Analyzing the Impact of Weather Variables on Monthly Electricity Demand

Ching-Lai Hor, *Member, IEEE*, Simon J. Watson, and Shanti Majithia

***Abstract*—The electricity industry is significantly affected by weather conditions both in terms of the operation of the network infrastructure and electricity consumption. Following privatization and deregulation, the electricity industry in the U.K. has become fragmented and central planning has largely disappeared. In order to maximize profits, the margin of supply has decreased and the network is being run closer to capacity in certain areas. Careful planning is required to manage future electricity demand within the framework of this leaner electricity network. There is evidence that the climate in the U.K. is changing with a possible 3°C average annual temperature increase by 2080. This paper investigates the impact of weather variables on monthly electricity demand in England and Wales. A multiple regression model is developed to forecast monthly electricity demand based on weather variables, gross domestic product, and population growth. The average mean absolute percentage error (MAPE) for the worst model is approximately 2.60% in fitting the monthly electricity demand from 1989 to 1995 and approximately 2.69% in the forecasting over the period 1996 to 2003. This error may reflect the nonlinear dependence of demand on temperature at the hot and cold temperature extremes; however, the inclusion of degree days, enthalpy latent days, and relative humidity in the model improves the demand forecast during the summer months.**

***Index Terms*—Climatic variables, forecasting, load pattern, monthly demand, multiple regression.**

I. Climate Change

THERE is evidence that the climate in the United Kingdom is showing long-term changes. The 1990s were the warmest decade of the last millennium in which six of these years had an average temperature above 10°C compared to only two years in the 1970s and 1980s. A total of 26 hot days (with a mean value > 20°C) were recorded in 1995, which was the highest total recorded in 225 years [1]. The average summer temperature for the last three decades has increased from 15.42°C in the 1970s to 15.96°C in the 1990s. This in itself is not necessarily firm evidence for climate change; however, this temperature increase is consistent with scenarios developed by the United Kingdom Climate Impacts Programme UKCIP02 report suggesting a gradual increase of annual temperatures of between $+0.5$°C to $+1.3$°C to be expected by 2020, $+0.8$°C to $+2.2$°C by 2050, and $+1.1$°C to $+3.0$°C by 2080 [2]. This means that very hot summers would be more frequent in the U.K. while winters would become milder.

The multiple regression model described in this paper is shown to accurately predict monthly demand under a wide range of weather conditions. Around 15 years of historical weather data from the U.K. Met. Office and the British Atmospheric Data Centre (BADC) as well as hourly electricity demand data from National Grid Transco have been used in order to develop the demand model. Global Circulation Models (GCMs) can be used to predict long-term seasonal weather patterns and it envisaged that output from these models could be used in conjunction with our regression model to produce future long-term monthly forecasts of electricity demand.

Manuscript received January 12, 2005; revised June 16, 2005. The work was supported by U.K.'s Engineering and Physical Science Research Council (EPSRC) under a joint initiative with the United Kingdom Climate Impacts Programme (UKCIP) entitled "Building Knowledge for a Changing Climate" (BKCC). Paper no. TPWRS-00011-2005.

C.-L. Hor and S. J. Watson are with the Centre for Renewable Energy Systems Technology (CREST), Angela Marmont Renewable Energy Laboratory, Loughborough University, Leicestershire LE11 3TU, U.K. (e-mail: c.hor@lboro.ac.uk; s.j.watson@lboro.ac.uk).

S. Majithia is with the National Grid Control Centre, St. Catherine's Lodge, Berkshire RG41 5BN, U.K. (e-mail: shanti.majithia@uk.ngrid.com).

Digital Object Identifier 10.1109/TPWRS.2005.857397

II. Load Demand Patterns

To establish the correlation between electricity demand and the weather-related parameters described above requires a sufficient period of historical data. In this paper, we consider the period from 1989 to 1995 to establish the parameters in the forecasting model and project forward to 2003 to compare with actual data in order to assess the accuracy of the models.

Initially, hourly demand data for England and Wales is aggregated to give a time series of monthly electricity demand values $E(m, y)$ for each month m in year y

$$E(m, y) = \sum_{d=1}^{N_d} \sum_{h=1}^{24} E^*(h, d) \tag{1}$$

where N_d is the number of days in month m, and $E^*(h, d)$ is the electricity demand in hour h on day d of month m.

The monthly demand values show a strong cyclic pattern reflecting the change in temperature throughout the year [3]. The scale of the demand curve may change depending on the prevailing weather conditions, e.g., may be reduced in the winter during a spell of unusually warm weather. In addition to the weather-related factors, economic parameters can affect the scale of the demand curves, e.g., during an economic boom, the demand may increase [3], [4]. Although more obvious when considering daily demand patterns, these factors will also affect the average monthly demand, and thus, we derive three models that relate the demand to weather and economic parameters as described below.

III. Regression Model

Regression models are often developed based on certain conditions that must be verified for the model to fit the data well

TABLE I
COMPARISON OF STATISTICAL ERROR INDICES

Error Type	ANN	S-E	B&J
ME	10026.79	25861.33	13764.58
MAE	13860.71	38012.25	27455.42
MSE	8.53×10^8	2.84×10^9	1.45×10^9
MAPE	4.77	14.29	10.98
SDE	9801.00	26878.72	19413.91
R^2	0.94	0.89	0.91

ME: mean error, MAE: mean absolute error, MSE: mean square error, MAPE: mean absolute percentage error, SDE: standard deviation of error, R^2: coefficient of determination.

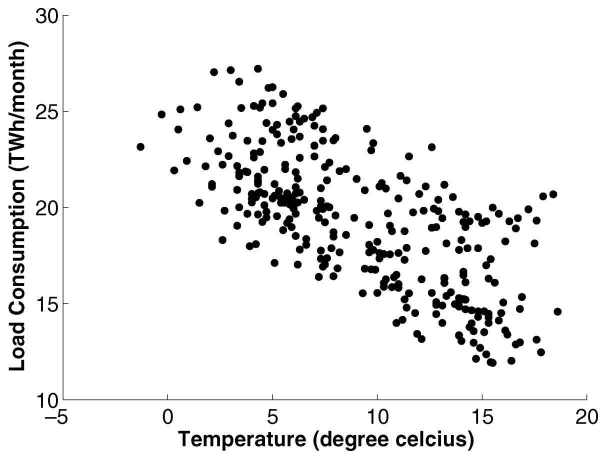


Fig. 1. Mean monthly demand as a function of monthly CET from 1970 to 1995.

and to be able to predict accurately. Parametric multiple regression is preferred in our analysis over neural network [5] or other non-parametric regression approaches [6] because the model is simple to use and easy to control since the inputs can be adjusted for each analysis. Unlike a black box type, e.g., neural network, it allows us to explore the relationship between the weather variables and the load demand. The chosen technique can also deal with random variation, missing observations, and measurement errors.

Table I shows a comparison of statistical error indices for three commonly used models for monthly load forecasting. These are an artificial neural network (ANN), a socioeconomic model (SE), and a Box and Jenkin's model (B&J) [3]. It will be shown later in this paper (see Table V) that our simpler multiple-regression model compares favorably with these models in terms of the error indices.

IV. Relationship Between Demand and Climatic Variables

A. Temperature Demand Relationship

Temperature is the main driving factor for load forecasting [7], [8]. Increased temperatures can also restrict the load carrying capacity of transmission and distribution lines [9]. In order to understand the relationship between temperature and demand, we first plot the average monthly demand values as a function of the so-called monthly Central England Temperature (CET) [10]. This is shown in Fig. 1.

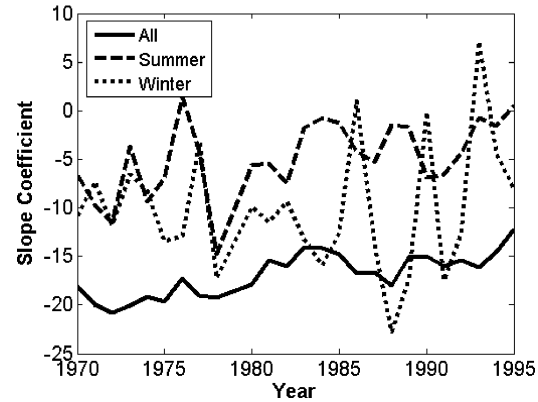


Fig. 2. Variation in the slope coefficient β_1 over the period 1970–1995 during winter, summer, and annually.

It can be seen that there is a strong inverse relationship between demand and temperature. In the winter time, there is a significant lighting and heating load that coincides with the lower temperatures. Conversely, in the summer, demand tends to be lower; however, above a critical temperature ($\sim$ 18°C), consumption tends to rise again due to an increase in cooling and air-conditioning loads [11]. Arguably between 14°C and 17°C, there is a dead zone/balance point in which the electricity consumption can be considered largely unresponsive to the temperature [12]. For mean daily demand, however, this balance point is shifted to 20°C, which can be used as the base temperature for cooling degree days.

Interestingly, there is also evidence for a reduced correlation at the extreme low temperatures where demand seems to saturate below a given temperature. This can probably be explained by the fact that the lighting load is not likely to change much even, at very low temperatures, and that there is a base level of comfort in the winter where an increase in heating is no longer a simple linear function of reducing temperature.

To look in more detail at the relationship between temperature and demand, a simple linear regression fit to the data was made of the form

$$\hat{E} = \beta_0 + \beta_1 T_{mm} \tag{2}$$

where $\hat{E}$ is the predicted electricity demand, β_0 and β_1 are constants fitted to the data, and T_{mm} is the CET mean monthly temperature value. This fit was subdivided by season.

Fig. 2 shows how the slope coefficient β_1 varies over the period 1970 to 1995 for winter and summer and annually. Fig. 3 shows a similar plot for autumn and spring as well as annually. It can be seen that the sensitivity to temperature is more significant during the autumn and spring periods when temperatures are in the mid-range. In the summer and winter periods, when the temperatures are at the extreme ends of the range, the correlation between demand and temperature is rather weaker. This seems to be consistent with the remarks above in relation to Fig. 1. At the same time, the correlation seems to be weakening with time. Part of this may be explained by a shift to gas heating during the period. It may also be that the increase in the use of double-glazing and better thermal insulation standards has reduced losses over the period and also weakened the relationship

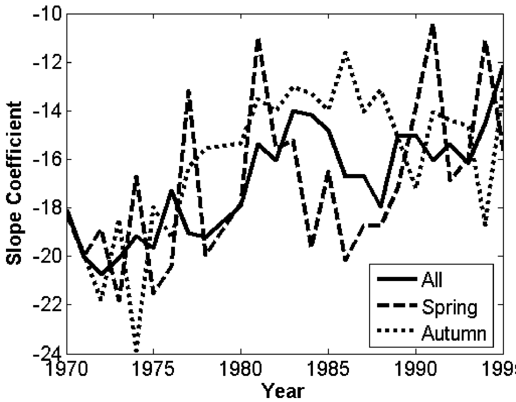


Fig. 3. Variation in the slope coefficient β_1 over the period 1970–1995 during autumn, spring, and annually.

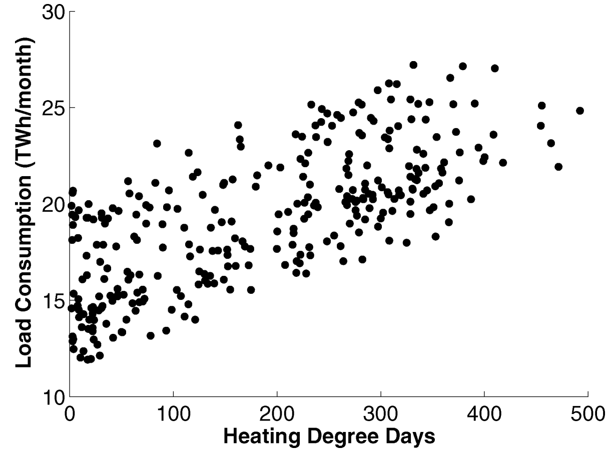


Fig. 4. Monthly electricity demand as a function of heating degree days from 1970 to 1995.

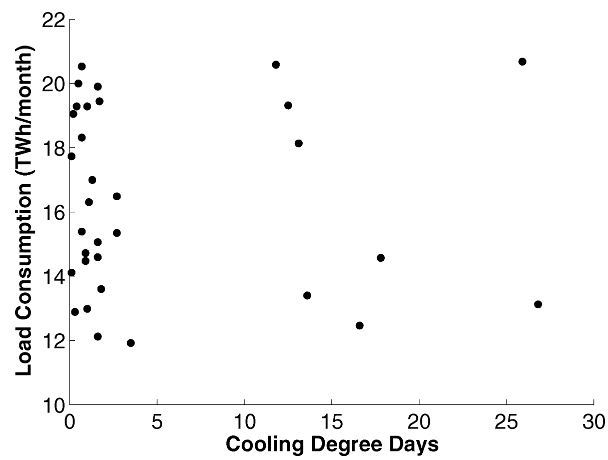


Fig. 5. Monthly electricity demand as a function of cooling degree days from 1970 to 1995.

with temperature. There is still quite a variation in the correlation with temperature over time that may be affected by other nonweather-related parameters, which we will discuss later in this paper.

B. Degree Days

The nonlinear and rather scattered relationship between temperature and load during the summer and winter months may introduce some problems for our multiple regression model. We therefore propose using the quantity known as degree days as more indicative than the temperature-load relationship. The parameter degree days allow us to quantify how cold and hot the weather has been using a single index. Degree days can be split into two categories:

- **Heating Degree Days** (HDD)—a measure of the severity and duration of cold weather used to quantify the heating requirement [13];
- **Cooling Degree Days** (CDD)—a measure of the severity and duration of hot weather used to quantify the cooling requirement [14].

If the HDD value is positive, then the energy is required for heating purposes, whereas if the CDD value is positive, then the energy is required for cooling. The HDD value can also be used to relate the cost of heating with temperature: The higher the HDD value, the colder the temperature, and therefore, more space heating will be required. To calculate the HDD value, a daily mean temperature T_{dm} is estimated based on the average of the minimum and maximum temperatures for that day [15]. If T_{dm} is above an outside dry bulb temperature of 15.5°C acting as a baseline temperature T_{base_H},[1] then HDD is set to 0. If T_{dm} is less than T_{base_H}, then the HDD value can be found by subtracting T_{dm} from T_{base_H} [15]. The HDD value for a month is thus given by

$$\text{HDD} = \sum_{d=1}^{N_d} (1 - \gamma_d) \left(T_{base_H} - T_{dm}\right) \tag{3}$$

[1]The base temperature is a value established from observations, which suggests that heating or cooling systems will not be required to operate. Heating and cooling systems generally have different baseline temperature.

where the indicator function, $\gamma_d = 1$ if $T_{base_H} - T_{dm} < 0$ and $\gamma_d = 0$ if $T_{base_H} - T_{dm} > 0$, and the HDD value for each day d is summed over the total number of days in the month N_d.

Fig. 4 shows a scatter plot of monthly demand data against HDD. The demand shows a strong linear correlation with HDD value, though there is some degree of scatter.

Unlike HDD, there is no official designation for the CDD baseline temperature T_{base_C}. In the UKCIP scientific report, 22°C was used on the basis of building industry energy management practice [2]. However, in our analysis, we use a value of $T_{base_C} = 20°\text{C}$ based on the position of the balance point for the mean daily demand (not shown). The definition of CDD by analogy to HDD is given by

$$\text{CDD} = \sum_{d=1}^{N_d} (\gamma_d) \left(T_{dm} - T_{base_C}\right) \tag{4}$$

where the indicator function is $\gamma_d = 0$ if $T_{dm} - T_{base_C} < 0$ and $\gamma_d = 1$ if $T_{dm} - T_{base_C} > 0$ [15].

Fig. 5 shows that there is a very low correlation between the value of CDD and the electricity demand, though the number of data points is relatively few. It is difficult to assess the cooling

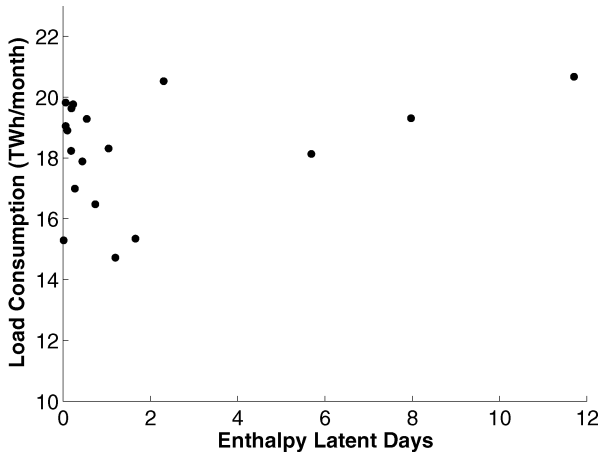


Fig. 6. Energy consumption and enthalpy latent days from 1983 to 1995.

load component on the demand since the daily average temperature in the U.K. seldom goes above 20°C. However, it is likely that the values of CDD will increase over time with the effect of climate change.

C. *Enthalpy Latent Days*

An alternative metric for assessing the cooling load is the quantity known as enthalpy latent days (ELD). The ELD value was established by the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) [15] and is used to account for the possible humidity effects on air-conditioning demand. It indicates the amount of energy required to remove excessive moisture from the outdoor air without reducing the indoor air temperature but lowering the indoor humidity to an acceptable level. The value of ELD over a period of a month can be defined as

$$\mathrm{ELD} = \frac{1}{24}\sum_{d=1}^{N_d}\sum_{h=1}^{24}(\gamma_h)\,(Q - Q_b) \qquad (5)$$

where Q is the hourly value of enthalpy (in kilojoules per kilogram) and Q_b is the enthalpy at the reference temperature of 25.6°C. We have used Q values from the meteorological station at Elmdon, Birmingham, U.K., as representative of England and Wales, being a fairly central location. The humidity ratio [15], [16] can be assumed constant throughout the process because for small changes in temperature, this assumption will give negligible or insignificant errors. The indicator function γ_h takes on a value of 1 if the hourly value of the temperature is above 25.6°C [15] and a value of 0 if the temperature is below this value or if $Q - Q_b < 0$.

Fig. 6 shows a plot of monthly demand as a function of ELD for the period 1983–1995. It can be seen that there are relatively few points in this plot (zero values have been omitted) as there are seldom periods of significant humidity in the U.K. but that there does seem to be some positive correlations between ELD and demand. Clearly further data would be required to define the relationship more precisely.

V. Load Demand Model

A. *Weather-Related Components*

We first propose a component of the forecasting model that includes the effect of the HDD, CDD, and ELD values as well as three other climate-related variables that are likely to have an effect on electricity demand, namely, mean monthly wind speed V_w, mean monthly sunshine hours M_s averaged over a number of sites in England and Wales, and monthly rainfall M_r in millimeters (mm) summed over a number of sites in England and Wales. The relationship of these climatic variables, i.e., degree days, windspeed, enthalpy latent days with respect to the load demand growth, can be treated linearly as suggested by [15].

Temperature is the most influential weather factor on electricity consumption [7]. Increased temperature during the winter will reduce electrical heating load. Conversely, in the summer, electrical heating is normally absent but high temperatures can give rise to a larger refrigeration and air-conditioning load [11]. On the other hand, **windspeed** may have some effects on air-conditioning and cooling fans; however, its principal effect on electricity demand is its effect on electrical heating. The effect of wind on the exterior walls of buildings is also to cool the walls, particularly if they are wet [7]. However, the windspeed is relatively localized. In this analysis, we have used data from a site that is representative of central England and Wales, namely, Elmdon.

Rainfall is very location specific and it affects mainly regional domestic consumption. In conjunction with wind speed, rainfall has also an impact on electrical heating demand. For simplicity, we assume that the relationship is linear.

Wet conditions can also affect the efficiency of air-vented dryers [7] in a similar way through increased **relative humidity**. This means that in periods of high humidity, electricity demand tends to increase [17]. Besides the relative humidity, the increase in cloud cover can increase lighting demand. A more direct indicator of cloud cover is the recorded hours of **sunshine** in a period [7]. Our historic data show the sunshine duration has a negative linear effects on load demand.

Therefore, the predicted electricity demand $\hat{E}_A$ can be formulated as follows:

$$\hat{E}_A = \alpha_0 + \alpha_1\mathrm{CDD} + \alpha_2\mathrm{HDD} + \alpha_3\mathrm{ELD} + \alpha_4 V_w + \alpha_5 M_s + \alpha_6 M_r \qquad (6)$$

where α_n are constants.

If we express the temperature and humidity dependence in the model more directly, the predicted electricity demand $\hat{E}_B$ is now given by

$$\hat{E}_B = \alpha_0' + \alpha_1'\theta_t + \alpha_3' H_r + \alpha_4' V_w + \alpha_5' M_s + \alpha_6' M_r \qquad (7)$$

where θ_t is the monthly CET, H_r is the mean monthly relative humidity at Elmdon, and α_n' are constants.

If we incorporate both the relative humidity and degree day variables, the predicted electricity demand $\hat{E}_C$ is then of the form

$$\hat{E}_C = \alpha_0'' + \alpha_1''\mathrm{CDD} + \alpha_2''\mathrm{HDD} + \alpha_3'' H_r + \alpha_4'' V_w + \alpha_5'' M_s + \alpha_6'' M_r \qquad (8)$$

where α_n'' are constants.

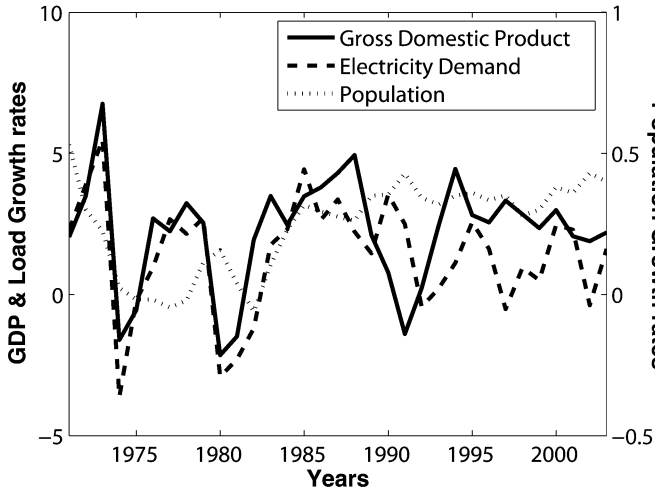


Fig. 7. Comparison between GDP, population growth, and electricity demand over the period 1971 to 2003.

B. Models Including Socioeconomic Factors

The final version of our forecasting model considers non-weather-related socioeconomic variables, namely, Gross Domestic Product (GDP) and population growth λ for England and Wales that are widely known to have an impact on electricity demand.

Fig. 7 shows a comparison between GDP, population, and annual electricity demand growth rates over the period 1971 to 2003. It can be seen that there is a very strong correlation between the electricity demand growth and the GDP growth and the electricity demand growth and the population growth.

We remove the population trend by dividing the monthly data by population interpolated from census data. The per capita consumption is then adjusted in order to isolate the influence of climate factors on electricity consumption. This can be done by first calculating an annual average per capita electricity consumption $\bar{E}(y)$ from the monthly data over the entire period of study [18]. The adjustment factor $F_{adj}(y)$ for each year was calculated from

$$F_{adj}(y) = \bar{E}(y)^{-1} \sum_{m=1}^{12} E(m, y). \tag{9}$$

Each month of energy data was adjusted by dividing it by the adjustment factor for that year [18], i.e.,

$$E_{adj}(m, y) = \frac{E(m, y)}{F_{adj}(y)}. \tag{10}$$

This adjustment was implemented to remove trend effects such as increased use of energy consuming technology. In fact, it inherently contains all necessary adjustment factors, including changes in appliance usage and efficiencies [18].

Though the monthly energy data have been adjusted, there is still a visible trend that could be caused by the GDP. Using the three forms of the weather-dependent components described in (6)–(8), we obtain three forecasting models. Model 1 for the predicted monthly electricity demand $\hat{E}_1$ is therefore of the form

$$\hat{E}_1 = \left(\hat{E}_A + \alpha_7 \mathrm{GDP}\right) \times F_{adj}(y) \tag{11}$$

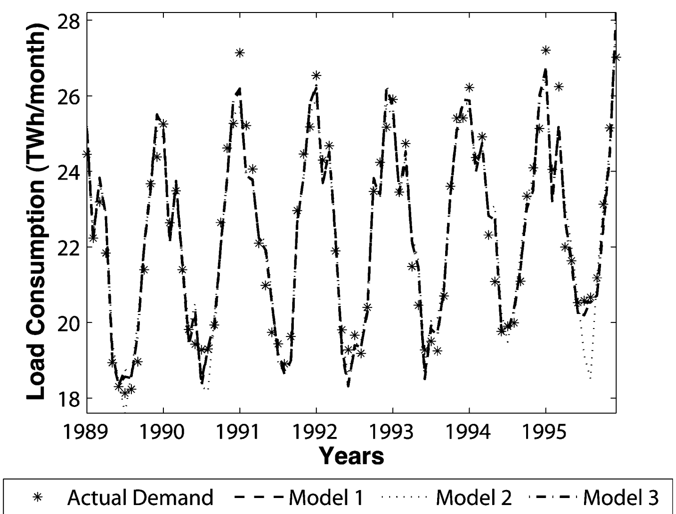


Fig. 8. Comparison between the actual and predicted electricity demand over the period 1989 to 1995.

and Model 2 for the predicted electricity demand $\hat{E}_2$ is given by

$$\hat{E}_2 = \left(\hat{E}_B + \alpha'_7 \mathrm{GDP}\right) \times F_{adj}(y). \tag{12}$$

The constants α_n and α'_n are determined by the best fit to the data in the "training period" 1989 to 1995.

Finally, Model 3 for the predicted electricity demand $\hat{E}_3$ is of the form

$$\hat{E}_3 = \left(\hat{E}_C + \alpha''_7 \mathrm{GDP}\right) \times F_{adj}(y). \tag{13}$$

VI. Model Accuracy

Fig. 8 shows a comparison between the actual and predicted monthly demands for Model 1, Model 2, and Model 3 during the training period 1989 to 1995.

It can be seen that there is a very close correspondence between the actual and predicted values and the difference between the model values is relatively small. If we quantify the overall accuracy of the three models, Model 1 and Model 3 perform slightly better where the average mean absolute percentage error (MAPE) is 2.23% and 2.19% compared to 2.60% for Model 2. In addition to that, Model 1 and 3 have values of goodness of fit equal to 94.5% and 94.6%, respectively. Fig. 9 compares the MAPEs between the three models over the fitting period from 1989 to 1995. From the figure, we can see that Model 1 and Model 3 perform better than Model 2. Because of industrial shutdowns and public holidays, May, June, and December are more problematic months for forecasting.

Looking in more detail at the fit to the data, we analyze the individual regression variables in the three models. Table II shows an analysis of the standard error for each of the parameters in Model 1, and Tables III and IV show the same for Model 2 and Model 3, respectively, following the fit to the period 1989 to 1995. An explanation of the parameters in Tables II and III is as follows.

The **t-ratio** is the ratio of the estimated parameter value to the estimated parameter standard deviation. The larger the ratio, the more significant the parameter is in the regression model [19]. Table II shows that α_2 has the largest ratio while α_3 has the

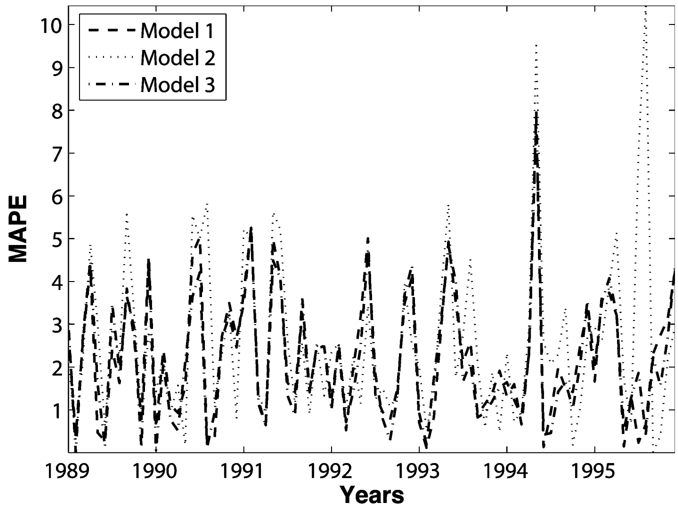


Fig. 9. Comparing the MAPEs between the three models over the period 1989 to 1995.

TABLE II
REGRESSION VARIABLES FOR MODEL 1

Variable	Value	Std. Error	t-ratio	p-value
α_0	19.03748	3.322991	5.729016	0.00000
α_1	0.121177	0.080677	1.502007	0.13724
α_2	0.023205	0.001730	13.41216	0.00000
α_3	*-0.07011*	0.171445	-0.40896	0.68372
α_4	0.273725	0.096235	2.844347	0.00572
α_5	-0.02466	0.003635	-6.78570	0.00000
α_6	*-0.01079*	0.005031	-2.14485	0.03516
α_7	0.094135	0.019642	4.792531	0.00001

smallest value. This suggests that the HDD value is the dominant variable in the model and the ELD value has very little influence on the demand.

The **p-value** is used to test the null hypothesis for each parameter. The smaller the p-value, the less likely the parameter is actually zero [19]. If the p-value is 0.95, there is a 95% chance that the actual parameter has a value of zero and the parameter can usually be removed from the model.

The **Durban–Watson** statistic measures the amount of autocorrelation between the error terms in the model with a scale of 0–4 [20]. A value of 2 means that there is no autocorrelation, while values below suggest positive autocorrelation and values above suggest negative autocorrelation [21].

It is argued that the impact of energy growth will depend on the structure of the economy and the stage of a country's economic development [4]. Our results obviously identify a clear pattern in this direction. In Table II, the negative α_3 coefficient for ELD could be caused by the weak correlation reflected in the relatively small amount of data in Fig. 6. Similarly, in the case of rainfall, we believe the negative α_6, α_6', and α_6'' coefficients are caused by a weak correlation between rainfall and demand at the national level.

If we examine the R^2 values in Table V, we see that Model 1 and Model 3 perform better than Model 2 for the period 1989–1995. The adjusted coefficient of determination Ra^2 is considered a better measure of goodness of fit because it incorporates a tradeoff between the increase in R^2 associated

TABLE III
REGRESSION VARIABLES FOR MODEL 2

Variable	Value	Std. Error	t-ratio	p-value
α_0'	21.43280	7.487790	2.86237	0.00541
α_1'	-0.60548	0.059656	-10.150	0.00000
α_3'	5.696283	6.399557	0.89011	0.37618
α_4'	0.314850	0.148045	2.12673	0.03665
α_5'	-0.01937	0.006085	-3.1828	0.00210
α_6'	*-0.01438*	0.006287	-2.2876	0.02491
α_7'	0.108347	0.022079	4.90722	0.00001

TABLE IV
REGRESSION VARIABLES FOR MODEL 3

Variable	Value	Std. Error	t-ratio	p-value
α_0''	11.53491	6.243164	1.847606	0.06855
α_1''	0.101058	0.036088	2.800326	0.00647
α_2''	0.022458	0.001793	12.52604	0.00000
α_3''	7.691949	5.577971	1.378987	0.17194
α_4''	0.388229	0.126455	3.070090	0.00297
α_5''	-0.01954	0.005173	-3.77664	0.00031
α_6''	*-0.01353*	0.005367	-2.52067	0.01381
α_7''	0.095406	0.019272	4.950638	0.00000

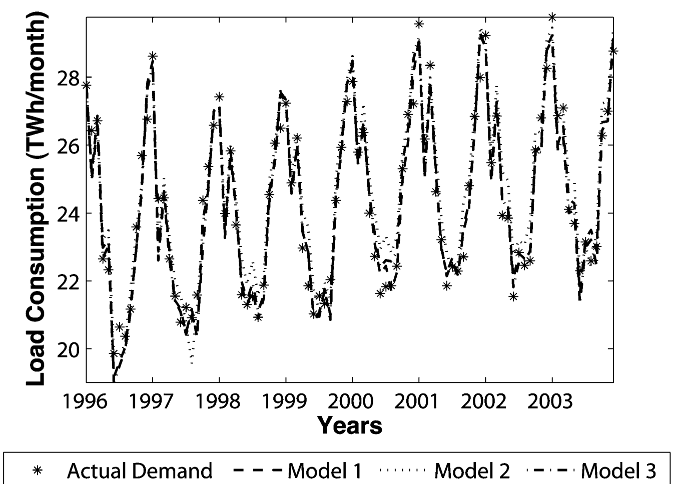


Fig. 10. Comparison between the actual and predicted electricity demand over the period 1996 to 2003.

with an additional parameter and the decreased degrees of freedom. However, the Ra^2 values differ little from R^2 values.

VII. FURTHER VERIFICATION—EIGHT-YEAR FORECAST

To evaluate the forecasting capability of the models, a test was conducted using the three models to forecast the electricity demand in the independent eight-year period from 1996 to 2003. The model output was then compared with the actual demand.

Fig. 10 shows the three model predictions compared with the actual data for the period from 1996 to 2003. Table V and Fig. 11 show once again that Model 1 and Model 3 perform better than Model 2, but now, Model 1 performs slightly better than Model 3 for out of sample prediction. The average MAPE values for the period 1996–2003 for the three models are 1.98, 2.69, and 2.10, respectively.

If we look in more detail at Fig. 10, we see that Model 1 and Model 3 exhibit similar results. This would imply that either ELD or relative humidity can be used in determining summer demand. Model 1 gives a slightly better prediction than Model 3

TABLE V
REGRESSION STATISTICS FOR PREDICTED DEMAND

		1989 - 1995			1996 - 2003		
Regression Statistics		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Sum of residuals	=	1.71×10^{-13}	4.74×10^{-12}	1.78×10^{-12}	3.84	−36.00	5.86
Average residual	=	2.03×10^{-15}	5.65×10^{-14}	2.12×10^{-14}	0.04	−0.375	0.06
Residual sum of squares	=	84.94	114.36	83.02	33.07	57.08	38.43
Standard error of the estimate	=	1.06	1.22	1.05	0.59	0.77	0.63
Proportion of variance explained	=	94.46%	92.54%	94.58%	91.16%	84.73%	89.72%
R^2	=	0.94	0.93	0.95	0.91	0.85	0.90
Adjusted Ra^2	=	0.94	0.92	0.94	0.90	0.83	0.88
Durbin-Watson statistic	=	1.86	1.53	1.80	2.10	1.14	1.86
MAPE	=	**2.23**	**2.60**	**2.19**	**1.98**	**2.69**	**2.10**

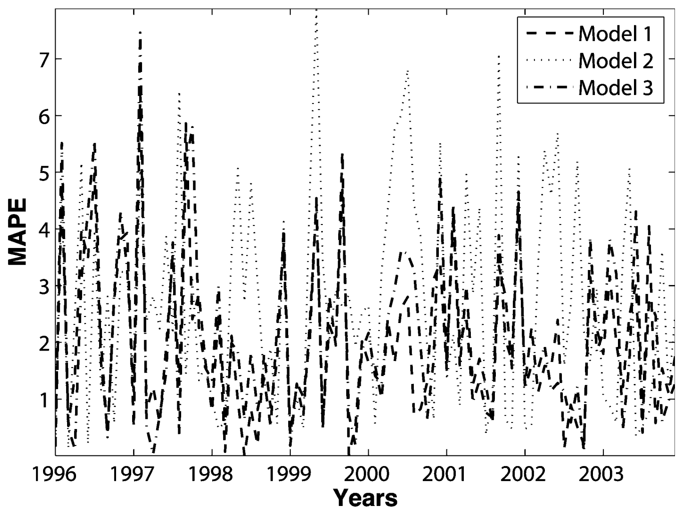


Fig. 11. Comparing the MAPE between the three models over the period 1996 to 2003.

for the summer demand in 2000, though the difference is small. Both Model 1 and Model 3 predict the increase in demand due to the U.K. record-breaking temperatures of August 2003, though the prediction given by Model 1 is slightly higher than Model 3. In general, compared to the other two models, Model 2 does not reflect the actual demand pattern well in the summer. It has largely underestimated the demand in the summer of 1997 and overestimated the demand in the summers of 1998, 2000, and 2002. Based on the results we have obtained, we can conclude that using degree days gives a more accurate forecast than using straight temperature when forecasting electricity demand. On the other hand, it is difficult to assess whether ELD is more efficient than relative humidity in predicting electricity demand during warm summer periods. Intuitively, ELD ought to be more indicative of the demand requirement for reducing humidity to comfortable levels; however, the problem remains that the values of ELD, CDD, and relative humidity tend to be too small to meaningfully derive the effects of these parameters. Most of the demand predictions were underestimated during those periods when the ELD is either small or zero. For example, the demand in August 2003 was unexpectedly less than the demand in June and July 2003 despite it being the hottest month of the three and having the highest ELD value.

The summer of 2003 was one of the hottest periods after 1976 and 1995 in the U.K., and arguably Model 1 and 3 are better able to forecast the summer load demand than Model 2 during positive temperature extremes. However, all three models underestimate the winter demand, particularly in 2001 and 2003. This may be due to the fact that the correlation of the demand with temperature at extreme low temperature is weaker, as suggested by the results in Section IV-A.

VIII. CONCLUSION

We have presented a relatively simple multiple-regression forecasting model including both climate-related and socioeconomic factors that can be used very simply by utility planners to assess long-term monthly electricity patterns using long-term estimates of climate parameters, GDP, and population growth. The model has been trained using monthly electricity demand data for England and Wales over the period 1989 to 1995 and the model accuracy tested against data for the period 1996 to 2003. Three basic models have been proposed using different parameterizations of temperature and humidity. The two models proposed that use the degree day parameterizations of temperature seem to perform better than the third model using a simple temperature variable. The importance of humidity has been suggested when looking at the accuracy of forecasting during the summer months, but due to a lack of data, it is difficult to define both the temperature and humidity dependence of the electricity demand during hotter months. The advantage of our approach is that it is more obvious to understand the relationship between climate-related and socioeconomic variables and electricity demand, unlike alternative "black-box" methods, e.g., neural networks. During the training period of 1989 to 1995, we have a number of extreme months in terms of temperature, namely, the hot summers of 1995 and the warm winter of 1989. Though the coefficients of the regression variables have been trained to take account of these extreme values, the relationship between electricity demand and extremes of temperature remains weak, though the effect of humidity is shown to be significant in the summer. To improve the accuracy of the model forecasts, more historical extreme cases are required. Such data may be obtained from analogs in other countries with larger temperature extremes though care must be exercised as the response may not be identical from country to country. The intention of our analysis was to start with a simple usable model that is able to explain a significant amount of the variability in the electricity demand. We believe our models to be robust to predict electricity demand a number of years into the future given reliable estimates of the relevant climate-related and socioeconomic parameters.

References

[1] M. Hedger, I. Brown, R. Connell, M. Gawith, and T. Downing, "Climate Xhange: Assessing the Impacts—Identifying Responses. Highlights of the First Three Years of the U.K. Climate Impacts Programme. Technical Report," The U.K. Climate Impacts Programme (UKCIP) and DETR, Oxford, U.K., 2000.

[2] M. Hulme, G. Jenkins, X. Lu, J. Turnpenny, T. Mitchell, R. Jones, J. Lowe, J. Murphy, D. Hassell, P. Boorman, R. McDonald, and S. Hill, "Climate Change Scenarios for the United Kingdom—The UKCIP02 Scientific Report. Technical Report," The U.K. Climate Impacts Programme (UKCIP), 2002.

[3] S. Al-Alawi and S. Islam, "Principles of electricity demand forcasting—Part I Methodologies," *Inst. Elect. Eng. Power Eng. J.*, vol. 10, no. 3, pp. 139–143, Jun. 1996.

[4] S. Islam and S. Al-Alawi, "Principles of electricity demand forecasting—Part II Applications," *Inst. Elect. Eng. Power Eng. J.*, vol. 11, no. 2, pp. 91–95, Apr. 1997.

[5] J. Taylor and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 626–632, Aug. 2002.

[6] W. Charytoniuk, M. Chen, and P. Olinda, "Nonparametric regression based short term load forecasting," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 725–730, Aug. 1998.

[7] B. Pitt, "Applications of data mining techniques to electric load profiling," Ph.D. dissertation, The Univ. Manchester Inst. Sci. Technol, (UMIST), Manchester, U.K., 2000.

[8] A. Douglas, A. Breipohl, F. Lee, and R. Adapa, "The impacts of temperature forecast uncertainty on Bayesian load forecasting," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1507–1513, Nov. 1998.

[9] J. Hackney, "Increasing the Value of Weather Information in the Operation of the Electric Power System. Technical Report," Environmental and Societal Impacts Group (ESIG), Nat. Center for Atmos. Res., Boulder, CO, 2002.

[10] (2002) Historical Central England Temperature Data. Data Sets are Available for Bona Fide Academic Research Only. Met. Office. [Online]. Available: http://www.badc.rl.ac.uk/data/cet/.

[11] "UKCIP. London's Warming—The Impacts of Climate Change on London. Technical Report," The U.K. Climate Impacts Programme (UKCIP), 2002.

[12] A. Amato, M. Ruth, P. Kirshen, and J. Horwitz, "Regional energy demand responses to climate change: Methodology and application to the commonwealth of Massachusetts," *Climate Change*, to be published.

[13] *U.K. Government Action Energy Programme (Carbon Trust). GPG310—Degree Days for Energy Management—A practical Introduction*. London, U.K.: New Action Energy, Carbon Trust, 2004.

[14] C. Lam, "Climatic and economic influences on residential electricity consumption," in *Energy Conversion Management*. New York: Elsevier, 1998, vol. 39(7), pp. 623–629.

[15] D. Sailor, "Relating residential and commercial sector electricity loads to climate—Evaluating state level sensitivities and vulnerabilities," in *Energy*. New York: Elsevier, 2001, vol. 26, pp. 645–657.

[16] Y. Huang, R. Ritschard, J. Bull, and L. Chang, "Climatic Indicators for Estimating Residential Heating and Cooling Loads," Lawrence Berkeley Laboratory, Berkeley, CA, Tech. Rep. LBL-21 101, 1986.

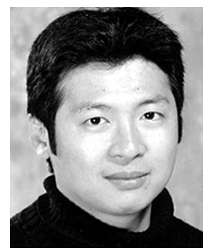
[17] I. Rüdenauer and C. Gensch, "Energy Demand of Tumble Driers with Respect to Differences in Technology and Ambient Conditions Commissioned by: European Committee of Domestic Equipment Manufacturers (CECED)," Inst. Appl. Ecol., Tech. Rep., 2004.

[18] D. Sailor and J. Muñoz, "Sensitivity of electricity and natural gas consumption to climate in the USA—Methodology and results for eight states," in *Energy*. New York: Elsevier, 1997, vol. 22(10), pp. 987–998.

[19] D. Montgomery and G. Runger, *Applied Statistics and Probability for Engineers*, 3rd ed. New York: Wiley, 2003.

[20] P. Berthouex and L. Brown, *Statistics for Environmental Engineers*, 2nd ed. Boca Raton, FL: CRC, 2002.

[21] G. Hewings, S. Changnon, and C. Dridi, "Testing for the Significance of Extreme Weather and Climate Events on State Economies. The Regional Economics Applications Laboratory (REAL) Discussion Paper," Univ. Illinois, (REAL-00-T-6):1–16, 2000.



Ching-Lai Hor (M'01) received the B.Eng (Hons) degree in electrical and electronic engineering from The University of Manchester Institute of Science and Technology, Manchester, U.K., in 1997 and the Ph.D. degree from Queen's University of Belfast, Belfast, U.K., in 2004.

He worked for two years as an Electrical Engineer in ALSTOM Power (formerly known as ABB Power Generation). In 2003, he joined the Centre for Renewable Energy Systems Technology, working on the impacts of climate change on the electricity supply industry.

Dr. Hor is a member of the IEE.

Simon J. Watson received the B.Sc. degree in physics from Imperial College, London, U.K., in 1987 and the Ph.D. degree from Edinburgh University, Edinburgh, U.K., in 1990.

He worked in the field of renewable energy research in conjunction with power systems at the Rutherford Appleton Laboratory, Oxfordshire, U.K., until 1999. He then worked at the green electricity supply company Good Energy, which provides electricity sourced from renewable energy generation to domestic and small commercial customers. In 2001, he was appointed as a Senior Lecturer in the Centre for Renewable Energy Systems Technology, Department of Electronic and Electrical Engineering, Loughborough University, Leicestershire, U.K.

Shanti Majithia is with National Grid Transco, Berkshire, U.K., as a Forecasting Development Manager. He is a Chartered Statistician and has over 30 years of experience in the field of energy forecasting, manpower planning, and load and market research. He has written and presented forecasting papers both in the U.K. and internationally. He has been involved in various projects on air-conditioning demand, climate change, and wind generation. He has recently been successful in applying adaptive logic network techniques for short-term gas forecasting.