Developing an Electric Load Forecast Uncertainty Curve

for the Installed Reserve Margin (IRM) Study

Kelvin Chu
Carlos Villalba

Con Edison of NY Inc.
Energy Management
Electric Resource Planning
Demand Forecasting

June, 2010

1.0 Introduction

Each year a set of load forecast uncertainty curves for each zone within the New York state is developed for use in the Installed Reserve Margin (IRM) study. The Load Forecast Uncertainty (LFU) curve is a probabilistic model representing the potential annual peak demands due to random weather fluctuation. The NYISO uses these curves as one of the inputs to the reliability study to determine the required installed reserve margin for the New York State each year. The potential peak demands will influence the Installed Reserve Margin. The IRM is the amount of generation that is required in order to maintain the designated reliability criterion.

This document presents a detailed description of the procedures, assumptions and statistical tests that can be applied to develop the LFU model.

Load Forecast Uncertainty curve is a probabilistic model representing the probability of occurrence of various potential annual peak demands. The uncertainty is considered to be mainly due to random weather fluctuation from year to year and does not include any long-term economic influence.

Figure 1 shows an example of a Load Forecast Uncertainty Curve. The curve is divided into 7 sections (or 7 bins), representing a range of potential annual peak demands. The MARS reliability program can accept up to ten potential peak demands, but a general practice used in the industry has been a seven-bin model as shown in Figure 1. Each bin includes a range of input data that is equivalent to one standard deviation of the data; consequently, the probability for each bin is defined, i.e. fixed. The MARS program calculates the Loss of Load Expectation (LOLE) for each potential annual peak demand and a weighted LOLE metric is determined using the calculated LOLEs and their respective probabilities. Annual peak demands in the higher bins, i.e. bins 6 & 7, impact the most the resulting NYCA LOLE, even though the probabilities of occurrence for these peak demands are very small.

2.0 Procedures for the Development of Zonal LFU Models

Developing the Load Forecast Uncertainty Curve

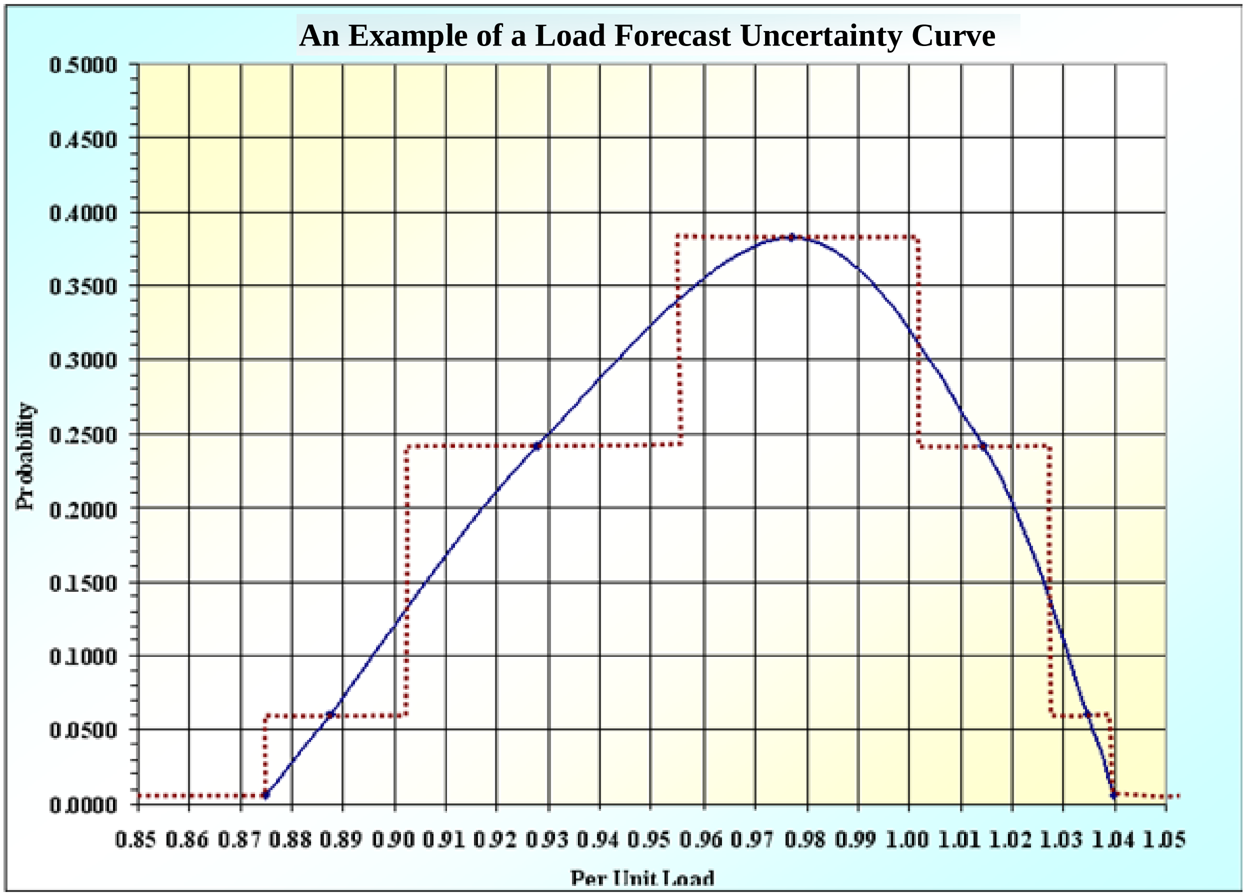

Figure 1: An Example of a Load Forecast Uncertainty Curve.

The following provides a detailed description of the approach and its respective assumptions. Figure 2 presents a flow diagram for these procedures.

Load Forecast Uncertainty Model Development Processes

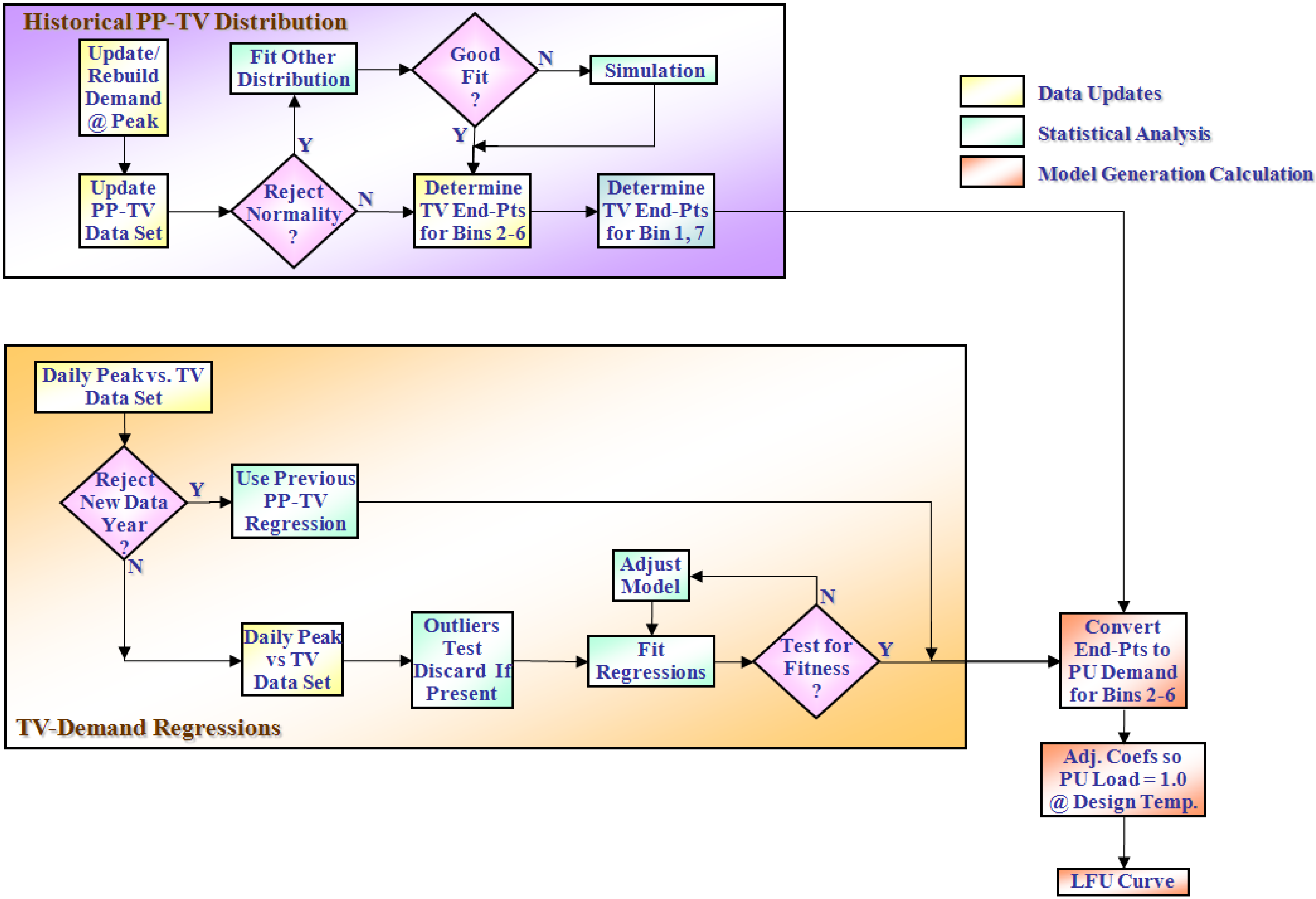

Figure 2: Load Froecast Uncertainty Model Development Processes.

2.1 Developing the Historical Peak-Producing Temperature Variable (PP-TV) Model

Each year a new temperature variable (TV) which corresponds to the peak load of the previous year is added to the historical peak-producing temperature variable (PP-TV) data set. The distribution of this set of data is assumed to be normal; however, due to the general lack of sufficient data, coupled with the characteristics of a particular zone, the normal assumption may not be valid. The normal assumption therefore needs to be confirmed whenever a new data is added to the data set. In the event that the normal assumption no longer holds, other distributions would need to be checked and tests performed to ensure a proper fitting distribution. In the event that the distribution cannot be determined, simulation would be conducted to enhance the dataset. The TV end points for Bins #2 to 6 will then be determined based on the distribution fitting or the simulation results.

Since Bins 1 and 7 represent extreme temperatures that rarely happened. There are usually very few data points available historically. As the tail ends of the normal distribution extend to infinity, it does not provide a good approximately for the cut-off points for Bins 1 and 7. There is currently no universal method of determining these points. A common consensus would be to assume that the end points are 3 standard deviations away from the mean, which represent about 99.9% of the available data. The following steps are taken to develop the PP-TV model.

1. Update the PP-TV dataset with the latest data, then test the data set for normality using several normality tests.
2. If normality cannot be rejected, then the data set is assumed to be normal; the average and the standard deviation of the new data set will be used to determine the end points from Bins 2 to 6 based on the designated probabilities for each bin.
3. If normality is rejected, then the data set will be fitted for other distributions and proceed to determine the end points for Bins 2 to 6
4. If the results of fitting the data set to other distributions are not acceptable, then simulation will be used to refine and enhance the new data set to determine the end points for Bins 2 to 6

The new data set should be tested for normality using several statistical normality tests such as the Kolmogorov-Smirnov, the Anderson-Darling, the Shapiro-Wilk and/or other similar tests. It should be noted that practically the goal of all of these tests is not to prove that a given dataset is indeed normally distributed; but rather, whether the hypothesis of normality of this data set can be rejected given sufficient evidence. The null hypothesis is that the data under consideration is not normally distributed. The alternative hypothesis, therefore, is that there is insufficient evidence to reject normality.

For example, Figure 3 shows a distribution of a hypothetical historical PP-TVs. A more detailed breakdown of the distribution shows that it does not visually resemble that of a normal distribution. However, extensive analysis using the tests mentioned above concluded that the normality of the current historical PP-TV data set cannot be rejected. The normal distribution was therefore used to model the peak producing TV.

These tests generally yield a test statistic that allows the user to determine a corresponding p-value. Given a set of data and assuming that normality is true, the p-value represents the probability that this result was observed from a normal distribution. The lower the p-value, the less likely the underlying data is normally distributed. If the p-value associated with the test statistic is lower than the designated boundary level, the test will reject the normality hypothesis. If the p-value is larger than the boundary alpha value, there is insufficient evidence to reject the normal hypothesis. For example, applying the Shapiro-Wilk test to the example historical PP-TV data set yields a W statistic of 0.9811 and a p-value of 0.7553. Since the p-value in this analysis is significantly higher than 0.05, it is not possible to reject the hypothesis that this data set is normal.

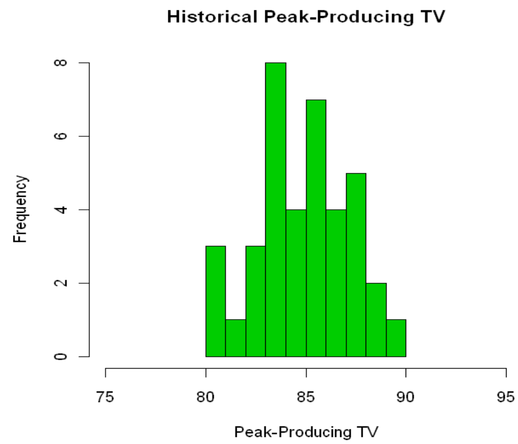

Figure 3: The Histogram of Historical Peak-Producing Temperature Variables (PP-TVs).

The ability of the Normality tests such as Shapiro-Wilk to reduce Type I (rejecting the hypothesis when the null hypothesis is true) and Type II (failing to reject the null hypothesis when the null hypothesis is false) errors depends very largely on the number of data available.

2.2 Determining the TV End Points for Bins 1 to 7

Once the distribution for the historical PP-TVs has been confirmed, the end points for the seven bins required for the Load Forecast Uncertainty calculation can now be determined.

A z transformation can be used as a convenient transformation on any distribution and has the property of centering and scaling the data to a zero mean and unit standard deviation. Using the standard z transformation and using bins separated by 1.0 standard deviation allows the construction of the table shown in Table 1. Since the distribution is assumed to be normal, the cumulative probability and the cumulative probability across bins can be determined.

Table 1: An Example Probability Distribution of PP-TVs for Various Bins.

Bin	Mid-Point	Z Begin	End	Cumulative Probability	Ind. Bin Probability	TV
1	-3.0	---	-2.5	0.00621	0.00621	77.77
2	-2.0	-2.5	-1.5	0.06681	0.06060	80.13
3	-1.0	-1.5	-0.5	0.30854	0.24173	82.49
4	0	-0.5	0.5	0.69146	0.38292	84.85
5	1.0	0.5	1.5	0.93319	0.24173	87.22
6	2.0	1.5	2.35	0.99379	0.06060	89.58
7	3.0	2.35	---	1.00000	0.00621	91.94

Since Bins 1 and 7 represent extreme temperatures that rarely happened. There are usually very few data points available historically. As the tail ends of the normal distribution extend to infinity, it does not provide a good approximation for the cut-off points for Bins 1 and 7. A different methodology is needed to determine the end points for Bins 1 and 7. As noted previously, there is no universal consensus on how these can be established. A simple solution is to set the end points for Bins 1 and 7 to be at three (3) standard deviations away from the mean. At this level, at least 99.9% of the available data would have been captured by the 7 bins.

2.3 Developing the Weather Response Curve

The distribution of the PP-TV data developed in previous section was based on a temperature metric such as the Temperature Variables (TV). The TVs must then be converted into the annual peak load demand.

1.

Once a TV-load response curve is developed, its slope can then be used to determine the appropriate MW/°F for different bins. The criterion for the selection of data and a new method of pooling data from different years to develop the response curve are discussed in the sections below. The resultant TV-Demand response curve is then used to translate the PP-TV to peak load demands in MW for each bin in the load forecast uncertainty curve.

2.3.1 Selection of TV-Demand Data for TV-Demand Regression

As noted previously, Load Demand vs. TV relationship may change over time due to changes in demography and technologies. Data collected long time ago may no longer represent the current trend of the consumption behavior. It is therefore necessary to limit the data to more recent years, i.e. 5 – 10 years. Furthermore, in order to model the high temperature saturation, it is necessary to have sufficient representation of these temperatures in the data so that the high temperature saturation effect will not be masked by the majority of the temperature data in the mid- and low-range. Since there is insufficient number of PP-TV information, the following assumptions are suggested:

1. Select only years that have TVs higher than the design TVs to serve as a proxy for the peak load demand – TV relationship.
2. Only summer month temperatures and loads for these selected years are used because these are months when peak load can occur, or when the temperatures can be close to the peak load temperatures.

2.3.2 Pooling of the TV-Load Demand Data for the TV-Demand Regression

Due to the demand growth from year to year, it is not possible to simply pool the load demands of different years into one data set and develop the regression. A simpler approach using a basic idea of "per unit" was used to pool these data together for the regression development. This approach maintains the relationship between various demand data points within a year to their respective annual peak demand. Each daily peak demand in the selected months for the selected year is rebuilt, i.e. including any load that may have been reduced due to any emergency operating procedures (EOPs), and then normalized using the weather-adjusted peak demand for that year. As a result, the growth factor from year to year is therefore neutralized and yet the relationship between the daily loads is maintained. The per-unitized demand vs. TVs can therefore be pooled together for as many years as necessary. This approach eliminates the more complicated statistical techniques that would be required if many years of demands are pooled.

The TV-demand data for each new year will be examined to determine if they should be added to the current dataset. If the year contains TVs higher than the designed TV, the year will be selected and the data will be normalized by its weather adjusted peak demand and added to the pool for regression analysis.

Figure 4 shows the data points for the TV vs. Demand in MW and TV vs. Demand in P.U.. It can be observed from the TV vs. Demand in MW graph (left) that the annual MW demands are separated from each other due to the annual demand growth. However, the per unit approach has indeed eliminated the effect of the growth and allowed the data from several years to be pooled together.

The selected and per-unitized data are then examined for outliers using various statistical tests. If outliers are present, they will be discarded from the dataset.

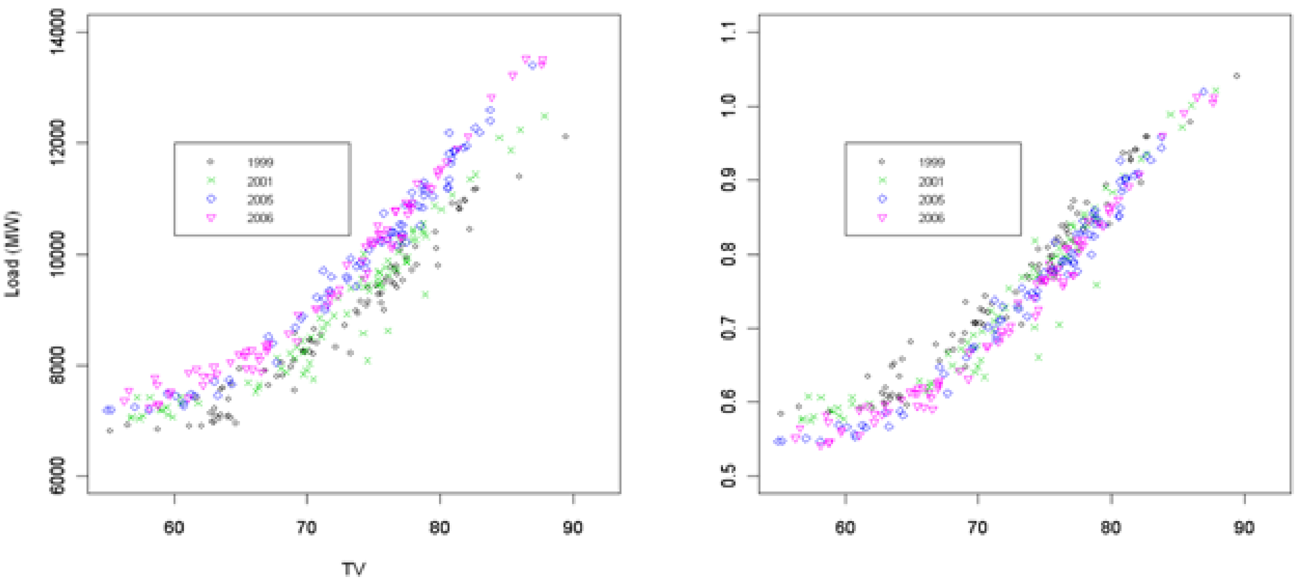

Figure 4: Demand-TV Relationship (MW Demand vs. TV & P.U. Demand vs. TV)

2.3.3 Developing the TV-Demand Regression and the Weather Response Curve

The updated dataset is then used to develop a regression model for the TVs and demands. It has been found that generally a third or fourth order polynomial model is sufficient to represent the dataset. Whether a 3^{rd} or 4^{th} order model is used depends completely on which one is a better statistically fitted model for the data at hand. The best fitted model is selected as the regression model.

Once the TV-Demand regression has been determined, the end points for the bins, which are expressed in TVs up to now, can be converted into per unit demand using the weather response curve. For example, the general equation for a 4^{th} order model is given as:

$$\text{P.U. Demand} = a_0 + a_1TV + a_2TV^2 + a_3TV^3 + a_4TV^4$$

The resulting weather response curve for a 4^{th} order polynomial is given by its derivative as:

$$\text{p.u./°TV} = a_1 + 2a_2TV + 3a_3TV^2 + 4a_4TV^3$$

The derivative of the TV-Demand regression represents the weather response curve (MW/°TV) at various TVs. The per unit value for the design TV determined from the equation above may not be exactly 1.0, therefore it is necessary to make a slight adjustment to ensure that the per unit at the design TV is unity. This can be achieved using a simple procedure of adjusting only the constant term of the model using the Excel solver. Figure 5 shows an example of a weather response function.

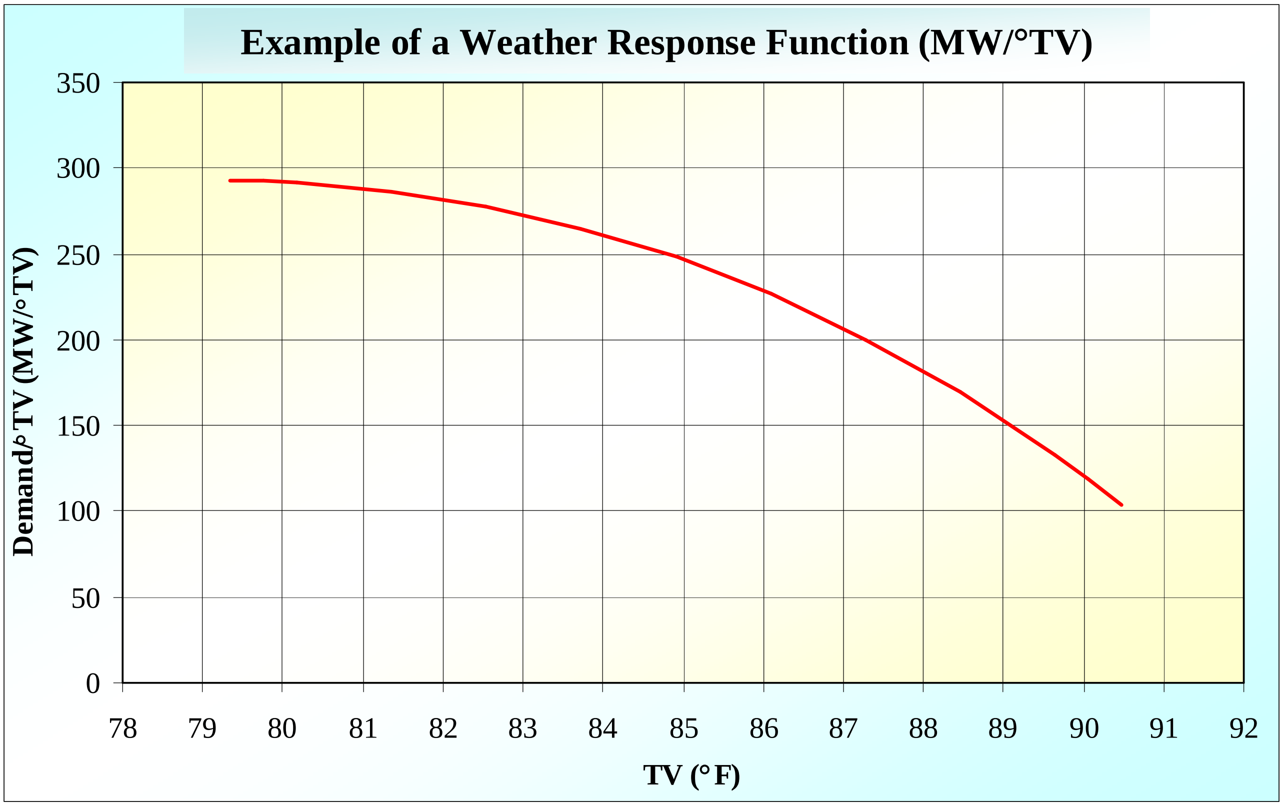

Figure 5: An Example Weather Response Curve (MW/°TV vs TV) for the LFU Calculation.

2.4 Finalizing the Load Forecast Uncertainty Curve

Using the weather response curve developed in Section 2.3.3, the per unit demand can then be calculated. Table 2 and Figure 6 show an example of the final per unit demand vs. probability.

Table 2: Per Unit Demands vs. Probabilities.

Bin	Mid-Point	Z Begin	End	Cumulative Probability	Ind. Bin Probability	TV	Per Unit Demand
1	-3.0	---	-2.5	0.00621	0.00621	77.77	0.8701
2	-2.0	-2.5	-1.5	0.06681	0.06060	80.13	0.8876
3	-1.0	-1.5	-0.5	0.30854	0.24173	82.49	0.9368
4	0	-0.5	0.5	0.69146	0.38292	84.85	0.9823
5	1.0	0.5	1.5	0.93319	0.24173	87.22	1.0211
6	2.0	1.5	2. 5	0.99379	0.06060	89.58	1.0501
7	3.0	2.5	---	1.00000	0.00621	91.94	1.0572

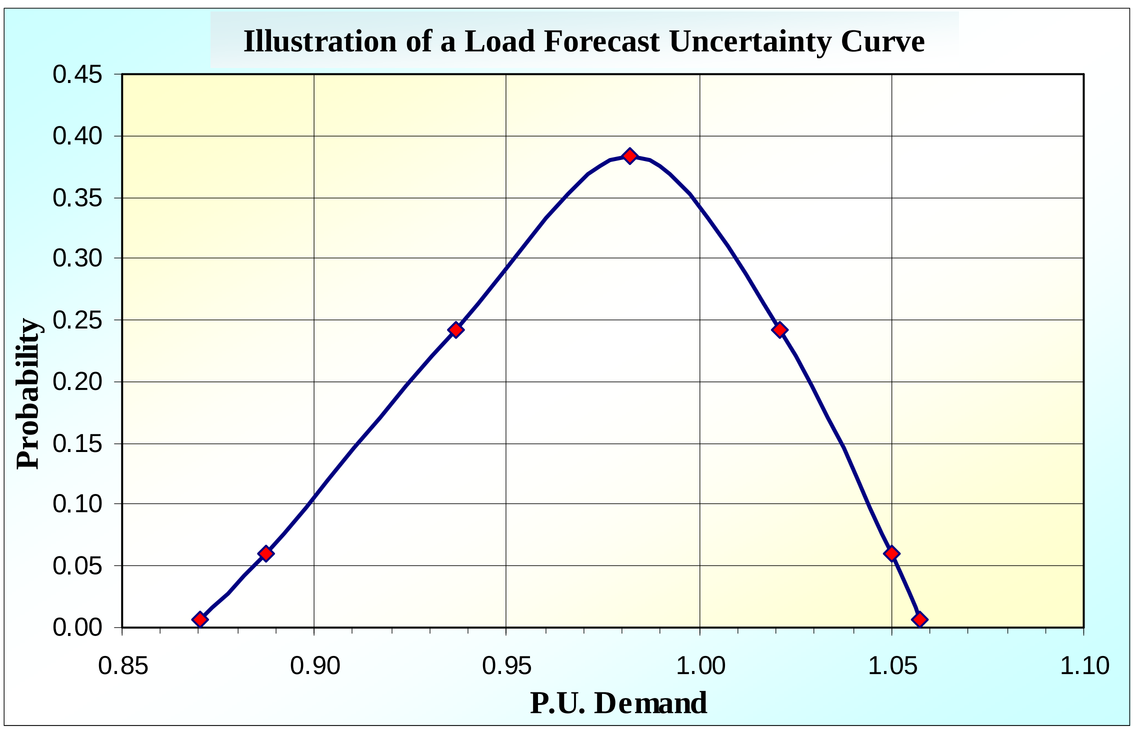

Figure 6: Illustration of a Load Forecast Uncertainty Curve – P.U. Demands vs. Probabilities

3.0 Conclusion

The above procedures can be summarized into three major steps:

1. Update the PP-TV data set for the new year and determine its distribution
2. Update the TV-Demand regression model and determine its respective weather response curve
3. Finalize the Load Forecast Uncertainty curve for the new year

The temperature variable (TV) as used in the example can be replaced by other temperature metrics that deemed appropriate by the utilities.