

Normalization of Peak Demand for an Electric Utility using PROC MODEL

Mark Harris, Jeff Brown, and Mark Gilbert*
Central and South West Services, Inc.
Tulsa, Oklahoma

Abstract

This paper discusses the isolation of abnormal events and corresponding impact on annual peak electricity demand, termed peak normalization in the electric utility industry. The primary drivers of peak electricity demand are weather and its effect on customer heating and cooling units. With tools such as SAS/INSIGHT® and PROC MODEL, an econometric model is defined and incorporated in a Monte Carlo simulation to normalize peak electricity demand in a representative and unbiased manner.

Introduction

Electricity is produced and consumed simultaneously because, unlike other commodities, electricity cannot be stored. The capacity to produce electricity must be sufficient to meet demand or portions of the system will fail. Potential supply of electricity, or capacity, is attained through expensive and time consuming power plant construction projects. Consequently, the electricity industry spends a great deal of time analyzing the historical demand for electricity and planning for the future demands of customers.

Peak demand normalization assists the electric utility in planning for future supply needs and capital expenditures. The operating efficiency of the electric provider is primarily determined by capacity to meet the cyclical demands of various customer segments. During periods of very high electricity demand (peak), capacity becomes limited. In short, the resource problem of the industry is having enough capacity to meet the annual peak demand, which typically occurs in the summer or winter.

In order to plan for future peaks, an understanding of historical peaks must be attained. The primary need for an accurate forecast is to first have the ability to adequately explain history. The objective of this analysis is to determine what the annual peak would have been under "normal" demand conditions by eliminating the random fluctuations in the historical data. The objective of this paper is to demonstrate the use of SAS® in an applied corporate research need. From methodology to analytical procedures, the relationship between electricity demand and causal factors is discussed along with a demonstration of the peak normalization modeling process.

*Economic Consultant, Project Manager, and Manager, respectively, Economic Forecasting and Analysis, Central and South West Services, Inc. The authors acknowledge the contributions of the Economic Forecasting and Analysis staff and Dr. Douglas C. Montgomery, Arizona State University.

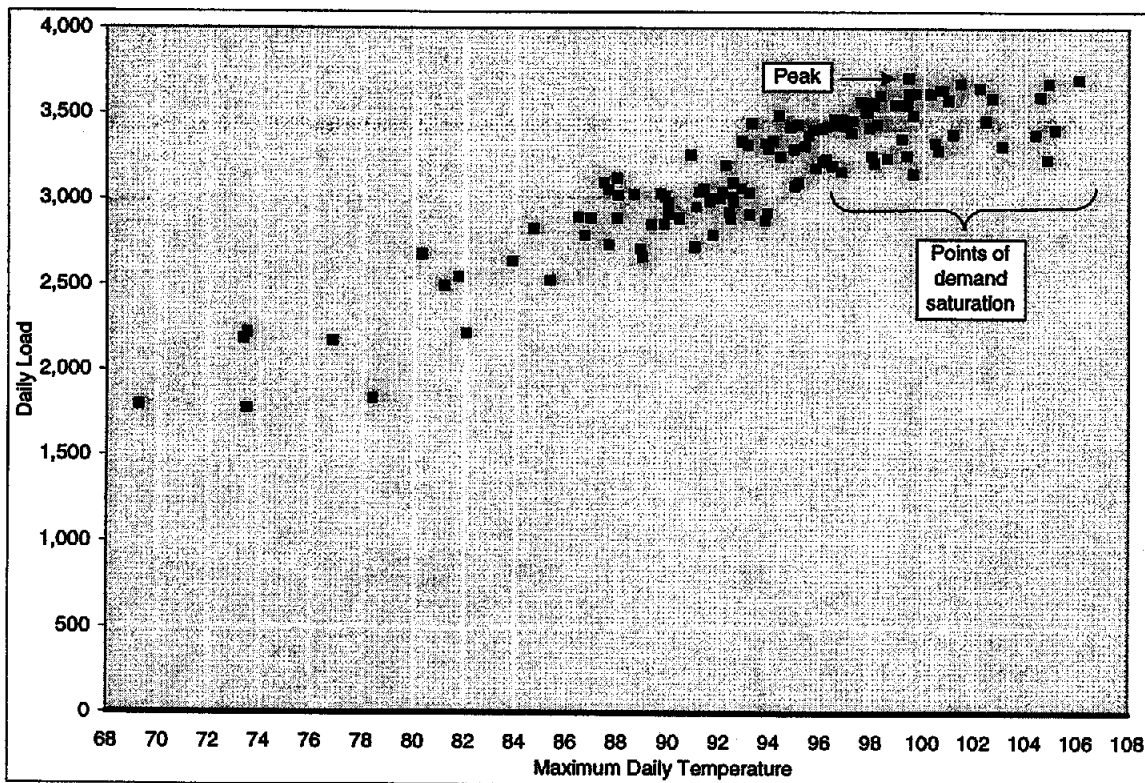
Uniqueness of Electricity Demand

Understanding annual peak electricity demand (peak) is difficult due to its unique characteristics. Peak occurs when various demand components coincide, causing electricity use to be high at that particular time. Electricity demand fluctuates constantly as customers switch various electrical appliances on and off. In addition, certain appliances such as refrigerators and air conditioners cycle on and off many times during a given period. The annual peak occurs when these cycles result in both significant and simultaneous demand. Significant demand generally occurs for an electric utility system when temperatures are extreme. Simultaneous demand occurs when the residential, commercial, and industrial sectors are all in need of large amounts of electricity at approximately the same hour of the day.

The analysis of annual peak demand is challenging because the peak only occurs once per year. Though it may seem as if there is ample data with 365 daily (8,760 hourly) data points each year, the data actually becomes limited once causal factors are identified. Causal factors include seasonal demand, days of the week, holidays, and temperature conditions. Once these factors are considered, a limited number of days per year have cause and effect conditions which are similar to conditions that actually cause the peak.

During the summer, most of the increases in daily peak load are driven by the need to cool homes and businesses. When outdoor temperatures become extremely hot, air-conditioners operate at or near capacity to maintain indoor comfort. The exact temperature that causes this higher demand may differ among customers depending on factors such as thermostat setting, efficiency, and insulation. Still, because the units are operating at or near capacity, the demand levels off at very high temperatures. This is shown in Figure 1, where peak is plotted against maximum daily temperature.

Figure 1. Daily Maximum Demand and Temperatures, June through September

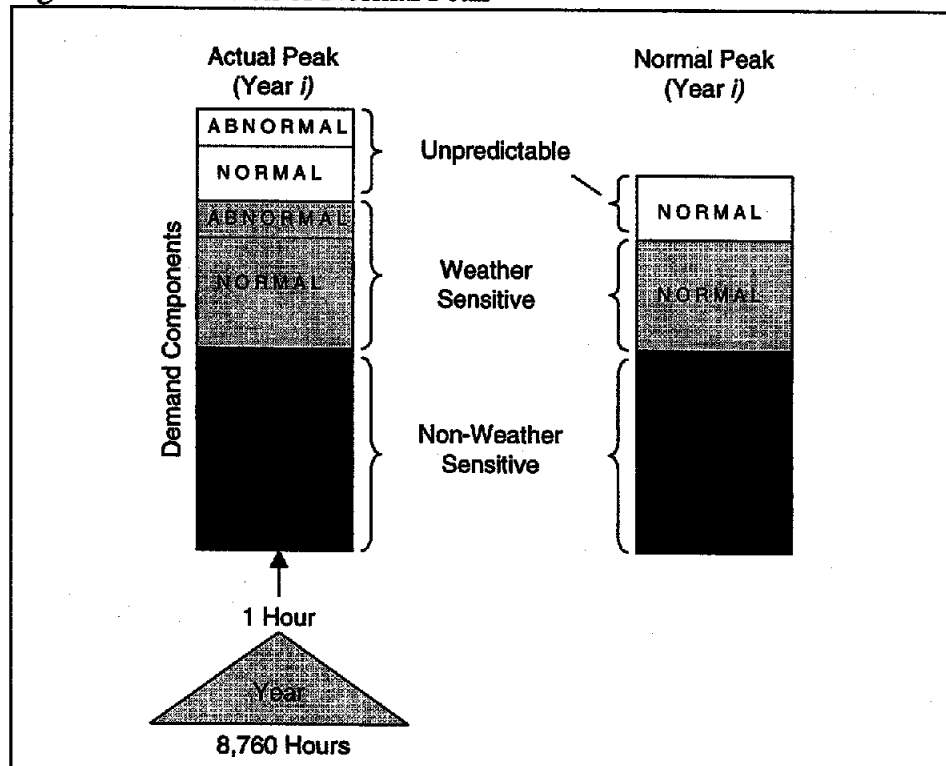


Normal Peak Derivation

Electricity has three primary demand components from an analytical perspective: non-weather sensitive, weather sensitive, and unpredictable. The non-weather sensitive demand component is not significantly influenced by temperature changes (e.g. lighting). Conversely, the weather sensitive demand component is highly responsive to changes in temperature (e.g. air-conditioners). The unpredictable demand component is made up of precarious situations that are difficult to both define and identify in a cause-effect model. Still, this unpredictable demand component is significant during peak conditions.

The weather sensitive and unpredictable demand components can be further delineated into categories of "normal" and "abnormal". The category of "normal" represents the portion of demand that can be predicted or expected based on historical occurrences, whereas the "abnormal" is more of a random event. Normal could be thought of as the average over a significant number of periods. Abnormal would then be an observation that is significantly different from the average. The normal annual peak is derived from actual annual peak by excluding the "abnormal" portions of demand. (Figure 2.)

Figure 2. Derivation of Normal Peak



The derivation of normal peaks identifies what the peaks would have been if normal weather and other conditions would have occurred. Normal peak establishes a benchmark from which to measure the additional peak demand caused by these extreme differences. Hence, strategic plans for the utility company can be based on demand estimates derived under “normal” conditions. These estimates can then form the basis for alternative scenario planning to incorporate “abnormal” conditions.

Peak Normalization Overview

Peak normalization is a fundamental process of evaluating annual peaks over time to isolate the impact of “abnormal” events. The difficulty in the process is the limited number of annual peaks and the changing conditions that affect those peaks. Since peak occurs only once each year, it is not feasible to make projections about the future based solely on these limited observations. Therefore, regression techniques are used to determine statistical relationships among a set of daily observations that are analogous with the annual peak. Then, this peak demand model is replicated numerous times in a stochastic simulation model.

Peak normalization occurs in three major steps. First, historical data is gathered and filtered. Second, estimates are obtained through econometric modeling and Monte Carlo (stochastic) model simulation. Lastly, the results are combined and evaluated in a probabilistic manner. SAS® is used extensively for each process step because of its data management and analytical capabilities.

Data Preparation

Daily temperature and daily peak load data was collected for the last several years. Each year, only the daily peak observations from June through September were used. This time period best reflects cause-effect relationships similar to those on the annual peak day. Over this historical period, the peak always occurred within these 4 months, most often in July or August. Specific patterns and relationships within the data were identified in SAS/INSIGHT®.

Due to less business demand, dummy variables were assigned to weekends, holidays, and Fridays. In addition to the temperature on the corresponding day, peak often occurs when temperatures have been high for several days (heat build up). Variables were created to reflect this phenomenon. Also, due to the non-linear relationship of peak and temperature, logarithms of temperature were applied.

Econometric Modeling / Monte Carlo Simulation

Regression techniques were employed to determine the cause-effect relationship of key variables and peak. Data exploration and correlation analysis uncovered the key independent variables that could be used to predict the peak. A general functional relationship of these key variables is shown in equation (1). Econometric techniques were used to fit this function in a way that resulted in an acceptable balance of statistical accuracy and theoretical logic. Preliminary coefficients and the autoregressive order were obtained using PROC AUTOREG.

(1) Peak = f (Minimum Daily Temp, Maximum Daily Temperature, Average Temperature over the Last Five Days, Dummy Variables)

Once the optimal model was obtained, the model structure was duplicated using PROC MODEL. Usually, this procedure is used for simultaneous equation models. However, this procedure was used in the normalization process because of its stochastic simulation ability. Essentially, the procedure estimates a single equation ordinary least squares model, and then uses the covariance matrices of the parameters and errors to generate simulation estimates.

In short, PROC MODEL provides a technique to address the two fundamental issues surrounding peak normalization. First, regardless of model accuracy, a certain amount of risk will remain. Second, the risk surrounding peak estimates should be addressed, thereby improving the utility planning process. The features in PROC MODEL allow for both of these issues to be measured.

The procedure specifications and variable definitions for the utility company in this study are shown in Table 1. The first step in the procedure reads in the NEWPREP dataset. The ID statement specifies the YEAR variable to identify observations in error messages or other listings and in the OUT= dataset. The single equation model is specified with the A and B $\hat{i}$'s identifying the intercept and parameter coefficients, respectively. Also, an autoregressive error process with a maximum likelihood procedure is specified with the %AR macro.

The FIT statement is used to estimate unknown parameters in the model by fitting the specified model equation to the observed data. The OLS option specifies that ordinary least squares will be used as the estimation method. The OUTPREDICT and OUTACTUAL options output the predicted and actual values of PEAK from the estimation to the MODLEST dataset. Similarly,

the OUTEST and OUTCOV combine to output the parameter estimates and the covariance matrix for the estimates to the COEFF dataset. Lastly, the OUTS option outputs the estimated covariance matrix of the equation errors to the S dataset. The S dataset is the covariance of the residuals computed from the parameter estimates.

Table 1. Model Specifications / Definitions

```
PROC MODEL DATA=NEWPREP;
  ID YEAR;
  PEAK = A + B1*LOGMIN + B2*LOGMAX + B3*LOGAT5 + B4*WEDUM +
    B5*DFRIDAY + Bi*DYRi ;
  %AR(PEAK,2,M=ML);
FIT PEAK / OLS DW OUT=MODLEST OUTPREDICT OUTACTUAL
  OUTEST=COEFF OUTCOV OUTS=S;
SOLVE PEAK / DATA=NEWPREP ESTDATA=COEFF SDATA=S RANDOM=30
  SEED=17 OUT=MONTE OUTACTUAL OUTPREDICT OUTRESID;
```

Where:

YEAR = year of the daily peak.

PEAK = daily peaks over time where excessive maximum temperatures occurred.

LOGMIN = natural logarithm of the daily minimum temperature.

LOGMAX = natural logarithm of the daily maximum temperature.

LOGAT5 = natural logarithm of the daily average temperature over the last 5 days.

WEDUM = dummy variable indicating weekends and holidays.

DFRIDAY = dummy variable indicating Fridays.

DYRi = dummy variables indicating the year of the daily peaks.

The SOLVE statement is used to simulate the model. This statement executes the Monte Carlo (stochastic) model simulation for input data values of PEAK. The DATA option specifies that the simulation will be based on the input dataset NEWPREP. Both the ESTDATA and SDATA options provide pseudo-random shocks during the simulation from the covariance matrices of the parameters and equation errors, respectively. The RANDOM option repeats the solution 30 times for each observation, with different random perturbations of the equations and parameters. The SEED option specifies the integer 17 to use as the seed in generating pseudo-random numbers. The final options specify that actual, predicted, and residuals will be output the MONTE dataset.

Initially, the model attained an adjusted R-square of .9836, with 1,669 daily observations over the historical period. The Monte Carlo simulation took each observation and created 30 “artificial” data points for each “actual” value. In theory, this produces a probability distribution for each estimate. Then, the maximum estimates for each year are evaluated to determine a “normal” peak. Once all of the simulation data is created, only the maximum peak estimate is kept for each simulation case ($n=30$) and each year ($n=i$). Averages and standard deviations are then computed by case, by year, and overall.

The final stage in the peak normalization process was to take the average and standard deviation information and implement another Monte Carlo simulation. This creates a normal probability distribution, which, in turn, depicts the risk and uncertainty underlying the normal peak estimate. The distribution suggests a percentage certainty that the average normal peak will fall within a certain demand range. As this derived annual normal peak is compared to the actual peak, any differences will be attributable to abnormal weather and/or an abnormal unpredictable component.

Conclusion

A utility's ability to normalize peak demand is key in understanding natural growth in the company as well as planning for future capacity needs. There are perhaps many techniques that could provide some understanding of peak demand. However, the aforementioned approach has proven to be successful. The peak normalization process has allowed the utility to accurately measure weather and abnormal impacts in demand, define statistical probabilities, and effectively plan for the company's future capacity needs.

SAS® has been an integral part of this analysis process. The data exploration tools and analytical procedures resulted in accurate and unbiased estimates. The techniques mentioned can apply to other utilities with peaks driven by temperature and non-temperature factors. The statistical methods available through PROC MODEL provide highly sophisticated scenarios that can be implemented for a wide variety of business applications.

References

SAS/ETS® User's Guide, Version 6, Second Edition, SAS Institute Inc., Cary, NC, 1993.

SAS/INSIGHT® User's Guide, Version 6, Second Edition, SAS Institute Inc., Cary, NC, 1993.

Author Contact Information

Central & South West Services, Inc.
Economic Forecasting & Analysis
Two West Second Street
Tulsa, OK 74103

Mark Harris	918-594-4246	joe_harris@csww.com
Jeff Brown	918-594-4208	jeff_brown@csww.com
Mark Gilbert	919-594-4084	mark_gilbert@csww.com

SAS, SAS/ETS, and SAS/INSIGHT are registered trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.