# Climate Change and the Demand for Electricity: A Non-Linear Time Varying Approach

Eshita Gupta
Indian Statistical Institute, Delhi
Working Paper

December 6, 2011

**Abstract**

This paper estimates climate sensitivity of electricity demand by examining the impact of apparent temperature on the electricity demand of Delhi using daily data for the period 2000-2009. A semiparametric variable coefficient model has been adopted to investigate the non linear time varying impact of climatic factors on electricity demand.

Electricity demand is a U-shaped function of temperature. I find that that the rising part of the temperature-electricity curve is becoming more pronounced over time implying increasing cooling demand per unit increase in summer temperatures. Increasing temperature dependence of cooling demand shifts the temperature-electricity curve of Delhi leftwards. Further, adverse effects of climate change will be asymmetrically distributed in different seasons resulting in serious disequilibrium in hot months in the future.

## 1 Introduction

A growing consensus regarding the plausibility of increases in the Earth's mean temperature has stimulated attempts to assess the possible impacts of such changes on different economic sectors. The goal of this paper is to quantify how climate change will affect electricity demand in the continental climate of Delhi ($28^0 30'$N), one of the biggest cities in India. I use a within-year identification strategy to estimate the effects of apparent temperature on daily electricity demand over a 10-year period (2000-2010). The estimated model is used to forecast the impact of $1^0$C, $2^0$C and $3^0$C increase in apparent temperature on electricity demand of Delhi in 2015 and 2021.

The literature highlights a U-shaped non-linear temperature-electricity curve (TEC). Starting from low levels, rising temperatures first decrease electricity demand due to lower heating demand in cold weather, and after the level of temperature exceeds the minimum electricity demand threshold, rising temperatures increase electricity demand due to higher cooling demand in hot weather. The expected net effect of climate change 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. As a result, climate warming would result in a decline in electricity demand in these countries. On the other hand, the reverse has been predicted for Germany with the cooling effect dominating the heating effect. 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. However, we do not know the size and nature of climate warming effects on electricity demand for India. Such quantification is attempted in this paper.

The key contribution of this paper is that it recognizes and addresses two special problems in the estimation of the TEC for developing countries. First, with rapid changes in economic structure in future, the relation is likely to be shifting over time. In this paper, I address this issue by estimating a semiparametric variable coefficient model that allows temperature-electricity relation to vary over time. Like in Engle et al. (1986), the temperature-electricity relation is modeled nonparametrically 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 nonparametric temperature-electricity relation to vary across years by interacting the non-parametric component with year dummies.

Second, black-outs or power-outages are common in many developing countries. This means that observed electricity use is typically less than the notional electricity demand (the object of interest). I adjust daily electricity consumption with daily shoratge data to obtain unrestricted demand of Delhi.

It is important to highlight the limitations of this study. First, this study takes a broad perspective, estimating the average TEC for the aggregate electricity demand of Delhi. The temperature-electricity sensitivities may differ across sectors significantly. In the residential and commercial sector, a large chunk of this demand is due to space conditioning and water heating that is highly sensitive to temperature. On the other hand, in agriculture and industrial sectors, electricity demand is determined by the level of economic activity and is largely temperature insensitive. Given data limitations it is not possible to obtain daily electricity demand data of different sectors and an aggregated approach has been adopted. Moreover, in Delhi (with 97.5% urban population), residential and commercial sector together account for about 80% of the total electricity demand.

Second, electricity demand can be modeled structurally, where electricity consumption is chosen to maximize the expected utility of the households and profits of the firms. In order to estimate such a model, I require data on prices, utilization and efficiency of electricity using equipment at the household and firm level over time. However, these data do not exist, and a structural model is hard to implement. Thus, like much of the literature this paper also works with a reduced form model.

There are three results from my analyses that have important implications for electricity-climate policy: first, it is observed that the rising part of the TEC is becoming more pronounced over time implying increasing cooling demand per unit increase in summer temperatures. For instance a $1^0$C increase in temperature at $30^0$C increased electricity demand by 3.2 Million kilowatt-hours (MKWH) in 2009 as compared to only 1.2 MKWH in 2000. On the other hand, a $1^0$C increase in temperature at $15^0$C decreased electricity demand by 0.8 MKWH in 2009 as compared to 0.7 MKWH in 2000. Second, increasing temperature dependence of cooling demand shifts the temperature-electricity curve of Delhi leftwards. In particular, minimum temperature threshold (TT) shifts from about 20-$23^0$C in 2000-2005 to about 17-$21^0$C in 2006-2009. Third, my study suggests that adverse effects of climate change will be asymmetrically distributed in different seasons. Higher temperature increases electricity demand in summers (led by April and May), monsoon (led by September) and post monsoons (led by October) and decreases demand in winters (led by January). Since, electricity saved in winters cannot be stored and used in summers, global warming could result in serious disequilibrium in hot months in the future.

The rest of the paper is organized as follows. Section 2 reviews existing studies and models that assess the impact of temperature on electricity demand. Section 3 discusses the estimation strategy. Section 4 describes the data sources and summary statistics of the major variables. Section 5 discusses the results of the empirical model. Section 6 forecasts future electricity demand impacts under three different climate scenarios, and evaluates the estimated model. Section 7 concludes the paper.

# 2 Understanding time varying TEC

Consider a hypothetical temperature-electricity curve (TEC) as represented in Fig.1. In this U-shaped curve, the minimum point is called the threshold. TEC is influenced by a large number of structural socio-economic developments, such as the growth in incomes, extent of electrification, energy efficiency improvements, cultural habits, and prevailing climatic conditions. Hekkenberg et al. (2009) argues that. over time temporal dynamics could influence the slopes as well as the threshold temperature of the TEC. For instance, increased internal heat gains in commercial buildings from increasing use of computers or decreasing tolerance for heat, leads to a general shift towards lower heating demand and higher cooling demand. Neglecting a downwards shifting threshold temperature results in the underestimation of the electricity demand resulting from a temperature increase. On the other hand, ignoring an upward shifting threshold temperature results in the overestimation of the electricity demand.

With increasing electricity access and rising income level, the number of households owning temperature control devices (such as air conditioners and air coolers) is increasing very rapidly in India. According to National Sample Survey Organization surveys (50th, 61st, 66th) the number of households owning an air cooler or an air conditioner doubled from 32.9% in 1993 to 60% in 2009 in urban Delhi (which represents 97.5% of total Delhi population as per census 2011) and increased from 20.6% to 26% in rural Delhi. In the case of refrigerators, the upward trend was even more impressive; saturation went from 29% in 1993 to 61.3% in 2009 in urban Delhi and from 17.7% to 38% in rural areas.

As per 61st (2004-05) NSSO survey ( which provides ownership of air coolers and air conditioners separately unlike other rounds) only 9% have access to air conditioners and 58% to air coolers in Urban Delhi. With growing incomes there is a very high probability that total air conditioning electricity demand could increase substantially. Further, with higher affordability sensitivity of households to higher temperatures is likely to increase which may further shift the location of the minimum point of the TEC (representing balance temperature with minimum comfort related heating and cooling demand). For instance, higher income households may want to switch on their air conditioners when average temperature is $19^0C$ in 2015 as compared to $22^0C$ in 2000.

According to a study (Kothawale et al.(2010)) done at IITM 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, significant warming trends in mean temperature were observed in two seasons characterized by high humidity-monsoon and post monsoon. Increasing night temperatures in these humid seasons could have signinificant implications for the usage of air conditioners and thus electricity demand. As market saturation of air conditioners is currently quite low the response of its diffusion (with growing incomes) to 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 will respond to global warming.

## 3 Relevant Literature

The simplest way to estimate a U-shaped TEC is to use a regression model that is quadratic in temperature. However, such a model assumes a symmetric relationship in the sense that 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), Mirasgedis et al. (2004)).

Nonetheless, a linear parametric model can still be used to estimate a non-linear relation by using the degree day approach (Douglas (1981), Al Zayer (1996), Sailor and Munoz (1997), Valor et al. (2001), Sailor (2001), Pardo et al. (2002), Mirasgedis et al. (2007)). This approach defines heating degree days (HDD) and cooling degree days (CDD). HDD and CDD quantify difference between the daily mean temperatures above or below a threshold temperature ( $18^0$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 TEC with ordinary least squares regression model using annual, monthly or daily data in the following manner:

$$\begin{aligned} e_{td} &= \beta_0 + \beta_1 trend + \beta_2 CDD_t + \beta_3 HDD_t + \beta_4 CDD_{t-1} + \beta_5 HDD_{t-1} + \\ & \sum_{k=1}^{11} \phi_k MONTH_{kt} + \sum_{b=1}^{6} \varphi_b WD_{td}^b + \beta_6 HOLIDAY_t + \beta_7 X_t + \varepsilon_t \end{aligned}$$

where $e$ is the demand for electricity on day $d$ of year $t$, $WD$ 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 $\varepsilon$ is the residual term. Although this approach estimates separate linear relationships of electricity demand due to heating and cooling demand, it relies on an arbitrary choice of threshold value ($18^0$C in most cases).

More recent papers such as Carcedo and Otero (2005) and Bessec and Fouquau (2008) 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:

$$e_t = \beta_0 + \beta_1 trend + \beta_2 (trend)^2 + \beta_3 (trend)^3 + \sum_{b=1}^{6} \varphi_b WD_{btd} + \beta_4 H_t + \beta_5 X_t + \beta_6 g(T_t; \gamma, c) + \varepsilon_t$$

where $g(T_t; \gamma, c)$ is a function of the temperature $T_t$ that allows a transition from a cold to a warm regime. In the literature, the transition function has been specified in different ways such 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 this literature.

Non-parametric methods, also known as smoothing models, have therefore been used to achieve greater flexibility in functional form. To estimate the functional form from data, such models replace global estimates of the electricity-temperature function with local estimates. For local estimators, a regression is estimated 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 $T$. This series of local estimates is then aggregated to summarize the relationship between the two variables. This resulting nonparametric estimate does not impose a particular functional form on the relationship between $E$ and $T$, and thus minimizes specification errors (Powell, 1994; Keele, 2008). 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 nonparametric regression models that rely on local estimates to estimate functional forms from data. Engle et al. (1986) estimated impact of weather on electricity sales of four US utilities with smoothing splines using monthly data for 7-8 years. The semiparametric partial linear regression model estimated by them is given by

$$\mathbf{E} = \mathbf{Z}\boldsymbol{\gamma} + \mathbf{f}(T) + \boldsymbol{\varepsilon}$$

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. Semiparametric model consists of a parametric and a nonparametric part at the same time. A fully nonparametric model is computationally complex in presence of numerous predictors. Some of the papers using semiparametric regressions to model electricity demand-temperature relationship are Hyndman and Fan (2008), Harvey and Koopman (1993) and Henley & Peirson (1997).

The only study that estimates the temperature-electricity relation for India is Ramesh et al. (1988) that estimated the impact of weather variables on peak electricity load for Delhi using ordinary least squares parametric regression separately for summers and winters, during the period 1980-1985. It should be noted that despite the fact that the relationship between electricity demand and climatic conditions in Delhi was investigated in the past for peak demand forecasting, this is the first study that focuses on non-linear time varying impact of climate on electricity demand using semiparametric variable coefficient model.

# 4 Estimation strategy

## 4.1 The reduced-form model

I estimate four models. The first model is based on simple linear regression. The second specifies a semiparametric additive model using unpenalized splines and the third estimates a semiparametric additive model with penalized splines. The fourth model estimates a variable coefficient model, where the smooth of temperature index is interacted with year dummies to capture the time varying impact of temperature on electricity demand.

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

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

where $e$ is electricity demand on day $d$ of year $t$, $MAJH$ is a dummy variable that takes 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,[1] $RAIN$ represents daily rainfall in millimeters(mm). $WD$ is the set of six day dummies to describe weekly periodicity of electricity demand (Wednesday is taken as the reference day). $\mathbf{y}$ is a set of nine year dummies (with 2000 as base year) to identify deterministic long term trend connected with the impact of demographic, technological, and socio-economic factors such as prices, urbanization, growing number of air conditioners and coolers on the 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 (1) can be rewriten in the following form

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

[1] A major holiday is one that is declared to be a holiday for all government employees (on account of national events or religious events). In addition, government employees are entitled to select 2 additional days of holidays from a list of holidays for minor religious festivals.

where, $E$ is an $n \times 1$ vector of electricity demands, $\varepsilon$ is an $n \times 1$ vector of errors, and $\mathbf{Z}$ is an $n \times p_1$ matrix of $p_1$ non-temperature predictors, $\gamma$ 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, $\eta$ is an $p_2 \times 1$ vector of coefficients of predictors in $\mathbf{T}$, $X$ is an $n \times p$ $(= p_1 + p_2)$ matrix of all predictors and $\beta$ is an $p \times 1$ vector of coefficients of $X$ predictors. The least squares and maximum likelihood estimator of $\beta$ is $\widehat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{E}$ and **Hat matrix** $H$ is an $n \times n$ matrix, such that $\widehat{E} = HE$. We can obtain $H = X(X^TX)^{-1}X^T$ and show that $trace(H) = trace(\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T) = tr(I_p) = p$ =estimated degrees of freedom (EDF) as measured by 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 transformation can overcorrect the nonlinearity between $E$ and $AT$ and thus, power transformations often cannot adequately capture the nonlinear relationship in the data. Model ll estimates a semiparametric model given by

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

Here, $\mathbf{f}(AT) = (f(AT_1), ....f(AT_n))'$ is an $n \times 1$ vector, where $f(AT)$ is an unknown smooth function i.e continuous and sufficiently differentiable function of $AT$. In this paper, I estimate $f(AT)$ by cubic regression spline using cardinal basis functions. Wood (2006) and Lancaster and Salkauskas (1986) gives full details. Such basis functions parameterize the spline in terms of its values at the knots and thus have advantages in terms of interpretability of the parameters along with good mathematical properties and numerical stability. It can be represented as a linear combination of the basis functions of regression splines. For instance,

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

where $b_j(AT)$ is the basis at the $jth$ 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 (3) becomes

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

where $\mathbf{X}$ is an $n \times (p_1 + (N-1))$ model matrix. One degree of freedom is lost due to identification constraint on $f(AT)$ i.e$\sum_{i=1}^{n} f(AT_i) = 0$. Based on Alkaike Information Criteria, I select twelve knots ($N = 12$ ) or eleven basis functions. Given knots, this model becomes a fully parametric model with an expanded model matrix and estimate predictor variable coefficients by minimizing $\parallel \mathbf{E}-\mathbf{X}\boldsymbol{\beta} \parallel^2$. The key limitation of this model is that the analyst must select the number of knots, and the location of these knots. Number of knots directly controls the degrees of freedom of a smooth term[2]. To deal with knot selection problem, I adopt penalized cubic spline approach. These models construct a penalty on $f()$ which will be large if $f$ is very wiggly and small if it is nearly flat. I estimate model III by adding a quadratic penalty as $\lambda\boldsymbol{\beta}^T\mathbf{P}\boldsymbol{\beta}$ and the following minimization problem is solved

$$\parallel \mathbf{E}-\mathbf{X}\boldsymbol{\beta} \parallel^2 + \lambda\boldsymbol{\beta}^T\mathbf{P}\boldsymbol{\beta} \tag{6}$$

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 (see appendix). $\lambda$ is the smoothing parameter that controls the trade-off between model fit and model smoothness. For $\lambda \to 0$ the minimization gives a wiggly function whereas letting $\lambda \to \infty$ gives a linear fit. The optimal $\lambda$ is selected by cross validation. It works as follows: for a given value of $\lambda$ , we omit the $ith$ observation from data and fit the penalized spline to this slightly truncated data set. I denote this prediction of $e_i$ as $\widehat{e}_{i-1}$. The model prediction errors are calculated, and this is repeated as each observation is droped in turn. The cross validation score is calculated as the average of the individual model prediction errors. One should choose the value of $\lambda$ 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 adjusted R-square GCV adjusts the average model prediction errors with the degrees of freedom (number of parameters estimated in the model). For penalized spline models, the GCV score is

$$GCV(\lambda) = \frac{\sum_{i=1}^{n} [e_i - \widehat{e}_i]^2 n}{[n - tr(H_\lambda)]^2}. \tag{7}$$

Minimizing $GCV(\lambda)$ with respect to $\lambda$ gives an estimate $\widehat{\lambda}$.Given $\lambda$ (6) is minimized w.r.t $\boldsymbol{\beta}$. We get $\widehat{\boldsymbol{\beta}} = \left[\mathbf{X}^T\mathbf{X} + \lambda\mathbf{P}\right]^{-1}\mathbf{X}^T\mathbf{E}$ and hat matrix $\mathbf{H}_\lambda = \mathbf{X}\left[\mathbf{X}^T\mathbf{X} + \lambda\mathbf{P}\right]^{-1}\mathbf{X}^T$.The trace of $H_\lambda$ , as in 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 basis dimension is not generally critical as actual effective degrees of freedom are controlled by $\lambda$. Number of knots should be selected to be

[2] In practice it is usual to use Alkaike Information Criteria for selecting optimal number of knots. A Model with a lower value of Alkaike Information Criteria is preffered

large enough to have enough degrees of freedom to represent the underlying true structure of data reasonably well, but small enough to maintain reasonable computational efficiency (Woods, 2006).

Model IV extends Model III by interacting $f(AT)$ by ten year dummies to capture time varying impact of climate on electricity demand. One can capture the time varying effect by estimating a seperate model (like model III) for each year. However, by pooling data for all 10 years I get more robust estimates for analyzing long term impact of climate on electricity demand. I first select number of knots for each year ( $N_t$) and corresponding basis functions $\mathbf{B}_t(AT)$ -

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

I select same 10 knots every year. Selected knots are $[k_0 = 12, 16, 20, 24, 28, 31, 33, 35, 37, 40 = k_{10}]\ \forall\ t$. $\mathbf{B}_t(AT)$ 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:

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

where, $\mathbf{f}_t(\mathbf{AT})$ is a vector of smooth function of the temperature index of year $t$ with dimension $n \times 1$.Here, $t$ indexes 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 $\mathbf{Y}$. $\mathbf{y}_t$ represents year dummy for year $t$. Degrees of freedom for $\mathbf{f}_t(\mathbf{AT})$ will be determined by the choice of $\lambda_t$. Note that same $\lambda_t$ is chosen for all years resulting in same degrees of freedom for each year. The fitting problem becomes:

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

where $X$ is an $n \times (p_1 + ((N-1) \times 10))$ model matrix (see appendix). Given $\lambda_t$, (**??** can be minimized w.r.t $\boldsymbol{\beta}$. We get $\widehat{\boldsymbol{\beta}} = \left[\mathbf{X}^T\mathbf{X} + \sum_t \lambda_t \mathbf{P}_t\right]^{-1} \mathbf{X}^T\mathbf{E} = \left[\mathbf{X}^T\mathbf{X} + \mathbf{K}\right]^{-1} \mathbf{X}^T\mathbf{E}$,with $\sum_t \lambda_t \mathbf{P}_t = \mathbf{K}$. Smoother matrix for penalized spines with interaction can be derived as $\mathbf{H}_\lambda = \mathbf{X}\left[\mathbf{X}^T\mathbf{X} + \sum_t \lambda_t \mathbf{P}_t\right]^{-1} \mathbf{X}^T = \mathbf{X}\left[\mathbf{X}^T\mathbf{X} + \mathbf{K}\right]^{-1} \mathbf{X}^T$.As discussed previously, one degree of freedom is lost due to identification constraint on $f_t(AT)$, which says $\sum_{i=1}^{N_t} f_t(AT_i) = 0\ \forall\ t$. From above electricity demand on a particular day is obtained as

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

where $e_{td}$ is electricity demand on day $d$ of year $t$. $z'_{td}$ is a row vector of parametric predictors for day $d$ of year $t$. The full form of eq(11) can therefore be written out as

$$e_{td} = \beta_0 + \beta_1 MAJH_{td} + \beta_2 MINH_{td} + \beta_3 RAIN_{td} + \sum_{t=1}^{10} f(AT_{td}) y_t + \sum_{t=1}^{9} \phi_t y_t + \sum_{b=1}^{6} \varphi_b WD^b_{td} + \varepsilon_{td} \tag{12}$$

As the errors from eq(12) are likely to be serially correlated, I carry out the following adjustment given in Li & Racine (2007) (chapter 18 section 18.2.2). By droping year dummies and estimating eq(12) separately for each year, $\widehat{\boldsymbol{\varepsilon}_d}$ for each $t$ is obtained. For each year $t$, a first order stationary auto-regressive model defined as

$$\boldsymbol{\varepsilon}_d = \rho_t \boldsymbol{\varepsilon}_{(d-1)} + \nu_d \tag{13}$$

where $\nu_d$ is white noise, is estimated . By regressing $\widehat{\varepsilon}_d$ on $\widehat{\varepsilon}_{d-1}$ of year $t$, an estimate of $\rho_t$ ($\widehat{\rho}_t$) is obtained. The model is then transformed to have serially uncorrelated disturbances by subtracting estimated previous day errors $\widehat{\varepsilon}_{d-1}$ from $e_d$ in the following manner:

$$e^*_d = e_d - \widehat{\rho}_t \widehat{\varepsilon}_{d-1} \tag{14}$$

By pooling estimated $e^*_d$ for each $t$, the final model becomes

$$e^*_{td} = \beta_0 + \beta_1 MAJH_{td} + \beta_2 MINH_{td} + \beta_3 RAIN_{td} + \sum_{t=1}^{10} f(AT_{td}) y_t + \sum_{t=1}^{9} \phi_t y_t + \sum_{b=1}^{6} \varphi_b WD^b_{td} + u_{td} \tag{15}$$

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

# 5 Data

## 5.1 Electricity Consumption and Shortage

The data on daily electricity consumption of Delhi has been obtained from the operator of the national electricity grid, the National Load dispatch centre[3]. In order to estimate the impact of global warming on electricity demand it is important to recognize that unlike developed countries (with high reliability of supply) electricity systems in India are continually inhibited with power shortages resulting in rationing and disrupted electricity usage pattern. When there are regular power failures, consumers are not able to consume the quantity they need, and they either substitute it with other alternatives such as diesel and kerosene, or resort to independent generation. As a result, electricity consumed reported by National Load Dispatch Centre (NLDC) is constrained electricity demand, and is equal to the electricity supplied by the utilities.

The unconstrained notional demand (sum of constrained and unmet demand) is only known for those periods during which the existing supply potential is in excess of demand. Thus, there is a demand function and a supply function, but demand is not always equal to the supply and the observed quantity is equal to the minimum of ex ante demand and supply quantity. In the past econometric disequilibrium models have been applied in various fields. In such situations, according to Fair and Jaffee (1972), the sample should either be separated into demand and supply regimes or the observed quantity should be adjusted for the effects of the rationing and then the demand and supply schedules should be estimated. In order to estimate true electricity demand of Nigeria (with low reliability of supply), Ojameruaye (1988) adjusted electricity consumption under rationing with a reliability index (based on frequency, time and power outage). In order to obtain unrestricted electricity demand for Delhi, I adjust daily total electricity consumption of Delhi with the observed daily shortage using daily electricity supply shortage data obtained directly from Delhi Transco Limited. Fig. 2 plots reliability index of electricity (electricity demand met as percentage of total demand including shortage). Apparently, the graph shows that there has been a significant improvement in the supply in post-2005 period.

## 5.2 Apparant Temperature and Rainfall

Data on all the climatic factors has been obtained from the website www.tutiempo.net/en/climate/India.?? This website gives the station wise data for all the major weather stations in India. I first

[3] The National Load Despatch Centre (NLDC) is a government body mandated to ensure integrated operation of the national power system.

constructed the apparent temperature index (AT) for Delhi using Steadman (1994) formula adjusting dry bulb temperature with humidity and wind speed. The formula is

$$AT_{td}\left({}^{0}C\right) = T_{td} + 0.33v_{td} + 0.07w_{td} - 4$$

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

where $T$ denotes average temperature in degree Celsius ($^{0}$C), $v$ denotes evaporation, $w$ denotes wind speed (m/s), and $h$ denotes relative humidity(%). Fig. 3 shows empirical density function of the apparent temperature by seasons, over two decades (1990-99 and 2000-2010). During this period, the empirical density function of the apparent temperature has shifted rightwards indicating increasing warming in all seasons.

# 6 Results: The effect of apparent temperature on electricity demand

## 6.1 Summary Statistics

In order to analyze the characteristics of the distributions of electricity demand and apparent temperature, the basic summary statistics are displayed in Table 2.Over the period, the average daily electricity demand (ED) increased from 49.76 MKWH in 2000 to 65.03 MKWH in 2009, with its maximum increasing from 64.6 MKWH in 2000 to 94.3 MKWH in 2009. At the same time, the standard deviation of the daily electricity demand increased from 6 MKWH in 2000 to 15 MKWH in 2009. During this period, the average daily apparent Temperature ranges from 26.45 and 27.7, with the peak occurring in 2009 & 2002, and trough occurring in 2005.

## 6.2 Main Results

Table 3 (a,b & c) summarizes the results of the estimated models. All models are estimated by the likelihood maximization approach or the penalized likelihood maximization (for Model III and Model IV) using mgcv package in R. For Model I and Model II usual frequestist approach is used to calculate standard errors and p-values for model coefficients. For Model III and Model IV Bayesian p-values and standard errors are reported. Wald test of significance of each parametric and smooth term are performed.

Goodness of fit daignostics shows that Model II is an significant improvement over model

I at 99% confidence levels. F-test based on residual values of the semiparametric model II and parametric model I yeilds an F statistics of 43.984, which has a p-value of .0. This implies that a local fit captures the nonlinearity between electricity demand and temperature much more accurately than the global fit of the parametric model I ( see the appendix).

Model III estimates penalized splines with 20 knots as compared to 12 knots used for unpenalized spline Model II. The results from Model III are not statistically different from model II. The F-test based on residual values of the model II and model III yeilds an F statistics of 2.04 and a p-value of 0.1229[4]. The advantage of using penalized splines is that the results are not influenced by the number of knots when fairly large number of knots are selected While in case of model II one has to proceed on trial and error basis. Since each knot represents additional parameters being added to the model, Eilers and Marx (1996) recommend using Akaike's Information Criteria (See the appendix for details).

Model IV when compared with Model III results in a significant improvement at the 99% confidence levels (with F statistics=59.9 and P-value=.000). Both GCV and AIC are much lower for model IV. It has a high adjusted R square of .938 implying it explains 93.8 % variation in electricity demand. The Durbin Watson statistic shows that the estimated model has no autocorrelation.

It is observed that rainfall has a significant negative impact on electricity demand, with 1 millimeter increase in rainfall reducing electricity demand by 0.05 MKWH. As expected, both holiday dummies turned out to be highly significant and negative. On a major holiday, electricity demand is estimated to be 3.32 MKWH lower than the average demand. A minor holiday, on the other hand, reduces demand by 0.42 MKWH. Estimates of parameters which model weekly cycle of electricity demand indicate that on Mondays, Saturdays and Sundays electricity demand is likely to be lower than the average level (with Wednesday as the reference day), and higher on Friday. These results might be expected as holiday and weekend loads have quite a different response to temperature than those on weekdays. Monday has a lower demand possibly due to the previous day holiday effect (also called holiday inertia) while Friday has a relatively higher demand, probably, due to the build up of the 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 the electricity demand is clearly non-linear. The estimated degrees of freedom (edf) for the temperature smooth term estimates ($\widehat{f_t}$ ) and

[4] The test statistic is defined as:

$$F = \frac{(RSS_{\text{smaller}} - RSS_{\text{larger}})/[df_{\text{res,larger}} - df_{\text{res,smaller}}]}{(RSS_{\text{larger}})/[df_{\text{res,larger}}}$$

their p-values support the hypothesis that the coefficients are statistically significant. Same smoothing parameter ($\lambda_t$) is chosen for all years resulting in equal degrees of freedom (approximately 6) for each year . Fig 5 plots all the estimated temperature dependence curves along with the Bayesian confidence intervals (dotted curves). Fig. 6 plots marginal effects for all years along with its 95% confidence intervals (dotted curves). Here we plot both types of confidence intervals obtained by normal plug in approach and Bayesian approach (See appendix).

Over time it is observed that the minimum temperature threshold is falling and the temperature dependence curves of Delhi are moving leftwards. During the period of analysis it has shifted from about 20-23$^0$C in 2000-2005 to about 17-21$^0$C in 2006-2009. This can be explained by the growing number of air conditioners and coolers with rising incomes. In other words, with higher affordability people's sensitivity towards hot temperatures is likely to increase, and they are expected to switch on cooling devices at a relatively lower temperatures.

In addition to the leftward shift of the TEC, it is observed that the rising part of the TEC is getting steeper over time, implying ever increasing cooling demand per unit increase in summer temperature. While the heating demand is declining per unit increase in winter's temperature the effect is much lower as compared to the increase in cooling demand in summers. For instance, a 1 $^0$C increase in temperature at 30 $^0$C increased electricity demand by 3.2 MKWH in 2009 as compared to only 1.2 MKWH in 2000. On the other hand, a 1 $^0$C increase in temperature at 15 $^0$C decreased electricity demand by .8 MKWH in 2009 as compared to .7 MKWH in 2000.

Previously, Carcedo and Otero (2005) estimated threshold transition model and found 15.5 $^0$C as the upper heating demand threshold and 18.4 $^0$C as the lower cooling demand threshold for Spain. The smooth transition model obtained 15.4 $^0$C as an optimal threshold temperature. Bessec and Fouquau (2008) found threshold temperature to be about 16 $^0$C for the whole sample of European countries and 14 $^0$C for the sample of cold European countries and 22.4 $^0$C for the sample of hot European countries. Although thresholds obtained in this paper are not directly comparable to previous studies (which are based on average temperature in contrast to apparent temperature used in the study), they give a fairly good idea about how threshold temperatures may vary both spatially and temporally with economic growth and cannot be assumed to be static.

# 7 Forecasting electricity demand and model evaluation

The above model has been used to forecast the impact of 1 $^0$C, 2 $^0$C and 3 $^0$C increase in apparent temperature on the electricity demand of Delhi in 2015 and 2021. Fig. 7 extrapolates marginal effects curves until 2021. It shows that threshold temperature is likely to shift from 21$^0$C in 2000 to 17$^0$C in 2021. The results are displayed in Table 4. A 1 $^0$C increase in apparent temperature (over average apparent temperature 2000-2009) increases net electricity demand[5] by about 405 MKWH (1.7%) in 2009 over its base electricity demand of 23809, 496 MKWH (1.82%) in 2015 over its base electricity demand of 27225 and 630 MKWH (2.05%) in 2021 over its base electricity demand of 30737. A 2 $^0$C increase in apparent temperature (over average apparent temperature 2000-2009) increases net electricity demand by about 822 MKWH (3.45%) in 2009, 1022 MKWH (3.75%) in 2015 and 1305 MKWH (4.25%) in 2021. Similarly, 3 $^0$C increase in apparent temperature (over average apparent temperature 2000-2009) increases net electricity demand by about 1248 MKWH (5.24%) in 2009, 1570 MKWH (5.77%) in 2015 and 2008 MKWH (6.53%) in 2021.

In addition Table 4 disaggregates the impacts by months. Higher temperature increases electricity demand in summers (led by April and May), monsoon (led by September) and post monsoons (led by October) and decreases demand in winters (led by January). It is observed that the maximum impact is likely to be felt in the hot month of April with average apparent temperature of 30 $^0$C, followed by, October and May. Marginal effect curve peaks at about 30 $^0$C indicating maximum sensitivity of electricity demand to temperature at this level. It is important to note that the average apparent temperature in April has shown maximum increase of 2.21 $^0$C in past years (see Table 5). Although a 1 $^0$C increase in temperature increases net electricity demand by 1.7% in 2009, demand increases by 4.2% in April, by 4% in October, by 2.3% in September and by 2% in May and March. On the other hand, a 1$^0$C increase in temperature decreases electricity demand by 1.5% in January, 0.5% in February and 1.2% in December. Since, electricity saved in winters cannot be stored and used in summers, global warming could result in serious disequilibrium in some of the months in the future.

To evaluate the forecasting performance, the actual demand of two years (2008-2009) has been compared with the predicted demand. In this evaluation predicted demand for the two years is calculated using coefficients of the estimated model (based on 2000-2007 data) and known temperatures and information on other drivers in these years. Data from the forecast period are not used for the model estimation. Fig. 8 illustrates the difference

[5] Net electricity demand increase means increase in electricity demand due to climate warming net of decrease in electricity demand in winters.

between observed and predicted electricity demand in year 2008, and 2009. These graphs demonstrate that the model predicts demand in both years remarkably well.

The above model can be used by electricity supply companies to predict electricity demand, both, in the short run (day to day basis) and in the long run (plan to plan basis). It is important to note that the electricity demand data used in the paper for establishing temperature-electricity correlation does not include electricity demand in industries out of captive generation and transmission and distribution losses[6] . Therefore, the projections of electricity demand are lower than the forecasts of the Seventeenth Electric Power Survey of India[7], which makes an adjustment for both in electricity demand data. For instance, the above survey forecasts total electricity demand (based on pre- 2005 annual data) for 2021-22 as 58759 MKWH as against 30737 MKWH in the current study (as no adjustment is made for transmission and distribution losses, and captive generation). Nonetheless, estimated temperature elasticities can be applied to the adjusted data for obtaining the impact of climate change under all three scenarios discussed above for managing and planning for future electricity supply .

# 8 Conclusions

Changing lifestyles and economic conditions make electricity demand increasingly more sensitive to temperatures over a period of years. This paper provides valuable insights regarding potential interactions between increasing cooling degree days and increasing incomes and the nature of resulting long term adjustments (such as higher stauration of air conditioners) in the electricity sector. The results from semiparametric variable coefficient model indicate that the variation in the slope of the TEC and threshold temperature is needed to allow for socio-economic dynamics in future electricity demand projections. An important contribution of the paper is the estimation of climate impacts by months. The model projects that the climate change can cause serious increase in future demand, particularly, during hot months such as May and April. These results can be extrermely useful for managing seasonal electricity disequilibrium situation in Delhi.

Further, the estimated threshold temperature in this paper can be used by HVAC[8] (Heating, Ventilation and Air Conditioning) designers to improve efficiency of electricity use. At

[6] On average, distribution and transmission losses account for about 30% of the electricity demand.

[7] The electric Power Survey Committee conducts surveys for the power demand, and makes demand forecasts for use in power sector planning.

[8] HVAC (Heating, Ventilation, and Air Conditioning) refers to technology of indoor or automotive environmental comfort. HVAC is important in the design of medium to large industrial and office buildings such as skyscrapers and in marine environments such as aquariums, where safe and healthy building conditions are regulated with temperature and humidity, as well as "fresh air" from outdoors.

present, the comfort standard practiced by HVAC designers is the same that is followed in the US for cooling the buildings (Air conditioned buildings). Large amount of electricity is consumed by HVAC systems in buildings and designing HVAC (for comfort) as per the changing climatic conditions in India can bring down the electricity demand drastically.

Important measures will be required to meet increased electricity demand due to climate change. For instance, there is a need to make a choice between fossil fuels and renewable sources for electricity generation. The results achieved in this work can be put to various practical uses by electricity production and sales companies, among which could be 1) understanding existing temperature-electricity sensitivity so as to manage risks related to unpredictable change in the energy demand under the extreme weather events, e.g., a heat wave, 2) quantifying the impact of projected climate change on electricity use and 3) forecasting required future capacity investments in the electricity sector.

Further, comprehensive assessment of impacts requires not only sound empirical research, but also more geographical coverage, especially in areas where severe climate change is likely to occur. Any work in the future should seek to extend the approach to other states within India to get an overall estimate of climate change on total electricity demand of India. It is extremely important to estimate state-wise TECs as states possess special socio-economic factors resulting in different TEC. Nonetheless, it is hoped that this analysis contributes to a better understanding of dynamic non linear TEC.

# Acknowledgements

First I would like to thank my advisors, Professors E Somanathan and Bharat Ramaswami, who have helped me tremendously at different stages of this paper. I also wish to thank Professor Probal Chaudhuri (ISI Kolkata), Dr. Swagata Nandi (ISI Delhi) and Professor Trevor Hastie (Stanford University) who have been kind enough to provide me with many helpful comments on the research as it developed. Dr. Simon woods (Bath University) has also been particularly helpful throughout the process in understanding the issues surrounding Generalized Additive Models (GAMs) estimation in R software with mgcv package. I thankfully acknowledge other contributors to the study- Dr. Abhiroop Mukhopadhyay, Dr. Indira Rajaraman, Dr. Nityananda Sarkar, Dr. Deepayan Sarkar, Mr. Rahul Mazumder, Dr. Ajay Mathur, Dr. TCA Anant, Mr. Prabir Sengupta, Dr Ligia Noronha and Ms Kiran Kumar- for their guidance and inputs.

Data to support the research in this study came from many sources. I am grateful to Mr. SK Sonee (CEO POSOCO), Mr. Ashok Rajput (CEA), Mr. Harish Rathore (Manager

NLDC), Mr Hemant Jain (Director CEA), Mr Ak Kaul (Director Delhi Transco) for their guidance and support in procuring electricity data. I am grateful to Mr. Shyam Kumar for his assistance in data operations. The earlier version of this paper has been presented at the Climate Policy Outreach meeting held at Xian in March 2011. The funding for the paper has been provided by the Planning & Policy Research Unit (PPRU) at ISI.

## References


[1] Al Zayer, J. and A.A. Al Ibrahim. 1996. "Modelling the Impact of Temperature on Electricity Consumption in the Eastern Province of Saudi Arabia." Journal of Forecasting, 15(2), 97-106.

[2] Amato, A.D. 2004. "Energy Demand Responses to Temperature and Implications of Climatic Change."

[3] Bessec, M. and J. Fouquau. 2008. "The Non-Linear Link between Electricity Consumption and Temperature in Europe: A Threshold Panel Approach." Energy Economics, 30(5), 2705-21.

[4] Dominici, F.; A. McDermott and T.J. Hastie. 2004. "Improved Semiparametric Time Series Models of Air Pollution and Mortality." Journal of the American Statistical Association, 99(468), 938-48.

[5] Eilers, P.H.C. and B.D. Marx. 2002. "Generalized Linear Additive Smooth Structures." Journal of Computational and Graphical Statistics, 11(4), 758-83.

[6] Engle F., Granger C.W.J, Rice J., Weiss A. (1986). "Semi parametric estimates of the relation between weather and electric sales". Journal of the American Statistical Association, 81(394), 310-120.

[7] Fair, R.C. and D.M. Jaffee. 1972. "Methods of Estimation for Markets in Disequilibrium." Econometrica: Journal of the Econometric Society, 497-514.

[8] Harvey, A. and S.J. Koopman. 1993. "Forecasting Hourly Electricity Demand Using Time-Varying Splines." Journal of the American Statistical Association, 88(424), 1228-36.

[9] Hastie, T. and R. Tibshirani. 1993. "Varying-Coefficient Models." Journal of the Royal Statistical Society. Series B (Methodological), 55(4), 757-96.

[10] Hekkenberg, M.; HC Moll and AJM Uiterkamp. 2009. "Dynamic Temperature Dependence Patterns in Future Energy Demand Models in the Context of Climate Change." Energy, 34(11), 1797-806.

[11] Henley, A. and J. Peirson. 1997. "Non Linearities in Electricity Demand and Temperature: Parametric Versus Non Parametric Methods." Oxford Bulletin of Economics and Statistics, 59(1), 149-62.

[12] Hyndman, R.J. and S. Fan. 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,

[13] Keele, L. Semiparametric Regression for the Social Sciences. Wiley Online Library, 2008.

[14] Li, Q., Racine, J.S., 2007. Nonparametric econometrics: Theory and practice. Princeton Univ Pr.

[15] Mirasgedis, S.; Y. Sarafidis; E. Georgopoulou; V. Kotroni; K. Lagouvardos and DP Lalas. 2007. "Modeling Framework for Estimating Impacts of Climate Change on Electricity Demand at Regional Level: Case of Greece." Energy Conversion and Management, 48(5), 1737-50.

[16] Moral-Carcedo, J. and J. Vicens-Otero. 2005. "Modelling the Non-Linear Response of Spanish Electricity Demand to Temperature Variations." Energy Economics, 27(3), 477-94.

[17] Ojameruaye, E.O. 1988. "An Econometric Disequilibrium Model for Electric Power System Planning in Nigeria." OPEC review, 12(4), 369-485.

[18] Pardo, A.; V. Meneu and E. Valor. 2002. "Temperature and Seasonality Influences on Spanish Electricity Load." Energy Economics, 24(1), 55-70.

[19] Peirson, J. and A. Henley. 1994. "Electricity Load and Temperature* 1:: Issues in Dynamic Specification." Energy Economics, 16(4), 235-43.

[20] Powell, J. (1994). "Estimation of Semiparametric Models." In: Engle,RF, and McFadden D.L (Eds). Handbook of Econometrics, Volume. 4. Elsevier Science B.V Amsterdam, pp 2443-2521.

[21] Ramesh, S.; B. Natarajan and G. Bhagat. 1988. "Peak Load Prediction Using Weather Variables." Energy, 13(8), 671-79.

[22] Ruppert, D.; Wand, M. P. and Carroll, R. J. Semiparametric Regression. Cambridge Univ Pr, 2003.

[23] Sailor, DJ and AA Pavlova. 2003. "Air Conditioning Market Saturation and Long-Term Response of Residential Cooling Energy Demand to Climate Change." Energy, 28(9), 941-51.

[24] Sailor, D.J. 2001. "Relating Residential and Commercial Sector Electricity Loads to Climate–Evaluating State Level Sensitivities and Vulnerabilities." Energy, 26(7), 645-57.

[25] Steadman, R.G. 1994. "Norms of Apparent Temperature in Australia." Australian Meteorological Magazine, 43(1).

[26] Valor, E.; V. Meneu and V. Caselles. 2001. "Daily Air Temperature and Electricity Load in Spain." Journal of Applied Meteorology, 40(8), 1413-21.

[27] Wadud, Z.; Noland, R. B. and Graham, D. J. "A Semiparametric Model of Household Gasoline Demand." Energy Economics, 32(1), pp. 93-101.

[28] Wangpattarapong, K.; S. Maneewan; N. Ketjoy and W. Rakwichian. 2008. "The Impacts of Climatic and Economic Factors on Residential Electricity Consumption of Bangkok Metropolis." Energy and Buildings, 40(8), 1419-25.

[29] Wood, S.N. 2006. Generalized Additive Models: An Introduction with R. CRC Press.

[30] Wood, S.N. and N.H. Augustin. 2002. "Gams with Integrated Model Selection Using Penalized Regression Splines and Applications to Environmental Modelling." Ecological modelling, 157(2), 157-77

[31] Yatchew, A. Semiparametric Regression for the Applied Econometrician. Cambridge Univ Pr, 2003.

Table 2-Summary Statistics for Elecricity Demand (E) and Apparent temperature(AT) 2000-2009

| Year | 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 |

# Tables

Table-3a Parameter estimates from Parametric and Semiparametric models

| Models | I | II | III | IV |
|---|---|---|---|---|
| Constant | 88** | 49.96** | 49.96** | 50.15** |
| | (1.45) | (.24) | (.24) | (.17) |
| 2001 | 2.22** | 2.29** | 2.29** | 2.08** |
| | (.28) | (.27) | (.27) | (.20) |
| 2002 | 4.7** | 4.8** | 4.8** | 4.7** |
| | (.281) | (.27) | (.27) | (.20) |
| 2003 | 5.9** | 6.2** | 6.2** | 5.88** |
| | (.28) | (.27) | (.27) | (.20) |
| 2004 | 8.5** | 8.64** | 8.64** | 8.45** |
| | (.281) | (.27) | (.27) | (.20) |
| 2005 | 10.11** | 10.27** | 10.27** | 10.11** |
| | (.28) | (.27) | (.27) | (.20) |
| 2006 | 12.58** | 12.86** | 12.84** | 12.57** |
| | (.28) | (.27) | (.27) | (.20) |
| 2007 | 11.75** | 12.07** | 12.05** | 11.80** |
| | (.28) | (.27) | (.27) | (.20) |
| 2008 | 12.69** | 13.09** | 13.06** | 12.81** |
| | (.28) | (.27) | (.27) | (.20) |
| 2009 | 16.14** | 16.30** | 16.30** | 16.03** |
| | (.28) | (.27) | (.27) | (.20) |
| Friday | 0.96** | 0.93** | 0.93** | 0.95** |
| | (0.23) | (0.22) | (0.22) | (0.16) |
| Monday | −0.58* | −0.54* | −0.54* | −0.53* |
| | (0.23) | (0.22) | (0.22) | (0.16) |
| Saturday | −0.99** | −0.98** | −0.98** | −0.94** |
| | (0.23) | (0.22) | (0.22) | (0.16) |
| Sunday | −3.62** | −3.55** | −3.55** | −3.56** |
| | (0.23) | (0.22) | (0.22) | (0.16) |
| Thursday | .28 | .24 | .25 | .33 |
| | (0.23) | (0.22) | (0.22) | (0.16) |
| Tuesday | −.11 | −.08 | −.08 | −.06 |
| | (0.24) | (0.22) | (0.22) | (0.16) |
| Major | −3.5** | −3.4*** | −3.4*** | −3.3** |
| | (0.32) | (0.31) | (0.31) | (0.22) |
| Minor | −.38 | −.27 | −.27 | −.42 * |
| | (0.26) | (0.25) | (0.25) | (0.18) |
| Rainfall | −.043** | −.052** | −.052** | −.052** |
| | (.01) | (.01) | (.01) | (.01) |
| AT | −5.91** | | | |
| | (0.19) | | | |
| $AT^2$ | .21** | | | |
| | (0.01) | | | |
| $AT^3$ | −.002** | | | |
| | (0.0001) | | | |

Notes:

1)Dependent variable for all models is electricity demand
2)Standard errors of coefficients are reported in parentheses
3) Wald test p-values significance is reported-
** significance of p-value at 99% significance level
* Significance of p-value at 95% significance level

Table 3b. Goodness of Fit Diagnostics

| | Models | | | |
|---|---|---|---|---|
| | **I** | **II** | **III** | **IV** |
| Adjusted $R^2$ | 0.873 | 0.884 | 0.884 | 0.938 |
| AIC | 20031.9 | 19709 | 19709 | 17464.2 |
| GCV | 14.301 | 13.09 | 13.091 | 7.09 |
| Durbin Watson Statistic | | | | 2.01 |
| N | 3643 | 3643 | 3643 | 3643 |
| Model DF (degrees of freedom) | 22 | 30 | 26.6759 | 81.316 |
| Residual DF (N-DF) | 3621 | 3613 | 3616.33 | 3561.68 |

Table 3c. Approximate significance of smooth terms in Semiparametric Models

| | EDF | F | p-value |
|---|---|---|---|
| **Model II** | | | |
| f(AT) | 11 | 1856 | 0.000 |
| **Model III** | | | |
| f(AT) | 7.67 | 2399 | 0.000 |
| **Model IV** | | | |
| f(at):2000 | 6.203 | 225.8 | 0.000 |
| f(at):2001 | 6.168 | 281.8 | 0.000 |
| f(at):2002 | 6.288 | 375 | 0.000 |
| f(at):2003 | 6.226 | 318.7 | 0.000 |
| f(at):2004 | 6.186 | 447.1 | 0.000 |
| f(at):2005 | 6.343 | 470.2 | 0.000 |
| f(at):2006 | 6.19 | 716.8 | 0.000 |
| f(at):2007 | 6.323 | 789.6 | 0.000 |
| f(at):2008 | 6.182 | 774.6 | 0.000 |
| f(at):2009 | 6.207 | 1410.7 | 0.000 |

Table 4 -Estimated impact of increase in AT on electricity consumption under three different scenarios (with 2009, 2015 and 2021 coefficients)

| | 2009 | | | 2015 | | | 2021 | | |
|---|---|---|---|---|---|---|---|---|---|
| **Base Electricity Demand*** | 23809 | | | 27225 | | | 30737 | | |
| Month | +1$^0$c | +2$^0$c | +3$^0$c | +1$^0$c | +2$^0$c | +3$^0$c | +1$^0$c | +2$^0$c | +3$^0$c |
| JAN | -24 | -44.2 | -63.9 | -22.9 | -41.2 | -58.2 | -22.5 | -39.2 | -53.7 |
| FEB | -7.7 | -9.7 | -8.8 | -2.4 | 1.4 | 8.2 | 3.5 | 13.7 | 27.3 |
| MAR | 32.6 | 72.1 | 124.4 | 44.2 | 97.9 | 168.6 | 56.6 | 125.6 | 216.8 |
| APR | 81.1 | 157.2 | 227.8 | 99.9 | 191.1 | 274 | 127.8 | 243 | 346.2 |
| MAY | 52.8 | 103.3 | 153.8 | 56.1 | 111.2 | 167.3 | 63.9 | 128.2 | 195.4 |
| JUN | 46.5 | 92.6 | 134 | 51.4 | 106.8 | 159.4 | 61.6 | 130.7 | 196 |
| JUL | 48 | 92.5 | 132.8 | 55.4 | 112.1 | 166.5 | 68.5 | 139.9 | 207.7 |
| AUG | 49.2 | 97.3 | 139.9 | 52.9 | 110.2 | 164.2 | 63.1 | 134.7 | 201.7 |
| SEP | 51.7 | 101.1 | 150.2 | 55.1 | 109 | 163.4 | 63 | 126.2 | 191.3 |
| OCT | 77.2 | 153.6 | 230.1 | 97.2 | 192.9 | 287.2 | 124.1 | 246 | 365.1 |
| NOV | 16 | 37 | 68 | 23.3 | 52.7 | 95.2 | 31.1 | 69.5 | 124.6 |
| DEC | -18.5 | -31 | -40.3 | -14.4 | -22.1 | -25.7 | -10.4 | -13 | -10.3 |
| Total | 404.7 | 822 | 1248 | 495.9 | 1022 | 1570.1 | 630.3 | 1305.4 | 2008.3 |
| Total as (%) of base electricity demand in the same year | 1.70% | 3.45% | 5.24% | 1.82% | 3.75% | 5.77% | 2.05% | 4.25% | 6.53% |
| Total as % of base electricity demand in 2009 | 1.70% | 3.45% | 5.24% | 2.08% | 4.29% | 6.59% | 2.65% | 5.48% | 8.43% |