2009 ERCOT Planning

Long-Term Hourly Peak Demand and Energy Forecast

May 1, 2009

Executive Summary

The 2009 Long-Term Demand and Energy Forecast (LTDEF) for the ERCOT region is presented in this report, including the methodology, assumptions and data upon which this forecast is based. The forecast is based on a set of econometric models describing the hourly load in the region as a function of certain economic and weather variables (primarily temperatures, heating and cooling degree-days). Economic and demographic data, including a county level forecast, are obtained on a monthly basis from Moody's Economy.com. Fourteen years of weather data are provided by DTN Meteorologix for 20 weather stations in ERCOT. The data provided by these vendors under contract with ERCOT are used as input to the energy and demand forecast models. The forecast does not account for load reductions under ancillary service programs since those programs are accounted-for in the ERCOT Capacity, Demand and Reserves report as reductions to demand for the purpose of reserve calculations.

The 2009 LTDEF reflects an initial economic decline in 2009, due to the current economic recession, and a recovery starting in 2010. For each year of the ten-year forecast period, the projected system peak demands are lower than those projected in last year's forecast. Figure 1 shows the historical peak demands from 2002 to 2008 and forecasts from 2009 until 2019. The 2009 summer peak demand forecast of 63,491 MW represents an increase of 2.11% from the 2008 actual peak demand of 62,179 MW, which was set with lower than normal temperatures (August). The historical compound growth rate for the last seven years (2002-2008) has been approximately 1.73%.

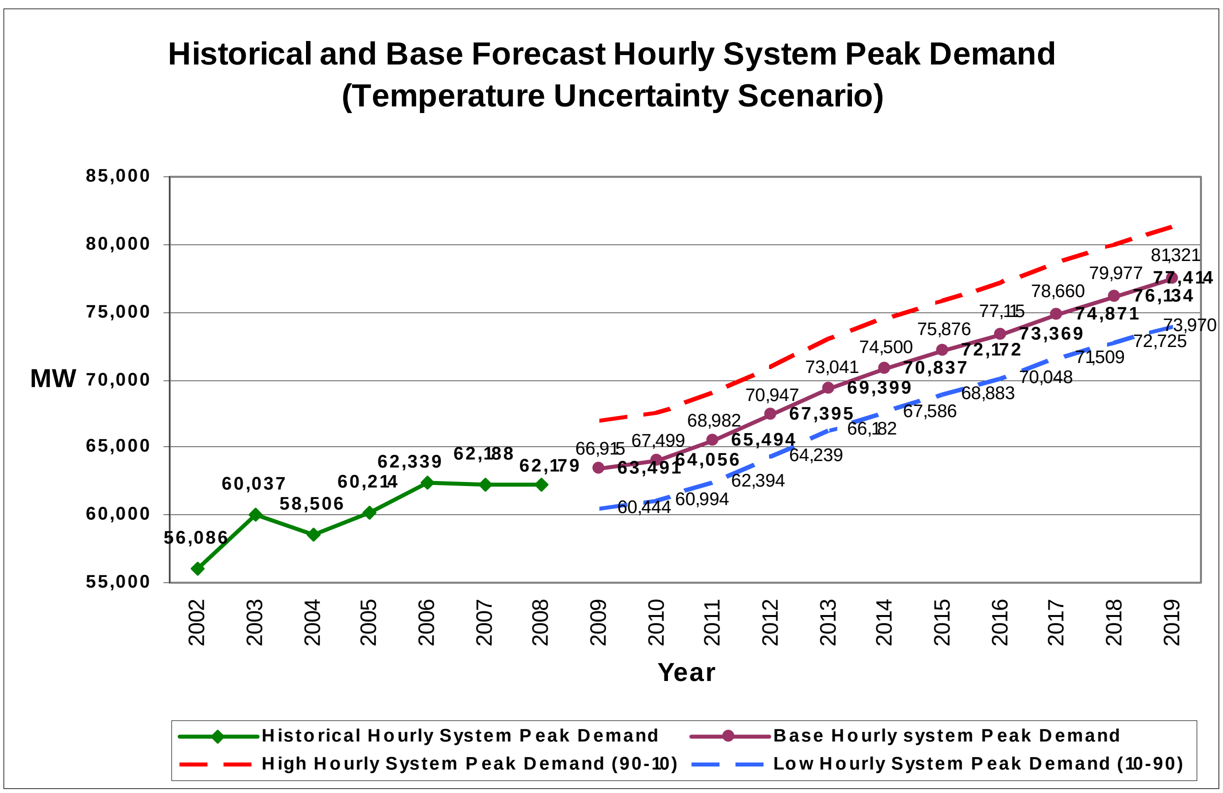


Figure 1 – Historical and Base Forecast Hourly Peak Demand

However, due the strong recovery reflected in the economic forecast, the ten-year growth rate for 2009-2019 is 2.00%, compared to last year's (2008 LTDEF) of 1.80% forecast growth rate for 2008 to 2018.

The key factor driving the lower peak demands and energy consumption (MWh), in comparison to the 2008 LTDEF, is the overall outlook of the economy, as measured by economic indicators such as the real per capita personal income, population, gross domestic product, and various employment measures including non-farm employment and total employment. The model was also recalibrated to include the effects of having an additional year of historical load data.

Also shown in Figure 1 are the forecast scenarios using statistical analysis and weather uncertainty profiles. The red dashed line on the top is a plot of the system peak demand forecasts using temperatures that exceed 90% of the historical temperatures (90^{th} percentile) experienced during the last fourteen years. This temperature uncertainty scenario forecast is referred to in the figure as the High hourly forecast 90-10. The low hourly forecast 10-90 refers to the forecasts obtained by using temperatures exceeding 10% of all temperatures during the last fourteen years. The forecast for 2009 is 63,491 MW and the 90% band is 66,915 MW or 5.39% higher than the forecast using normal weather.

The energy consumption forecast is shown in Figure 2. The energy forecast for 2009 to 2019 is lower in the first four years (2009-2012) and overtakes the levels projected in last year's forecast starting in 2013. The key factor in the decline in energy consumption is the downturn projected in the economic outlook for Texas as a result of the current recession, which is captured by economic indicators such as the real per capita personal income, gross domestic product, and various employment measures including non-farm employment and total employment.

The energy consumption forecast for 2009 of 312,204 GWh represents a decrease of 0.07% from the 2008 actual energy consumption of 312,437 GWh. The ERCOT Long-Term Demand and Energy Forecast (2009 LTDEF) energy growth rate for 2009 to 2019 is 2.04% per year, compared to last year's (2008 LTDEF) 1.79% forecast growth rate for 2008 to 2018.

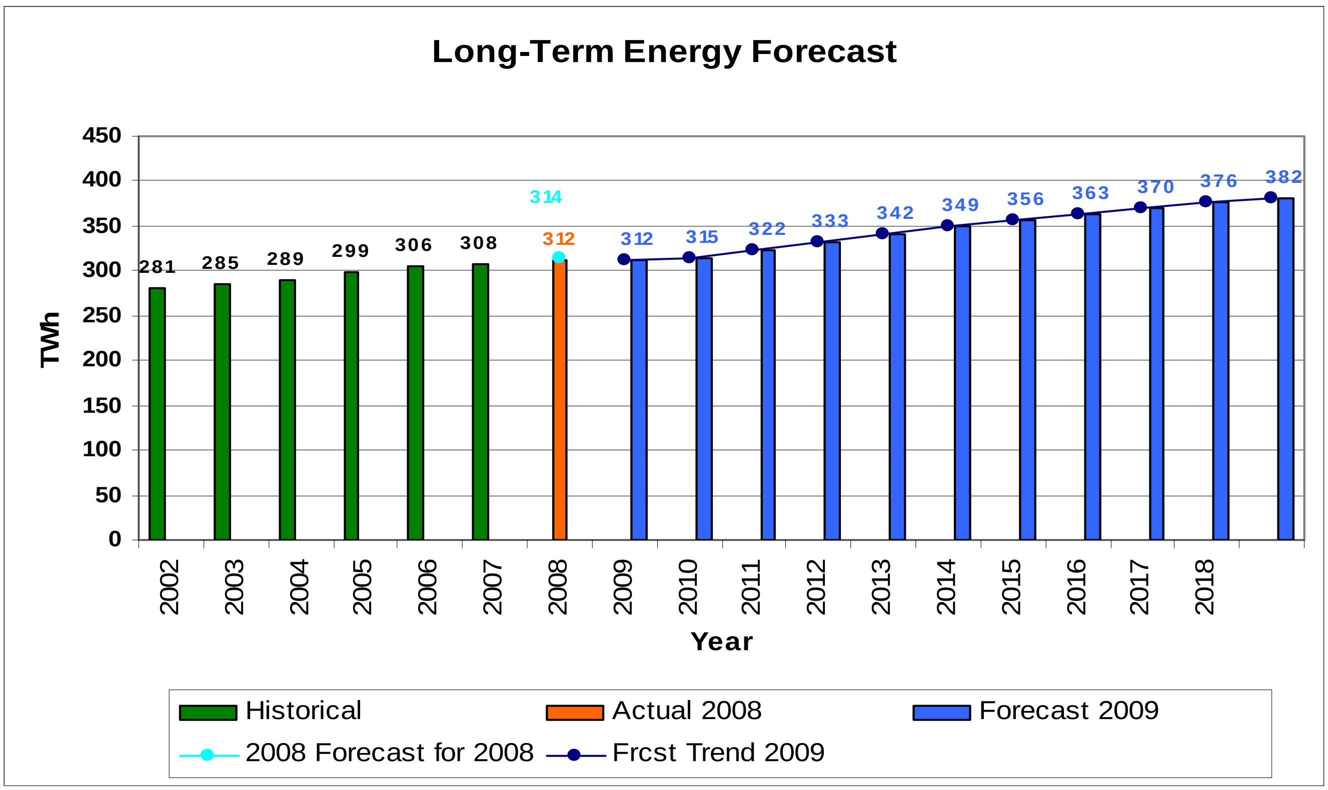


Figure 2 – Historical and Forecast Energy (TWh) Consumption

Introduction

This report gives a high level overview of the forecasts obtained from the 2009 Long-Term Forecast Model. The methodology is briefly described, highlighting the major aspects involved in producing the forecast, including the data inputs used in the process. Second, a historical perspective of the load growth in the ERCOT's territory is provided and final results of the forecast peak demands and energy from 2009 to 2019 are presented in a graphical form and summarized in a table summary format. Third, a discussion of the major drivers of peak demands and energy consumption is included, along with the uncertainties associated with the forecast, and the differences with last year's forecast. The final hourly load shape forecast is presented in a graphical form giving a perspective or comparison of the actual and forecast trends out into the period 2009 to 2019. Finally, a more detailed description of the econometric forecasting methodology used by ERCOT is provided in Appendix 3.

General Background: Forecast Development Description

The 2009 Long-Term Demand and Energy forecast was produced with a set of econometric models that use weather, economic and demographic data and calendar variables to capture and project the long-term trends in the historical load data for the past six years.

First, a representative hourly load shape by weather zone is forecasted using an average weather profile of temperatures and Cooling Degree Hours (CDH) and Heating Degree Hours (HDH) obtained from historical data to project the load shape into the future. Other factors such as seasonal daily, weekly, monthly and yearly load variations and holidays, in addition to exogenous variable interactions, such as of weather and weekends and weekdays are also considered. This hourly ERCOT Load Shape only describes the hourly load fluctuations within the year and in itself does not reflect the long-term trend.

The long-term trend is provided by the energy forecast. The monthly energy forecast models for each weather zone use Cooling Degree Days (CDD) and Heating Degree Days (HDD), economic and demographic data, and indicator variables for special events to project the monthly energy for next eighteen years (2009 - 2019).

Data Sources

Economic and demographic data, including a 20-year forecast at the county level, are obtained on a monthly basis from Moody's Economy.com. These data are used as input to the monthly energy models.

Fourteen years of weather data are available from DTN Meteorologix for 20 weather stations in ERCOT. Data from these weather stations are used to develop weighted hourly weather profiles for each of the eight weather zones. These data are used in ERCOT's Load Shape models. Monthly CDD and HDD are used in the monthly energy models.

The economic and demographic, and weather data are provided by the vendors above, and as such, are proprietary data and under contracts which require that these data not be released to the public.

Historical load data are available on an hourly basis from ERCOT's data aggregation systems since July 31, 2001 when ERCOT began operations under a single control area. Prior to 2001, ERCOT obtained hourly load data from Transmission and Distribution Service Providers (TDSP) going back to 1995. Historical weather zone load data have only been collected from July 31, 2001.

ERCOT's Historical and Forecasted Peak Demands and Average Load Growth

The Figure 3 (below) compares the ERCOT's average hourly load with the annual system peak demand. The growth of the average hourly load is considered almost as a fixed amount that can be estimated with a reasonable degree of accuracy. The peak demand growth, however, is a much more volatile variable and more difficult to predict. The many factors affecting peak demand and the high degree of uncertainty in the long run make it a challenging variable, in term of assessing its behavior in the future.

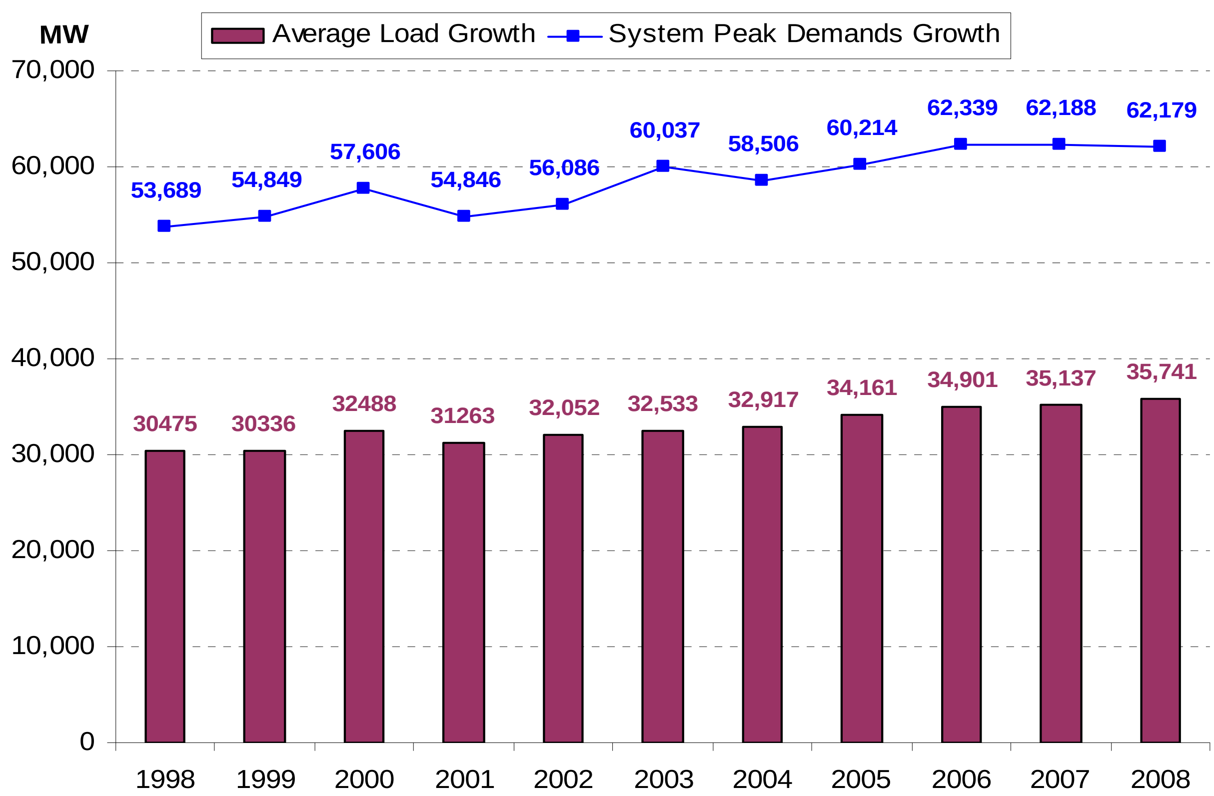


Figure 3 – ERCOT Historical Average Load versus System Peak Growth

Over the period from 1998 to 2008, ERCOT's average hourly load grew 17.28%. By comparison, ERCOT's system peak grew 15.81%. The average annual growth rate of the system peak was 1.58% over this period.

From 2002 to 2008, a similar pattern can be detected. The average load growth rate was 11.51% versus 10.86% for the system peak. The average growth rate of the system peak demand below the average load growth over the period from 2002 to 2008 was 0.95%.

The actual system peak demand from 1998 to 2006 experienced a high growth rate which can be attributed to the specific weather for that period. The same cannot be said for the growth in system peak demand for 2006/2007 and 2007/2008. It is not likely that these specific weather patterns will be reproduced in the future, or that the relationship between average load and peak demand growth will be kept the same as in either of these periods. Although the system peak demand is affected by economic and demographic factors, it is predominantly determined by weather. On the other hand, the average load growth intrinsically reflects growth associated with economic and demographic factors.

The 2009 Long-Term peak demand and average load forecast is graphed below in Figure 4. Over the ten year period (2009-2019) the average load is projected to grow 22.35% or at a 2.23% growth rate. The total system peak demand growth over the same period is 21.93%, equivalent to a 2.19% average annual growth rate. The equivalent compounded growth rate equates to 2.00%.

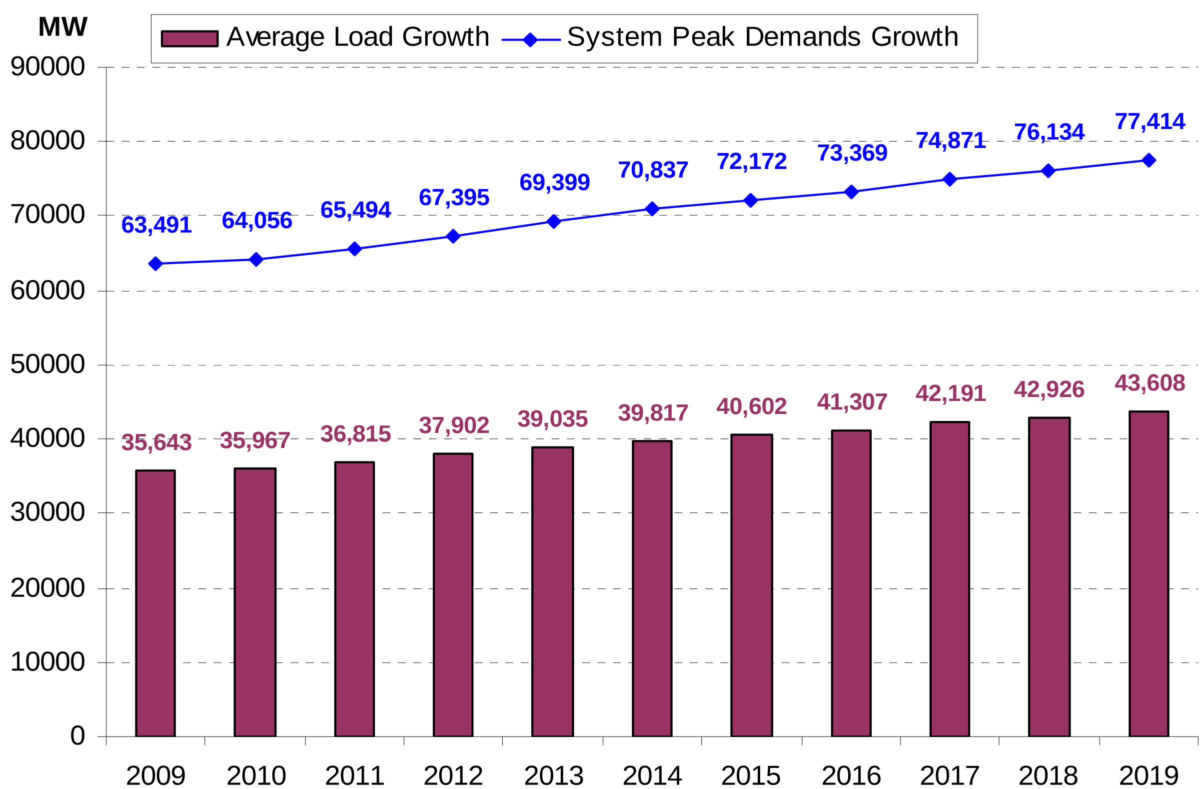


Figure 4 – ERCOT Forecast Average Load versus System Forecast Growth

ERCOT's Peak Demand and Energy Forecasts

The annual historical and forecast peak demands, and the energy consumption, are displayed in figure 5 and 6, below. The historical peak demand compound growth rate from 2002 to 2008 was 1.73% and the energy growth rate over the same period was 1.80%. The 2009 LTDEF peak demand and energy forecast produced compounded growth rates of 2.00% for the peaks from 2009 to 2019 and 2.04% for the energy over the same period.

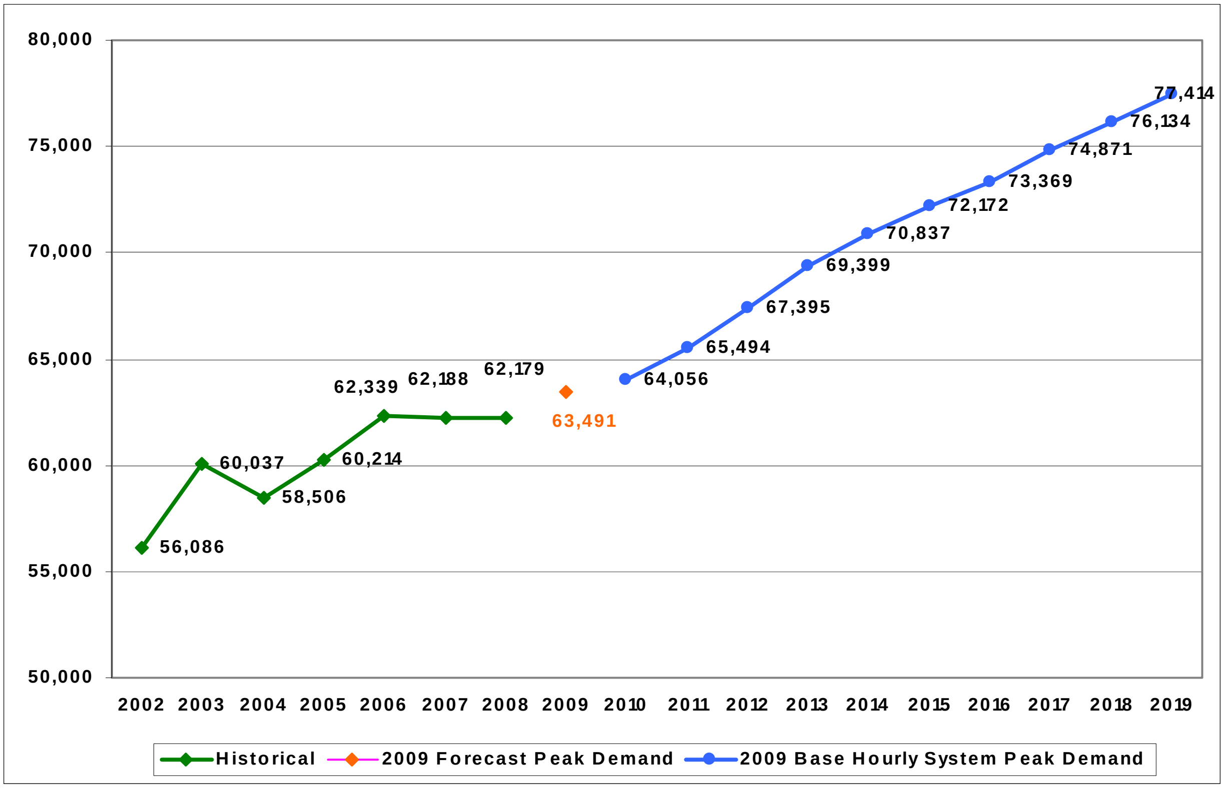


Figure 5 – Historical and Forecast Hourly Peak Demands

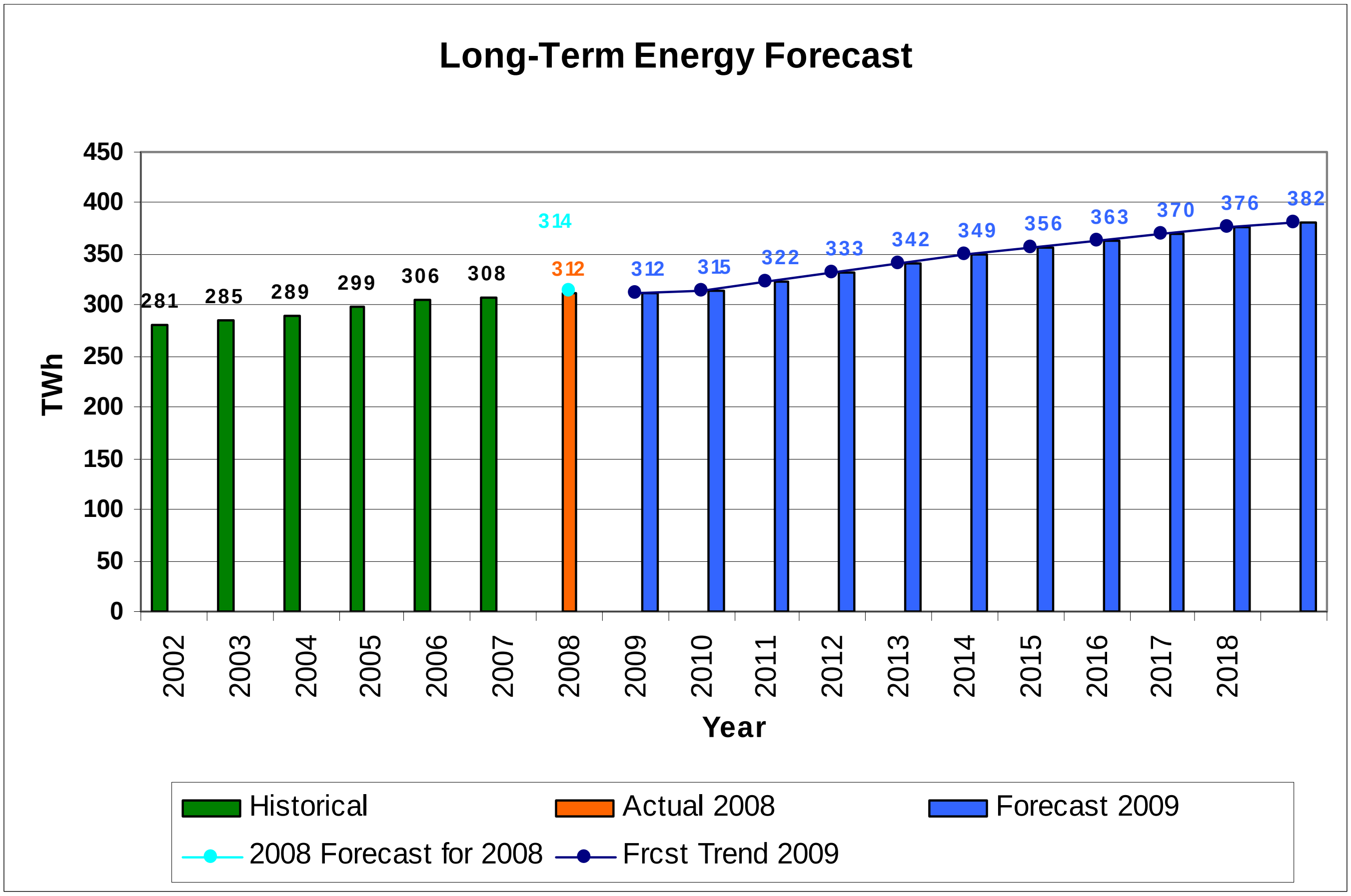


Figure 6 – Historical and Forecast Energy Consumption

Economic Outlook and Factors Driving Peak Demand and Energy

Growth in electricity demand and consumption is closely correlated with three main factors: 1) Weather, 2) Economics, and 3) Demographics. Economic and demographic changes can affect the characteristics of electrical demand in the medium to the long-run. Weather, on the other hand, drives most of the variation in electric demand in the short-run. Thus, since weather also affects the variation in the electric demand in the long-run, long-term forecasting using historical average weather profiles to indicate the future variation in weather.

In general, the economic variables used in the models throughout the eight weather zones in the ERCOT electric grid, are various forms of employment indicators, such as total non-farm employment and total employed, real personal per-capita personal income, gross domestic product and population. Employment is a measure of the growth in the commercial and industrial areas. Population is a proxy for capturing customer formation, and income addresses overall standard of living which translates into increase in comfort and convenience and in many instances leads directly to an increase in electricity demand. GDP is an important measure of economic activity in a country or an area, such as the ERCOT territory. The gross domestic product is the synthesis of three sides of the economy: expenditure, output, and income. GDP is thought to capture the overall health of the economy and shows a high correlation with the growth in electricity use. These key factors are driving the lower peak demand and energy consumption forecasts, reflecting the overall state of the economy. The graphs of each indicator are presented further down this document in Figures 7 – 11.

The 2009 forecast is lower than last year's forecast for 2009 due to the national economic recession that started in December 2007, and developed into the current deep recession and financial meltdown at the US and global level. The result has been that growth has slowed to some extent at the state level, here in Texas, which affects the state's outlook for growth in employment, income and gross domestic product (GDP). Additionally, there are some shorter term effects, derived from the housing sub-prime loans and the credit liquidity issues, which will prevail over the next two to four years. Ultimately, the economy is forecasted to rebound in 2010 and return to its normal trend in 2013. In addition to the depressed economy due to the recession, there were energy efficiency and conservation effects indirectly triggered by the escalation of gas prices that started at the beginning of the summer 2008 and lasted throughout the summer and much of the remainder of the year. The resulting impact on consumer's budgets led to a higher level of awareness and energy conservation by households and commercial establishments.

There has been a deceleration in the Texas employment, and near-term decline is forecasted. However, Texas will continue to perform better than the US. Even though the decline in housing permits is similar to the US as a whole and existing home sales slowdown considerably, the decline in home prices has been less than everywhere else in the country. The decline in high energy prices has provided some relief to the state's consumers as a whole. However, this decline has negated the energy-related boom in the Houston economy that took place in 2008. Longer-term, growing global energy demand and decreasing energy supply will raise the energy prices, but not to the peak levels seen in 2008.

The effects of these economic indicators, used in the 2009 forecast result in an impact of about 2700 MW in the summer forecast in 2009, or approximately 4.3% decline from last year. The energy forecast for 2009 is around 2.30% below the forecast produced in 2008. In the long-term the energy forecast is higher than last year's forecast due to a strong recovery fueled by a catch-up effect in housing starts and unit car sales. The system peak demand also recovers after 2013, but never reaches the peak levels of the projection in 2008.

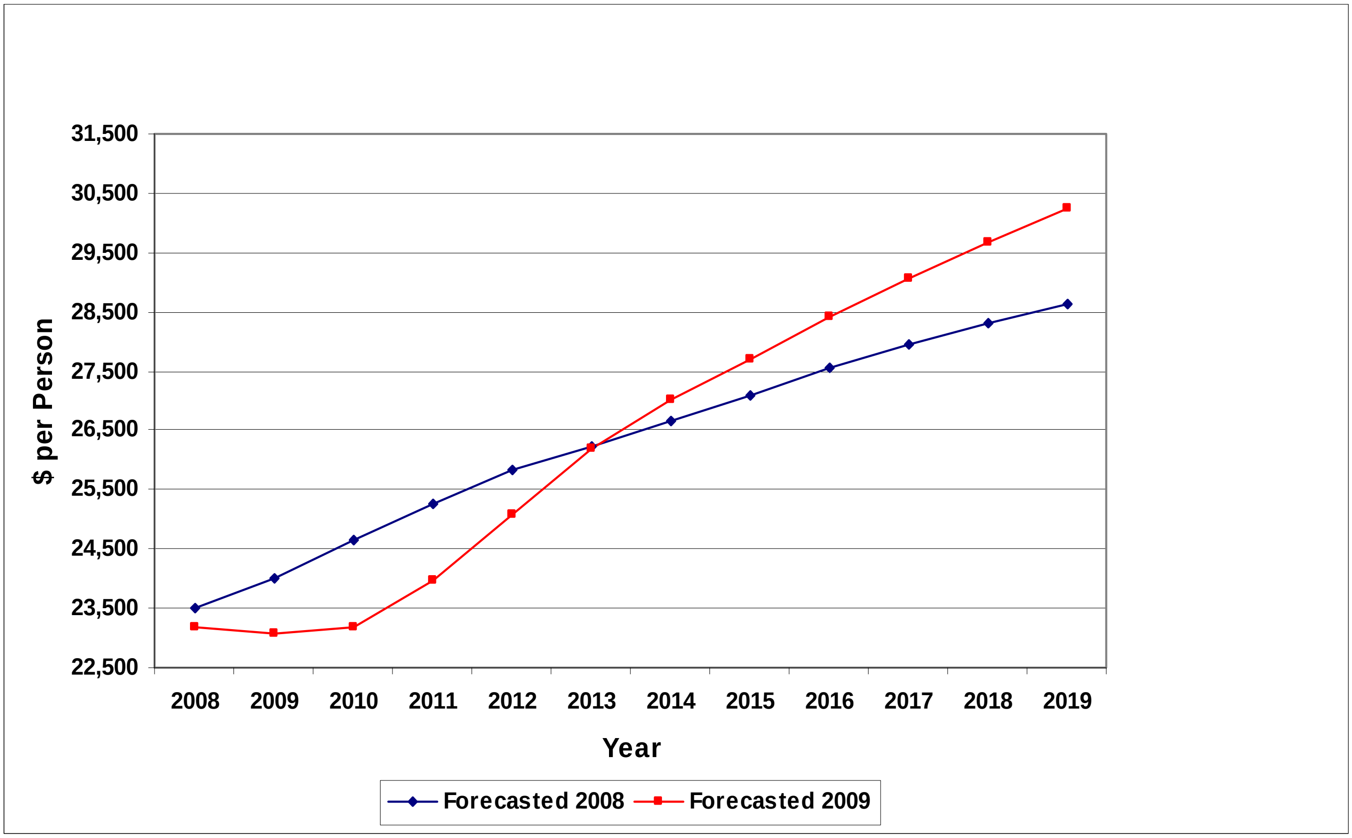


Figure 7- Real Personal Per-Capita Income

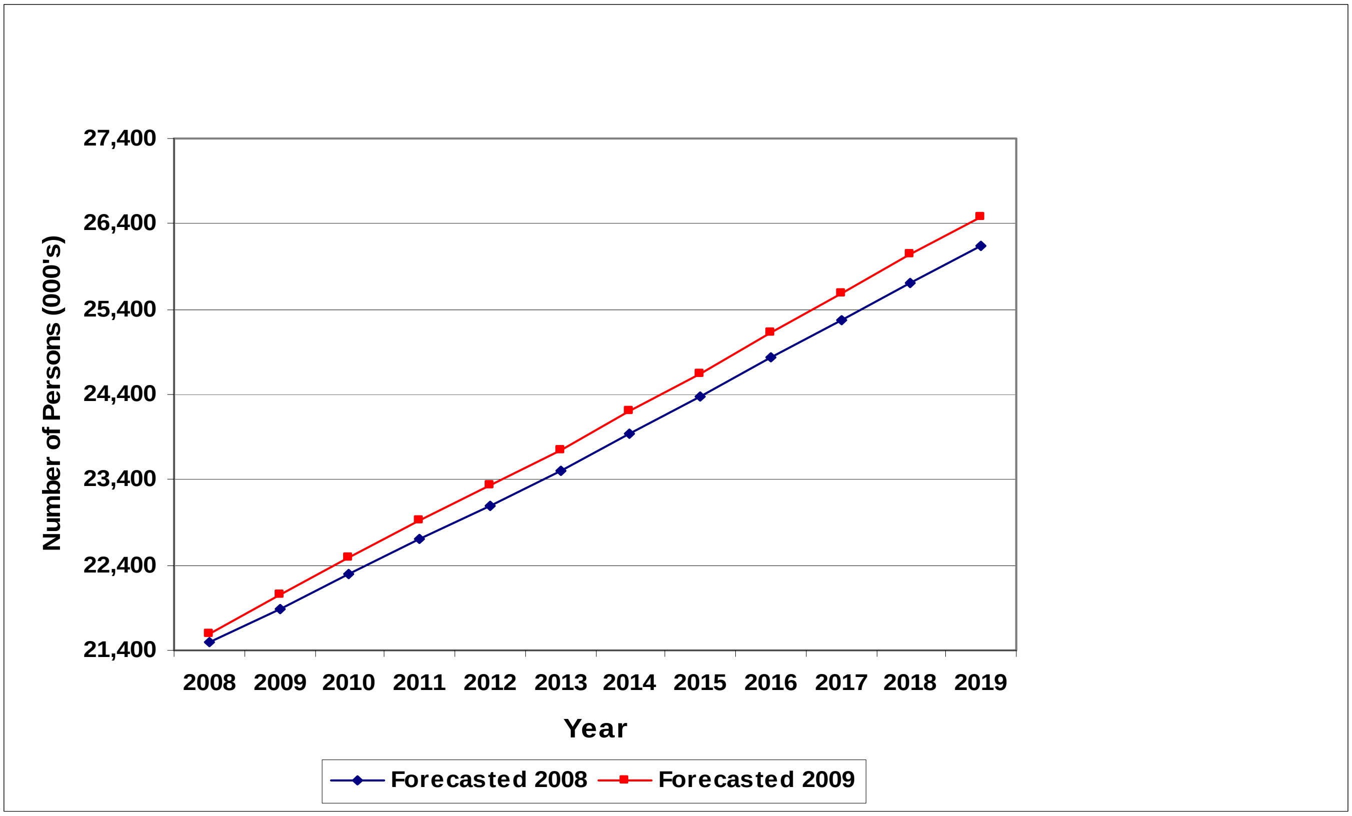


Figure 8 – Population in the ERCOT Territory

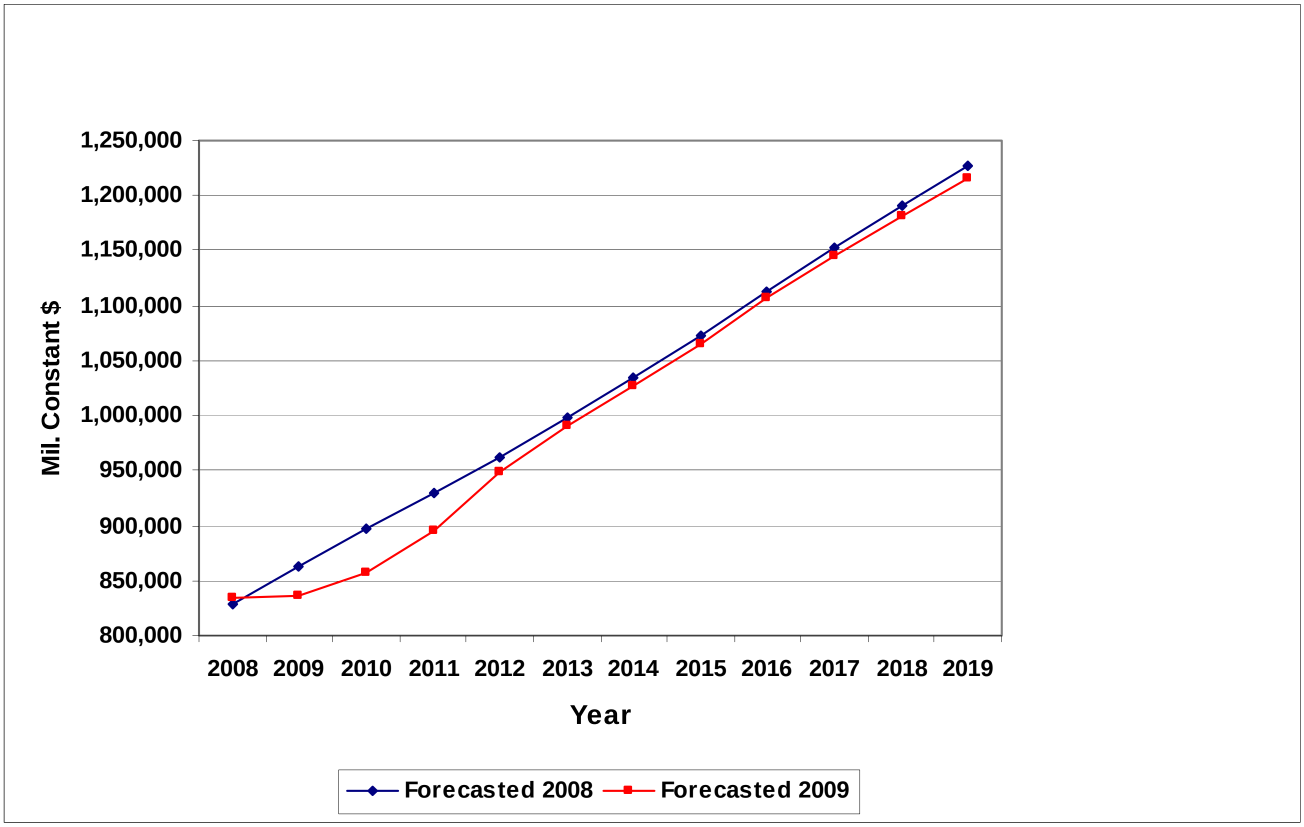


Figure 9 – Gross Domestic Product in the ERCOT territory

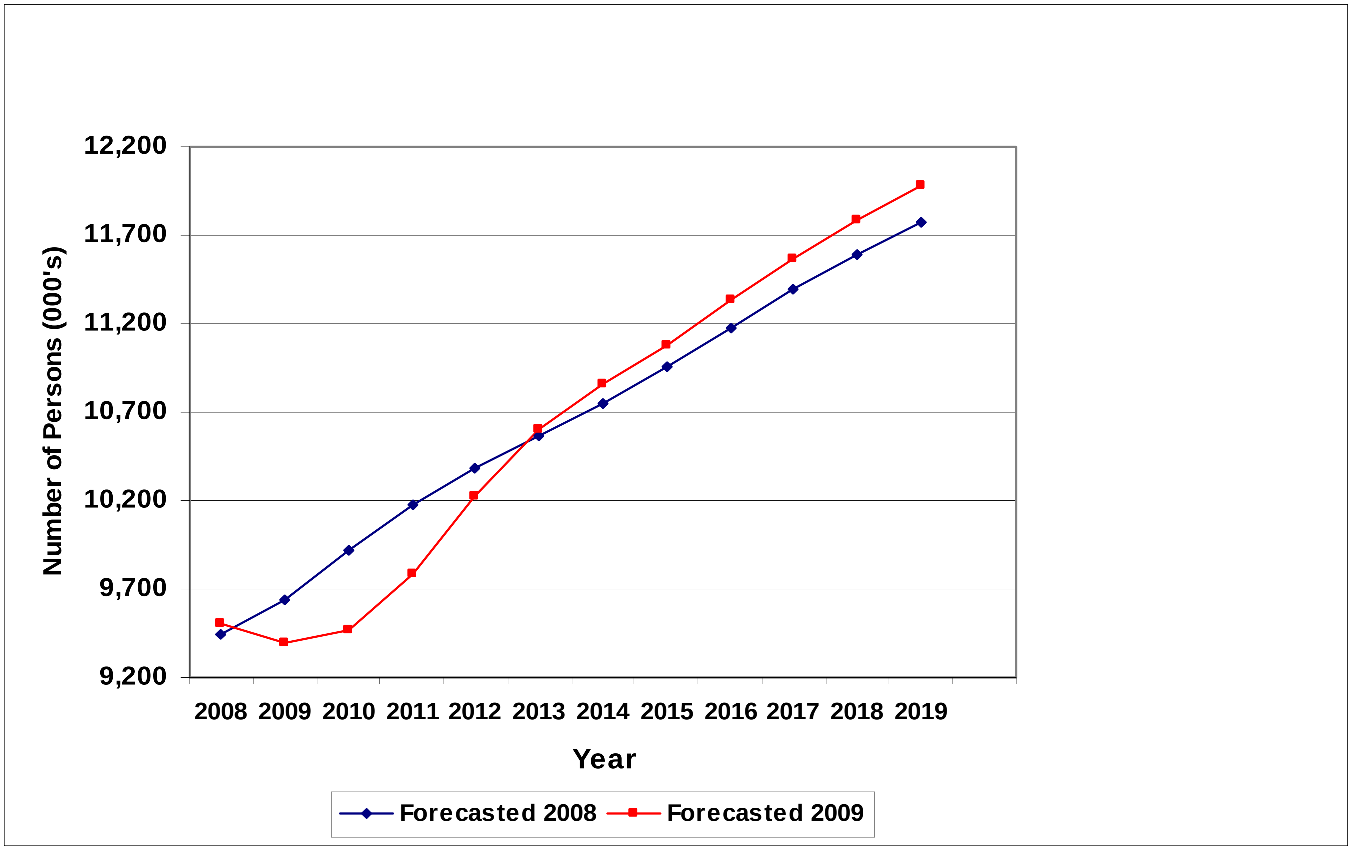


Figure 10 – Total Non-Farm Employment

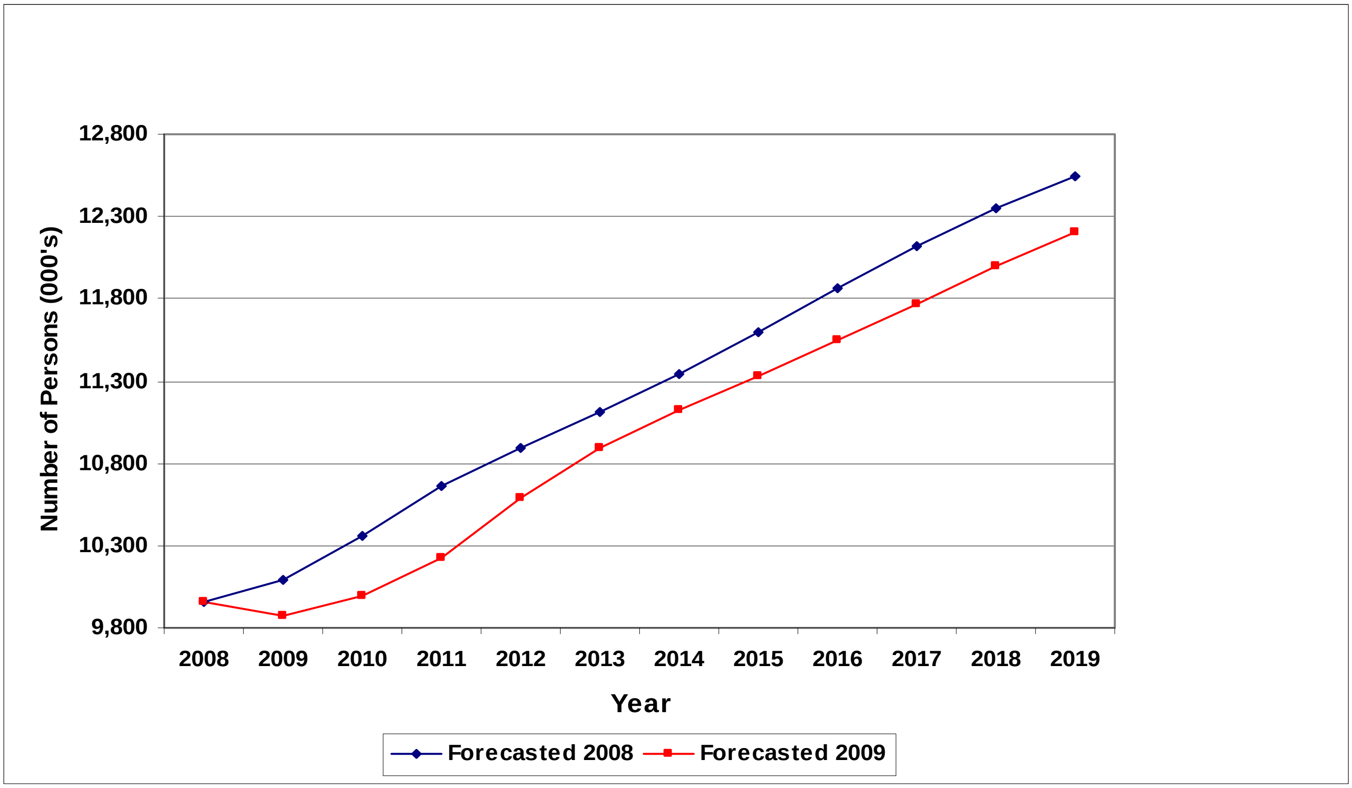


Figure 11 – Total Persons Employed

ERCOT's Peak Demand and Energy Uncertainty

One measure of the uncertainty associated with extreme weather impacts on the peak demands can be obtained by using a more extreme weather profile to obtain the forecasts. ERCOT developed weather profiles that rank at the 90^{th} percentiles of all the temperatures in its hourly temperature database and did the same to develop with the 10^{th} percentile of all temperatures. Strictly speaking these are not confidence bands in the statistical sense, but common use has been to use this term to refer to the results. A more appropriate term would be to use scenarios associated with the 90^{th} percentile temperature distribution or 90^{th} percentile scenario forecasts. ERCOT has also, in the past, run Monte Carlo simulation to assess the extreme temperatures on the peak demands.

For the 2009 LTFM the 90% Confidence Bands were developed and are depicted in the figures below. The high forecast for 2009 is 5.39% higher than the 2009 forecast with an average weather profile.

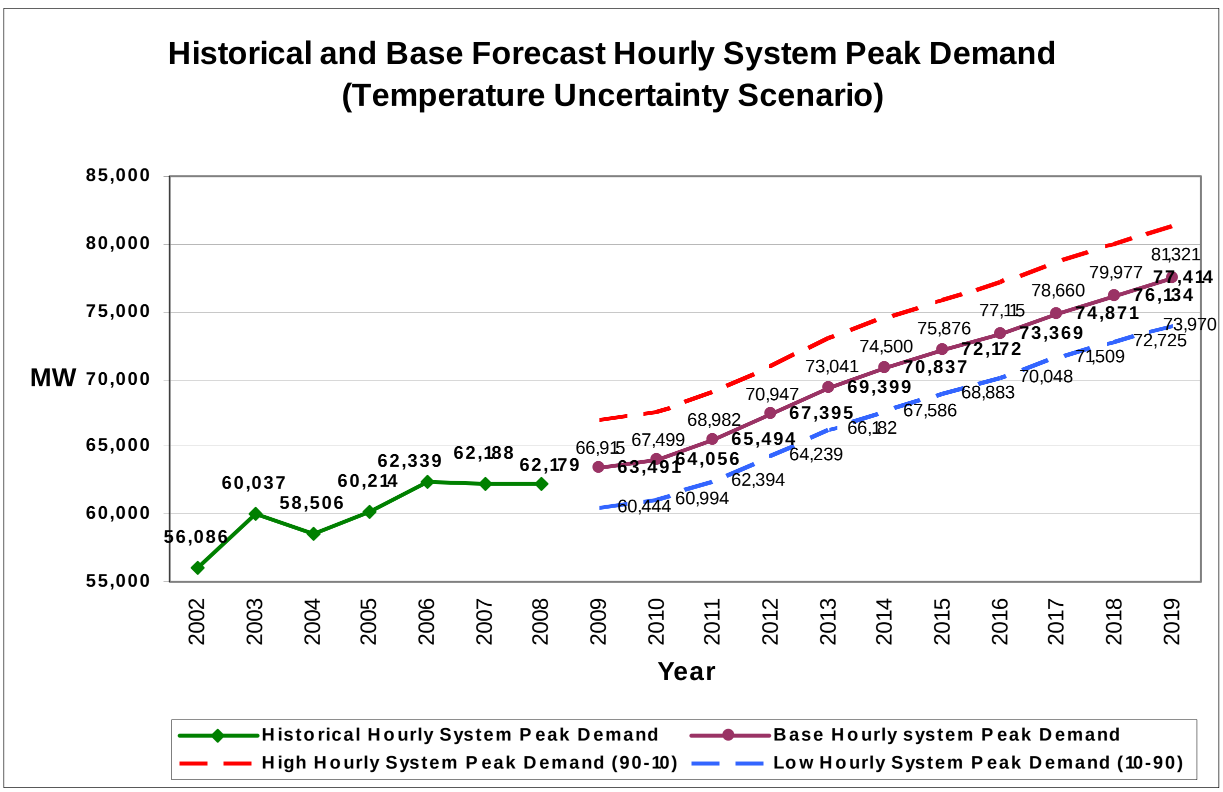


Figure 12 – Historical and Forecast Hourly Peak Demand

Differences with Last Year's Forecast

In the near term, the forecast differs significantly from last year's forecast. Overall, the forecast is lower due to the effects of a national recession that are having an impact on the Texas economy. The forecasting models were recalibrated based on having an additional year of actual data. The figure below shows the two forecasts over the 2009 to 2019 time frame.

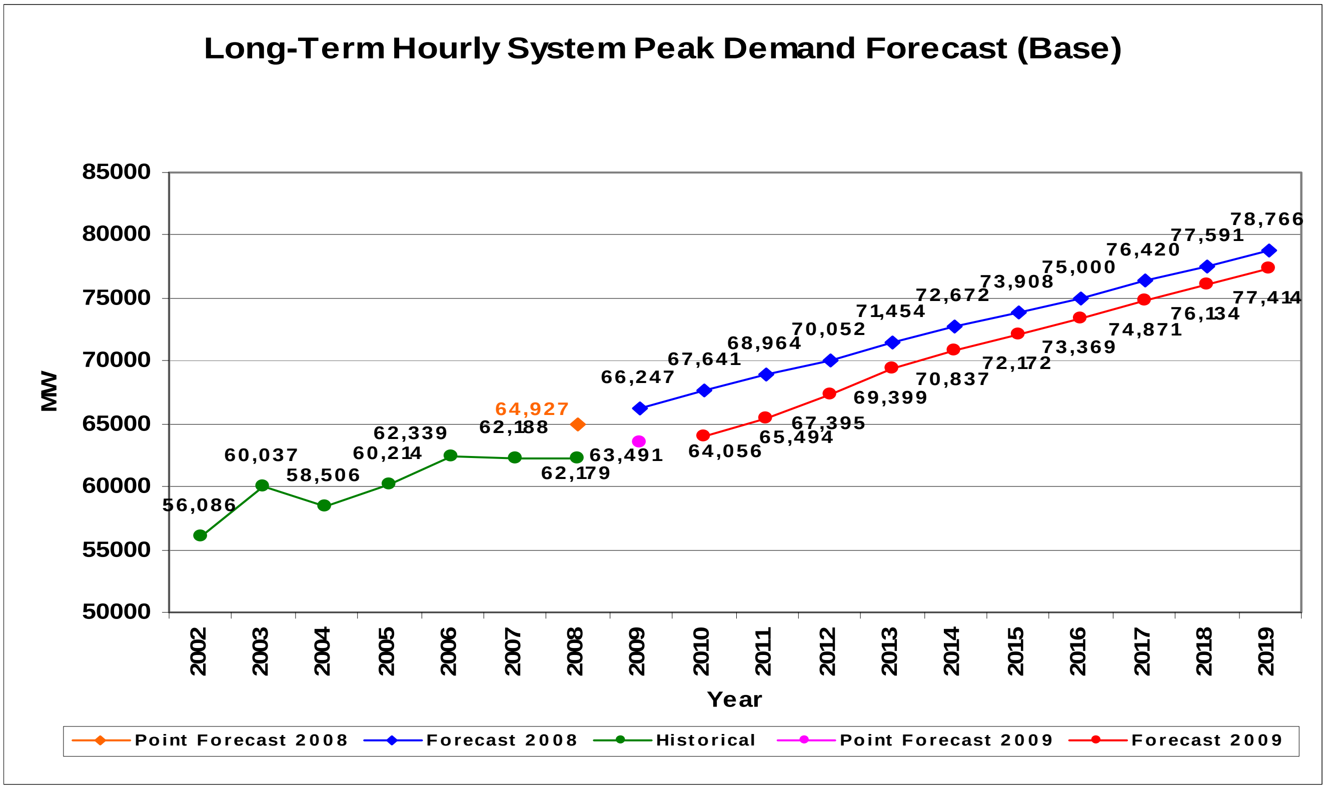


Figure 13- Comparison of 2008 and 2009 Forecast

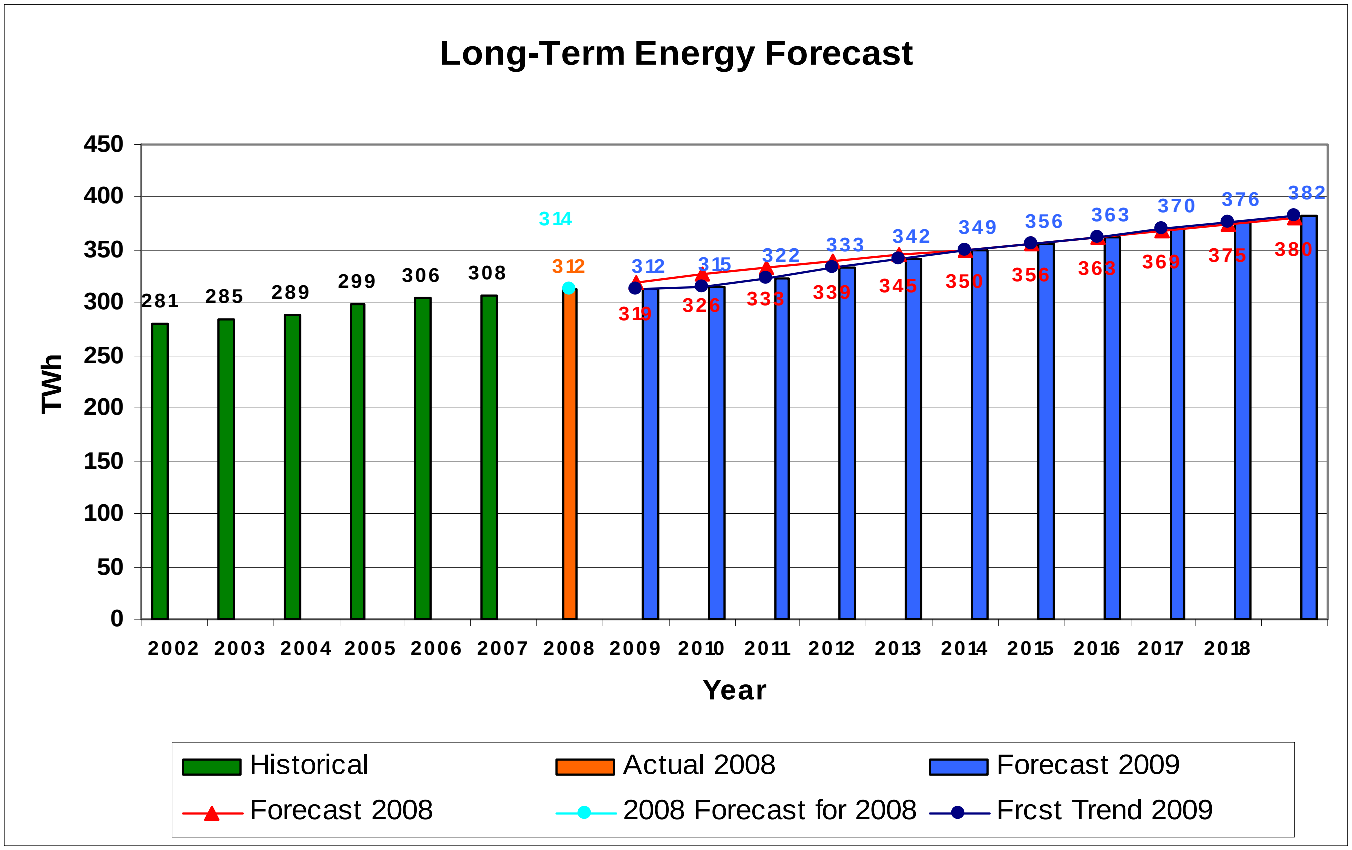


Figure 14 - Comparison of 2008 LTDEF and 2009 LTDEF

ERCOT's Load Shape Forecast

The process used to develop ERCOT's peak demand forecast produces an hourly Load Shape for each weather zone. The hourly load forecast also contributes the annual system peak demands that are used in the resource adequacy assessment, NERC summer and Long-Term assessments, and other reports. The 2009 Long-Term System Hourly Load forecast over the next five years (2009-2015) and the forecast (fitted) results are shown in the figure below.

Figures 15 and 16 depict the forecast load shapes for 2009 to 2015. Each of these load shapes is derived using an average weather profile. Because of this, the load shapes are basically the same for each forecast year. The upward trend comes from the economic forecasts that drive the energy consumption forecasts. Figure 17 shows one 24 hour day for the peak forecast day in 2009.

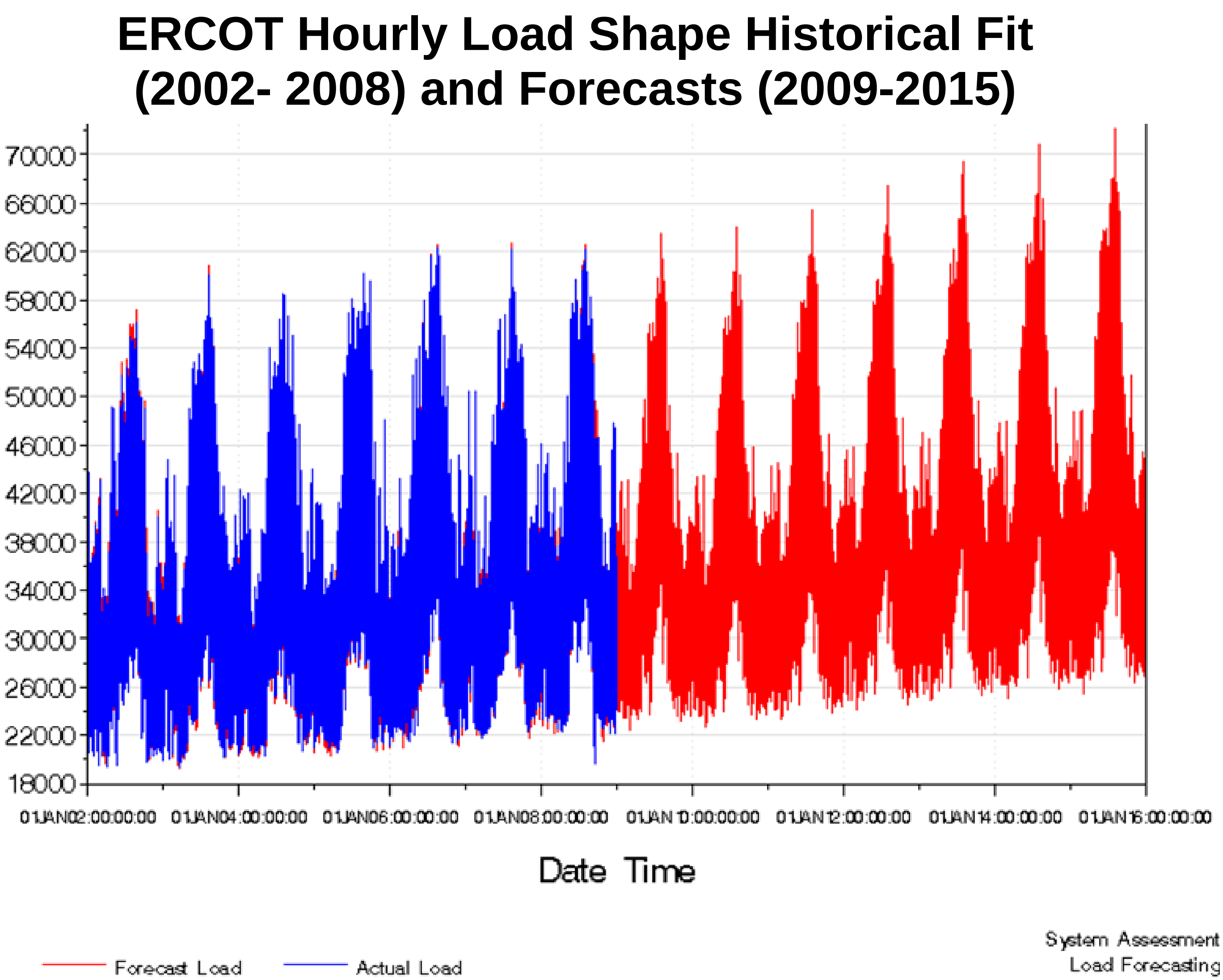


Figure 15 – Hourly Load Forecast including Historical Fit

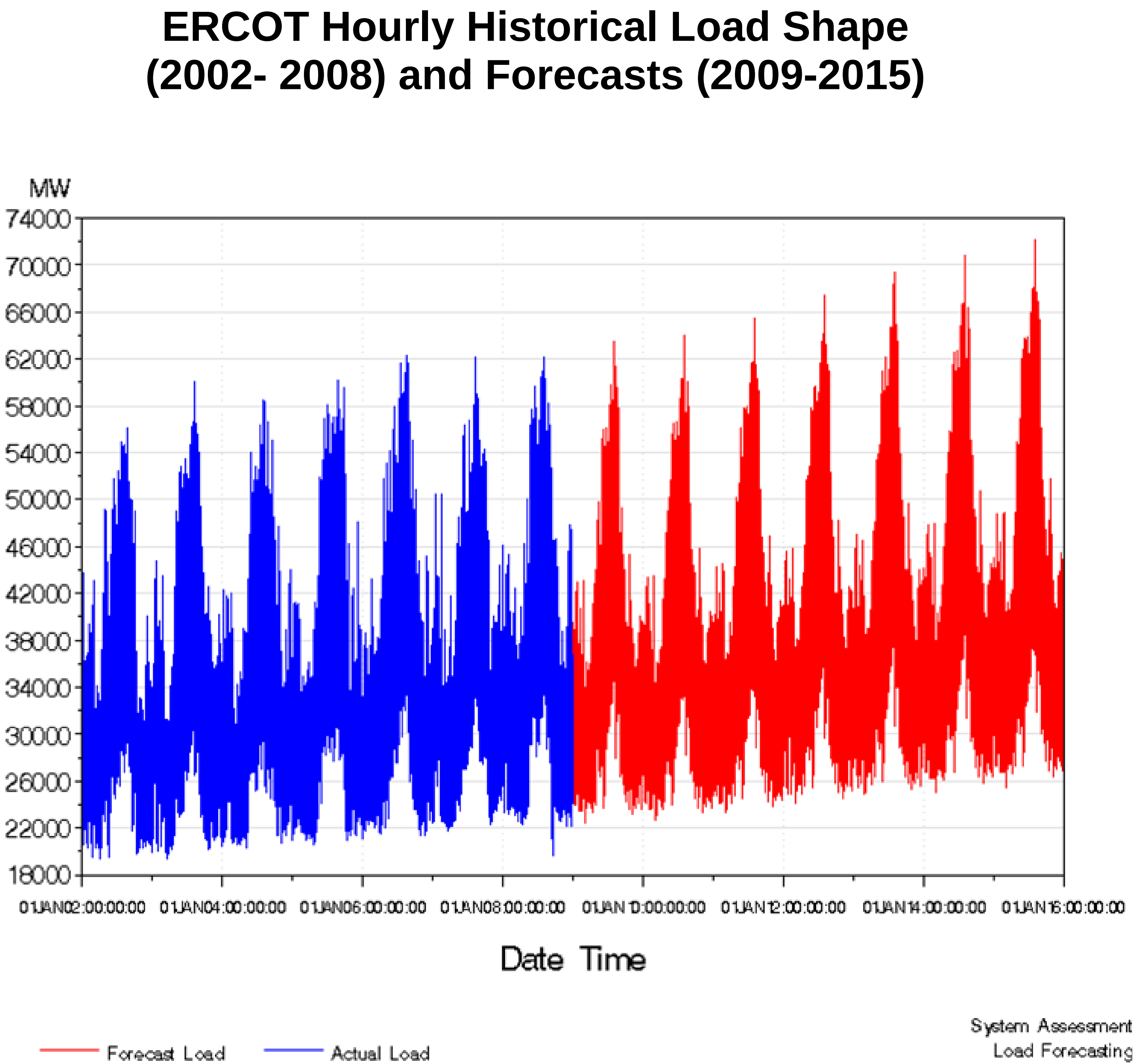


Figure 16 – Hourly Load Forecast and Actual

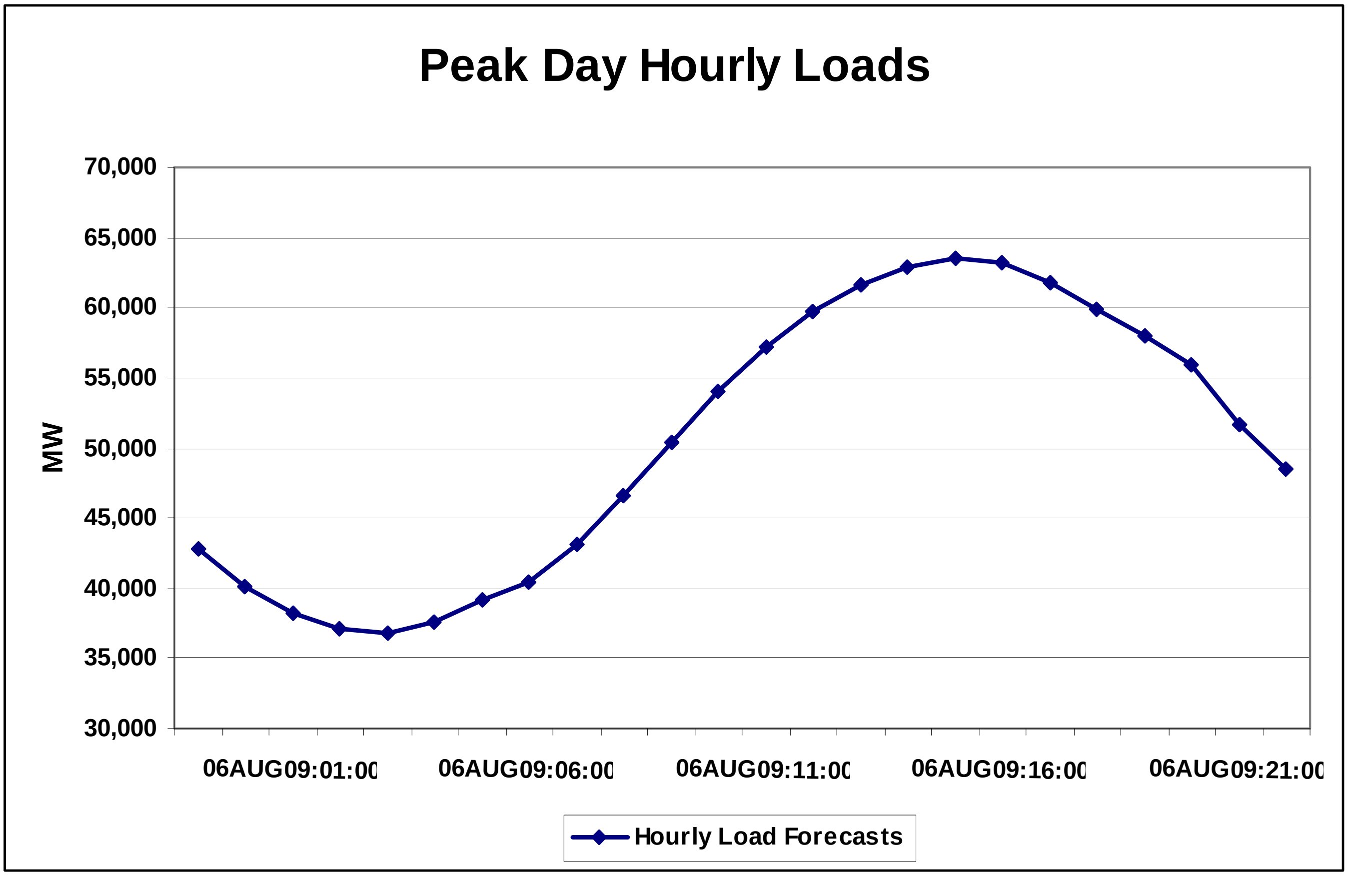


Figure 17 – Hourly Peak Loads for August 6, 2009

ERCOT's Peak Demand and Energy Forecast by Weather Zone

There are eight defined weather zones at ERCOT. The weather zones are: 1) North, 2) North Central, 3) East, 4) Far West, 5) West, 6) South Central, 7) Coastal, and 8) South. The largest MSAs are located in the North Central, South Central and Coastal zones. The Dallas/FW area is in the North-Central, and the Austin and San Antonio areas are contained within the South-Central and Houston is in the Coastal zone. All three areas have been affected by substantial slowdowns in nonresidential construction and property markets. This cyclical weakness is most apparent in Dallas and Austin, where office vacancy rates are particularly high and rising. However, the credit crisis has also meant that weakness is emerging on San Antonio and Houston. Further, although service-producing industries had previously offsets declines in manufacturing employment, in the short-term that is no longer the case in these major MSAs.

Corporate restructuring in finance, retail, and high tech has meant that many jobs in professional services, banking, and retail have been lost. Moreover, the Houston area has slowed down due to the decline in energy prices. However, the overall effect of lower energy prices, which are a major factor for large industrial manufacturing industries, is difficult to assess, because lower energy costs improve the profitability of industries which use a lot of energy as an input. Thus, the forecasts for these major zones vary in terms of near-term economic performance. Longer term, after the current cycle finally ends, the various fundamentals which drive above-average long-term performance of the largest, compared to the U.S as a whole, remain in place. These include above average population growth, relatively lower costs of doing business when contrasted with comparable metropolitan areas elsewhere in the country, energy resources, concentration of high tech companies, and growing transportation and distribution capacity. The forecasts for the smaller zones show an average or below average trend in growth.

The annual forecasts data by weather zone are included in Tables 2 and 3 of appendix 2.

APPENDIX 1: PEAK DEMANDS AND ENERGY CONSUMPTION DATA

A summary of the 2009 Long-Term Forecast Model (LTFM) results is condensed below. This table includes forecast energy, forecast energy for the load shape, the MWh historical values, the coincident and zonal peaks, the diversity, coincident, and load factors and the diversity in % terms. For reference, historical data for 2002-2008 included. The MW peak is a coincident peak and the zonal peak refers to the aggregate of individual non-coincident peaks. The Energy MWh column, from 2002-2008, contains the forecasted values for that period. The MWh_Hist contains the historical energy consumption for 2002-2008. The following quantities in the table below can be defined as follows (numbers are rounded):

Load Factor: (energy/(peak*number of hours)
Diversity: (Non-Coincident Peak – Coincident Peak)
Diversity Percent: (Diversity Factor/Coincident Peak)
Coincident Factor: (1-Diversity Percent)

Year	Actual/Forecast MWh	MWh Peak	Zonal Peak	Diversity	Coincident Factor	Diversity %	Load Factor
2002	280,772,959	56,086	57,233	1,146	97.96%	2.04%	57.15%
2003	284,983,916	60,037	60,376	339	99.44%	0.56%	54.19%
2004	289,140,984	58,506	59,316	810	98.62%	1.38%	56.42%
2005	299,253,971	60,214	61,364	1,150	98.09%	1.91%	56.73%
2006	305,740,287	62,339	63,352	1,013	98.37%	1.63%	55.99%
2007	307,800,947	62,188	63,570	1,382	97.78%	2.22%	56.50%
2008	312,437,873	62,179	64,379	2,200	96.46%	3.54%	57.36%
2009	312,232,038	63,491	64,134	643	98.99%	1.01%	56.14%
2010	315,064,738	64,056	64,701	645	98.99%	1.01%	56.15%
2011	322,498,252	65,494	66,152	658	99.00%	1.00%	56.21%
2012	332,936,113	67,395	68,072	677	99.00%	1.00%	56.39%
2013	341,949,514	69,399	70,019	620	99.11%	0.89%	56.25%
2014	348,795,556	70,837	71,543	706	99.00%	1.00%	56.21%
2015	355,679,122	72,172	72,889	717	99.01%	0.99%	56.26%
2016	362,834,886	73,369	74,096	727	99.01%	0.99%	56.45%
2017	369,590,447	74,871	75,594	724	99.03%	0.97%	56.35%
2018	376,037,309	76,134	76,896	762	99.00%	1.00%	56.38%
2019	382,006,828	77,414	78,088	673	99.13%	0.87%	56.33%

Table 1 – Forecast Results of the 2009 Long-Term Forecast Model

APPENDIX 2: WEATHER ZONE LOAD DATA

Year	North	North Central	East	Far West	West	South Central	Coast	South
2002	1,904	20,527	2,175	1,830	1,595	9,492	14,578	3,985
2003	2,070	22,303	2,319	1,805	1,675	10,016	15,823	4,025
2004	2,047	20,749	2,265	1,658	1,562	9,619	16,611	3,996
2005	2,080	21,975	2,351	1,661	1,542	10,162	16,282	4,159
2006	2,361	22,698	2,433	1,599	1,613	10,718	16,728	4,189
2007	2,166	22,034	2,248	1,637	1,469	10,419	18,240	3,976
2008	2,333	22,589	2,300	1,803	1,564	10,884	16,826	3,881
2009	2,027	22,903	2,278	1,828	1,651	10,478	18,082	4,244
2010	2,003	23,072	2,305	1,821	1,651	10,646	18,308	4,251
2011	2,038	23,344	2,389	1,850	1,696	11,061	18,790	4,326
2012	2,088	23,656	2,525	1,888	1,786	11,571	19,449	4,432
2013	2,134	24,026	2,738	1,928	1,864	12,106	20,042	4,563
2014	2,171	24,315	2,735	1,948	1,912	12,562	20,528	4,666
2015	2,195	24,597	2,805	1,968	1,947	12,895	21,025	4,740
2016	2,215	24,814	2,870	1,987	1,983	13,200	21,498	4,801
2017	2,248	25,160	2,946	2,014	2,031	13,552	22,041	4,877
2018	2,272	25,416	3,024	2,034	2,085	13,863	22,501	4,941
2019	2,301	25,661	3,204	2,053	2,136	14,135	22,925	5,001

Table 2 – Historical and Forecast Coincident Peak Demands by Weather Zones (MW)

Year	North	North Central	East	Far West	West	South Central	Coast	South
2002	9,852	96,765	11,544	10,345	7,349	45,082	79,026	20,810
2003	9,836	96,671	11,747	10,199	7,601	45,797	82,483	20,649
2004	10,196	96,149	11,686	10,154	7,939	46,217	85,698	21,103
2005	10,517	100,639	12,204	10,401	8,128	49,032	86,562	21,772
2006	10,666	102,215	12,324	10,795	8,077	50,956	88,743	21,963
2007	10,575	101,866	12,670	10,907	8,085	51,815	89,864	22,020
2008	10,871	102,967	13,125	11,447	8,336	54,411	88,822	22,459
2009	10,578	102,753	12,647	11,348	8,020	53,675	90,654	22,528
2010	10,467	103,534	12,808	11,290	8,025	54,582	91,776	22,581
2011	10,663	104,781	13,285	11,471	8,249	56,789	94,254	23,006
2012	10,963	106,472	14,072	11,737	8,712	59,593	97,764	23,624
2013	11,198	107,895	14,717	11,954	9,071	62,287	100,538	24,290
2014	11,332	109,042	15,181	12,083	9,285	64,236	102,881	24,754
2015	11,457	110,361	15,579	12,207	9,460	66,035	105,419	25,162
2016	11,612	111,659	16,000	12,351	9,670	67,871	108,092	25,579
2017	11,773	112,922	16,382	12,488	9,883	69,642	110,556	25,944
2018	11,913	114,128	16,820	12,609	10,144	71,280	112,856	26,290
2019	12,075	115,237	17,221	12,731	10,395	72,727	114,998	26,622

Table 3 – Historical and Forecast Energy by Weather Zones (GWh)

APPENDIX 3: METHODOLOGY

A Modified Approach to Long-Term Load And Energy Forecasting: Its Uses In An ISO's Environment For Resource Adequacy And Transmission Planning

Introduction

The main focus of this paper is the benefits of a modified approach to long-term demand and energy forecasting model in an ISO's setting. The forecasts that were produced by a regression model are input into several planning processes that are important in the long-term planning of an electrical grid. The development of this forecasting methodology was designed to address the needs for forecasts in several processes. The load forecasting methodology that was adopted is discussed and its results are outlined. The objective of this methodology is to determine a long-term view of the peak demands that ERCOT (total load served in the ERCOT region including exports across DC ties and excluding private use network loads) can expect to face, in order to secure sufficient resources in the next five to ten years. The discussion covers the success experienced in using this methodology and details of the process involved in producing the forecasts. More specifically, this paper details:

- A methodology developed specifically for ERCOT to meet its specific needs.
- How the methodology chosen has been used to successfully meet ERCOT's planning objectives.

Why it is needed

The development of a long-term trend outlook uses a regression model that forecasts peak demands that are most likely to occur under normal weather conditions to determine the approximate timing for scheduling the building of transmission lines to balance the supply and demand for electric power in the ERCOT electrical grid. The load forecast is an input to the reserve margin calculation. As such, the load forecast is a key component necessary for meeting this objective, which is used to ensure a balanced system.

A resource adequacy assessment begins with the calculation of a reserve margin as,

Reserve margin= ((Resources – Firm Load Forecast)/Firm Load Forecast) *100

This calculation is the foundation of the process for determining the adequacy of the system. The review of resource adequacy is an annual process that ensures that enough resources will be available to meet demand in the medium-to long-term time frame.

The forecast is also used in the medium-range planning of resources by the outage coordinators to schedule plant outages for the next year.

Another aspect of system adequacy, where the load forecast plays an important role, entails performing a load sensitivity assessment. This assessment is related to the risk associated with the volatility of the load due to weather. The 90% approximate forecast limits due to the volatility associated with forecasting the load, using temperatures at the 90th percentile of the distribution, are calculated for the next ten to fifteen years to assess the risks of extreme weather volatility on the peak demands. These load volatility estimates are an input into the loss-of-load-probability studies (LOLP), which are used to determine the target reserve margin.

Reviews of the reserve margin to ensure its adequacy are performed every few years through a LOLP study. In this study, expected load, load forecast error, the load volatility due to weather, generation fleet, maintenance schedules, and unit forced outage rates are input into a unit commitment and dispatch model in order to simulate the interrelationships between these variables over a number of replications. This simulation yields an expected un-served energy value. Then, the target reserve margin is obtained by finding the minimum point of the intersection where the LOLP is the ERCOT/NERC standard of one event every ten years.

Load volatility estimates derived from the load forecast are also used by NERC in the summer and winter reliability assessments. These load forecasts feed into the reporting requirements of FERC 714.

The long-term hourly load forecast by weather zones also serves an important function in performing economic analyses. It is an input to the UPLAN software which determines whether or not to undertake transmission projects.

As described above, the load forecast is a major input to several planning processes. The long-term forecast can affect the adequacy of the system grid. Some of the consequences of load forecast errors and their impact on system adequacy can be:

- Building excessive additional generation capacity and/or transmission facilities
- Inadequate levels of resources and generation leading to blackouts and price spikes
- Sending incorrect signals to the market regarding the value of ancillary payments and energy

Finally, the energy consumption forecast provides the means to determine the annual $/MWh ERCOT fee for the annual budget review, conducted by the Texas PUC.

Availability of methods

There are a wide variety of methods that can be used to forecast system peak and energy consumption. Such methods range from simple trending methods to more complex ones such as end-use forecasting or hybrid end use and econometric techniques, sophisticated Box-Jenkins Transfer function (Dynamic Regression) models and neural network models that can be adapted to produce long-term forecasts

For ERCOT, data requirements were a major determinant of which method was feasible and appropriate to implement. There were specific requirements to be met in terms of the end product. The following describes the specific nature of these data needs.

Forecast Level of Detail

An hourly forecasted load shape by weather zones for the next five to ten years was needed as an input into UPLAN for economic analysis of transmission projects. The hourly loads from the load shape, combined with the results of a monthly energy forecast, were considered a feasible way to produce a system peak forecast for each year in the five-to-ten-year horizon. The system peaks and energy consumption forecasts were thought to be a high priority for this important process as these forecasts could as well be used as inputs into the resource adequacy process.

Load and Weather Data level of Detail

ERCOT Staff decided to produce long-term forecasts for eight major areas in Texas where weather data was available and coincided with the available data appropriate for load analysis. Thus, from ERCOT‘s standpoint, weather zones were the logical choice. In addition, these zones also coincided with the major areas of interest for the analysis of transmission projects. In summary, the total load by weather zone was chosen as meeting the objective of the forecast needs. These forecasts then could be aggregated to a system level.

Economic, Demographic and Price Data Level of Detail

Besides hourly load, ERCOT also secures weather data, economic and demographic data from outside providers. In regard to prices, which are considered an important driver for inclusion in a demand equation, it is not clear as to whether the wholesale prices that ERCOT collects are really the most relevant for a forecasting application, in terms of being the prices ultimately faced by the consumer. Since the wholesale prices are collected on an hourly basis, and retail prices are better reflected by an average over a longer time period, such as a month, wholesale hourly prices do not capture the correlation with the MWh consumption correctly. Several attempts to include market clearing prices of energy (MCPE s) in the forecasting models were made but were unsuccessful. The models obtained showed price to be insignificant or to indicate a nonsensical relationship regarding the direction of the effect of price (wrong sign on the coefficient) and thus should not be included in a long-term demand equation. To make matters more challenging in this respect, an objective and credible forecast of these prices would represent a major accomplishment in itself. Inclusion of a price variable in the forecasting models could potentially provide a means to calculate an unbiased and credible forecast of the price effect on the long-term load response.

Method Selection

There is no single best forecasting method. The choice of a forecasting method in this case was based on the specific circumstances of the situation being faced. Given the requirements at the time, in terms of available data, the capabilities needed of any chosen method, and the intended use of the resulting forecasts, a regression with capabilities of performing a correction for autocorrelated errors was deemed as the most appropriate choice available to meet ERCOT's objectives. This methodology is unique in that it directly and successfully forecasts an hourly load shape using a regression model estimated by seasons. This methodology could potentially be applied to other entities facing similar requirements.

Forecast Process --- General Description

The forecast process starts with the development of regression equations from historical data for demand peaks and energy. These use the following input drivers:

Trend Variables
- Population
- Income
- Economic

Calendar Variables
- Seasonal Variation
- Daily Variation
- Weekly Variation
- Holidays

Weather profiles from actual data that use an average representation of weather not prediction of weather
- Temperature
- Humidity
- Cooling Degree Days (CDD)
- Heating Degree Days (HDD)

The results are forecasts for energy and peak.

The data used to prepare the forecast came from the following sources:

1. Economic Data
   - Economic data obtained from Economy.com
   - Data includes economic and demographic data (such as income, employment, housing permits, GDP, population and migration patterns) for Texas at the state, county, metropolitan statistical areas (MSAs). Some of these data is also available at the national level

2. Weather Data
   - Ten years of weather data obtained from Weather Bank for 20 weather stations

- The data is first weighted by individual weather stations using ERCOT's standard factor, and then for the total system using weights proportional to the load in each weather zone

3. Load Data
   - Settlement load data available on an hourly basis since July 31, 2001
   - Prior to 2001, we have Transmission and Distribution Service Providers (TDSP) hourly data

The weather data is used in the development of weather normalized profiles by weather zone and is accomplished by calculating the normalized temperature profile by weather zone. The weather profiles use the rank-average method which involves the following steps:

1) Rank the hourly temperatures for each year for each weather zone from highest to lowest
2) Determine the median temperature from all years for every hour
3) Calculate the sum of the absolute values of the difference of the median and the hourly temperatures for all hourly temperatures in each year
4) Determine the year with the minimum summed value and select this year as the typical year profile
5) Use this year's profile to re-sort the median temperatures

A major issue in the preparation of the long-term forecast relates to the variable selection process. The process in this case generally entails performing the following analyses with the following considerations:

- Multiple regression analysis is used to develop the forecasting equations
- Initial selection of variables comes from a variation of the stepwise procedure using a combination of the Least Absolute Shrinkage and Selection Operator (LASSO) and the Least Angle Regression (LAR) to determine those that were the most statistically significant
- A methodical process and pre-specified strategy of selecting a subset of those variables using empirical results and informed judgment
- Variables selected for inclusion had to meet the following: 1) justifiable on a logical basis , 2) historically measurable and 3) must have an available forecast
- Ordinary least squares techniques with models that can selectively include autoregressive error terms, are used to calculate the appropriate coefficients on each variable and to choose the best equations

Load shape and Energy forecasts were developed from monthly energy and hourly load shape equations for each season of the following form:

- The general formulation of the energy equations include the following variables:

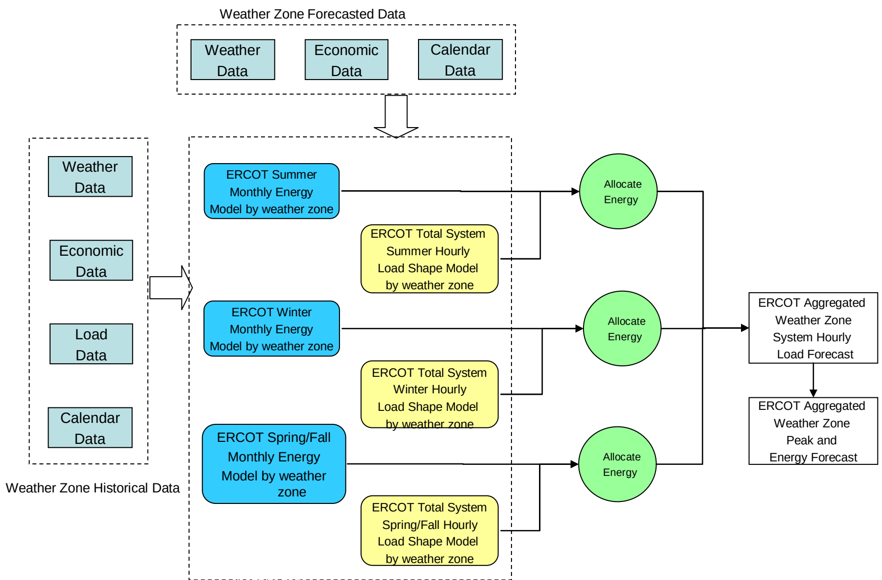
  Energy Month i = f {CDD, HDD, Income, Population, Employment, GDP, Monthly Indicators, AR terms}

- The general formulation of the load shape equations include selected variables from some of the following:

  Load hour i =f {Max Temps, Lagged Temps, Heat Index, Non-Linear Temp Components (square and cube), Temp Gains (diff between daily high and low temps), Temp Build-up, Dew Point, Month*Temp Interactions, CDD, HDD, Hour of Day Indicators, Weekday/Weekend, Holidays, AR terms}

Putting it all together

Weather Zone Forecasting Process



The Weather Zone forecasting process flow is as follows:

1. Obtain weather and economic variables by weather zone (historical and forecast)
2. Develop regression equations by weather zone describing the historical actual:
--Monthly Energy
* Using a different equation for each season
-- Hourly Load Shape
* Using a different equation for each season or a single model for all seasons
3. Incorporate forecasted values of economic and normalized temperatures for 2008-2019 by weather zone into monthly energy equation to produce forecasted monthly energy
4. Incorporate normalized temperatures for 2008-2019 by weather zone into monthly load shape equation to produce forecasted load shape
5. Produce hourly demand forecast by weather zone by fitting forecasted monthly energy under projected hourly load shape

Hourly Forecast

The calculation of an hourly forecast is a result of the process described above and yields the following results:

- The forecasted hourly shape from the load shape equations is scaled to produce the final hourly forecast
  - Each hour's load is scaled so that the amount of energy under the load shape for a month is equal to the amount of energy projected for that month by the energy forecast from the energy equations
  - The percent of a month's energy that is contained in each hour from the load shape equation is maintained
- The peak forecast is the highest hourly load from this final hourly forecast

Mathematical/statistical rigor

(A) Derivation:

There are instances in which the models may require to perform a correction for auto correlated error terms. The mathematical/statistical intricacies of the models are presented below. The peak demand forecasts are obtained by combining the results of two models: an hourly model that forecasts the load shape and a monthly energy forecast which includes economic and demographic variables to determine the long-term trend. The hourly load shape model is of the following form:

$$Y_t = \alpha_o + \sum_{i=1}^{23} \beta_i HR_{i,t} + \sum_{i=1}^{n} \gamma_i W_{i,t-s} + \sum_{i=1}^{n} \Omega_i DT_{i,t} + \sum_{i=1}^{n} \Theta_i WI_{i,t} + \sum_{i=1}^{n} \delta_i SV_{i,t} + \sum_{i=1}^{n} \nu_i \mathrm{E}_{i,t} + \frac{\varepsilon_{i,t}}{\left(1 - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p\right)}$$

Where:

- Y_t is the hourly load (MW)
- $HR_{i,t}$ are hourly indicator variables
- $W_{i,t-s}$ are weather variables and their lags
- $DT_{i,t}$ are day type variables
- $WI_{i,t}$ are weather interaction variables
- $SV_{i,t}$ are sunlight variables
- $E_{i,t}$ are special events variables
- $\varepsilon_{i,t}$ is a random error term
- $\Phi' s$ are autocorrelation terms specified with a lag (backshift) operator, $L^s = \mathrm{X}_{t-s}$

This model specified in mathematical form can be generalized as follows:

$$Y_t = \beta_o + \sum_{i=1} \beta_\kappa \mathrm{X}_{\kappa,t} + \frac{\varepsilon_t}{\Phi\left(L\right)}$$

Where:

$\beta_o, \beta_{1,} \ldots, \beta_\kappa$ = coefficients to be estimated

$\mathrm{X}_{\mathrm{K,t}} = {}_{\mathrm{K}}$ regressor variables, K=l, …, m

ε_t = a random error term

Φ (L) = an autoregressive structure of order ρ where ρ = 24 or an AR(ρ) process

" $= \left(1 - \Phi_1 L - \Phi_2 L^2 - \Phi_3 L^3 - \ldots - \Phi_p L^p\right)$

Φ_j = autoregressive coefficients

L^j = Lag operator, $L^j = X_{t-j}$

Thus, the model to be estimated can be derived as follows:

(1) $$\Phi(L)Y_t = \Phi(L)\beta_o + \Phi(L)\sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t} + \varepsilon_t$$

Where the constant term $\alpha_o = \Phi(L)\beta_o$.

Expanding the expression on the right hand side,

$$\Phi(L)\sum_{\kappa=1}^{m}\beta_{\kappa,t} = \left(1 - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p\right) - \left(\beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_m X_{mt}\right)$$ and gathering common terms together we obtain

$$\Phi(L)\sum_{\kappa=1}^{m}\beta_{\kappa,t} = \sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t} - [\ \Phi_1 + \Phi_2 + \ldots + \Phi_p\ ]\bullet\sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t-p}$$

Or more succinctly,

(2) $$\Phi(L)\sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t} - \sum_{j=1}^{p}\Phi_j\sum_{\kappa=1}^{m}\beta_{\kappa,t-j}.$$

The expression on the left hand side of the equation is

$$\Phi(L)Y_t = \left(1 - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p\right)Y_t$$
$$\Phi(L)Y_t = Y_t - \Phi_1 Y_{t-1} - \Phi_2 Y_{t-2} - \ldots - \Phi_p Y_{t-p}$$

Or more compactly stated,

(3) $$\Phi(L)Y_t = Y_t - \sum_{j=1}^{p}\Phi_j Y_{t-j}$$

Substituting (2) and (3) into (1) we get,

$$Y_t - \sum_{j=1}^{p}\Phi_j Y_{t-j} = \gamma_o + \sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t} - \sum_{j=1}^{p}\Phi_j\sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t-j} + \varepsilon_t$$

or

$$Y_t = \gamma_o + \sum_{j=1}^{p}\Phi_j Y_{t-j} + \sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t} - \sum_{j=1}^{p}\sum_{\kappa=1}^{m}\beta_\kappa X_{\kappa,t-j} + \varepsilon_t$$

Where

$$\varepsilon_t = \mu_t - \sum_{j=1}^{p}\Phi_j L^j\mu_{t-j}.$$

(B) Estimation:
In vector notation [1],

$$y_t = x_t'\beta + \mu_t$$

Where $\left(x_t = x_{1t,} x_{2t,} ..., x_{Kt} \right)'$

$$\mu_t = \varepsilon_t + \varphi_1 \mu_{t-1} + \varphi_2 \mu_{t-2} + ... + \varphi_p \mu_{t-p}$$ [2]

And $\varepsilon_t = N\left(o, \sigma^2\right)$, normally and independently distributed with mean o and variance of σ^2

y_t	=	dependent values
x_t'	=	a column vector of regressor variables
β	=	a column vector of structural parameters

The autoregressive parameter vector, $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_p)'$ and its variance covariance matrix:

$$\mu = \left(\mu_{1,} \mu_{2,...,} \mu_p\right)' = \Sigma$$
$$\mathrm{E}\left(\mu \mu'\right) = \Sigma \sigma^2 U$$

Since the stepwise-like procedure BACKSTEP is specified for testing the statistical significance of the φ's, the TOEPLITZ matrix is used, with the (i,j)th

element $\gamma_{|i-j|}$ is equal to $R\hat{\varphi} = r$

Where r = $\left(r_1, r_2, ..., r_p\right)'$ and r_i is the lag i sample autocorrelation. The matrix [R, r] is treated as sum-of-squares cross products matrix coming from a simple regression using N-K observations, where K = number of estimated parameters.

This method of estimation is known as the Yule-Walker (YW) method. It alternates the estimation of β using generalized least squares (GLS) with the estimation of the φ's using the YW equations applied to the sample autocorrelation function (SA).

The steps are:

1) Form OLS estimates of β.
2) Estimate φ from the SAC function of the OLS residuals using the YW equations.
3) Estimate U from the estimate of φ and Σ from U and the OLS estimate of σ^2.

---

[1] This material comes from the SAS Autoreg Procedure in the ETS manual.

[2] SAS parametization computes the signs of the autoregressive parameters reversed from what is presented in most of the literature. The parametization shown here is in agreement with most of the literature.

The second model forecasts the long-term trends in energy consumption (MWh) utilizing economic, demographic, weather, and season variables and possibly autoregressive terms. The form of the model is as follows:

$$Y_t = \beta_o + \sum_{i=1}^{n} \gamma_i CDD_{n,t} + \sum_{i=1}^{n} \Theta_i HDD_{n,t} + \sum_{i=1}^{s} \delta_i \mathrm{E}_{it} + \sum_{i=1}^{11} \alpha_i m_{it\,it} + \varepsilon_{i,t}$$

Where: Y_t = Monthly energy consumption (MWh)
$CDD_{n,t}$ = Cooling Degree Days (n terms using different basis)
$HDD_{n,t}$ = Heating Degree Days (n terms using different basis)
E_{it} = Economic and Demographic variables
m_{it} = Monthly indicator variables

$\varepsilon_{i,t}$ is a random error term

This model represented in general form is as follows:

$$Y_t = \beta_o + \sum_{i=1}^{p} \beta_\kappa \mathrm{X}_{\kappa,t} + \frac{\varepsilon_{i,t}}{\left(1 - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p\right)}$$

Where,

$\beta_o, \beta_1, \ldots, \beta_p$ = coefficients to be estimated

$\varepsilon_{i,t}$ is a random error term

$\Phi's$ are autocorrelation terms specified with a lag (backshift) operator,

$L^s = \mathrm{X}_{t-s}$

This energy equation is estimated using the Yule-Walker method as described above..

(C) <u>Allocation of Energy Under Load Shape:</u>

Let $Y_{LSi,t}$ = hourly load shape forecast from the first model,

Y_{Et} = monthly energy forecast from the second model,

Then, the long-term load forecast is obtained as follows:

$$Y_{LSi,t} = Y_{LSj} \cdot \frac{\Sigma Y_{Et}}{\Sigma Y_{LSi}}$$

Where:

Y_{LSj} is the load at hour j, j=1, …,8760

Thus, the annual system peak demand is obtained as,

$$\text{Y peak} = \max \left\{ Y_{LSit\,;}\, i = 1,\ldots 8760; t = 1,\ldots 12 \right\}$$

Conclusions-- Forecast Performance, Results, Findings and Properties

Model validation using actual temperatures in the forecast period – The validation of the model is done by using the actual temperatures experienced during the year, instead of the "50-50 normal profile" temperatures that were used to produce the forecast. The forecasting model is estimated with the same data used in the forecasting process and with the same mix of variables as originally formulated for each equation.

The result of this validation reveals the forecasting error due to the inaccuracy of the model itself and its formulation (misspecification, incorrect functional form, irrelevant variables, lacking important variables, etc) the error in forecasting the independent variables that serve as drivers, except for actual temperatures which reflect the exact temperatures that produced the loads. Thus, this is way to take out the effects of weather to evaluate the accuracy of the model and other input variables.
.
The forecasting model can be used to perform weather scenarios by looking at 90th percentile temperatures (90-10). Thus, it can be used to look at load volatility using the model with a wide variety of weather profiles – including extreme weather profiles.

There are strengths and weaknesses associated with the process described in this paper. They are:

ERCOT's model strengths

- The methodology is statistical and mathematical in nature, but it still allows for judgment to be incorporated into the results by selecting variables that contribute to the generation of a forecast that passes, not only statistical tests, but common sense criteria.
- This approach was implemented in an automated fashion using macro routines in SAS. With so many models to maintain (8 zones * 3 seasons per zone = 24 models total), it is advantageous to have the ability to make changes and produce normal or extreme weather or any other type of forecasts very quickly.
- The chosen methodology remains consistent in the face of changes in the structural pattern of new incoming data. This is an indication of the robustness of the approach and the model.

ERCOT's model weaknesses

- The initial set-up for the infrastructure for using this approach is time consuming and complex.
- The model was developed from a top-down approach analyzing total ERCOT (system) load. Thus, it does not allow analysis at a more disaggregated level such as focusing at the class level, i.e., residential, business commercial, large industrial customers, etc.

An important aspect associated with any forecasting model is the robustness of its forecasts. Another related consideration is whether these forecasts can be considered reliable enough to lend the model some credibility. In this case, there are forecasts produced with a very similar model for 2005, using the same methodology but, with system load data instead of disaggregated data for weather zones. The model presented here aggregates across zones can be used to obtain the system peak. The results produced by the model for 2005 are very similar in terms of the magnitude of the percent forecast errors. The overall error was between 0 and + 0.5%. This pattern of successful forecasting gives this methodology some credibility and shows its robustness.