ornl

OAK
RIDGE
NATIONAL
LABORATORY

UNION
CARBIDE

NUREG/CR-12[illegible]6
ORNL/NUREG/TM-379

A Conditional Model of Peak and Minimum Loads and the Load Duration Curve for Electricity

John L. Trimble
Dennis E. Stallings
Ben Thomas

Prepared for the
U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Under Interagency Agreement DOE 40-550-75

OPERATED BY
UNION CARBIDE CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 , ..rt Royal Road, Springfield, Virginia 22161

Available from
GPO Sales Program
Division of Technical Information and Document Control
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

NUREG/CR-1256
ORNL/NUREG/TM-379
Dist. Category RE

Contract No. W-7405-erg-26

Energy Division

A CONDITIONAL MODEL OF PEAK AND MINIMUM LOADS AND THE LOAD DURATION CURVE FOR ELECTRICITY

John L. Trimble
Dennis E. Stallings
Ben Thomas

Manuscript Completed — January 1980

Date Published — May 1980

Prepared for the

U.S. Nuclear Regulatory Commission
Office of Nuclear Reg tory Research
Washington, D.C 20555

Under Interagency Agreement DOE 40-550-75

NRC FIN No. B-0190-8

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37830
operated by
UNION CARBIDE CORPORATION
for the
DEPARTMENT OF ENERGY

CONTENTS

ACKNOWLEDGMENTS

We thank W. S. Chern and R. C. Tepel of ORNL, T. Takayama of DOE, D. L. Kaserman of the University of Tennessee, and D. R. Bonney of the Tennessee Valley Authority for helpful commentary on earlier versions of this report. We thank all the cited utilities for providing tapes of hourly loads, but special thanks go to Charles J. Roncaioli and Margaret Brown of Northeast Utilities and to Michael C. Calimano of Central Hudson Gas and Electric Corporation. We also thank C. A. Gallagher and B. D. Holcomb for computational assistance and Lana Melton and Lillian Cochran for typing assistance.

ABSTRACT

This report presents a model that extends the traditional model of electricity demand to account for intra-period load variation, the kind of variation that is important for evaluating marginal-cost-reflecting price structures. The time-of-day rate is one such price structure. The traditional model of electricity demand explains inter-period demand variation. It says nothing about load variation. The report explains how a model that integrates with previous studies of electricity demand might be formulated. It specifies two concrete models within this framework and estimates them for a number of different utility companies.

The model's within-sample-period performance in predicting peak loads is presented for one version of the model extension along with estimations for other variations. In addition a number of plots of actual load distributions, a summation of load variation information, against the actual load distributions, are presented and used to evaluate the performance of specific models.

1. INTRODUCTION

Before proceeding to the topic of this report, it may forestall confusion later if a few potentially misleading terms are clarified now. Economists' relatively recent interest in electricity consumption, historically a domain for engineers, has resulted in confusing uses of the same terms. Economists traditionally use the term 'demand' to mean one thing while engineers use it to mean another. To avoid potential misunderstanding, 'demand' in this report shall be used in the economist's sense to refer to 'kilowatt hours (kWh) of consumption per unit of time.' Another potentially confusing term 'load' shall refer to 'instantaneous consumption measured in kilowatts (kW).' And hourly units shall be considered to be adequate to measure instantaneous consumption; hence, we shall use the term 'hourly load' not 'hourly demand'; thus 'peak load' will refer to 'maximum hourly load' and 'minimum load' to 'minimum hourly load'; no other discrepancies between 'demand' and 'load' will be sanctioned.

Since demand refers to consumption per unit of time, it will prove useful later to attach a name to the time unit. The term 'accounting period' will be used to refer to 'the unit of time used to measure demand in a given application.' Thus accounting periods of a month, a quarter or a year might arise depending on what sort of model is being considered. In this report the accounting period is a year because a model of electricity demand per annum is being considered. Hence, in this report, when the term 'peak load' is used, for example, it will refer to 'peak load per annum.' In portions of the report, however, the term 'accounting period' will be explicitly used. This is to emphasize that the statements can apply to any model of electricity demand not simply a model of annual electricity demand.

Having hopefully laid aside potential terminological difficulties, the remainder of this section addresses the purpose of the report and the motivation behind it.

Evaluation of the necessity to build additional electric power generation capacity has been a recurring problem for the Nuclear Regulatory

Commission (NRC), which regulates nuclear power installations. From the NRC's point of view, "need for power" assessments, as they are called, must be independent of an applicant utility's assessment. In May 1976, the NRC commissioned ORNL to develop a forecasting model of electricity demand to support their independent assessments. The resulting model was first published in October 1978;[1] it provides forecasts of state-level electricity demand (SLED) and average price.

However, the SLED model (as it will be called in this report) is a compromise between the NRC's need to have a model with wide geographic coverage and one that forecasts at the service-area level. The SLED model also does not forecast peak load which is important for determining capacity needs. Yet a model with all these capabilities that has been validated for all service areas in the continental U.S. (as was done for states with SLED) is out of the question. To satisfy the NRC's needs, a compromise solution was undertaken: extend the SLED model to forecast demand, average price and peak load at the service-area level but validate the model for only a few service areas geographically dispersed to represent climatic differences in the U.S. The resulting model can then be estimated on a case-by-case basis as required.

These extensions to SLED have been developed along two lines. One extension focuses on disaggregating SLED predictions of demand and average price to service-area predictions of the same quantities. It constructs models of demand ratios (service-area to state) and average price ratios.[2] The other extension which is presented in this report focuses on estimating peak load. But the model presented in this report also estimates minimum load and the corresponding load duration curve.

The process by which load duration curve, peak[3] and minimum[4] load estimates are produced is depicted in Fig. 1. The SLED model estimates state-level demand and average price for each consuming sector (denoted, respectively, by Y_R, Y_C, Y_I and P_R, P_C, P_I). These estimates are then disaggregated into service-area estimates of demand and average price (denoted, respectively, by y_R, y_C, y_I and p_R, p_C, p_I) by the service-area disaggregation model.[5] An estimate of aggregate demand ($y = y_R + y_C + y_I$) is then computed. Aggregate demand, y, is then fed

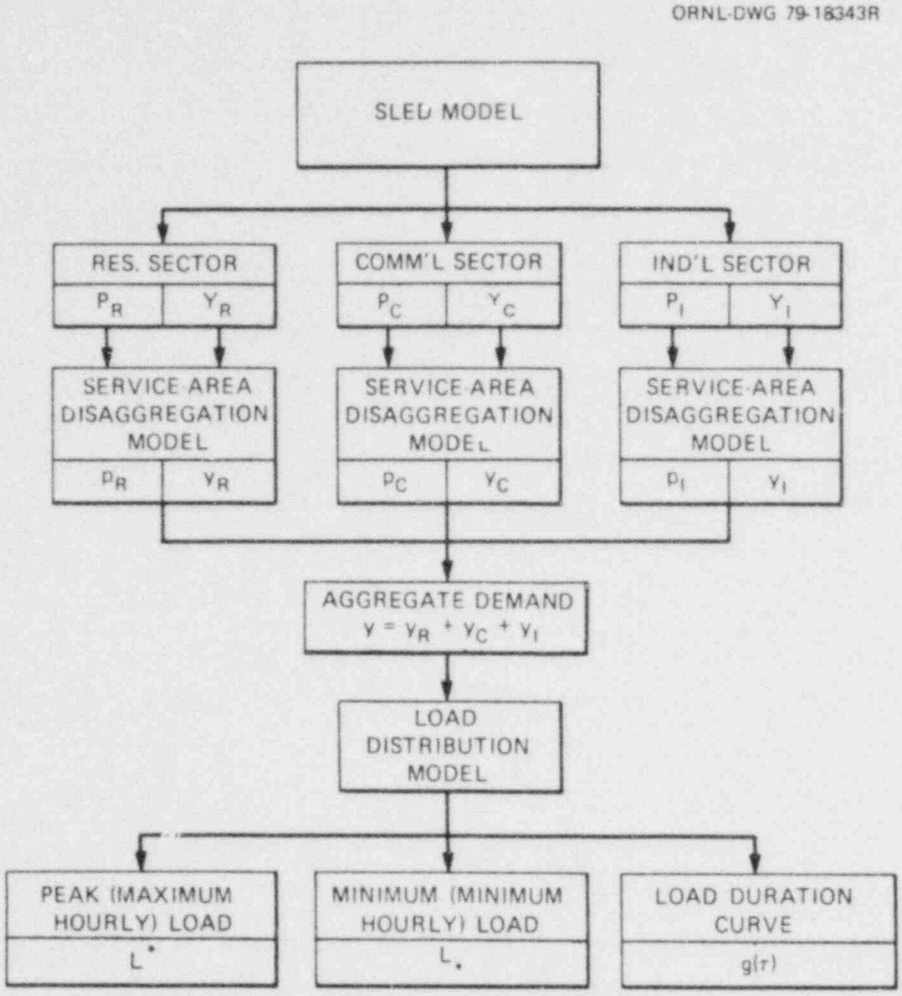

Fig. 1. The LD model extends the SLED model to forecast peak and minimum loads and the load duration curve.

into the load distribution model (LD) which estimates the service-area load duration curve, peak and minimum loads conditioned on the aggregate demand estimate.

The LD model is named for its relationship to the load duration curve; the relationship between the two is described in the next section.

Before proceeding on, it may be useful to note that the LD model can be used in conjunction with almost any model of electricity demand. The essential requirement is that the conditional forecast of aggregate demand (y) be expressed in the same time unit (i.e., accounting period) as the LD model. For instance, an annual aggregate demand estimate (y) would not be useable if the LD model were set up to take a monthly aggregate demand estimate. The electricity demand and LD models must be conformable in the sense that they both use the same accounting period.

2. BACKGROUND

This section deals with two topics: (1) load duration curves and how they can be estimated and (2) the fundamental loss of important information that always occurs in models of annual electricity demand. A major point in the report is that a valuable portion of the lost information can be recovered by estimating the load duration curve in conjunction with aggregate electricity demand. Thus topics (1) and (2) above are closely related to one another. This section provides an overview of how (1) and (2) can be related in a model framework; but, details are left to section four.

2.1 Load Duration Curves and Their Estimation

The load duration curve is an analytical tool widely used in the electric power industry. It provides a powerful summary of how load has been distributed over an accounting period, usually a year. Figure 2 exhibits a typical annual load duration curve. The curve is designated by $g(\tau)$ where τ is the proportion of elapsed time in the accounting period (i.e., $0 \leqslant \tau \leqslant 1$); and L [identical to $g(\tau)$] designates continuously measured electricity load which is always greater than or

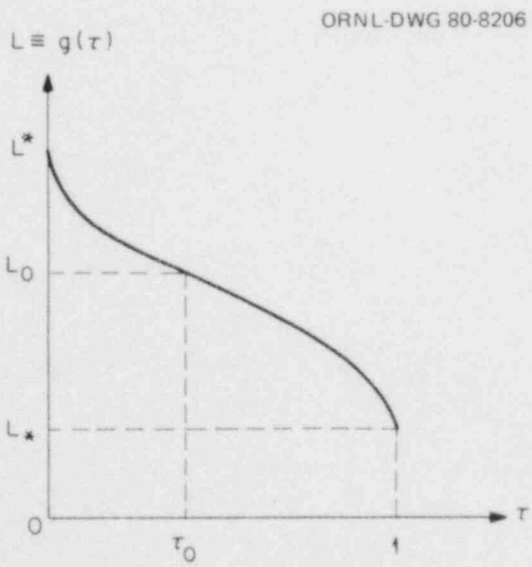

Fig. 2. Typical load duration curve.

equal to minimum load L_*, and less than or equal to peak load L^*. Thus, for any proportion of the accounting period's total time, τ_0, the load duration curve tells what level of load was equaled or exceeded.

It can be shown that the area under the load duration curve, $g(\tau)$, in Fig. 2 can be expressed as a function of aggregate demand, y, of Fig. 1. In fact, this link is required to estimate a load duration curve from the SLED ι as outlined in Fig. 1. To establish this relationship, it is necessary to re-express the load duration curve as a cumulative probability distribution, F(L). In the remainder of the report, 'F(L)' shall be referred to as the 'load distribution' which will be abbreviated as LD. Note then that as a distribution function F(L) must satisfy the following condition:

$$0 \leq F(L) \leq 1 \quad \text{for} \quad 0 \leq L \leq L^* . \tag{1}$$

Since $L \equiv g(\tau)$ is the expression for the load duration curve, F(L) can be expressed in terms of the inverse function for $g(\tau)$. This inverse function is

$$g^{-1}(L) = \begin{cases} 1 \text{ if } 0 \leq L < L_* \\ G(L) \text{ if } L_* \leq L \leq L^* \end{cases} \tag{2}$$

where G(L) is the inverse of the load duration curve between minimum and peak loads and is assumed to be twice differentiable. The distribution function F(L) satisfying condition (1) may then be expressed in terms of G(L)[6] as:

$$F(L) = 1 - g^{-1}(L) = \begin{cases} 0 \text{ if } 0 \leq L < L_* \\ 1 - G(L) \text{ if } L_* \leq L \leq L^* . \end{cases} \tag{3}$$

It is easy to see that the portion of F(L) below L_* being equal to zero contributes nothing to F(L) and hence can be dropped.

The load distribution is illustrated in Fig. 3.

ORNL-DWG 80-8208

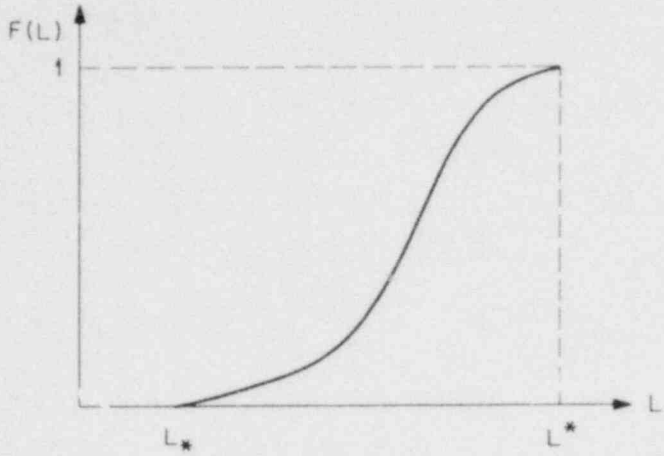

Fig. 3. Load distribution.

The probability density function f(L) corresponding to F(L) — defined simply as the first derivative of F(L), i.e., $f(L) = \frac{d}{dL}F(L)$ — is exhibited in Fig. 4. This density is required for establishing the relationship between aggregate demand, y, in Fig. 1 and the load duration curve g(τ) of Figs. 1 and 2. Their relationship is given by:

$$\int_0^1 g(\tau)d\tau \equiv \int_{L_*}^{L^*} Lf(L)dL = E(L) \ . \tag{4}$$

ORNL-DWG 80-8207

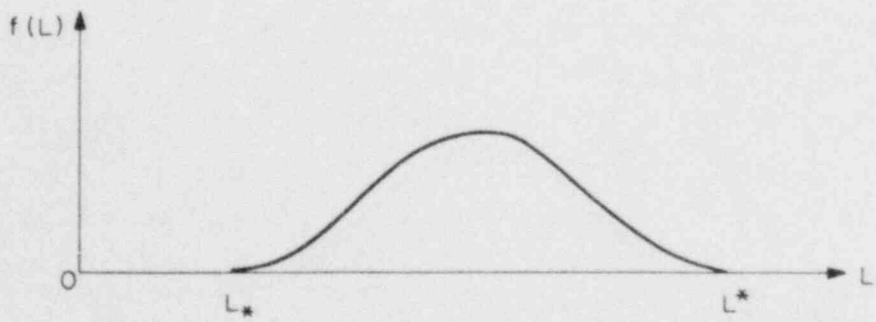

Fig. 4. Probability density of load.

The term to the right of the identity sign in (4) is derived by making the transformation L ≡ The term on the far right designates the expectation (or average of L which $\int_{L_*}^{L^*} Lf(L)dL$ defines.

When L is observed in finite units (hourly in this case), then aggregate demand, y, is equal to average hourly load, E(L), times the number of observations, N, of hourly load in the accounting period — i.e.,

$$y = NE(L) \tag{5}$$

where E(·) designates the mathematical expectation or average of L. Thus aggregate demand, y, is related to the corresponding load duration curve through E(L), average hourly load. One could think of average hourly load as a summary measure characterizing the load duration curve; indeed, this would be true under a very restrictive assumption. A more adequate representation of the load duration curve requires additional summary measures however.

Equations (4) and (5) link, respectively, the load duration curve to the load distribution and aggregate demand to average hourly load. We can thus address the problem of estimating the load duration curve as the problem of estimating the load distribution. The latter problem can then be addressed by specifying a parametric family of probability distributions to represent the load distribution. The problem then becomes one of estimating the parameters of the family, a problem for which a number of widely known methods are available. We choose from among these methods one known as the "method of moments."

Estimating the load distribution (and hence the load duration curve) involves estimating the parameters of the family of probability distributions as well as the parameters designating peak load (L^*) and minimum load (L_*). The problem of estimating the family's parameters by the method of moments is greatly simplified if hourly loads are normalized to separate L^* and L_* from the calculations of the other parameters. The following definitions to normalize loads will serve this purpose:

$$Z = (L^* - L_*)^{-1}(L - L_*) \tag{6a}$$

$$z_i = (L^* - L_*)^{-1}(y_i - L_*) \; ; \qquad (6b)$$

L designates continuous load measured between L_* and L^*; Z designates continuous normalized load measured between zero and one; y_i designates the ith hourly load and is considered to be an observation of L; and z_i designates the ith normalized hourly load and must therefore be considered as an observation of Z.

We shall refer to the load distribution arising from the transformation (6a) as the normalized load distribution.

Estimation of the normalized load distribution (and hence the normalized load duration curve) by the method of moments involves equating theoretical moments of the normalized load distribution with their corresponding sample moments. The number of theoretical moments equated to sample moments must equal the number of parameters in the normalized load distribution. Although any type moments may be used, the customary practice is to use either raw moments or central moments (i.e., moments about the mean). We use central moments.

Suppose that the normalized load distribution has n parameters to be estimated. (Recall that the parameters L^* and L_* are eliminated from these calculations.) Estimates of the n parameters may be calculated by solving the following n equations (LF designates load factor):

$$N\mu_r = \sum_{i=1}^{N} (z_i - \bar{z})^r \quad (r = 2, \ldots, n) \qquad (6c)$$

$$N\mu_1 = N\bar{z} = \sum_{i=1}^{N} z_i = [N/(L^* - L_*)][L^*(LF) - L_*] \; . \qquad (6d)$$

As long as N, the sample size of hourly loads, remains relatively close to the 8760 hours in a year (or in whatever accounting period is being used), the estimates of parameter values obtained by solving (6c)-(6d) may be treated as population values rather than as statistics. As the sample size N decreases, the need to account for the sampling distribution of the normalized load distribution parameter estimates increases. This procedure is not a very suitable one in small samples mainly because method-of-moments estimators are not as efficient as ones obtained by other methods; and, moreover, their efficiency tends to decrease as higher moments (i.e., higher r) are used.

In this report system hourly loads are used to estimate normalized load distribution parameters. Typically ninety-five percent or more of each year's observations are included. Thus the small sample problem is not a problem here; yet, it is still important to recognize its potential presence.

So far we have established how a normalized load duration curve can be estimated from hourly load data. Two additional estimation problems still need to be considered: (1) how to estimate peak (L^*) and minimum (L_*) loads and (2) how to estimate year-to-year changes in the normalized load distribution in terms of explanatory variables of policy interest. Both problems, it turns out, are solved in the same way: by specifying an appropriate number of econometric equations equal to the number of parameters in the specified normalized load distribution plus one additional econometric model for estimating peak load. This will yield n+1 estimates of moments of normalized load and n+2 parameters; the last parameter is identified by using the annual estimate of total energy (i.e., aggregate demand), y; recall that the load distribution model assumes that an estimate of y is conditionally given (see Fig. 1).

From these n+1 econometric models and the conditionally given prediction of aggregate demand, a total of n+2 predicted values are obtained for each year. Each predicted value is a prediction of one of the moments discussed earlier. Thus, load distribution predictions are calculated by substituting the predicted moments for the sample moments in equations (6c)-(6d) and solving simultaneously. Load distribution (and hence load duration curve) estimates are predicted for each year in this way.

In section four the specification of the econometric models and a parametric family of distributions to represent the load distribution are discussed in more detail.

2.2 Temporal Aggregation

We are concerned in this report with a model (i.e., the SLED model) that predicts annual electricity demand based, of course, on historical observations of electricity consumption. Such a model provides useful information on the underlying determinants and pattern of electricity consumption over time — that is, from year to year in this case. But such a model is virtually useless if information about intra-year variation in electrical load is required. Evaluation of time-varying — e.g., seasonal and time-of-day — rates requires such information.

One might say that knowledge of intra-year variation is lost since annual consumption is the sum of hourly loads in a given year. We refer to this process of summing over time (i.e., summing hourly loads in this case) as temporal aggregation. Thus temporal aggregation poses a fundamental problem if all that is available on the one hand is annual consumption data and, on the other, one needs, for example, to evaluate the policy implications of a time-varying rate.

The data typically available from published sources to estimate annual electricity demand is of two forms: unit averages and point samplings. A unit average is the aggregation of a variable that varies continuously over time. Annual electricity consumption, for example, is the unit average of electrical load per year. Point sampling refers to variables that do not vary continuously over time; measurement of the variable is made at a particular point in time. The housing stock is a variable for which point samples might be taken. The annual housing stock might refer to the stock of houses in place at the beginning of the year, at the end or at some intermediate point.

Perhaps some insight can be gained into the temporal aggregation problem if we first specify a continous time model of electricity load then compare it to the annual model often estimated. A model of electricity load varies continuously in time and, hence, theoretically is not subject to distortion arising from temporal aggregation. Let aggregate (i.e., cross-sectionally aggregate) electricity load be expressed as follows as a functional in continuous time, δ:

$$y(\delta) = x(\delta)'b + u(\delta) , \quad 0 \leq \delta \leq T \tag{7}$$

where δ denotes continuous real time in an interval of length T; $y(\delta)$ is aggregate electricity load; $x(\delta) = [x_1(\delta), x_2(\delta), \ldots, x_M(\delta)]'$ is a vector of M explanatory variables measured in continuous time; $b = (b_1, b_2, \ldots, b_M)'$ is a set of M parameters; and $u(\delta)$ is a stochastic residual.

The model of equation (7) has no temporal aggregation problem. It conforms to our intuitive conception of how electricity is consumed. It can be used to evaluate the policy implications of time-varying rates. Yet, even with these appealing properties, the continous time model of equation (7) can onl· be used in very special circumstances. The data required to estimate its parameters rarely, if ever, are available from published sources. Such a model is useful primarily in experiments where the required data is assembled. In the residential sector, for example, one important explanatory variable requires the continuous measurement of usage for the electric space heating system, a variable which is rarely available. Even, if more realistic compromise variables, like temperature near the household, are considered, they practically never can be adequately matched to hourly usage not to mention that the data collection effort required for such variables is a obviously formidable one.

Because its burdensome data requirements are rarely met, the continuous time model of electricity load is usually replaced in much applied work with the following model:

$$Y_t = X_t'B + U_t \quad t = 1, 2, \ldots, S . \tag{8a}$$

In this model electricity load is unit averaged, i.e.,

$$Y_t = \int_{t+a-1}^{t+a} y(\delta)d\delta , \quad 0 \leq a < 1 \quad t = 1, \ldots, S \tag{8b}$$

where t enumerates an accounting period of length T and "a" determines where in continuous time the measurements are made; we refer to unit averaged electricity load as electricity consumption if it is an observation and electricity demand if it is a theoretical concept used in a model of electricity demand. In the electricity demand model of equation (8a), Y_t is electricity demand; but the parameter vector $B = (B_1, \ldots, B_M)$ is estimated with electricity consumption data. The vector of explanatory variables is comprised of both unit-averaged values, i.e.,

$$X_{jt} = \int_{t+a-1}^{t+a} x(\delta)d\delta \ , \quad t = 1, \ldots, S \ ; \quad j = 1, \ldots, M \qquad (8c)$$

and point-sampled values, ie.,

$$X_{jt} = x(\delta_{0_t}) \ , \quad t = 1, \ldots, S \ , \quad j = 1, \ldots, M$$

where δ_{0_t} is the point in real time that X_{jt} is measured (e.g., at the beginning of the year).

Obviously, the model of equations (8) provides no information about intra-period variation in aggregate electrical load. In this sense it suffers from a loss of information due to temporal aggregation. As a result it cannot address policy questions having to do with the intra-year variation in electricity load. The most efficient pricing structures, however, are those that vary with the marginal cost of generating electricity; and, since the marginal cost of generation tends to vary continuously, time-varying price structures that reflect marginal cost variation are more efficient and most actively discussed as instruments to stimulate conservation of electrical energy resources.

An extension of a model of annual electricity demand to deal with intra-year load variation adds, therefore, an important policy-considering capability. We attempt in this report to move in this direction by extending the SLED model to predict peak and minimum loads and the load duration curve.

3. LITERATURE REVIEW

Existing studies that attempt to provide a means for predicting peak electricity load are sparse. The state of development of a demand model with this capability is rudimentary at best. On one hand is a long-standing industry practice that extrapolates peak load from load factor and demand estimates and on the other is the practice of estimating peak load directly by applying the peak-load pricing model. In addition there have been a number of efforts to estimate continuous-time models of electricity load which are also capable of estimating peak load. Related to this literature are the attempts by some to estimate load duration curves but the link between this work and empirical studies of electricity demand and peak load is not clear.

A recent Charles River Associates (CRA) report[7] reviews forecasting procedures for peak load. Since that report appeared, there have been some additions to the existing literature but none that were not covered in the taxonomy presented therein. A review of the literature in the time-of-day pricing area appears in Aigner and Poirier.[8] Related to this literature but beyond the scope of the CRA report is a small body of literature of methods for estimating load duration curves. Since that work is pertinent to the work presented in this report, it is included in the following review.

The CRA report classifies peak load forecasting methods into three generic groups which for convenience we re-name: load factor methods, direct methods, and load curve methods. Two of these methods, the load factor and direct methods, depend in some way on an estimate of aggregate electricity demand; the continuous-time model of electricity load does not. The load factor method requires an estimate of aggregate demand to calculate the peak load estimate; the direct method requires an aggregate demand estimate if the load factor is to be estimated too. The first of these classifications, the load factor method, indicates that the load factor is being modeled, the second indicates that peak load is being modeled directly or endogenously; and, the third indicates that the load curve is being modeled.

3.1 Load Factor Methods

Load factor methods are popular in the electric utility industry.[9] Given an estimate of aggregate demand and an estimate of the corresponding load factor, peak load is estimated by substituting the aggregate demand estimate into the load factor equation (formed by equating the load factor estimate to its definition) and solving for peak load.

The simplicity of this method accounts for its popularity in the utility industry; its weak link to well-accepted theory accounts for its infrequent use in the economics literature. Actually, the load factor and aggregate demand presents the same information as peak load and aggregate demand. And, if peak load, the load factor and aggregate demand are all estimated by double logarithmic or semi-logarithmic models with normally distributed residuals, transforming from one to the other is very simple because of the closure property of the normal distribution under subtraction.

3.2 Direct Methods

The direct method expresses peak load as an endogenous variable allowing it to be estimated directly. This method is very popular amongst economists because it fits the paradigm of the peak-load pricing model.[10] There peak and off-peak periods have been treated as separate commodities. There are, however, two important differences between the direct method and the peak load pricing (PLP) model. In the PLP model, peak is defined over some period of usually several hours; in the direct method, peak is an instantaneous value. And, in the PLP model, peak and off-peak demand are priced separately. But in the direct method, this is not necessarily the case.

Moving to specific studies, the early work comes from the forecasting literature primarily in engineering. These studies classify into weather-related models and behavioral and weather-related models. An illustration of the first model type is given in Galiana.[11] Because historical data appear to show that a certain portion of electricity demand is sensitive to weather changes, the weather-related model is built around the maintained hypothesis that peak load is composed of additive non-weather-sensitive and weather-sensitive components plus a random component. Often in the early work, a non-weather component was added as something of an afterthought and thus treated as constant.[12] But, in such cases, the model is little more than a model of peak-hour weather changes between accounting periods. Gupta[13] extended the crude weather-related/behavioral model somewhat by allowing annual non-weather-sensitive peak load to be "influenced by economic conditions, energy conservation and annual kilowatthour sales." However, even with this extension, his approach is not a very satisfactory one. It fails, for example, to take account of the interaction between the behavioral use of electricity-using equipment in responding to weather changes. That aspect of electricity consumption alone suggests that weather-sensitive and non-weather-sensitive demand are not additively separable, cf. Hausman, McFadden, and Kinnucan.[14]

Spann and Beauvais,[15] Murray, Spann, Pulley, and Beauvais,[16] Betancourt and Habermann,[17] and Uri[18] provide examples of more sophisticated attempts to estimate weather-related/behavioral models. The Spann and Beauvais[19] work is straightforward in applying the direct method. Peak load is assumed to be a function of a set of explanatory variables (some of them computed from other data) in a single-equation model; in their model, peak load is not explicitly related to aggregate demand. Murray et al.[20] and Betancourt and Habermann[21] do relate aggregate demand and peak load to one another. In both of these studies both aggregate demand and peak load are estimated by statistically independent single-equation models. Neither attempted to tie the two together through the stochastic residual despite the fact that peak loads are contained in aggregate consumption. Murray et al.[22] an expansion of the Spann and Beauvais study,[23] assumed stochastic independence between the peak and aggregate demand models despite the fact that Mitchell[24] criticized their earlier study[25] for not having the price elasticities for peak load and aggregate demand interrelated.

Uri's work[26] is somewhat different from the studies cited above. His models are based upon a mixture of stochastic time series and econometric methods of estimation. However, beyond the addition of a stochastic process to a direct method model, which introduces new estimation as well as other problems, this work seems to offer little else that is new.

3.3 Load Curve Methods

By far the bulk of the literature having to do with the estimation of electricity load is in this classification. The studies fall into two subcategories: hourly load forecasting models and models designed to analyze the data from time-of-day (TOD) rate experiments. The latter models, while important in that they help to provide a substantive empirical base for developing a more complete theory of electricity demand, are not very useful outside the environment in which they were developed because most of the explanatory variables measured in the TOD rate experiments are not available in nonexperimental time series records.[27]

For hourly load forceasting models, the data problems cited above do not exist since these efforts have been directed toward the forecast environment and use in many cases pure stochastic time series models or a mixture of the stochastic time series and econometric models. However, a lack of explanatory variables in pure time series models[28] and parameter instability in mixed econometric time series models[29] make these studies not very helpful for addressing policy questions related to peak load, an important requirement for the purposes of this report.

In addition to the forecasting models mentioned above, other single-equation regression models with special estimation techniques have been applied by Einhorn[30] and Platt[31] to estimate load curves. Both efforts are, however, overly ambitious for the data used; and, as a consequence, are not too helpful for the present study.

3.4 Related Literature: Load Duration Curves

Utility companies have historically used load duration curves because they provide powerful and convenient presentations of the variation in electrical loads. As such they are useful for the planning of future optimal capacity mix and for optimal load dispatch. Until very recently, however, estimates of load duration curves were only very crudely done. For planning and dispatch models in fact a customary industry practice was and still is to assume a specific shape of load duration curve to hold for all periods included in the optimization process.

For the work presented in this paper, load duration curves are important because they represent the variation in continuous-time load that can occur each year, because they are widely used in the electric power industry, and because they translate very easily into a load distribution. Sant[32] and Trimble[33] describe the mathematical relationship between load duration curves and load distributions.

The first effort to improve upon the crude method mentioned above was done by Loney[34] in terms of a dynamic optimization model. That effort turned out to be cumbersome and very costly to compute. Subsequently, several improvements to Loney's approach were proffered in the literature by Uri and Maybee,[35] Maybee,[36] and Maybee, Randolph, and Uri.[37]

Synthesized, these latter five articles proffer two types of approximations to a given load duration curve and two methods of estimating parameters for the load duration curve approximations. Maybee, Randolph, and Uri[38] proposes an optimal step-function approximation to a given load duration curve. Uri and Maybee[39] proposes an econometric method of forecasting the heights of this step-function over time in terms of economic and weather-related variables. Uri and Maybee[40] proposes a four-parameter linear exponential (in continuous-time magnitude) smooth approximation to a given load duration curve with econometric estimation over time of the parameters of the approximation; and Uri and Maybee[41] proposes a stochastic time series method of approximating the parameters of the linear exponential function.

For all the models for approximating load duration curves a two-stage estimation procedure is applied. First an approximation of the load duration curve is carried out then the parameters of the approximation are estimated. Thus, these methods will require, at a minimum, that the heights of a given step-function approximation or the estimates of the parameters of the linear exponential approximation be computed for every accounting period in a given sample.

In addition, even though these methods provide reasonably good approximations and forecasts for a load duration curve, they are only very loosely linked to previous empirical work on electricity demand. Consequently, in their present form, the usefulness of these methods is quite limited.

3.5 Lessons from the Literature

Being deficient in certain respects, the existing literature offers lessons for constructing models of peak load and the load distribution when estimation must be carried out with primarily temporally aggregated data. First, none of the previous models attempts to integrate peak and minimum loads, the load duration curve and aggregate demand into a unified model. Second, all of these models are incapable of treating a rate structure that redistributes load any differently from one that does not with the inevitable result that rates of the former kind cannot be adequately evaluated. Third, the existing literature is comprised of models of essentially two types: those that integrate well with previous work in electricity consumption but do not enlighten in the area of load redistribution and those that do the latter but are not integrated with previous work in electricity consumption. In addition with the complexities of a unified model comes an increasing complexity in estimation thus limiting what may be proposed as a unified model.

In the next section we present a unified model which adds only system hourly load data to improve upon the deficiencies of previous models. This model can evaluate the redistributive effects of a new rate structure but its usefulness has limitations that most likely can only be solved with more detailed disaggregate data. One example where this is true is described by Mitchell[42] as "a pressing task for new research..." because "...there is so far no forecasting model that is able to analyze the effect of peak load rate structures that have time-differential changes for both energy (kWh) and power (kW)." When applied differentially to individuals such rates can only be properly evaluated in a model estimated with individual data.

4. A MODEL OF PEAK AND MINIMUM LOADS AND THE LD CONDITIONED ON AGGREGATE DEMAND

Thus far we have implicitly assumed that load duration curves must be inherently interesting simply because utility companies have found them to be useful. The load duration curve — or equivalently the load distribution — is, however, a description of the extent of variation in electricity load in an accounting period as determined by customer behavior. As such it is interesting in the sense that behavior and hence the load distribution can be modified with appropriate incentives.

4.1 The Load Distribution Model

In the traditional view of electricity as a quantity demanded (say) per year, prices can be classified as financial incentives that alter demand or those that alter load (i.e., the load distribution). Alternately, one might redefine demand in temporal units small enough to obviate any need to distinguish between 'demand' and 'load.' The latter would seem to be the preferable approach because it maintains quantity demanded as the fundamental quantity responding to the financial incentive. As noted earlier, however, data limitations rarely permit one to estimate demand in such small temporal units.

If demand must be estimated in (say) annual units leaving load variation unaccounted for, an important dimension of consumption is lost. Moreover, any capability to evaluate the impact of load shifting incentives is also lost. The load distribution captures some of the lost dimension. It measures, in addition, load variation in the same time units that demand is measured.

One way to capture the load distribution in an annual model of electricity demand is to specify additional econometric models as suggested in the last section. Such models follow quite naturally if the load factor and moments of normalized load enter the utility functions of individual households or the decision-making agents in firms as arguments along with aggregate electricity, E; that is they follow if agents are viewed as maximizing a utility function

$u(E, LF, \mu_1, \mu_2, x)$

subject to the budget constraint $\rho_1 E + \rho_2 LF + \rho_3 \mu_1 + \rho_4 \mu_2 + \rho_5 x = M$ where x designates all other goods and M total expenditure; the ρ_i $(i = 1, \ldots, 5)$ each designate the price of the variable to which they are attached in the budget constraint. As expressions of a previously unaccounted for dimension of demand, the moments of normalized load and the load factor would logically be viewed as demand functions and as such would be determined by the same exogenous variables as demand. Thus, price, income, electricity-using equipment holdings, weather and socioeconomic variables would determine the load distribution just as they determine demand. The primary difference will be that moments of normalized load and the load factor will differ from demand and one another in the degree to which a given exogenous variable influences them. But, as we shall argue in Sect. 5, aggregate electricity, load factor and moment of normalized load demands are technologically constrained. The structure of electricity prices and income, historically, have had little influence on this technology in such a way that these demands would be affected by them to any significant degree. Consequently, historical observations of average price and income will not show much influence on load factor and moment of normalized load demands.

To envision how adjustment works in the LD model, consider the impact of a newly imposed time-of-day electricity rate for residences that ultimately leaves annual consumption at the same level at a lower cost so that the primary effect of the rate is to redistribute load. Usage adjustment is of course essentially fixed in the short run by household holdings of electricity-using and electricity-use-control appliances. The latter holdings are not likely to exist at the outset and the former will cost more to operate.

As households seek ways to reduce operating costs of electricity-using appliances perhaps in some areas by burning wood stoves during peak periods, new appliances begin to appear on the market. These appliances reduce the cost of operating electricity-using appliances by better providing households with the capability to choose when they consume electricity. Thus in the longer term heat and cold storage units with time-of-day controls and time-of-day controls for appliances and lighting become part of household appliance holdings. Consumption shifts to off-peak periods from peak periods.

Ultimately, as consumption shifts from peak to off-peak periods, peak load decreases while average load remains constant causing the load factor to rise. In addition, assume that minimum load remains relatively stable causing the difference between peak and minimum load to decline. As a result the first moment of normalized load rises too. But even as load is now spread more evenly over the accounting period, the variance of the load distribution may decrease if the decline in peak load dominates; otherwise it will increase.

Thus the partial effect of a time-of-day rate comprised of peak and offpeak rates would be (1) to increase the load factor and first moment of normalized load and decrease the second moment of normalized load for an increase in the peak price and/or decrease in the off-peak price; and (2) to cause just the opposite effect for opposite movement in these prices.

Let us turn now to consideration of specific models of the load distribution.

The previous section outlined how a load duration curve could be related to aggregate demand and how it could be estimated for each year via the method of moments. Peak and minimum loads it was explained could be estimated by expressing the moments of the load distribution as econometric models with annually measured explanatory variables. This section specifies a model that can estimate annual peak and minimum loads and the load duration curve via this procedure.

First annual moments of normalized load and an econometric equation for peak load[43] (the load factor in this model) are specified and estimated. Then the normalized load distribution parameters, peak and minimum loads are estimated: take predicted values of the moments of normalized load and the load factor and solve simultaneously for the normalized load distributions parameters, peak and minimum loads conditioned upon (i.e., given) an estimate of aggregate demand.[44]

In order to implement the model outlined here and in the previous section, we must first specify a parametric family of probability distributions to represent the load distribution, F(L), of equation (3). In addition, we must choose a family that can be estimated using the method outlined above. Specifically, this means we must choose a distribution for which the central moments are easily calculated and which is defined on a bounded domain so that peak and minimum demand can be expressed as parameters. The beta distribution[45] satisfies these requirements and in addition can assume a wide variety of shapes. We shall refer to load distributions using the beta as version I of the load distribution (LD) model. Another distribution which also satisfies the above requirements is a mixture of two beta distributions. We shall refer to this version of the LD model as version II.

4.2 The LD Model: Version I

This subsection presents a complete specification of verison I of the LD model. In addition in section four, estimates of peak load and the load duration curve under alternative assumptions are also presented, for this version.

In version I of the LD model, we specify the following beta family of probability distributions to represent the load distribution:

$$f_t(L_t) = \frac{(L_t - L_{*t})^{a_t - 1} (L_t^* - L_t)^{b_t - 1}}{B(a_t, b_t)(L_t^* - L_{*t})^{a_t + b_t - 1}}, \quad L_{*t} \leq L_t \leq L_t^* \qquad (9)$$

$$B(a_t, b_t) = \Gamma(a_t)\Gamma(b_t)/\Gamma(a_t + b_t)$$

$$\Gamma(\gamma) = \int_0^\infty u^{\gamma-1} e^{-u} du, \quad \gamma > 0 .$$

where $a_t > 0$, $b_t > 0$. Note that a t-subscript has been added to emphasize that the entire load distribution including peak and minimum load changes every year. We shall, however, omit the t-subscript elsewhere in the report in order to simplify the notation. Exceptions will be made where its omission will be confusing.

The normalized version of version I of the LD model of equations (9) is found by transforming according to equation (6a):

$$f(Z) = \frac{Z^{a-1}(1 - Z)^{b-1}}{B(a, b)}, \quad 0 \leq Z \leq 1 \tag{10}$$

$a > 0$, $b > 0$. It is the standardized beta family and does not include L^* and L_*. The moments of normalized load $\mu_r (r=1,2)$ for this version I normalized LD are:

$$\mu_1 = a/(a + b) \tag{11a}$$

$$\mu_2 = ab/(a + b)^2(a + b + 1) \tag{11b}$$

Recall that values of μ_1 and μ_2 in the LD model are predicted each year in terms of a set of exogenous variables. Let X_1 and X_2 represent, respectivley, K×1 and L×1 vectors of these exogenous variables and let β_1 and β_2 represent K×1 and L×1 vectors of parameters. We specify the econometric models for these normalized moments as linear models, i.e.,

$$\mu_1 = X'\beta_1 + \varepsilon_1 \tag{12a}$$

$$\mu_2 = X_2'\beta_2 + \varepsilon_2 \; ; \tag{12b}$$

ε_1 and ε_2 are normally distributed residuals with zero means and constant variance. Both ε_1 and ε_2 can be correlated perhaps by being treated as simultaneous equations to account for their interrelationship.[46]

The relations of equations (12) predict μ_1 and u_2 (call these predicted values $\hat{\mu}_1$ and $\hat{\mu}_2$) for each year. Substituting these predicted values $\hat{\mu}_1$ and $\hat{\mu}_2$ into equations (11) and solving for a and b, predictions of a and b are obtained for each year. This yields a prediction of the normalized load distribution of equation (10) for each year. Estimates of peak and minimum loads are obtained by specifying an additional econometric model[47] for the load factor — defined as the ratio of average hourly load to peak load per annum — similar to thos for the moments of normalized load in equations (12). An econometric model of peak load might alternately be specified; however, the load factor model is more compatible with estimation by moments.

Let X_3 be an M×1 vector of exogenous variables that determine the year-to-year variation in the load factor, LF; let β_3 be an M×1 vector of parameters and ε_3 a normally distributed residual with mean zero and constant variance, ie.,

$$LF = X_3'\beta_3 + \varepsilon_3 \ . \tag{13}$$

Also let $\widehat{LF}$ designate predicted values of LF.

Summarizing, the model of moments of normalized load and the load factor is

$$\begin{aligned} \mu_1 &= X_1'\beta_1 + \varepsilon_1 \\ \mu_2 &= X_2'\beta_2 + \varepsilon_2 \\ LF &= X_3'\beta_3 + \varepsilon_3 \end{aligned} \tag{14}$$

where $\underline{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ has zero mean vector and in general covariance matrix

$$\sum = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ & & \sigma_3^2 \end{bmatrix} .$$

As noted earlier the normalized load distribution (or equivalently the normalized load duration curve) is estimated by substituting the predicted values $\hat{\mu}_1$ and $\hat{\mu}_2$ into the following equations:

$$a = \mu_1 \mu_2^{-1}[\mu_1(1-\mu_1) - \mu_2]$$
$$b = (1-\mu_1)\mu_2^{-1}[\mu_1(1-\mu_1) - \mu_2] \qquad (15)$$

Estimation of peak and minimum loads from this model requires additional explanation.

Note that μ_1, μ_2, and LF, each designating moments of some sort, are defined as:

$$\mu_1 = E(Z) = E\left(\frac{L - L_*}{L^* - L_*}\right)$$
$$\mu_2 = Var(Z) = E\left(\frac{L - L_*}{L^* - L_*}\right)^2 - \left\{E\left(\frac{L - L_*}{L^* - L_*}\right)\right\}^2 \qquad (16)$$
$$LF = E(L/L^*)$$

where $E(\cdot)$ designates the expectation operator and equation (6a) has been used. Both L_* and L^* are parameters in the load distribution hence they are treated as constants inside the expectation operator. Thus equations (16) may be re-written as:

$$\mu_1 = (L^* - L_*)^{-1}[E(L) - L_*]$$
$$\mu_2 = (L^* - L_*)^{-2}[E(L^2) + L_*^2 - 2L_*E(L)] \qquad (17)$$
$$LF = E(L)/L^* .$$

Solving now equations (17) for L_*, L^*, and $E(L^2)$ we have

$$L_* = (\widehat{LF}-\hat{\mu}_1)\widehat{E(L)}/(1-\hat{\mu}_1)\widehat{LF} \qquad (18a)$$

$$L^* = \widehat{E(L)}/\widehat{LF} \qquad (18b)$$

$$E(L^2) = (1-\hat{\mu}_1)^{-2}\widehat{LF}^{-2}\widehat{E(L)}^2 \cdot [(1-\widehat{LF})^2 - (\widehat{LF}-\hat{\mu}_1)^2 + 2(\widehat{LF}-\hat{\mu}_1)(1-\hat{\mu}_1)\widehat{LF}] \; . \tag{18c}$$

where "hats" have been placed over each variable to indicate "predicted values." Equations (18a) and (18b) provide the equations for calculating predicted values of peak and minimum loads. Recall that E(L) is given by equation (5) and the estimate of aggregate demand from the SLED and service area disaggregation models. It is in this sense that the LD model is conditioned on aggregate demand. All other variables appearing on the right hand side of equations (18) are provided from equations (14) as predicted values. Although an estimate of $E(L^2)$ is provided by this procedure, it will not be used in this report.

Load duration curve, peak and minimum load estimates using version I of the LD model are presented in the next section. Referring ahead though to Fig. 8, it can be seen that the "shape" of the load distribution is not captured as completely as is possible. In the next subsection we present an extension of the version I of the LD model which we call version II which can better account for the "shape" of the load duration curve. This additional capability does not, however, come without additional computational complexity.

The chief advantage of version I is that it provides the required estimates in a computationally simple framework. Obtaining estimates from version II, on the other hand, requires a numerical solution. And a solution that falls within a priori bounds is not guaranteed unless some complicating and somewhat arbitrary modifications are made to the method-of-moments estimation procedure. Version I, on the other hand always yields solutions that fall within the appropriate bounds.

4.3 The LD Model: Version II

A natural extension of version I of the LD model to better account for the "shape" of the load distribution is made by estimating additional moments of normalized load. In this section we specify a more complex family of distributions to represent the load distribution. This family of distributions has more parameters than the family of equation (9) and hence requires additional moments.

The family of distributions we specify is a mixture of two of the beta densities of equations (9). By specifying a mixture of two beta densities, we preserve the required properties of the family: (1) simply calculated moments and (2) peak and minimum load parameters that define the maximum and minimum values which instantaneous load can take. The family comprised of mixing two beta families is:

$$g(L) = \phi \frac{(L - L_*)^{a_1-1}(L^* - L)^{b_1-1}}{B(a_1, b_1)(L^* - L_*)^{a_1+b_1-1}}$$

$$+ (1-\phi) \frac{(L - L_*)^{a_2-1}(L^* - L)^{b_2-1}}{B(a_2, b_2)(L^* - L_*)^{a_2+b_2-1}}, \quad L_* \leq L \leq L^* \tag{19}$$

where $0 < \phi \leq 1$, $a_i > 0$ and $b_i > 0$ $(i = 1, 2)$ comprise the five parameters that will remain after transforming the load distribution (19) to the normalized load distribution using equation (6a). Recall that t-subscripts to signify that the load distribution is different for each year have been suppressed on all parameters as well as L.

The moments of normalized load for the version II LD of equation (19) are:

$$\mu_r = \sum_{j=0}^{r} \binom{r}{j} \alpha_{r-j}(-\alpha_1)^j, \quad r = 1, \ldots, 5 \tag{20a}$$

$$\alpha_{r-j} = \phi(a_1)_{r-j}(a_1+b_1)^{-1}_{r-j} + (1-\phi)(a_2)_{r-j}(a_2+b_2)^{-1}_{r-j} \tag{20b}$$

$$r-j = 1, \ldots, 5$$

$$(\gamma)_k = \gamma(\gamma+1) \ldots (\gamma+k-1) \quad k \text{ a positive, finite integer.} \tag{20c}$$

Predicted values of the moments of normalized load and the load factor are obtained as before from a set of econometric models of the following form:

$$\mu_r = X_r'\beta_r + \varepsilon_r \quad r = 1, \ldots, 5 \tag{21}$$

$$LF = X_{r+1}\beta_{r+1} + \varepsilon_{r+1} .$$

Substituting predicted values $\hat{\mu}_r$ for the μ_r in equation (20a), one can then solve numerically for ϕ, a_i, and b_i $(i = 1, 2)$ in each year. The procedure for solving for these parameters is a bit tedious; we thus present it in the Appendix. Obtaining estimates of peak, L^*, and minimum, L_*, loads is done as outlined for the version I LD model.

Since numerical difficulties frequently arose with estimating the load distribution in version II of the LD model, we present in the next section only a few selected results for it. The potential gain in estimating the load duration curve is pictured graphically. And moments of normalized load are estimated using several different estimation techniques.

5. PEAK LOADS AND LOAD DISTRIBUTIONS FOR SEVERAL ELECTRIC UTILITIES

This section presents estimates of peak loads and in selected cases normalized load distributions with minimum loads as a part of them for various electric utilities in the continental United States. Peak loads are estimated for twenty geographically dispersed utilities; the impact of minimum load on normalized load distribution estimates in conjunction with normalized load distribution estimates are presented for selected utilities in selected years.

Assuming rational consumers of electricity, we might expect the real prices of electricity and substitute fuels to be important determinants of the load factor and moments of normalized load. If so, it would mean that these prices influence the way consumers distribute their loads over the year. But there are cogent reasons to believe that historical prices have had a minimal impact on load distributions. First, being for the most part undifferentiated as to when consumption occurs, historical electricity prices have offered little incentive to affect the way consumers distribute their loads leaving that instead to the influence of other factors. Second, consumption in the short run is essentially fixed by the stock of electric equipment in place. The impact on load distributions is then fixed by the technical characteristics and operating patterns of this equipment. Operating costs for the equipment, being a primary determinant of operating patterns, are largely determined by electricity price and hence have little impact on load distributions. Thus changes in the composition of electric equipment in place has historically defined pretty much the limits to the shape of load distributions. Third, real prices of substitute fuels can influence the composition of electric equipment in place if their levels relative to electricity sufficiently lower operating costs so that consumers switch to other-fueled equipment or are sufficiently higher so that consumers switch from other-fueled to electric equipment. However, the historical differentials in the relative prices of other fuels has not produced switching of this sort on any grand scale.

This does not mean that electricity and substitute-fuel prices do not affect load distributions. Indeed, if structured to do so, they will and the impact can be substantial. But historical electricity price levels have not been set according to the timing of consumption and substitute fuel price levels have not induced switching to or from electric equipment. Consequently, while there is good reason to expect properly structured prices to affect load distributions in the future, there is correspondingly little reason to expect historical prices to have done so in the past.

As a result, holdings of electric equipment, their usage levels as affected by weather and socioeconomic factors and changes in the composition of electric equipment holdings have been the historical determinants of load distributions. Moreover, there still tends to be differences in the impact of these determinants on the load distribution in each of the three consuming sectors. Residential and commercial loads tend to be weather sensitive while industrial loads tend to vary more with the technical requirements of production. Industrial loads therefore tend to be more stable. And when changes do occur, they tend to be "once and for all" changes arising from plant relocations, changes in production processes or commitments and the like. Residential and commercial consumption, on the other hand, tends to be set in large part by holdings of electric equipment. Part of those holdings, tending to be for uses like refrigeration, home lighting, cooking, home food freezing, clothes drying, and commercial lighting, are not very weather sensitive. But uses for space heating, air-conditioning, and water heating tend to be very weather sensitive and to account for a substantial share of both residential and commercial electricity consumption and load variation.

Space heating and air-conditioning equipment holdings have been calculated for the residential but not the commercial sector. However, since residential consumption accounts for most of the consumption in the two sectors, electric equipment in this sector will influence load distributions more. In this section we specify single equation models for load factors and moments of normalized load with residential space heating and air-conditioning stocks, weather and a measure of jumps in industrial sales as the exoge us variables.

Four subsections compr e the body of this section. The first describes the data compiled. The second estimates peak loads and in selected cases normali   oad distributions using state-level measurements for exogenous variables under two sets of simplifying restrictions of version I of the LD model as well as version I without simplifying restrictions. And the third estimates peak and minimum loads and normalized load distributions for two utilities using state- and service-area-level measurements of exogenous variables and version II of the LD model.

5.1 Data

The data used for estimation in this report consists of hourly loads obtained directly from the utilities studied, state-level estimates[48] of annual electric space heating and central air-conditioning saturations, state-level and service-area-level measurements of annual heating and cooling degree days[49] and annual measurements of the proportion of "all but industrial" to "total" electricity sales by utility for each utility system studied.[50] F each utility, hourly system loads were recorded for as many years a were available in machine readable form. Inevitably, for each year of data some hourly loads were missing but the proportion of missing values has been relatively small (i.e., less than three percent) for each year of data for each utility. The years for which hourly loads were compiled for a given utility is indicated in Table 1 under the column headed "years observed."

Utilities were chosen for analysis based upon their geographic location with the intent that most geographic regions of the U.S. would be covered. Since the purpose of this work was not to construct a model for an "average" utility in the nation, no attempt was made to design a sample with some sort of geographic strata. Rather, the purpose in selecting geographically dispersed utilities was to check for geographic peculiarities that might invalidate the model in some way.

Table 1. Single-Equation Estimates: Combined Model of Load Factors*
(standard errors in parentheses)

Utility	Dependent variable	CDD	LNAC	CDD x LNAC	HDD	LNHEL	HDD x LNAC	dw	R^2	Years observed	Largest maximum absolute difference	Smallest maximum absolute difference
American Electric Power	N/A									1970-74		
Carolina Power and Light	LF	.1119 (.2566)	-100.3 (152.6)	.0957 (.0885)	.1590 (.1228)	-205.9 (153.1)	.0586 (.0421)	1.225	.9988	1962-75	.1623 1962	.0521 1966
Central Hudson	LF	.1326 (.1930)	-1.457 (44.97)	.0372 (.047)	.1075 (.0220)	-130.359 (49.74)	.0221 .006	2.326	.9998	1960-74	.1016 1963	.0496 1964
Central Illinois Public Service	LF	-.4712 (.5245)	230.6 (220.2)	.0670 (.1140)	.0752 (.0961)	-49.62 (238.4)	-.0282 (.0212)	1.196	.9996	1965-72	.2958 1967	.0545 1973
Commonwealth Edison of Chicago	LF	-12.99 (19.35)	2629.0 (2570.0)	2.418 (3.629)	.2454 (1.189)	-2293.0 (2373.0)	-.1409 (.3550)	2.714	.9965	1968-74	.2662 1971	.0715 1973
Florida Power	LF	.2243 (.0936)	-32.82 (324.8)	.0465 (.0751)	.0789 (.2396)	-161.9 (281.0)	.0408 (.1890)	3.349	.9985	1961-74	.1609 1962	.0593 1971
Iowa Public Service	LF	.0808 (.1949)	-25.23 (95.71)	.0524 (.0669)	.0454 (.0244)	-113.7 (71.83)	.0098 (.0071)	1.874	.9995	1962-74	.1412 1965	.0596 1970
Jacksonville Electric Authority	N/A									1969-74		
Jersey Central Power and Light	LF	-1.503 (2.302)	-295.5 (195.0)	-.7030 (1.022)	.4608 (.4795)	240.6 (243.1)	.1169 (.1448)	1.878	.9998	1967-74	.0932 1967	.0555 1974
Metropolitan Edison	LF	3.257 (1.676)	212.1 (160.5)	.2746 (.1419)	.3163 (.2106)	-770.7 (517.4)	.0740 (.0634)	1.805	.9999	1967-74	.1289 1967	.0632 1974

Table 1. (cont.)

Utility	Dependent variable	CDD	LNAC	CDD x LNAC	HDD	LNHEL	HDD x LNAC	dw	R^2	Years observed	Largest maximum absolute difference	Smallest maximum absolute difference
New Jersey Public Service Electric and Gas	LF	-.0190 (.1502)	-37.19 (74.10)	-.0223 (.0563)	.0837 (.0216)	-65.09 (48.20)	.0140 (.0091)	2.684	.9999	1963-74	.1294 1963	.0724 1972
Niagara Mohawk	LF	.8443 (.2213)	157.6 (89.85)	.2519 (.0630)	.0873 (.0228)	-296.1 (77.66)	.0202 (.0062)	3.145	.9999	1962-74	.1143 1963	.0627 1967
Northeast Utilities	LF	2.464 (.9411)	-278.4 (125.1)	.6566 (.2605)	-.0651 (.07[illegible])	205.9 (172.2)	-.0320 (.0271)	2.499	.9999	1966-74	.0838 1968	.0555 1971
Pennsylvania Electric	LF	2.574 (1.049)	223.97 (100.4)	.2180 (.0888)	.3697 (.1318)	-874.3 (323.7)	.0998 (.0397)	1.653	.9999	1967-74	.0939 1967	.0481 1974
Pennsylvania Power and Light	LF	.6443 (1.836)	-89.71	.0612	-.0425	58.62	-.0046	1.790	.9999	1965-74	.0862 1965	.0479 1969
Power Authority of the State of New York	LF	2.320 (1.073)	279.2 (435.5)	.6389 (.3054)	-.0235 (.1104)	-433.0 (376.4)	-.0009 (.0300)	2.450	.9982	1962-74	.1126 1964	.2924 1963
San Diego Gas and Electric	LF	.2457 (.4657)	-192.9 (120.5)	.1184 (.1586)	.2069 (.1220)	-62.68 (134.4)	.0890 (.0460)	2.477	.9992	1961-74	.1614 1972	.0669 1961
Southern California Edison	LF	.0841 (.5219)	38.55 (121.9)	.0367 (.1947)	.1996 (.1276)	-352.0 (129.3)	.0996 (.0498)	2.099	.9998	1964-74	.0978 1973	.0566 1966
Wisconsin Electric Power Company	LF	.1634 (.1932)	124.0 (40.13)	.0060 (.0560)	.0687 (.0131)	-207.9 (28.89)	.0149 (.0034)	2.241	.9999	1961-74	.1251 1972	.0593 1965
Wisconsin Michigan Power Company	LF	.7320 .2351	11.54 (48.83)	.1817 (.0680)	.0494 (.0159)	-162.0 (35.[illegible]6)	.0126 (.0041)	2.260	.9999	1961-74	.1109 1970	.0550 1963

5.2 The LD Model: Version I

The emphasis of previous work has been on peak loads and to a lesser degree on the load duration curve. Minimum loads have received very little attention. To some extent this is justified because peak load gives a clear signal of how much capacity is required and the load duration curve tells how to distribute capacity among the broad categories of equipment. Minimum load does not even indicate how much base load generation capacity is needed because the optimal allocation among equipment types dictates that base load equipment serve more than minimum load. Minimum load, however, as we shall see, is very important in positioning the load distribution and we might expect that it would be highly correlated with the optimal allocation to base load equipment. Nevertheless, we shall follow the lead of previous work and emphasize peak load leaving our discussion of minimum load to showing its impact in positioning the load distribution (or equivalently the load duration curve).

As noted earlier in the opening commentary of this section, we specify in this subsection models of load factors and moments of normalized load as functions of state-level aggregates. This eases the burden of compiling data from a large number of utilities but might at first blush seem to bias the individual effects of the exogenous variables. Actually, this is not the case; however, a state-level aggregate of an exogenous variable may not be as good a predictor of service-area moments as would be a service-area aggregate of the same variable. For example, consider the influence of temperature measured in degree days on the load factor of a northern California utility. Because the warmer climate of southern California so influences, the number of heating degree days for the state will be larger than for the utility's service area. And likewise the number of cooling degree days for the state will be less than for the service area. But this in itself does not make service-area rather than state degree days better predictors of the utility's load factor. What matters is which is more closely correlated with the utility's load factor. Intuitively, one might expect service-area degree days to be more closely correlated.

Also it is just as intuitively reasonable to expect service-area degree days to be closely correlated with state degree days making, therefore, state degree days nearly as good a predictor as service-area degree days. Consequently, for state-level exogenous variables closely correlated with corresponding service-area aggregates, useful information about the importance of the service-area aggregates as determinants of the load factor or other moments can be gleaned from models estimating these moments as functions of the state-level aggregates.

Let us consider now the impact on the estimated load distribution of imposing some restrictions that simplify version I. The restrictions we consider will have no effect on the calculation of peak load estimates.

5.2.1 Version I with two simplifying restrictions

Referring again to the matter of emphasizing peak load and the load distribution in version I, we may do this quite easily by removing minimum load. This is done by setting L_* equal to zero in the equations of (9)-(14). Peak load calculation as we just noted is unaffected by this restriction but this is not true for the positioning of the load distribution. The other simplifying restriction fixes the parameter "b" in equation (9) at a constant value which makes estimation the determination of the estimators' sampling properties easier. We consider first then the joint impact of these two restrictions and consider later the impact of setting L_* at zero.

Table 1 lists ordinary least squares estimates of a load factor model using state-level aggregates of the exogenous variables. The exogenous variables are: cooling degree days for the state (CDD), heating degree days for the state (HDD), the natural logarithm of the state's proportion of centrally air-conditioned homes (LNAC), and the natural logarithm of the state's proportion of electrically heated homes (LNHEL). These latter two variables are natural logarithms of what are known as the saturation levels of centrally air-conditioned and electrically heated homes for the state. Because all heating and cooling exogenous variables are combined in this model, we refer to it as the combined model of load factors. Table 1 also presents the Durbin Watson test statistic (dW), the years for which data was available and for each of these spans of years the largest and smallest maximum absolute difference between the actual (empirical) and estimated load distributions. These latter two statistics measure how well this particular version of version I of the LD model estimates the load distribution; they are the test statistics for the Kolnogorov-Smirnov test of goodness of fit; the largest and smallest of these indicate how well and how poorly the load distributions were estimated over the span of years indicated. These test statistics indicate that all of the estimated load distributions calculated under the simplification for version I of fixing one parameter are not significantly different from the empirical (actual) load distribution at the one-percent level. Despite the apparently good fit we will see shortly that the one-parameter version of the LD model is likely to encounter difficulties outside the range of the sample.

Clearly the coefficients of determination (R^2) indicate that weather and space conditioning measurements account very well for the year-to-year variation in load factors. The Durbin Watson tests do not indicate the presence of serial correlation. A problem though with this particular model is the presence of linear relationships amongst some of the exogenous variables. This is suggested by the poor performance of individual coefficient standard errors — shown in parentheses - in conjunction with a high R^2 value. As it turns out heating degree days are highly correlated over time with cooling degree days for every region causing a multicollinearity problem. The problem is similar for central air-conditioning and heating measurements. Consequently, in later load factor and moments of normalized load models we keep cooling measurements separate from heating measurements.

Notwithstanding that the models in Table 1 have collinear exogenous variables, it is still useful to examine how well this load factor model predicts peak load in the sample period. This is true because, on the one hand, how well a model "fits" the data is unaffected by multicollinearity in the exogenous variables and, on the other the state-level exogenous variables are highly correlated with their service-area counterparts. Accordingly, Tables 2, 3, and 4 present information that

Table 2. Ratio of actual to estimated peak load for selected utility systems[a]

Year	Iowa Public Service	Jersey Central Power & Light	Metropolitan Edison	Niagara Mohawk
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1962	0.98	N/A[b]	N/A[b]	1.00
1963	0.98	N/A[b]	N/A[b]	1.01
1964	1.00	N/A[b]	N/A[b]	1.06
1965	1.04	N/A[b]	N/A[b]	1.00
1966	1.03	N/A[b]	N/A[b]	1.00
1967	0.96	0.99	1.00	0.98
1968	1.03	1.01	1.01	1.01
1969	1.01	1.01	1.01	0.98
1970	0.99	0.98	0.99	1.01
1971	0.98	1.00	0.99	0.99
1972	0.98	1.02	1.01	1.02
1973	0.99	1.01	1.01	1.00
1974	1.02	0.98	1.00	0.99

Table 2. (cont.)

Year	Northeast Utilities	Pennsylvania Electric	Pennsylvania Power & Light	PASNY	Public Service Electric & Gas of New Jersey
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1962	N/A[b]	N/A[b]	N/A[b]	0.99	N/A[b]
1963	N/A[b]	N/A[b]	N/A[b]	1.03	1.00
1964	N/A[b]	N/A[b]	N/A[b]	0.99	1.00
1965	N/A[b]	N/A[b]	1.00	0.95	0.99
1966	1.00	N/A[b]	1.00	1.02	1.00
1967	1.01	1.00	0.99	0.98	0.99
1968	0.99	1.00	1.01	1.12	1.02
1969	1.00	0.99	0.99	0.96	N/A[c]
1970	0.99	1.00	0.98	0.96	0.98
1971	1.01	1.01	1.01	0.95	1.01
1972	0.99	1.00	1.02	1.05	1.01
1973	0.99	1.00	1.01	1.03	1.00
1974	1.01	1.00	0.99	0.98	0.99

Table 2. (cont.)

Year	San Diego Gas & Electric Co.	So. California Edison	Wisconsin Electric Power Co.	Wisconsin Michigan
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	0.99	N/A[b]	N/A[b]	N/A[b]
1962	1.01	N/A[b]	1.00	1.01
1963	1.00	N/A[b]	0.99	1.00
1964	1.03	0.98	1.01	0.98
1965	0.97	1.02	1.06	1.00
1966	1.04	1.00	1.02	1.01
1967	1.00	1.02	0.99	1.01
1968	0.96	1.02	0.99	1.00
1969	1.03	0.98	0.99	1.00
1970	1.01	0.98	1.01	0.98
1971	0.98	1.01	0.99	1.01
1972	0.95	1.00	1.00	0.99
1973	1.03	0.99	1.00	1.01
1974	1.02	1.02	1.01	1.00

[a]Estimates of peak load computed using estimated load factors and actual aggregate demand calculated from system load data.

[b]Data not available for this year.

[c]Incomplete data for this year.

Table 3. Actual and predicted peak loads[a]

Year	Carolina Gas & Electric Power		Central Illinois Public Service		Commonwealth Edison of Chicago	
	Peak	$\hat{\text{Peak}}$	Peak	$\hat{\text{Peak}}$	Peak	$\hat{\text{Peak}}$
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1962	1787	1623.66	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1963	1638	1651.49	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1964	1749	1784.28	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1965	1931	2062.94	771	788.37	N/A[b]	N/A[b]
1966	2184	2212.00	894	891.38	N/A[b]	N/A[b]
1967	2270	2247.61	868	870.78	N/A[b]	N/A[b]
1968	2834	2817.87	1061	1028.50	8950	8954.71
1969	3055	3109.90	1126	1098.67	9265	9408.61
1970	3484	3525.90	1210	1189.82	10027	9003.45
1971	3625	3749.18	1252	1287.76	8180	8915.89
1972	4119	4003.57	1394	1431.56	11750	12211.99
1973	4711	4596.84	1503	1505.17	12462	12322.35
1974	4771	4666.75	1500	1488.32	12270	11898.38

Table 3. (cont.)

	Central Hudson Gas & Electric Co.		Florida Power Co.		Iowa Public Service	
Year	Peak	$\widehat{\text{Peak}}$	Peak	$\widehat{\text{Peak}}$	Peak	$\widehat{\text{Peak}}$
1960	241	244.44	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	261	260.33	805	809.13	N/A[b]	N/A[b]
1962	277	280.12	960	890.37	225	229.49
1963	296	287.56	897	963.69	252	256.58
1964	323	324.67	1034	1006.06	266	265.62
1965	362	354.39	1003	1047.98	293	280.50
1966	393	403.21	1242	1219.22	317	308.59
1967	407	414.32	1254	1262.50	313	325.01
1968	454	442.99	1551	1549.10	376	366.39
1969	491	484.58	1710	1679.24	392	387.33
1970	521	519.79	1990	1953.39	422	423.83
1971	549	551.67	2152	2232.68	437	446.88
1972	587	585.82	2501	2459.79	486	491.[illegible]3
1973	632	632.48	2862	2991.69	5[illegible]	554.99
1974	584	592.08	2970	2844.01	595	584.20

Table 3. (cont.)

	Jersey Central Power & Light		Metropolitan Edison		Niagara Mohawk	
Year	Peak	$\hat{\text{Peak}}$	Peak	$\hat{\text{Peak}}$	Peak	$\hat{\text{Peak}}$
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1962	N/A[b]	N/A[b]	N/A[b]	N/A[b]	3201	3207.83
1963	N/A[b]	N/A[b]	N/A[b]	N/A[b]	3316	3273.08
1964	N/A[b]	N/A[b]	N/A[b]	N/A[b]	3498	3512.95
1965	N/A[b]	N/A[b]	N/A[b]	N/A[b]	3701	3700.02
1966	N/A[b]	N/A[b]	N/A[b]	N/A[b]	3987	3992.91
1967	1227	1234.91	912	915.11	3971	4031.88
1968	1455	1446.50	1021	1013.49	4335	4292.27
1969	1604	1592.97	1139	1131.63	4442	4512.93
1970	1716	1748.92	1188	1199.84	4614	4555.75
1971	1880	1881.29	1271	1286.23	4551	4598.84
1972	2122	2070.91	1361	1352.37	4827	4740.80
1973	2456	2436.77	1474	1466.04	4896	4881.36
1974	2396	2440.54	1378	1378.[illegible]3	4787	4821.76

Table 3. (cont.)

	Northeast Utilities		PASNY		Pennsylvania Electric Co.	
Year	Peak	$\widehat{Peak}$	Peak	$\widehat{Peak}$	Peak	$\widehat{Peak}$
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1962	N/A[b]	N/A[b]	601	605.39	N/A[b]	N/A[b]
1963	N/A[b]	N/A[b]	609	588.62	N/A[b]	N/A[b]
1964	N/A[b]	N/A[b]	602	506.20	N/A[b]	N/A[b]
1965	N/A[b]	N/A[b]	618	646.90	N/A[b]	N/A[b]
1966	2367	2376.68	734	722.73	N/A[b]	N/A[b]
1967	2414	2432.34	801	821.52	1201	1199.59
1968	2740	2706.17	858	765.21	1312	1312.77
1969	2918	2920.81	880	913.03	1462	1472.33
1970	3172	3154.02	870	905.23	1535	1533.80
1971	3223	3263.93	858	900.63	1645	1628.63
1972	3520	3468.10	990	942.13	1711	1709.24
1973	3645	3620.56	962	931.48	1790	1792.70
1974	3496	3546.16	1031	1053.86	1766	1772.78

Table 3. (cont.)

	Pennsylvania Power & Light		Public Service Electric & Gas of New Jersey		San Diego Gas & Electric Co.	
Year	Peak	$\widehat{\text{Peak}}$	Peak	$\widehat{\text{Peak}}$	Peak	$\widehat{\text{Peak}}$
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]	727	719.25
1962	N/A[b]	N/A[b]	N/A[b]	N/A[b]	745	752.65
1963	N/A[b]	N/A[b]	3370	3357.99	805	802.40
1964	N/A[b]	N/A[b]	3665	3663.40	835	857.13
1965	1853	1851.71	3953	3996.95	925	893.82
1966	2085	2079.97	4100	4090.96	950	984.96
1967	2202	2213.96	4308	4336.09	1091	1087.55
1968	2493	2471.90	4828	4749.80	1163	1115.86
1969	2702	2724.15	N/A[c]	N/A[c]	1218	1258.18
1970	2897	2954.32	5398	5492.67	1343	1355.26
1971	3157	3116.95	5925	5857.46	1470	1435.98
1972	3483	3420.87	6201	6119.05	1579	1497.64
1973	3598	3576.75	6816	6810.36	1518	1563.33
1974	3662	3715.84	6316	6397.26	1498	1530.26

Table 3. (cont.)

Year	So. California Edison Peak	So. California Edison $\hat{Peak}$	Wisconsin Electric Power Co. Peak	Wisconsin Electric Power Co. $\hat{Peak}$	Wisconsin Michigan Peak	Wisconsin Michigan $\hat{Peak}$
1960	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1961	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]	N/A[b]
1962	N/A[b]	N/A[b]	1218	1218.47	224	221.17
1963	N/A[b]	N/A[b]	1365	1347.70	259	257.87
1964	5335	5427.47	1413	1421.10	267	272.67
1965	5863	5742.93	1484	1491.00	279	279.37
1966	6173	6196.71	1603	1639.46	309	305.74
1967	7001	6889.07	1655	1646.04	303	300.56
1968	7425	7298.86	1901	1887.22	320	321.58
1969	7804	7965.27	1984	1968.25	329	330.55
1970	8274	8444.97	2100	2120.44	337	343.22
1971	9350	9297.70	2268	2231.65	358	353.85
1972	9815	9846.13	2423	2420.10	381	384.42
1973	10253	10349.58	2633	2623.01	426	421.23
1974	9997	9839.94	2573	2600.42	421	420.83

[a]Estimates of peak load computed using estimated load factors and actual aggregate demand calculated from system load data.

[b]Data not available for this year.

[c]Incomplete data for this year.

Table 4. Percentage Difference Between Actual and Predicted Peak Loads[a]

Utility		LF = f (cooling variables)	Year	LF = f (heating variables)	Year	LF = f (heating and cooling)	Year
American Electric Power	best	.096	1970	.102	1971	N/A	
	worst	1.89	1972	1.56	1972		
Carolina Power and Light	best	.128	1967	.762	1973	.570	1968
	worst	8.54	1965	15.8	1962	9.14	1962
Central Hudson	best	.033	1960	.121	1970	.0764	1973
	worst	5.35	1973	2.97	1966	2.851	1963
Central Illinois Public Service	best	.957	1970	1.32	1974	.320	1967
	worst	7.43	1967	6.89	1967	3.06	1968
Commonwealth Edison of Chicago	best	.097	1974	.290	1973	N/A	
	worst	3.545	1969	2.330	1969		
Florida Power	best	.058	1970	1.43	1966	.051	1961
	worst	9.98	1965	18.5	1963	7.43	1963
Iowa Public Service	best	.014	1964	.210	1972	.433	1970
	worst	5.71	1965	7.46	1967	4.27	1965
Jacksonville Electric Authority	best	.459	1969	.142	1971	N/A	
	worst	3.19	1971	2.44	1970		
Jersey Central Power and Light	best	.413	1967	.598	1974	.069	1971
	worst	4.08	1970	3.67	1967	1.92	1970
Metropolitan Edison	best	.097	1971	.011	1973	.017	1974
	worst	2.51	1969	2.70	1969	1.20	1971
New Jersey Public Service Electric and Gas	best	.326	1968	.127	1964	.044	1964
	worst	5.81	1974	2.61	1972	1.75	1970

Table 4. (cont.)

Utility		LF = f (cooling variables)	Year	LF = f (heating variables)	Year	LF = f (heating and cooling)	Year
Niagara Mohawk	best	.324	1971	.086	1962	.026	1965
	worst	3.53	1962	9.23	1966	1.79	1972
Northeast Utilities	best	.314	1970	.012	1972	.409	1966
	worst	2.42	1967	3.65	1967	1.47	1972
Pennsylvania Electric	best	.270	1972	.139	1970	.059	1968
	worst	1.79	1971	2.18	1971	.995	1971
Pennsylvania Power and Light	best	.085	.973	.028	1966	.070	1965
	worst	2.46	1972	2.20	1967	1.98	1970
Power Authority of the State of New York	best	.059	1969	.073	1962	.335	1963
	worst	13.7	1968	14.5	1968	10.8	1968
San Diego Gas and Electric	best	.723	1962	.281	1969	.316	1967
	worst	11.7	1973	6.94	1962	5.15	1972
Southern California Edison	best	1.54	1975	.125	1966	.3[illegible]7	1972
	worst	4.43	1970	2.93	1969	2.07	1969
Wisconsin Electric Power Company	best	.157	1972	.284	1964	.038	1962
	worst	7.41	1963	4.84	1967	2.27	1966
Wisconsin Michigan Power Company	best	.285	1973	.350	1966	.039	1974
	worst	4.55	1963	4.97	1973	2.12	1964

[a]Heating variables and cooling variables results taken from material presented in Table 5.

is useful for evaluating how well the combined load factor model predicts peak load <u>given actual aggregate demand</u>.

Table 2 presents ratios of actual peak to estimated peak (calculated from the models of Table 1) for a representative selection of the twenty utility systems studied. In this selection, the worst prediction was for the Power Authority of the State of New York (PASNY) in 1968 where peak load was underestimated by 12 percent. Generally though the predictions are reasonably good. Ninety percent fall within three percent of the actual value. Table 3 presents for comparison these actual and estimated peaks (also from Table 1) for all but two of the twenty utilities — American Electric Power (AEP) and Jacksonville Electric Authority (JEA). The system load data from AEP covered too few years to do much with it and the JEA data had discrepancies which we were unable to reconcile. The effect of separating heating variables (heating degree days and the stock of electric central heat homes) and cooling variables (cooling degree days and the stock of centrally air-conditioned homes) is depicted in Table 4 in terms of the percentage difference between actual and predicted peak loads. Although multicollinearity tends to confound the effects of individual exogenous variables, the percentage difference between actual and estimated peaks is in the relatively small combined load factor model. For the models where cooling or heating variables alone were used, the cases estimated — some had an insufficient number of observations for estimation of the combined model — had larger such differences than did the combined model.

Figures 5 and 6 present plots of the one-parameter estimate of the load distribution superimposed on the empirical (actual) load distribution. The plots are for Central Hudson Gas and Electric Corporation of New York for 1963 and Southern California Edison for 1974. Although each of these estimated load distributions is a good "fit" by the Kalnogorov-Smirnov nonparametric test, careful inspection of the plots suggests two potential problems: first, since minimum load is set at zero, the estimated load distribution will always misrepresent the actual load distribution at this extremity. In a sense, we are counting on one parameter to capture both the "shape" of the load distribution and the location of minimum load. Of course, this simplified version of the

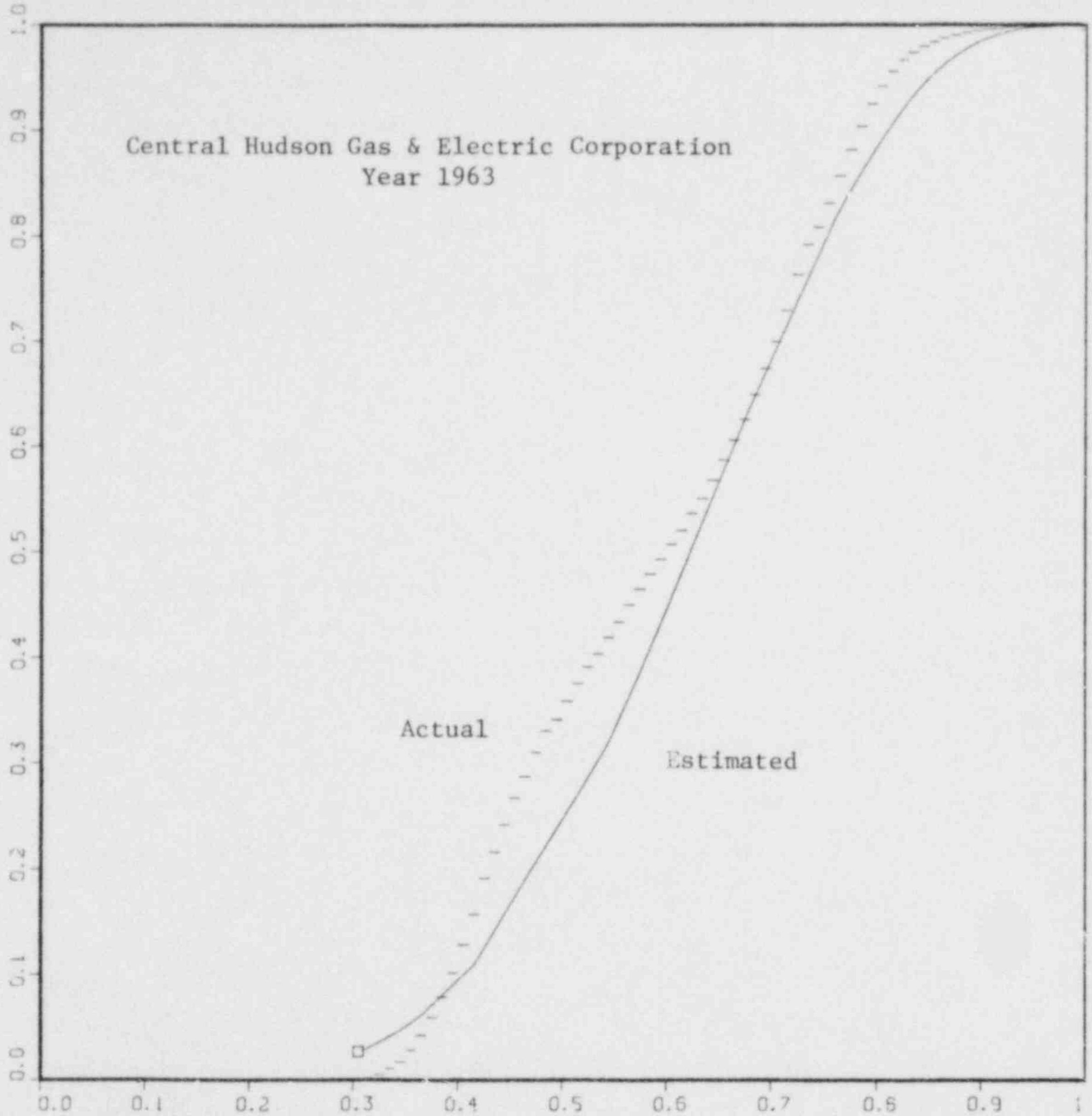

Fig. 5. Estimated and actual load distributions: version I: $L_* = 0$, one parameter.

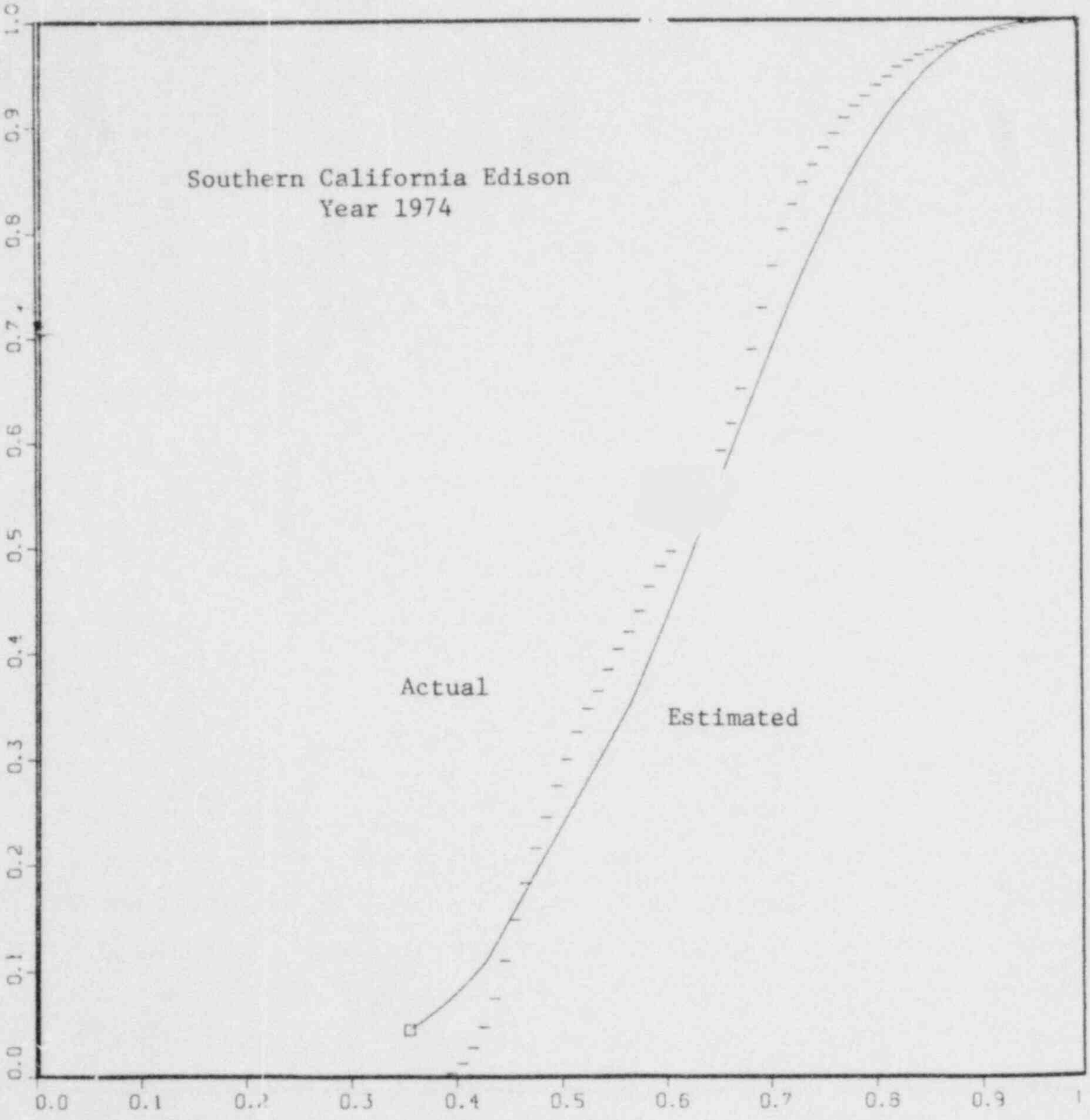

Fig. 6. Estimated and actual load distributions: version I: $L_* = 0$, one parameter.

LD model cannot produce an estimate of m·ıimum load. Second, there appears to be a tendency for the estimated load distribution to lie below the actual one. Plots for other years for these utilities exhibit this tendency in opposite: the estimated load distribution tends to lie above the actual one. This phenomenon is a consequence of fixing one parameter in each utility's load distributions for all years. In computing "b" of equation (10), we computed the method of moments estimator of b using equations (15) for each year then averaged them over all years. Thus, by analogy, just as an average tends to be either larger or smaller than any actual value used to compute it, so the one-parameter load distribution estimate tends to lie above or below the actual one. This tendency will present a problem outside the range of the sample for which the single value of "b" was estimated because the tendency of the estimated load distribution to lie above or below the actual one will be accentuated. Consequently, since future load distributions are expected to change considerably, the one-parameter method will not likely give even an adequate estimate of the load distribution outside this range.

5.2.2 Version I with only one restriction: minimum load set at zero

The effect of imposing the restriction $L_* = 0$ can be seen by relaxing the assumption that one of the load distribution parameters is fixed. Table 5 presents estimates of the first two moments of normalized load imposing $L_* = 0$ and assuming that the stochastic residuals of equations (14) are uncorrelated. Note that the load factor and first moment of normalized load are identical under this restriction [see equations (16)]; the second moment of normalized load will be the load factor variance. Table 5 presents the model estimates of the load factor and load factor variance with the same exogenous variables as were used in Table 1 except that heating and cooling variables are now separated. The same statistical measures as presented in Table 1 are presented here too.

Note that the coefficients of determination (R^2) again indicate that these models "fit" the data well. In only one case, F SNY, is the

Table 5. Single-Equation Estimates: Load Factor and Variance*
(standard errors in parentheses)

Utility	Dependent variable	CDD	CDD x LNAC	LNAC	HDD	LNHEL x HDD	LNHEL	dw	R^2	Years of data
American Electric Power	LF	-.8642 (.5089)	-.0914 (.0459)	-58.06 (4.104)				2.484	.9999	1970-74
	LF				.0694 (.0090)	-341.7 (46.01)	-.0420 (.0071)	1.511	.9999	
	Var	-.1313 (.0440)	-.0123 (.0040)	-1.289 (.3550)				2.493	.9973	
	Var				-.0005 (.0007)	-4.398 (3.512)	-.0002 (.0005)	2.755	.9981	
Carolina Power and Light	LF	.4513 (.0196)	-284.4 (30.83)	.2068 (.0228)				1.065	.9984	1962-74
	LF				.1882 (.0090)	-281.3 (75.09)	.0860 (.0219)	1.994	.9967	
	Var	.0093 (.0015)	-11.69 (2.294)	.0061 (.0017)				1.198	.9933	
	Var				.0052 (.0006)	-6.939 (4.950)	.0016 (.0015)	1.275	.9889	
Central Hudson	LF	1.077 (.0528)	-150.1 (11.41)	.2607 (.0252)				2.449	.9991	1960-74
	LF				.1198 (.0027)	-123.7 (14.72)	.0237 (.0023)	2.536	.9997	
	Var	.0027 (.0015)	-5.434 (.3312)	.0014 (.0007)				1.711	.9992	
	Var				.0007 (.0002)	-3.809 (.8134)	.0001 (.0001)	1.783	.9992	
Central Illinois Public Service	LF	.8524 (.0975)	141.7 (14.88)	-.2034 (.0235)				2.815	.9986	1965-74
	LF				.1109 (.0064)	146.7 (40.63)	-.0286 (.0072)	2.818	.9989	
	Var	.0314 (.0054)	4.245 (.8276)	.0078 (.0013)				2.329	.9940	
	Var				.0037 (.0004)	2.988 (2.313)	.0008 (.0004)	2.634	.9953	

Table 5. (cont.)

Utility	Dependent variable	CDD	CDD x LNAC	LNAC	HDD	LNHEL x HDD	LNHEL	dw	R^2	Years of data
Commonwealth Edison of Chicago	LF	.7284 (.4886)	78.44 (47.84)	-.1049 (.1171)				2.597	.9930	1968-74
	LF				.1116 (.0413)	197.1 (129.1)	-.0365 (.0265)	2.084	.9921	
	Var	.0468 (.0264)	1.113 (2.587)	-.0086 (.0063)				2.178	.9808	
	Var				.0056 (.0023)	7.155 (7.172)	-.0018 (.0015)	1.743	.9772	
Florida Power	LF	.1862 (.0074)	-369.9 (158.6)	.1179 (.0485)				2.102	.9972	1961-74
	LF				.9111 (.0576)	-417.4 (45.29)	.6469 (.0706)	2.108	.9954	
	Var	.0072 (.0004)	2.028 (8.410)	-.0002 (.0026)				2.804	.9941	
	Var				.0360 (.0032)	-17.54 (2.497)	.0287 (.0039)	1.969	.9896	
Iowa Public Service	LF	.5446 (.0224)	-201.8 (15.47)	.2086 (.0185)				2.314	.9990	1962-74
	LF				.0621 (.0043)	-105.6 (27.55)	.0120 (.0041)	2.301	.9992	
	Var	.0118 (.0014)	-9.979 (.9562)	.0063 (.0011)				1.936	.9980	
	Var				.0002 (.0003)	-4.627 (2.083)	-.0000 (.0003)	2.352	.9976	
Jacksonville Electric Authority	LF	.1440 (.0179)	.0498 (.1160)	-223.2 (345.6)				3.374	.9997	1969-74
	LF				.6426 (.0888)	-575.4 (40.88)	.6551 (.0483)	2.182	.9997	
	Var	.0032 (.0020)	-.0088 (.0132)	18.20 (39.23)				3.053	.9981	
	Var				.0168 (.0085)	-27.44 (3.921)	.0204 (.0046)	2.709	.9988	

Table 5. (cont.)

Utility	Dependent variable	CDD	CDD x LNAC	LNAC	HDD	LNHEL x HDD	LNHEL	dw	R^2	Years of data
Jersey Central Power and Light	LF	.4604 (.0658)	-267.9 (21.80)	.2306 (.0414)				1.780	.9996	1967-74
	LF				.0822 (.0236)	-72.85 (58.29)	.0069 (.0068)	2.362	.9995	
	Var	-.0041 (.0056)	-8.003 (1.866)	-.0005 (.0035)						
	Var				-.0011 (.0020)	-1.815 (5.035)	-.0010 (.0006)			
Metropolitan Edison	LF	.8752 (.1859)	-59.34 (2.622)	.0788 (.0172)				2.926	.9999	1967-74
	LF				.1033 (.0082)	-231.9 (46.74)	.0366 (.0064)	2.464	.9998	
	Var	-.1118 (.0244)	-2.735 (.3435)	-.0087 (.0023)				2.632	.9977	
	Var				.0003 (.0012)	2.006 (6.586)	-.0014 (.0009)	1.856	.9969	
New Jersey Fublic Service Electric and Gas	LF	.5392 (.0441)	-199.8 (24.39)	.1936 (.0392)				1.616	.9992	1963-74
	LF				.0843 (.0027)	-88.28 (19.08)	.0115 (.0033)	1.973	.9998	
	Var	.0032 (.0024)	-5.656 (1.348)	-.0009 (.0022)				1.285	.9983	
	Var				.0012 (.0004)	.8921 (2.565)	-.0008 (.0004)	2.122	.9977	
Niagara Mohawk	LF	1.303 (.0730)	-198.5 (12.50)	.3771 (.0360)				1.319	.9996	1962-74
	LF				.1214 (.0056)	-164.5 (26.48)	.0289 (.0037)	1.650	.9995	
	Var	.0256 (.0050)	-6.042 (.8627)	.0088 (.0025)						
	Var				.0024 (.0004)	-3.510 (2.030)	.0004 (.0003)	1.424	.9960	

Table 5. (cont.)

Utility	Dependent variable	CDD	CDD x LNAC	LNAC	HDD	LNHEL x HDD	LNHEL	dw	R^2	Years of data
Northeast Utilities	LF	1.575	-172.4	.4314				3.224	.9998	1966-74
		(.0861)	(8.952)	(.0319)						
	LF				.1007	-238.6	.0388	1.993	.9995	
					(.0032)	(58.82)	(.0090)			
	Var	.0241	-6.861	.0081				2.360	.9988	
		(.0090)	(.9305)	(.0033)						
	Var				.0035	-.3123	-.0000	2.765	.9986	
					(.0002)	(3.690)	(.0006)			
Pennsylvania Electric	LF	1.586	-59.31	.1427				3.130	.9999	1967-74
		(.1589)	(2.241)	(.0147)						
	LF				.1227	-234.2	.0433	1.852	.9999	
					(.0072)	(40.76)	(.0056)			
	Var	-.0724	-1.984	-.0061				2.172	.9980	
		(.0209)	(.2941)	(.0019)						
	Var				.0004	-4.891	-.0001	2.077	.9987	
					(.0007)	(4.008)	(.0006)			
Pennsylvania Power and Light	LF	.9205	-56.14	.0823				1.749	.9999	1965-74
		(.1033)	(2.266)	(.0098)						
	LF				.1020	-231.1	.0377	2.106	.9998	
					(.0040)	(34.36)	(.0052)			
	Var	-.0234	-1.839	-.0020				1.994	.9984	
		(.0117)	(.2560)	(.0011)						
	Var				.0023	-3.838	.0003	1.916	.9983	
					(.0004)	(3.530)	(.0005)			
Power Authority of the State of New York	LF	1.117	.3072	-220.2				2.055	.9972	1962-74
		(.2275)	(.1122)	(38.98)						
	LF				.1062	.0335	-241.2	2.072	.9966	
					(.0178)	(.0118)	(83.94)			
	Var	-.0065	-.0091	2.611				2.264	.4530	
		(.0454)	(.0224)	(7.785)						
	Var				.0018	-.0002	1.811	2.407	.4360	
					(.0033)	(.0022)	(15.37)			

Table 5. (cont.)

Utility	Dependent variable	CDD	CDD x LNAC	LNAC	HDD	LNHEL x HDD	LNHEL	dw	R^2	Years of data
San Diego Gas and Electric	LF	1.007	-206.3	.3757				1.168	.9971	1961-74
		(.0751)	(31.07)	(.0600)						
	LF				.2680	-20.0	.1002	1.155	.9982	
					(.0141)	(53.64)	.0203			
	Var	.0136	-8.572	.0061				1.432	.9967	
		(.0031)	(1.273)	(.0025)						
	Var				.0030	-7.812	.0011	1.822	.9973	
					(.0007)	(2.572)	(.0010)			
Southern California Edison	LF	.7864	-214.5	.2575				1.976	.9992	1964-74
		(.0780)	(23.52)	(.0528)						
	LF				.2069	-321.2	.1062	.1526	.9997	
					(.0082)	(26.57)	.0099			
	Var	.0104	-5.628	.0018				1.638	.9984	
		(.0032)	(.9524)	(.0021)						
	Var				.0036	-9.658	.0023	2.313	.9991	
					(.0004)	(1.426)	(.0005)			
Wisconsin Electric Power Company	LF	.9948	-163.7	.2733				1.927	.9990	1962-74
		(.0725)	(11.61)	(.0320)						
	LF				.0703	-153.2	.0184	1.574	.9995	
					(.0032)	(27.08)	(.0039)			
	Var	-.0005	-7.972	-.0011				1.133	.9963	
		(.0073)	(1.164)	(.0032)						
	Var				.0008	-9.216	.0005	2.231	.9982	
					(.0003)	(2.713)	(.0004)			
Wisconsin Michigan Power Company	LF	1.318	.3673	-191.5				1.885	.9995	1962-74
		(.0591)	(.0260)	(9.459)						
	LF				.0899	-171.3	.0224	1.718	.9996	
					(.0035)	(30.13)	(.0043)			
	Var	.0200	.0042	-4.783				1.153	.9942	
		(.0058)	(.0026)	(.9325)						
	Var				.0017	-6.048	.0006	1.284	.9940	
					(.0004)	(2.147)	(.0005)			

*All values stated x 10^{-3} except dw and R^2.

unexplained variance not less than three percent and this is explained by the fact that PASNY, lacking much of a residential sector, has historically sold electricity by contractual arrangements with municipalities, particularly New York City's municipal subway, arrangements which tend to cause jumps in the load distribution. Also, although the sample sizes are too small to give an accurate picture in many cases, the Durbin-Watson tests do not indicate that serial correlation is present to any significant degree.

Plots of estimated — superimposed on actual — load distributions in Figs. 7-11 does, however, reveal some consequences of restricting minimum load to be zero. Of course, the lower end of the load distribution will not be well represented with this model. Figures 7 and 8 exhibit cases where the version I with $L_* = 0$ produces a reasonably good estimate of the load distribution. But Figs. 9, 10, and 11 give examples of how poor an estimate one can get using this restriction. Figure 9 — where the estimated load distribution lies just below the actual one — gives a hint of what can happen. Figures 10 and 11 are more blatant. Even though the Kolmogorov-Smirnov test indicates a good fit, clearly there is room for improvement. What happens is that forcing the estimated load distribution through the origin restricts how "steep" the load distribution can be. Consequently, one gets an estimated load distribution in certain cases that lies almost entirely below the actual one.

5.2.3 Version I without the restrictions

From what has been shown above it would appear that simplification of version I along the lines discussed leads to undesirable results. How then does version I fare without the restrictions? Figure 12 exhibits a plot of estimated and actual load distributions where the estimated load distribution is forced through both peak and minimum loads. Clearly this estimate of the load distribution is superior to those obtained under the restriction(s). Of course the "shape" of the actual load distribution is not captured exactly but the estimate does not exhibit the tendency to lie entirely above or below the actual load distribution

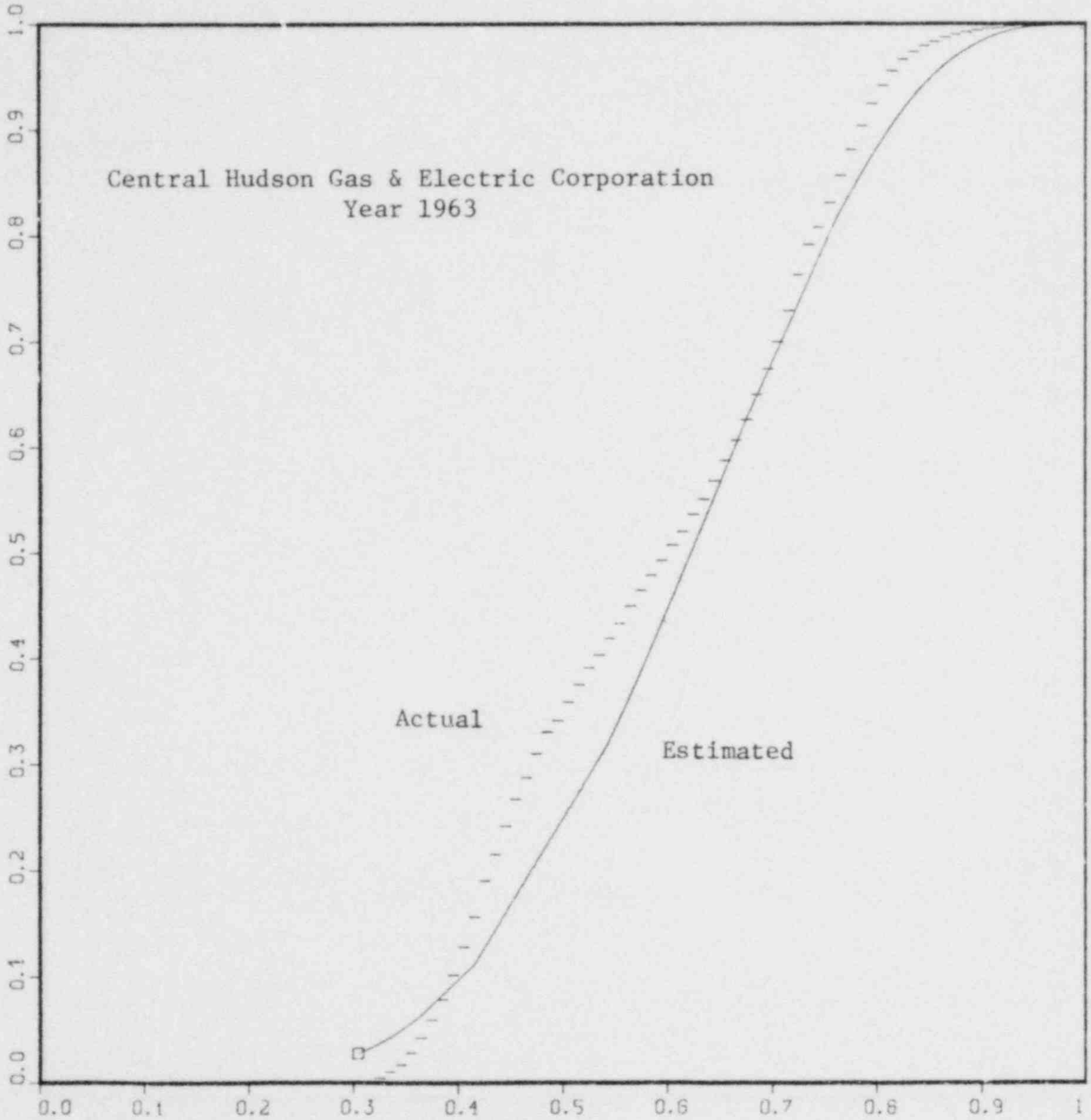

Fig. 7. Estimated and actual load distributions: version I; $L_* = 0$, two parameters.

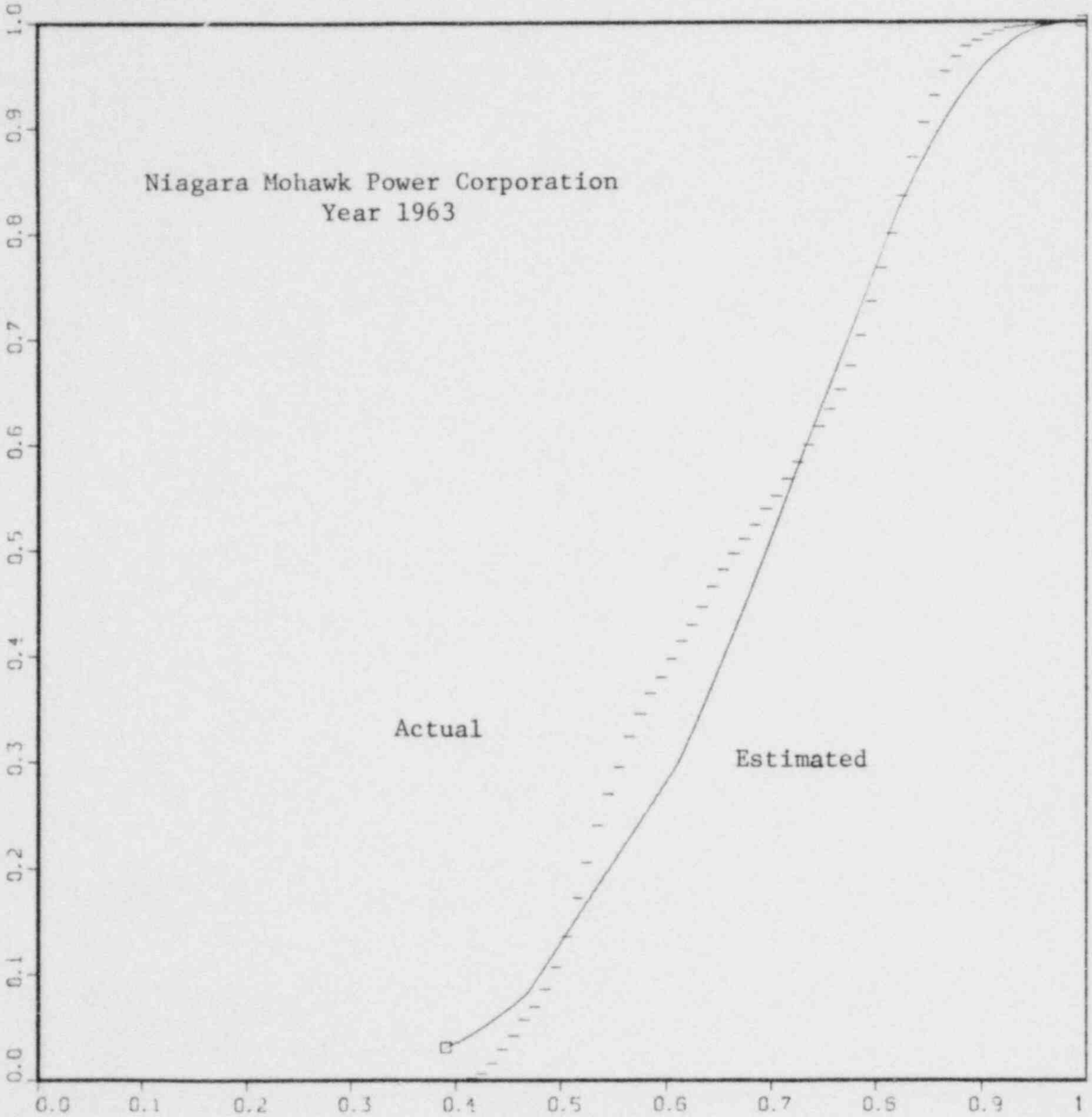

Fig. 8. Estimated and actual load distributions: version I: $L_* = 0$, two parameters.

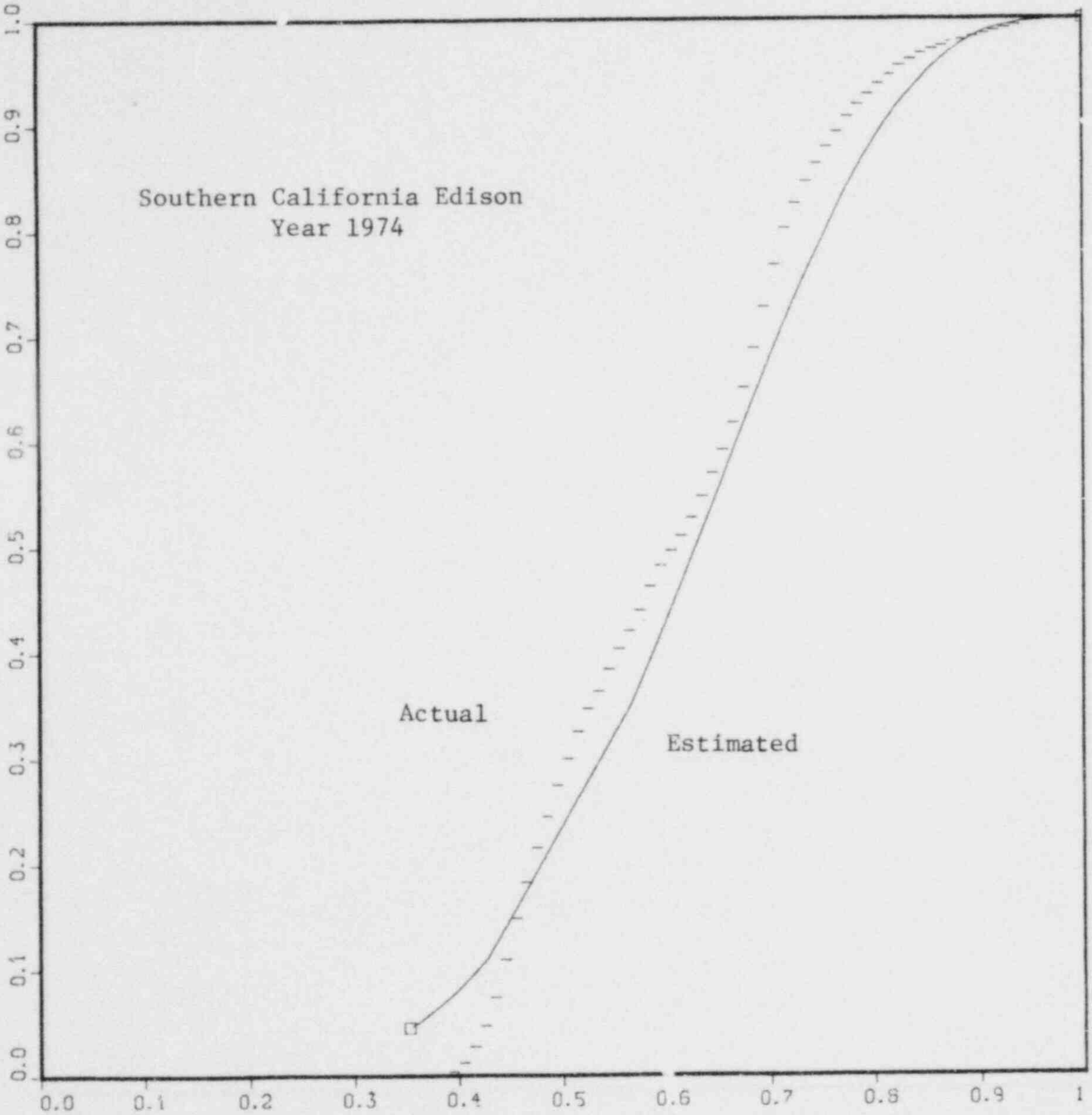

Fig. 9. Estimated and actual load distributions: version I: $L_* = 0$, two parameters.

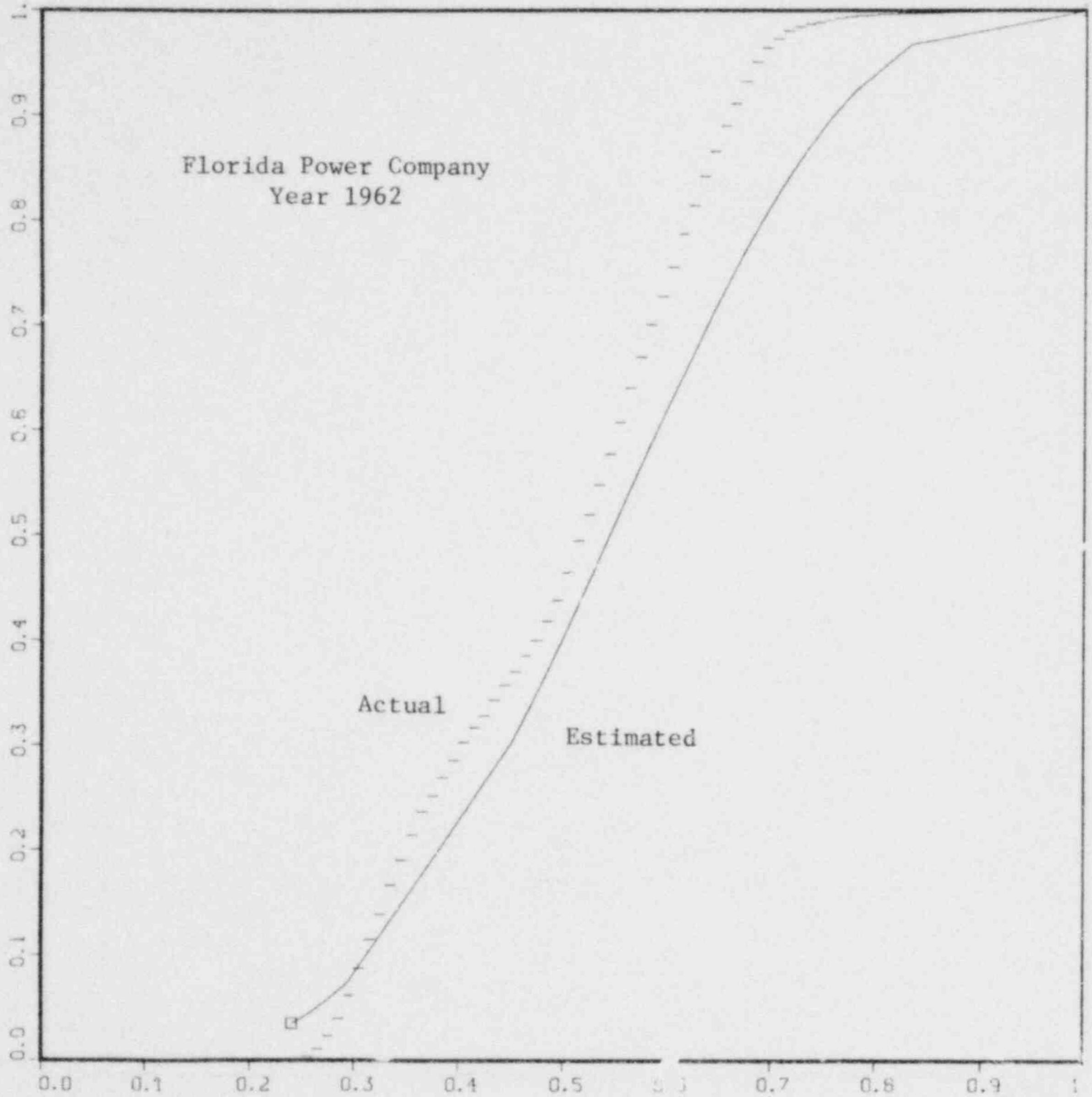

Fig. 10. Estimated and actual load distributions: version I: $L_* = 0$, two parameters.

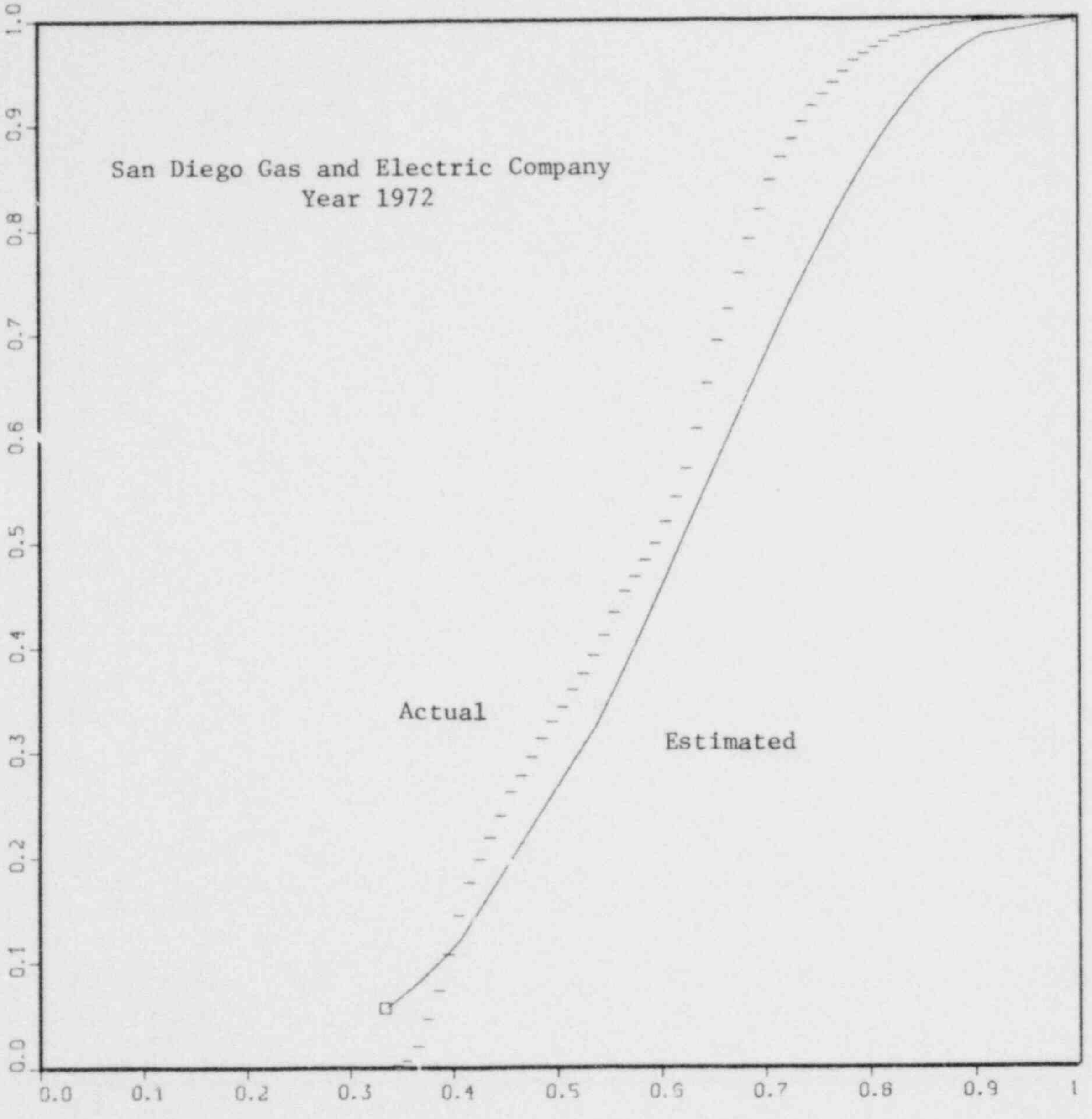

Fig. 11. Estimated and actual load distributions: version I: $L_* = 0$, two parameters.

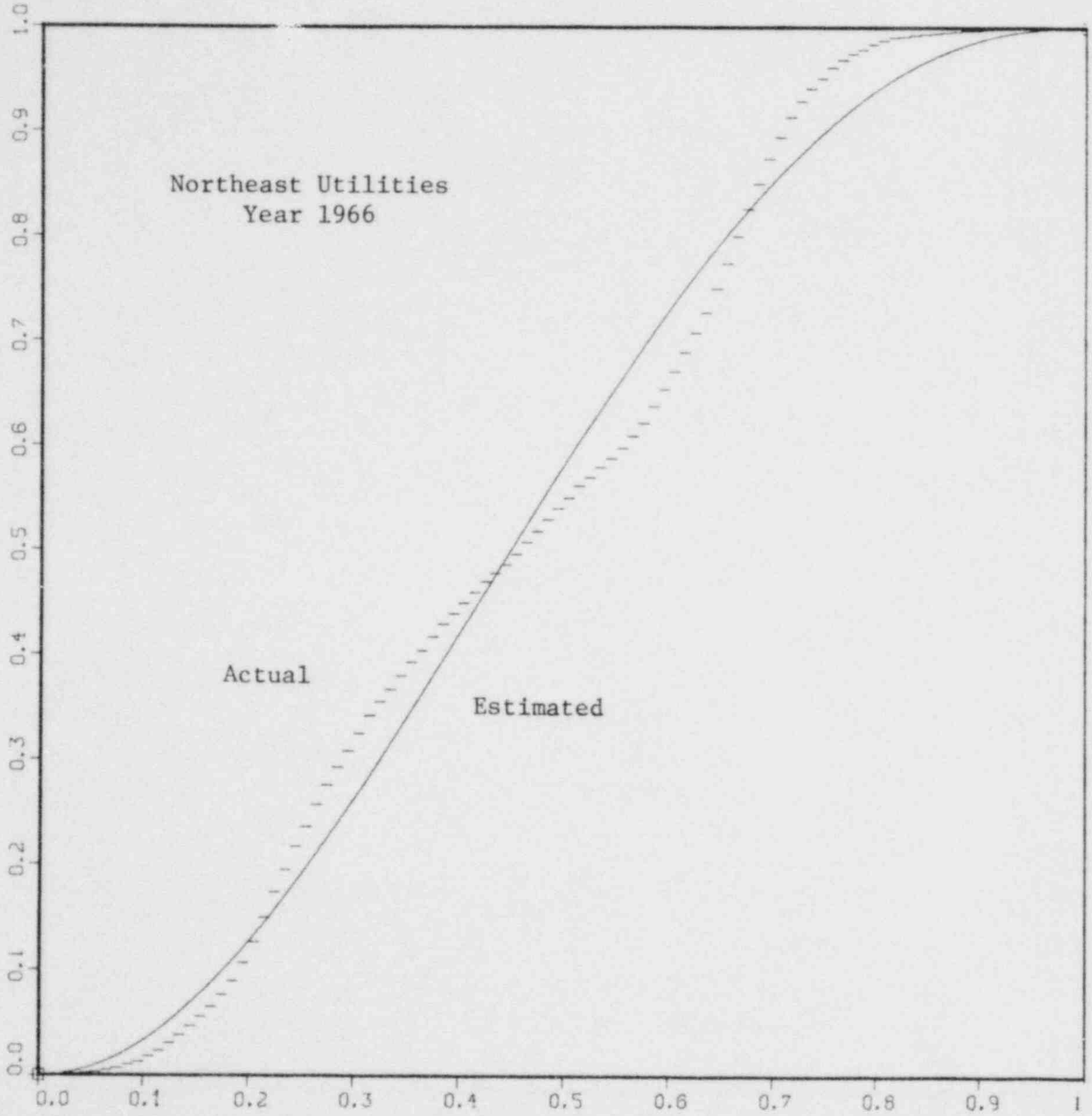

Fig. 12. Estimated and actual load distributions: version I: no restrictions.

and, when estimated, the endpoints — peak and minimum loads — will be better represented.

If the "shape" of the load distribution is important enough to seek a more accurate estimate, this can be done by estimating additional moments. Doing so allows the "bends" in the actual load distribution to be better reflected in the estimate. The next section takes up this matter of estimating additional moments; in addition, estimates presented in that section have been calculated using service-area aggregates as exogenous variables — that is, to the extent that such aggregates are available from published sources.

5.3 The LD Model: Version II

Before discussing the estimates made under version II, we will look at the potential gain in estimating the load distribution. Referring again to Fig. 12, note that the estimated load distribution is straighter than the actual load distribution which curves back and forth across it. At normalized load (Z) values of approximately 0.32, 0.68, and 0.75, the actual load distribution exhibits "bends" where a change of direction occurs which the estimated one does not. In principle, by estimating higher moments of normalized load, the estimated load distribution can be made to better reflect these "bends." Figure 13 exhibits an estimated load distribution where this has been done. Figure 14 presents the actual probability density and estimated probability density associated, respectively, with the actual and estimated load distributions of Figure 13. For these plots, actual values of peak and minimum loads were used and the normalized load distribution parameters were estimated directly from hourly loads by the method of moments procedure described in the previous section (i.e., actual rather than predicted values of sample moments of normalized load were used). Clearly the estimated load distribution of Fig. 13 better represents the "shape" of the actual load distribution than any of those presented earlier. It would thus seem that a better load distribution estimate might be got from version II.

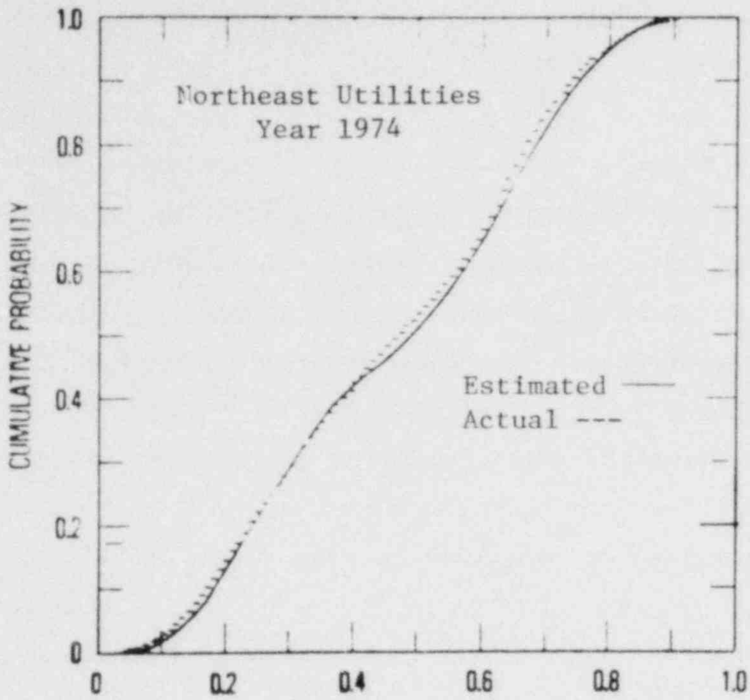

Fig. 13. Estimated and actual load distributions: version II.

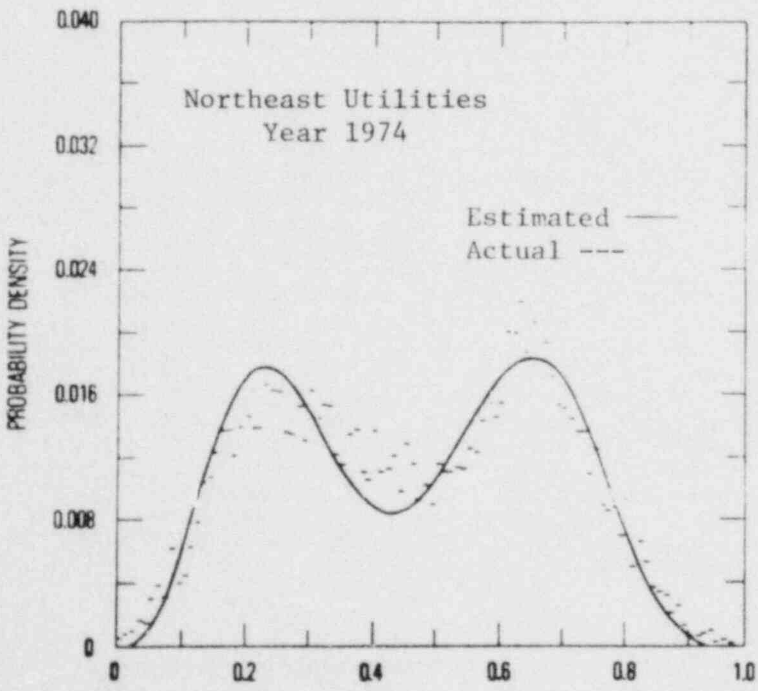

Fig. 14. Estimated probability density and actual probability distribution of load: version II.

Looking at version II performance for two utilities where service area data was readily available, Tables 6-10 present estimates of the required six moments of version II using three alternative estimation methods.

More specifically, estimates of these six moments were obtained using three distinct methods for the Central Hudson and San Diego Gas and Electric Utilities: ordinary least squares, generalized least squares and instrumental variable estimation. Additionally, each moment was also estimated with the inclusion of a seventh explanatory variable not previously used, the ratio of industrial sales to total sales (IND). IND is thought to be a significant determinant of annual demand variation by accounting for uneven growth in the industrial sector (i.e., jumps in the load distribution arising because, for example, a new industrial user of electricity moves into the service area or a new manufacturing process requiring substantially more or less electricity is introduced).

The ordinary least squares (OLS) estimates are presented in Tables 5 and 6. Each moment of normalized load (MU 1,..., MU 5) and the load factor (MU 6) is first regressed on the heating and cooling variables (cooling degree days (CDU) and heating degree days (HDU) now pertain to the service area) separately in Table 6 followed by estimates of the same models with IND included. Results are better, in terms of R^2, in the four variable model, of course, but only slightly and predominantly in MU 1 through MU 5. IND did not prove to be significant at the 10% level for any of the estimated models for Central Hudson. For San Diego, it was significant everywhere for MU 1-MU 5 only. Obviously, the importance of IND will vary with demand characteristics of individual utilities, but it should be observed that either the heating or cooling variables alone explains usually well over 90% of the variation in the first five moments of normalized load.

The major reason for not using the heating variables (HDU, LNHEL, LNHELHDU) in the same equations with the cooling variables (CDU, LNAC, LNACCDU) is, as noted previously, because of multicollinearity. This very proble owever, presents an opportunity to use instrumental variable estimation (IVE). IVE requires a matrix of variables, Z, highly correlated with the independent variables of the regression equation,

Table 6. Ordinary least squares estimates of moment functions[a]

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	LNHEL	LNHELHDU	HDU[b]	R^2
Central Hudson	MU 1	-154.8 (50.98)	.2474 (.1396)	.7353 (.2789)				.98918
	MU 1				-66.07 (108.5)	.0095 (.0173)	.0699 (.0172)	.98988
	MU 2	-93.46 (45.23)	.1366 (.1238)	.3627 (.2475)				.97275
	MU 2				-20.54 (97.24)	.0014 (.0155)	.0334 (.0154)	.97400
	MU 3	-60.42 (35.24)	.0809 (.0965)	.1955 (.1928)				.95704
	MU 3				-4.663 (75.82)	-.0011 (.0121)	.0177 (.0120)	.95892
	MU 4	-41.00 (27.08)	.0512 (.0741)	.0036 (.1482)				.94200
	MU 4				1.197 (58.26)	-.0018 (.0093)	.0101 (.0092)	.94455
	MU 5	-28.84 (20.96)	.0342 (.0574)	.0706 (.1147)				.92738
	MU 5				3.058 (45.11)	-.0018 (.0072)	.0062 (.0071)	.93055
	MU 6	-184.3 (17.82)	.3581 (.0488)	1.242 (.0975)				.99930
	MU 6				-127.8 (23.10)	.0250 (.0037)	.1223 (.0037)	.99976

Table 6. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	LNHEL	LNHELHDU	HDU[b]	R^2
San Diego Gas and Electric Co.	MU 1	-144.4	.2685	.7358				.98572
		(32.88)	(.0658)	(.1078)				
	MU 1				-194.9	.1644	.3655	98878
					(42.02)	(.0291)	(.0461)	
	MU 2	-79.51	.1661	.4341				.96776
		(25.49)	(.0510)	(.0836				
	MU 2				-108.2	.0996	.2132	.97285
					(33.72)	(.0234)	(.0370)	
	MU 3	-48.63	.1085	.2733				.94871
		(18.54)	(.0371)	(.0608)				
	MU 3				-64.67	.0633	.1337	.95562
					(24.87)	(.0172)	(.0273)	
	MU 4	-31.78	.0740	.1806				.92870
		(13.52)	(.0270)	(.0444)				
	MU 4				-40.67	.0419	.0883	.93727
					(18.29)	(.0127)	(.0201)	
	MU 5	-21.82	.0521	.1237				.90910
		(9.955)	(.0199)	(.0327)				
	MU 5				-26.55	.0286	.0604	.91892
					(13.55)	(.0094)	(.0149)	
	MU 6	-205.0	.3355	.9009				.99673
		(21.54)	(.0431)	(.0706)				
	MU 6				-204.1	.1725	.4699	.99763
					(26.45)	(.0183)	(.0290)	

Table 6. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	IND	LNHEL	LNHELHDU	HDU[b]	IND	R^2
Central Hudson	MU 1	-571.7 (151.9)	1.057 (.3038)	3.150 (.8764)	1841.5 (649.2)					.99460
	MU 1					-175.8 (135.5)	.0296 (.0229)	.1530 (.0672)	-706.2 (552.8)	.99159
	MU 2	[illegible]64.2 (134.5)	.8561 (.2690)	2.509 (.7759)	-1637.3 (574.7)					.98647
	MU 2					-123.7 (120.2)	.0203 (.0204)	.1116 (.0596)	-663.9 (490.4)	.97884
	MU 3	-353.1 (103.3)	.6489 (.2067)	1.890 (.5962)	-1292.5 (441.6)					.97926
	MU 3					-88.58 (92.80)	.0142 (.0157)	.0813 (.0460)	-540.2 (378.5)	.96726
	MU 4	-268.6 (78.34)	.4930 (.1567)	1.432 (.4521)	-1005.4 (334.9)					.97273
	MU 4					-65.60 (70.65)	.0104 (.0120)	.0608 (.0350)	-430.0 (288.1)	.95662
	MU 5	-206.6 (60.06)	.3791 (.1202)	1.100 (.3466)	-784.9 (256.8)					.96650
	MU 5					-50.07 (54.27)	.0080 (.0092)	.0465 (.0269)	-342.0 (221.4)	.94651
	MU 6	-277.7 (67.01)	.5394 (.1341)	1.783 (.3867)	-412.6 (286.4)					.99945
	MU 6					-154.8 (27.85)	.0299 (.0047)	.1428 (.0138)	-173.8 (113.6)	.99981

Table 6. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	IND	LNHEL	LNHELHDU	HDU[b]	IND	R^2
San Diego Gas and Electric Co.	MU 1	-110.6 (103.1)	.2379 (.1116)	.6467 (.2802)	158.1 (455.2)					.98589
	MU 1					-240.8 (97.81)	.1871 (.0528)	.4280 (.1286)	-202.4 (386.9)	.98908
	MU 2	-48.09 (79.72)	.1377 (.0863)	3512 (.2166)	147.1 (352.0)					.96831
	MU 2					-126.8 (79.29)	.1088 (.0428)	.2386 (.1043)	82.10 (313.6)	.97303
	MU 3	-22.50 (57.84)	.0848 (.0626)	.2043 (.1572)	122.3 (255.4)					.94986
	MU 3					-69.65 (58.64)	.0657 (.0[illegible]17)	.1405 (.0771)	-22.00 (231.9)	.95566
	MU 4	-10.96 (42.09)	.0551 (.0455)	.1257 (.1144)	97.47 (185.8)					.93061
	MU 4					-39.83 (43.14)	.0415 (.0233)	.0871 (.0567)	3.700 (170.7)	.93727
	MU 5	-5.495 (30.93)	.0373 (.0335)	.0806 (.0840)	76.40 (136.5)					.91186
	MU 5					-23.31 (31.96)	.0270 (.0173)	.0560 (.0420)	14.32 (126.4)	.91902
	MU 6	-100.5 (58.23)	.2409 (.0630)	.6251 (.1582)	489.1 (257.1)					.99760
	MU 6					-123.4 (55.52)	.1325 (.0300)	.3599 (.0730)	356.3 (219.6)	.99812

[a]All coefficients x 10^{-3}
[b]Service area heating and cooling degree days

Table 7. Instrumental variable estimates of moment functions [a]

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	IND[c]	LNHEL	LNHELHDU	HDU[b]	IND	R^2
San Diego Gas and Electric Co.	MU 1[d]	-65.96 (171.0)	.1517 (.2115)	.6590 (.4099)	44.94 (61.38)					.97995
	MU 1[e]					-138.8 (146.4)	.1319 (.0801)	.4313 (.1719)	-141.0 (507.3)	.98450
	MU 2[d]	-3.332 (128.6)	.0615 (.0703)	.3247 (.1510)	108.1 (445.5)					.95509
	MU 2[e]					-46.31 (121.9)	.0644 (.1507)	.2425 (.2922)	-41.10 (437.5)	.9617[illegible]
	MU 3[d]	13.62 (95.19)	.0270 (.1177)	.1747 (.2282)	114.7 (341.7)					.92980
	MU 3[e]					-11.21 (83.00)	.0330 (.0481)	.1449 (.1033)	.6474 (304.8)	.93678
	MU 4[d]	16.74 (66.97)	.0125 .0366	.1000 (0568)	101.8 (232.0)					.90431
	MU 4[e]					2.714 (66.60)	.0174 (.0824)	[illegible] (.1597)	15.89 (239.1)	.91020
	MU 5[d]	15.48 (50.16)	.0060 (.0620	.0597 (.1202)	85.09 (180.1)					.88016
	MU 5[e]					7.964 (48.41)	.0091 (.0265)	.0597 (.0568)	20.30 (167.7)	.88240
	MU 6[d]	-77.67 (81.45)	.2246 (.0455)	.6128 (.0956)	548.5 (282.2)					
	MU 6[e]					-100.2 (94.00)	.1151 (.1162)	.4230 (.2253)	182.5 (337.4)	.99676

Table 7. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	IND[c]	LNHEL	LNHELHDU	HDU	IND	R^2
Central Hudson	MU 1[d]	-507.9	9225.6	2.856	-1676.6					.99423
		(349.5)	(7137.)	(1.247.)	(1296.)					
	MU 1[e]					78.98	-.000009	.1914	-924.9	.98324
						(279.5)	(.0452)	(.0934)	(759.5)	
	MU 2[d]	-399.1	.7200	2.199	-1456.1					.98543
		(145.7)	(.0236)	(.0487)	(396.1)					
	MU 2[e]					109.7	-.0155	.1438	-840.8	.95769
						(349.5)	(.7137)	(1.888)	(1295.5	
	MU 3[d]	-303.0	543.9	1.651	-1152.8					.97767
		(157.5)	(321.7)	(.8512)	(584.0)					
	MU 3[e]					95.74	-.0141	.0157	-671.3	.93270
						(194.3)	(.0314)	(.0650)	(528.1)	
	MU 4[d]	-230.8	.4136	1.251	-900.5					.94580
		2683.2	(.0004)	(.0009)	(7292.6)					
	MU 4[e]					77.11	-.0115	.0791	-527.4	.91065
						(208.4)	(.4255)	(.1126)	(772.5)	
	MU 5[d]	-177.1	.3173	.9593	.703.8					.96373
		(91.86)	(.1876)	(.4963)	(340.5)					
	MU 5[e]					60.80	-.000009	.0605	-415.8	.88675
						(115.3)	(.0187)	(.0386)	(313.5)	
	MU 6[d]	-381.9	.7488	2.375	-833.2					.99925
		(81.26)	(.0131)	(.0272)	(220.9)					
	MU 6[e]					-143.7	.0283	.1454	-189.8	.99977
						(63.16)	(.1290)	(.3413)	(234.1)	

[a]All coefficients x 10^{-3}

[b]Service area heating and cooling degree days

[c]Coefficients for IND are not estimated by the instrumental variables technique.

[d]Instrument variables are LNHEL, LNHELHDU and HDU

[e]Instrumental variables are LNAC, LNACCDU and CDU

Table 8. Instrumental variable estimates of moment functions[a]

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	LNHEL	LNHELHDU	HDU[b]	R^2
San Diego Gas and Electric Co.	MU 1[c]	-75.53 (83.39)	.1610 (.1483)	.6760 (.1660)				.97977
	MU 1[d]				-110.8 (33.27)	.1193 (.1989)	.3898 (.1983)	.98462
	MU 2[c]	-26.36 (29.51)	.0838 (.1764)	.3897 (.1760)				.95452
	MU 2[d]				-38.16 (58.99)	.0607 (.1049)	.2303 (.1174)	.96212
	MU 3[c]	-10.82 (46.70)	.0507 .8310	.2437 (.9300)				.92991
	MU 3[d]				-11.34 (20.06)	.0331 (.1199)	.1451 (.1196)	.93685
	MU 4[c]	-4.949 (15.46)	.0335 (.0924)	.1612 (.0922)				.90195
	MU 4[d]				-.4342 (32.53)	.0188 (.5790)	.0960 (.6480)	.91070
	MU 5[c]	-2.647 (24.72)	.0236 (.4400)	.1108 (.4920)				.86256
	MU 5[d]				3.945 (11.10)	.0109 (.6640)	.0658 (.6620)	.88111
	MU 6[c]	-194.5 (221.2)	.3377 (.1322)	.9424 (.1319)				.99637
	MU 6[d]				-136.4 (49.6[illegible])	.1315 (.8840)	.4773 (.9890)	.99617

Table 8. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	LNHEL	LNHELHDU	HDU[b]	R^2
Central Hudson	MU 1[c]	-182.9 (123.1)	.3089 (.2691)	.8035 (.4055)				.98878
	MU 1[d]				25.49 (273.6)	-.0045 (.0430)	.0751 (.0220)	.98907
	MU 2[c]	-116.9 (241.7)	.1870 (.0380)	.4170 (.0190)				.97187
	MU 2[d]				61.09 (108.9)	-.0111 (.2660)	.0381 (.4010)	.97014
	MU 3[c]	-79.51 (85.07)	.1220 (.2080)	.2396 (.3130)				.95551
	MU 3[d]				56.92 (187.8)	-.0105 (.030)	.0213 (.0150)	.95588
	MU 4[c]	-56.21 (145.1)	.0839 (.0230)	.1487 (.0110)				.93981
	MU 4[d]				46.61 (64.92)	-.000009 (.1500)	.0128 (.2390)	.94077
	MU 5[c]	-40.71 (50.70)	.0597 (.1240)	.0978 (.1870)				
	MU 5[d]				36.75 (111.2)	-.000007 (.018)	.0082 (.0880)	.92619
	MU 6[c]	-220.4 (113.5)	.4439 (.1790)	1.355 (.0900)				.99898
	MU 6[d]				-154.7 (26.73)	.0293 (.065)	.1216 (.098)	.99972

[a] All coefficients x 10^{-3}

[b] Service area heating and cooling degree days

[c] Instrument variables are LNHEL, LNHELHDU and HDU

[d] Instrumental variables are LNAC, LNACCDU and CDU

Table 9. Joint generalized least square estimates of moment functions: pairwise correlation[a]
(asymptotic standard errors in parentheses)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	LNHEL	LNHELHDU	HDU[b]	R^2
San Diego Gas and Electric Co. (sample size = 14)	MU 1	-124.5	.2121	.6486				
		(18.43)	(.0439)	(.0918)				
	MU 2				-85.91	.0775	.1919	.97959
					(18.76)	(.0161)	(.0325)	
	MU 2	-86.65	.0750	.1841				
		(17.96	(.0157)	(.0320)				
	MU 3				-37.19	.0723	.2132	.96386
					(9.961)	(.0241)	(.0509)	
	MU 3	-35.68	.0745	.2240				
		(9.336)	(.0231)	(.0497)				
	MU 4				-29.40	.0301	.0758	.93103
					(9.111)	(.0082)	(.0170)	
	MU 4	-30.53	.0296	.0725				
		(9.227)	(.0082)	(.0170)				
	MU 5				-14.67	.0320	.0931	.90909
					(5.074)	(.0125)	(.0268)	
	MU 5	-15.93	.0431	.1188				
		(8.672)	(.0180)	(.0316)				
	MU 6				-194.4	.1656	.4679	.99772
					(23.01)	(.0167)	(.0284)	
	MU 6	-201.1	.1688	.4618				
		(25.44)	(.0179)	(.0288)				

Table 9. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	LNHEL	LNHELHDU	HDU[b]	R^2
Central Hudson (sample size = 12)	MU 1	-124.3 (14.67)	.1639 (.0732)	.5981 (.2164)				
	MU 2				-68.28 (25.29)	.0086 (.0047)	.0296 (.0137)	.98507
	MU 2	-67.41 (24.42)	.0091 (.0046)	.0326 (.0136)				
	MU 3				-40.17 (.9856)	.0329 (.0500)	.1313 (.1485)	.96833
	MU 3	-40.82 (8.965)	.0308 (.0481)	.1206 (.1445)				
	MU 4				-26.59 (13.01)	.0026 (.0025)	.0089 (.0079)	.95011
	MU 4	-27.02 (13.59)	.0029 (.0026)	.0101 (.0080)				
	MU 5				-17.40 5.524	.0080 .0290	.0381 (.0869)	.87374
	MU 5	-29.89 (20.88)	.0365 (.0572)	.0731 (.1145)				
	MU 6				-128.4 (23.0)	.0251 (.0037)	.1224 (.0037)	.99878
	MU 6	-131.1 (22.91)	.0255 (.0037)	.1223 (.0037)				

[a] All coefficients x 10^{-3}

[b] Service area heating and cooling degree days

Table 10. Joint generalized least square estimates of moment functions: pairwise correlation[a]
(asymptotic standard errors in parentheses)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	IND	LNHEL	LNHELHDU	HDU[b]	IND	R^2
San Diego Gas and Electric Co. (sample size = 14)	MU 1	-380.8 (118.8)	.6695 (.2347)	2.095 (.7085)	-1114. (543.3)					
	MU 2					-97.90 (94.40)	.0149 (.0164)	.0788 (.0534)	-400.6 (440.7)	.97959
	MU 2	-99.82 (93.16)	.0154 (.0162)	.0813 (.0530)	-419.0 (437.3)					
	MU 3					-223.4 (79.77)	.3863 (.1574)	1.177 (.4762)	-801.1 (365.7)	.96386
	MU 3	-224.0 (80.00)	.3866 (.1579)	1.184 (.4775)	-809.9 (366.6)					
	MU 4					-54.02 (54.89)	.0080 (.0095)	.0441 (.0312)	-296.4 (257.4)	.93103
	MU 4	-52.84 (55.07)	.0078 (.0095)	.0439 (.0312)	-293.0 (257.9)					
	MU 5					-131.6 (46.67)	.2271 (.0921)	.6897 (.2784)	-503.9 (213.7)	.86600
	MU 5	-215.2 (57.46)	.3973 (.1147)	1.146 (3.338)	-814.4 (248.9)					
	MU 6					-159 [illegible] 26.66	.0307 .0045	.1432 .0136	-178.5 111.9	.94737
	MU 6	-154.8 (27.85)	.0299 (.0047)	.1428 (.0138)	-173.8 (113.6)					

Table 10. (cont.)

Utility	Dependent Variable	LNAC	LNACCDU	CDU[b]	IND	LNHEL	LNHELHDU	HDU[b]	IND	R^2
Central Hudson (Sample size = 12)	MU 1	-49.84	.1496	.4571	360.7					
		(79.38)	(.0803)	(.2244)	(373.6)					
	MU 2					-47.63	.0610	.1476	165.8	
						(62.89)	(.0341)	(.0860)	(257.7)	.99074
	MU 2	-81.04	.0746	.1836	23.24					
		(60.00)	(.0326)	(.0825)	(246.5)					
	MU 3					-7.741	.0486	.1431	138.8	
						(42.42)	(.0426)	(.1202)	(200.7)	.97738
	MU 3	3.441	.0416	.1226	190.2					
		(40.97)	(.0410)	(.1163)	(194.4)					
	MU 4					-9.151	.0212	.0524	86.37	
						(31.55)	(.0172)	(.0435)	(129.9)	.96825
	MU 4	-21.55	.0262	.0640	37.63					
		(31.57)	(.0172)	(.0435)	(129.9)					
	MU 5					1.570	.0187	.0527	77.92	
						(21.91)	(.0219)	(.0622)	(104.0)	.95455
	MU 5	21.35	.0078	.0156	175.6					
		(28.63	(.0301)	(.0792)	(130.0)					
	MU 6					-88.85	.1137	.3254	465.0	
						(52.13)	(.0281)	(.0696)	(209.5)	.99561
	MU 6	-119.4	.1284	.3505	369.1					
		(54.96)	(.0297)	(.0725)	(218.1)					

[a] All coefficients x 10^{-3}

[b] Service area heating and cooling degree days

$\underline{X}$, to estimate the coefficients, β_r $(r = 1, ..., 6)$ of equations (21) in the following manner:

$$\beta_r = (Z_r'X_r)^{-1}Z_r'\mu_r \quad (r = 1, ..., 5)$$

$$\beta_6 = (Z_6'X_6)^{-1}Z_6'LF$$

where $\underline{Z}$ is an appropriately dimensioned matrix of heating (cooling) variables and $\underline{X}$ is a conformable matrix of cooling (heating) variables. The advantage in using this technique is in obtaining consistent estimators which do not result under OLS when the necessary conditions are met.

IVE results are presented in Tables 7 and 8, where Table 8 is the estimation with the IND variable added. Each moment is regressed on both the heating and cooling model with the alternate one as the instrumental variable matrix. The results are not remarkably different from OLS although R^2 is noticeably reduced in the case of the instrumental variable estimation. IND does not appear significant for either utility. IVE is primarily used when there are errors in measuring both the dependent and independent variables and it is also necessary that these errors be independent of one another, of the true values of the variables and of the disturbance term. If the data do not fit these requirements then it isn't clear that the estimators will be superior to those obtained by OLS.

Errors do exist in the hourly load information from the utilities and in the appliance stock data. The stock of houses, used in both LNHEL and LNAC, is calculated from the number of electricity customers, and the number of electrically heated and centrally air conditioned households is also estimated. It is not certain, then, that the error terms are uncorrelated across variables or that the disturbance terms are independent. In fact, considering that MU 1-MU 6 represent moments of the same probability distribution, it is highly likely that the disturbance terms are mutually correlated.

In a situation where a series of equations have mutually correlated disturbance terms, it is possible to treat the series as seemingly unrelated regression equations. The only link between these equations is an assumed constant covariance over all observations. This type of link, if it exists, would mean that the application of OLS to each equation would produce coefficients that are unbiased and consistent, but not efficient. By combining equations, it is possible to use the information concerning their correlation that is lost when each equation is regressed separately.

One way of determining which equations to combine is to check the six moments in each utility for the degree of correlation. For both Central Hudson and San Diego, the first five moments were all highly positively correlated with one another, but each was correlated to a much lesser extent with MU 6. For simplicity, then, and for ease of comparison with the other methods, consecutive overlapping pairs (i.e., MU 1 and MU 2, MU 2 and MU 3, etc.) were combined and estimated by the joint-generalized least squares (GLS) method. These results are presented in Table 9 and 10, again, with the latter being the case where IND is included.

In comparing the results in Tables 5 and 9, it is evident that the variance of the regression coefficients is generally smaller for the GLS estimates. The more efficient estimators evidences that the disturbance terms in the moment functions are indeed correlated regardless of whether the heating or cooling terms are used. Using the GLS estimates combines two moment of normalized load models and thus also allows information from all six variables to be used simultaneously, although the value of doing so for predictive purposes has yet to be investigated. The R^2 terms in the OLS single equation regressions are higher for both the heating and cooling equations than the corresponding joint equations in estimating each moment for both utilities. The differences in the R^2 for both methods for MU 6 were, however, slight. The addition of the IND variable in Table 10 produced no significant improvements. In some cases, R^2 even declined with the inclusion of the additional variable due to the method of calculating a joint correlation coefficient.

By all of the estimation methods tried, moments of normalized load estimates under version II seem to fit the data very well and to otherwise produce reasonable statistical results. However, as it turns out, small relative differences between moment estimates can lead to unacceptable results when attempting to numerically solve for the normalized load distribution parameters. Such differences always arise, of course, because the regression of a moment of normalized load on a set of exogenous variables (such as was done in Tables 6-10) will always produce moment estimates that differ from the actual moments, a difference that is acknowledged and accounted for in the specification of a stochastic residual.

What happens when one or more of these differences is larger than can be tolerated is either no numerical solution for the normalized load distribution parameters is possible or, if a solution exists, it produces negative parameter estimates. The load distribution specification of equation (19) is undefined for negative parameters. If proper estimates of the load distribution parameters cannot be obtained, the estimation procedure breaks down and must be modified. For this reason we favor version I since no such problem arises and note that, in a more complete specification of econometric models for the moments of normalized load that takes account of inter-moment correlation, this sort of problem may not arise. However, this seems a high price to pay since it complicates the estimation procedure still further.

To deal with one final matter, we have not presented any discussion in this report of the signs of coefficients in the moment equations. This was intentional. The purpose of the report was to present the results of our data explorations and to suggest that the information contained in the load duration curve might be captured by estimating moments. A complete integration of this line of analysis with consumer theory has yet to be done. We leave that to future work.

6. CONCLUSION

We have presented in this report a model that extends the traditional model of electricity demand to account for intra-period load variation, the kind of variation that is important for evaluating prices which reflect the marginal cost of producing electricity. All that is required to use this model in addition to that for estimating a traditional model of electricity demand is a compilation of hourly loads (preferably by class of customer) and some additional time in estimation. Apart from hourly loads, the model uses only temporally aggregated data.

The background from which load duration curves merge with economic studies of electricity demand was presented. One procedure by which load duration curves might be integrated with aggregate electricity demand and which leads to an estimate of the load duration curve was also discussed. Inherent in this discussion is that the model extension presented accounts for some of the information lost in temporal aggregation. The appropriate literature is also reviewed. Two load duration curve models are specified and estimated for as many as twenty utilities. The results indicate that one of the two model extensions considered, version I, is, in terms of its ability to provide the requisite information at a reasonable cost, recommended over the other.

APPENDIX

In this appendix we present the method used to calculate normalized load distribution parameter estimates for version II. We estimate first the mean and variance of each of the univariate beta densities in equation (19) using the method of moments as typically applied to the normal distribution.[51] The estimates of the a_i and b_i $(i = 1, 2)$ are then obtained by using the equations (15).

The first step in finding method-of-moments estimators requires that a negative root be found by numerically solving the following nonic equation:

$$C_9\ell^9 + C_8\ell^8 + C_7\ell^7 + C_6\ell^6 + C_5\ell^5 + C_4\ell^4 + C_3\ell^3 + C_2\ell^2 + C_1\ell + C_0 = 0 \quad \text{(A1)}$$

where $C_9 = 24$, $C_8 = 0$, $C_7 = 84K_4$, $C_6 = 36\mu_3^2$, $C_5 = 90K_4^2 + 72K_5\mu_3$, $C_4 = 444K_4\mu_3^2 - 18K_5^2$, $C_3 = 288\mu_3^4 - 108\mu_3K_4K_5 + 27K_4^3$, $C_2 = -(63K_4^2 + 72\mu_3K_5)\mu_3^2$, $C_1 = -96\mu_3^4K_4$, $C_0 = -24\mu_3^6$, and where K_j $(j = 4, 5)$ are the fourth and fifth sample cumulants

$$K_4 = \mu_4 - 3\mu_2^2$$

$$K_5 = \mu_5 - 10\mu_3\mu_2 \ .$$

If equation (A1) does not have a negative root, then the method of moments fails for the version II LD model. It may still, however, fail even if one exists for it is possible for equation (A1) to have a negative root and find that one or more of the a_i and b_i $(i = 1, 2)$ are negative. The beta density of equation (9) is not defined for negative values of any of these four parameters hence neither is the mixture density of equation (19).

Let the mean and variance of each of the beta densities comprising the mixture density of equation (19) be designated, respectively, μ_{1i} and μ_{2i}. Thus, we have

$$\mu_{1i} = a_i/(a_i+b_i) \quad (i = 1,2)$$

$$\mu_{2i} = a_i b_i/(a_i+b_i)^2(a_i+b_i+1) \; .$$

Now, define the differences of the μ_{ij} (j = 1, 2) from μ_1, the first moment of normalized load

$$d_1 = \mu_{i1} - \mu_1$$

$$d_2 = \mu_{i2} - \mu_1 \; .$$

If $\hat{\ell}$ is a negative root of (A1) and S is defined by

$$S = \frac{-8\hat{\mu}_3\hat{\ell}^3 + 3\hat{K}_5\hat{\ell}^2 + 6\hat{\mu}_3 K_4\hat{\ell} + 2\hat{\mu}_3{}^2}{\hat{\ell}(2\hat{\ell}^3 + 3\hat{K}_4\hat{\ell} + 4\hat{\mu}_3{}^2)} \; ,$$

where "hats" over the K's and μ's indicate "predicted values" then the following are estimates of d_1 and d_2:

$$\hat{d}_1 = [S - (S^2-4\hat{\ell})^{1/2}]/2$$

$$\hat{d}_2 = [S + (S^2-4\hat{\ell})^{1/2}]/2 \; .$$

This then yields

$$\hat{\mu}_{i1} = \hat{d}_i + \hat{\mu}_1 \quad (i = 1, 2)$$

where $\hat{\mu}_1$ is the predicted value for the first moment of normalized load. Then the μ_{i2}'s are estimated by

$$\hat{\mu}_{i2} = [\hat{d}_i(2S-\hat{\mu}_3/\hat{\ell})/3] + \hat{\mu}_2 - \hat{d}_i{}^2 \quad (i = 1, 2).$$

And we get estimates of a_i and b_i $(i = 1, 2)$ from the following equations:

$$\hat{a}_i = \hat{\mu}_{i1}\hat{\mu}_{i2}^{-1}[\hat{\mu}_{i1}(1-\hat{\mu}_{i1}) - \hat{\mu}_{i2}] \quad (i = 1, 2)$$

$$\hat{b}_i = (1-\hat{\mu}_{i1})\hat{\mu}_{i2}^{-1}[\hat{\mu}_{i1}(1-\hat{\mu}_{i1}) - \hat{\mu}_{i2}] \ .$$

The mixture parameter ϕ is estimated by

$$\hat{\phi} = \hat{d}_2/(\hat{d}_2-\hat{d}_1) \ .$$

FOOTNOTES AND CITATIONS

1. See Chern et al. (1978).

2. See Tepel et al. (1980).

3. Recall that peak (L^*) and minimum (L_*) loads refer to the maximum and minimum instantaneous load in a given service area per year; recall also that maximum and minimum hourly loads per year are considered in this report to be adequate observations of peak and minimum loads.

4. See footnote 3.

5. See footnote 2.

6. Recall that F(L) is the load distribution which is a more convenient way of expressing the load duration curve.

7. Charles River Associates, Inc. (1978).

8. Aigner and Poirier (1979); see also Boyd (1976), Electric Power Research Institute (forthcoming), and Lawrence (1977).

9. See New York Power Pool (1979).

10. For an early discussion of the peak load pricing model, see Steiner (1957).

11. Galiana (1976).

12. See Charles River Associates, Inc. (1978) for a discussion of these models.

13. Gupta (1976).

14. Hausman, McFadden, and Kinnucan (1978).

15. Spann and Beauvais (1977).

16. Murray, Spann, and Pulley (1978).

17. Betancourt and Habermann (1978).

18. Uri (1976, 1977).

19. See footnote 15.

20. See footnote 16.

21. See footnote 17.

22. See footnote 16.

23. See footnote 15.

24. Mitchell (1977).

25. See footnote 15.

26. Uri (1976, 1977).

27. This holds unless regulatory agencies begin to require regular and frequent surveys of utility company customers to provide this information in a published time series record.

28. Pure time series models do not contain explanatory variables. As a consequence, they are useful as a forecasting tool but not as a tool of policy analysis; see Parzen and Pagano (1977) for an example of a pure time series model.

29. One attempt to introduce policy variables into the time series model was made by Uri (1977). However, this procedure has come under suspicion because the structural parameters so estimated tend to be unstable; see Uri (1977).

30. Einhorn (1978).

31. Platt (1978).

32. Sant (1979).

33. Trimble (1978).

34. Loney (1971).

35. Uri and Maybee (1977/8, 1978a, 1978b).

36. Maybee (1978).

37. Maybee, Randolph, and Uri (1979).

38. See footnote 37.

39. Uri and Maybee (1978b).

40. Uri and Maybee (1977/8).

41. Uri and Maybee (1978a).

42. Mitchell (1977).

43. Either peak load or the load factor maybe modeled in the general scheme presented in the second section. Because it is more compatible with method-of-moments estimation, we use the load factor.

44. Recall that an annual aggregate demand for electricity estimate is required for the estimation method described in this report.

45. See Rao (1973) p. 150.

46. The correlations could arise from a simultaneous equations specification of the model of equations (12)-(13).

47. See footnote 43.

48. See Verleger and Iascone (1977).

49. See National Oceanic and Atmospheric Administration (1960-1974).

50. See Federal Economic Regulatory Commission (1960-1974).

51. See Quandt and Ramsey (1978), p. 731.

REFERENCES

Aigner, Dennis J. and Dale J. Poierier (1979). *Electricity Demand and Consumption by Time of Use: A Survey*, Electric Power Research Institute, Palo Alto, CA; May.

Betancourt, R. and R. Habermann, Jr. (1978). "An Econometric Analysis of Peak Electricity Demand in the Short Run," presented at the joint meeting of the Allied Social Science Associations and Econometric Society, August; revised December 1978.

Boyd, J. W. (1976), ed. *Proceedings on Forecasting Methodology for Time-of-Day and Seasonal Electric Utility Loads*, Electric Power Research Institute, Palo Alto, March.

Charles River Associates, Inc. (1978). *Procedures for Independent Analyses of Electric Energy and Peak Loads by State Agencies*, Vol. 1, Charles River Associates, Boston, July, pp. 7-1-7-32.

Chern, W. S., R. E. Just, B. D. Holcomb, and H. D. Nguyen (1978). *Regional Econometric Model for Forecasting Electricity Demand by Sector and by State*, ORNL/NUREG-49, Oak Ridge National Laboratory, Oak Ridge, TN, October.

Einhorn, M. A. (1978). "A Two Stage Methodology to Investigate the Determinants of the Hourly Demand for Electricity," presented at the Electric Power Research Institute conference: Modelling and Analysis of Electricity Demand bv Time of Day, June, mimeo; forthcoming in conference *Proceedings*.

Electric Power Research Institute (forthcoming). *Proceedings* for conference: Modelling and Analysis of Electricity Demand by Time of Day (June 1978).

Galiana, F. D. (1976). "Short Term Load Forecasting," in J. W. Boyd, ed., *Proceedings on Forecasting Methodology for Time-of-Day and Seasonal Electric Utility Loads*, Electric Power Research Institute, Palo Alto, March.

Gupta, P. C. (1976). "Model Identification in Peak Power Demand Forecasting," in J. W. Boyd, ed., *Proceedings on Forecasting Methodology for Time-of-Day and Seasonal Electric Utility Loads*, Electric Power Research Institute, Palo Alto, March.

Hausman, J. A., M. Kinnucan, and D. McFadden (1978). "A Two Level Electricity Demand Model: Evaluation of the Connecticut Time of Day Pricing Test," presented at the Electric Power Research Institute conference: Modelling and Analysis of Electricity Demand by Time of Day, June, mimeo; forthcoming in conference *Proceedings*.

Loney, S. T. (1971). "A Dynamic Programming Algorithm for Load Duration Curve Fitting," in Lootsma, F. A. (ed.), *Numerical Methods for Non-Linear Optimization,* Academic Press, New York.

Lawrence, A. C. (1977), ed. *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demands,* Electric Power Research Institute, Palo Alto, December.

Maybee, J. S. (1978). "Forecasting and Approximating Utility Load Duration Curves," *Symposium Papers: Energy Modeling and Net Energy Analysis,* Institute of Gas Technology, Colora: Springs, Colorado, August 21-25, pp. 721-734.

Maybee, J. S., P. Randolph, and N. D. Uri (forthcoming 1979). "Optimal Step Function Approximations to the Load Duration Curve," *Engineering Optimization.*

Mitchell, B. M. (1977). "Econometric Estimation of Peak Electricity Demands: Comments," in A. C. Lawrence, ed., *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demands,* Electric Power Research Institute, Palo Alto, December, pp. 2-23-2-30.

Murray, M. P., R. Spann, L. Pulley, and E. Beauvais (1978). "The Demand for Electricity in Virginia," *The Review of Economics and Statistics,* LX, pp. 585-600.

New York Power Pool (1979). *Report of Member Electric Systems of the New York Power Pool and the Empire State Electric Energy Research Corporation,* Vol. 1, New York Power Pool, New York, April.

Parzen, E. and M. Pagano (1977). "An Approach to Modeling Seasonally Stationary Time Series," in A. C. Lawrence, ed., *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demands,* Electric Power research Institute, Palo Alto, December, pp. 2-55-2-72; also in A. Lawrence and D. Aigner, eds., *Annals of Applied Econometrics 1979-* ., 9, pp. 137-153.

Platt, H. D. (1978). "Modeling System Load Curves by Time of Day," presented at the Electric Power Research Institute conference: Modelling and Analysis of Electricity Demand by Time of Day, June, mimeo; forthcoming in conference *Proceedings.*

Quandt, R. E. and J. B. Ramsey (1978). "Estimating Mixtures of Normal Distributions and Switching Regressions," *Journal of the American Statistical Association,* 73, December, pp. 730-738.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications,* John Wiley, New York.

Sant, D. J. (1979). "The Implication of Prices for Load Management in Electric Power Generation," *Southern Economic Journal,* 45, pp. 743-759.

Spann, R. M. and E. C. Beauvais (1977). "Econometric Estimation of Peak Electricity Demands," in A. C. Lawrence, ed., *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demands*, Electric Power Research Institute, Palo Alto, December, pp. 2-1-2-22; also in (1979), A. Lawrence and D. Aigner, eds., *Annals of Applied Econometrics 1979-1*, 9, pp. 119-136.

Steiner, P. O. (1957). "Peak Loads and Efficient Pricing," *Quarterly Journal of Economics*, 71, pp. 585-610.

Tepel, Richard C., Donald R. Alvic, Jonathon M. Jay, and Alexander D. Thorne (1980 forthcoming). *Econometric Model for the Disaggregation of State-Level Electricity Demand Forecasts to the Service Area*, ORNL/NUREG/TM-359, Oak Ridge National Laboratory, Oak Ridge, Tenn.

Trimble, J. L. (1978). "Translating Annual Electrical Generation Demand into Peak Demand Using Hourly Load Variation," *Proceedings of the American Statistical Association: Business and Economics Section*, August, pp. 600-603; revised September 1978, mimeo.

Uri, N. D. (1976). "Intermediate Term Forecasting of System Loads Using Box-Jenkins Time Series Analysis," in J. W. Boyd, ed., *Proceedings on Forecasting Methodology for Time-of-Day and Seasonal Electric Utility Loads*, Electric Power Research Institute, Palo Alto, March.

Uri, N. D. (1977). "A Mixed Time-Series Econometric Approach to Forecasting Peak System Load," in A. C. Lawrence, ed., *Forecasting and Modeling Time-of-Day and Seasonal Electricity Demands*, Electric Power Research Institute, Palo Alto, December, pp. 2-31-2-49; also in A. Lawrence and D. Aigner, eds., *Annals of Applied Econometrics 1979-1*, 9, pp. 135-171.

Uri, N. D. and J. S. Maybee (1977/8). "Smooth Approximation and Forecasting of the Load Duration Curve," *Electric Power Systems Research*, 1, pp. 213-223.

Uri, N. D. and J. S. Maybee (1978a). "Forecasting the Load Duration Curve Using Box-Jenkins Time Series Analysis," *Engineering Optimization*, 3, pp. 193-200.

Uri, N. D. and J. S. Maybee (1978b). "A Methodology for Forecasting Discrete Approximations to the Load Duration Curve," *Energy Sources*, 4, pp. 125-133.

Verleger, P. K., Jr. and F. Iascone (1977). *Residential Demand for Energy: Estimates of Residential Stocks of Energy Using Capital*, Vol. II, Electric Power Research Institute, Palo Alto, January.

NUREG/CR-1256
ORNL/NUREG/TM-379
Dist. Category RE

INTERNAL DISTRIBUTION

1. M. Baker
2. D. J. Bjornstad
3. R. S. Carlsmith
4. S. A. Carnes
5. W. S. Chern
6. W. Fulkerson
7. N. E. Hinkle
8. R. B. Honea
9. H. M. Long
10. J. B. Mills
11. W. R. Mixon
12. F. R. Mynatt
13. J. P. Stovall
14. J. L. Trimble
15. D. P. Vogt
16. T. J. Wilbanks
17. D. J. Wilkes

18-19. Central Research Library
20. Document Reference Section
21. Laboratory Records (RC)
22-23. Laboratory Records
24. ORNL Patent Office
25. Nuclear Safety Information Center
26-100. Economic Analysis Section

EXTERNAL DISTRIBUTION

101. Institute of Energy Analysis, ORAU Library, Oak Ridge, Tennessee 37830

102-103. Technical Information Center, Department of Energy, Oak Ridge, Tennessee 37830

104 Assistant Manager, Office of Energy Research and Development, DOE, P. O. Box E, Oak Ridge, Tennessee 37830

105-201. Energy Demand Analysis Group Distribution List, Economic Analysis Section, Energy Division

202-427. Distribution Category RE

☆U.S. GOVERNMENT PRINTING OFFICE: 1980-640-245/035