used to train the models are from 2004 to 2008. Although historical data back to 1997 are provided by the system operator, we found that inclusion of more training data does not improve the forecasting accuracy. Basically, larger training data sets are effective antidotes to fight against over-fitting. However, increasing the size of the data set by aggregating data way back in the past is not always helpful because the relationships between load and the predictor variables are slowly changing over time. A few missing load and temperature data were filled in by interpolating between neighboring values. The test sets are completely separate from the training sets and are not used for model estimation or variable selection. Based on the selected variables, several hundreds of sub-functions in the proposed additive model have been estimated for the 48 half-hourly periods.

Although the forecasting model is developed to generate forecasts up to seven days ahead, we focus on day-ahead forecasting performance in this paper, as forecasts beyond one day ahead are derived iteratively using the model. Moreover, being a key input for power system operations and market analysis, day-ahead load forecasting is broadly investigated in the field of STLF. In this paper, we use the mean absolute error (MAE) and MAPE to measure the forecasting performance.

An Artificial Neural Network (ANN) based model has been built for comparative study. Specifically, three-layer feed-forward networks have been used to define the functional relationship between the load and the explanatory factors. Following the proposed semi-parametric regression approach, a separate ANN model has been estimated for each half-hourly period. The hyperbolic tangent function in (5) is used for hidden neurons and output neurons since it can produce both positive and negative values, which helps speeding up the training process compared with the logistic function whose output is only positive.

$$f(x) = \tanh(1.5x) = \frac{e^{1.5x} - e^{1.5x}}{e^{1.5x} + e^{1.5x}} \tag{5}$$

The Levenberg-Marquardt approach is used to train the ANN models. This approach is suitable for training medium-size ANN (containing up to one hundred weights) with low Mean Square Error (MSE). Network weights and biases are updated in the batch mode; i.e., weights are updated after the entire training set is presented to the model. The input variables of the ANN have been selected through a similar cross-validation procedure that is used for the regression model.

In addition to the ANN based model, the hybrid model in [16] has also been used to compare the forecasting performance.

Table I shows the MAE and MAPE values for each half-hourly period. According to these results, the proposed model performs generally better than the ANN based model, and also better than the hybrid model with comparable results for some periods. The overall MAE and MAPE values of the proposed models are 0.11 GW and 1.88% respectively, which can be considered to be a satisfactory performance compared with state-of-art load forecasting techniques.

Figure 4 illustrates the load time plots for January 2009. It can be seen that the forecast values (dashed line) follow the actual demand (solid line) remarkably well. The proposed method has also been used to forecast electricity demand in months other than January 2009, and the forecasting performances in the other months are generally better.

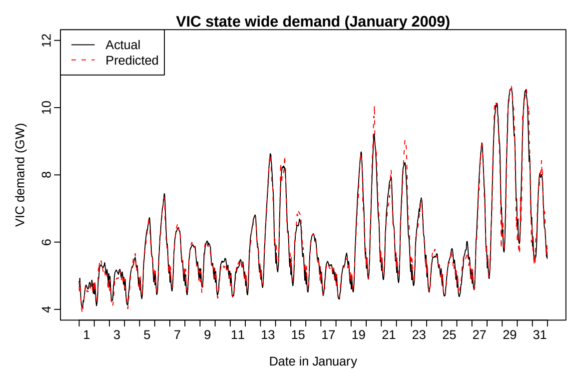

Fig. 4. Actual and predicted demand for January 2009.

Forecasting software based on the proposed model has been developed for AEMO to forecast the short-term electricity demand of Victoria for system operation and planning. The forecasting performance has been evaluated with two different types of predictions: ex ante forecasts and ex post forecasts. Specifically, *ex ante forecasts* are those that are made using only the information that is available in advance, which represents the *real* forecasts. On the other hand, *ex post forecasts* are those that are made using information on the “driver variables” that becomes available during the forecast period; i.e., actual temperatures in this case. The difference between the ex ante forecasts and ex post forecasts allows us to isolate model error from errors in the temperature forecasts.

Figure 5 compares the 24-hour-ahead demand and temperature forecasts (as made at 7am on 11 January) with actual

TABLE I
RESULTS FOR OUT-OF-SAMPLE TEST FOR JANUARY 2009 (UNIT OF INDEX: MAE(MW); MAPE(%))

	Additive model		**ANN**		**Hybrid**	
Period	MAE	MAPE	MAE	MAPE	MAE	MAPE
1	92.92	1.84	96.00	1.90	92.50	1.83
2	101.11	2.07	96.40	1.95	98.27	2.01
3	110.34	2.10	97.20	1.84	104.60	2.00
4	99.52	1.96	96.06	1.88	96.10	1.90
5	83.59	1.72	98.88	2.00	87.42	1.80
6	82.43	1.75	96.78	2.02	90.58	1.88
7	74.94	1.62	97.02	2.05	88.78	1.89
8	74.35	1.61	104.19	2.22	87.79	1.87
9	66.36	1.42	105.17	2.22	81.84	1.71
10	66.93	1.41	120.38	2.47	76.09	1.57
11	78.65	1.56	121.55	2.44	117.60	2.38
12	92.62	1.78	137.13	2.62	122.15	2.36
13	122.05	2.21	154.94	2.80	147.32	2.67
14	130.24	2.25	166.70	2.86	135.00	2.39
15	126.94	2.18	169.53	2.88	124.27	2.17
16	111.58	1.91	162.95	2.71	126.25	2.12
17	121.73	2.01	171.11	2.78	163.39	2.64
18	121.08	1.96	170.06	2.73	173.89	2.66
19	111.88	1.81	169.43	2.69	136.50	2.13
20	114.53	1.83	171.47	2.71	150.77	2.32
21	118.56	1.88	180.30	2.79	153.69	2.32
22	112.78	1.78	187.13	2.89	133.22	2.00
23	121.38	1.89	190.88	2.95	144.11	2.08
24	130.21	2.00	192.41	2.98	175.94	2.48
25	130.52	2.00	207.10	3.20	173.73	2.42
26	124.05	1.91	211.66	3.25	151.13	2.14
27	135.69	2.08	221.22	3.42	145.70	2.22
28	129.31	2.02	229.48	3.52	120.72	1.93
29	123.20	1.93	227.03	3.50	126.54	1.96
30	128.52	1.99	226.86	3.48	127.13	1.97
31	137.81	2.08	227.26	3.45	128.96	1.98
32	147.34	2.23	238.86	3.61	141.83	2.13
33	158.04	2.37	260.73	3.90	164.70	2.45
34	145.24	2.25	249.75	3.77	173.02	2.52
35	142.57	2.23	251.53	3.89	179.63	2.61
36	136.52	2.18	234.60	3.68	173.13	2.53
37	134.06	2.14	227.43	3.64	195.56	2.79
38	138.27	2.21	230.03	3.73	149.12	2.38
39	112.65	1.83	198.92	3.28	160.93	2.46
40	97.70	1.60	186.27	3.06	132.65	2.04
41	89.44	1.50	176.38	2.89	132.38	2.09
42	88.74	1.52	162.43	2.73	118.41	1.81
43	90.08	1.58	150.45	2.60	119.47	1.86
44	89.64	1.64	140.51	2.55	108.79	1.85
45	79.56	1.50	121.52	2.29	111.61	1.94
46	89.12	1.68	126.78	2.41	90.01	1.71
47	91.05	1.68	110.60	2.02	91.66	1.69
48	84.33	1.62	94.95	1.82	81.75	1.58
Average	110.21	1.88	168.04	2.81	126.73	2.14

temperatures and demand. January 11 is selected for illustration because the maximum temperature in this day is above 40 Celsius degree, presenting a hard case for load forecasting and the possible high demand is of concern for the system operator. The upper graph shows the actual demand (solid line) compared with both the ex-ante (dashed line) and ex-post predictions (dotted line). The bottom graph shows the actual temperature (solid line) and forecast temperature (dashed line). It can be seen that temperature forecasts sometimes deviated from the true values (significantly lower beyond 15 hours ahead in this case), although the temperature profiles had reasonably accurate figures for the daily maximum temperature. According to the upper graph, both the ex-ante and ex-post

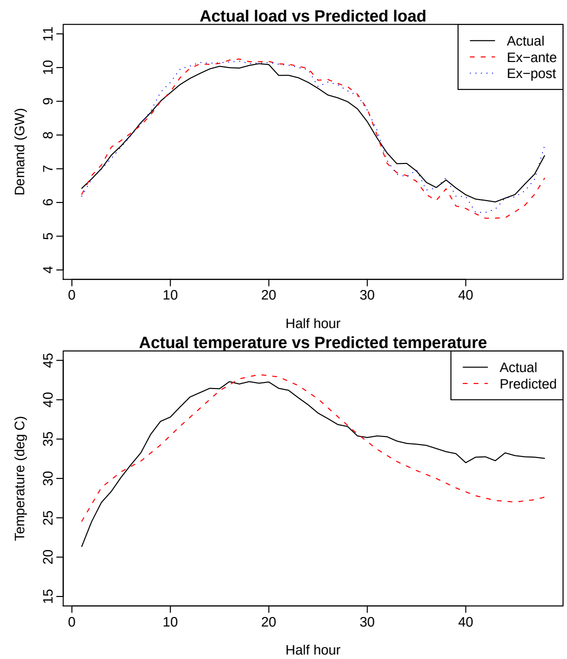

Fig. 5. Actual and predicted demand as made at 7:00am on 11 January 2010.

forecasts follow the actual demand well, providing good evidence of the performance of the model. As we would expect, the load forecasting models are sensitive to the within-day temperature profiles used to drive the model, and the ex post predictions based on actual temperatures give better accuracy. The results indicate that the demand forecasts would have been considerably better if the temperature profile forecasts were more accurate.

The forecast distributions have been evaluated in Figure 6, in which the probability density function of the half-hourly demand was estimated using kernel density estimation [29] for six half-hour periods on 11 January 2010, with the actual demands shown as dot points. In Figure 6, the actual demand values at the six half-hour periods all fall within the region predicted from the forecast distribution, although the actual demand at 4:00 deviates to the right tail due to the large deviation of the temperature forecasts as indicated in Figure 5.

All the numerical studies have been performed using R [30], running on a PC with 4GB of RAM and 2.33GHz clock frequency (duo core CPU). The computation time including variable selection and model training in this case is under 30 minutes.

V. CONCLUSION

The major contribution of this paper is to propose a new statistical methodology to forecast the short-term electricity demand. The proposed additive model allows nonlinear and nonparametric terms within the regression framework, which can capture the complex non-linear relationship between electricity demand and its drivers. The proposed bootstrap method can generate forecasting distribution incorporating the randomness from the model and exogenous variables, suitable for industrial application with acceptable computation burden. The forecasting results demonstrate that the model

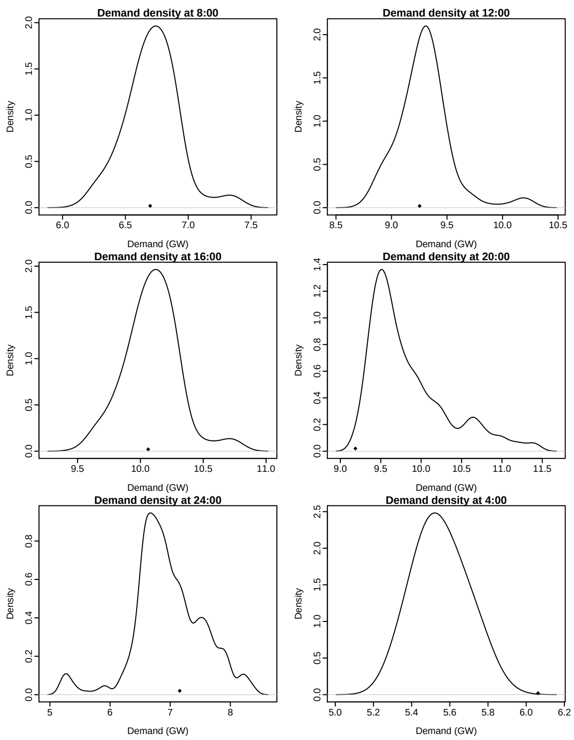

Fig. 6. Actual and predicted demand as made at 7:00am on 11 January 2010.

performs remarkably well on both the historical data and in real-time on-site implementation. The proposed model has been used by AEMO to forecast the short-term loads of two regions in Australian National Electricity Market. In general, the proposed method can be applied for different power systems with appropriate model estimation for each case, and the models can be constructed following the processes described in the paper. Some areas for possible future improvement may include the following.

- Incorporating more temperature sites into the model. For Victoria's geographically concentrated population, two temperature sites may be adequate. But for regions with a widely distributed population, a larger number of temperature sites would be desirable, or multi-region models may be considered.
- Including more drivers in the model to improve the prediction accuracy. For example, humidity could be included if the system is in tropical areas, and operational and market factors may be appropriate in some circumstances.

Acknowledgment

The authors would like to thank the Australian Energy Market Operator (AEMO) for supplying the data used in this paper, and their valuable comments and feedbacks on this research work.

References

[1] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44 – 55, 2001.
[2] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley Finance, 2006.
[3] S. J. Huang and K. R. Shih, "Short-term load forecasting via ARMA model identification including nonGaussian process considerations," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 673–679, 2003.
[4] J. W. Taylor, "Short-term load electricity demand forecasting using double seasonal exponential smoothing," *Journal of the Operational Research Society*, vol. 54, pp. 799–805, 2003.
[5] H. M. Al-Hamadi and S. A. Soliman, "Long-term/mid-term electric load forecasting based on short-term correlation and annual growth," *Electric Power Systems Research*, vol. 74, pp. 353–361, 2005.
[6] P. E. McSharry, S. Bouwman, and G. Bloemhof, "Probabilistic forecast of the magnitude and timing of peak electricity demand," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1166–1172, 2005.
[7] J. Taylor and P. McSharry, "Short-term load forecasting methods: An evaluation based on european data," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2213–2219, 2007.
[8] B. Kermanshahi, "Recurrent neural network for forecasting next 10 years loads of nine Japanese utilities," *Neurocomputing*, vol. 23, pp. 125–133, 1998.
[9] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 887–896, 2006.
[10] ——, "Short-term bus load forecasting of power systems by a new hybrid method," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 333 –341, 2007.
[11] M. S. Kandil, S. M. El-Debeiky, and N. E. Hasanien, "Long-term load forecasting for fast developing utility using a knowledge-based expert system," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 491 – 496, 2002.
[12] K. B. Song, Y. S. Baek, D. H. Hong, and G. Jang, "Short-term load forecasting for the holidays using fuzzy linear regression method," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 96–101, 2005.
[13] V. Ferreira and A. Alves da Silva, "Toward estimating autonomous neural network-based electric load forecasters," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1554–1562, 2007.
[14] Z. Yun, Z. Quan, S. Caixin, L. Shaolan, L. Yuming, and S. Yang, "RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 853 – 858, 2008.
[15] B. J. Chen, M. W. Chang, and C.-J. Lin, "Load forecasting using support vector machines: a study on EUNITE competition 2001," *IEEE Trans. Power Syst.*, vol. 19, no. 4, p. 18211830, 2004.
[16] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 392–401, 2006.
[17] H. Morita, T. Kase, Y. Tamura, and S. Iwamoto, "Interval prediction of annual maximum demand using grey dynamic model," *International Journal of Electrical Power & Energy Systems*, vol. 18, no. 7, pp. 409–413, 1996.
[18] A. Hinojosa, V.H.; Hoese, "Short-term load forecasting using fuzzy inductive reasoning and evolutionary algorithms," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 565 – 574, 2010.
[19] R. J. Hyndman and S. Fan, "Density forecasting for long-term peak electricity demand," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 1142–1153, 2010.
[20] T. Hastie and R. Tibshirani, *Generalized additive models*. London: Chapman & Hall/CRC, 1990.
[21] R. Ramanathan, R. F. Engle, C. W. J. Granger, F. Vahid-Araghi, and C. Brace, "Short-run forecasts of electricity loads and peakss," *International Journal of Forecasting*, vol. 13, pp. 161–174, 1997.
[22] D. Fay, J. V. Ringwood, M. Condon, and M. Kelly, "24-h electrical load data—a sequential or partitioned time series?" *Neurocomputing*, vol. 55, no. 3–4, pp. 469–498, 2003.
[23] G. E. P. Box and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society, Series B*, vol. 26, no. 2, pp. 211–252, 1964.
[24] B. Mizrach, "Forecast comparison in L_2," Rutgers University, Department of Economics, Departmental Working Papers 1995-24, 1995. [Online]. Available: ftp://snde.rutgers.edu/Rutgers/wp/1995-24.pdf
[25] H. Thode Jr., *Testing for Normality*. Marcel Dekker, New York, 2002.
[26] M. Feldstein, "The error of forecast in econometric models when the forecast-period exogenous variables are stochastic," *Econometrica*, vol. 39, p. 5560, 1971.
[27] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40–79, 2010.
[28] AEMC, "Annual electricity market performance review 2008," Australian Energy Market Commission, Final report, 2008.
[29] B. W. Silverman, *Density estimation for statistics and data analysis*. London: Chapman and Hall, 1986.

[30] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2009, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org

Shu Fan (M'08, SM'09) received his B.S., M.S. and Ph.D. degrees in the Department of Electrical Engineering, from China's Huazhong University of Science and Technology (HUST), in 1995, 2000 and 2004 respectively. He conducted postdoctoral research sponsored by the Japanese Government in Osaka Sangyo University from 2004 to 2006. He was a visiting assistant professor at the Energy Systems Research Center at the University of Texas at Arlington from 2006 to 2007. Presently, he is a senior research fellow at Monash University, Australia. His research interests include energy system forecasting, power system control and high-power power electronics.

Rob J Hyndman has a B.Sc.(Hons) and Ph.D. (1993) from the University of Melbourne, Australia. He is Professor of Statistics and Director of the Business and Economic Forecasting Unit at Monash University. He is also Editor-in-Chief of the *International Journal of Forecasting* and author of over 100 research papers and two books. He received the 2007 Moran Medal from the Australian Academy of Science for his contributions to statistical research. His research interests include statistical forecasting, nonparametric modelling and computational statistics.