Forecasting Demand for Electricity: Some Methodological Issues and an Analysis

Pillai N., Vijayamohanan

Centre for Development Studies, Prasanth Nagar, Ulloor, Trivandrum, Kerala, India

29 May 2008

Forecasting Demand for Electricity:
Some Methodological Issues and an Analysis

Vijayamohanan Pillai N.

Associate Professor
Centre for Development studies
Trivandrum – 695 011
Kerala
India

e-mail: vijayamohan@cds.ac.in

Forecasting Demand for Electricity:
Some Methodological Issues and an Analysis

Vijayamohanan Pillai N.

Abstract

Electricity demand projection is of utmost importance as electricity has become a vital input to the wellbeing of any society, driving the demand for it from an ever-expanding set of diverse needs to grow on an increasing rate, which in turn places increasing demands on scarce resources of capital investment, material means, and man-power. More specifically, the continuing 'energy crisis' has made crucial the need for accurate projection of electricity demand; hence the importance of the forecasting methods. The present paper critically evaluates the electricity demand forecasting methodology and proposes a methodology in the classical time series framework.

Forecasting Demand for Electricity: Some Methodological Issues and an Analysis

Vijayamohanan Pillai N.

1. Introduction

"Soviets + Electricity = Socialism", that was one of Lenin's development slogans. And that reflects the significance of electricity as a vital input to the wellbeing of any society and that is why the demand for it from an ever-expanding set of diverse needs is growing on an increasing rate. This in turn places increasing demands on scarce resources of capital investment, material means, and man-power. Prognosis and forecasting of electricity consumption has thus become a significant element of utmost necessity of the planning exercise in the power sector. More specifically, the continuing 'energy crisis' has made crucial the need for accurate projection of electricity demand; hence the importance of the forecasting methods.

In what follows we critically evaluate the electricity demand forecasting methodology and propose a methodology in the classical time series framework. This paper is divided into four sections including this introduction. The next section presents a brief theoretical discussion on forecasting and demand analysis, followed by the third section on the methodological issues involved. In the last section are presented the empirical results of our attempt to forecast electricity demand in Kerala in a simple, objective and theoretically sufficient manner.

2. Electricity demand: Forecasting Methods

The forecasting methods used for electricity demand in general may be divided into formalized and non-formalized methods.[1]

The first set includes extrapolation and correlation (regression) methods. For more distant periods of time in which considerable changes of the structure of the power sector must be considered, attempts have been made to use non-formalized methods such as some variants of Delphi method.

In the case of the formalized forecasting methods, two approaches may be distinguished in their scope:

i) an approach in which we try to penetrate the internal structure, and the internal and external linkages of the observed object and to explain its response to input impulses; and

ii) a statistical approach in which the object is treated as a 'black box' whose internal workings are unknown.

The first approach is partly represented by the method of structural analysis. It proceeds from a relatively detailed description of the most important parts of the power system, from energy sources through processes of refining and of transforming individual forms of energy, to the final consumption of energy. This method, however, is used to a limited extent only.

More common are the statistical approaches that take the object as a 'black box' and try to explain its mechanism on the basis of the interconnections of the individual elements of the observed path of the system. Here the analysis of time series and one-dimensional or multi-dimensional correlation (regressions) are used.

---

1 This section largely draws on Pillai (2001).

Time series analysis is used to predict the future path of the system by extrapolation on the basis of past observations. However, though the individual mathematical functions may characterize the past development of the system well, extrapolation of future trends is very risky and the mathematical function to be used must be chosen very carefully.

Extrapolation Models

The extrapolation of energy demand may in general be carried out using a number of mathematical functions such as:

i) linear extrapolation : $y_t = a + bt$;

ii) parabola (second degree) : $y_t = a + bt + ct^2$;

iii) exponential line : $y_t = ae^{bt}$;

iv) k-transformation : $y_t = (a + bt)^{1/k}$;

v) Growth curves; and

vi) (First order) auto-regressive trend model: $y_t = a + by_{t-1}$.

A variation of the previous model is the logarithmic auto-regressive trend model:

vii) Logarithmic auto-regressive trend model: $\log y_t = a + b \log y_{t-1}$.

Growth curves include logistic function and Gompertz function.

a) Logistic function : $y_t = L / (1 + ae^{-bt})$; and

b) Gompertz function : $y_t = L \exp(-ae^{-bt})$;

where L is the prescribed upper limit, and 'a' and 'b' are the parameters to be estimated.

In linear extrapolation, the variable to be forecast, y_t, is linearly plotted against time (t), and the resulting plot is extrapolated into reasonable future time spans. The parameter 'b' gives the rate of change (slope) of the line, and dividing the rate of change coefficient by the average value of y_t gives an average (arithmetic) growth rate per time unit. While in a linear trend the rate of change is constant, in a second degree polynomial (parabola), it increases linearly with time as 'b + 2ct', 'c' giving the acceleration coefficient. In extrapolation using exponential trend, the logarithm of y_t is plotted against time. These semi-log plots, which are frequently linear, are then extrapolated into the future to make forecasts. In this case, the parameter 'b' directly gives the (exponential) growth rate of y_t . In k-transformation trend method, y_t values are transformed using an appropriate power coefficient 'k' lying between zero and unity. (If k = 1, we get a linear trend.) The growth rate of this function is obtained by dividing the 'b' coefficient by the product of 'k' and the linear trend, 'a + bt' (= y_t).

Growth curves are used to predict the time path of a variable for which there is a limit. The curves trace the time path of the variable in an 'S'-form; and range from zero at 't = minus infinity', to the upper limit, L, at 't = plus infinity'. The Gompertz curve, however, is not symmetrical, while the logistic one is. With the Gompertz curve, the growth in the variable in the initial stages is comparatively faster than with the logistic curve

The first-order autoregressive trend model is generally useful for only short-term forecast, since long-term forecast obtained from this model is nothing but the stationary mean of the series. A variant of this model is the logarithmic (log-linear) autoregressive trend model. In these models one has the option of setting the intercept term (a) = 0; then 'b' in the simple autoregressive model represents the rate of change of the series y_t , and in the logarithmic model, the compounded rate of growth of the series. Both linear and compounded extrapolations based on these two autoregressive models are commonly used as a simple means of forecasting. Note that these models involve regression with a lagged dependent variable. If the additive error process is serially correlated, the coefficient estimates will be inconsistent.

It should be remembered that these simple extrapolation methods are at best useful as only a quick way for casual forecasts, as they usually provide little forecasting accuracy. It is advisable to consider some forecast statistics such as standard error of forecast, root-mean-square simulation error, etc. The time series analysis in autoregressive model may be extended to more sophisticated ARIMA (autoregressive-integrated-moving average) model for forecast.

In any given case, such regression functions need not be equally convenient. For instance, in the case of electricity consumption, numerous analysis have revealed that three very different time intervals can be defined for the long-run development of individual countries. The first corresponds to low values of energy consumption per capita and is marked by a considerable variation in annual increments. After having reached a certain value of the per capita consumption, the development becomes steadier and its trend begins to conform to an exponential pattern. Annual increments stabilize and assume a normal distribution. Having achieved a certain development level, the development gradually slows down. That is why the extrapolation requires the utilization of functions with decreasing annual increments.

The time trend analysis explains only the most important basic components of the development. The explanation of other regular residuals about the line of the trend requires the application of correlation (regression) methods, mainly those of a multi-dimensional nature. To a certain extent, they also help consumption forecasts for more distant periods of time.

Using correlation (regression) methods, we try to consider the influence of internal and external factors affecting energy consumption, such as:

a) demographic factors: population development, manpower development, etc.;

b) level of economic activity: total national product (income), material consumption, size of public sector, price level, etc.;

c) climatic conditions: number of frosty days, degree days, etc.

The relation between electricity consumption and social-economic factors may be described in terms of either simple linear models or log-linear models. In some cases incremental method, with all the variables taken in increments, also is used.

It should be pointed out that the use of such models is connected with the prognosis of the independent variables. This in turn may involve macro-econometric modeling.

Econometric Studies

A classical survey of the studies on the demand for electricity (in the U. S.) was given by Taylor (1975) in the Bell Journal of Economics, and it was later on updated and extended to natural gas, heating fuels, diesel and aviation fuels, coal, and gasoline (Taylor, 1977). In summarizing the empirical results on the demand for electricity in his survey paper in 1975, Taylor concluded:

(a) The price elasticity of demand for electricity, for all classes of consumers, is much larger in the long run than in the short run.
(b) This holds for the income elasticity of demand.

(c) The long run price elasticity of demand is indicated to be elastic.

(d) The evidence on the magnitude of the long run income elasticity is much more mixed. Estimates range from 0 to 2, and clearly depend on the type of model employed.

No econometric study of electricity demand had dealt with the decreasing block pricing in a completely satisfactory way, and the estimates of price (as also income) elasticities probably contained biases of indeterminate sign and magnitude as a consequence. Taylor's suggestions (1975) to deal with this problem were two-fold:

(a) Multi part tariffs require the inclusion of marginal price and intramarginal expenditure as arguments in the demand function, and

(b) The prices employed should be derived from actual rate schedules.

All the studies (in the US) had used either *ex post* average prices or prices derived from *Typical Electric Bills*, an annual publication of the US Federal Power Commission. In response to Taylor's suggestions, most of the studies since then have utilized 'the wisdom of employing electricity prices from actual rate schedules'. Several studies have also sought to improve modeling of the dynamics of electricity demand through inclusion of stocks of electricity consuming appliances in the demand function, and also the possibilities of inter-fuel substitution. Some other studies have utilized date on individual households, small geographical areas, or the area served by a utility, in a bid to utilize a data set of higher quality than that provided by data at the state, or national level, as well as to avoid (or at least reduce) aggregation bias in estimates of price and income elasticities.

To account for the dynamic characteristic of demand, a lagged dependent variable is usually used as a regressor in the log-linear model with a partial adjustment mechanism (Koyck distributed lag with geometrically declining weights). This specification facilitates to distinguish between short run and long run elasticities. Thus while the coefficient of the price variable in this model represents the short run price elasticity of demand, the long run price elasticity is obtained by dividing the short run coefficient by one less the coefficient of the lagged dependent variable used as a regressor, i.e., by the rate of adjustment. However, the presence of the lagged dependent variable, as already noted, makes the OLS estimator inconsistent due to the possible correlation between the lagged endogenous variable and the random variable, as well as the serial correlation among the successive values of the latter.

Demand Forecasting in India

At the all-India level, forecasts of electricity requirement and demand are made by the Planning Commission and by the Annual Electric Power Surveys (APS) convened by the Ministry of Energy, the Central Electricity Authority (CEA) being the Secretariat to the APS. These are not independent forecasts to the extent that the two bodies do have extensive discussion that usually leads to a reconciliation of results. Still, subtle differences exit between the methodologies employed by them.

The Planning Commission projects electricity demand as part of its macro-economic analysis for all the sectors of the economy. Industrial requirements are derived for a set of 'major' industries – mostly large scale or very electricity intensive – by applying consumption norms to production targets. In the Draft Plan, consumption in the rest of the industrial sector is assumed to be proportional to that in the major industries (60 per cent of that for major industries). Railway and irrigation requirements are also estimated by using sectoral targets and norms of electricity use. Consumption in other sectors – domestic, commercial, public lighting, water works, and miscellaneous – is based on trend growth rates or on regression analysis that relates sector growth rates to electricity demands. Using the input-output model, the Planning Commission is able to check the consistency of the resultant macro-forecast with its expectations for the economy at large.

On the other hand, the APS begins with a detailed survey of industrial establishments that demand 1 MW or more of electricity to solicit their views on how rapidly their requirements for power will increase. The methods used to derive forecasts for other consumer categories are much the same as those employed by the Planning

Commission – many of the coefficients relating output to electricity are identical – but with two differences. Regression analysis as opposed to trend analysis has not been used, and the exercise is conducted state-wise and then region-wise before being aggregated to the national level. The views the APS forms about the prospects for power in each state form the basis for the involvement of the SEBs and state governments in the demand forecasting exercises. Since the forecasts underpin the state-wise investment programs and these in turn influence the case for Central Plan assistance, there is a fundamental concern with the APS position.

Both the Planning Commission and the APS rely heavily for forecasting on the targets for other sectors – major industries, agriculture (irrigation), railway traction – and then use electricity consumption norms to derive electricity forecasts. Thus, they are both often characterized as end-use forecasting methods. The major difference is that the Planning Commission's analysis begins with macro-economic considerations and the APS' with completely disaggregated data. Forecasts based on the end-use methods, in the sectors where end-use coefficients are applied, have the advantage that they are not affected by uncertainties about the extent of shortages, past or present. But the regression and trend analyses of the APS and Planning Commission approaches are affected, primarily in the estimation of the regression coefficients and secondly in the choice of base period values. The Planning Commission approach has the virtue of strong consistency with the overall Plan, but the disadvantage of not providing clues to levels of consumption in individual states or in particular load centres. The APS, while providing disaggregated data, suffers from its length of preparation and the considerable cost involved in organizing a detailed survey of so many units throughout India. Furthermore, it has been found that the APS may often be upwardly biased. The APS forecasts exceed the demand met by between 20 and 80 per cent, and the divergence generally increases in the later years, as might be expected. Thus the energy consumption forecasts for Kerala by the successive APS since the 12th APS, for 1994 are in the order of 12466, 9328, 9409, and 8567 million units (MU) respectively (by the 12th, 13th, 14th, and 15th, the latest, APS). It should, however, be pointed out that it does no good to compare these forecasts with the actual demand met (about 7027.7 MU of energy internally sold) in Kerala, fraught with severe power cuts and load shedding. One way to account for such upward divergence is to regard it as reflecting the unsuppressed demand more faithfully than the realized demand.

Demand Projections: Kerala

Demand projections for Kerala based on the above (12^{th}, 13^{th}, 14^{th} and 15^{th}) APS results are in consideration now in the state. A steady decrease in the peak demand/energy consumption requirements is discernible in each of these forecasts that is attributed to some restrictions and revisions in the trends relative to the base year, (reflecting the increasing quanta of suppressed demand due to lack of generation capability). The State has accepted the 14^{th} APS as 'more dependable' (Government of Kerala, 1997: 12); whereas the Balanandan Committee (to study the development of electricity in Kerala) finds the 15^{th} APS 'as the better estimates for future planning' (Government of Kerala, 1997: 37). Considering the divergences in these forecasts of the APS for Kerala (as shown above, for example, for 1994), the State Planning Board constituted a working group to study the demand forecasts for Kerala. The committee used a log-linear model and growth rates of 4.72, 10, and 15 per cent for the HT and EHT industries to arrive at three different demand projections. The domestic demand projections in all these exercises were based on the growth of population as per the Census report. For the other sectors, the projections were made based on the trends (using semi-log scale). It should be noted that the energy demand forecast for 1994 by the Committee is only 8945 MU.

A number of computer software packages of energy planning models are available at present for energy demand forecasts, such as LEAP (Long range Energy Alternative Planning), BEEAM-TEESE (Brookhaven Energy Economy Assessment model-TERI Economy Simulation and Evaluation), MEDEE-S, ELGEM, etc. International Energy Initiative (IEI), Bangalore, has put forward a Development Focused End-Use Oriented, Service directed methodology (DEFENDUS) for estimating demand and supply of energy in an energy system (see, Amulya Kumar N. Reddy 1990), and an exercise based on this has been done for the KSEB. This methodology, with its twin focus of developed living standard and improved end-use efficiency, seeks to estimate demand for a particular energy source/carrier in a given year based on two variables – the number of energy users and their actual energy requirement in any base year as well as the expected changes in the subsequent years. The total energy demand is then equal to the aggregate demand of all the categories of users for every end-use.

The trend analysis is simple and viable, and logical enough also in that the time trend incorporates the effects of the technological advance on the supply side (that facilitates availability upon demand) and the progress in requirement level on the demand side over time. The latter may in general be approximated to the improvement in the living standard of the customers, and to this extent the trend analysis is sufficient in itself. However, from the perspective of a more general and comprehensive social-economic framework, a closer examination of the influence of other, more immediate determinants is in order. The variables of primary significance here are income of the consumer, and price of energy he is to pay, along with the relevant social, demographic, environmental or climatic, and other economic factors. This in turn involves an econometric analysis.

3. Some Methodological Issues

An unfortunate consequence of the widespread fascination for econometrics in the context of the Indian power sector analysis has been the fatality of a consistent, logical approach to power demand analysis, in the particular confines of its objective environment. Correlating aggregate electricity consumption and gross domestic product in the industrialized, advanced, countries where electricity service contributes significantly to everyday, social and work, life has become a standard tool of simple analysis for some obviously general conclusions and is beyond any methodological error as such there. However, the story is different and turns vicious with attempts at mapping this methodology on to an alien range in an underdeveloped power system where the contribution of the service of electricity is insignificant. This is so even in the industrial sector in India. Foe example, in 1994-95, the cost of fuels, electricity, and lubricants consumed in the production process in all the industries in the factory sector in India was just 9.64 per cent of the total inputs costs (and in Kerala, 6.01 per cent only). The cost of electricity purchased was only 4.02 per cent (and in Kerala, 2.85 per cent). Again, this was only 14.8 per cent of the net value added (in Kerala, 10.41 per cent). (Annual Survey of Industries, Factory Sector, 1994-95, Vol. 1, Tables 3 and 9.) In 1995-96, the percentage share of fuels, electricity, and lubricants consumed in the ASI factory sector of India in the value of total inputs was 9.56 per cent only and that in the value of products, 7.81 per cent; in Kerala, it was respectively 5.77 and 4.72 per cent only. (Statistical Abstract – India, 1998, CSO, GOI, Table 9.1, pp. 85-86.) During the 90s (1991-92 to 1997-98), power and fuel expenses of the whole manufacturing sector in India remained at about 6 per cent of the net sales and at about 7.5 per cent of the total production costs. The percentage share of electrical energy consumption in net sales as well as in total production costs (given in brackets below) in some of the major industries in India in 1997-98 is as follows - Textiles: 8.8 (10.2); Chemicals: 4.5 (5.8); Fertilizer: 13.7 (17.6); Pesticides: 4.4 (5.8); Polymers: 11.2 (14.2); Petroleum Products: 0.9 (1.2); Cement: 30.1 (44.0); Metal and Metal Products: 10.1 (12.1); Steel: 9.8 (11.6); Ferrous Metals: 9.1 (10.7); Non-Ferrous Metals: 15.3 (19.8); Machinery: 2.0 (2.6); Electrical Machinery: 3.0 (3.9); Paper: 22.6 (26.2). (CMIE, Industry – Financial Aggregates and Ratios, June, 1999, Company Finance.)

The methodology is questionable even in the industrialized sector power demand analysis in a less industrialized region like Kerala, that too with very limited number of electricity-intensive industries. Adoption of this methodology here then amounts to correlating the national/state domestic product or industrial product exclusively with an insignificant input, in violation of the ethics of a consistent and logical analytical exercise, and results in gross specification error. Moreover, there are a large number of small scale and cottage industries that use practically little electricity but together contribute significantly to the industrial product. The percentage share of the unregistered firms in the manufacturing sector's contribution to net domestic product (at current prices) in India was 35.4 per cent; in 1994-95, it was 34.99 per cent (and in 1970-71, 1980-81, and in 1990-91, it was respectively 46.66, 46.25, and 39.08 per cent). In the case of Kerala, about 41 per cent of the net State domestic product that originated in the manufacturing sector was contributed by the unregistered firms in

1997-98. It was found in 1994-95 that about 66 per cent of the enterprises in rural India and about 52 per cent in urban India) did not use any energy in their manufacturing process (Unorganized Manufacturing Enterprises in India – NSS 51st Round, 1994-95, NSSO, GOI, July 1998, p.ii). Only 7.9 per cent of the rural firms and 30.4 per cent of the urban firms in India (and 11.9 and 17.7 per cent respectively in Kerala) are reported to have used some electrical energy in their production process in that year (ibid., pp. 35-36). The contribution of the unorganized manufacturing sector, on the other hand, in terms of gross value added to the national economy in 1994-95 was estimated at Rs. 322748.9 million, out of which 41 per cent came from the rural sector; and that to the (Kerala) State's economy was at Rs. 6466.4 million with 72.07 per cent from the rural enterprises (ibid., pp. 57-64). It is worth noting that among the EHT consumers in Kerala only 37.5 per cent were consuming above 50 million units (MU) of electricity in 1995. All these are in addition to the problem of using aggregate variables that conceal everything of the characteristics of the units into which the analysis is paradoxically intended to make a look.

Despite this fact, the craze for econometrics has made a fetish of this methodology, correlating the less correlatable aggregates in almost all the energy demand studies at the national/State levels. The Fuel Policy Committee of India (1974), P. D. Henderson 91975), R. K. Pachauri (1975), Nirmala Banerjee (1979), World Bank (1979) J. K. Parikh (1981), and P. P. Pillai (1981) are some of the forerunners of this accurst tradition. P. P. Pillai (1981) analyses the aggregate as well as the sector-wise demand in Kerala considering   constant net domestic product, average price of electricity, time trend, and the one-period lagged dependent variable and computes short- and long-run elasticities.

P. P. Pillai's work, 'the outcome of an attempt for a quantitative study of the supply and demand aspects of electricity in Kerala using some known analytical tools' (as the author prefaces, p. ix), may be the first of its kind in Kerala. However, the simplicity of the approach whether in terms of aggregation or by means of mechanical adoption of econometrics has in fact detracted from its value of analytical rigour and useful exercise. Average price as used in this study as an explanatory variable in demand analysis especially using time series data contains dangers of aggregation bias as well as measurement error. Proper estimation of price elasticity of demand requires, (in conformity with the suggestion made by Taylor long back, 1975, 1977), the use of data on actual rate paid by a cross-section of consumers, rather than the aggregate average revenue (to the utility) over time; (pooled time series cross section data can also be used). Average revenue might be an indicator of the supply price in the aggregate, but never a representative of the demand price, the particular price a customer is faced with at a decision making juncture, especially in the context of the block rate tariff system. Moreover, the use of time series data simply ignores the possibility of changes of the intercept (that on an average accounts for the influence of factors other than those considered in the model) and of the slope of the line (that reflects the average intensity of energy consumption with respect to that variable); when using pooled data, it should also be checked that the intercept and slope have not changed significantly over time. One problem in using the time series average price is that it has the same (increasing) trend over time as electricity consumption has, thus misrepresenting price-demand relationship and the corresponding elasticity measure, unless the deflator series grows faster than the price series. Choice of a suitable deflator also poses problems. Thus, P. P. Pillai's use of non-deflated average price series must have ipso facto produced a positive price demand relationship, even though he reports negative price elasticity in all cases. This negativity could as well be due to severe multi-collinearity among the regressors, all having the same (increasing) trend along with time, used as one of the regressors, and the author may have unknowingly accepted it as true sign. Moreover, the use of OLS in the presence of lagged dependent variable as a regressor and of possible serial correlation may have marred the credibility of the estimates.

An apt example of mechanical adoption and use of econometrics against its grain usual in the academic circles is P. P. Pillai's Cobb-Douglas production function approach to Kerala's hydro-electric power system, with capital and labor as 'variable' inputs. It is common sense that labor is not at all a variable factor of production in hydro-electric power generation, it being a part of sunk capital, and capital is 'variable' only in the long-run in a power system, while production function is used to depict short-run stories, with *ceteris paribus* assumptions.

Even the very elasticity of demand analysis is open to serious questioning. The price elasticity of demand loses its relevance in an underdeveloped power system such as ours. Demand for electricity remains largely unresponsive or less responsive to its price as it has almost become a necessity in the absence of important fuel substitutes, and, at the same time, it commands a substantially lower budget share (due both to lower unit price and to low consumption level) in the case of most of the consumers. Thus for example, the share of fuel and power in the total private final consumption expenditure in the domestic market (at current prices) in India in 1997-98 was just 3.29 per cent; and electricity consumption accounted for only 0.68 per cent in the total (National Income Statistics, CSO, GOI). In 1980-81 and 1990-91, the share of fuel and power was 4.64 and 4.52 per cent respectively and that of electricity, 0.40 and 0.62 per cent respectively. The growth in electricity consumption has not been strong enough to facilitate a pronounced rate of substitution for other fuels, especially, the traditional one, kerosene oil. The percentage share of electricity in the private final consumption expenditure on total fuel and power grew from 8.63 per cent in 1980-81 to 20.57 per cent in 1997-98, marking an average annual compound growth rate of 5.24 per cent, while that of kerosene oil fell from 15.22 per cent to 10.5 per cent over the same period at a decay rate of (–)2.16 per cent p.a. This in turn suggests a very weak marginal rate of substitution of electricity for kerosene oil (or elasticity) of just about (–) 0.40; i.e., one percentage increase in the share of electricity consumption expenditure could on an average substitute for (or induce a fall of) 0.4 percentage in the share of kerosene oil consumption expenditure. In short, electricity could not yet make an effective inroad upon the economic life in India in general to the extent it should have done.

The results of the National Sample Survey on consumer expenditure also present the same picture of an input of insignificance. The percentage share of monthly per capita consumption expenditure on fuel and light in the total has been hovering about a virtually frozen range of 6-7 per cent over the whole second half of the last century in India. The expenditure share in 1995-96 (52nd round of NSS) for the rural consumers was just 7.1 per cent and for the urban ones, 6.4 per cent (Household Consumer Expenditure and Employment Situation in India, 1995-96, NSSO, GOI, Sept., 1998). The proportions were respectively 6.2 and 5.6 per cent in 1952-53 (5th round), 5.8 and 6.0 in 1960-61 (16th round), and 7.5 and 6.8 per cent in 1987-88 (43rd round). Implicit in this almost stationary consumption share scenario might be some stunning explanations of a stunted expression of Engel's law.

For a more concrete example, let us consider the case of the connected consumers themselves. The per capita electricity consumption of the connected domestic customers (that made up about 75 per cent of the total customers) in India in 1995-96 was 772.32 units at an average tariff rate of Ps. 95.94 per unit (for 18 State Electricity Boards), thus giving in general an average per capita electricity consumption expenditure of Rs. 740. 97 (or, Rs. 61.75 per month) – only 7.04 per cent of the per capita income (of Rs. 10524.8) of that year. In the case of Kerala State, the electricity consumption per (electrified) domestic consumer in 1997-98 was 953.72 units at an average rate of Ps. 76.96 per unit, that indicates an average domestic consumption expenditure of Rs. 733.99 (or Rs. 61.66 per month) on electricity. The domestic sector that made up about 75 per cent of the total customers nearly consumed 50 per cent of the electricity sold in Kerala in that year. In general, the consumption of electricity per connected consumer in Kerala in 1997-98 was 1480.71 units at an average tariff rate of Ps. 123.74 per unit, giving an average electricity consumption expenditure of Rs. 1832.16 (or, 152.68 per month) – only 15.35 per cent of the per capita income (of Rs. 11936) of that year. Similarly, as a substantial share of residential and commercial electricity consumption goes to serve basic need of lighting which is fairly unresponsive to income rather than to more income elastic, luxury end uses, power demand remains less income elastic also to this extent. Moreover, the whole edifice of demand analysis crumbles to dust in an encounter with power cuts and load shedding, that restrict actual consumption to availability rather than to actual requirement which is the long run experience of Kerala.

Till the turn of the 1980s, Kerala had been a power surplus state, exporting power to neighbouring states (Period I). Since the draught year of 1982-83, unprecedented power shortage has become a part of life in the state (Period II). Recurring draught coupled with inadequate installed capacity has thus unleashed a reign of power cuts and load shedding, constraining the actual demand down. Reliance on past demand data for forecasting purposes thus becomes grossly erroneous and highly questionable. If some measurement of these shortages is possible to be made, the constrained demand can be adjusted accordingly to arrive at a probable measure of unsuppressed demand, which in turn can be used as data base for forecasting. One method is to assume first that when restrictions are imposed on consumers, their level of consumption is held at some fraction of their consumption during an earlier base period. Then the shortfall in supply equal to these percentage restrictions can be found and inflated by a factor that reflects suppressed growth in demand since the base period and the impact of unscheduled load shedding. This in turn can be used to adjust the suppressed demand data (World Bank, 1979: 13). Another method uses as demand inflative factor, the fraction of customers affected by load shedding during peak period and thus deprived of chance to contribute to peak period demand. The main problem with all such methods is the non-availability of accurate data and information.

We have seen that the economic relationship demand is hypothesized to have with (per capita) income and unit price is weak and hence unwarranted in the case of an underdeveloped power system such as in India/Kerala. This then leaves us with a bare option of demographic variables to consider in the power demand analysis. However, with nearly 50 per cent of the households in Kerala (and nearly 60 per cent in the rural areas) remaining unelectrified (as per 1991 Census; the restrictions imposed on providing connections on account of the serious power shortage may not have altered this situation substantially in the later years), population may not be a sufficiently strong variable. The search for the appropriate power demand determinants may thus end in the domain of technical factors such as number of electricity consumers, etc., which are more immediate and direct causatives. The growth of demand for power is generally assumed to be determined by the growth of number of (connected) consumers and that of intensity of their power consumption (i.e., electricity consumption per customer), as also the interaction between these two factors. (See P. P. Pillai, 1981: 81 – 82; Henderson (1975) uses sectoral output in the place of number of consumers.) Another immediate factor of influence is connected load, the total of the rating (in kilowatts) of all the electricity using appliances installed on a consumer's premises. This also may be considered along with the relevant intensity of energy consumption (electricity consumption per kilowatt (KW) of connected load) and the interaction between the two. However, the demand function we specify correlates power (energy) demand with number of customers only. Other variables, connected load, consumption intensity, and interaction, are left out in view of the serious multicollinearity problem. Moreover, number of customers (N) is more immediate and direct than connected load (CL) in determining energy demand, as not only is N in fact the causative of CL, but also a customer may not use all his electric devices simultaneously or continuously; there are times, on the other hand, when *all* the consumers together exert demand pressure on the system. There is yet another significant reason. A growing power system is expected to become more and more electricity intensive in that its CL grows faster than N (so that the electricity intensification factor, i.e., connected load per customer (CL/N), increases over time). Despite the restrictions imposed on providing new connections since 1982-83, the domestic, commercial, and LT industrial consumers in Kerala have behaved in the expected line, becoming more electricity intensive (in terms of appliances installations), but the HT-EHT industry and 'others' (agriculture, public services, licensees, etc.) have not. This surprising tendency of a faster decaying electricity intensity in the State's HT-EHT industrial and agricultural sectors has overshadowed the normal growth in the other sectors and been reflected in the aggregate, the growth of CL trailing behind that of N. Thus, considered on all counts, N is preferred to CL in this study.

Any demand analysis in the context of the Kerala power system must distinguish between the Period I of no-power-shortage (till 1981-82), and the Period II of power cuts (since 1982-83), and hence adjust the supply–constrained demand data series of Period II upward in order to ensure its continuity and conformity. This we accomplish here by assuming that the unconstrained demand of Period II is basically a continuum of the demand function of Period I, but for changes in the demand intensity rate. That is, we assume that the Period I demand relationship continues unhampered into Period

II also, but with time-varying slope coefficient, which is obtained in effect by allowing the slope of the Period I demand function to grow over time at a rate at which the 'actual' rate of change of demand in Period II grows in relation to change in number of consumers.

The so-called 'Gulf boom' of increasing remittances of the non-resident Keralites from the Gulf has triggered an unprecedented growth of the housing sector and encouraged an increasing demand for electricity intensive appliances in Kerala especially since the mid-seventies. Number of houses in the electrified group must also have increased (in absolute terms) as a result of the social security schemes of the government (IRTC and IEI, 1996). Though the serious power shortage situation has however entailed restrictions on providing new connections since 1982-83, energy consumption in the domestic and commercial (as also in the LT industrial and agriculture) sectors have outgrown the number of connections, resulting in higher consumption intensity in this period (Period II). Consumption intensity in relation to CL also has increased during this period in the domestic, commercial, and agricultural sectors (but declined with the industrial customers). Thus, in the aggregate, while N has grown at an exponential rate of 6.68 per cent and CL at 6.1 per cent during Period II, consumption of electricity (C) has increased at 7.14 per cent per annum. That is both the 'horizontal' (growth of N or CL) and 'vertical' (growth of consumption intensity, i.e., growth of C/N or C/CL) growth components contributed to the total growth of consumption during Period II. (Note that $r(C) = r(N) + r(C/N) = r(CL) + r(C/CL)$, where 'r' represents the growth rate.) This dramatic growth in consumption intensity (i.e., consumption per customer, C/N as well as per KW of CL, C/CL) in Period II even in the face of power cuts and restrictions is ample evidence to back our assumption of an increasing slope (dC/dN – marginal electricity consumption intensity of customers) of the energy consumption function during Period II in relation to Period I function. Now, below we discuss how we obtain and apply this time-growing slope coefficient.

We can approximate the ratio of the annual increment of consumption to that of number of customers ($\Delta C/\Delta N$, where $\Delta C = C_t - C_{t-1}$, and similarly for ΔN) in Period II to the slope (dC/dN) of the actual consumption function obtainable in that period.

Next we take the annual average growth rate of this $\Delta C/\Delta N$ series and then allow the slope of the Period I consumption function to grow at this rate during Period II. The intercept of this function (that stands to account for the influence of all other factors on consumption) may also be allowed to vary in Period II; however, we can justifiably assume that the changes in the influences of these other factors are only minimal and that the Period II consumption function stands to keep the continuum upon the Period I function intercept, but with a time-growing slope. Thus the adjusted consumption ($C_{adjusted}$) data series of Period II is obtained as:

$$C_t^{adjusted} = a + b(1+r)^t N_t$$

where a = Period I function intercept,

b = Period I function slope,

r = Annual growth rate of $\Delta C/\Delta N$ of Period II,

t = 1, 2, ……, for years 1982-83, ……, 1997-98, and

N = number of customers.

This then provides us with a measure of unconstrained energy consumption of the actually existing (i.e., constrained) number of customers in Period II –consumption model I. We can also have another scenario of unconstrained consumption of unconstrained number of customers – consumption model II, where we assume no restriction on providing connections, that grow unhampered at the same rate as in Period I of no-power-shortage. Here both C and N are adjusted upwardly in terms of Period I growth structure, and the growth rate of $\Delta C/\Delta N$ is applied to the slope of Period I consumption function along with the adjusted N.

Now that we have the complete consumption data base, consisting of the original consumption data till 1981-82 and the estimated, unconstrained data since 1982-83, we can feed it into an appropriate model for forecast purposes. We follow the logic of multi-variate autoregressive –moving average (MARMA) model, also called transfer function model, that correlates a dependent variable to its own lagged values, current and lagged values of one or more explanatory variables, and an error term whose behaviour is partially 'explained' by a time series model. In our bi-variate model $C_t = \alpha + \beta N_t + U_t$, the error term U_t representing unexplained variations in C_t , can be analyzed and 'explained' by constructing an autoregressive-integrated-moving average (ARIMA) model to it. The combined regression-time series model is then $C_t = \alpha + \beta N_t + \phi^{-1}(B)\,\theta(B)\,\eta_t$, where η_t is an independently and identically distributed (iid) normal error variate (white noise), ϕ and θ are respectively the autoregressive (AR) and moving average (MA) parameters and B is the backward shift operator (for example, $BC_t = C_{t-1}$, $B^2C_t = C_{t-2}$, etc. As the combined model yields a structural explanation of that part of the variance of C_t that can be explained in terms of a structural regression, and a time series explanation of that part of the variance of C_t that cannot be explained structurally, it is likely to provide much better forecasts (see Box and Jenkins, 1970, chaps, 10, 11; Makridakis and Wheelwright, 1978, chap. 11; Pindyck and Rubinfeld, 1981, chap. 20).

It should be acknowledged that forecasting is essentially uncertain, and forecasts should define a range (confidence interval) around a central estimate (expected value, mean) within which it is agreed there is a high probability that the unsuppressed demand will actually fall. Only the central estimates are considered in evaluating the APS and Planning Commission forecasts, but for capacity planning purposes, a range of values would more accurately reflect our extent of knowledge about the future. Moreover, forecasts should be principally concerned with estimating demand more than five years ahead. Generation projects in particular take a long time to gestate; initial decisions on such large commitments of capital must be made at least 4 to 8 years before the project can be commissioned, and generation projects typically have a life of at least 30 years. Hence, the significance of long-range forecasting. Also, there is a need, at the official level, to develop a single, generally acceptable demand forecast (a single exercise that defines a single range of demand to be expected). A first step would be to reach a consensus on the methodology to be employed for short-run forecasts. For long range planning, a macro econometric regression model would suit better.

An advantage of the regression approach is its flexibility, but an equally strong disadvantage is its practicality in the face of possible intrusion of errors. Any decision maker using an analytical tool would like to know where in the whole exercise errors are likely to arise and be able to make allowances for them. In econometric forecasting models, there are five commonly recognized reasons why forecasts may go astray: i) wrong equation or variables are used (specification error) ii) coefficients are inaccurately estimated (estimation error), iii) one-off shocks to the economy may alter the underlying situation (eg., the oil crisis or a change in state boundaries), iv) more gradual change s in the economic structure may not be anticipated (e.g., an acceleration of rural electrification with a new government), and v) the exogenous variables, the weather, level of economic activity, etc., may be incorrectly forecast. In a relatively simple model, it is easier to make allowances for errors of types (iii) to (v). Again, a simple forecasting model is to be preferred to taking the risk of making specification errors with a complex model.

Uncertainty accompanies every forecast. The range of uncertainty of the forecast increases with the length of the period of forecast. For very long forecast periods, non-formal methods are utilized, such as some variant of the Delphi method. The basic features of this approach are the use of questionnaires, anonymous collective work, iterative precisioning of results, evaluation of the experts' qualifications, and a statistical evaluation of the obtained answers.

4. A Time Series Analysis

We start with an analysis of the time series data on the internal maximum demand (MD) for power (in megawatts (MW)) in the Kerala system from 1957-58 to 1997-98 in the framework of the common extrapolation models explained earlier. Remember the given demand series is in fact supply constrained, i.e., suppressed, demand, especially since the infamous draught year of 1982-83. However, for the initial forecasting analysis using all the available models as we listed above, we take this demand series as such, since our aim is just an exposition of the use and strength of simulation of these models. (We have not tried the growth models, primarily because of the problem of the choice of an upper limit L.) Table 1 reports the OLS estimates of the parameters along with other statistics of these models. Time trend is the explanatory variable in all but the two first-order automatically-regressive trend models, in which the one period lagged dependent variable (MD or log MD, as the case may be) explains the relationship. We have also fitted the historical simulation of the demand series based on the estimated models, and report in Table 1 the corresponding measures of the performance of the model, to see how closely the simulation tracks the demand series. These simulation error measures, signifying the deviation of the simulated variable from its actual time path, are root-mean-square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and Theil inequality coefficient (TIC) along with its 3 components.

RMSE is the square root of the mean of the squared deviations between the simulated and the actual values. MAE is just the arithmetic mean of the deviations between the two, while MAPE gives MAE in percentage terms. Evidently, the smaller the error, the better and more reliable the simulation.

TIC is a very useful simulation statistic related to the RMSE and applied to the evaluation of historical simulations or ex post forecasts. It is given by the ratio of the RMSE to the sum of the square roots of the mean squared values of the simulated and the actual data series, such that it will always fall between 0 and 1. If TIC =0, the simulated and the actual series coincide for all t and there is a perfect fit. If TIC = 1, on the other hand, the predictive performance of the model is the worst. The TIC is decomposed into 3 components, bias proportion (BP), variance proportion (VP), and covariance proportion (CP), with BP + VP + CP = 1. The BP is an indication of systematic error, since it measures the extent to which the mean values of the simulated and actual series deviate from each other. Whatever be the value of the TIC, we would hope to obtain a BP much closer to zero for a good fit. The VP indicates the ability of the model to replicate the degree of variability in the variable under study. If

VP is large, it means that the actual series has fluctuated considerably while the simulated series shows little fluctuation, or vice versa, which is quite undesirable. we would hope to see minimum variability between the two. The CP measures the unsystematic error, i.e., it represents the remaining error after deviations from average values and average variabilities have been accounted for. Since it is unreasonable to expect simulations perfectly correlated with actual series, this component of error is less problematic. In fact, it is generally accepted that for any value of $TIC > 0$, the ideal distribution of inequality over the 3 components is $BP = VP = 0$, and $CP = 1$.

Now, it can be seen that all the function fits are highly significant in terms of the t-values of the parameter estimates, adjusted R^2, and F-values (Table 1). The 'k-transformation' model has been turned out to be either defined or significant only for the values of $k = 0.4$ and $k = 0.5$, out of a range of values tried; both these results are reported. For the second-degree polynomial (parabola) and the k-transformation, the DW statistic is around 2, signifying no first-order serial correlation. The DW statistic cannot be used if the regression equation contains a lagged dependent variable, as in the case of the first-order auto-regressive trend models (both linear and logarithmic), with DW statistic around 2. The correct procedure is to use the Durbin h test statistic that can be estimated using the given DW statistic and the variance of the parameter estimate. In the case of the simple auto-regressive trend model, Durbin h statistic turns out to be -1.217, the absolute value of which is less than the normal critical value of 1.645 at 5 per cent significance level. Hence we cannot reject the null hypothesis of no serial correlation, and the use of the OLS estimation procedure is thus validated. Similarly, in the logarithmic auto-regressive model, $h = -1.074$, and this model also comes out to be significantly reliable.

However, in the other two cases, linear and exponential models, the DW statistic is very low, indicating the presence of positive first-order serial correlation. The existence of serial correlation can be observed in the graphs by noticing that regression residuals tend to be highly correlated. When the simulated value associated with one observation lies below the corresponding actual value, it is vary likely that the neighbouring simulated values will follow suit. For a fairly long span of time, (nearly

25 years since 1965 in a 41-year sample), the fitted values from the exponential model tend to lie much below the actual ones. What is more discouraging from the forecasting perspective is that the last 10 or so fitted values from the linear model lie below the actual values, suggesting that the forecasts based on this model are likely to understate the actual demand time series, whereas the forecasts based on the exponential model are likely to considerably overstate the actual series, as the corresponding fitted values to the end of the sample series in this latter case tend to shoot up at an ever-increasing rate. This calls for suitable corrections for the presence of serial correlation.

To improve the results, we have re-estimated the two models using correction procedure for first-order auto-regression, AR(1). The results are also reported in Table 1, where 'AR(1)' stands for the estimate of the first-order autocorrelation (autoregression) coefficient . The estimation statistics, adjusted R^2 and F have in fact improved very much now, and the DW's are around 2, though in the case of the exponential model, with AR(1) correction, the t-values of the time parameter estimate has dropped considerably, it now being significant at 17 per cent only. However, for our purposes, what is important is the ability of the model to forecast. In fact, the correction for serial correlation substantially improves the fit of the original model, as seen by examining the plot of the fitted and actual values (not presented here). The results on the simulation error measures also point to this fact which we discuss below.

For all the models, we have also estimated and report in Table 1 itself the measures of simulation error - root-mean-square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), as well as Theil inequality coefficient (TIC) along with its 3 components – bias proportion (BP), variance proportion (VP), and covariance proportion (CP). By these measures, it appears that all the models in general have very good fitting performance/ability, with very low TIC, along with an almost zero BP in most cases and a close to zero VP. In particular, in the case of linear and exponential models, the AR(1) correction has yielded much lower simulation error measures; and in fact, the corrected exponential model has the lowest TIC and its components and the highest adjusted R^2 and F values among all the models, and the logarithmic autoregressive trend model and the second-degree polynomial (parabola) model are very close in their performance results to this (corrected exponential) model, with very low TIC, and bias and variance proportions – in the case of the parabola model, for example, BP is zero to 16 places while VP, to 3 places, and CP is very close to unity. Considering the simplicity in use, then, these models emerge as the best choice for forecasting purposes, given the data base.

Similarly, the results of the analysis of energy consumption (C) using the data on internal sales of electricity (in million units (MU)) from 1957-58 to 1997-98 in the framework of these models are given in Table 2. In the light of the above discussion, the results are self-explanatory.

Remember both the data bases we have used here are the ones actually constrained by power shortages – power cuts and load shedding. It goes without saying that reliance on and use of these data for forecasting purposes just involves high risk of errors of underestimation, as already explained – for example, the steady decrease that is observed in the APS demand forecasts for Kerala is an apt case in point. Hence the significance of the need to adjust these data upward to account for the suppressed part of the actual demand. The methodology explained earlier for accomplishing this is illustrated in detail with the data series of total energy consumption and of maximum demand in the two scenarios (models). Then we take up an analysis of the sectoral energy consumption in the framework of the 'bottom-up' method of forecasting total consumption. Once the constrained part of the data series is adjusted upward, the whole data base is used to develop the bi-variate ARIMA model for forecasting.

1. Total (Internal) Energy Consumption

The electricity consumption (C) – number of customers (N) linear relationship during Period I (1957-58 to 1981-82) is found to be highly significant and with very small simulation errors:

$$C_t = 382.959 + 1.625\ N_t$$
$$(5.755)\quad(19.224)$$
Adjusted R^2 = 0.9389; SER = 197.119; F = 369.575;

DW = 0.3073;

RMSE = 189.07; TIC = 0.0587;

where figures in brackets are t-values, SER is standard error of regression, DW is Durbin-Watson statistic, RMSE is root-mean-square error, and TIC is Theil inequality coefficient. The low DW statistic indicates the presence of positive first-order serial correlation. However, no correction is made here and we carry over these vestiges on to the next period; the whole data base is purged of time-series disturbances at the stage of the development of our BARIMA model.

During 1982-83 to 1997-98, electricity consumption in the State grew at an exponential rate of 7.14 per cent per annum. The widely fluctuating consumption data series of this period is smoothed out using this annual growth rate in order to obtain an increasing trend over the period. The series of the ratio of the annual increment of this consumption to that of number of customers ($\Delta C/\Delta N$) is then obtained, and its growth rate is found to be 1.734 per cent per annum.

Now, the slope coefficient (1.625) of the Period I relationship is made to grow annually at this rate during Period II, and the corresponding consumption series is generated for the respective number of customers upon the same Period I function intercept which gives the adjusted, unconstrained, consumption. Thus we find that in 1982-83, the first year of the worst draught era, the actual (i.e., constrained) consumption fell by 71.5 MU over the previous year (in 1981-82, consumption increased by 144.1 MU over 1980-81), whereas if the above consumption – number of customers relationship continued, unhampered by the restricting draught, but with an increased consumption intensity of the actual number of customers, into that year also, it should have been about 3637.6 MU. Similarly, in 1997-98, the last year of the period of our study, the energy consumption should have been 11528.5 MU, given the actual number of customers, instead of the restricted 7716.23 MU.

The whole data series, consisting of the original data up to 1981-82 and the adjusted one since 1982-83, when correlated with the actual number of consumers, gives the following relationship:

$$C_t = 59.468 + 2.0501\ N_t$$
$$(0.755)\quad (60.093)$$

Adjusted $R^2 = 0.989$; SER = 336.956; F = 3611.22;
DW = 0.1685;
RMSE = 328.635; TIC = 0.0343;

To improve upon the forecasting performance of this model, we construct and apply an ARIMA noise model to its residual series (U_t). Based on the examinations of the patterns of the sample autocorrelation and partial autocorrelation functions, we have tried a number of specifications and tentatively chosen an ARIMA(1, 0, 2) model for these residuals. That is,

$$U_t = (1 - \phi_1 B)^{-1} (1 - \theta_1 B - \theta_2 B^2) V_t$$

where ϕ_1 is AR parameter, θ_1 and θ_2 are MA parameters and V_t is an iid normal variate. The parameter estimates are:

Estimate of ϕ_1 = 0.893; t-statistic = 13.519;

Estimate of θ_1 = 0.240; t-statistic = 1.731;

Estimate of θ_2 = 0.582; t-statistic = 4.21;

Adjusted $R^2 = 0.856$; SER = 127.755; F = 78.278;
DW = 1.9404; χ^2 (3, 40) = 8.53;

RMSE = 121.199; TIC = 0.1885.

The parameters θ_1 and θ_2 of a MA(2) model constrained to lie in the triangular region in the θ_1, θ_2 plane defined by the inequalities:

$\theta_2 + \theta_1 < 1$;

$\theta_2 - \theta_1 < 1$; and

$-1 < \theta_2 < 1$.

Similarly, the parameters ϕ_1, ϕ_2 of an AR(2) process must lie within the same triangular region as do the MA(2) parameters. The restriction is necessary for the AR(2) model to ensure stationarity while for an MA(2) process, it is to ensure invertibility, since *any* MA process of *infinite* order is necessarily stationary. (For an AR(1) process, the necessary condition for stationarity is $-1 < \phi_1 < 1$.) The parameter estimates of our ARIMA(1, 0, 2) model do obey the restrictions.

A significant stage in the ARIMA model building is diagnostic checking, testing the time-series model for 'goodness of fit', by checking if any 'non-randomness' that is explainable is still present in the residuals from the model. By assumption, these residuals are uncorrelated at any lag, having mean zero and common variance, and therefore if the model is correct and the parameter estimates are close to the true values of the parameters, then within sampling variation both the residual mean and the residual autocorrelations should be zero. In particular, the null hypothesis of zero residual autocorrelations for a given lag (as well as of zero residual mean) is not rejected (at the 5 per cent significance level), if the sample autocorrelation function (as well as sample mean) of the residuals lie in absolute value within the limits of 2 standard error (SE) of the residual autocorrelations (or 2 standard deviation (SD)). An overall test of model adequacy is provided by the chi-square test (to test the hypothesis that the residuals are white noise, i.e., they are not correlated with each other). Since the χ^2 statistic is a measure of correlation among the residuals (it being the sum of the squared sample residual autocorrelations weighted by their number), its value must be as small as possible for a correctly specified model.

For the above ARIMA(1, 0, 2) model, the sample residual autocorrelations are found to be well within the ± 2 SE limits and the χ^2 statistic has 37 degrees of freedom (40 lags minus 3 estimated parameters) and a small value of 8.53, falling much below the 99.9 per cent point (69.34) on the distribution, so that it is almost certain that the residuals are white noise.

Thus our complete model is $C_t = \alpha + \beta N_t + (1 - \phi_1 B)^{-1} (1 - \theta_1 B - \theta_2 B^2) V_t$. The parameters are then estimated simultaneously estimated and are given below:

$$C_t = -4483.602 + 2.3175 N_t + (1 - 0.998 B)^{-1} (1 - 0.212 B - 0.536 B^2) V_t$$

(–0.029) (5.041) (12.484) (1.470) (3.596)

Adjusted R^2 = 0.998; SER = 133.105; F = 5689.43;

DW = 2.0799; χ^2 (4, 40) = 8.87;

RMSE = 124.508; TIC = 0.0129.

The estimated AR and MA coefficients obey the restrictions and the very low χ^2 value asserts (with almost certainty) the model adequacy. The residual autocorrelations are also found to be almost zero within the sampling variation, confirming the white noise character of the residuals. The maximum (absolute) value of the residual autocorrelation function (ACF) is 0.178 at the 18^{th} lag, well within the 2SE limits, the SE of the autocorrelations being 0.158. Also the arithmetic mean of the residuals (–1.6207) is very small relative to the estimated (sample) SD (19.94).

The full model is also to be diagnosed if the residuals be independent of the causor time series. To test this hypothesis of independence, the estimated cross correlation function (CCF) between the residuals and the input time series (number of customers) is to be examined to see if the values lie in absolute value within the limits of 2 SE of the cross correlations. The CCF from the above model are found to have no significant values (the maximum absolute value of the CCF being 0.314 at the 16^{th} lead, with SE = 0.158), signaling that the causor series and the model residuals are uncorrelated. Hence the tentative model is accepted as adequate, and can now be used for forecasting purposes.

Note that the simulation performance of the model also has improved significantly. Both the RMSE and the TIC have dropped by more than two-and-a-half times; RMSE from 328.6 (of the bi-variate regression model) to 124.5 (of the BARIMA model), and TIC from 0.034 to 0.013.

We have also compared the forecast values from this consumption model I with the actual (constrained) ones since 1980-81 (the first and the last series are up to 1997-98 only). The number of customers since 1998-99 is assumed to have the same (constrained) growth rate as during Period II (1982-83 to 1997-98), i.e., 6.68 per cent per annum on an average. The unconstrained (simulated) consumption of 1982-83 is about 13 per cent above the constrianed one; however, by 1997-98, it becomes nearly 50 per cent above the latter (growing at about 9 per cent p. a.), indicating a possible measure of the increasing extent of the suppressed demand.

The comparison shows that our forecasts (from consumption model I) are only slightly below the 15^{th} APS series; for example, for 1997-98, our projection (11440.54 MU) is about 97.2 per cent of the 15^{th} APS figure (11770 MU) and for 2009-10, about 95.63 per cent (25600.14 MU / 26770 MU); whereas with respect to the 14^{th} APS, our 1997-98 projection is about 10 per cent below the APS forecast (12861 MU) and that for 2009-10 is about 22 per cent lower (than the APS 32894 MU). It can, however, be seen that if we smoothed out the actual consumption series of period II at a growth rate of 7.34 per cent, instead of at the actual 7.14 per cent per annum, the resultant forecast would almost equal the 15^{th} APS series, and at a growth rate of 7.93 per cent, forecast series almost equal to the 14^{th} APS ones would also result. It should be noted that the unconstrained consumption series generated from consumption model I is in fact an underestimate in that it is based on the actual number of customers which falls in a constrained set; since 1982-83, the Board has been restricting the number of connections as well. During 1957-58 to 1981-82, the number of customers of the Kerala power system had registered an average exponential growth rate of 11.15 per cent per annum, whereas for the period from 1982-83 to 1997-98, the growth rate was only 6.68 per cent per annum. Thus the above ex post simulation is just the estimate of the 'unrestricted' power demand of the 'restricted' i.e., actual number of electricity consumers in Kerala, and therefore does not represent the potential, market, demand of the potential number of customers. That is, the actual period II growth rate of consumption must have been much above the recorded (constrained) rate of 7.14 per cent. This then lead us to considering a second scenario – that of the unrestricted, potential, power demand that would occur, if the number of connections could grow unhindered as during the period prior to 1982-83.

During the Period II, when the number of customers grew at an annual rate of 6.68 per cent, electricity consumption outgrew it at a rate of 7.14 per cent (reflecting increased intensity of consumption even in the face of shortage restrictions). Given this norm, if the number of customers in Period II were allowed to grow at the same rate (11.15 per cent) as in period I, then consumption should have grown at (7.14) × (11.15) / (6.68) = 11.9 per cent p. a. In developing consumption model II, we smooth out the Period II consumption series at this rate and estimate the $\Delta C/\Delta N$ series, where both C and N are unconstrained, growing at the rates respectively of 11.9 and 11.5 per cent p. a. These annual increment ratios, in turn are found to grow at 0.677 per cent p. a., and the slope of the Period I C-N function is then made to grow annually at this rate during Period II, upon the Period I intercept, in order to generate the series of the unconstrained consumption of the unconstrained number of customers. The complete data base then yields the following relationship (with t-values in brackets):

$$C_t = 234.271 + 1.7896\, N_t$$
$$(5.573) \quad (150.109)$$

Adjusted $R^2 = 0.9982$; SER = 201.004; F = 22532.64;

DW = 0.3390;

RMSE = 196.039; TIC = 0.0152;

The residuals from this model are tentatively diagnosed to follow an ARIMA(1, 0, 2) model, with the following parameter estimates:

Estimate of ϕ_1 = 0.673; t-statistic = 7.305;

Estimate of θ_1 = 0.240; t-statistic = 1.717;

Estimate of θ_2 = 0.590; t-statistic = 4.268;

Adjusted R^2 = 0.709; SER = 108.324; F = 32.653;

DW = 1.9615; χ^2 (3, 40) = 9.04;

RMSE = 102.765; TIC = 0.2783.

And the complete model is estimated simultaneously and is given below:

$$
\begin{array}{l}
C_t = 190.517 + 1.8166\ N_t + (1 - 0.706\ B)^{-1}(1 - 0.216\ B - 0.570\ B^2)V_t \\
\qquad\quad (2.322) \qquad (90.469) \qquad\quad (7.501) \qquad\qquad\quad (1.505) \qquad (4.02)
\end{array}
$$

Adjusted R^2 = 0.9995; SER = 108.86; F =18872.83;

DW = 2.0214; χ^2 (4, 40) = 8.42;

RMSE = 101.83; TIC = 0.0078;

Residual mean = – 0.00766; Estimated SD = 16.31;

Maximum (absolute) value of the residual ACF = 0.193 at 18th lag; SE = 0.158;

Maximum (absolute) value of the CCF = 0.300 at 16th lead; SE = 0.158.

Notice the substantial fall in both SER and simulation errors, compared with the structural model results. All the tests on the white noise null hypothesis of the residuals are also positive. The model is thus acceptable for forecast purposes and we have compared the forecast values from this consumption model II with other projections. The number of customers since 1998-99 is assumed to grow at the same unconstrained rate as during Period I (actual) and Period II (by assumption). It is seen that the 14th APS figure for 1997-98 (12861 MU) is only 71 per cent of our unconstrained estimate of consumption (18166.9 MU) for that year, and for 1009-10 (30418 MU), is only 48 per cent (of 63962.2 MU). Evidently, the unconstrained consumption estimates from our two models (one with the actual, constrained, number of customers, and the other with unconstrained number) may be taken as the lower and upper limits of the 'actual market' demand. Assigning appropriate weights to the two scenarios, we can have a weighted mean series to represent this market demand.

2. (Internal) Maximum Demand

Next we take up the modeling of the maximum demand (MD). The crippling effect of the severe shortage has heavily hampered the growth of MD especially on account of the cyclic load shedding, designed, since 1989-90, to curtail MD on the system during peak load hours in the evening, that affects all the consumers except the major industrial customers, who are already subjected to power cuts. During Period I, MD had an annual average exponential growth rate of 9.6 per cent, while in Period II of shortages, it fell to 5.16 per cent. Both N and CL grew faster than MD in both the periods, leaving decaying MD intensity factors. That is, the contribution to the growth of MD by the 'horizontal' growth components (growth of N or CL) was cut down by the decaying 'vertical' growth components (growth of MD/N or MD/CL). The increasing trend that characterized the consumption intensity of Period II was not reflected in MD intensity. Hence we have to bring up Period II MD in line with C. Below we build up bi-variate ARIMA models, correlating MD to N, and constructing and applying appropriate ARIMA model to the residuals from this relationship, for the two scenarios of constrained and unconstrained number of customers.

During Period I, when N grew at 11.15 per cent p. a., MD registered an annual exponential growth rate of 9.6 per cent. Based on this norm, 6.68 per cent growth in N should have involved (9.60) × (6.68) / (11.15) = 5.76 per cent growth in MD during Period II, instead of the actual 5.16 per cent. This (5.76 per cent) growth rate we use to smooth out the MD series during Period II, and from the resultant series and the corresponding (actual) N series, we derive the annual $\Delta MD/\Delta N$ coefficients, found to grow at a rate of 0.438 per cent p. a. At this rate is then allowed to increase, during Period II, the slope coefficient of the Period I MD-N relationship, given below (with t-values in brackets):

$$MD_t = 50.256 + 0.449\ N_t$$
$$(6.061) \quad (42.670)$$

Adjusted R^2 = 0.987; SER = 24.563; F = 1820.747;
DW = 0.6040;
RMSE = 23.560; TIC = 0.0298.

In continuum of this Period I relationship, the Period II MD time series data are generated, upon the above intercept, from the actual Period II N, weighted by the above time-growing slope coefficient. This unconstrained MD of the actual (constrained) number of customers, for example, for 1982-83 is found to be 938.96 MW against the actual 823.2 MW. In that year, it should be noted, MD increased by just 20 MW over the previous year, whereas in 1981-82, it increased by 77.8 MW over 1980-81. The estimated unconstrained MD for 1997-98 is 2560.96 MW against the actual 1785.8 MW.

Now it is straightforward to have an estimate of the installed capacity required to meet this unconstrained demand, assuming a certain percentage reserve margin (PRM). Let us for simplicity take the true market demand for power in 1982-83, if the same maximum demand – number of customers norm continued into that year, at an average 900 MW, with a given (constrained) number of consumers in that year, and assume a minimum PRM of 20. This would then give us an installed capacity of 900 × 1.20 = 1080 MW required in 1982-83, given the technical availability, to meet the unconstrained demand of these customers, against an actual capacity of 1011.5 MW. It should be noted seriously that the Kerala power system had been virtually pinned up to this level of capacity for a long period of 9 years since the commissioning of the first stage of Idukki hydro-power project in 1976-77, as the complacency out of a comfortable power surplus situation had started to corrode the system planning, designing and implementing exercises with grave consequences for the future. For 1997-98, with a 20 PRM, the installed capacity required, given the availability factor, to meet the unconstrained market demand of an average 2500 MW, is estimated at 3000 MW, against an actual capacity of 1775.8 MW. Note that this capacity level itself was below the recorded maximum demand of 1785.8 MW in that year, which was met through 'liberal' import of energy, amounting to an astounding 55 per cent of the total internal energy sales in Kerala in that year. (The available power, determined by storage, outage, etc., would be much below.)

The complete set of data, consisting of the Period I actual MD series and the Period II adjusted MD series, yields the following relationship with N:

$$MD_t = 29.620 + 0.477\, N_t$$
$$(4.763) \quad (177.058)$$

Adjusted $R^2 = 0.999$; SER = 26.595; F = 31349.5;

DW = 0.4664;

RMSE = 25.938; TIC = 0.0115.

To improve upon the forecasting power of this model, the residuals from it are couched in an ARIMA(1, 0, 2) model, tentatively chosen after examining the ACF and PACF of the residuals and several tries with different specifications. Thus the residuals

$$U_t = (1 - \phi_1 B)^{-1} (1 - \theta_1 B - \theta_2 B^2) V_t ,$$

where ϕ_1 is AR parameter, θ_1 and θ_2 are MA parameters and V_t is pure white noise. The parameter estimates are:

Estimate of ϕ_1 = 0.568; t-statistic = 5.752;

Estimate of θ_1 = 0.239; t-statistic = 2.087;

Estimate of θ_2 = 0.756; t-statistic = 6.345;

Adjusted R^2 = 0.654; SER = 15.651; F = 25.533;

DW = 2.0295; χ^2 (3, 40) = 11.05;

RMSE = 14.848 ; TIC = 0.3073.

The parameter estimates satisfy the restrictions on them, and the low chi-square value confirms, in almost certain terms, the residuals from the models as pure white noise, containing nothing 'explainable'.

The ARIMA(1, 0, 2) model of the residuals is now combined with the structural model of the MD-N relationship, and the estimates of the full model are:

$$MD_t = 25.366 + 0.4808\, N_t + (1 - 0.593\, B)^{-1}\, (1 - 0..228\, B - 0.752\, B^2) V_t$$

(2.607) (120.709) (5.846) (1.915) (6.137)

Adjusted R^2 = 0.9995; SER = 15.842; F =21505.72;

DW = 2.0770; χ^2 (4, 40) = 10.81;

RMSE = 14.819; TIC = 0.00652;

The ARIMA parameter estimates continue to obey the theoretical restrictions on them. The arithmetic mean of the residuals (-0.0566) is statistically zero in comparison with the estimated (sample) SD (2.373); so also are the ACF, the maximum absolute value being 0.230 at the 11th lag (SE = 0.158), and the CCF, the maximum absolute value being 0.240 at 9th lead (SE = 0.158). And above all, the very low chi-square value reasserts with almost certainty the model adequacy. The fall in SER (from 26.595 of the structural model to 15.842 of the complete BARIMA model) and in the simulation error measures (from 25.938 RMSE and 0,0115 TIC of the structural model to 14.819 RMSE and 0.0065 TIC of the BARIMA model) is also substantial.

A comparison of the forecast values of MD from this MD model I, with the actual (constrained) values since 1980-81 is in order now. The number of customers since 1998-99 is assumed to grow at the same rate as during Period II of power shortage, i.e., at 6.68 per cent p. a. on average. There is a striking similarity between our estimates from the above MD model I and the APS series. For the years 1997-98 and 1998-99, our estimates of MD (in MW) are: 2545.9, and ***2717.5***; while the 14th APS ones (accepted by the State) are: 2514, and ***2713***. However, the two estimates diverge thereafter, the 14th APS series leaping ahead of our series; which in 1999-2000 is about 99 per cent of the former, and in 2009-10, about 87 per cent. On the other hand, the 15th APS estimates (recommended by the Balanandan Committee) come much closer to our forecasts in later years: for example, our estimates since 2004-05 (in MW) are: 39657.1, 4229.7, 4510.2, 4809.7, 5129.2, 5470.2 and 5834.0, and the 15th APS estimates are: 3952, 4229, 4505, 4811, 5139, 5488, and 5801. Needless to repeat here, the forecasts from our MD BARIMA model I are in fact underestimates, as they are so designed as to represent the unconstrained demand of a 'constrained' set of customers; and to this extent the APS forecasts also are underestimates.

Next we turn to building up our MD model II to estimate the unconstrained MD of the unconstrained number of customers, growing in Period II and thereafter in the near future at the same rate (11.15 per cent) as in Period I of no-power-shortage. We have found that consumption intensity in Period II was growing despite the restrictions and assumed in developing the consumption model II that were the number of customers allowed to grow at the unconstrained rate of 11.15 per cent, then consumption would increase by 11.9 per cent p. a. during Period II. it is not unreasonable to assume the same growth rate for unconstrained MD in Period II also in view of the lack of other relevant information. Thus we smooth the MD series out in Period II at an annual average growth rate of 11.9 per cent. This gives us an average growth rate of 0.677 per cent p. a. of the $\Delta MD/\Delta N$ series, as in consumption model II; this growth rate is then applied to the slope coefficient of the Period I MD-N relationship, and thence the Period II adjusted MD series is generated. The full data base makes up the following relationship:

$$MD_t = 9.130 + 0.4950\, N_t$$
$$(1.061) \quad (202.780)$$

Adjusted $R^2 = 0.999$; SER = 41.154; F = 41121.10;

DW = 0.2445;

RMSE = 40.137; TIC = 0.115.

N_t represents the unconstrained number of customers.

The residuals from the model are found and tentatively accepted to have an ARIMA(2, 0, 2) structure. That is,

$$U_t = (1 - \phi_1 B - \phi_2 B^2)^{-1} (1 - \theta_1 B - \theta_2 B^2) V_t ,$$

where ϕ_1 and ϕ_2 are AR parameters, θ_1 and θ_2 are MA parameters, and V_t is pure white noise. The parameter estimates are:

Estimate of ϕ_1 = 0.716; t-statistic = 4.358;

Estimate of ϕ_2 = 0.131; t-statistic = 2.773;

Estimate of θ_1 = 0.273; t-statistic = 1.980;

Estimate of θ_2 = 0.694; t-statistic = 5.299;

Adjusted $R^2 = 0.797$; SER = 18.663; F = 38.37;

DW = 1.9032; χ^2 (4, 39) = 10.55;

RMSE = 17.426 ; TIC = 0..2228.

The parameter estimates satisfy the restrictions on them, and the low chi-square value confirms, in almost certain terms, the residuals from the models as pure white noise, containing nothing 'explainable'.

The complete model, estimated simultaneously is:

$$MD_t = \underset{(-0.414)}{-203.078} + \underset{(37.839)}{0.5204}\, N_t + (1 - \underset{(2.065)}{0.417}\, B - \underset{(2.485)}{0.542}\, B^2)^{-1} (1 - \underset{(2.463)}{0.480} B - \underset{(2.699)}{0.428}\, B^2)\, V_t$$

Adjusted R^2 = 0.9998; SER = 18.093; F =41071.43;

DW = 2.0262; χ^2 (5, 39) = 11.15;

RMSE = 16.643; TIC = 0.00463;

Residual mean = – 0.0395; estimated SD = 2.666;

Maximum (absolute) value of the residual ACF = 0.266 at 10th lag; SE = 0.160;

Maximum (absolute) value of the CCF = 0.210 at 13th lead; SE = 0.160.

The fall in both the SER and the simulation errors (compared with the structural model results) is highly significant. All the diagnostic checking tests confirm the white noise character of the model residuals, and assert the model adequacy. Thus the MD model II is acceptable for forecasts purposes. In the context of a comparison with the forecast values, the number of customers since 1998-99 is assumed to have an annual growth rate of 11.15 per cent, the same unconstrained rate as during Period I (actual) and Period II (by assumption). We find that the 14th APS projection for 1997-98 (2514

MW) is only 50.5 per cent of our unconstrained estimate of 4892.2 MW for that year, and for 2009-10 (6264 MW), is only 34.6 per cent (of 18101.2 MW). This potential estimate from MD model II deviates by more than 2500 MW in 1997-98 from the adjusted demand of the actual number of consumers of MD model I. From this, now, the installed capacity required to meet the unconstrained maximum demand of the unconstrained (in number) market agents in 1997-98, assuming an average demand of 4900 MW and minimum 20 PRM, is estimated at 5880 MW, given the availability factor. Considering the tempo of the progress in project additions in the Kerala power sector, it is very much unlikely ever to reach this level of capacity in the foreseeable future. Thus in the face of this 'perfectly' unconstrained forecast of maximum demand, we would just continue to be reeled under the constraints of power shortage for a very long period.

As already mentioned, several trials have been run before finalizing the above models; we have tried with variable transformation (logarithmic specification) and series differencing to ensure stationarity. However, differencing the series has been found to be not necessary and the results have also not been satisfactory. The diagnostic checking of all the four of our models has affirmed the adequacy, the 'goodness of fit', of the models. Moreover, linear specification has faired much better here than logarithmic transformation. On the other hand, logarithmic models have turned out to be preferable in the sector-wise analysis, which we discuss below.

3) Sector-wise analysis

In the disaggregated, 'bottom-up', method, energy consumption forecast exercise for each individual sector is carried out separately and independently of each other, and the sum of the individual sectoral forecasts makes up the total consumption forecast.

KSEB caters to 8 customer groups, viz., domestic, commercial, LT industry, HT-EHT industry, public lighting, agricultural, water works, and licensees; since 1997-98, railway traction also has become a customer sector in Kerala. We take up for individual analysis the first 4 sectors and 'others' including the remaining agricultural, public lighting, water works, and licensees. Data are available for domestic and industrial customers since 1960-61, and for commercial and others since 1964-65. Only one scenario is considered – the unconstrained consumption function of the actual (constrained) number of customers in each sector. In all the cases, logarithmic models have been found to have much better forecast performance. Below we briefly discuss these models:

i) Domestic sector

In 1997-98, the domestic customers made up about 76 per cent of the total number of energy customers of the KSEB, and accounted for about 49 per cent of energy consumption, 30.4 per cent of sales revenue, and for 45.1 per cent of total connected load (CL). In 1970-71, the percentage shares of this sector in customer number, consumption, sales revenue, and CL were respectively 71, 4.3, 18.8 and 25.4 per cent. There is no gainsaying the fact that the growth of the Kerala power system largely hinges upon that of this sector. Energy consumption (C) of the domestic consumers grew on an average at an impressive 11.6 per cent p. a. during Period II, despite restrictions that cut down the growth of the number of customers from 9.7 per cent in Period I to 6.7 per cent in Period II, and the growth of CL from 10.1 per cent to 7 per cent respectively. The sector could thus maintain significant levels of consumption intensity such that both the 'horizontal' (N or CL) and the 'vertical' (consumption intensity, C/N or C/CL) growth components contributed to the total growth in C. Electricity intensification behaviour (connected load per customer, CL/N) was also on the rise in this sector, despite restrictions, indicating increasing investment in energy intensive appliances per customer. All this in fact validates our assumption and application of time-growing slope coefficient (representing increasing consumption intensity) of the consumption function.

The Period I (1960-61 to 1981-82) consumption-number of customers (C-N) relationship is found to be highly significant and is given below:

$$C_t = -67.491 + 0.4326 \ N_t$$
$$(-9.718) \quad (37.718)$$

Adjusted $R^2 = 0.986$; SER = 17.452; F = 1422.67;
DW = 0.1987;
RMSE = 16.639; TIC = 0.0399.

The Period II actual domestic consumption series growing at 11.6 per cent p. a. has a steadily increasing trend and is therefore not required to be smoothed out. $\Delta C/\Delta N$ series is derived from these actual values of Period II, and is found to have a growth rate of 7.71 per cent p. a. on an average, which is then applied to the slope coefficient of the above relationship in order to generate the adjusted consumption of Period II, based upon the same intercept. The complete data base is then analyzed in logarithmic terms and the residuals from the resultant C-N structural relationship are found to have an ARIMA(1, 0, 1) process. The full model is obtained, with t-values in brackets, as:

$$\ln C_t = 41.564 + 0.5644 \ \ln N_t + (1 - 0.9978 \ B)^{-1} (1 - 0.6058 \ B) \ V_t$$
$$(0.491) \quad (22.801) \qquad (151.025) \qquad (4.560)$$

Adjusted $R^2 = 0.9994$; SER = 0.0396; F =20779.39;
DW = 2.011; χ^2 (3, 36) = 14.48;
RMSE = 0.0374; TIC = 0.00304;

Residual mean = 0.002195; Estimated SD = 0.006304;
Maximum (absolute) value of the residual ACF = 0.261 at 9th lag; SE = 0.164;
Maximum (absolute) value of the CCF = 0.311 at 12th lead; SE = 0.164.

The residuals from this model are found to have zero ACF and CCF within the sampling variation. This along with the low chi-square value confirms the model adequacy. Projections from this model presume the same actual (constrained) Period II growth rate of 6.74 per cent for N in the forecast horizon.

ii) Commercial sector

The commercial consumers, who bear the major brunt of the tariff cross-subsidization of the KSEB, constituted about 16 per cent of the number of customers, and partook 8.5 per cent of the total consumption, 19.1 per cent of the sales revenue, and 9.2 per cent of the CL in 1997-98. In 1970-71, the percentage shares were 22.2, 3.5, 16.6, and 9.2 per cent respectively. Along with the domestic consumers, the commercial customers too maintained increasing levels of consumption intensity over the whole period under consideration such that both the horizontal and the vertical growth components made up the overall growth in consumption. Intensification effort, CL/N, also had a rising trend in this sector.

The estimated Period I (1964-65 to 1981-82) C-N relationship is given below:

$$C_t = -22.437 + 0.6913\ N_t$$
$$\quad (-3.325) \quad (19.660)$$

Adjusted R^2 = 0.960; SER = 10.824; F = 386.507;
DW = 0.4930;
RMSE = 10.168; TIC = 0.0452.

The time-series data on the commercial customers' period II energy consumption shows a steadily increasing trend (except in 1987-88, when it fell by about 4 MU over 1986-87) up to 1994-95; in 1995-96, consumption decreased by about 265 MU to 689 MU and again by 40 MU the next year. During 1982-83 to 1994-95, C grew at an average annual rate of 11.41 per cent. However, as smoothing out the entire Period II consumption data at this growth rate would result in unjustifiably substantial inflation of C in the initial years of Period II, this growth rate is applied only for the last 3 years (1995-96 to 1997-98). During the first 5 years of the nineties, C registered an annual average growth rate of 17.6 per cent, and hence these consumption figures are accepted as such without any adjustment. During the remaining period, in the eighties (1982-83 to 1989-90), C is found to have grown at a rate of 9.90 per cent p. a., and at this rate therefore these data are smoothed out. From these Period II consumption data and the corresponding number of customers data, are derived the $\Delta C/\Delta N$ series, found to have an annual average growth rate of 7.13 per cent, which in turn is applied to the slope coefficient of the Period I C-N relationship. The Period II adjusted consumption data thus generated and the Period II actual consumption data are then combined and analyzed after logarithmic transformation in the BARIMA framework to obtain the following model (with t-values in brackets):

$$\ln C_t = \underset{(0.738)}{7.608} + \underset{(23.555)}{0.7124} \ln N_t + (1 - \underset{(60.540)}{0.9905} B)^{-1} V_t$$

Adjusted R^2 = 0.998; SER = 0.0572; F =6197.282;

DW = 1.9382; χ^2 (2, 31) = 18.28;

RMSE = 0.0545; TIC = 0.00486;

Residual mean = 0.007823; estimated SD = 0.00984;

Maximum (absolute) value of the residual ACF = 0.339 at 5th and 15th lag; SE = 0.177;

Maximum (absolute) value of the CCF = 0.311 at 11th lead; SE = 0.177.

iii) LT industrial sector

The LT industrial customers of the KSEB made up only 1.76 per cent of the total number of consumers, and accounted for 6.7 per cent of the energy consumption, 8.4 per cent of the sales revenue, and 20.1 per cent of the CL in 1997-98. In 1970-71, these shares were respectively 3.1, 6.0, 13.1, and 19.6 per cent. Even though they maintained a little higher growth in consumption during Period II than that in their number, their overall period history of consumption intensity remained negative; so was consumption intensity per KW of CL also, in spite of an overall positive trend in investment in electricity intensive appliances as evidenced in increasing intensification per customer (CL/N).

The Period I (1960-61 to 1981-82) C-N relationship in this sector is found to be:

$$C_t = 20.238 + 0.00551 \; N_t$$
$$(3.768) \quad (22.017)$$

Adjusted R^2 = 0.958; SER = 12.007; F = 484.728;
DW = 0.6414;
RMSE = 11.448; TIC = 0.0419.

The Period II consumption series is smoothed out at the actual (Period II) growth rate of 7.26 per cent p. a., and the resultant $\Delta C/\Delta N$ series yields an annual average growth rate of 3.75 per cent, with which we adjust the slope coefficient of Period I C-N function. The adjusted and the actual data are then combined and analyzed; the residuals from the bivariate (C-N) logarithmic structural regression are diagnosed to follow an ARIMA(1, 0, 0) process. The complete model estimated is:

$$\ln C_t = 14.451 + 0.369 \; \ln N_t + (1 - 0.9957\, B)^{-1} \; V_t$$
$$(0.236) \quad (31.933) \qquad\qquad (40.506)$$

Adjusted R^2 = 0.994; SER = 0.0657; F =3054.99;
DW = 2.0465; χ^2 (2, 36) = 15.82;
RMSE = 0.0630; TIC = 0.00580;

Residual mean = 0.00201; estimated SD = 0.0107;
Maximum (absolute) value of the residual ACF = 0.282 at 3rd lag; SE = 0.164;

Maximum (absolute) value of the CCF = 0.228 at 9th lead; SE = 0.164.

iv) HT & EHT industrial sector

With just 0.03 per cent of the total number of the KSEB customers and 14.3 per cent of the total CL, the HT & EHT industrial sector in Kerala took in 26 per cent of the total energy sold, and yielded 35 per cent of the sales revenue in 1997-98; the percentage shares in 1970-71 were 0.04, 28.5, 60.0, and 25.4 per cent respectively. Where the industrial sector has fallen, the domestic sector has risen. Consistently deteriorating consumption intensity of these industrial customers and of their CL marks this sector over the entire period. Surprisingly, even though they could, unlike others, keep up in Period II a growth rate in their number (9.2 per cent) much closer to that of Period I (9.9 per cent), the growth in CL of Period II was more than half-way behind that of Period I, in sharp contrast to the practice in LT industry – are the few major industrial units in Kerala dispensing with electricity intensive techniques in view of power shortage? In this light, energy forecasts presuming higher rates of industrial requirement, as carried out by the State planning Board, appear to be groundless. Hence the tome-growing slope coefficient modeling is not resorted to for this sector. Instead, we just smooth out the Period II consumption series.

When, in period I of no-power-cuts, number of HT-EHT industrial connections grew at an annual rate of 9.93 per cent, then consumption increased by 7.58 per cent p. a. Using this norm, if, in Period II, number of customers could grow at 9.18 per cent p. a., their consumption must have grown at $(9.18) \times (7.58)/(9.93) = 7.01$ per cent p. a., instead of at the actual 2.82 per cent. The Period II consumption series is smoothed out at this unconstrained rate.

The complete data base, comprising the Period I (1960-61 to 1981-82) actual series and the Period II (1982-83 to 1997-98) smoothed out series, when analyzed and fed in logarithmic form into a bivariate (C-N) structural regression, has yielded residuals tentatively found to behave as in an ARIMA(1, 0, 1) process. The full model estimated simultaneously is:

$$\ln C_t = 8.481 + 0.1465 \ \ln N_t + (1 - 0.973\ B)^{-1} (1 - 0.591\ B)\ V_t$$
$$(5.266) \quad (31.373) \qquad\qquad (49.733) \qquad\qquad (4.872)$$

Adjusted R^2 = 0.9902; SER = 0.0707; F =1218.361;

DW = 2.0552; χ^2 (3, 36) = 13.94;

RMSE = 0.0668; TIC = 0.00451;

Residual mean = 0.00239; estimated SD = 0.0113;

Maximum (absolute) value of the residual ACF = 0.292 at 1st lag; SE = 0.164;

Maximum (absolute) value of the CCF = 0.174 at 18th lead; SE = 0.164.

v) Others

Agriculture, public lighting, water works, and licensees (as well as railway traction since 1997-98) constitute ‘others’ here. With 6.3 per cent of the total number of customers and 11.2 per cent of the total CL, this sector consumed 10 per cent of the total energy sold and contributed 7.4 per cent of the sales revenue in 1997-98. The percentage shares in 1970-71 were respectively 3.7, 17.3, 8.0 and 15.6 per cent. Agriculture alone had 6.2 per cent of the total number of customers and 9.4 per cent of CL, accounting for 4.4 per cent total consumption and 1.95 per cent of total sales revenue in 1997-98. The 1970-71 percentage shares were 3.5, 8.8, 2.2 and 3.0 per cent respectively. Consumption intensity in Period II of these customers (C/N) and of their CL (C/CL) was on the rise, but the former had a decaying trend over the whole period and the latter, a growing trend. A steadily deteriorating trend of per customer investment in electricity intensive devices (CL/N) in this sector also marks the whole period. In agriculture, the atmosphere has been as gloomy as in HT & EHT industry over the entire period with fast plummeting consumption intensity (of both the customers and of their CL) and intensification investment (CL/N), even though in Period II consumption intensity (in both terms) gained ground. The time-growing slope coefficient modeling is thus in place in this sector.

The Period I (1964-65 to 1981-82) structural relationship is given by:

$$C_t = 125.284 + 0.00158\ N_t$$
$$(9.655)\quad (6.143)$$

Adjusted $R^2 = 0.697$; SER = 30.340; F = 37.738;
DW = 0.3893;
RMSE = 28.499; TIC = 0.0722.

During Period II consumption of this sector increased by 7.87 per cent p. a., on an average, and the Period II consumption series is smoothed out at this rate. The resultant $\Delta C/\Delta N$ coefficients show an average annual growth rate of 5.43 per cent; and the Period I slope estimate is adjusted at this rate to generate the Period II unconstrained consumption data. The residuals chaffed out from the complete set of data after determining the logarithmic bivariate (C-N) structural relationship are then observed to follow tentatively an ARIMA(1, 0, 0) process. The full model, simultaneously estimated is:

$$\ln C_t = 6.921 + 0.2893\ \ln N_t + (1 - 0.9906\ B)^{-1}\ V_t$$
$$(0.452)\quad (29.426)\qquad\qquad (23.941)$$

Adjusted $R^2 = 0.998$; SER = 0.0800; F =1298.037;
DW = 1.9662; χ^2 (2, 31) = 9.34;
RMSE = 0.0762; TIC = 0.00646;

Residual mean = 0.00723; estimated SD = 0.0138;

Maximum (absolute) value of the residual ACF = 0.275 at 12^{th} lag; SE = 0.177;

Maximum (absolute) value of the CCF = 0.306 at 3^{rd} lag; SE = 0.177.

The simulated (from 1980-81 to 1997-98) and the forecast (since 1998-99, assuming the actual, constrained, Period II growth rate of N) series of each sector in this 'bottom-up' approach are compared with their sum total. The historical ascent of the share of the domestic sector's consumption in total consumption at the (historically validated) cost of the HT-EHT industrial sector blazingly outstands in the whole scenario. By 2004-05, the former outstrips the latter in consumption magnitude. It should be noted that the domestic sector with its 37.5 per cent consumption share had surpassed the HT-EHT industry (with its 36.6 per cent share in consumption) by 1995-96 itself, in terms of constrained consumption, the former growing over the whole period at 13.8 per cent p. a. on an average, and the latter at just 4.9 per cent p. a. However, in the unconstrained scenario (though of the constrained set of number of customers, that grow on an average at 9.2 per cent p. a. in the HT-EHT industrial sector and at 6.7 per cent p. a. in the domestic sector, HT-EHT industrial consumption exceeds the domestic one till 2003-04, as may be expected. The declining trend of consumption intensity in HT-EHT industry persists or rather continues to worsen in the near future, while its steady ascent in the domestic sector ensures the sector's excellence.

It should be remarked that the total consumption obtained from the disaggregated approach to forecasting is almost comparable with that from our consumption model I. APS disaggregated forecasts are not available for a comparison. The State Planning Board's (SPB) bottom-up forecasts, as already mentioned, suffer from estimation biases: the SPB's consumption forecasts for HT-EHT industry, based on groundless optimistic assumption of higher growth rates of consumption, turn out to be overestimates, while those for other sectors, derived from the historical trends of consumption, that remains in a constrained environment, happen to be utter underestimates. An objective approach to and analysis of the diverse ramifications the concrete historical contexts unfold as they evolve over time must have identified the supply-constrained demand vis-à-vis an upsurging intensity-intensification behaviour in the domestic-commercial sectors and that vis-à-vis a lethargic, sluggish, process in the industrial sectors in Kerala, that is a wide expanse of a metropolitan city, not an industrial estate.

Table 1
Estimation Results of the Forecast Models
Period: 1957-58 to 1998-99

1. Linear Trend: Consumption = f (Time)

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	-899.54	-3.76	0.895	348.62	0.157
Time	181.02	18.67	Parameter instability: 3.159**		
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.099	4.90E-17	0.027	0.973	
Residual Analysis					
Normality (χ^2)	Skewness	Kurtosis	SD		
3.35 (0.1870)	0.681	3.25	752.21		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
30.03	9.11	12.95			18.38
(0)	(0.0006)	(0)			(0)

2. Quadratic Trend: Consumption = f (Time, Time Squared)

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	725.61	4.68	0.982	1088.1	0.706
Time	-40.595	-2.44	Parameter instability: 2.371**		
$Time^2$	5.1540	13.750			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.041	8.05E-16	0.004	0.996	
Residual Analysis					
Normality (χ^2)	Skewness	Kurtosis	SD		
0.631 (0.7294)	-0.056	3.59	311.11		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
10.94	++	1.39			10.70
(0)		(0.2548)			(0.0002)

3. Semi-Log Trend: ln Consumption = f (Time)

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	6.082	141.94	0.977	1773.7	0.317
Time	0.073	42.12	Parameter instability: 3.437*		
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.0086	2.10E-12	0.0056	0.994	
Residual Analysis					
Normality (χ^2)	Skewness	Kurtosis	SD		
1.44 (0.4864)	0.119	2.12	0.135		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
10.34	10.14	5.69			13.28
(0)	(0.0003)	(0)			(0)

4. k-transformation (k = 0.5): Consumption = f (Time)

	Estimate	t-value	Adj. R2	F-value	DW statistic
Constant	7.26	3.67	0.973	1479.5	0.487
Time	1.937	31.530			
	Simulation Error Analysis				
	TIC	BP	VP	CP	
	0.05	0.017	0.0074	0.976	
	Residual Analysis				
Normality ($\chi 2$)	Skewness	Kurtosis	SD		
2.08 (0.3541)	-0.29	3.92	377.14		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
+	+	2.16			+
		(0.0845)			

5. First-order Auto-regressive: C_t = f (One-period lagged C_t)

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	15.296	0.22	0.987	3087.1	1.837
$C_{t\text{-}1}$	1.068	55.560	Durbin h:	0.533	
			Parameter instability: 1.149*		
Simulation Error Analysis					
	TIC 0.034	BP 1.33E-14	VP 0.003	CP 0.997	
Residual Analysis					
Normality (χ^2) 17.50 (0.0002)	Skewness -0.387	Kurtosis 6.11	SD 260.94		
Autoregression (F) 7.25 (0.0001)	Heteroscedasticity (F) 10.82 (0.0002)	ARCH (F) 2.78 (0.0355)			RESET (F) 0.076 (0.9132)

6. Logarithmic Auto-regressive: ln C_t = f (ln $C_{t\text{-}1}$)

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	0.231	2.24	0.993	5327.46	1.873
ln $C_{t\text{-}1}$	0.9799	72.99	Durbin h:	0.414	
			Parameter instability: 0.216		
Simulation Error Analysis					
	TIC 0.005	BP 9.99E-14	VP 0.002	CP 0.998	
Residual Analysis					
Normality (χ^2) 1.37 (0.5039)	Skewness 0.411	Kurtosis 2.64	SD 0.075		
Autoregression (F) 1.45 (0.2322)	Heteroscedasticity (F) 0.215 (0.8073)	ARCH (F) 0.336 (0.8872)			RESET (F) 0.369 (0.6936)

7. Linear Trend with AR(1) correction: Consumption = f (Time)					
	Estimate	t-value	Adj. R2	F-value	DW statistic
Constant	-1983.66	-0.19	0.987	1510.03	1.810
Time	-100.570	-0.25	Parameter instability: 1.545*		
AR(1)	1.045	16.40			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.034	4.95E-11	0.003	0.997	
	Residual Analysis				
Normality ($\chi2$)	Skewness	Kurtosis	SD		
14.81 (0.0006)	-0.334	5.87	260.4		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
6.95	8.78	3.64			3.35
(0.0002)	(0.0007)	(0.0109)			(0.0457)
8. Quadratic Trend with AR(1) Correction: C = f (Time, Time Squared)					
	Estimate	t-value	Adj. R2	F-value	DW statistic
Constant	1413.96	1.80	0.989	1162.47	1.625
Time	-110.12	-1.55	Parameter instability: 2.185*		
Time2	6.65	4.72			
AR(1)	0.709	4.99			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.031	5.60E-13	0.003	0.997	
	Residual Analysis				
Normality ($\chi2$)	Skewness	Kurtosis	SD		
11.34 (0.0034)	-0.481	5.39	239.37		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
5.76	++	4.99			3.67
(0.0007)		(0.0019)			(0.0350)
9. Semi-log Trend with AR(1) correction: ln C = f (Time)					
	Estimate	t-value	Adj. R2	F-value	DW statistic
Constant	6.25	39.54	0.993	2927.36	1.766
Time	0.068	12.82	Parameter instability: 0.810		
AR(1)	0.798	9.540			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.004	1.13E-13	0.002	0.998	
	Residual Analysis				
Normality (χ^2)	Skewness	Kurtosis	SD		
0.066 (0.9674)	0.095	3.05	0.07		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
1.36	0.363	0.202			0.091
(0.2655)	(0.6977)	(0.9591)			(0.9132)

10. Partial Adjustment (Short-run Growth Rate) Model: ln C_t = f (ln C_{t-1},Time)

<table>
<tr><td rowspan="4">Constant
Time
ln C_{t-1}</td><td>Estimate</td><td>t-value</td><td>Adj. R^2</td><td>F-value</td><td>DW statistic</td></tr>
<tr><td rowspan="3">1.31
0.0136
0.798</td><td rowspan="3">2.61
2.20
9.54</td><td>0.993</td><td>2927.36</td><td>1.766</td></tr>
<tr><td colspan="3">Durbin h: 0.887</td></tr>
<tr><td colspan="3">Parameter instability: 0.810</td></tr>
<tr><td colspan="6">Simulation Error Analysis</td></tr>
<tr><td></td><td>TIC
0.004</td><td>BP
1.13E-13</td><td>VP
0.002</td><td>CP
0.998</td><td></td></tr>
<tr><td colspan="6">Residual Analysis</td></tr>
<tr><td>Normality (χ^2)
0.066 (0.9674)</td><td>Skewness
0.095</td><td>Kurtosis
3.05</td><td>SD
0.07</td><td colspan="2"></td></tr>
<tr><td>Autoregression (F)
1.36
(0.2655)</td><td>Heteroscedasticity (F)
1.77
(0.1562)</td><td>ARCH (F)
0.202
(0.9591)</td><td colspan="3">RESET (F)
2.36
(0.1082)</td></tr>
</table>

Note:
1. *' and '**' indicate statistical significance at 5 and 1 per cent respectively.
2. + = not available in non-linear least squares
3. ++ = near singular matrix
4. Figures in brackets are the corresponding p-values.
5. C = Electricity Consumption in the State (Million Units)
6. Adj. R-squared = Adjusted R-squared;
7. TIC = Theil inequality coefficient
8. ln = Natural log
9. BP = Bias proportion; VP = Variance proportion
10. CP = Covariance proportion
11. AR(1) = Estimate of first order auto-regression coefficient
12. Parameter instability = Joint (F-) test statistic for parameter constancy
13. ARCH (F) = Autoregressive Conditional Heteroscedasticity (F) statistic (5 lags)
14. RESET (F) = Regression Specification Test (F) statistic

Table 2
Multi-variate Econometric Models
Period: 1960-61 to 1998-99

Model 1. $C_t = f(N_t, PCI_t, AR_t)$

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	-129.66	-0.23	0.993	1686.1	1.306
N_t	1.175	18.53	Parameter instability: 1.591*		
PCI_t	1.127	4.23			
AR_t	-30.53	-2.30			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.024	1.64E-14	0.002	0.998	
Residual Analysis					
Normality (χ^2)	Skewness	Kurtosis	SD		
14.19 (0.0008)	-1.161	4.83	191.94		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
3.08	2.64	0.443			3.44
(0.0233)	(0.0340)	(0.8145)			(0.0434)

Model 2. $\ln C_t = f(\ln N_t, \ln PCI_t, \ln AR_t)$

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	0.237	0.27	0.990	1266.21	0.635
$\ln N_t$	0.683	29.61	Parameter instability: 2.501**		
$\ln PCI_t$	0.539	4.71			
$\ln AR_t$	-0.398	-3.26			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.005	1.03E-13	0.002	0.998	
Residual Analysis					
Normality (χ^2)	Skewness	Kurtosis	SD		
1.006 (0.6048)	-0.048	2.22	0.077		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
6.08	0.695	0.791			6.95
(0.0005)	(0.6552)	(0.5652)			(0.0029)

Model 3. C_t = f(N_t, PCI_t, AR_t, Time)

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	-1261.12	-1.92	0.994	1505.83	1.609
N_t	0.876	7.12	Parameter instability: 1.837*		
PCI_t	1.567	5.37			
AR_t	-18.287	-1.41			
Time	30.180	2.76			
	Simulation Error Analysis				
	TIC 0.022	BP 7.54E-15	VP 0.001	CP 0.999	
	Residual Analysis				
Normality (χ^2) 232.32 (0.0)	Skewness -1.22	Kurtosis 5.91	SD 173.47		
Autoregression (F) 3.44 (0.0147)	Heteroscedasticity (F) 3.03 (0.0129)	ARCH (F) 1.10 (0.3821)			RESET (F) 0.457 (0.6373)

Model 4. ln Ct = f(ln Nt, ln PCIt, ln ARt, Time)

	Estimate	t-value	Adj. R2	F-value	DW statistic
Constant	-4.52	-1.73	0.991	1023.51	0.748
ln Nt	1.040	5.53	Parameter instability: 2.155**		
ln PCIt	0.916	4.07			
ln ARt	-0.295	-2.29			
Time	-0.040	-1.92			
	Simulation Error Analysis				
	TIC	BP	VP	CP	
	0.005	2.56E-13	0.002	0.998	
	Residual Analysis				
Mean	SD	Skewness	Kurtosis	JB-Normality	
3.67E-08	0.073	-0.165	2.46	0.652 (0.7219)	
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
4.88	0.756	1.05			7.76
(0.0023)	(0.6430)	(0.4115)			(0.0017)

Model 5. ln C_t = f(ln N_t, ln PCI_t, ln AR_t, ln C_{t-1})

	Estimate	t-value	Adj. R^2	F-value	DW statistic
Constant	-0.004	-0.01	0.994	1476.77	1.37
ln N_t	0.301	3.45	Durbin h: 2.788		
ln PCI_t	0.300	2.83	Parameter instability: 1.445		
ln AR_t	-0.217	-2.05			
ln C_{t-1}	0.537	4.50			
Simulation Error Analysis					
	TIC	BP	VP	CP	
	0.004	3.69E-13	0.001	0.999	
Residual Analysis					
Normality (χ^2)	Skewness	Kurtosis	SD		
0.1.65 (0.4379)	-0.110	3.26	0.060		
Autoregression (F)	Heteroscedasticity (F)	ARCH (F)			RESET (F)
2.77	0.461	0.0004			5.92
(0.0774)	(0.8718)	(0.9833)			(0.0206)

Note:
1. *' and '**' indicate statistical significance at 5 and 1 per cent respectively.
2. C = Electricity Consumption in the State (Million Units)
3. N = Number of Electricity Consumers
4. PCI = Per Capita State Income (at 1980-81 prices)
5. AR = Average Price (Revenue) (at 1981-82 prices)
6. Figures in brackets are the corresponding p-values.

REFERENCES

1. Banerjee, Nirmala (1979) *Demand for Electricity*, Centre for Studies in Social Sciences, Calcutta, K. P. Bagchi and Co., Calcutta.

2. Box, G. E. P. and Jenkins, G. M. (1970) *Time Series Analysis*, Holden-day, San-Francisco.

3. Government of India (1974), The Report of the Fuel Policy committee of India.

4. Government of Kerala (1977) *Report of the Steering Committee on Energy and Power, Ninth Five Year Plan, 1997-2002*, State Planning Board, Feb., Thiruvananthapuram,, Kerala, India.

5. Henderson, P. D. (1875) *India: The Energy Sector*, Oxford University Press.

6. IRTC and IEI, (1996) *Exercises for Integrated Resource Planning for Kerala: End-Use Analysis – An Empirical Study: Technical Report I – Electricity*, Palakkad, Kerala).

7. Lencz, Imrich (1977) 'Prognoses of the consumption of energy, The Czechoslovak experience', in William D. Nordhaus (ed.)(1977) pp. 83 – 94.

8. Makridakis, S. and Wheelwright, S. C. (1978) *Forecasting Methods and Applications*, Wiley, New York.

9. Nordhaus, William D. (ed.) (1977) *International Studies of the Demand for Energy*, North-Holland, Amsterdam.

10. Pachauri, R. K. (1975) *Dynamics of Electrical Energy Supply and Demand: An Economic Analysis,* Praeger Publishers, New York.

11. Parikh, Jyoti K. (1981) *Modeling Energy Demand for Policy Analysis*, Planning Commission, Government of India, New Delhi.

12. Pillai, P. P. (1981) *Dynamics of Electricity Supply and Demand in Keralam – A Macro-Econometric Analysis*, Agricole publishing Academy, New Delhi.

13. Pillai N. Vijayamohanan (2001) *Electricity Demand Analysis and Forecasting – The Tradition is Questioned!* Working Paper No. 312, Centre for Development Studies, Thiruvananthapuram, Kerala, India.

14. Pindyck, R. S. and Rubinfeld, D. L. (1981) *Econometric Models and Economic Forecasts*, 2nd ed., Mc-Graw-Hill.

15. Reddy, Amulya Kumar N. (1990) *Wanted Immediately: An Energy Strategy For India (I): DEFENDUS Energy Strategy Offers Hope*, Bangalore: http://amulya-reddy.org.in/Publication/nodate_chapter4-020503.pdf

16. Taylor, L. D. (1975) 'The demand for electricity: a survey', *Bell Journal of Economics and Management Science*, Vol. 6, No. 1, (Spring).

17. Taylor, L. D. (1977) 'The demand for energy: a survey of price and income elasticities', in William D. Nordhaus (ed.) (1977).

18. World Bank (1979) *India: Economic Issues in the Power Sector*, A World Bank Country Study, November.