Exhibit No.
Issue: Weather Normalization
Witness: Mr. Mark Quan
Type of Exhibit: Direct Testimony
Sponsoring Party: Empire District Electric.
Case No.
Date Testimony Prepared: October 2009

**Before the Public Service Commission**
**Of the State of Missouri**

**Direct Testimony**

**of**

**Mark Quan**

**October 2009**

**DIRECT TESTIMONY
OF
MR. MARK QUAN
ON BEHALF OF
THE EMPIRE DISTRICT ELECTRIC COMPANY
BEFORE THE
MISSOURI PUBLIC SERVICE COMMISSION**

**Q. PLEASE STATE YOUR NAME, TITLE, AND BUSINESS ADDRESS FOR THE RECORD.**

A. My name is Mark Quan. I am a Principal Consultant for Itron's Forecasting Solutions group. My business address is 11236 El Camino Real, San Diego, California 92130.

**Q. WOULD YOU PLEASE DESCRIBE YOUR EDUCATIONAL BACKGROUND AND PRIOR ACADEMIC EXPERIENCE?**

A. I graduated from the University of California at Los Angeles with a Bachelor's Degree in Applied Mathematics with a specialization in Computer Studies. I graduated from Stanford University with a Master's Degree in Operations Research.

From 1989 to 1997, I was employed by Pacific Gas & Electric in San Francisco, California. My responsibilities at PG&E were in the areas of resource planning, gas supply planning, power contracts, and revenue requirements.

In 1997, I joined the consulting staff of Regional Economic Research ("RER"). RER was acquired by Itron in 2002. My responsibilities at

RER/Itron include performing and managing statistical analysis of client loads for the purpose of long-term forecasting and short-term forecasting. The analysis includes developing time series, multivariate regression, and neural network models for use in short term system dispatch forecasts and long-term budget, planning, and rate setting forecasts. In addition to performing analysis for clients, I am responsible for portions of Itron's forecasting training curriculum teaching introduction to forecasting, load modeling, and statistical software training classes.

**Q. HAVE YOU PREVIOUSLY FILED TESTIMONY BEFORE THIS COMMISSION?**

A. Yes. I submitted testimony on behalf of The Empire District Electric Company ("Empire") in Case No. ER-2008-0093 on the subject of weather normalization.

**Q. WHAT IS THE PURPOSE OF YOUR TESTIMONY?**

A. The purpose of my testimony is to support work I conducted to develop weather-normalized sales estimates for Empire. Using a statistical-based modeling approach, I developed weather-normalized sales for the historical test year. The test year is from July 1, 2008 through June 30, 2009. Weather-normalized sales are estimated for the following five classes: Residential, Commercial, General Power, Small Heating, and Total Electric Building.

**Q. WHAT ARE THE RESULTS FROM THE WEATHER NORMALIZATION?**

A. Applying the method described in my testimony, the normal values I calculated are show in **Table 1** to **Table 5** for each class.

**Table 1: Residential Normal Values**

| Month | Actual Billed Sales (kWh) | Normal Billed Sales (kWh) | Normal Calendar Sales (kWh) |
|---|---|---|---|
| Jul 2008 | 146,864,124 | 148,903,378 | 173,015,938 |
| Aug 2008 | 170,819,723 | 176,565,123 | 170,243,365 |
| Sep 2008 | 141,332,660 | 151,990,318 | 122,911,968 |
| Oct 2008 | 96,815,175 | 102,737,897 | 95,012,874 |
| Nov 2008 | 101,414,636 | 99,876,750 | 120,358,211 |
| Dec 2008 | 168,479,701 | 161,972,255 | 182,784,489 |
| Jan 2009 | 214,536,500 | 206,238,811 | 198,612,398 |
| Feb 2009 | 177,206,962 | 177,374,254 | 163,825,311 |
| Mar 2009 | 140,142,971 | 144,805,560 | 137,703,072 |
| Apr 2009 | 122,552,244 | 123,885,310 | 103,057,895 |
| May 2009 | 98,713,072 | 98,249,227 | 109,326,674 |
| Jun 2009 | 106,839,072 | 105,484,178 | 126,828,117 |

## Table 2: Commercial Normal Values

| Month | Actual Billed Sales (kWh) | Normal Billed Sales (kWh) | Normal Calendar Sales (kWh) |
|---|---|---|---|
| Jul 2008 | 30,787,785 | 30,938,845 | 34,364,554 |
| Aug 2008 | 33,071,410 | 33,692,647 | 32,766,931 |
| Sep 2008 | 31,300,924 | 32,601,002 | 28,431,313 |
| Oct 2008 | 24,957,520 | 25,818,405 | 24,010,588 |
| Nov 2008 | 22,264,502 | 22,201,894 | 23,317,541 |
| Dec 2008 | 27,498,398 | 27,057,578 | 29,088,058 |
| Jan 2009 | 32,819,406 | 32,261,673 | 30,372,746 |
| Feb 2009 | 27,730,334 | 27,740,864 | 26,377,198 |
| Mar 2009 | 25,346,428 | 25,635,483 | 25,837,148 |
| Apr 2009 | 24,235,284 | 24,314,309 | 22,685,509 |
| May 2009 | 23,035,973 | 23,052,665 | 25,579,826 |
| Jun 2009 | 24,709,934 | 24,556,584 | 28,225,849 |

**Table 3: GP Normal Values**

| Month | Actual Billed Sales (kWh) | Normal Billed Sales (kWh) | Normal Calendar Sales (kWh) |
|---|---|---|---|
| Jul 2008 | 76,531,239 | 76,715,637 | 80,493,657 |
| Aug 2008 | 80,483,450 | 81,035,700 | 82,019,195 |
| Sep 2008 | 79,369,471 | 80,293,105 | 73,972,194 |
| Oct 2008 | 67,489,055 | 68,020,378 | 68,263,294 |
| Nov 2008 | 59,131,793 | 59,085,058 | 56,325,220 |
| Dec 2008 | 62,720,413 | 62,408,551 | 67,435,722 |
| Jan 2009 | 68,586,350 | 68,167,521 | 65,128,741 |
| Feb 2009 | 59,213,445 | 59,152,995 | 58,256,146 |
| Mar 2009 | 59,380,880 | 59,463,609 | 60,865,767 |
| Apr 2009 | 60,624,799 | 60,681,404 | 58,468,518 |
| May 2009 | 61,245,770 | 61,427,696 | 64,235,079 |
| Jun 2009 | 66,674,479 | 66,585,665 | 70,793,952 |

**Table 4: SH Normal Values**

| Month | Actual Billed Sales (kWh) | Normal Billed Sales (kWh) | Normal Calendar Sales (kWh) |
|---|---|---|---|
| Jul 2008 | 8,520,114 | 8,554,363 | 9,139,703 |
| Aug 2008 | 9,225,310 | 9,355,454 | 9,249,744 |
| Sep 2008 | 8,573,457 | 8,804,365 | 7,956,660 |
| Oct 2008 | 7,183,533 | 7,322,287 | 7,043,413 |
| Nov 2008 | 6,407,277 | 6,367,521 | 6,747,032 |
| Dec 2008 | 9,432,199 | 9,227,093 | 10,432,661 |
| Jan 2009 | 12,106,349 | 11,712,778 | 11,114,670 |
| Feb 2009 | 10,230,498 | 10,188,651 | 9,541,841 |
| Mar 2009 | 8,403,014 | 8,511,887 | 8,396,766 |
| Apr 2009 | 7,192,731 | 7,227,071 | 6,743,078 |
| May 2009 | 6,559,590 | 6,575,094 | 7,093,359 |
| Jun 2009 | 7,112,699 | 7,086,065 | 7,701,092 |

**Table 5: TEB Normal Values**

| Month | Actual Billed Sales (kWh) | Normal Billed Sales (kWh) | Normal Calendar Sales (kWh) |
|---|---|---|---|
| Jul 2008 | 33,989,355 | 34,104,589 | 36,896,290 |
| Aug 2008 | 37,364,177 | 37,847,237 | 37,082,963 |
| Sep 2008 | 35,663,588 | 36,579,236 | 33,423,342 |
| Oct 2008 | 30,698,663 | 31,282,984 | 29,887,523 |
| Nov 2008 | 26,907,814 | 26,744,901 | 29,593,140 |
| Dec 2008 | 34,758,359 | 34,072,788 | 36,184,430 |
| Jan 2009 | 40,760,904 | 39,661,876 | 37,058,839 |
| Feb 2009 | 33,047,863 | 32,980,067 | 31,549,500 |
| Mar 2009 | 29,577,150 | 30,010,450 | 29,721,231 |
| Apr 2009 | 27,303,394 | 27,476,233 | 25,616,147 |
| May 2009 | 27,732,553 | 27,761,883 | 30,234,777 |
| Jun 2009 | 30,033,072 | 29,929,442 | 32,376,084 |

Q. **WHAT IS WEATHER NORMALIZATION?**

A. Weather Normalization is the process of determining what historical consumption would have been if normal weather conditions existed. The process is a mathematical method to adjust the existing monthly sales for a class based on a statistical model and normal weather conditions.

Q. **CAN YOU DESCRIBE THE WEATHER NORMALIZATION PROCESS?**

A. The weather-normalization process entails adjusting actual sales based on the difference between what would have happened under normal weather conditions versus what happened under actual weather conditions. The fundamental equation used in the process is shown below.

$$NormalSales_{month} = \frac{ModelNormalSales_{month}}{ModelActualSales_{month}} \times ActualSales_{month}$$

In this equation, actual monthly sales are multiplied by the ratio of modeled sales under normal conditions to modeled sales under actual conditions. For example, if the ratio of the $\text{ModelNormalSales}_{\text{month}}$ to $\text{ModelActualSales}_{\text{month}}$ is 0.90, then the $\text{ActualSales}_{\text{month}}$ should be mulitiplied by 0.90 because the model estimates that sales under normal conditions are lower than sales under actual weather conditions by approximately 10%. The method is more fully described in Schedule MQ-2.

**Q. HOW DO YOU OBTAIN THE MODELED SALES UNDER ACTUAL CONDITIONS?**

A. To obtain modeled sales under actual conditions, I developed a multivariate regression model for each class and used the model to estimate sales for using actual weather data over the test period. The regression model predicts daily load as a function of actual daily weather. The regression model is developed using customer class load research data. The independent variables include weather splines for heating and cooling responses, daytype and holiday variables for seasonal variations, and sunlight variables for lighting effects. These variables capture the changing customer consumption patterns throughout the year. The weather spline variables capture the nonlinear interaction between load and weather. I have included the regression model specifications and results for the five classes in Schedule MQ-1.

**Q. HOW DO YOU OBTAIN THE MODELED SALES UNDER NORMAL CONDITIONS?**

A. To obtain modeled sales under normal conditions, I used the same multivariate regression model mentioned above and forecast the sales using normal weather data through the test period.

**Q. IN THE MODELS, WHAT ARE THE MWH PER DEGREE CHANGE IMPACTS?**

A. Because the load-weather relationship is non-linear, a single kWh/degree number is not applicable for any class. Instead, the kWh/degree change depends on the degree at which the value is calculated. Embedded in the regression model for each class are heating and cooling degree day variables that describe the kWh/degree change at different temperature points.

In the Residential Class model, I use CDD65 and CDD70 temperature splines for cooling impacts. Associated with these variables are model coefficients that describe the kWh/degree change when temperature increases above 65 degrees. Between 66 and 70 degrees, a one degree change results in a 1.04958 kWh increase. The 1.04958 is the coefficient on the CDD65 variable. Above 70 degrees, a one degree change results in a 1.66793 (1.04958 + 0.61835) kWh increase. The 1.66793 is the sum of the coefficients on the CDD65 and CDD70 variables.

In the Residential Class, I use HDD55, HDD60, and HDD55Trend temperature splines for heating impacts. Excluding the HDD55Trend variable, a one degree change between 56 and 60 degrees results in a

0.51805 kWh increase and a one degree change below 60 degrees, a one degree change results in a 0.54727 (0.51805 + 0.02922) kWh increase. When accounting for the HDD55_Trend variable, the impact increases below 55 degrees by 0.03480 kWh multiplied by a trend factor (Year-2002 + days in year/366) based on 2002. For example, on January 1, 2008, the impact is 6.00273 (2008-2002 + 1/366) multiplied with 0.03480 kWh, or 0.20890 kWh.

For the other Classes, the model coefficients are interpreted the same way. These coefficients are shown in Schedule MQ-1.

**Q. HOW DID YOU DEVELOP NORMAL WEATHER CONDITIONS FOR THE SALES MODEL?**

A. Normal weather conditions are developed using a 30-year average of historical weather from 1979 through 2008. The averages are obtained by a Rank and Average method. In this method, historical daily average temperatures are ranked from the highest value to the lowest value in each month. For each historical day, the corresponding heating degree day (HDD) and cooling degree day (CDD) values are calculated for multiple temperature reference points. Next, the normal HDD and CDD values are calculated as the average across the 30 historical years within a month. This defines the normal hottest day of each month as the average across the hottest days in the past 30 historical years in the same month. The final step in this method is to map the ranked averages to the test year actual weather. The final result maps the normal hottest day of the month to the hottest historical day in the corresponding test year month.

**Q. WHAT ADJUSTMENT DID YOU MAKE FOR BILLING CYCLES?**

A. The fundamental equation includes billing cycle variations in the calculation. The variation is implicit in the "month" subscript. To calculate billed normal sales, I forecast the daily consumption under normal and actual conditions and aggregated the consumption based on monthly billing cycle dates. To calculate calendar normal sales, I aggregated consumption under normal conditions based on the calendar dates. However, the ratio denominator of ModelActualSales remains calculated over billing cycle dates. This ratio embeds the conversion from billing cycle sales to calendar sales as well as the conversion to normal sales.

**Q. DOES THIS CONCLUDE YOUR TESTIMONY?**

A. Yes, it does.

# REGRESSION MODEL SPECIFICATIONS AND RESULTS

## RESIDENTIAL MODEL

Model fit statistics

| | |
|---|---|
| ▪ R-Squared | 0.964 |
| ▪ Adjusted R-Squared | 0.963 |
| ▪ Mean Abs. Dev. (MAD) | 1.60 |
| ▪ Mean Abs. % Err. (MAPE) | 4.63% |
| ▪ Durbin-Watson Statistic | 2.073 |

Variable Statistics

| Variable | Coefficient | T-Stat |
|---|---|---|
| CONST | 27.06959 | 33.547 |
| DailyAverageTemperature.HDD60 | 0.02922 | 0.658 |
| DailyAverageTemperature.HDD55 | 0.51805 | 10.42 |
| WeatherTransforms.HDD55_Trend | 0.0348 | 6.598 |
| DailyAverageTemperature.CDD65 | 1.04958 | 19.743 |
| DailyAverageTemperature.CDD70 | 0.61835 | 9.127 |
| MonthlyBinary.Jan | 5.30198 | 8.286 |
| MonthlyBinary.Feb | 4.93771 | 9.672 |
| MonthlyBinary.Mar | 2.75547 | 5.817 |
| MonthlyBinary.May | 0.95295 | 2.036 |
| MonthlyBinary.Jun | 4.32128 | 8.371 |
| MonthlyBinary.Jul | 6.9458 | 12.804 |
| MonthlyBinary.Aug | 7.06532 | 13.03 |
| MonthlyBinary.Sep | 2.99283 | 5.989 |
| MonthlyBinary.Oct | 0.08486 | 0.163 |
| MonthlyBinary.Nov | 1.48822 | 2.286 |
| MonthlyBinary.Dec | 4.1413 | 4.775 |
| DOWBinary.Monday | -1.39497 | -9.443 |
| DOWBinary.Tuesday | -1.61436 | -9.154 |
| DOWBinary.Wednesday | -1.58461 | -8.417 |
| DOWBinary.Thursday | -1.6996 | -8.989 |
| DOWBinary.Friday | -2.01884 | -11.37 |
| DOWBinary.Saturday | -0.40541 | -2.807 |

| | | |
|---|---|---|
| SunTimes.FracDark17 | 6.14008 | 2.606 |
| SunTimes.FracDark8 | 1.427 | 0.919 |
| US_Holidays.NYHol | 0.57733 | 0.707 |
| US_Holidays.MLKing | 0.4613 | 0.559 |
| US_Holidays.PresidentDay | 1.17494 | 1.555 |
| US_Holidays.MemorialDay | 3.09446 | 3.738 |
| US_Holidays.July4thHol | 1.421 | 1.735 |
| US_Holidays.LaborDay | 3.97419 | 4.416 |
| US_Holidays.Thanksgiving | 0.4875 | 0.532 |
| US_Holidays.FriAftThanks | 0.87403 | 0.954 |
| US_Holidays.XMasHol | 1.17186 | 1.429 |
| MonthlyBinary.Year2006 | -2.05118 | -2.864 |
| MonthlyBinary.Year2005 | -2.38873 | -3.282 |
| MonthlyBinary.Year2004 | -2.83854 | -3.823 |
| MonthlyBinary.Year2003 | -2.70576 | -3.56 |
| MonthlyBinary.Year2002 | -2.79349 | -3.59 |
| AR(1) | 0.53902 | 27.235 |

**COMMERICAL MODEL**

Model fit statistics

- R-Squared 0.958
- Adjusted R-Squared 0.957
- Mean Abs. Dev. (MAD) 1.88
- Mean Abs. % Err. (MAPE) 3.93%
- Durbin-Watson Statistic 2.072

Variable Statistics

| Variable | Coefficient | T-Stat |
|---|---|---|
| CONST | 29.3746 | 25.046 |
| DailyAverageTemperature.HDD55 | 0.35782 | 31.027 |
| DailyAverageTemperature.CDD65 | 0.98347 | 14.273 |
| DailyAverageTemperature.CDD60 | 0.28477 | 5.103 |
| MonthlyBinary.Jan | 3.80909 | 4.875 |
| MonthlyBinary.Feb | 2.97765 | 3.917 |

| | | |
|---|---|---|
| MonthlyBinary.Mar | 0.8844 | 1.275 |
| MonthlyBinary.May | 2.51247 | 3.637 |
| MonthlyBinary.Jun | 6.34948 | 8.229 |
| MonthlyBinary.Jul | 8.88684 | 11.099 |
| MonthlyBinary.Aug | 8.86773 | 11.026 |
| MonthlyBinary.Sep | 5.70427 | 7.441 |
| MonthlyBinary.Oct | 2.50012 | 3.335 |
| MonthlyBinary.Nov | 2.29713 | 2.445 |
| MonthlyBinary.Dec | 3.34735 | 3.063 |
| DOWBinary.Monday | 11.72026 | 71.591 |
| DOWBinary.Tuesday | 12.42913 | 61.708 |
| DOWBinary.Wednesday | 12.72496 | 58.369 |
| DOWBinary.Thursday | 12.43354 | 56.812 |
| DOWBinary.Friday | 11.99073 | 59.176 |
| DOWBinary.Saturday | 3.34102 | 20.838 |
| SunTimes.FracDark17 | 4.90751 | 1.494 |
| US_Holidays.NYHol | -7.83879 | -8.727 |
| US_Holidays.MLKing | -2.12486 | -2.491 |
| US_Holidays.PresidentDay | -0.67173 | -0.809 |
| US_Holidays.MemorialDay | -11.9097 | -13.075 |
| US_Holidays.July4thHol | -14.5499 | -16.169 |
| US_Holidays.LaborDay | -12.3772 | -13.586 |
| US_Holidays.Thanksgiving | -14.2561 | -13.881 |
| US_Holidays.FriAftThanks | -5.26901 | -5.128 |
| US_Holidays.XMasHol | -8.97082 | -9.956 |
| MonthlyBinary.Year2006 | 0.7224 | 0.672 |
| MonthlyBinary.Year2005 | -1.3445 | -1.233 |
| MonthlyBinary.Year2004 | -3.46917 | -3.191 |
| MonthlyBinary.Year2003 | -1.84833 | -1.697 |
| MonthlyBinary.Year2002 | 0.51661 | 0.471 |
| AR(1) | 0.67937 | 39.112 |

**GENERAL POWER MODEL**

Model fit statistics

- R-Squared 0.968
- Adjusted R-Squared 0.965
- Mean Abs. Dev. (MAD) 215.16
- Mean Abs. % Err. (MAPE) 2.75%
- Durbin-Watson Statistic 2.076

Variable Statistics

| Variable | Coefficient | T-Stat |
|---|---|---|
| CONST | 4992.98853 | 36.363 |
| DailyAverageTemperature.HDD50 | 22.49200 | 6.858 |
| DailyAverageTemperature.CDD70 | 28.62724 | 2.546 |
| DailyAverageTemperature.CDD55 | 34.35936 | 5.685 |
| MonthlyBinary.Jan | 328.47564 | 1.939 |
| MonthlyBinary.Feb | 483.49800 | 2.883 |
| MonthlyBinary.Mar | -27.88954 | -0.163 |
| MonthlyBinary.May | 43.76640 | 0.256 |
| MonthlyBinary.Jun | 322.27457 | 1.718 |
| MonthlyBinary.Jul | 395.71891 | 1.987 |
| MonthlyBinary.Aug | 897.24985 | 4.56 |
| MonthlyBinary.Sep | 294.72798 | 1.621 |
| MonthlyBinary.Oct | 297.89237 | 1.659 |
| MonthlyBinary.Nov | -121.45020 | -0.648 |
| MonthlyBinary.Dec | 229.96350 | 1.15 |
| DOWBinary.Monday | 3219.42390 | 71.04 |
| DOWBinary.Tuesday | 3606.86919 | 66.402 |
| DOWBinary.Wednesday | 3693.74723 | 63.178 |
| DOWBinary.Thursday | 3707.34409 | 63.008 |
| DOWBinary.Friday | 3380.34001 | 61.311 |

| | | |
|---|---|---|
| DOWBinary.Saturday | 1105.57628 | 25.05 |
| US_Holidays.NYHol | -2648.81910 | -13.93 |
| US_Holidays.MLKing | -727.69327 | -3.643 |
| US_Holidays.PresidentDay | -448.70797 | -2.429 |
| US_Holidays.MemorialDay | -3133.79174 | -12.1 |
| US_Holidays.July4thHol | -2900.47786 | -10.04 |
| US_Holidays.LaborDay | -2833.05410 | -10.94 |
| US_Holidays.Thanksgiving | -3845.45498 | -12.83 |
| US_Holidays.FriAftThanks | -2932.75898 | -8.732 |
| US_Holidays.SatAftThanks | -828.74718 | -2.765 |
| US_Holidays.XMasHol | -3117.82976 | -10.97 |
| US_Holidays.XMASAft | -1845.91375 | -7.268 |
| US_Holidays.July4thMonFri | -2236.14065 | -7.733 |
| AR(1) | 0.61467 | 15.256 |

**SMALL HEATING MODEL**

Model fit statistics

- R-Squared 0.937
- Adjusted R-Squared 0.935
- Mean Abs. Dev. (MAD) 3.44
- Mean Abs. % Err. (MAPE) 3.75%
- Durbin-Watson Statistic 1.866

Variable Statistics

| Variable | Coefficient | T-Stat |
|---|---|---|
| CONST | 70.46335 | 31.605 |
| DailyAverageTemperature.HDD40 | 1.00446 | 8.986 |
| DailyAverageTemperature.HDD50 | 0.85452 | 11.455 |
| DailyAverageTemperature.CDD55 | 0.43562 | 4.514 |

| | | |
|---|---|---|
| DailyAverageTemperature.CDD65 | 1.01158 | 5.665 |
| DailyAverageTemperature.CDD75 | 0.23868 | 1.181 |
| MonthlyBinary.Jan | 6.10703 | 3.306 |
| MonthlyBinary.Feb | 5.3342 | 2.876 |
| MonthlyBinary.Mar | 1.47144 | 0.837 |
| MonthlyBinary.May | 1.13053 | 0.641 |
| MonthlyBinary.Jun | 4.74225 | 2.391 |
| MonthlyBinary.Jul | 5.54922 | 2.679 |
| MonthlyBinary.Aug | 7.22274 | 3.486 |
| MonthlyBinary.Sep | 2.92826 | 1.56 |
| MonthlyBinary.Oct | 0.07514 | 0.042 |
| MonthlyBinary.Nov | 0.91909 | 0.498 |
| MonthlyBinary.Dec | 9.23013 | 4.881 |
| DOWBinary.Monday | 17.99493 | 33.963 |
| DOWBinary.Tuesday | 18.09031 | 28.611 |
| DOWBinary.Wednesday | 18.66931 | 27.635 |
| DOWBinary.Thursday | 18.06553 | 26.482 |
| DOWBinary.Friday | 17.72768 | 27.768 |
| DOWBinary.Saturday | 8.66486 | 16.627 |
| US_Holidays.NYHol | -14.33675 | -4.768 |
| US_Holidays.MLKing | -1.53982 | -0.519 |
| US_Holidays.PresidentDay | -4.65574 | -1.905 |
| US_Holidays.July4thHol | -17.08476 | -5.796 |
| US_Holidays.MemorialDay | -14.08384 | -4.754 |
| US_Holidays.LaborDay | -17.68261 | -5.949 |
| US_Holidays.Thanksgiving | -23.50372 | -5.281 |
| US_Holidays.FriAftThanks | -4.76621 | -1.355 |
| US_Holidays.XMasHol | -10.44897 | -3.524 |
| MonthlyBinary.Year2005 | -10.0629 | -5.683 |
| MonthlyBinary.Year2006 | -10.31014 | -5.863 |
| AR(1) | 0.55519 | 18.591 |

**TOTAL ELECTRIC MODEL**

Model fit statistics

- R-Squared 0.938
- Adjusted R-Squared 0.936
- Mean Abs. Dev. (MAD) 37.91
- Mean Abs. % Err. (MAPE) 3.16%
- Durbin-Watson Statistic 1.914

Variable Statistics

| Variable | Coefficient | T-Stat |
|---|---|---|
| CONST | 889.58488 | 38.711 |
| DailyAverageTemperature.HDD55 | 5.84107 | 9.548 |
| DailyAverageTemperature.HDD45 | 10.53311 | 13.313 |
| DailyAverageTemperature.CDD60 | 8.38791 | 6.193 |
| DailyAverageTemperature.CDD65 | 8.69391 | 4.547 |
| DailyAverageTemperature.CDD75 | 2.14015 | 1.299 |
| MonthlyBinary.Jan | 62.17141 | 3.167 |
| MonthlyBinary.Feb | 58.69167 | 3.602 |
| MonthlyBinary.Mar | 13.47397 | 0.901 |
| MonthlyBinary.May | 36.89043 | 2.475 |
| MonthlyBinary.Jun | 79.04523 | 4.803 |
| MonthlyBinary.Jul | 131.03695 | 7.678 |
| MonthlyBinary.Aug | 122.29715 | 7.147 |
| MonthlyBinary.Sep | 88.57696 | 5.482 |
| MonthlyBinary.Oct | 56.44813 | 3.361 |
| MonthlyBinary.Nov | 81.01401 | 3.945 |
| MonthlyBinary.Dec | 93.87024 | 3.633 |
| DOWBinary.Monday | 156.31077 | 40.43 |
| DOWBinary.Tuesday | 162.88448 | 34.694 |
| DOWBinary.Wednesday | 181.88871 | 35.824 |

| | | |
|---|---|---|
| DOWBinary.Thursday | 177.22321 | 34.76 |
| DOWBinary.Friday | 188.44942 | 39.752 |
| DOWBinary.Saturday | 70.40244 | 18.664 |
| SunTimes.FracDark17 | 114.81395 | 1.576 |
| SunTimes.FracDark8 | -124.38005 | -2.737 |
| US_Holidays.NYHol | -136.16071 | -6.511 |
| US_Holidays.MLKing | -6.16519 | -0.297 |
| US_Holidays.PresidentDay | -22.20004 | -1.153 |
| US_Holidays.MemorialDay | -128.21783 | -5.931 |
| US_Holidays.July4thHol | -126.34881 | -5.935 |
| US_Holidays.LaborDay | -168.14571 | -7.793 |
| US_Holidays.Thanksgiving | -220.59886 | -9.116 |
| US_Holidays.FriAftThanks | -30.98892 | -1.279 |
| US_Holidays.XMasHol | -102.76309 | -4.814 |
| MonthlyBinary.Year2006 | -45.39835 | -2.226 |
| MonthlyBinary.Year2005 | -10.08161 | -0.488 |
| MonthlyBinary.Year2004 | -33.27005 | -1.62 |
| MonthlyBinary.Year2003 | 0.43624 | 0.021 |
| AR(1) | 0.63663 | 32.2 |

# Weather Normalization Method For Empire District Electric Company

Itron, Inc.
11236 El Camino Real
San Diego, California 92130
(858) 724-2620

October, 2009

# Table of Contents

# 1

# Project Summary

In 2007, the Empire District Electric Company (Empire) engaged Itron's forecast consulting services to develop a weather normalized forecast for July 1, 2006 to June 30, 2007. The weather normalized forecast was developed for the following five Empire classes.

- Residential (Res)
- Commercial (Com)
- Small Heating (SH)
- General Power (GP)
- Total Electric (TEB)

The weather normalization method and forecast was submitted to the Missouri Public Service Commission in 2007.

In 2009, Empire engaged Itron's forecast consulting services to update the weather normalization forecast for July 1, 2008 through June 30, 2009 using the same method as in 2007. This report summarizes the method developed in 2007 and modified for the 2009 project. The weather normalization process employed by Itron uses load research data provided by Empire and is described in Section 2. This method includes the development of daily statistical models (Section 3) and daily normal weather (Section 4).

# 2

# Normalization Method

Weather normalization is the process of mathematically adjusting actual energy sales so that it represents energy typically used under a normal year condition. This process accounts for weather differences from between actual conditions and normal conditions.

Because the process is mathematical, two key assumptions are necessary to account for the differences between actual and normal sales. First, energy consumption is modeled based on historical relationships between actual consumption and historical weather. The model incorporates a set of descriptive variables to capture a statistical correlation between the variables and consumption. Second, normal conditions are assumed based on historical weather data. In this section, Itron describes the steps used to normalize historical sales based on the models and the normal weather developed by Itron in Sections 3 and 4. This method was employed in 2007.

***Step 1. Daily Sales Models.*** In this step, Itron developed five regression models to capture the relationship between actual consumption and historical weather. The regression models were developed for the following classes.

- Residential (Res)
- Commercial (Com)
- General Power (GP)
- Small Heating (SH)
- Total Electric (TEB)

The models utilize Empire's Load Research data to articulate the models in Section 3.

***Step 2. Simulate Daily Sales With Actual Weather.*** In this step, Itron used the five regression models developed in Step 1 to forecast the historical daily sales using actual weather. This step results in the model prediction of sales under actual weather conditions.

***Step 3. Simulate Daily Sales With Normal Weather.*** In this step, Itron used the five regression models developed in Step 1 to forecast the historical daily sales using normal weather. This step results in the model prediction of sales under normal weather conditions.

***Step 4. Calculate the Normal Revenue Cycle Month Sales.*** In this step, Itron adjusts the historical monthly revenue cycle sales provided by Empire for normal weather conditions. The result of this step is a monthly series of revenue cycle sales under normal conditions.

To calculate the normal revenue cycle sales, the following steps were taken.

1. Calculate the model sales with actual weather over the revenue cycle (***Model Actual Sales***). This step estimates the model predicted monthly revenue sales with actual weather.
2. Calculate the model sales with normal weather over the revenue cycle (***Model Normal Sales***). This step estimates the model predicted monthly revenue sales with normal weather.
3. Calculate the ***Normal Revenue Cycle Sales*** by adjusting the actual revenue sales over the revenue cycle (***Actual Revenue Cycle Sales***) using the ratio of the (1) and (2)

$$NormalRevenueCycleSales_{month} = \frac{ModelNormalSales_{month}}{ModelActualSales_{month}} \times Actual\,RevenueCycleSales_{month}$$

In calculating ***Normal Revenue Cycle Sales***, ***Model Actual Sales***, and ***Model Normal Sales*** are summed over the historic **billing cycle month** provided by Empire. Because the meter read schedule does not contain fixed read dates, the "**Last Read Date**" is used to define the meter read schedule for the purposes of calculating the ***Normal Revenue Cycle Sales***.

In this approach, the use of the ratio of ***Model Actual Sales*** to ***Model Normal Sales*** removes the model bias from the normal calculation and directly adjusts the ***Actual Revenue Cycle Sales*** using normalization models developed with load research data.

***Step 5. Calculate the Normal Calendar Month Sales.*** In this step, Itron uses the same adjustment in Step 4 to adjust the ***Actual Revenue Cycle Sales*** to calendar month sales. The calculation is identical except that the ***Model Normal Sales*** is summed over the **calendar month** instead of the billing cycle month. This approach embeds into the ***Model Actual Sales*** (summed over the revenue month) and ***Model Normal Sales*** (summed over the calendar month) ratio the adjustment from revenue cycle sales to calendar month sales.

The final products of the weather normalization method are monthly normal sales based on both billing (revenue) cycles and calendar months.

# 3

# Models

The energy consumption models capture the load response to weather and other conditions. In developing these models, historical load research data were examined and used to estimate linear regression models using daily data. This section discusses the regression models.

## 3.1 Residential Model

The Residential Daily Sales model was developed to articulate the relationship between the Residential class consumption and actual weather patterns. Hourly load research data (load research means) were provided by Empire from January 1, 1995 through February 28, 2007. These hourly data are shown in Figure 1. The hourly data aggregated to daily energy data are shown in Figure 2. Upon inspecting these data, data from January 2002 through February 2007 are used in the residential model.

**Figure 1: Residential Hourly Load Research Data**

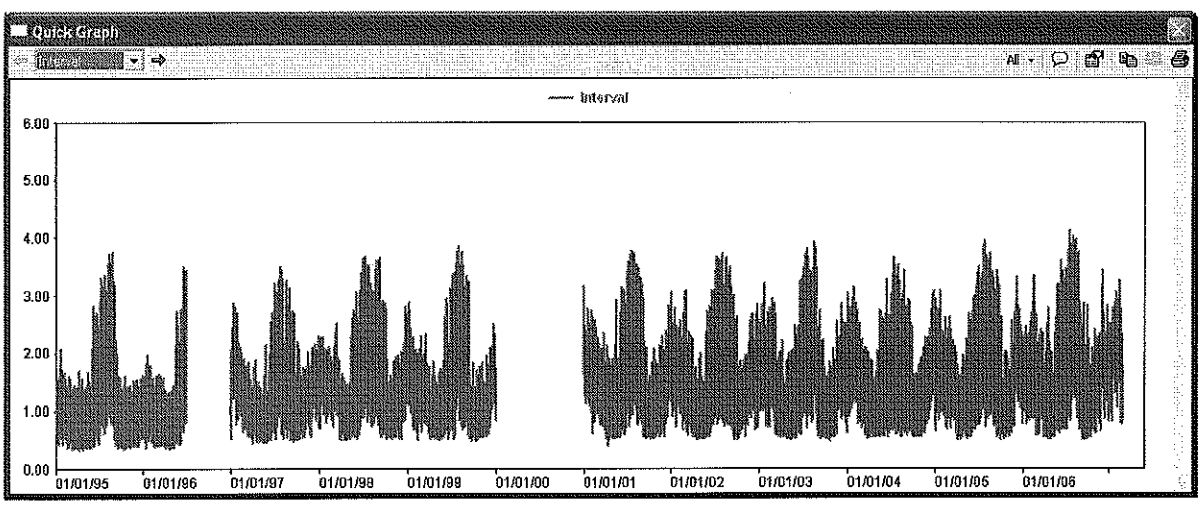
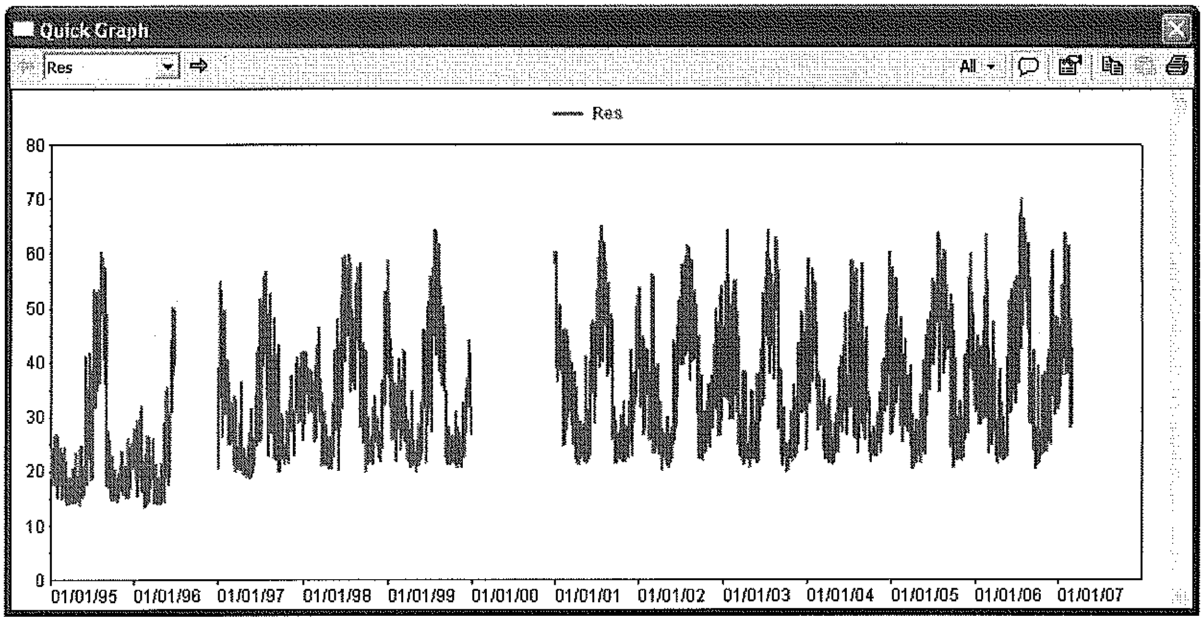
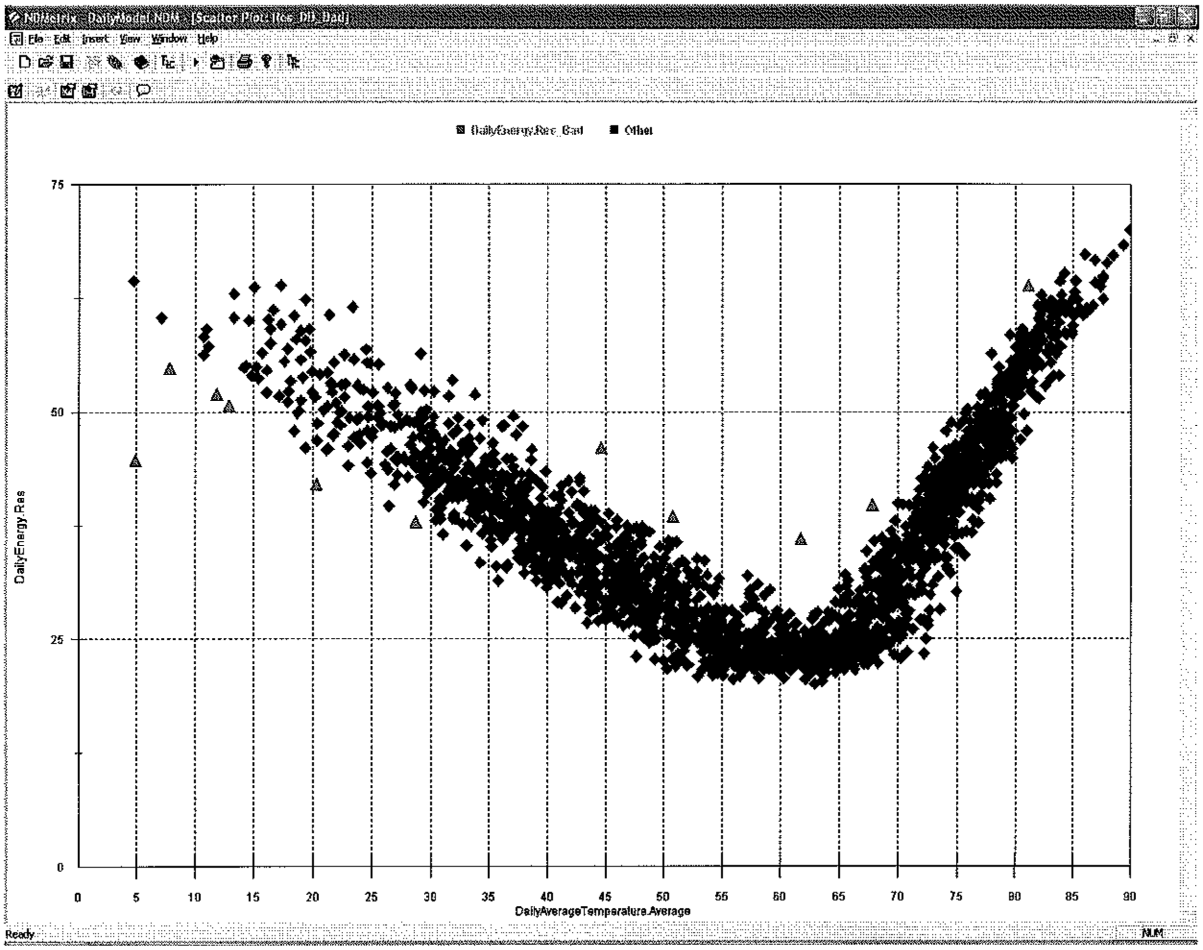


**Figure 2: Residential Daily Energy**

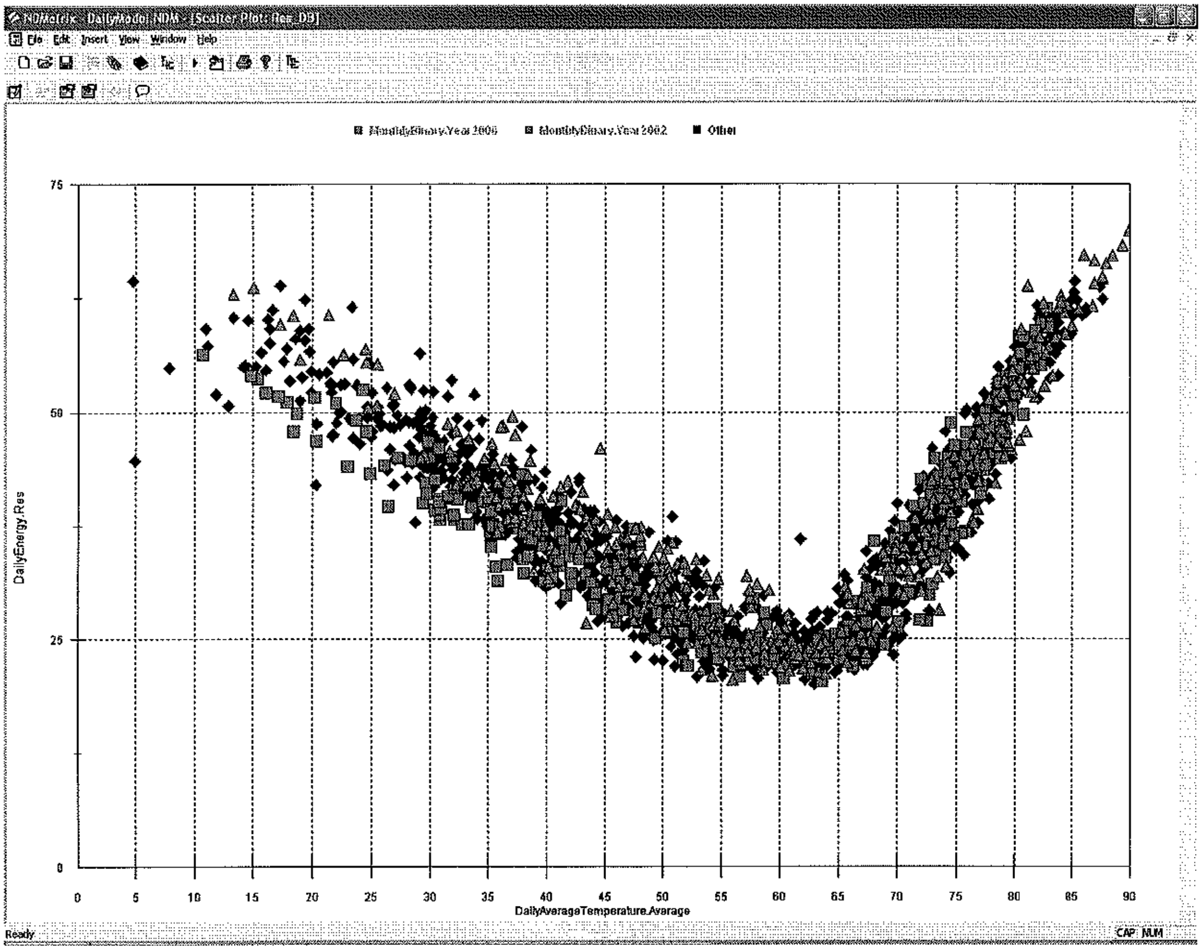



The load-weather relationship is best viewed using the scatter plots shown in Figure 3 and Figure 4. In these figures, daily energy is shown in the Y-axis and daily average temperature is shown on the X-axis. These figures demonstrate the non-linear load response to actual weather. Two main observations are seen in these figures. In Figure 3, data outside the general load-weather relationship are show in red triangles. These data points are removed from model estimation. In Figure 4, the heating response is seen as changing between 2002 (brown squares) and 2006 (green triangles). The model is constructed to account for this changing heating response.

## Figure 3: Residential Bad Data (Red Triangles)



## Figure 4: Residential Energy Temperature Scatter Plot




***Residential Model.*** A linear regression model is used to articulate the load-weather relationship. This model contains the following classes of variables and their function in the model context (Table 1). A full description of the model can be viewed in the *MetrixND* project file.

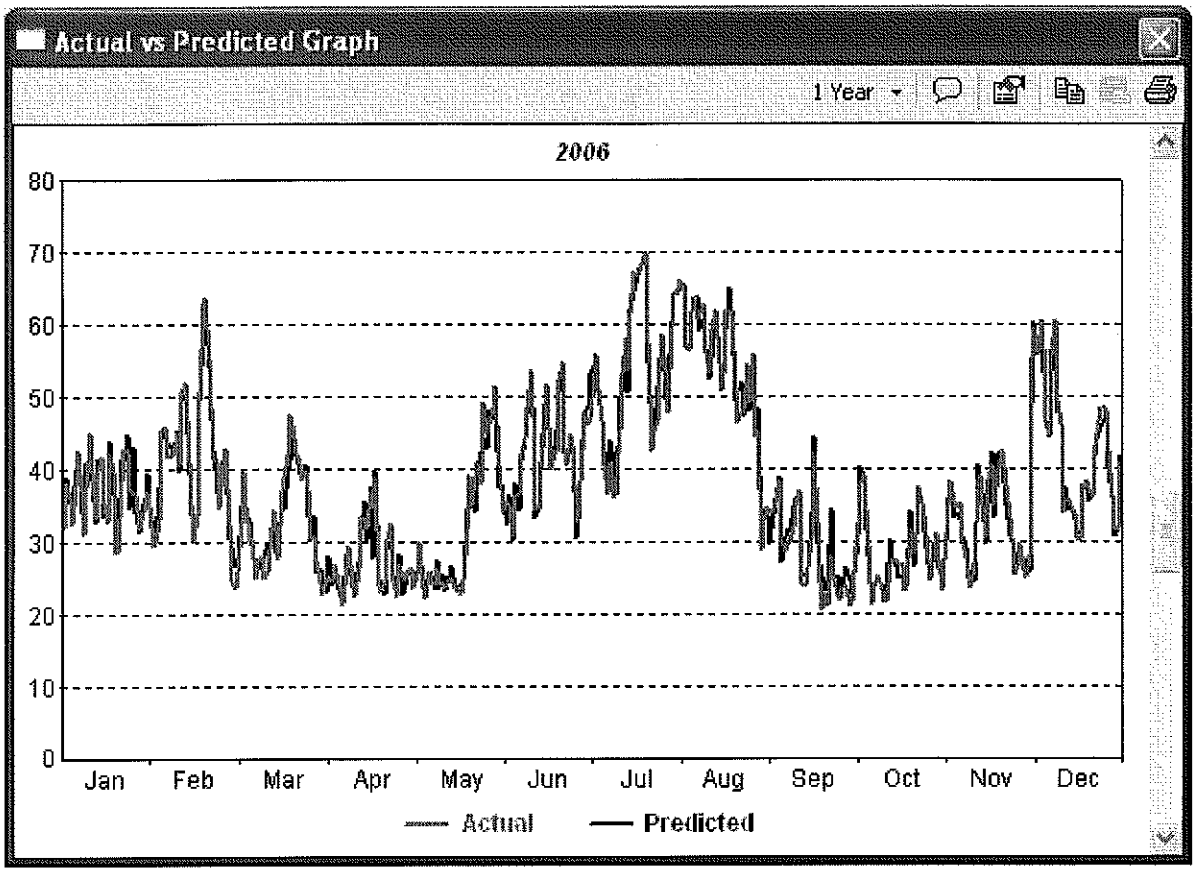
## Table 1: Residential Model Variables

| Variable Class | Purpose |
|---|---|
| Monthly Binaries | These variables account for changing seasonal consumption pattern for year. |
| Day of Week Binaries | These variables account for changing consumption pattern for each day of the week. |
| Sunlight | These variables account for the changing time of sunrise and sunset. |
| Holidays | These variables account for changes in consumption as a result of national holidays. |
| Annual Binaries | These variables account for changes in the load research samples and load growth over the estimation period. |
| Temperature Splines | These variables account for the nonlinear load response to weather and the changing heating response. |
| AR Term | This term removes the remaining serial correlation and clarifies the remaining model coefficients. |

The overall fit of the regression model can be seen graphically in Figure 5 and numerically in the statistics below.

- R-Squared 0.964
- Adjusted R-Squared 0.963
- Mean Abs. Dev. (MAD) 1.60
- Mean Abs. % Err. (MAPE) 4.63%
- Durbin-Watson Statistic 2.073

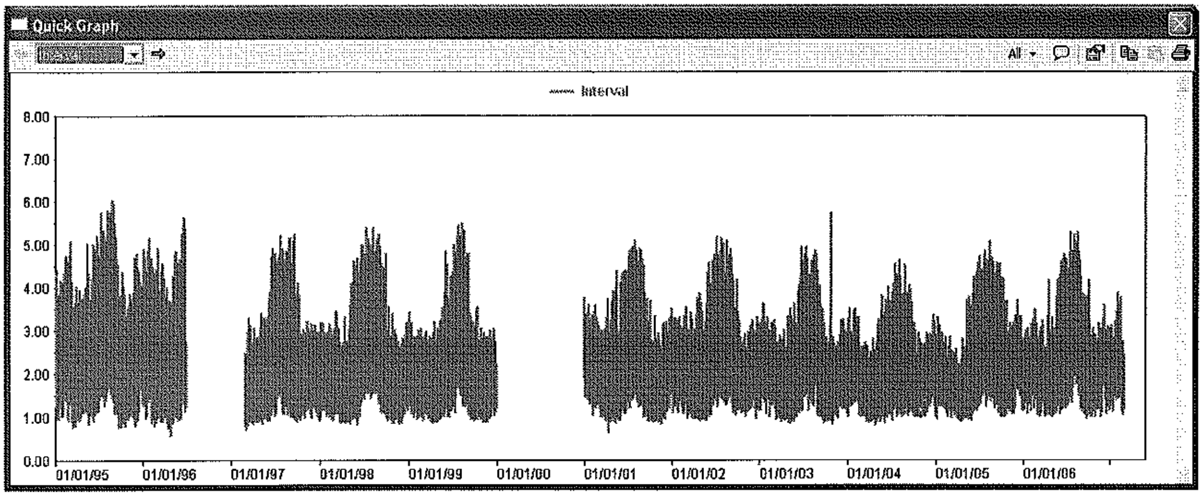
**Figure 5: Residential Model Fit – Actual Versus Predicted Values**



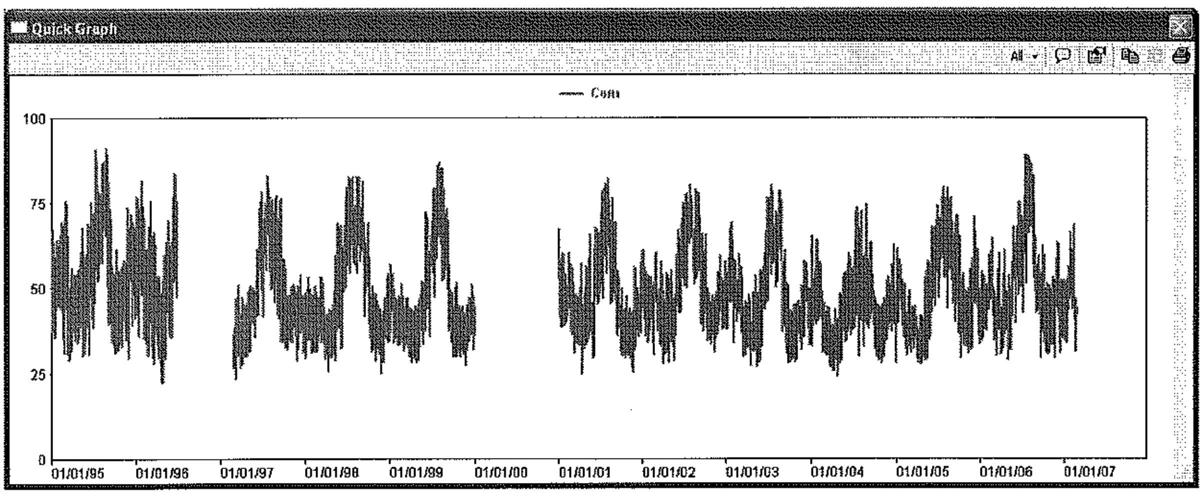

## 3.2 Commercial

The Commercial Daily Sales model was developed to articulate the relationship between the commercial class consumption and actual weather patterns. Hourly load research data (load research means) were provided by Empire from January 1, 1995 through February 28, 2007. These hourly data are shown in Figure 6. The hourly data aggregated to daily energy data are shown in Figure 7. Upon inspecting these data, data from January 2002 through February 2007 are used in the commercial model.

## Figure 6: Commercial Hourly Load Research Data

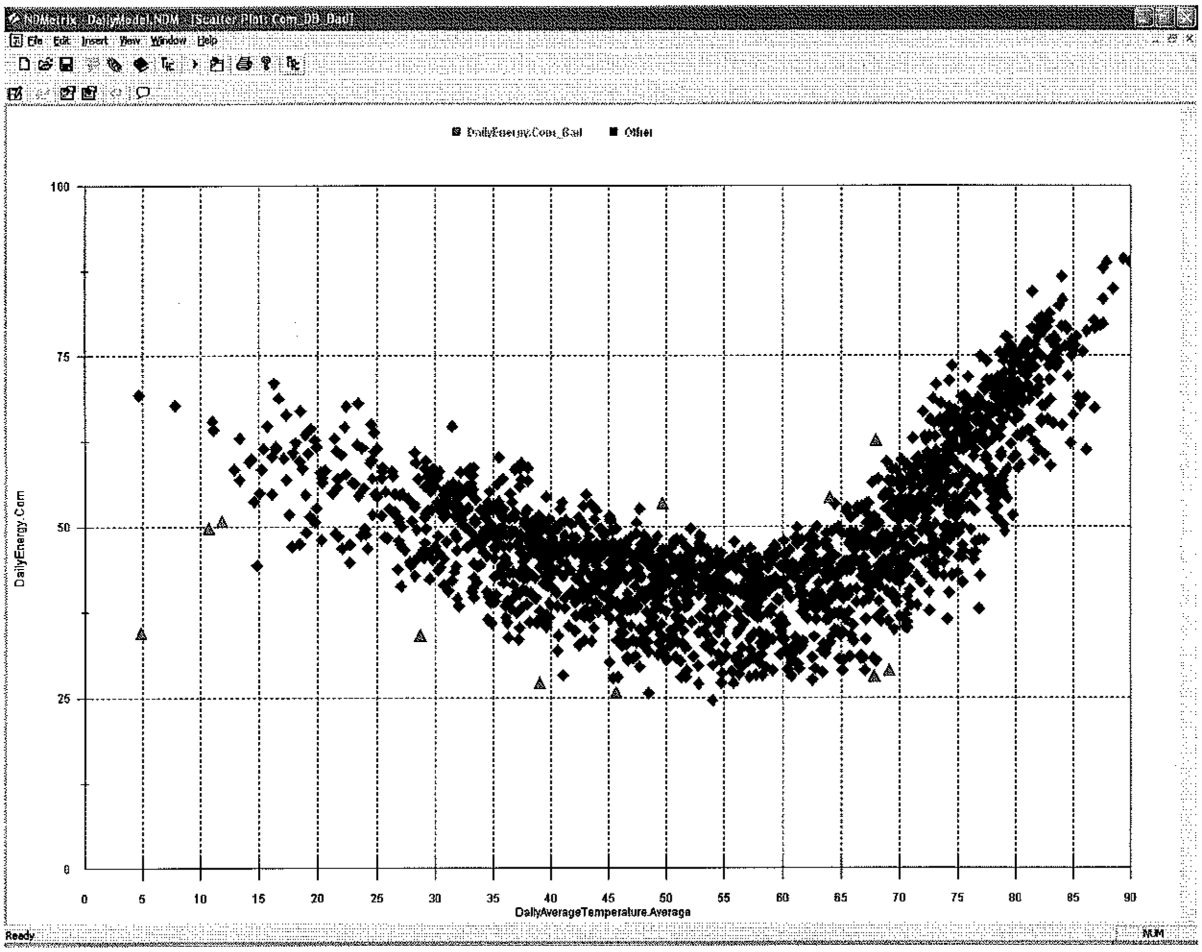



## Figure 7: Commercial Daily Energy

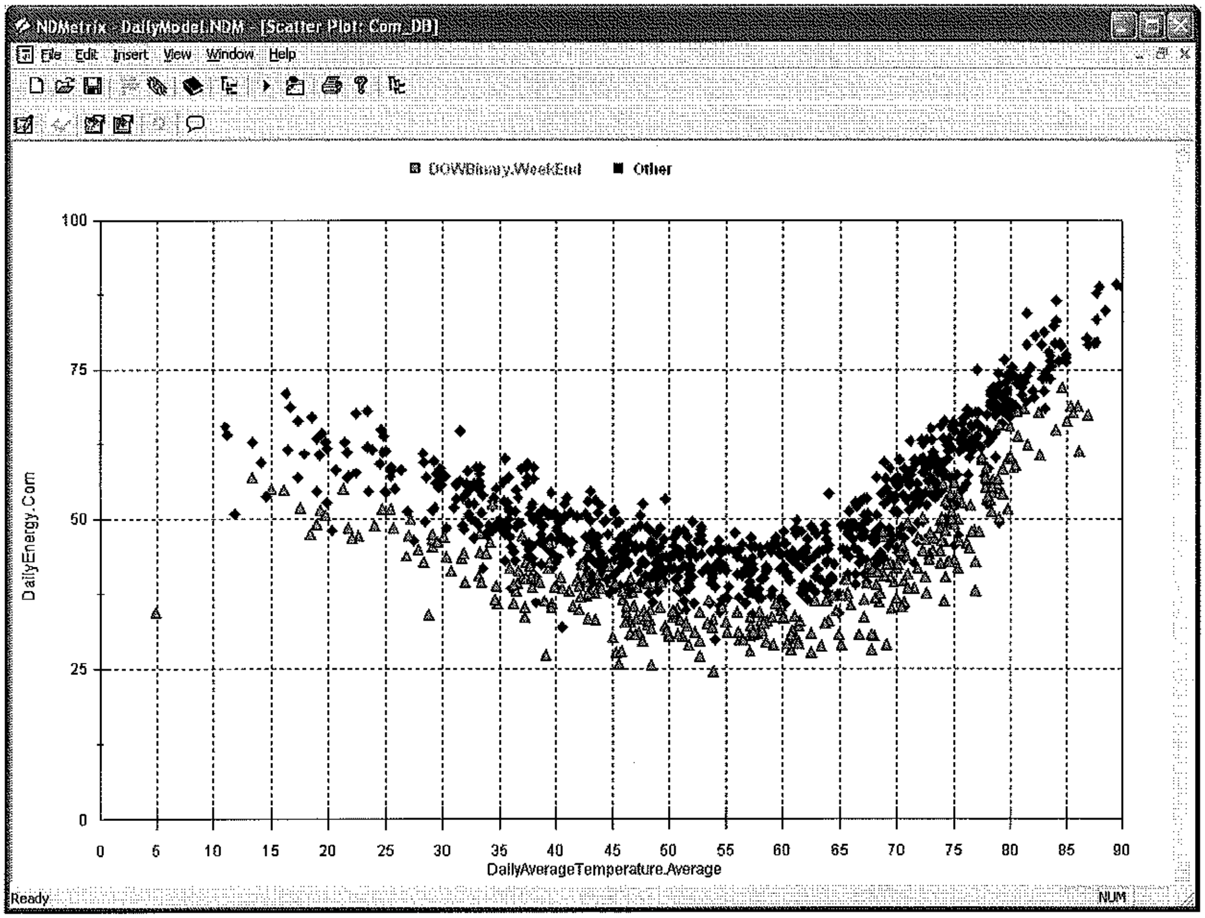



The load-weather relationship is best viewed using the scatter plots shown in Figure 8 and Figure 9. In these figures, daily energy is shown in the Y-axis and daily average temperature is shown on the X-axis. These figures demonstrate the non-linear load response to actual weather. Two main observations are seen in these figures. In Figure 8, data outside the general load-weather relationship are show in red triangles. These data points are removed from model estimation. In Figure 9, the weekend response (green triangles) is clearly lower than the weekday response (blue diamonds).

## Figure 8: Commercial Bad Data



## Figure 9: Commercial Energy Temperature Scatter Plot

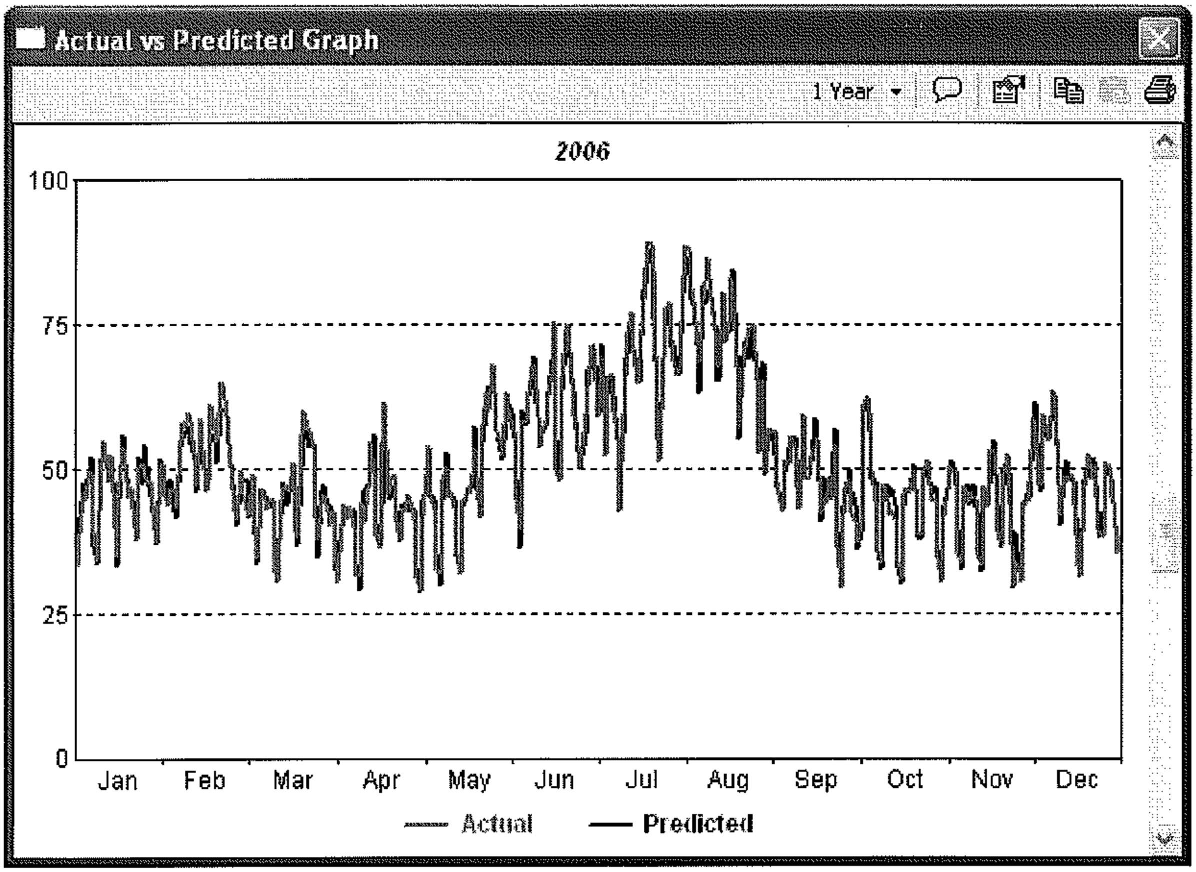



***Commercial Model.*** The commercial model is built with the same classes of variables used in the residential model (Table 1). However, temperature splines have been adjusted for the commercial weather response and no changing weather response is modeled.

The overall fit of the regression model can be seen graphically in Figure 10 and numerically in the statistics below. A full description of the model and the associated model statistics can be viewed in the *MetrixND* project file.

- R-Squared 0.958
- Adjusted R-Squared 0.957
- Mean Abs. Dev. (MAD) 1.88
- Mean Abs. % Err. (MAPE) 3.93%
- Durbin-Watson Statistic 2.072

**Figure 10: Commercial Model Fit – Actual Versus Predicted Values**




## 3.3 General Power

The General Power (GP) Daily Sales model was developed to articulate the relationship between the GP class consumption and actual weather patterns. Hourly load research data (load research means) were provided by Empire from January 1, 1995 through February 28, 2007. These hourly data are shown in Figure 11. The hourly data aggregated to daily energy data are shown in Figure 12. Upon inspecting these data, data from January 2006 through February 2007 are used. The shortened historical series accounts for the significant drop in consumption beginning in 2006.

**Figure 11: General Power Hourly Load Research**

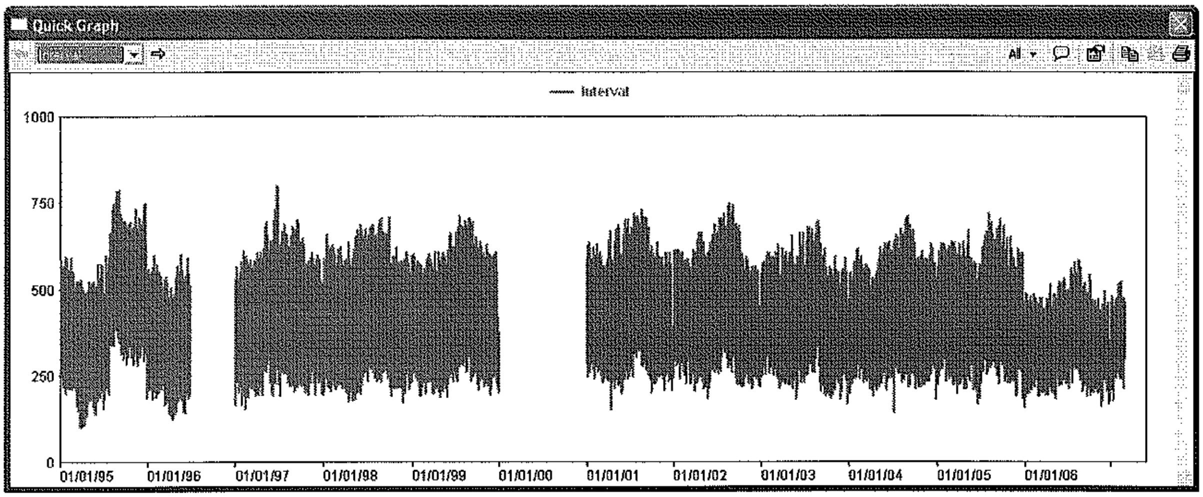
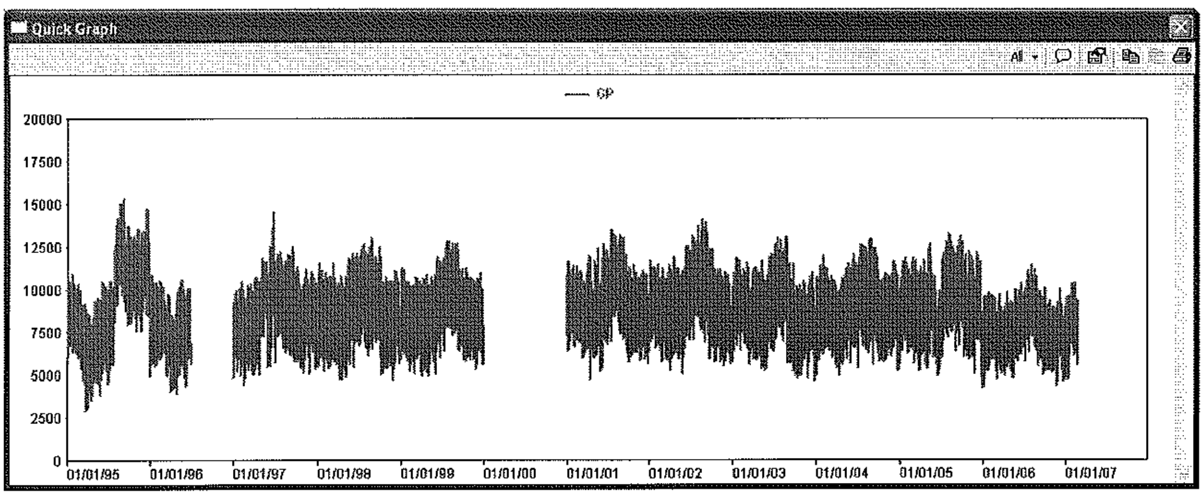
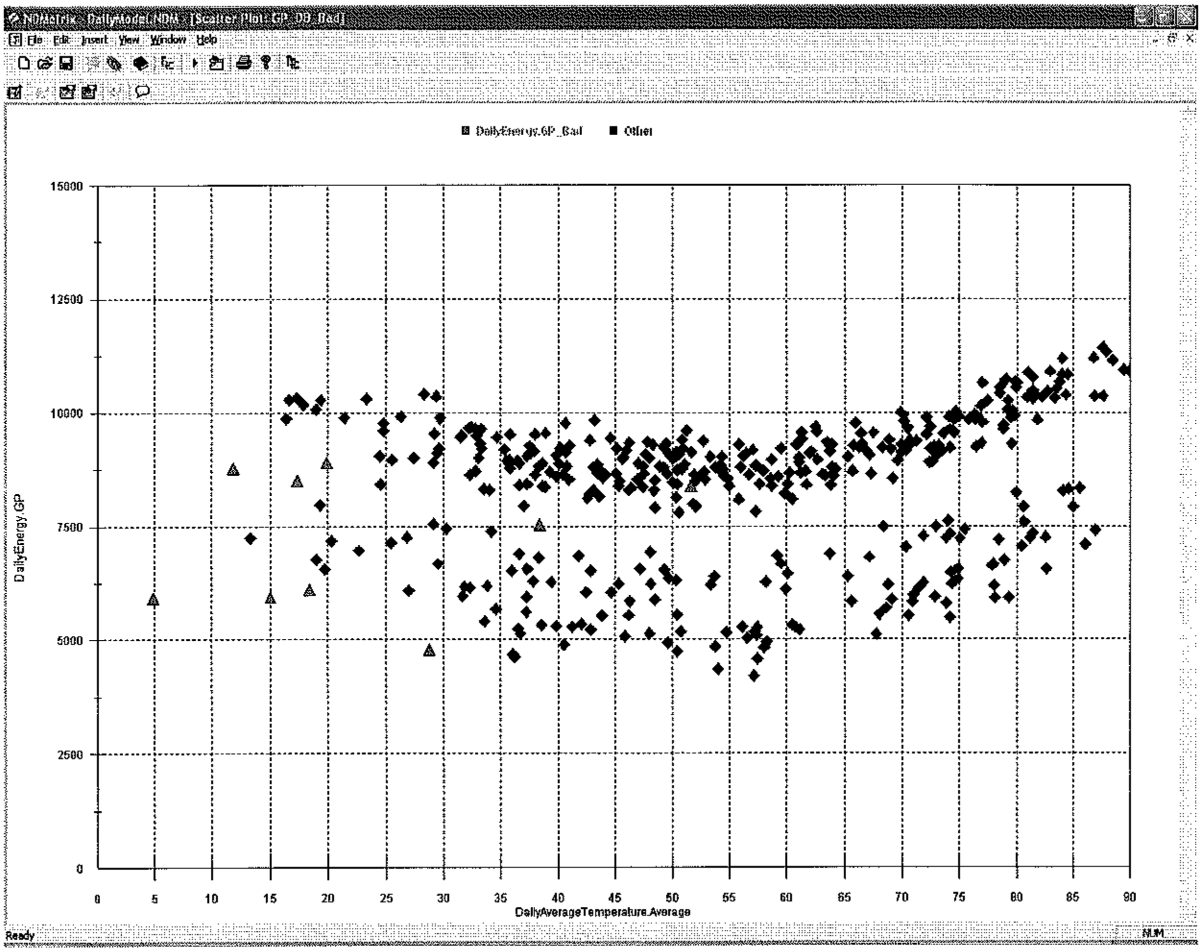



**Figure 12: General Power Daily Energy**

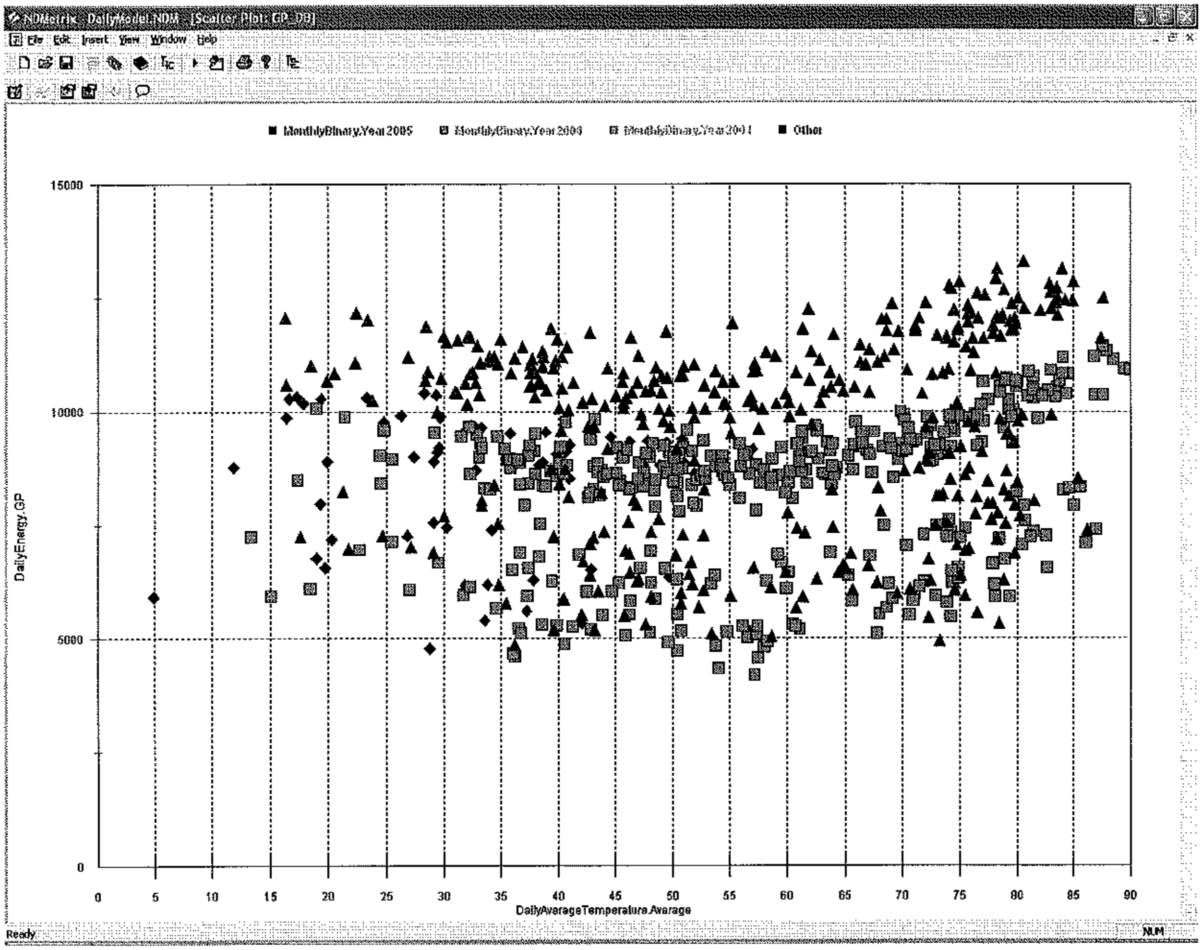



The load-weather relationship is best viewed using the scatter plots shown in Figure 13 and Figure 14. In these figures, daily energy is shown in the Y-axis and daily average temperature is shown on the X-axis. These figures demonstrate the non-linear load response to actual weather. Two main observations are seen in these figures. In Figure 13, data outside the general load-weather relationship are show in red triangles. These data points are removed from model estimation. In Figure 14, the 2005 data points (red triangles) and the 2006 data points (green squares) are highlighted. Based on visual inspection, the cooling response between 2005 and 2006 clearly changing further demonstrating the need to remove pre-2006 data.

## Figure 13: General Power Bad



## Figure 14: General Power Energy Temperature Scatter Plot

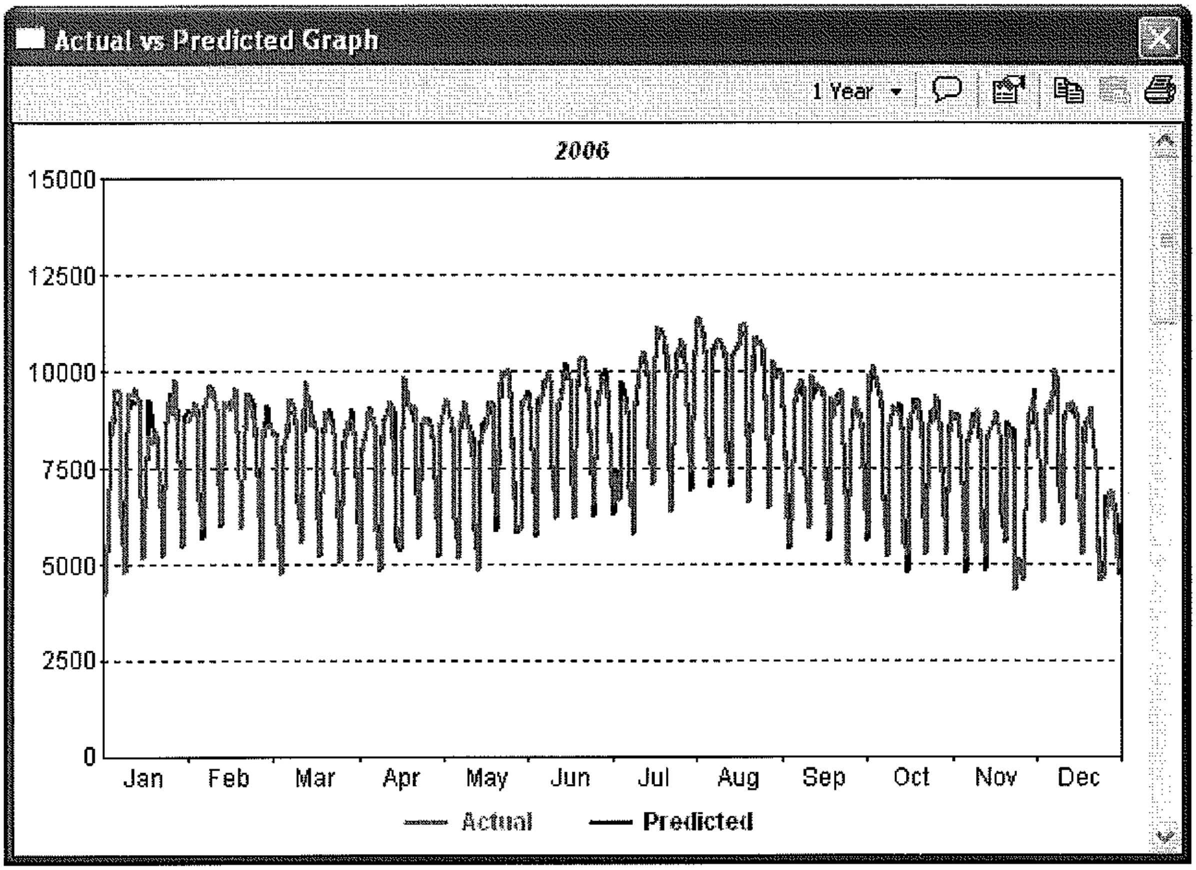



***GP Model.*** The GP model is built with the same classes of variables used in the residential model (Table 1). However, temperature splines have been adjusted for the GP weather response and no changing weather response is modeled.

The overall fit of the regression model can be seen graphically in Figure 15 and numerically in the statistics below. A full description of the model and the associated model statistics can be viewed in the *MetrixND* project file.

- R-Squared 0.968
- Adjusted R-Squared 0.965
- Mean Abs. Dev. (MAD) 215.16
- Mean Abs. % Err. (MAPE) 2.75%
- Durbin-Watson Statistic 2.076

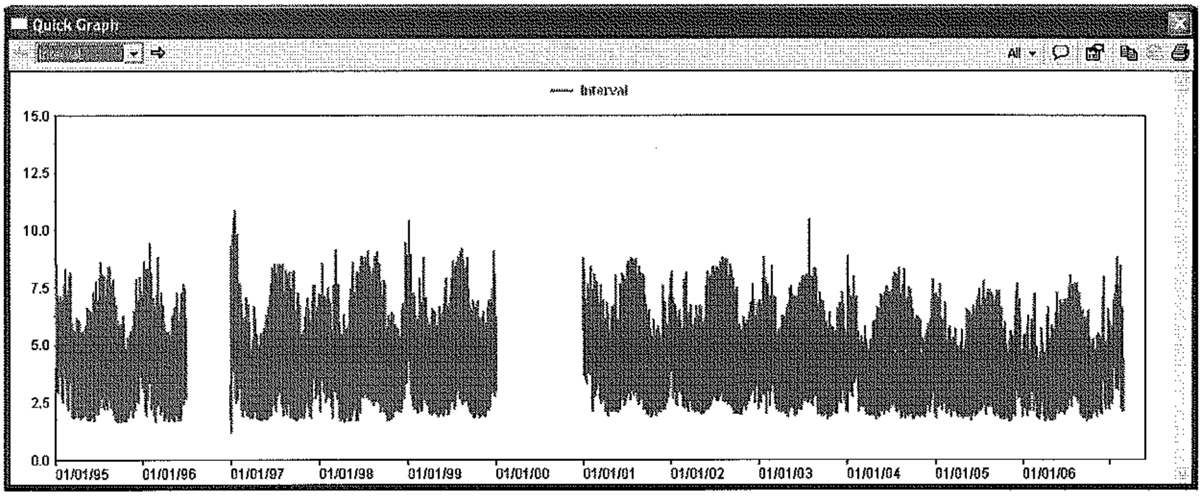
**Figure 15: GP Model Fit – Actual Versus Predicted Values**



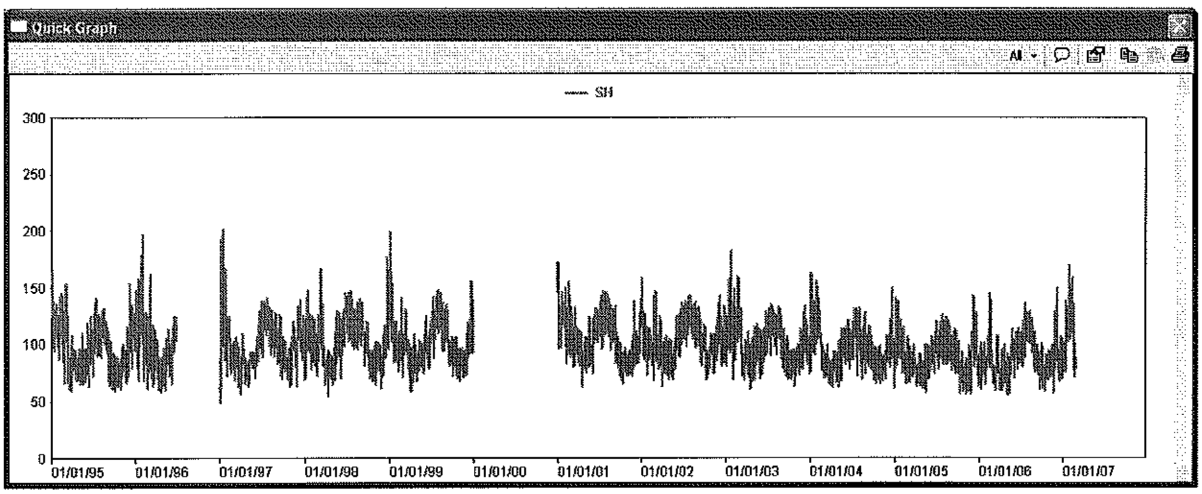

## 3.4 Small Heating

The Small Heating (SH) Daily Sales model was developed to articulate the relationship between the SH class consumption and actual weather patterns. Hourly load research data (load research means) were provided by Empire from January 1, 1995 through February 28, 2007. These hourly data are shown in Figure 16. The hourly data aggregated to daily energy data are shown in Figure 17. Upon inspecting these data, data from January 2005 through February 2007 are used. The shortened historical series removes the downward sloping trend that begins in 2001 and stabilizes in 2005.

## Figure 16: Small Heating Hourly Load Research




## Figure 17: Small Heating Daily Energy




The load-weather relationship is best viewed using the scatter plots shown in Figure 18 and Figure 19. In these figures, daily energy is shown in the Y-axis and daily average temperature is shown on the X-axis. These figures demonstrate the non-linear load response to actual weather. Two main observations are seen in these figures. In Figure 18, data outside the general load-weather relationship are show in red triangles. These data points are removed from model estimation. In Figure 19, the 2004 data points (purple triangles) clearly have a different temperature responses than 2005 (red squares) and 2006 (green circles). The different temperature response demonstrates the need to remove the pre-2005 data.

## Figure 18: Small Heating Bad

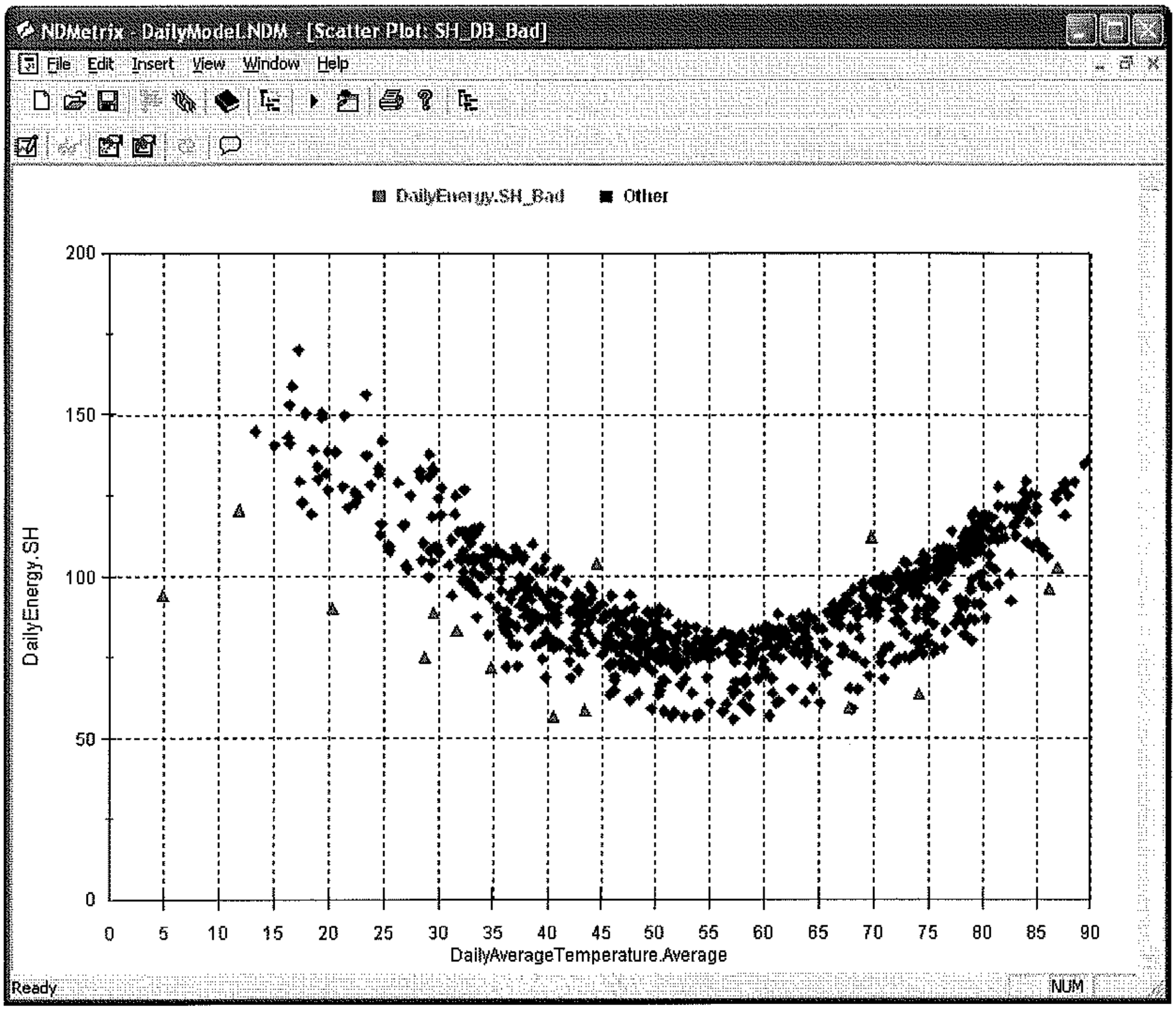
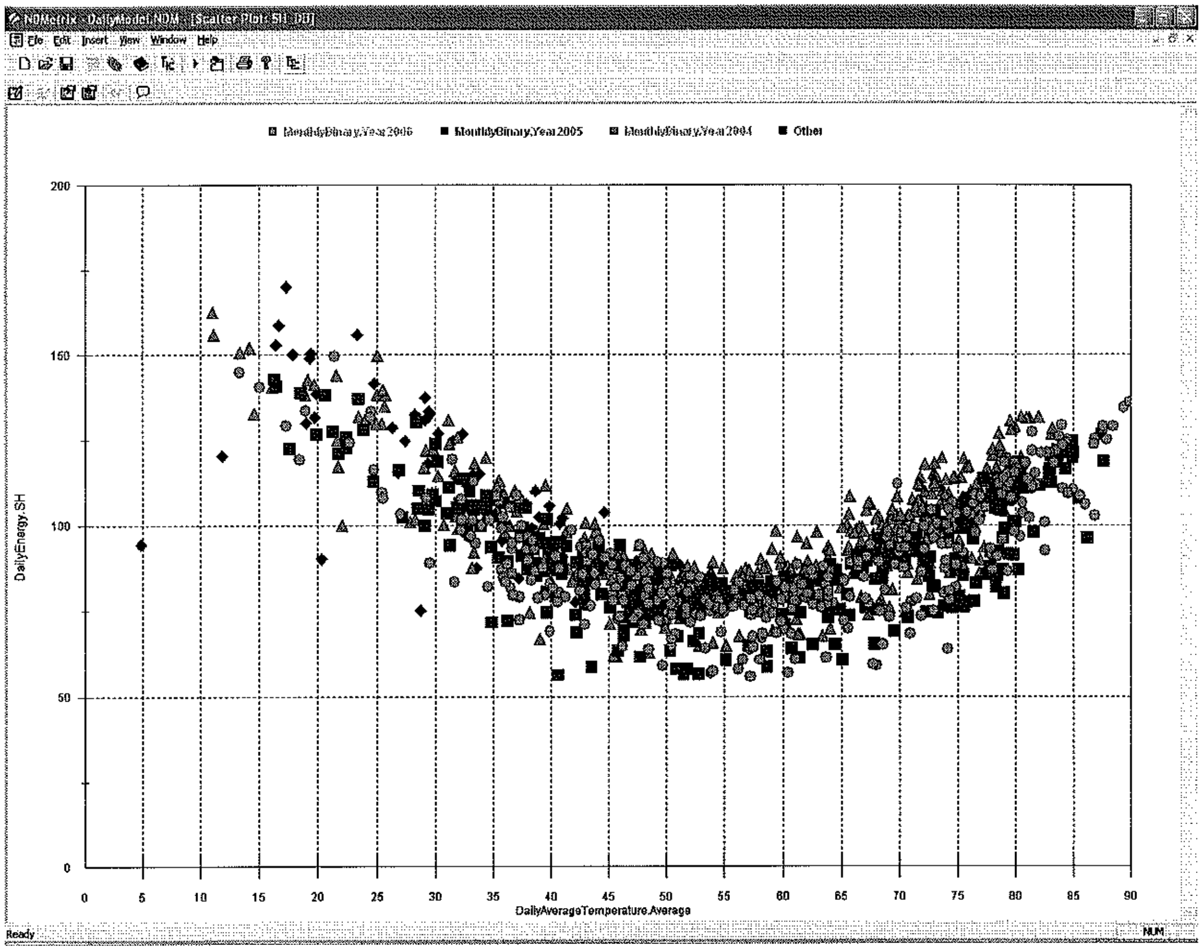


**Figure 19: Small Heating Energy Temperature Scatter Plot**

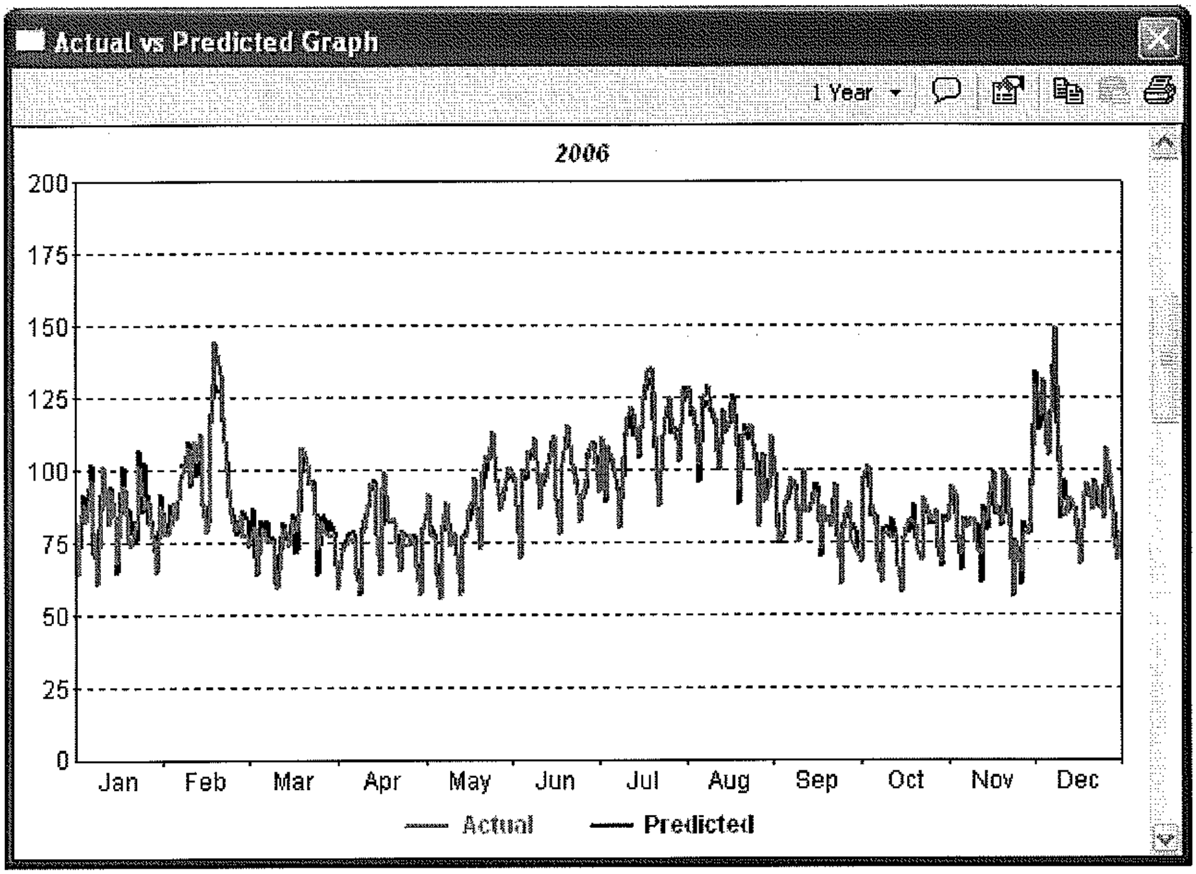



***SH Model.*** The SH model is built with the same classes of variables used in the residential model (Table 1). However, temperature splines have been adjusted for the SH weather response and no changing weather response is modeled.

The overall fit of the regression model can be seen graphically in Figure 20 and numerically in the statistics below. A full description of the model and the associated model statistics can be viewed in the *MetrixND* project file.

| | |
|---|---|
| R-Squared | 0.937 |
| Adjusted R-Squared | 0.935 |
| Mean Abs. Dev. (MAD) | 3.44 |
| Mean Abs. % Err. (MAPE) | 3.75% |
| Durbin-Watson Statistic | 1.866 |

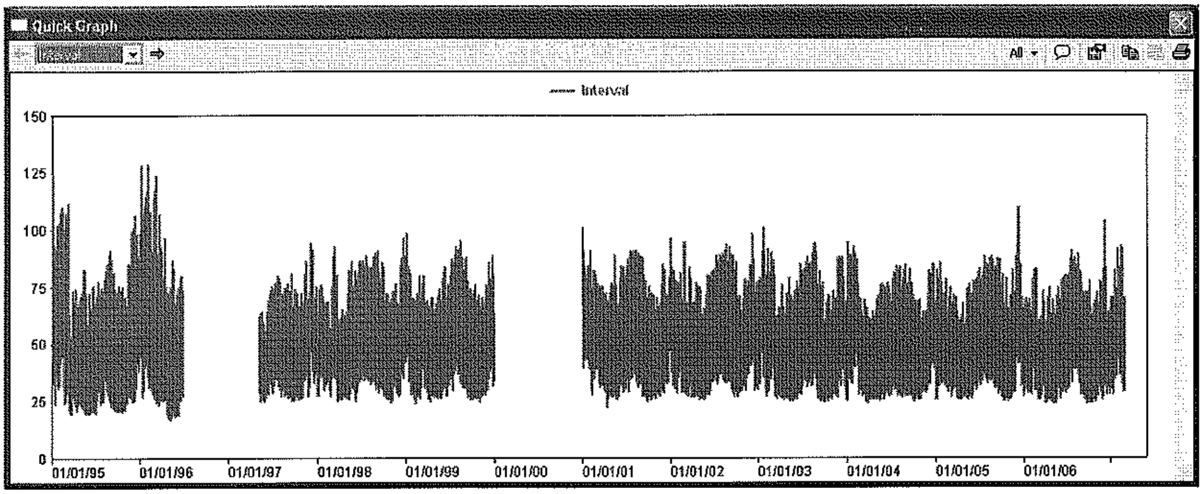
**Figure 20: SH Model Fit – Actual Versus Predicted Values**



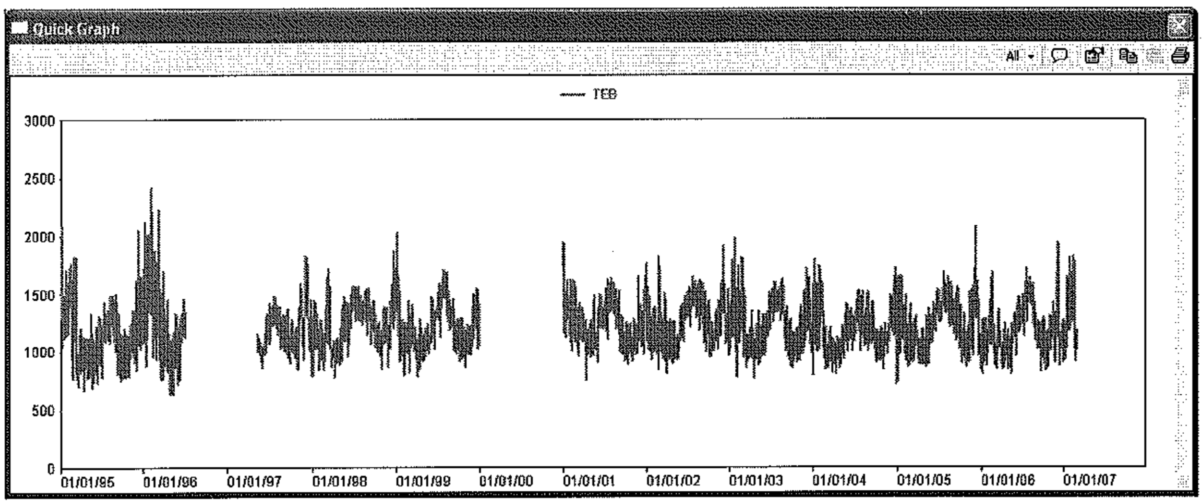

## 3.5 Total Electric

The Total Electric (TEB) Daily Sales model was developed to articulate the relationship between the TEB class consumption and actual weather patterns. Hourly load research data (load research means) were provided by Empire from January 1, 1995 through February 28, 2007. These hourly data are shown in Figure 21. The hourly data aggregated to daily energy data are shown in Figure 22. Upon inspecting these data, data from January 2003 through February 2007 are used. The shortened historical series captures the stable level of loads that appears after the beginning of 2003.

## Figure 21: Total Electric Hourly Load Research




## Figure 22: Total Electric Daily Energy




The load-weather relationship is best viewed using the scatter plots shown in Figure 23 and Figure 24. In these figures, daily energy is shown in the Y-axis and daily average temperature is shown on the X-axis. These figures demonstrate the non-linear load response to actual weather. Two main observations are seen in these figures. In Figure 23, data outside the general load-weather relationship are show in red triangles. These data points are removed from model estimation. In Figure 24, the 2002 data points (red triangles) are shown against the 2003 through 2007 data (blue diamonds). This view shows the 2002 data with a higher load and higher cooling weather response, which results in the data being excluded from the model.

## Figure 23: TEB Bad

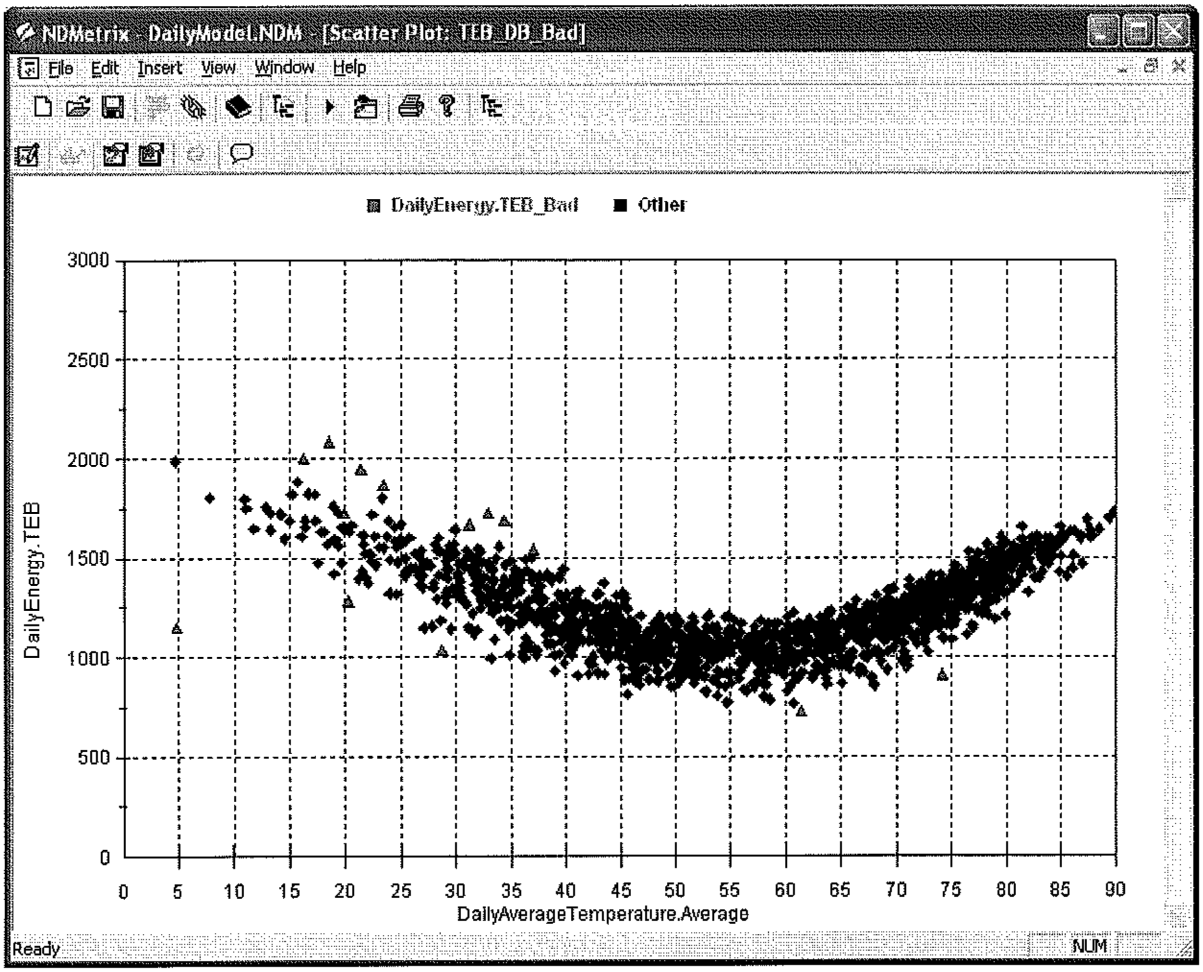
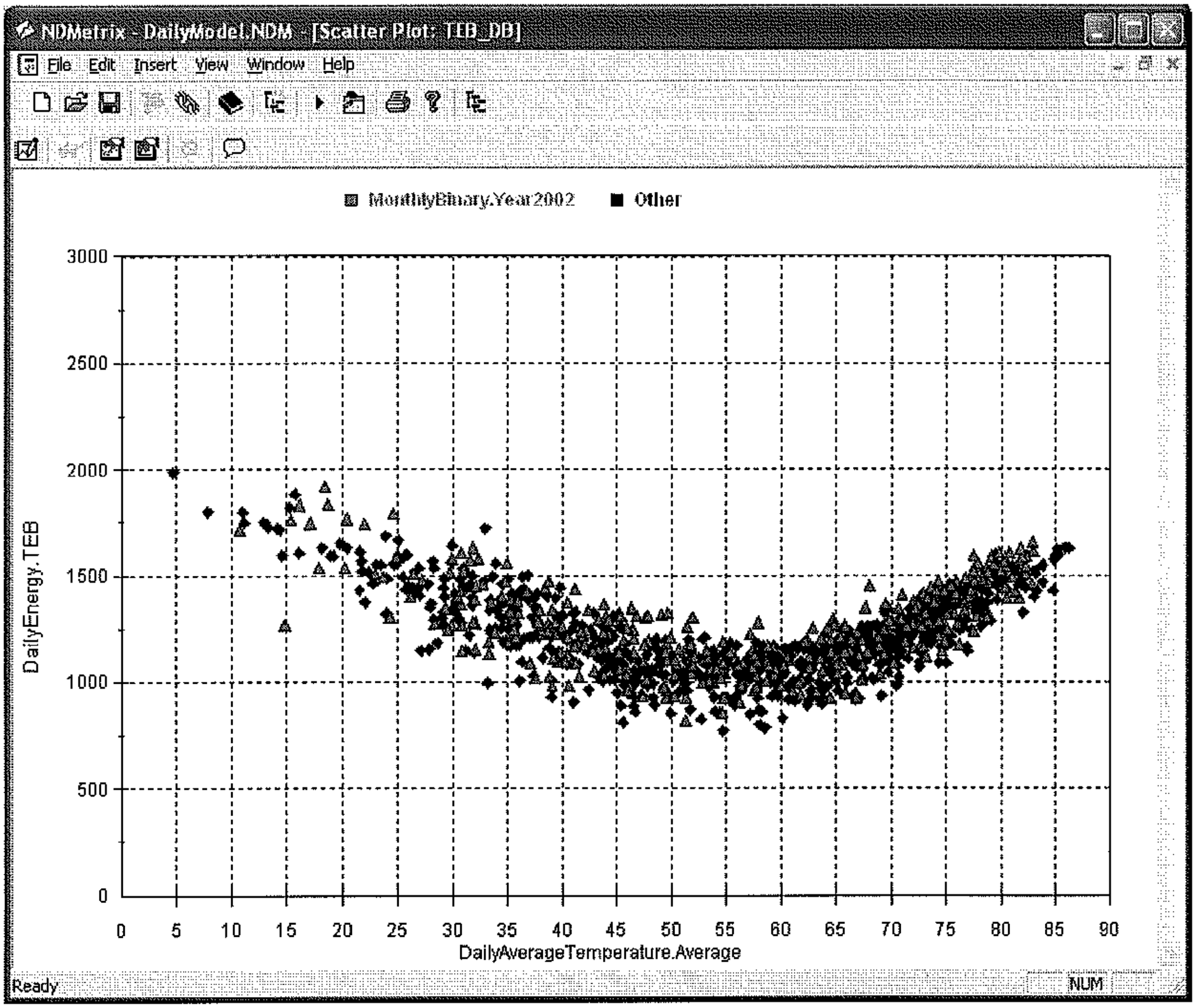


**Figure 24: TEB Energy Temperature Scatter Plot**

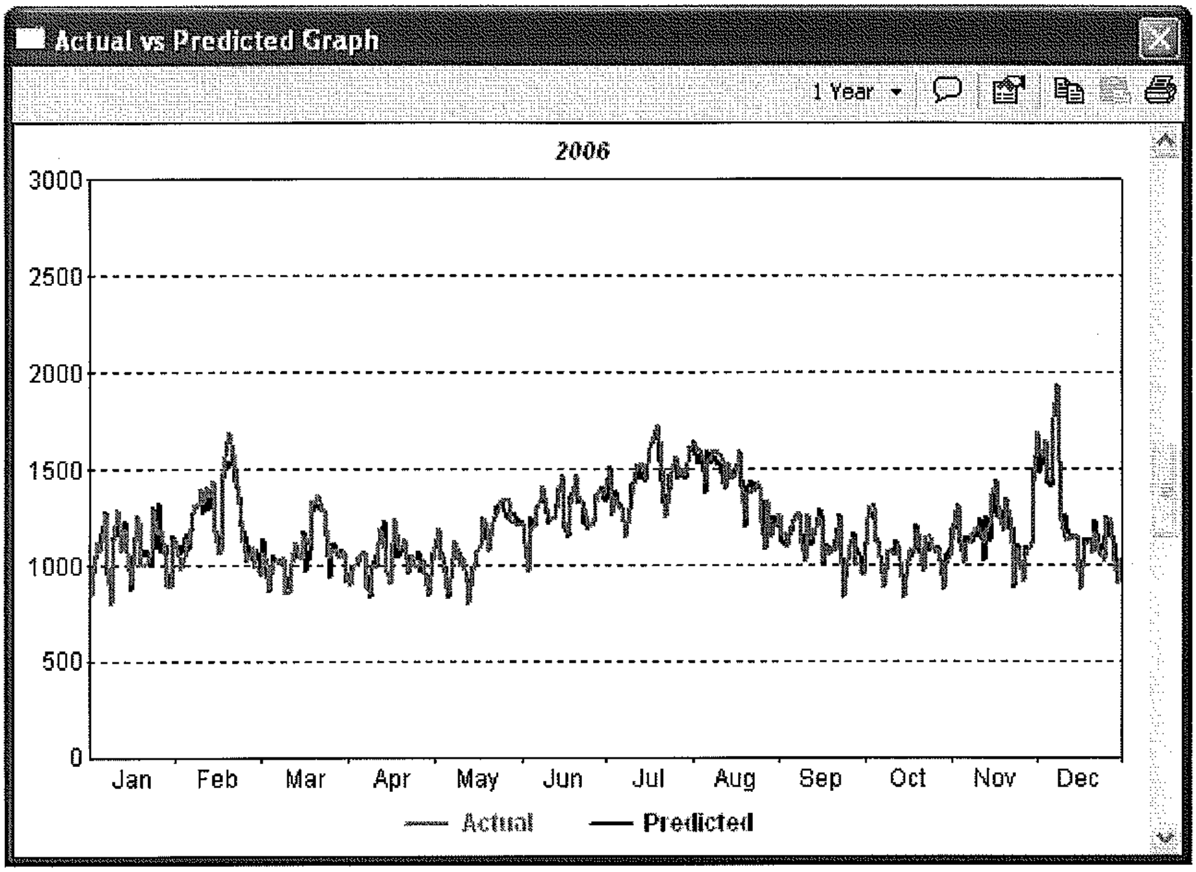



***TEB Model.*** The TEB model is built with the same classes of variables used in the Residential model (Table 1). However, temperature splines have been adjusted for the TEB weather response and no changing weather response is modeled.

The overall fit of the regression model can be seen graphically in Figure 25 and numerically in the statistics below. A full description of the model and the associated model statistics can be viewed in the *MetrixND* project file.

| | |
|---|---|
| ▪ R-Squared | 0.938 |
| ▪ Adjusted R-Squared | 0.936 |
| ▪ Mean Abs. Dev. (MAD) | 37.91 |
| ▪ Mean Abs. % Err. (MAPE) | 3.16% |
| ▪ Durbin-Watson Statistic | 1.914 |

**Figure 25: TEB Model Fit – Actual Versus Predicted Values**



# 4

# Weather Data

Normal weather conditions are a key component in the weather normalization process. In this section, the method to calculate the normal weather is discussed.

***Data.*** Historical hourly weather data from 1979 through 2008 for Springfield, Missouri were provided by Empire. These data were used to develop the daily normal weather used in the weather normalization process.

***Method.*** A rank and average method is used to develop daily normal weather. In this method, the historical data are ranked from the highest to lowest daily temperature value in each month[1]. For each historical day, corresponding heating degree days (HDD) and cooling degree days (CDD) are calculated. The normal HDD and CDD values are calculated by averaging the HDD and CDD values after they have ranked based on average daily temperature. In this method, the hottest days in the month are averaged across the 30-years of data. Similarly, the second hottest days in the month are averaged across the 30-years of data. The normal HDD and CDD values are then mapped back to the historical test year based on average temperature rankings in each month. Four steps are used to develop the daily normal HDD and CDD values.

> ***Step 1. Calculate Daily Values.*** The historical hourly values for each data were used to create the daily average temperatures.
>
> $$AverageTemperature_{day} = \frac{\sum_{hour} Temperature_{hour}}{24}$$
>
> ***Step 2. Calculate HDD and CDD Values.*** For each historical day, the HDD and CDD were calculated based on the Average Temperature in Step 1. CDD values were calculated based on temperature reference points of 60, 65, 70, 75, and 80 degrees. HDD values were calculated based on temperature reference points of 40, 45, 50, 55, 60, and 65 degrees.
>
> ***Step 3. Calculate Rank and Average based on Average Temperature.*** For each historical month, temperatures were ranked from highest to lowest value.

[1] In the Rank and Average calculation, February 29$^{th}$ values are excluded.

The corresponding HDD and CDD values on each day were averaged to calculate the normal HDD and CDD values.

***Step 4. Map Normal HDD and CDD to Calendar Year.*** In this step, the Normal HDD and CDD values calculated (Step 3) are mapped to the test year period based on rank in the test year month. The result is shown for average temperatures in Figure 26. In this figure, the bold blue line is the normal temperatures.

**Figure 26: Normal Average Temperatures**