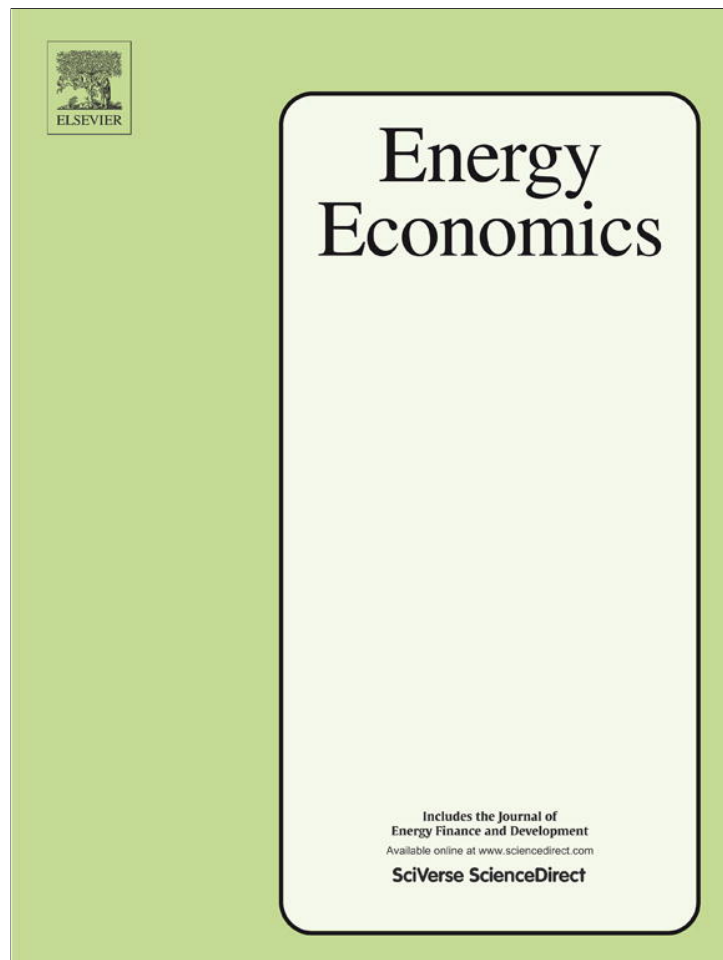


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Global warming and electricity demand in the rapidly growing city of Delhi: A semi-parametric variable coefficient approach

Eshita Gupta

Indian Statistical Institute, Delhi, India

ARTICLE INFO

Article history:

Received 18 July 2011

Received in revised form 22 April 2012

Accepted 25 April 2012

Available online 13 June 2012

JEL classification:

A10

C14

C50

O10

Q41

Q43

Keywords:

Temperature–electricity curve

Variable coefficient model

ABSTRACT

This paper estimates the climate sensitivity of electricity demand in Delhi using daily data on electricity demand and apparent temperature for the period 2000–09. The study adopts a semi-parametric variable coefficient model in order to investigate the impact of climatic factors on electricity demand. As evident from previous studies, electricity demand is a U-shaped function of temperature. We find the rising part of the temperature–electricity curve to become more pronounced over time implying an increase in cooling demand per unit increase in summer temperatures. The study therefore predicts the adverse effects of climate change on electricity demand to be asymmetrically distributed in different seasons in the future, resulting in a serious disequilibrium in the hot months.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

There is growing consensus among scholars on the plausibility of increases in the Earth's mean temperature. It has stimulated attempts to assess the impact of such changes on different sectors. In light of this interest, the present paper attempts to quantify how climate change will affect electricity demand in the continental climate of Delhi ($28^{\circ}30'N$), which is one of the most populous cities in India. For this purpose, we use a semi-parametric variable coefficient approach to estimate the effects of apparent temperature¹ on daily electricity demand over a 10-year period (2000–09). We use the estimated model to simulate the impact of 1 °C, 2 °C and 3 °C increases in apparent temperature on the electricity demand of Delhi up to 2030.

Existing literature highlights a U-shaped non-linear temperature–electricity curve (TEC) where, starting from low levels, rising temperatures first decrease electricity demand due to lower heating requirements in cold weather but where the demand begins to increase due to the higher cooling demand in hot weather once, the level of temperature exceeds the minimum electricity demand threshold. The expected net effect of global warming on electricity demand is therefore ambiguous *prima facie*. Previous studies have shown that the heating effect

dominates the cooling effect in cold countries such as Sweden, which means that global warming would result in a decline in electricity demand in these countries. However, scholars have predicted the reverse for Germany with the cooling effect dominating the heating effect (Bessec and Fouquau, 2008). This suggests that much warmer countries such as India are also likely to experience a net increase in their electricity demand due to climate warming. No studies exist yet of the nature and extent of the climate warming effects on electricity demand in case of India. The present paper attempts such a quantification.

This is the first study to estimate a temperature–electricity curve for India, the key contribution of this paper being that it recognizes and addresses two special problems in the estimation of temperature–electricity curves for developing countries. Firstly, with rapid changes in the economic structure of such countries in future, the relation is likely to shift over time. In this paper, we address this issue by estimating a semi-parametric variable coefficient model that allows the temperature–electricity relation to vary over time. As in Engle et al. (1986), we model the temperature–electricity relation non-parametrically using cubic regression splines, so that weather extremes can have relatively larger impacts on electricity demand, while the other predictor variables enter the regression linearly. The innovation of this paper is to allow the non-parametric temperature–electricity relation to vary across years by interacting the non-parametric component with year. Secondly, black-outs or power-outages are common in many developing countries. This means that observed electricity use is typically less than the notional

E-mail address: egupta13@gmail.com.

¹ 'Apparent temperature' refers to what various combinations of temperature, humidity and windspeed feel like based on human physiology and clothing science and the need for the body to maintain a thermal equilibrium.

electricity demand (which is the object of interest in this study). We adjust daily electricity consumption using daily shortage data in order to obtain the unrestricted electricity demand in Delhi.

One important limitation of this study is that it takes a broad perspective, estimating the average temperature–electricity curve for the aggregate electricity demand of Delhi whereas the temperature–electricity sensitivities may differ across sectors significantly. For instance, while a large chunk of this demand is due to space conditioning and water heating in the residential and commercial sectors, which is highly sensitive to temperature, in agriculture and industrial sectors, electricity demand is determined by the level of economic activity, which is thus largely temperature insensitive. Since, given the data limitations, it is not possible to obtain the daily electricity demand data for different sectors, we have adopted an aggregated approach in this study. In the case of Delhi where 97.5% of the population is urban, the residential and commercial sectors taken together account for approximately 80% of the total electricity demand.

A second limitation is that while electricity demand can be modeled structurally, where electricity consumption is chosen to maximize the expected utility of the households and profits of the firms, we have not adopted this model because a) the data on prices, utilization and efficiency of electricity using equipment at the household and firm level over time is not available and b) a structural model is hard to implement. Thus, as is the case with much of the literature, our study too works with a reduced form model.

Three results from our analysis have important implications for electricity–climate policy: Firstly, we observe that the rising part of the temperature–electricity curve is becoming more pronounced over time, implying an increase in the cooling demand per unit increase in summer temperatures. For instance, a 1 °C increase in temperature at 30 °C increased the electricity demand by over 3 M kWh in 2009 as compared to only over 1 M kWh in 2000. On the other hand, a 1 °C increase in temperature at 15 °C decreased the electricity demand by only 0.8 M kWh in 2009 as compared to 0.7 M kWh in 2000. Secondly, the increasing temperature dependence of the cooling demand shifts the temperature–electricity curve of Delhi leftwards. In particular, the minimum temperature threshold (TT) shifts from about 20–22 °C in the first half (2000–05) to about 18.5–20.5 °C in the second half (2006–09) of the period. Thirdly, the results suggest that the adverse effects of global warming will be asymmetrically distributed in the different seasons. While higher temperatures would increase electricity demand in all seasons except winters, the maximum impact is likely to be felt in the hot month of April, with average apparent temperatures of 30 °C, followed by October and May. Given the dominance of summer electricity demand in the Indian electricity consumption pattern, increasing temperature dependence in the summer months with extreme temperature events may lead to capacity problems.

The rest of the paper is organized as follows. Section 2 explains time-varying temperature–electricity curves. Section 3 reviews existing studies and models that assess the impact of temperature on electricity demand. Section 4 discusses the estimation strategy while Section 5 describes the data sources. In Section 6, we discuss summary statistics and results of the empirical model and in Section 7 we simulate future electricity demand impacts under three different climate scenarios. Section 8 concludes the paper.

2. Understanding the time-varying temperature–electricity curve

Let us consider a hypothetical temperature–electricity curve as represented in Fig. 1. In this U-shaped curve, the minimum point is the threshold. A large number of socio-economic and physical factors such as the growth in incomes, extent of electrification, energy efficiency improvements, cultural habits, and prevailing climatic conditions influence the temperature–electricity curve. Hekkenberg et al. (2009) argue that over time temporal dynamics could influence the slopes as well as the threshold temperature of the temperature–electricity curve. For instance,

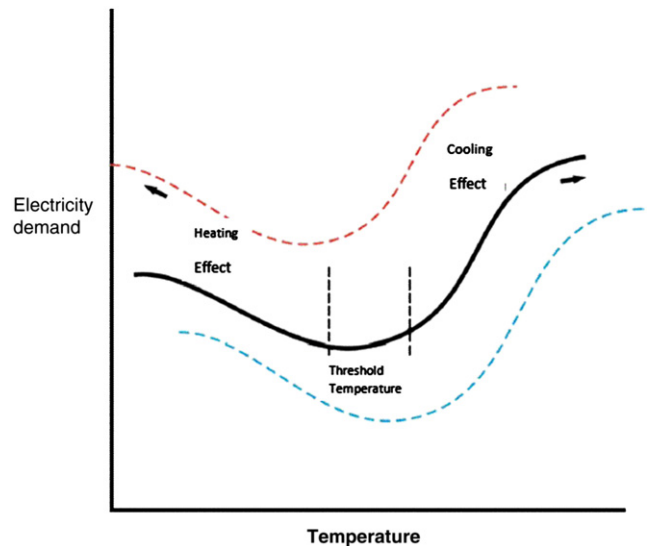


Fig. 1. Temperature–electricity curve.

increased internal heat gains in commercial buildings from an increase in the use of computers, or even a decrease in the tolerance for heat with higher income levels, lead to a general shift towards a lower heating demand and a higher cooling demand. Thus, neglecting a downward shifting threshold temperature results in the underestimation of the electricity demand that arises from a temperature increase. On the other hand, ignoring an upward shifting threshold temperature results in the overestimation of the electricity demand.

The number of households owning temperature control devices (such as air conditioners and air coolers) is increasing very rapidly in India with increasing electricity access and income. According to the National Sample Survey Organization (NSSO) surveys (50th, 61st and 66th), the number of households owning an air cooler² or an air conditioner doubled from 32.9% in 1993 to 60% in 2009 in urban Delhi (which represents 97.5% of the total Delhi population as per the Census 2011) while it increased from 20.6% to 26% in rural Delhi. In the case of refrigerators, the upward trend was even more impressive, with penetration increasing from 29% in 1993 to 61.3% in 2009 in urban Delhi and from 17.7% to 38% in rural areas. In the 2004–05 NSSO survey (which provides data on the ownership of air coolers and air conditioners separately unlike in the previous rounds) only 9% have access to air conditioners and only 58% to air coolers in Urban Delhi. However, with increasing incomes, there is a very high probability that the total air conditioning electricity demand could increase substantially. Further, with increased purchasing power, the sensitivity of households to higher temperatures is likely to increase, which may further shift the location of the minimum point of the temperature–electricity curve. For instance, higher income households may want to switch on their air conditioners when the average temperature is just 19 °C in 2015 as compared to 22 °C in 2000.

According to Kothawale et al. (2010) temperatures (mean, maximum and minimum) increased by about 0.2 °C per decade for the period 1971–2007, with a much steeper increase in minimum temperature than maximum temperature. On a seasonal scale, they observed significant warming trends in mean temperature in two seasons characterized by high humidity: i.e., monsoon and post-monsoon periods. Moreover, increasing night temperatures in these humid seasons could have significant implications for the use of air conditioners and thus for electricity demand. Since the market saturation of air conditioners is currently quite low, the response of its diffusion (along with the rising standard

² Air coolers based on a fan for cooling consumes much less power than air conditioners that operate on the principle of gas compression.

of living) to a long-term increase in the number of hot days and extreme temperature events may play an important role in determining how electricity consumption on the whole would respond to global warming.

3. The temperature–electricity curve: ‘The studies so far’

The simplest way to estimate a U-shaped temperature–electricity curve is to use a regression model that is quadratic in temperature. However, such a model assumes a symmetric relationship because, at any point in the curve, upward and downward changes in temperature of equal magnitude would lead to identical changes in electricity demand. This is an extremely strong assumption and many past studies have shown that the sensitivity of electricity demand to temperature changes depends on initial temperature levels (Valor et al., 2001; Henley and Peirson, 1997). Nonetheless, a linear parametric model can still be used to estimate a non-linear relation by using the degree day approach (Al Zayer and Al Ibrahim, 1996; Mirasgedis et al., 2007; Pardo et al., 2002; Sailor, 2001; Sailor and Pavlova, 2003; Sailor and Munoz, 1997; Valor et al., 2001; Wangpattarapong et al., 2008). This approach defines heating degree days (HDD) and cooling degree days (CDD). CDD and HDD quantify the difference between the daily mean temperatures above or below a threshold temperature (where 18 °C is used as a common threshold temperature), respectively. The HDD index is calculated on the basis of the relation: $HDD = \max(0, 18 - T_d)$, where T_d is the average daily air temperature on day d . The CDD index is calculated on the basis of the relation: $CDD = \max(0, T_d - 18)$. These studies estimated the temperature–electricity curve with the ordinary least squares regression model using annual, monthly or daily data in the following manner:

$$e_d = \beta_0 + \beta_1 TREND_d + \beta_2 CDD_d + \beta_3 HDD_d + \beta_4 CDD_d + \beta_5 HDD_d + \sum_{k=1}^{11} \phi_k MONTH_{kd} + \sum_{b=1}^6 \varphi_b WD_{bd} + \beta_6 HOLIDAY_d + \beta_7 X_d + \varepsilon_d$$

where e is the demand for electricity on day d , WD is a set of week data dummies, $MONTH$ is a set of month dummies, $HOLIDAY$ is dummy for holidays, X includes socio-economic factors such as income and population, and ε is the residual term. Although this approach estimates separate linear relationships of electricity demand due to the heating and cooling demand, it relies on an arbitrary choice of threshold value (18 °C in most cases).

However, more recent studies such as those of Moral-Carcedo and Vicens-Otero (2005) and Bessec and Fouquau (2008) have estimated the above non-linear relationship by obtaining these thresholds endogenously rather than choosing it a priori using different types of non-linear threshold regression models. These studies estimated the above relationship in the following manner:

$$e_d = \beta_0 + \beta_1 TREND_d + \beta_2 (TREND_d)^2 + \beta_3 (TREND_d)^3 + \sum_{b=1}^6 \varphi_b WD_{bd} + \beta_4 HOLIDAY_d + \beta_5 X_d + \beta_6 g(T_d; \gamma, c) + \varepsilon_d$$

where $g(T_d; \gamma, c)$ is a function of the temperature T_d that allows a transition from a cold to a warm regime. In the literature, the transition function has been specified in different ways as piece-wise linear or as a smooth function (exponential or logistic). The assumption of particular functional forms for the transition function is a limitation of such models.

Other researchers have attempted to address this limitation by using non-parametric methods, also known as smoothing models, to achieve greater flexibility in the functional form. To estimate the functional form from data, such models replace global estimates of the electricity–temperature function with local estimates. Local methods estimate a regression between electricity demand (E) and temperature (T) for some restricted range of E and T . This local estimate of the dependency is repeated across the range of E and T and the series of local estimates is then aggregated to summarize the relationship between the two

variables. The resulting non-parametric estimate does not impose a particular functional form on the relationship between E and T , and thus minimizes specification errors (Keele, 2008; Powell, 1994; Ruppert et al., 2003). The estimates are also consistent under more general conditions than are parametric estimates (Wadud et al., 2010; Yatchew, 2003). Both loess and splines are common non-parametric regression models that rely on local estimates to estimate functional forms from data. Engle et al. (1986) estimated the impact of weather on the electricity sales of four US utilities with smoothing splines using monthly data for 7–8 years. The semi-parametric partial linear regression model estimated by them is given by

$$E = Z\gamma + f(T) + \varepsilon.$$

In the above regression, temperature (T) is assumed to affect electricity sales non-linearly by an unknown cubic smoothing spline function f . However, other important variables (Z) such as income and prices enter linearly in the model. The semi-parametric model consists of a conventional parametric and a non-parametric part at the same time. A fully non-parametric model is computationally complex in the presence of numerous predictors. Hyndman and Fan (2008), Harvey and Koopman (1993), and Henley and Peirson (1997) are some studies that use semi-parametric regressions in order to model the temperature–electricity relationship.

Ramesh et al. (1988) is the only study that estimates the temperature–electricity relation for Delhi, assessing the impact of weather variables on the peak electricity load separately for summers and winters, separately during the period 1980–85, using the ordinary least squares parametric regression. However, as mentioned earlier, electricity demand has increased greatly since then. Moreover, while this study investigates the relationship between electricity demand and climatic conditions in Delhi in the past for the purpose of peak demand forecasting, ours is the first study that derives the non-linear dynamic temperature–electricity curve of Delhi and focuses on the time-varying impact of global warming on electricity demand using a semi-parametric variable coefficient model. Not only did the previous study not control for important climatic factors such as rainfall and windspeed, it also did not make any adjustment for the unmet electricity demand.

4. Estimation strategy: ‘The reduced-form model’

We estimate four models in the study. While the first model is based on simple linear regression, the second specifies a semi-parametric additive model using unpenalized splines. The third estimates a semi-parametric additive model with penalized splines while the fourth model is a variable coefficient model where a smooth function of the temperature index is interacted with year to capture the time-varying impact of temperature on electricity demand.

Model 1 estimates the non-linear relationship between electricity demand (E) and apparent temperature (AT) by including a global cubic polynomial in AT in the regression equation. This model takes the following form:

$$e_{td} = \beta_0 + \beta_1 MAJH_{td} + \beta_2 MINH_{td} + \beta_3 RAIN_{td} + \sum_{t=1}^9 \phi_t y_t + \sum_{b=1}^6 \varphi_b WD_{bd} + \beta_4 AT_{td} + \beta_5 AT_{td}^2 + \beta_6 AT_{td}^3 + \varepsilon_{td} \quad (1)$$

where e is electricity demand on day d of year t , $MAJH$ is a dummy variable that takes the value one for the major holiday, and zero otherwise, $MINH$ is a dummy variable that takes value one for the minor holiday, and zero otherwise,³ $RAIN$ represents daily rainfall in

³ A major holiday is one that is declared to be a holiday for all government employees (on account of national events or religious events). Minor holidays are the 2 additional days of holidays that government employees are entitled to select for minor religious festivals from a list of scheduled holidays.

millimeters (mm). WD is the set of six day dummies to describe the weekly periodicity of electricity demand where Wednesday is taken as the reference day. \mathbf{y} is a set of nine-year dummies with 2000 as the base year to identify the deterministic long-term trend connected with the impact of demographic, technological, and socio-economic factors such as prices, urbanization, and the increasing number of air conditioners and air coolers on electricity demand. The inclusion of year-fixed effects accounts for any fixed differences across years that may be correlated with all unobservable factors. In matrix notation Eq. (1) can be rewritten in the following form

$$\mathbf{E} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{T}\boldsymbol{\eta} + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

where, \mathbf{E} is an $n \times 1$ vector of electricity demands, $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of errors, and \mathbf{Z} is an $n \times p_1$ matrix of p_1 non-temperature predictors, $\boldsymbol{\gamma}$ is an $p_1 \times 1$ vector of coefficients of predictors in \mathbf{Z} , \mathbf{T} is an $n \times p_2$ matrix of AT temperature predictors, $\boldsymbol{\eta}$ is an $p_2 \times 1$ vector of coefficients of predictors in \mathbf{T} , \mathbf{X} is an $n \times p$ ($= p_1 + p_2$) matrix of all predictors and $\boldsymbol{\beta}$ is an $p \times 1$ vector of coefficients of \mathbf{X} predictors. The least squares and maximum likelihood estimator of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{E}}$ and *Hat matrix* \mathbf{H} is a $n \times n$ matrix, such that $\hat{\mathbf{E}} = \mathbf{H}\mathbf{E}$. We can obtain $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ and show that $trace(\mathbf{H}) = trace(\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T) = tr(I_p) = p =$ estimated degrees of freedom (EDF) as measured by the number of parameters in the model. This model assumes that the relationship between E and AT is strictly cubic regardless of whether this is true or not. When it is not, the power transformations often cannot adequately capture the nonlinear relationship in the data.

Model 2 estimates a semi-parametric model given by

$$\mathbf{E} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{f}(AT) + \boldsymbol{\varepsilon}. \quad (3)$$

Here, $\mathbf{f}(AT) = (f(AT_1), \dots, f(AT_n))'$ is a $n \times 1$ vector, where $f(AT)$ is an unknown smooth function, i.e., continuous and sufficiently differentiable function of AT . In this paper, we estimate $f(AT)$ by cubic regression splines⁴ using cardinal basis functions.⁵ Such basis functions parameterize the spline in terms of its values at the knots and thus have advantages in terms of the interpretability of the parameters along with good mathematical properties and numerical stability. $f(AT)$ can be represented as a linear combination of the basis functions of regression splines. For instance,

$$f(AT_i) = \sum_{j=1}^N b_j(AT_i)\eta_j = \mathbf{B}(AT_i)\boldsymbol{\eta} \quad (4)$$

where $b_j(AT)$ is the basis at the j th point (commonly known as a knot), $\mathbf{B}(AT)$ is the model matrix containing N cubic spline basis for $f(AT)$ and $\boldsymbol{\eta}$ is the corresponding regression parameter vector. Thus Eq. (3) becomes

$$\mathbf{E} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{B}(AT)\boldsymbol{\eta} + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (5)$$

where \mathbf{X} is an $n \times (p_1 + (N - 1))$ model matrix. One degree of freedom is lost due to the identification constraint on $\mathbf{f}(AT)$ i.e. $\sum_{i=1}^n f(AT_i) = 0$. Using

⁴ Suppose there is a knot sequence $K, AT_{min} = k_1 < \dots < k_N = AT_{max}$, where k_2, \dots, k_{N-1} are interior knots, and k_1, k_N are two knots at the boundaries of the data $[AT_{min}, AT_{max}]$, dividing the data into $N - 1$ subintervals $[k_1, k_2] \dots [k_{N-1}, k_N]$. A spline is a piecewise-polynomial real function: $f : [AT_{min}, AT_{max}] \rightarrow \mathbb{R}$ on an interval $[AT_{min}, AT_{max}]$ composed of $N - 1$ ordered disjoint subintervals $[k_1, k_2] \dots [k_{N-1}, k_N]$. The restriction of f to an interval i is a polynomial P_i

$P_i : [k_{i-1}, k_i] \rightarrow \mathbb{R}$

⁵ For full details of cardinal basis functions, see Wood (2006) and Lancaster and Salkauskas (1986).

the Akaike Information Criterion⁶, we select twelve knots ($N = 12$) or eleven basis functions. Given knots, this model becomes a fully parametric model with an expanded model matrix. We estimate predictor variable coefficients by minimizing $\|\mathbf{E} - \mathbf{X}\boldsymbol{\beta}\|^2$. The key limitation of this model, however, is that it requires the analyst to select the number and location of the knots. The number of knots directly controls the degrees of freedom of a smooth term. In order to deal with the knot selection problem, we adopt the penalized cubic spline approach. These models construct a penalty on $f(\cdot)$ which will be large if f is very wiggly and small if it is nearly flat.

Model 3 adds a quadratic penalty as $\lambda\boldsymbol{\beta}^T\mathbf{P}\boldsymbol{\beta}$ and solves the following minimization problem:

$$\|\mathbf{E} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\boldsymbol{\beta}^T\mathbf{P}\boldsymbol{\beta} \quad (6)$$

where \mathbf{P} is the penalty matrix whose coefficients depend on the second derivatives of f , a measure used commonly to represent the roughness of the smooth terms.⁷ λ is the smoothing parameter that controls the trade-off between model fit and model smoothness. For $\lambda \rightarrow 0$ the minimization gives a wiggly function whereas letting $\lambda \rightarrow \infty$ gives a linear fit. The optimal λ is selected by cross validation where it works as follows: for a given value of λ , we omit the i th observation from data and fit the penalized spline to this slightly truncated data set. We denote this prediction of e_i as \hat{e}_{-i} . The model prediction errors are calculated, and this is repeated as each observation is dropped in turn. The cross-validation score is calculated as the average of the individual model prediction errors. One should choose the value of λ with the smallest cross-validation score. In practical applications one replaces the cross validation (CV) criteria by the generalized cross validation (GCV) as the CV is computationally very intensive and has other problems (Woods, 2006). Like the adjusted R-square, the GCV adjusts the average model prediction errors with the degrees of freedom (the number of parameters estimated in the model). For penalized spline models, the GCV score is

$$GCV(\lambda) = \frac{\sum_{i=1}^n [e_i - \hat{e}_{-i}]^2}{[n - tr(H_\lambda)]^2} \quad (7)$$

Minimizing $GCV(\lambda)$ with respect to λ gives an estimate $\hat{\lambda}$. Given λ , Eq. (6) is minimized w.r.t $\boldsymbol{\beta}$. We get $\hat{\boldsymbol{\beta}} = [\mathbf{X}^T\mathbf{X} + \lambda\mathbf{P}]^{-1}\mathbf{X}^T\mathbf{E}$ and the hat matrix $\mathbf{H}_\lambda = \mathbf{X}[\mathbf{X}^T\mathbf{X} + \lambda\mathbf{P}]^{-1}\mathbf{X}^T$. The trace of \mathbf{H}_λ , as in the linear regression, represents the degrees of freedom in the spline model and is nearly equivalent to the number of parameters in the spline fit. Due to shrinkage from the penalty term, the degrees of freedom for a penalized spline model will not be an integer. With penalized splines the exact choice of the basis dimension is not generally critical as actual effective degrees of freedom are controlled by λ . It is necessary to select the number of knots to be large enough to have enough degrees of freedom to represent the underlying true structure of the data reasonably well but small enough to maintain reasonable computational efficiency (Woods, 2006).

Model 4 extends Model 3 to a variable coefficient model (See Hastie and Tibshirani (1993)) to capture the time-varying impact of climate on electricity demand. Model 4a estimates the least

⁶ The Akaike Information Criterion is a measure of the relative goodness of fit of a statistical model: $AIC = 2k - 2\ln(L)$, where k is the number of parameters in the model, and L is the maximized value of the likelihood function. In practice, it is customary to use the Akaike Information Criterion in order to select the optimal number of knots. Researchers prefer a model with a lower value of the Akaike Information Criterion. In practice, it is customary to use the Akaike Information Criterion in order to select the optimal number of knots. Researchers prefer a model with a lower value of the Akaike Information Criterion.

⁷ Wood and Augustine (2002) derive the following wiggleness measure

$$j(f) = \int [f''(AT)]^2 dAT$$

$$f''(AT) = \sum_{j=1}^N b''_j(AT)\eta_j = \mathbf{B}''(AT)\boldsymbol{\eta}$$

constrained factor variable coefficient model by interacting $f(AT)$ by ten year dummies. It estimates a different smooth function of temperature for each year. Model 4b estimates a simple numeric variable coefficient model by adding an additional term in Model 3 that interacts $f(AT)$ with the year number. It assumes that the coefficients of the smooth function of temperature change linearly with the year. One can capture the time-varying effect by estimating a separate model (like Model 3) for each year. However, by pooling data for all 10 years we get more robust estimates that are preferable for the purpose of analyzing the long-term impact of climate on electricity demand.

Model 4a: We first select the number of knots for each year (N_t) and corresponding basis functions $\mathbf{B}_t(AT)$

$$f_t(AT) = \sum_{j=1}^{N_t} b_{jt}(AT)\eta_{jt} = \mathbf{B}_t(AT)\boldsymbol{\eta}_t. \quad (8)$$

In general, knots are placed at evenly spaced quantiles of the unique data. We select the same 10 knots every year. The selected knots are $[k_0 = 6.17, 14.39, 18.47, 22.67, 26.84, 30.98, 33.87, 35.78, 37.59, 43.83 = k_{10}] \forall t$. $\mathbf{B}_t(AT)$ is a row vector of basis functions for year t . $\boldsymbol{\eta}_t$ is the coefficient vector of the basis functions of year t . The model becomes:

$$\mathbf{E} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{f}(AT)\mathbf{Y} + \boldsymbol{\varepsilon} = \mathbf{Z}\boldsymbol{\gamma} + \sum_{t=1}^{10} \mathbf{f}(AT)\mathbf{y}_t + \boldsymbol{\varepsilon} = \mathbf{Z}\boldsymbol{\gamma} + \sum_{t=1}^{10} \mathbf{f}_t(AT) + \boldsymbol{\varepsilon} \quad (9)$$

where, $\mathbf{f}_t(AT)$ is a vector of the smooth function of the temperature index of year t with dimension $n \times 1$. Here, t indexes the year with $t = 1$ for year 2000 and $t = 10$ for year 2009. \mathbf{Y} is an $n \times 10$ matrix of year dummies. \mathbf{y}_t is the t th column of \mathbf{Y} . The \mathbf{y}_t represents the year dummy for year t . The degrees of freedom for $\mathbf{f}_t(AT)$ will be determined by the choice of λ_t . Note that the same λ_t is chosen for all years resulting in the same degrees of freedom for each year. Thus, the fitting problem becomes similar to any other generalized additive models:

$$\min \|\mathbf{E} - \mathbf{X}\boldsymbol{\beta}\|^2 + \sum_t \lambda_t \boldsymbol{\beta}^T \mathbf{P}_t \boldsymbol{\beta} \quad (10)$$

where \mathbf{X} is a $n \times (p_1 + ((N - 1) \times 10))$ model matrix. Given λ_t , Eq. (10) can be minimized w.r.t $\boldsymbol{\beta}$. We get $\boldsymbol{\beta} = [\mathbf{X}^T \mathbf{X} + \sum_t \lambda_t \mathbf{P}_t]^{-1} \mathbf{X}^T \mathbf{E} = [\mathbf{X}^T \mathbf{X} + \mathbf{K}]^{-1} \mathbf{X}^T \mathbf{E}$, with $\sum_t \lambda_t \mathbf{P}_t = \mathbf{K}$. A smoother matrix for penalized splines with interaction can be derived as $\mathbf{H}_\lambda = \mathbf{X}[\mathbf{X}^T \mathbf{X} + \sum_t \lambda_t \mathbf{P}_t]^{-1} \mathbf{X}^T = \mathbf{X}[\mathbf{X}^T \mathbf{X} + \mathbf{K}]^{-1} \mathbf{X}^T$. As discussed previously, one degree of freedom is lost due to the identification constraint on $\mathbf{f}_t(AT)$, which requires $\sum_{i=1}^N \mathbf{f}_t(AT_i) = 0 \forall t$. From the above, we obtain the electricity demand on a particular day

$$e_{td} = \mathbf{z}'_{td} \boldsymbol{\gamma} + f_t(AT_{td}) + \varepsilon_{td} \quad (11)$$

where e_{td} is electricity demand on day d of year t . \mathbf{z}'_{td} is a row vector of parametric predictors for day d of year t . We can write the full form of Eq. (11) therefore as

$$e_{td} = \beta_0 + \beta_1 MAJH_{td} + \beta_2 MINH_{td} + \sum_{t=1}^9 \phi_t y_t + \sum_{b=1}^6 \varphi_b WD_{td}^b + \beta_3 RAIN_{td} + f_t(AT_{td}) + \varepsilon_{td}. \quad (12)$$

As the errors from Eq. (12) are likely to be serially correlated, we carry out the following adjustment given in Li and Racine (2007).⁸

By dropping year dummies and estimating Eq. (12) separately for each year, we obtain $\hat{\varepsilon}_d$ for each t . For each year t , a first order stationary auto-regressive model is defined as

$$\varepsilon_d = \rho_t \varepsilon_{(d-1)} + \nu_d \quad (13)$$

where ν_d is white noise, is estimated. By regressing $\hat{\varepsilon}_d$ on $\hat{\varepsilon}_{d-1}$ of year t , we obtain an estimate of ρ_t ($\hat{\rho}_t$). The model is then transformed in order to have serially uncorrelated disturbances by subtracting estimated previous day errors $\hat{\varepsilon}_{d-1}$ from the electricity demand on a given day e_d in the following manner:

$$e_d^* = e_d - \hat{\rho}_t \hat{\varepsilon}_{d-1}. \quad (14)$$

By pooling estimated transformed electricity demand e_d for each t , the final model becomes

$$e_{td}^* = \beta_0 + \beta_1 MAJH_{td} + \beta_2 MINH_{td} + \beta_3 RAIN_{td} + \sum_{t=1}^9 \phi_t y_t + \sum_{b=1}^6 \varphi_b WD_{td}^b + f_t(AT_{td}) + u_{td} \quad (15)$$

where u_{td} are serially uncorrelated disturbances and we get consistent estimates of the coefficients.⁹

Model 4b on the other hand interacts $f(AT)$ with year as a numeric rather than as a factor variable (as was the case in Model 4a) and thus assumes that the coefficients of the smooth of temperature change linearly with year. The model becomes:

$$\mathbf{E} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{f}(AT) + \mathbf{Y}_0 \mathbf{f}(AT) + \boldsymbol{\varepsilon}. \quad (16)$$

Here, \mathbf{Y}_0 is a $n \times n$ diagonal matrix with year numbers $\mathbf{y}_0 = (1, 2, 3, \dots, 10)$ on its leading diagonal. In this model each row of the model matrix of $\mathbf{f}(AT)$ is multiplied by the corresponding value of the year number. The fitting problem as in any other generalized additive model is

$$\min \|\mathbf{E} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda_1 \boldsymbol{\beta}^T \mathbf{P}_1 \boldsymbol{\beta} + \lambda_2 \boldsymbol{\beta}^T \mathbf{P}_2 \boldsymbol{\beta} \quad (17)$$

where λ_1 and \mathbf{P}_1 correspond to $\mathbf{f}(AT)$ and λ_2 and \mathbf{P}_2 correspond to $\mathbf{Y}_0 \mathbf{f}(AT)$. Thus this model is also estimated as a generalized additive model as discussed in detail above. The selected knots for this model are the same as in Model 4a [$k_0 = 6.17, 14.39, 18.47, 22.67, 26.84, 30.98, 33.87, 35.78, 37.59, 43.83 = k_{10}$]. Electricity demand on a particular day is obtained as

$$e_d^* = \beta_0 + \beta_1 TREND_y + \beta_2 MAJH_d + \beta_3 MINH_d + \beta_4 RAIN_d + \sum_{b=1}^6 \varphi_b WD_d^b + f(AT_d) + f(AT_d) \times y_o + u_d. \quad (18)$$

In this model, we replace the year dummies with linear year trend $TREND_y$, which ranges from 1 for the year 2000 to 10 for 2009. Further, instead of correcting for within year first order autocorrelation as was the case in Model 4a, we correct for the first order autocorrelation after pooling all the years for Model 4b.

5. Data

5.1. Electricity consumption and shortage

We obtained the data on daily electricity consumption in Delhi for 2000–09 from the operator of the national electricity grid, the National Load Dispatch Centre (NLDC). In order to estimate the impact of global warming on electricity demand, we need to recognize that the electricity systems in India are continually inhibited with power shortages, which result in rationing and disrupted electricity usage

⁸ For more details, refer Li and Racine (2007) in chapter 18, section 18.2.2.

⁹ In this process we lose one observation per year and thus the total number of observations falls to 3643 from 3653.

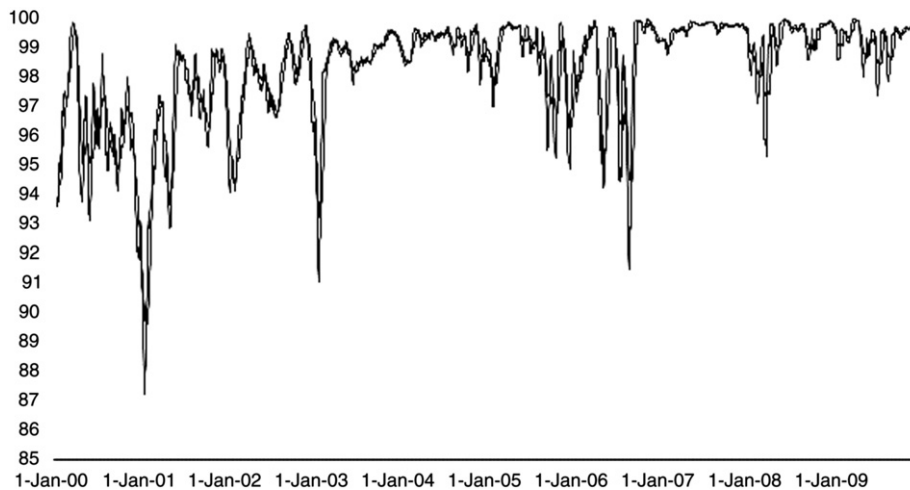


Fig. 2. Daily reliability index of electricity supply (%) of Delhi (2000–2009).

patterns. When there are regular power failures, consumers are not able to consume the quantity they need forcing them thereby to either substitute electricity with alternative energy sources such as diesel and kerosene, or resort to independent generation. As a result, the electricity consumption reported by the NLDC is constrained electricity demand, which is equal to the electricity supplied by the utilities. In order to obtain the unrestricted electricity demand for Delhi, we adjust the daily total electricity consumption of Delhi with the observed daily shortage¹⁰ using daily electricity supply shortage data obtained directly from the Delhi Transco Ltd. This gives us

$$e_d = c_d + s_d$$

where e_d denotes electricity demand on day d , c_d denotes electricity consumption on a given day d and s_d is the shortage (or the unmet demand) on day d . Fig. 2 plots the reliability index of electricity (i.e., the total electricity demand met as a percentage of the total demand including the shortage). The graph shows that there has been a significant reduction in shortages in the post-2005 period.

5.2. Apparent temperature and rainfall

We obtained data on all the climatic factors from the website www.tutiempo.net/en/climate/India, which gives station-wise data for all the major weather stations in India. We first constructed the apparent temperature index (AT) for Delhi (Safdarjung station) using Steadman's (1994) formula by adjusting dry bulb temperature with humidity and wind speed, which is given below:

$$AT_{td}(^{\circ}C) = T_{td} + 0.33v_{td} - 0.07w_{td} - 4$$

$$v_{td} = \frac{h_{td}}{100} \times 6.105 + e^{\left(\frac{17.27T}{237.7 + T}\right)}$$

where T denotes average temperature in degree Celsius ($^{\circ}C$), v denotes evaporation, w denotes wind speed (m/s), and h denotes relative humidity (%).

¹⁰ The daily shortage data comprises five components: shedding due to transmission and distribution constraints; shedding by discoms in theft-prone areas; shedding in order to restrict over drawal; shedding due to grid constraints; and shedding in order to restrict under-frequency operations.

6. Results: 'The effect of apparent temperature on electricity demand'

6.1. Summary statistics

Table 1 displays the basic summary statistics that will be used to analyze the salient characteristics of the distribution of electricity demand and apparent temperature in the city of Delhi. Over the period, the average daily electricity demand (ED) increased from 50 M kWh in 2000 to 65 M kWh in 2009, with the highest daily demand increasing from 65 M kWh in 2000 to 94 M kWh in 2009. At the same time, the standard deviation of the daily electricity demand increased from 6 M kWh in 2000 to 15 M kWh in 2009. During this period, the average daily apparent temperature ranged from 26.5 and 27.7, with the peak occurring in 2009 and 2002 and the trough occurring in 2005. Fig. 3a presents the box plot of daily electricity demand and daily apparent temperature by months. It shows that the greatest consumption occurs during the summer (led by May) and monsoon months (led by July) and the lowest consumption occurs in the post-monsoon (led by November) and winter months. Fig. 3b presents a box plot of daily electricity demand by week days in order to examine the variation in electricity demand according to week days.

Table 1
Summary statistics for electricity demand (E) and apparent temperature (AT) 2000–2009.

	Variable	No of days	Mean	Standard deviation	Min	Max
2000	E	366	49.77	6.16	32.52	64.64
	AT		27.23	9.45	7.62	42.53
2001	E	365	51.57	6.90	36.20	64.34
	AT		27.24	9.05	9.06	40.82
2002	E	365	54.47	8.02	39.47	70.47
	AT		27.70	9.25	9.33	42.53
2003	E	365	55.06	7.59	38.58	72.21
	AT		27.07	9.04	8.71	40.90
2004	E	366	57.73	8.50	39.80	74.66
	AT		27.37	8.79	7.67	39.91
2005	E	365	58.78	8.94	40.55	77.88
	AT		26.45	9.08	8.61	43.45
2006	E	365	61.44	10.61	40.01	83.26
	AT		27.30	8.42	6.17	40.27
2007	E	365	61.35	11.26	39.72	85.30
	AT		27.27	9.35	7.15	43.83
2008	E	366	62.05	11.16	41.75	84.62
	AT		27.28	8.89	6.97	40.53
2009	E	365	65.31	14.70	38.41	94.31
	AT		27.69	8.81	11.06	42.54

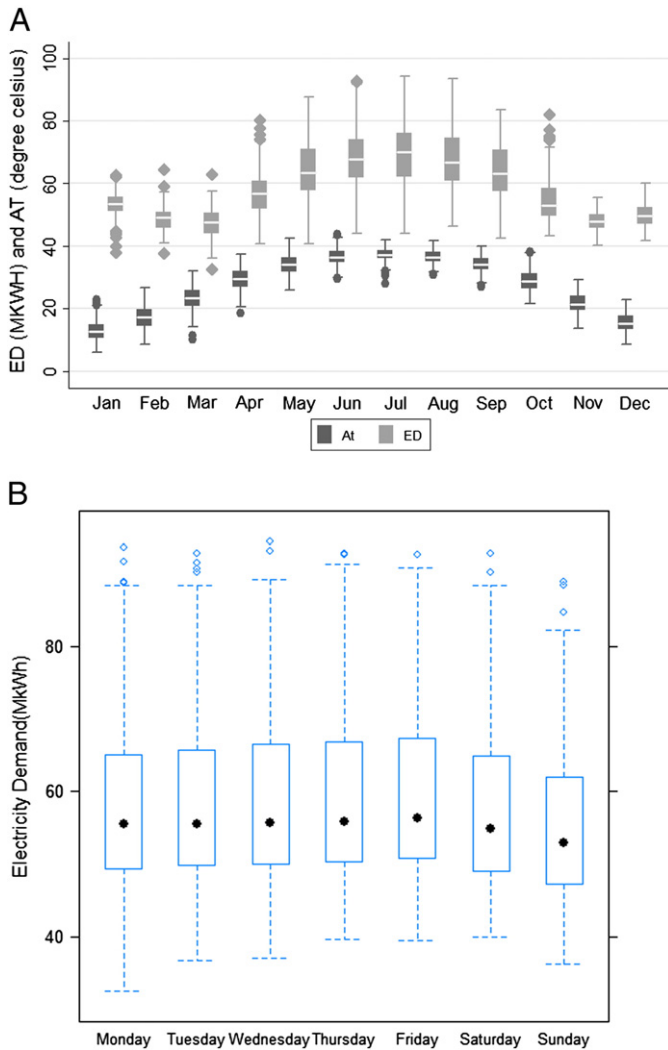


Fig. 3. A. Box plot of daily electricity demand and apparent temperature by months (2000–2009). B. Box plot of daily electricity demand by week days (2000–2009).

According to the chart, non-working days – Sundays and Saturdays – record lower consumption than working days.

6.2. Main results

Tables 2a, 2b and 2c summarizes the results of the estimated models. We estimate all models by the likelihood maximization approach or the penalized likelihood maximization (for Models 3 and 4) using the mgcv package in R. For Models 1 and 2, we use the usual frequentist approach in order to calculate standard errors and p-values for model coefficients. In case of Models 3 and 4, we report the Bayesian p-values and standard errors. We perform the Wald tests of significance for each parametric and smooth term.

Model 2 estimates the non-linear relationship by unpenalized splines using 12 knots selected by the Akaike Information Criterion. The goodness of fit diagnostics and Table 3a show that Model 2 is a significant improvement on Model 1 at 99% confidence levels. An F-test based on the residual values of the semi-parametric Model 2 and the parametric Model 1 yields an F-statistic of 43.984, which has a p-value of .0. This implies that a local fit captures the non-linearity between electricity demand and temperature much more accurately than the global fit of the parametric model.

Model 3 estimates penalized splines with 20 knots as compared to the 12 knots used for the unpenalized spline in Model 2. The goodness of fit diagnostics and Table 3b show that the results from Model 3 are not statistically different from Model 2. The F-test based on residual values of

Models 2 and 3 yields an F statistic of 2.04 and a p-value of 0.12.¹¹ The advantage of using penalized splines is that the results are not influenced by the number of knots when a fairly large number of knots are selected.

Model 4a, in comparison with Model 3 in Table 3c, shows a significant improvement at the 99% confidence level (with F-statistic = 59.31 and p-value = 0). Both the generalized cross-validation (GCV) and Akaike Information Criterion (AIC) are much lower for Model 4. It has a high adjusted R square of .94 implying that it has the ability to explain 94% variation in the electricity demand. The Durbin–Watson statistic (2.01) shows that the estimated model has no autocorrelation.

Model 4a performs only marginally better than Model 4b in respect of adjusted R and other goodness of fit diagnostics. The Durbin–Watson statistic (2.1) shows that the estimated model has no autocorrelation. Moreover, Model 4a has the limitation that we cannot forecast for the future from this model. Fig. 4a plots year-fixed effects from Model 4a which shows that the yearly trend is nearly linear. Similarly, the basis coefficients of Model 4a seem to change linearly over 2000–09 in Fig. 4b. We therefore estimate the much simpler Model 4b imposing linearity assumptions on these coefficients. The advantage of using Model 4b is that we can forecast the time-varying temperature–electricity curves for the future.

We find Models 4a and 4b to give similar results with regard to linear predictors. We observe rainfall to have a significant negative impact on electricity demand, with a 1 mm increase in rainfall reducing electricity demand by 0.05 MkWh in both models. As expected, both holiday dummies turned out to be highly significant and negative. On a major holiday, electricity demand is estimated to be about 3 MkWh lower than the average demand. On a minor holiday, on the other hand, the reduction in demand was only 0.4 MkWh. The estimates of parameters which model the weekly cycle of electricity demand indicate that on Mondays, Saturdays and Sundays electricity demand tends to be lower than the average level (with Wednesday as the reference day) while it is higher on Fridays. These results are as expected since holiday and weekend loads show quite a different response to temperature than those on weekdays. Mondays show a lower demand possibly due to the holiday effect of the previous day (also called holiday inertia) while Fridays show a relatively higher demand, due possibly to the build-up of work at the end of the week. Thus, most of the parametric results are in line with previous studies done in this context.

The effect of apparent temperature on electricity demand is clearly non-linear. The estimated degrees of freedom (edf) for the temperature smooth term estimates and their p-values support the hypothesis that the coefficients are statistically significant. The same smoothing parameter λ_t is chosen for all years resulting in equal degrees of freedom (approximately 6) for each year $f_t(AT)$ in Model 4a. In the case of Model 4b, the estimated degrees of freedom are also about 6 for $f(AT)$ and about 9 for $f(AT) \times y_0$. Fig. 5 plots all the estimated temperature–electricity curves along with the Bayesian confidence intervals¹² (the shaded region)

¹¹ The test statistic is defined as:

$$F = \frac{(RSS_{smaller} - RSS_{larger}) / (df_{res.larger} - df_{res.smaller})}{(RSS_{larger}) / df_{res.larger}}$$

¹² As the penalized splines fit is a trade-off between bias and variance one should account for possible bias in the estimate of f in the determination of variability bands. Wahba (1983) and Nychka (1988) have demonstrated how the Bayesian approach take into account possible bias in the estimate of \hat{f} . Models 3 and 4, effectively impose prior beliefs about the likely characteristics of the correct model by imposing a particular penalty (Woods, 2006). In this approach, we specify a prior distribution on the parameters β such that it reflects our belief that smooth models are more likely than wiggly models. The prior distribution on β is chosen to give a convenient form for the posterior distribution of β given the assumed normal distribution of electricity demand (E) i.e., $(\beta|E)$. For Model 4a, it is $\beta|E \sim N\left(\hat{\beta}, \left[X^T X + \sum_t \lambda_t P_t\right]^{-1} \sigma^2\right)$ and for Model 4b it is $\beta|E \sim N\left(\hat{\beta}, \left[X^T X + \lambda_1 P_1 + \lambda_2 P_2\right]^{-1} \sigma^2\right)$, with $\hat{\sigma}^2 = \frac{\|E - X\hat{\beta}\|^2}{|n - tr(\hat{H}_t)|}$.

Table 2a
Parametric results.

Models	1	2	3	4a	4b
Constant	88.11** (1.45)	49.96** (.24)	49.96** (.24)	50.15** (.17)	49.35** (.14)
2001	2.22** (.28)	2.29** (.27)	2.29** (.27)	2.08** (.20)	
2002	4.7** (.281)	4.8** (.27)	4.8** (.27)	4.7** (.20)	
2003	5.9** (.28)	6.2** (.27)	6.2** (.27)	5.88** (.20)	
2004	8.5** (.281)	8.64** (.27)	8.64** (.27)	8.45** (.20)	
2005	10.11** (.28)	10.27** (.27)	10.27** (.27)	10.09** (.20)	
2006	12.58** (.28)	12.86** (.27)	12.84** (.27)	12.57** (.20)	
2007	11.75** (.28)	12.07** (.27)	12.05** (.27)	11.82** (.20)	
2008	12.69** (.28)	13.09** (.27)	13.06** (.27)	12.82** (.20)	
2009	16.14** (.28)	16.30** (.27)	16.30** (.27)	16.03** (.20)	
Friday	0.96** (0.23)	0.93** (0.22)	0.93** (0.22)	0.95** (0.16)	.98** (.16)
Monday	-0.58** (0.23)	-0.54** (0.22)	-0.54** (0.22)	-0.54** (0.16)	-0.54** (0.16)
Saturday	-0.99** (0.23)	-0.98** (0.22)	-0.98** (0.22)	-0.95** (0.16)	-0.95** (0.16)
Sunday	-3.62** (0.23)	-3.55** (0.22)	-3.55** (0.22)	-3.56** (0.16)	-3.56** (0.16)
Thursday	.28 (0.23)	.24 (0.22)	.25 (0.22)	.32 (0.16)	.31 (.16)
Tuesday	-.11 (0.24)	-.08 (0.22)	-.08 (0.22)	-.06 (0.16)	-.08 (0.16)
Major	-3.5** (0.32)	-3.4*** (0.31)	-3.4*** (0.31)	-3.3** (0.22)	-3.4** (0.22)
Minor	-.38 (0.26)	-.27 (0.25)	-.27 (0.25)	-.42* (0.18)	-.47* (0.18)
Rainfall	-.043** (.01)	-.052** (.01)	-.052** (.01)	-.053** (.01)	-.053** (.01)
AT	-5.91** (0.19)				
AT ²	.21** (0.01)				
AT ³	-.002** (0.0001)				
Trend (year)					1.56** (.02)

Dependent variable for all models is electricity demand. Standard errors are given in brackets ().

Notes:

- ** Significant at 99% significance level.
- * Significant at 95% significance level.

Table 2b
Approximate significance of smooth terms in semiparametric models.

	EDF	F	p-value
Model 2			
f(AT)	11	1856	0.000
Model 3			
f(AT)	8.73	1899	0.000
Model 4a			
f(AT): 2000	6.4	218.6	0.000
f(AT): 2001	6.2	279.3	0.000
f(AT): 2002	6.3	369	0.000
f(AT): 2003	6.3	314.1	0.000
f(AT): 2004	6.2	448.9	0.000
f(AT): 2005	6.5	460.4	0.000
f(AT): 2006	6.3	712.5	0.000
f(AT): 2007	6.5	764.5	0.000
f(AT): 2008	6.2	767.3	0.000
f(AT): 2009	6.1	1417.6	0.000
Model 4b			
f(AT)	5.8		0.000
f(AT): year no	8.9		0.000

for both models. For Model 4a it plots $\hat{f}_t(AT_i)$ for each t and i along with $\hat{f}_t(AT_i) \pm 2 \times SE(\hat{f}_t(AT_i))$. For Model 4b it plots $\hat{f}(AT_i) + \hat{f}(AT_i) \times y_0$ with $(\hat{f}(AT_i) + \hat{f}(AT_i) \times y_0) \pm 2 \times SE(\hat{f}(AT_i) + \hat{f}(AT_i) \times y_0)$.

Table 2c
Goodness of fit diagnostics.

Model	1	2	3	4a	4b
Adjusted R ²	0.873	0.884	0.884	0.94	0.93
AIC	20031.87	19,709	19,709	17474.5	17,638
GCV	14.301	13.09	13.091	7.09	7.3
No obs	3643	3643	3643	3643	3652
Model DF (degrees of freedom)	22	30	26.67585	81.918	25.7
Residual DF (N-DF)	3621	3613	3616.33	3561.08	3626.3

Table 3a
Comparing Model 1 and Model 2 (F-test).

	Residuals (1)	Res. df (2)	Diff. df (3)	Diff. res (4)	F = $\frac{(4)/(3)}{\text{Model2}(1)/(2)}$	p-value
Model 1	51,472	3621				
Model 2	46,904	3613	8	4568	43.984	000

Table 3b
Comparing Model 2 and Model 3 (F-test).

	Residuals (1)	Res. df (2)	Diff. df (3)	Diff. res (4)	F = $\frac{(4)/(3)}{\text{Model3}(1)/(2)}$	p-value
Model 2	46,904	3613				
Model 3	46,964	3615	-2.26	-60.177	2.0436	0.1229

Fig. 6 plots the marginal effect (first derivative) curve from both models of each temperature–electricity curve in Fig. 5 with 95% Bayesian confidence intervals (see shaded region). For Model 4a it plots $\hat{f}'_t(AT_i)$ for each t along with $\hat{f}'_t(AT_i) \pm 2 \times SE(\hat{f}'_t(AT_i))$. For Model 4b it plots $\hat{f}'(AT_i) + \hat{f}'(AT_i) \times y_0$ with $(\hat{f}'(AT_i) + \hat{f}'(AT_i) \times y_0) \pm 2 \times SE(\hat{f}'(AT_i) + \hat{f}'(AT_i) \times y_0)$.

Table 3c
Comparing Model 3 and Model 4a (F-test).

	Residuals (1)	Res. df (2)	Diff. df (3)	Diff. res (4)	F = $\frac{(4)/(3)}{\text{Model4}(1)/(2)}$	p-value
Model 3	46,964	3615.3				
Model 4	24,684	3561.1	54.2	22,277	59.31	000

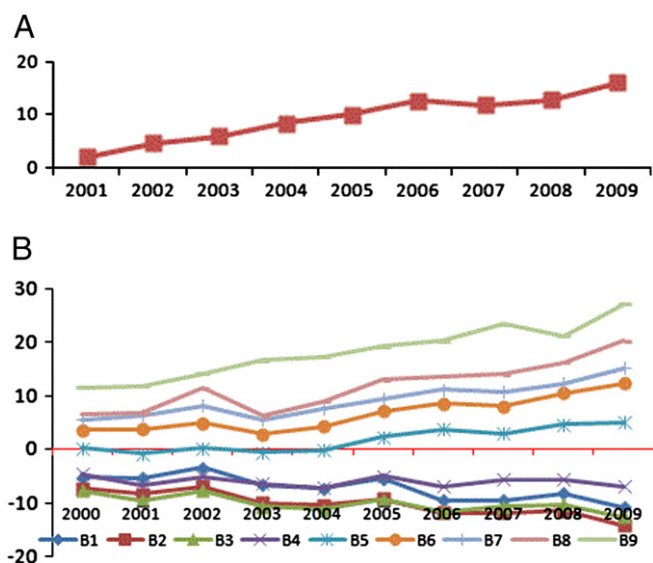


Fig. 4. A. Year-fixed effects from Model 4a. Note: base year is 2000 and it plots coefficients of year dummies from 2001 to 2009 as obtained from the factor variable coefficient model. B. Basis coefficients from Model 4a. Note: it plots nine coefficients corresponding to nine basis functions of each year from 2000 to 2009 as obtained from the factor variable coefficient model. Although there are 10 knots for each year but as discussed in the paper one degree of freedom is lost due to the identification constraint.

We obtain the minimum temperature threshold for the corresponding year when a marginal effect curve cuts the zero line on the y-axis. Over time, it is evident that the minimum temperature threshold is falling and the temperature dependence curves of Delhi are moving leftwards. Fig. 7 plots these estimated threshold intervals represented by the shaded region (at zero line) in Fig. 6. It is clear that threshold intervals have shifted from approximately 20–22 °C in 2000–05 to about 18.5–20 °C in 2006–09 during the period of analysis.

It is possible to explain this shift through reference to the increase in the use of air conditioners and air coolers with rising incomes. In other words, people's sensitivity to hot temperatures is likely to increase with their ability to afford expensive cooling devices, which would in turn result in their switching of such devices at relatively lower temperatures.

In addition to the leftward shift of the temperature–electricity curve, we observe that the rising part of the curve is becoming steeper over time, implying an ever increasing cooling demand per unit increase in summer temperatures. As discussed previously, this may partly be attributable to the increasing penetration of energy-intensive cooling devices such as air conditioners that give greater control over rising temperatures to the residents, especially in the humid summer climate of Delhi. The effect of the decline in heating demand per unit increase in winter temperatures, however, is much lower than the increase in cooling demand in summers. For instance, in both models, a 1 °C increase in temperature at 30 °C in the summer increased electricity demand by over 3 M kWh in 2009 as compared to only over 1 M kWh in 2000. On the other hand, a 1 °C increase in temperature at 15 °C in the winter decreased electricity demand by only 0.8 M kWh in 2009 as compared to 0.7 M kWh in 2000.

In an earlier study which estimated a threshold transition model, Moral-Carcedo and Vicens-Otero (2005) found 15.5 °C as the upper heating demand threshold and 18.4 °C as the lower cooling demand threshold for Spain. The smooth transition model obtained 15.4 °C as an optimal threshold temperature. In another study, Bessec and Fouquau (2008) found the threshold temperature to be about 16 °C

for the whole sample of European countries while it was 14 °C for the sample of cold European countries and 22.4 °C for the sample of hot European countries. Although the thresholds obtained in this paper are not directly comparable with the previous studies, due to the average temperatures that those studies are based on in contrast with the apparent temperatures used in the present study, the thresholds give a fairly good idea about how threshold temperatures may vary both spatially and temporally with economic growth and that therefore they cannot be assumed to be static.

7. Global warming and electricity demand

The leftward shifting temperature–electricity curve and the rightward shifting temperature distribution may have significant implications for electricity demand in India in future.

7.1. Key assumptions

Based on the existing global warming projections for India by the Intergovernmental Panel on Climate Change (IPCC) and the Hadley Centre, we assume three hypothetical scenarios for global warming: a uniform 1 °C, 2 °C and 3 °C increase in daily apparent temperatures between 2009 and 2030 over the daily apparent temperature for 2009. Assuming that other predictor variables are the same as in the year 2009, we therefore use the estimated model to predict the impact of 1 °C, 2 °C and 3 °C increase in daily apparent temperatures on the electricity demand in Delhi till the year 2030.

7.2. Results: the impact of global warming

As shown in the previous section, the impact of temperature change or global warming on electricity demand is likely to be time-varying. Fig. 8 therefore plots the forecasted temperature–electricity curve, the marginal effect curve and the threshold temperature up to 2030 under the baseline scenario (i.e., assuming that the temperature remains the same as in 2009) using Model 4b.

The results show that the rising part of the curve is likely to become much steeper in the future. A 1 °C increase in apparent temperature at 30 °C is expected to increase electricity demand by over 5 M kWh in 2015, 6 M kWh in 2020, 7 M kWh in 2025 and about 8 M kWh in 2030. We observe that the threshold apparent temperature is likely to fall from about 19 °C in 2009 to 17.7 °C by 2030.

Fig. 9 displays the results for the three global warming scenarios. A 1 °C increase in apparent temperature increases the electricity demand by about 402 M kWh (or 1.7%) in 2009 over its base electricity demand of 23,827 M kWh, 564 M kWh (or 2.1%) in 2015 over its base electricity demand of 27,522 M kWh, and 968 M kWh (or 2.6%) in 2030 over its base electricity demand of 36,761 M kWh. A 2 °C increase in apparent temperature, on the other hand, increases net electricity demand by about 842 M kWh (or 3.5%) in 2009, 1182 M kWh (or 4.3%) in 2015 and 2032 M kWh (or 5.5%) in 2030. Similarly, a 3 °C increase in the apparent temperature increases the net electricity demand by about 1321 M kWh (or 5.5%) in 2009, 1856 M kWh (or 6.7%) in 2015 and 3191 M kWh (or 8.7%) in 2030.

In addition, Fig. 9 also disaggregates the impacts of global warming by months. Higher temperatures increase electricity demand in the summers (led by April and May) and in the monsoon period (led by September) and post-monsoon period (led by October) while lower temperatures decrease the electricity demand in winters (led by January). It is evident that the maximum impact is likely to be felt in the hot month of April with an average apparent temperature of 30 °C, followed by the months of October and May. The marginal effect curve peaks at about 30 °C, indicating the maximum sensitivity of electricity demand to temperature at this level. As demand peaks during the months of June and July, the additional increments in

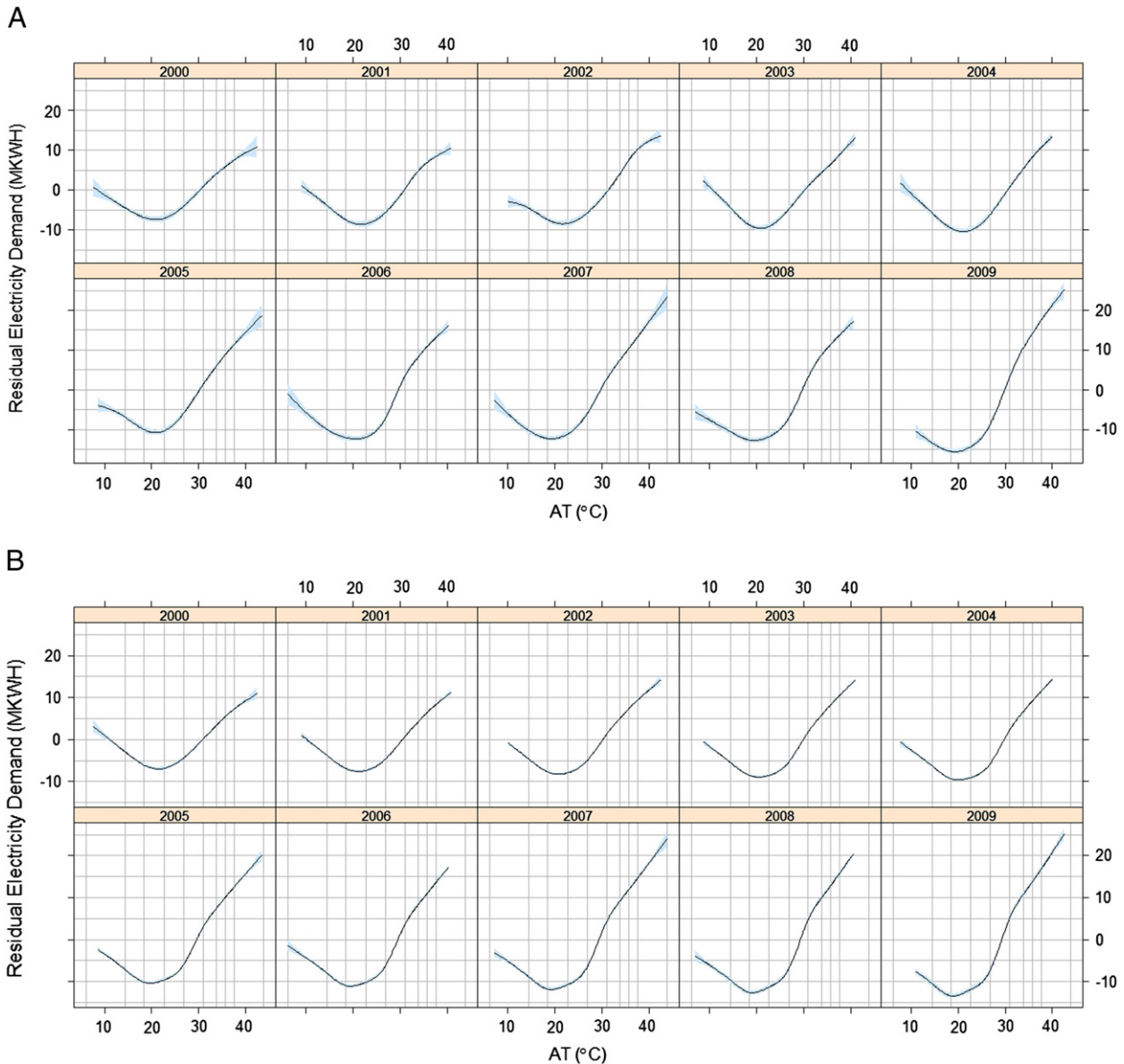


Fig. 5. A. Estimated temperature–electricity curves along with the 95% Bayesian confidence intervals from Model 4a. Note: It plots $\hat{f}_t(AT_t)$ for each t along with $\hat{f}_t(AT_t) \pm 2 \times SE(\hat{f}_t(AT_t))$ as represented by the shaded region. B. Estimated temperature–electricity curves along with the 95% Bayesian confidence intervals from Model 4b. Note: it plots $\hat{f}_t(AT_t) + \hat{f}_t(AT_t) \times y_0$ for each t along with $(\hat{f}_t(AT_t) + \hat{f}_t(AT_t) \times y_0) \pm 2 \times SE(\hat{f}_t(AT_t) + \hat{f}_t(AT_t) \times y_0)$ as represented by the shaded region.

electricity demand due to higher temperatures (over 30 °C) slow down and the marginal effect curve stabilizes.

Although a 1 °C increase in temperature increases the total electricity demand by 1.7% in 2009, the demand increases by about 3.6% in April, by 3.1% in October, and by 2.4% in September, May, and March. On the other hand, a 1 °C increase in temperature decreases electricity demand by 1.2% in January, 0.4% in February and 0.9% in December. Moreover, although at present higher temperatures reduce electricity demand in the month of February, this trend may be reversed in the future years with a leftward shift of the temperature threshold shifting the balance between the decreasing electricity demand for heating and the increasing electricity demand for cooling resulting from global warming. With a 2 °C and 3 °C increase in apparent temperature we may witness this trend much faster. Since electrical energy saved in winters cannot be easily stored for use in summers, global warming could result in a serious disequilibrium in electricity supply and demand during some months of the year in the future.

In order to evaluate the prediction performance, we have compared the actual demand of two years (2008–09) with the predicted demand. In this evaluation, we have calculated the predicted demand for the two years using coefficients of the estimated model (based on 2000–07 data), known temperatures, and information on other drivers in these years. We have not used data from the forecast period for the model estimation. Fig. 10 illustrates the difference between observed and predicted electricity demand in 2008 and 2009. These graphs demonstrate that the model predicts demand in both years remarkably well. For instance, in 2008 the mean observed electricity consumption is about 62 M kWh while its standard deviation is about 11 M kWh. The root mean square error of predicted electricity consumption is 4.3 M kWh and the mean absolute error is 3.4 M kWh. Both these measures are much lower than the standard deviation of the electricity demand.

There are limitations to our research design when it comes to measuring the impacts of global warming on electricity demand in

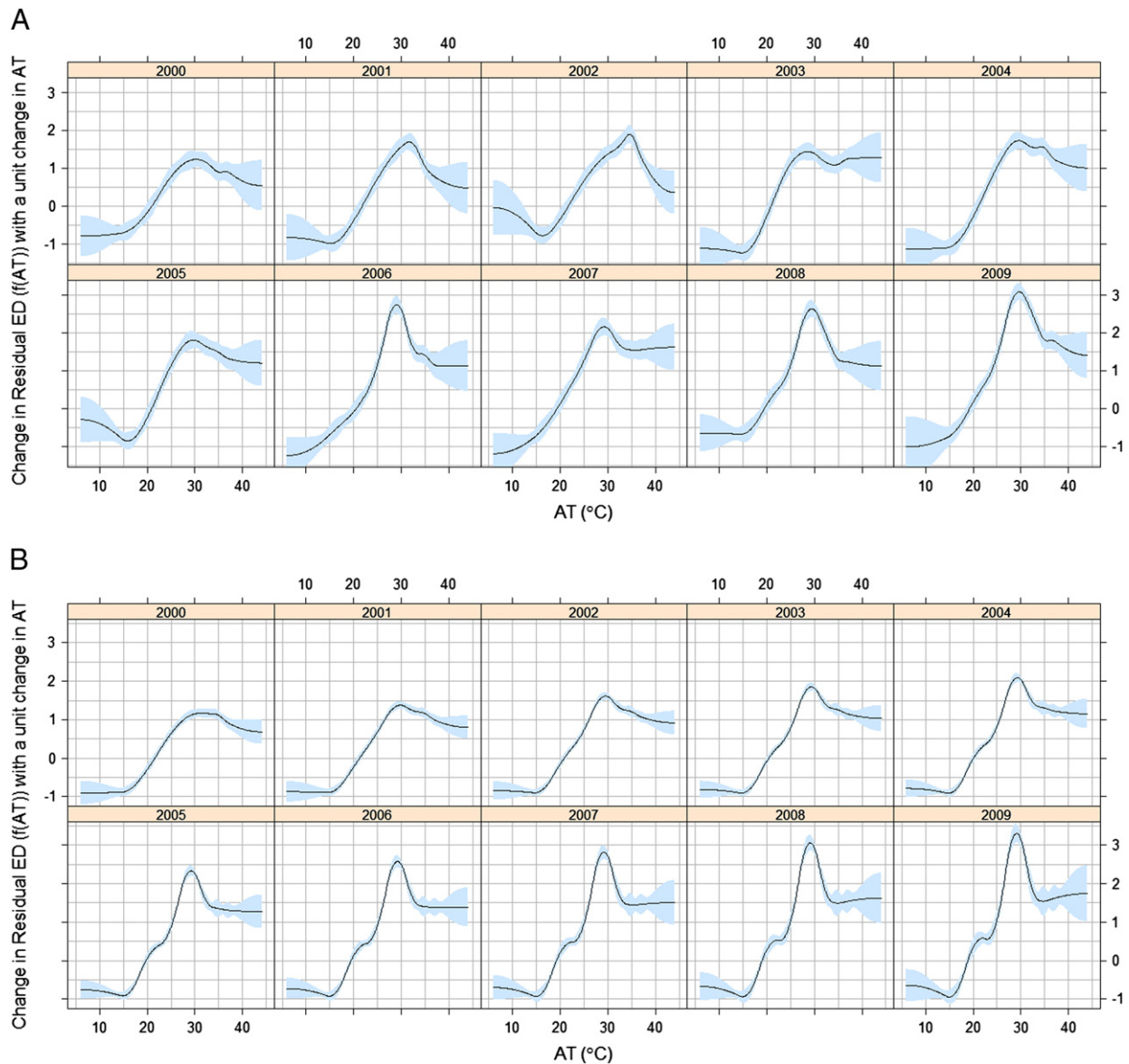


Fig 6. A. Estimated marginal effect curves along with the 95% Bayesian confidence intervals from Model 4a. Note: it plots $\hat{f}'_t(AT_t)$ for each t along with $\hat{f}'_t(AT_t) \pm 2 \times SE(\hat{f}'_t(AT_t))$ as represented by the shaded region. B. Estimated marginal effect curves along with the 95% Bayesian confidence intervals from Model 4b. Note: It plots $\hat{f}'_t(AT_t) + \hat{f}'_t(AT_t) \times y_0$ for each t along with $(\hat{f}'_t(AT_t) + \hat{f}'_t(AT_t) \times y_0) \pm 2 \times SE(\hat{f}'_t(AT_t) + \hat{f}'_t(AT_t) \times y_0)$ as represented by the shaded region.

Delhi. Firstly, the impact of global warming on electricity demand is addressed by making a prediction for 20 years ahead. The model that is used for the prediction is based on the past 10 years of data and validated using 2-year-ahead prediction checks. Clearly, the prediction will be correct only if there is not too much structural change in the next 20 years. For instance, using the factor variable coefficient model we observe that the basis coefficients change almostly linearly with time over the past 10 years and we thus imposed the assumption that these coefficients will change linearly with time in the numeric variable coefficient model. However, these assumptions may not hold true in 2030 if there are significant structural changes in the economy. Hence, the results obtained in this study should not be interpreted as exact forecasts but as simply indicative of the direction and magnitude of the effects that might be expected from climate change. Secondly, as against the uniform global warming scenarios considered in the paper, there is a large body of evidence which suggests

that temperature change has been non-homogeneous across months and seasons in the past. As discussed in Section 2, summers and post-monsoon seasons have shown the maximum shift. For instance, the average apparent temperature in April has shown a maximum increase of 2.21 °C over two decades (1990–99 and 2000–10). Rising temperatures during this season coupled with the peaked marginal effect curve during this period (Fig. 6) will intensify the disequilibrium problem. A monthly disaggregation of climate impacts will therefore enhance the policy relevance of the scenarios considered in this paper. Thirdly, this study only investigates the total electricity demand pattern for possible changes related to apparent temperature. Apart from the effect of global warming on total electricity demand, global warming will have significant impacts on future peak electricity demands and its variability. Colombo et al. (1999), using Canadian data, have shown that with climate change, the average peak demand may not increase drastically; instead, the number of high-electricity-consumption days may increase

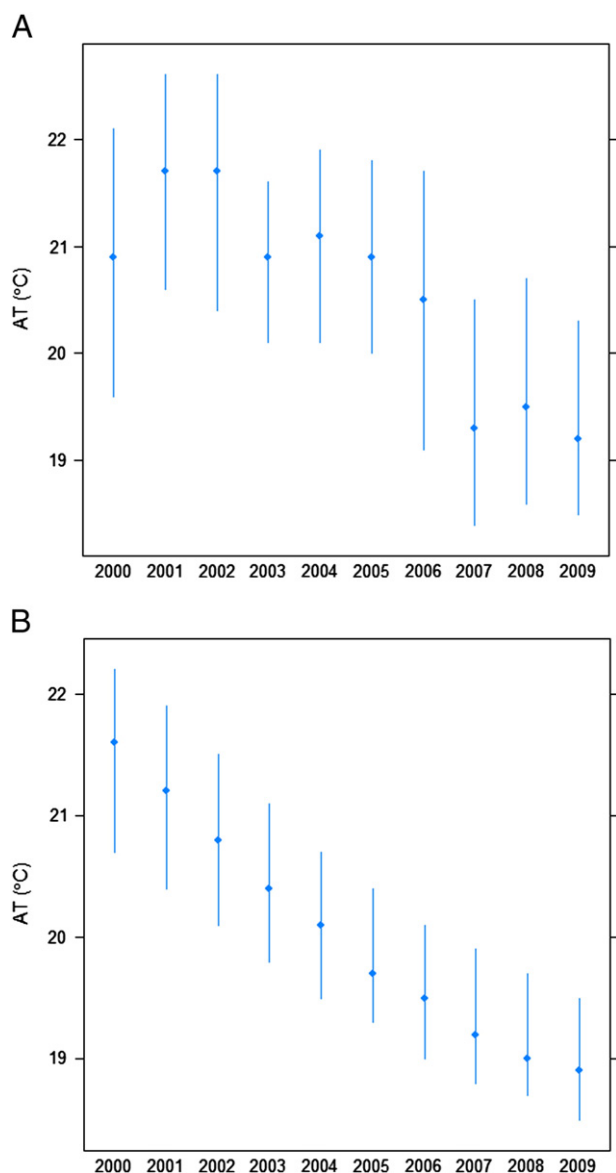


Fig. 7. A. Minimum apparent temperature thresholds for 2000–2009 from Model 4a. B. Minimum apparent temperature thresholds for 2000–2009 from Model 4b.

appreciably due to higher variability, placing a stress on power utility to meet this higher demand. However, due to data limitations, we have not considered peak demand, which is left for future research.

8. Conclusions

Both changing lifestyles and economic conditions in India have made electricity demand increasingly sensitive to temperature. This paper provides valuable insights into potential interactions between increasing cooling degree days and increasing incomes, and the impact of the resulting long-term adjustments (such as the higher penetration of air cooling devices) in the electricity sector. The results from a semi-parametric variable coefficient model indicate that the variation in the slope of the temperature–electricity curve and the threshold temperature is important for future electricity demand projections. An important contribution of the paper is the estimation of climate impacts by months. The model projects that the warming can result in significant increases in future demand, particularly

during the hot months of April and May. These results can be extremely useful in managing the seasonal electricity disequilibrium situation in Delhi. For instance, demand-side management by shifting electricity loads from periods of deficit supply to surplus supply via dynamic pricing and other control mechanisms may help those in charge of energy policy-making and policy-implementation to cope with the problem.

Policy makers will moreover need to come up with new measures to meet increased electricity demand due to global warming. They would, for instance, have to make a choice between fossil fuels and renewable energy sources for electricity generation. The results obtained in our study will be of use to electricity production and sales companies too in order to a) understand existing temperature–electricity sensitivity so as to manage risks related to unpredictable changes in energy demand under extreme weather events, for e.g., a heat wave; b) quantify the impact of projected global warming on electricity use; and c) forecast required future capacity investments in the electricity sector. The estimated threshold temperature of our study would be of use to HVAC¹³ (heating, ventilation and air-conditioning) designers for the purpose of improving the efficiency of electricity use in their products. At present, the comfort standard practiced by HVAC designers is the same as that adopted in the U.S.A. for cooling buildings (i.e., air-conditioned buildings). Since a large amount of electricity is consumed by HVAC systems in buildings, designing HVAC (for comfort) as per the changing climatic conditions in India could bring down the electricity demand drastically.

Comprehensive assessment of impacts however requires not just sound empirical research but more geographical coverage, especially in areas where severe global warming is likely to occur. Hence, any future work on the topic should seek to extend the approach to other states in India in order to get an overall estimate of global warming on total electricity demand in India since the different socio-economic profiles of the states would lead to different temperature–electricity curves. It is our hope nevertheless that the present study would contribute to a better understanding of the dynamic non-linear temperature–electricity curve in a rapidly growing city.

Acknowledgments

I thank Professors E Somanathan and Bharat Ramaswami for their guidance and invaluable help at different stages of this paper; Professors Probal Chaudhuri and Trevor Hastie and Drs. Swagata Nandi and Deepayan Sarkar for help with developing the research; and Dr. Simon Woods for his assistance with understanding the issues surrounding Generalized Additive Models (GAMs) estimation in R software with mgcv package. I also acknowledge other contributors to the study: Drs. Abhiroop Mukhopadhyay, Indira Rajaraman, Nityananda Sarkar, Rohini Somanathan and Ajay Mathur and Mr. Rahul Mazumder, Mr. Prabir Sengupta, and Ms Kiran Kumar. I am thankful to two anonymous referees for extremely useful comments. The data for the research came from many sources. I thank Messrs. SK Sonee, Ashok Rajput, Harish Rathore, Hemant Jain and AK Kaul for their support in procuring electricity data and Mr. Shyam Kumar for assistance with data operations. An earlier version of this paper was presented at the Climate Policy Outreach meeting held in Xian in March 2011, Second National Research Conference on Climate Change

¹³ HVAC (heating, ventilation, and air conditioning) refers to technology for indoor or automotive environmental comfort. HVAC is important in the design of medium to large industrial and office buildings such as skyscrapers or marine environments such as aquariums, where safe and healthy building conditions take into consideration temperature and humidity, including “fresh air” from outdoors.

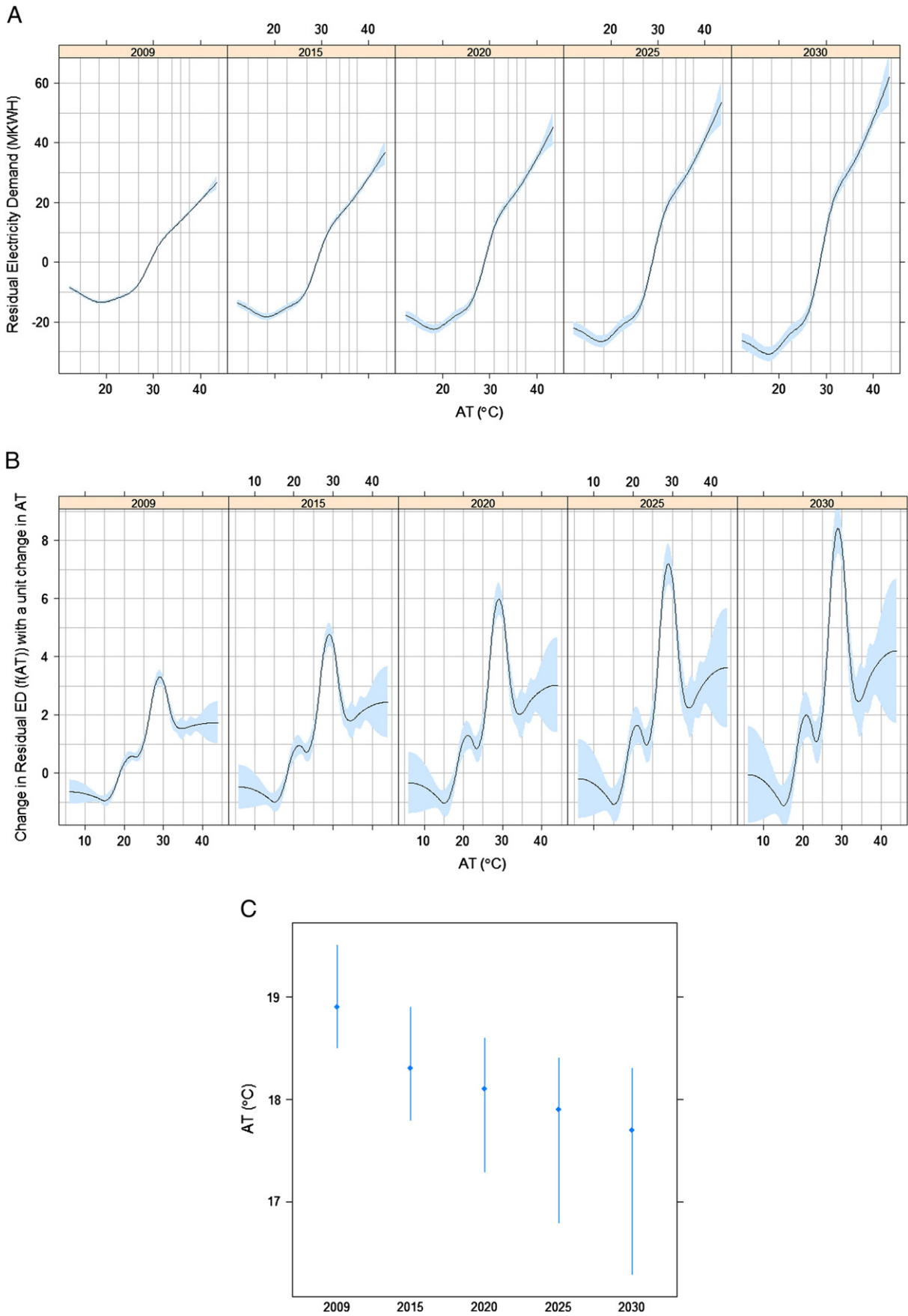


Fig. 8. A. Change in electricity demand (%) over the base electricity demand in the corresponding year due to 1 °C increase in apparent temperature. B. Change in electricity demand (%) over the base electricity demand in the corresponding year due to 2 °C increase in apparent temperature. C. Change in electricity demand (%) over the base electricity demand in the corresponding year due to 3 °C increase in apparent temperature.

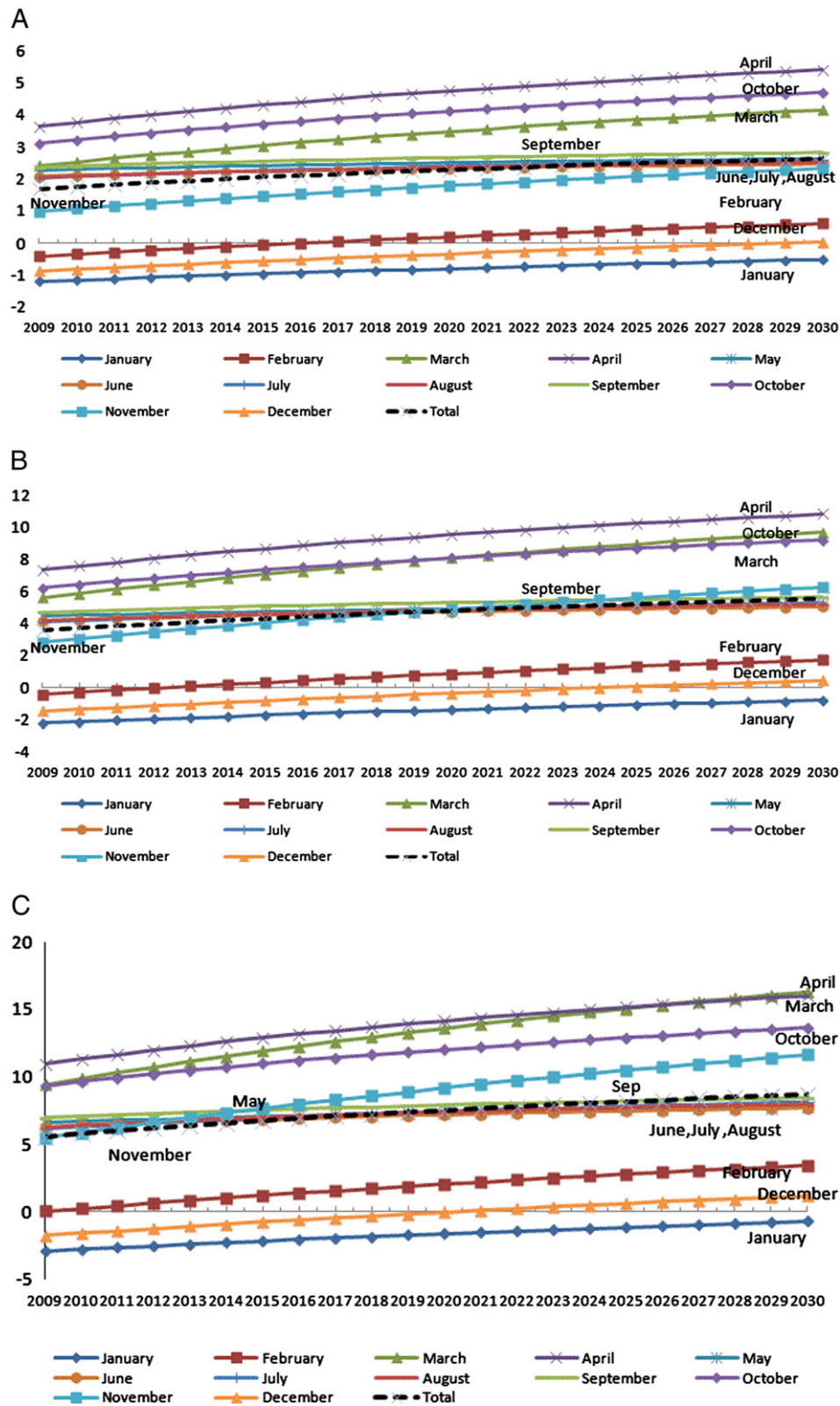


Fig. 9. A. Forecasted temperature–electricity curves till 2030. B. Forecasted temperature–electricity marginal effect curves till 2030. C. Forecasted minimum apparent temperature thresholds till 2030 from Model 4b.

November 2011 at IIT Delhi and Seventh Annual Conference on Economic Growth and Development December 2011 at ISI Delhi. We gratefully acknowledge the funding received from the European

Commission's Climate Policy Outreach project (SERVICE CONTRACT n. 21.040100/2009/551481/SER/C1) and the Planning & Policy Research Unit (PPRU) at ISI.

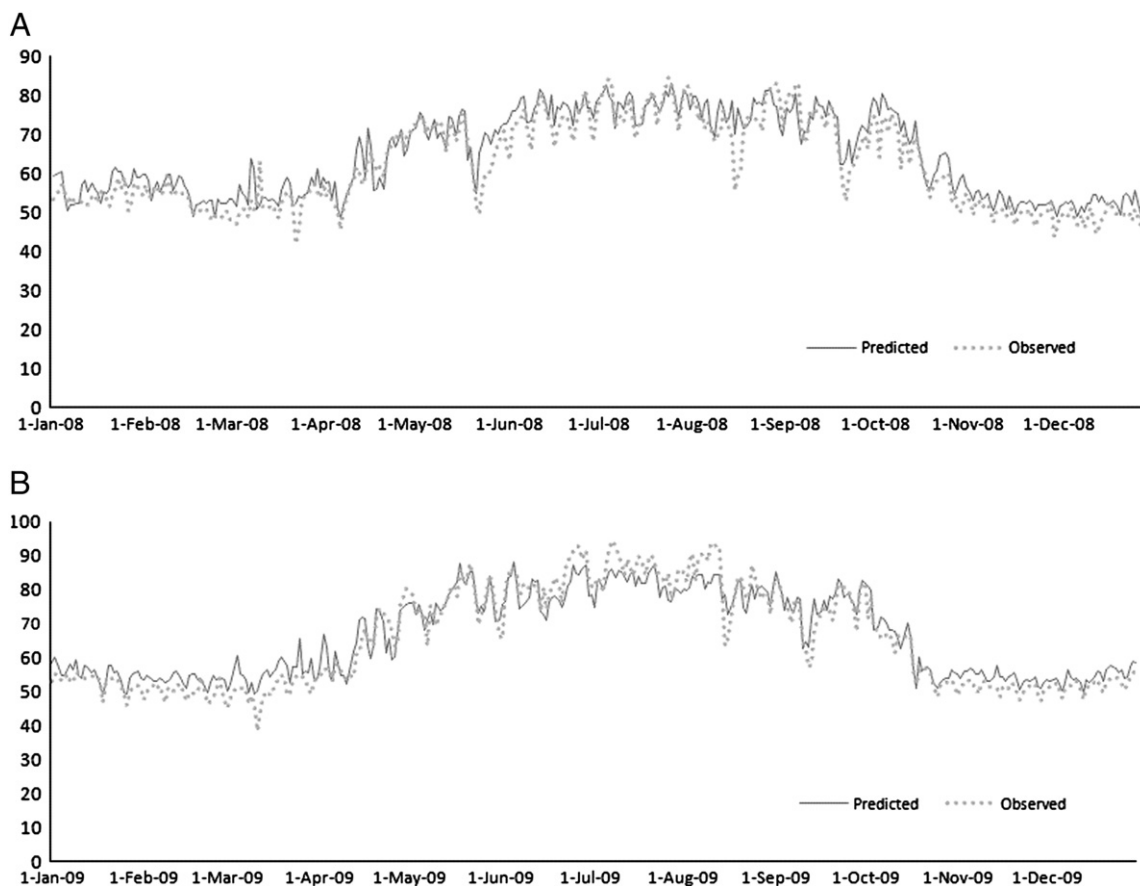


Fig. 10. A. Model evaluation: Predicted and observed electricity demand in 2008. B. Model evaluation: predicted and observed electricity demand in 2009.

References

- Al Zayer, J., Al Ibrahim, A.A., 1996. Modeling the impact of temperature on electricity consumption in the Eastern Province of Saudi Arabia. *J. Forecast.* 15 (2), 97–106.
- Bessec, M., Fouquau, J., 2008. The non-linear link between electricity consumption and temperature in Europe: a threshold panel approach. *Energy Econ.* 30 (5), 2705–2721.
- Colombo, A.F., Etkin, D., Karney, B.W., 1999. Climate variability and the frequency of extreme temperature events for nine sites across Canada: implications for power usage. *J. Clim.* 12, 2490–2502.
- Engle, F., Granger, C.W.J., Rice, J., Weiss, A., 1986. Semi parametric estimates of the relation between weather and electric sales. *J. Am. Stat. Assoc.* 81 (394), 310–320.
- Harvey, A., Koopman, S.J., 1993. Forecasting hourly electricity demand using time-varying splines. *J. Am. Stat. Assoc.* 88 (424), 1228–1236.
- Hastie, T., Tibshirani, R., 1993. Varying-coefficient models. *J. R. Stat. Soc. B Methodol.* 55 (4), 757–796.
- Hekkenberg, M., Moll, H.C., Uiterkamp, A.J.M., 2009. Dynamic temperature dependence patterns in future energy demand models in the context of climate change. *Energy* 34 (11), 1797–1806.
- Henley, A., Peirson, J., 1997. Non linearities in electricity demand and temperature: parametric versus non parametric methods. *Oxf. Bull. Econ. Stat.* 59 (1), 149–162.
- Hyndman, R.J., Fan, S., 2008. Forecasting long-term peak half-hourly electricity demand for South Australia. Report for Electricity Supply Industry Planning Council (SA). Monash University Business and Economic Forecasting Unit, Australia.
- Keele, L., 2008. Semiparametric Regression for the Social Sciences. Wiley Online Library. (<http://onlinelibrary.wiley.com/doi/10.1002/9780470998137.ch8/summary> (ACCESSED ON: 31st Jan 2012)).
- Kothawale, D.R., Munot, A.A., Kumar, K.K., 2010. Surface air temperature variability over India during 1901–2007 and its association with ENSO. *Clim. Res.* 42, 89–104.
- Lancaster, P., Salkauskas, K., 1986. *Curve and Surface Fitting: An introduction*. Academic Press, London.
- Li, Q., Racine, J.S., 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, New Jersey.
- Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Kotroni, E., Lagouvardos, Y., Lalas, D.P., 2007. Modeling framework for estimating impacts of climate change on electricity demand at regional level: case of Greece. *Energy Convers. Manage.* 48 (5), 1737–1750.
- Moral-Carcedo, J., Vicens-Otero, J., 2005. Modeling the non-linear response of Spanish electricity demand to temperature variations. *Energy Econ.* 27 (3), 477–494.
- Nychka, D., 1988. Bayesian Confidence Intervals for Smoothing Splines. *J. Am. Statist. Assoc.* 83 (404), 1134–1143.
- Pardo, A., Meneu, V., Valor, E., 2002. Temperature and seasonality influences on Spanish electricity load. *Energy Econ.* 24 (1), 55–70.
- Powell, J., 1994. Estimation of semiparametric models. In: Engle, R.F., McFadden, D.L. (Eds.), *Handbook of Econometrics*, vol. 4. Elsevier Science B.V., Amsterdam, pp. 2443–2521.
- Ramesh, S., Natarajan, B., Bhagat, G., 1988. Peak load prediction using weather variables. *Energy* 13 (8), 671–679.
- Ruppert, D., Wand, M.P., Carroll, R.J., 2003. *Semiparametric Regression*. Cambridge University Press, Cambridge, UK.
- Sailor, D.J., 2001. Relating residential and commercial sector electricity loads to climate-evaluating state level sensitivities and vulnerabilities. *Energy* 26 (7), 645–657.
- Sailor, D.J., Muñoz, J.R., 1997. Sensitivity of electricity and natural gas consumption to climate in the U.S.A.—Methodology and results for eight states. *Energy* 22 (10), 987–998 (October, ISSN 0360-5442, 10.1016/S0360-5442(97)00034-0).
- Sailor, D.J., Pavlova, A.A., 2003. Air conditioning market saturation and long-term response of residential cooling energy demand to climate change. *Energy* 28 (9), 941–951.
- Steadman, R.G., 1994. Norms of apparent temperature in Australia. *Aust. Meteorol. Mag.* 43 (1).
- Valor, E., Meneu, V., Caselles, V., 2001. Daily air temperature and electricity load in Spain. *J. Appl. Meteorol.* 40 (8), 1413–1421.
- Wadud, Z., Noland, R.B., Graham, D.J., 2010. A semiparametric model of household gasoline demand. *Energy Econ.* 32 (1), 93–101.
- Wahba, G., 1983. Bayesian "Confidence Intervals" for the cross-validated smoothing spline. *J. R. Stat. Soc. Series B Methodol.* 45 (1), 133–150.
- Wangpattarapong, K., Maneewan, S., Ketjoy, N., Rakwichian, W., 2008. The impacts of climatic and economic factors on residential electricity consumption of Bangkok metropolis. *Energy Build.* 40 (8), 1419–1425.
- Wood, S.N., 2006. *Generalized Additive Models: An Introduction with R*. CRC Press, NW.
- Wood, S.N., Augustin, N.H., 2002. Gams with integrated model selection using penalized regression splines and applications to environmental modeling. *Ecol. Model.* 157 (2), 157–177.
- Yatchew, A., 2003. *Semiparametric Regression for the Applied Econometrician*. Cambridge University Press, Cambridge, UK.