Load Forecasting Model Whitepaper

Resource Adequacy Planning Department

PJM Interconnection

Last Updated: April 27, 2016

Contents

Figures

Equations

Tables

I. INTRODUCTION

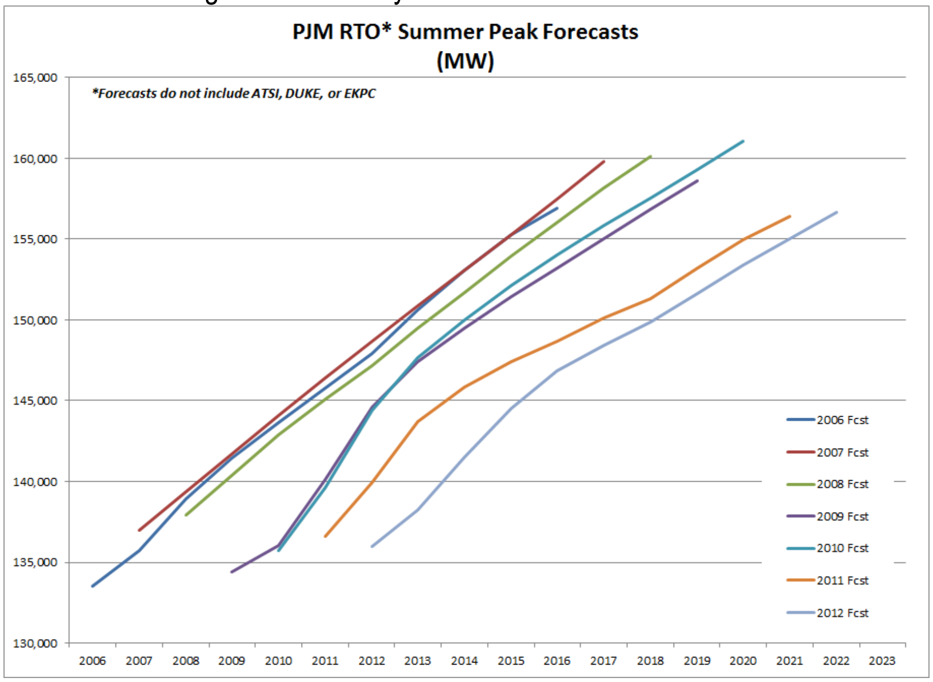
This document discusses the impetus for, and development of, significant changes to the load forecasting models maintained by the PJM Interconnection. These changes were implemented with the release of the 2016 PJM Load Forecast Report. It is intended to serve as documentation of the implemented peak and energy forecast models. Its intended audience is members of the PJM Load Analysis Subcommittee and the Planning Committee.

II. BACKGROUND OF FORECAST INVESTIGATION

The PJM Interconnection established an independent load forecast, produced by PJM staff, with the release of the 2006 PJM Load Forecast Report. The initial PJM load forecast model was an econometric model that produced estimates of non-coincident and coincident peak loads for each PJM zone, locational deliverability area (LDA) and the RTO. It used local economic activity, weather, and day-type variables as explanatory variables/drivers. Weather data and economic data and forecasts were procured from outside vendors. The model featured simulation of historical weather patterns and regional diversities to develop a distribution of forecasts which were then used to produce monthly and seasonal forecasts across a range of weather conditions. In ensuing years, PJM made numerous changes to the model based on internal review, stakeholder input and recommendations from outside consultants (e.g., adding a net energy for load forecast, extending the forecast horizon from ten to fifteen years, directly modeling coincident peak, adopting an indexed economic variable with six components, etc.), but the basic structure of the model remained consistent.

Beginning with the Great Recession of 2007-2009, the accuracy of the PJM model decayed noticeably, with a trend towards over-forecasting. Each successive load forecast tended to be lower for a given year than the forecast produced one year prior, as illustrated in Figure II-1 below. Initially, the majority of the inaccuracy could be attributed to revisions to the economic forecast used to develop the PJM load forecast. While the U.S. economy recovered at a slow and unsteady pace from the recession, successive economic forecasts continued to predict a period of robust growth in the near-term followed by a resumption of trend growth in the long-term. Since the economic forecast is acquired from a third party, PJM's ability to correct the bias was extremely limited. But as the economy continued to recover and economic forecasts stabilized, the model's accuracy did not improve commensurately and it became apparent that other factors were also contributing to model error. The declining peak load forecasts had significant impacts: previously approved transmission expansion projects were postponed or canceled and resources were over-procured through the forward capacity market. As a result, by the end of 2012 there were growing calls both internally and externally for PJM to address the issue. The latest efforts are discussed in Section IV, along with commentary on earlier efforts in APPENDIX E and APPENDIX F.

Figure II-1 : History of PJM Peak Load Forecasts



III. FORECAST MODEL STRUCTURE

PJM uses regression models with daily load as the dependent variable and independent variables including calendar effects, weather, economics, and end-use characteristics. The model is estimated over historical data back to 1998 and is used to produce a 15-year forecast for PJM zones, LDAs, and the RTO. LDA and RTO forecasts are produced using a bottom-up approach, forecasting zonal contributions and aggregating.

Dependent Variables

PJM starts with hourly metered load data collected via Power Meter, and then makes adjustments based on estimated load drops[1] and estimated distributed solar generation to obtain hourly unrestricted loads. Individual daily zonal peaks (non-coincident peaks or NCPs) are identified, as well as zonal contributions to LDA and RTO peaks.

In the case of daily energy, hourly metered load is only supplemented with estimated distributed solar generation and not estimated load drops. Daily energy is then the sum of these hourly values.

Independent Variables

Calendar

The forecast model includes a number of variables to capture calendar effects, represented primarily as either binary variables or fuzzy binary variables. Binary variables take a value of 1 or 0, whereas fuzzy binary variables have values ranging from 0 to 1. There is also a graduated variable to take into account the effect of Christmas lights.

Day of Week (=1 when that day, 0 otherwise)

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday

Month (=1 when that month, 0 otherwise)

January	February	March	April	May	June
July	August	September	October	November	

Holiday variables are coded such that they have values for more than one day, and for some holidays these values can differ year to year depending on the day of the week the holiday is observed. These variables are included because generally these days would be expected to have loads below norm.

[1] Load Drops are described in Attachment A of PJM Manual 19.

MLK (Martin Luther King Day), *PresDay* (Presidents' Day), *MemDay* (Memorial Day), and *LaborDay* (Labor Day)

	Value
Day before Holiday	0.2
On Holiday	1
All other days	0

GoodFri (Good Friday) and *Thanks* (Thanksgiving Day)

	Value
On Holiday	1
All other days	0

July4th (Independence Day)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
July 2	0.10	0.00	0.00	0.15	0.15	0.10	0.15
July 3	0.70	0.25	0.15	0.20	0.80	0.20	0.20
July 4	1.00	1.00	0.80	1.00	1.00	0.40	0.30
July 5	0.80	0.15	0.15	0.25	0.70	0.30	0.15
July 6	0.00	0.00	0.00	0.00	0.10	0.20	0.00
All other days	0.00	0.00	0.00	0.00	0.00	0.00	0.00

FriAThanks (After Thanksgiving Day)

	Value
On Holiday	1
Day After	0.2
All other days	0

XMasWkB4 (Week before Christmas Day)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
December 21	0.33	0.33	0.33	0.50	0.50	0.50	0.33
December 22	0.50	0.50	0.67	0.67	0.80	0.50	0.50
December 23	1.00	0.67	0.67	1.00	1.00	0.67	0.67
All other days	0.00	0.00	0.00	0.00	0.00	0.00	0.00

XmasEve (Christmas Eve)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
December 24	1.00	1.00	0.80	0.67	1.00	0.50	0.33
All other days	0.00	0.00	0.00	0.00	0.00	0.00	0.00

XMasDay (Christmas Day)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
December 25	1.00	1.00	1.00	1.00	1.00	0.50	0.50
All other days	0.00	0.00	0.00	0.00	0.00	0.00	0.00

XMasWk (Time around Christmas and New Year's)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
December 26	1.00	0.67	0.67	0.67	1.00	0.20	0.25
December 27	0.25	0.33	0.33	0.33	0.50	0.20	0.15
December 28	0.25	0.33	0.33	0.33	0.33	0.20	0.15
December 29	0.33	0.33	0.33	0.33	0.33	0.20	0.15
December 30	0.80	0.50	0.33	0.50	0.50	0.25	0.25
January 2	0.80	0.15	0.33	0.33	0.67	0.25	0.15
January 3	0.00	0.15	0.00	0.15	0.15	0.15	0.15
January 4	0.00	0.00	0.00	0.00	0.15	0.00	0.00
All other days	0.00	0.00	0.00	0.00	0.00	0.00	0.00

NYEve (New Year's Eve)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
December 31	0.8	0.8	0.8	0.8	1.0	0.4	0.4
All other days	0.0	0.0	0.0	0.0	0.0	0.0	0.0

NYDay (New Year's Day)

	Value by Day of Week						
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
January 1	1.0	1.0	1.0	1.0	1.0	0.5	0.4
All other days	0.0	0.0	0.0	0.0	0.0	0.0	0.0

DLSav_EPA2005 (Daylight Savings Time Indicator)

	Value
Within Daylight Savings Time Period	1
All other days	0

XMasLights (Christmas Lights Indicator)

This is a graduated variable that starts at 1 on the Friday after Thanksgiving, and increases by 1 each day until December 23.

Economics

PJM uses a collection of economic variables to represent different sectors (Residential, Commercial, and Industrial). Variables are combined into a single Economic Index using a weighting scheme that reflects each sector's share of a zone's total electricity sales. Moody's Analytics provides the historic and forecast data for all economic variables.

Each zone is assigned one or more metropolitan statistical areas (MSAs) [2] that represent economics in its footprint[3]. Each individual variable is summed across MSAs to derive the zonal figure to be used in later calculations.

Residential	Commercial	Industrial
Households	Non-manufacturing Employment	Gross Domestic Product
Population	Gross Domestic Product	Gross Metropolitan Product
Real Personal Income	Gross Metropolitan Product	
	Population	

Each variable is then divided by its value in 1998 to convert each into an index. This conversion is necessary because the variables are expressed in different terms, with some being count-based measures (i.e. population) while others are dollar-based (i.e. Gross Domestic Product). Three sector economic indexes are then constructed based on survey-driven weights[4] (see Equation III-1).

[2] The lone exception to the use of metropolitan statistical areas is the Dominion zone. Statewide economics are used in its case.
[3] The MSAs used for each zone are listed in Manual 19
[4] These weights originate from a 2010 forecasters survey by Itron which ascertained the relative importance of each variable to sector energy sales.

Equation III-1 : Sectoral Economic Indexes

$$\mathbf{ResEconIndex_{zone,t} = (HH_{zone,t}/HH_{zone,1998})^{0.47} \times (Pop_{zone,t}/Pop_{zone,1998})^{0.26} \times (PInc_{zone,t}/PInc_{zone,1998})^{0.27}}$$

$$\mathbf{ComEconIndex_{zone,t} = (NMEmp_{zone,t}/NMEmp_{zone,1998})^{0.47} \times (GDP_{zone,t}/GDP_{zone,1998})^{0.20} \times (GMP_{zone,t}/GMP_{zone,1998})^{0.16} \times (Pop_{zone,t}/Pop_{zone,1998})^{0.17}}$$

$$\mathbf{IndEconIndex_{zone,t} = (GDP_{zone,t}/GDP_{zone,1998})^{0.47} \times (GMP_{zone,t}/GMP_{zone,1998})^{0.53}}$$

Where
HH = Households
Pop = Population
PInc = Real Personal Income
NMEmp = Non-manufacturing Employment
GDP = Gross Domestic Product
GMP = Gross Metropolitan Product

The three individual sector economic indexes are then combined into a single economic index for each zone, using each sector's share of the zone's total electricity sales over a 5-year period from FERC Form 1 (see Equation III-2)[5]. Weights are published annually with the Load Forecast Statistical Appendix.

Equation III-2 : Aggregate Economic Index

$$\mathbf{EconIndex_{zone,t} = ResWt \times ResEconIndex_{zone,t} + ComWt \times ComEconIndex_{zone,t} + IndWt \times IndEconIndex_{zone,t}}$$

Where
ResWt = (Residential Sales)/(Total Sales)
ComWt = (Commercial Sales)/(Total Sales)
IndWt = (Industrial Sales)/(Total Sales)

The economic index is then used in several locations in the forecast model to help the model explain different phenomena, namely, base load growth and the relationship of economics with weather sensitive energy demand.

End-Use Characteristics

The load forecast model includes variables to capture trends in end-use characteristics of equipment saturation and efficiency. Terms, analysis, and index construction are described in detail in Section IV - Equipment/Appliance Saturation and Efficiency. The result is three equipment indexes that describe different activity types: Heating, Cooling and Other. Each index is a weighted combination of various equipment types across the Residential and Commercial sectors. These variables are then used in several locations in the forecast model as laid out in Equation IV-4, Equation IV-5, and Equation IV-6.

[5] Data for EKPC is collected from the Kentucky Public Service Commission (http://psc.ky.gov/).

Weather

The forecast model includes several different variables to capture the impact of weather on load across the seasons. Weather variables are specified as splines over defined ranges, to allow for different relationships of load to weather depending on conditions.

For the winter and colder periods in the shoulder months, wind-adjusted temperature (or oft-called winter weather parameter (WWP)) is used as the weather parameter. For the summer and hotter periods in the shoulder months, the temperature-humidity index (THI) is used as the weather parameter. These concepts are defined in Equation III-3.

Equation III-3 : Weather Parameters (WWP and THI)

Wind-Adjusted Temperature

WWP = Temp – (0.5 x (Wind – 10)), if Wind > 10
WWP = Temp, if Wind <= 10

Where
Wind = Wind velocity in MPH
WWP = Wind-Adjusted Temperature
Temp = Dry bulb temperature

Temperature-Humidity Index

THI = Temp – 0.55 x (1 – Hum) x (Temp – 58), if Temp >= 58
THI = Temp, if Temp < 58

Where
THI = Temperature-Humidity Index
Temp = Dry bulb temperature
Hum = Relative Humidity (where 100% = 1)

These weather concepts are then used to create four-section splines each for the summer[6] (May – September) and winter (January, February, and December) seasons. The shoulder months use a combination of WWP and THI to reflect the contrasting weather patterns that these periods contain. These combinations are defined and described in detail in Section IV - Refined Weather Treatment.

In addition to the seasonal specific splines, CDD, HDD and one day lags of both terms are used in the forecast model as well. These variables are used year-round, and are defined in Equation III-4.

[6] This definition of summer for THI differs from that used in the Load Forecast to define summer, which includes only June through August. This is because while May and September are noticeably milder than June through August and very unlikely to contain summer peaks, they do share similar weather characteristics.

Equation III-4 : Weather Parameters (CDD and HDD)

$$\mathbf{CDD = Maximum\ (Avg\_Temp - 65\ ,\ 0)}$$
$$\mathbf{HDD = Maximum\ (60 - Avg\_Temp\ ,\ 0)}$$

Where
Avg_Temp = Daily Average Temperature

Load Adjustments

Load adjustments are variables introduced to individual zones to reflect a change that has already occurred that is not adequately captured by the forecast model. Examples of situations potentially requiring a load adjustment would be a large customer shutdown/addition or service territory shift in a zone. These adjustments are intended to cover sudden shifts that the load forecast model would otherwise overlook. Load adjustment variables are a 1 for a specified time period and a 0 otherwise.

Autoregressive Error Term

The load forecast employs an autoregressive model for the errors with a one period lag, or an AR(1) error structure. This is a statistical method to account for errors being correlated with a lag of themselves. The rationale and investigation into the use of an autoregressive term is detailed in Section IV. Autoregressive error parameter definitions can vary depending on the software being used. PJM uses SAS software to develop the load forecast model, and as such the coefficient on the AR(1) parameter is defined in Equation III-5.

Equation III-5 : Autoregressive Error Term Specification

$$\mathbf{OrigError_t = NewError_t - \varphi \times OrigError_{t-1}}$$

Where
OrigError = Model Error prior to AR(1)
New Error = Model Error after AR(1)
φ = AR(1) coefficient

Model Summary

The following tables are a summary of the dependent and independent variables used in the load forecast model. This is intended as a quick reference guide with variable names as they are listed in the Statistical Appendix released with each annual Load Forecast.

Dependent Variables

Variable	Description
ENERGY	Daily Energy (GWh)
NCP	Daily Non-coincident Peak (MW)
CP_RTO_EKPC	Daily contribution to PJM Peak (MW)
CP_RFC_ATSI_DUKE	Daily contribution to RFC portion of PJM Peak (MW)
CP_WEST_EKPC	Daily contribution to PJM West Peak (MW)
CP_PJM_MA	Daily contribution to PJM Mid-Atlantic Peak (MW)
CP_MA_CENT	Daily contribution to PJM Central Mid-Atlantic Peak (MW)
CP_MA_EAST	Daily contribution to PJM Eastern Mid-Atlantic Peak (MW)
CP_MA_SOUTH	Daily contribution to PJM Southern Mid-Atlantic Peak (MW)
CP_MA_WEST	Daily contribution to PJM Western Mid-Atlantic Peak (MW)
CP_PJM_SOUTH	Daily contribution to PJM South Peak (MW)
CP_GPU	Daily contribution to FE-East Peak (MW)
CP_PLGRP	Daily contribution to PLGRP Peak (MW)

Independent Variables

Variable	Description
Intercept	Model Intercept term
Monday	Daily binary
Tuesday	Daily binary
Wednesday	Daily binary
Thursday	Daily binary
Friday	Daily binary
Saturday	Daily binary
MLK	Holiday fuzzy binary
PresDay	Holiday fuzzy binary
GoodFri	Holiday binary
MemDay	Holiday fuzzy binary
July4th	Holiday fuzzy binary
LaborDay	Holiday fuzzy binary
Thanks	Holiday binary
FriAThanks	Holiday fuzzy binary
XMasWkB4	Holiday fuzzy binary
XMasEve	Holiday fuzzy binary
XMasDay	Holiday fuzzy binary
XMasWk	Holiday fuzzy binary
NYEve	Holiday fuzzy binary
NYDay	Holiday fuzzy binary
XMasLights	Graduated variable
January	Monthly binary
February	Monthly binary
March	Monthly binary
April	Monthly binary
May	Monthly binary
June	Monthly binary
July	Monthly binary
August	Monthly binary
September	Monthly binary
October	Monthly binary
November	Monthly binary
DLSav_EPA2005	Daylight Savings binary
Heat_IN2_HDD	Heating Trend variable: Heating Equipment Index x Economic Index x Heating Degree Days
Cool_IN2_CDD	Cooling Trend variable: Cooling Equipment Index x Economic Index x Cooling Degree Days
Heat_IN2_Lag1HDD	One-Day Lag Heating Trend variable: Heating Equipment Index x Economic Index x Heating Degree Days (-1)

Cool_IN2_Lag1CDD	One-Day Lag Cooling Trend variable: Cooling Equipment Index x Economic Index x Cooling Degree Days (-1)
S1_THI	Temperature-Humidity Index Spline 1 (in effect months 5-9)
Cool_S2_THI	Cooling Equipment Index x Temperature-Humidity Index Spline 2 (in effect months 5-9)
Cool_S3_THI	Cooling Equipment Index x Temperature-Humidity Index Spline 3 (in effect months 5-9)
Cool_S4_THI	Cooling Equipment Index x Temperature-Humidity Index Spline 4 (in effect months 5-9)
Heat_S1_WWP	Heating Equipment Index x Wind-adjusted Temperature Spline 1 (in effect months 1, 2, 12)
Heat_S2_WWP	Heating Equipment Index x Wind-adjusted Temperature Spline 2 (in effect months 1, 2, 12)
Heat_S3_WWP	Heating Equipment Index x Wind-adjusted Temperature Spline 3 (in effect months 1, 2, 12)
Heat_S4_WWP	Heating Equipment Index x Wind-adjusted Temperature Spline 4 (in effect months 1, 2, 12)
Heat_Shldr_WAT19_50lt	Heating Equipment Index x Wind-adjusted Temperature (in effect months 3, 4, 10, 11 when less than 50)
Shldr_WAT19_Base	Wind-adjusted Temperature (in effect months 3, 4, 10, 11 when greater than or equal to 50 and less than or equal to 70)
Cool_Shldr_THI	Cooling Equipment Index x Temperature-Humidity Index (in effect months 3, 4, 10, 11 when wind-adjusted temperature is greater than 70)
Other_DailyIN2	Other Equipment Index x Economic Index
LA_YYYY	Load Adjustment identified by year it takes effect, Binary variable
AR1	Autoregressive Error Term

Forecast Simulation – Non-Coincident Peaks

Once the load model is estimated, forecasts for each PJM transmission zone are produced by solving the zonal NCP equations, moving through the year day by day, using forecasted economic variables, equipment indexes, and historical weather patterns for each day.

For forecasting purposes, values for economic variables and equipment indexes are drawn from the forecasts obtained from third party vendors.

To model the most likely weather conditions (often referred to as normal or peak-eliciting weather) a weather rotation technique is used to simulate a distribution of daily load scenarios generated by historical weather observations, representing actual weather patterns that occurred across the PJM control region.

To enhance the simulation process, each yearly weather pattern is shifted by each day of the week moving forward six days and backwards six days, providing 13 different weather scenarios for each historical year. For early January and late December dates, data from the same calendar year is applied. Table III-1 below illustrates the shift of weather data across the scenarios.

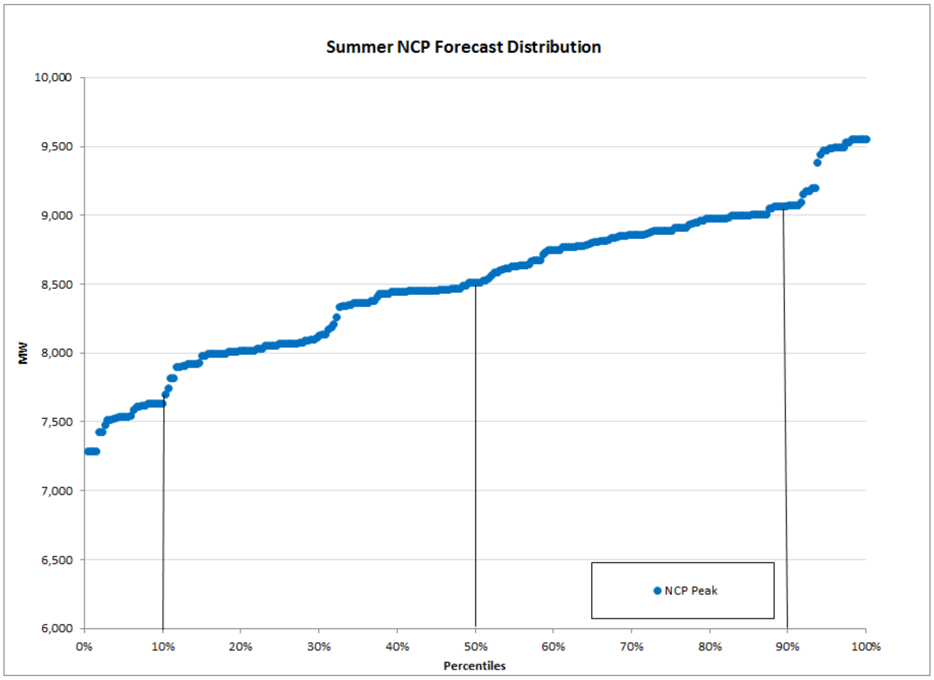
Table III-1 : Mapping of Weather Scenarios to Dates

	Weather Scenarios												
		Rotate Forward						Rotate Backward					
Date	A1995	B1995	C1995	D1995	E1995	F1995	G1995	H1995	I1995	J1995	K1995	L1995	M1995
1-Jan	1/1/1995	1/2/1995	1/3/1995	1/4/1995	1/5/1995	1/6/1995	1/7/1995	12/31/1995	12/30/1995	12/29/1995	12/28/1995	12/27/1995	12/26/1995
2-Jan	1/2/1995	1/3/1995	1/4/1995	1/5/1995	1/6/1995	1/7/1995	1/8/1995	1/1/1995	12/31/1995	12/30/1995	12/29/1995	12/28/1995	12/27/1995
3-Jan	1/3/1995	1/4/1995	1/5/1995	1/6/1995	1/7/1995	1/8/1995	1/9/1995	1/2/1995	1/1/1995	12/31/1995	12/30/1995	12/29/1995	12/28/1995
4-Jan	1/4/1995	1/5/1995	1/6/1995	1/7/1995	1/8/1995	1/9/1995	1/10/1995	1/3/1995	1/2/1995	1/1/1995	12/31/1995	12/30/1995	12/29/1995
5-Jan	1/5/1995	1/6/1995	1/7/1995	1/8/1995	1/9/1995	1/10/1995	1/11/1995	1/4/1995	1/3/1995	1/2/1995	1/1/1995	12/31/1995	12/30/1995
6-Jan	1/6/1995	1/7/1995	1/8/1995	1/9/1995	1/10/1995	1/11/1995	1/12/1995	1/5/1995	1/4/1995	1/3/1995	1/2/1995	1/1/1995	12/31/1995
7-Jan	1/7/1995	1/8/1995	1/9/1995	1/10/1995	1/11/1995	1/12/1995	1/13/1995	1/6/1995	1/5/1995	1/4/1995	1/3/1995	1/2/1995	1/1/1995
-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-
25-Dec	12/25/1995	12/26/1995	12/27/1995	12/28/1995	12/29/1995	12/30/1995	12/31/1995	12/24/1995	12/23/1995	12/22/1995	12/21/1995	12/20/1995	12/19/1995
26-Dec	12/26/1995	12/27/1995	12/28/1995	12/29/1995	12/30/1995	12/31/1995	1/1/1995	12/25/1995	12/24/1995	12/23/1995	12/22/1995	12/21/1995	12/20/1995
27-Dec	12/27/1995	12/28/1995	12/29/1995	12/30/1995	12/31/1995	1/1/1995	1/2/1995	12/26/1995	12/25/1995	12/24/1995	12/23/1995	12/22/1995	12/21/1995
28-Dec	12/28/1995	12/29/1995	12/30/1995	12/31/1995	1/1/1995	1/2/1995	1/3/1995	12/27/1995	12/26/1995	12/25/1995	12/24/1995	12/23/1995	12/22/1995
29-Dec	12/29/1995	12/30/1995	12/31/1995	1/1/1995	1/2/1995	1/3/1995	1/4/1995	12/28/1995	12/27/1995	12/26/1995	12/25/1995	12/24/1995	12/23/1995
30-Dec	12/30/1995	12/31/1995	1/1/1995	1/2/1995	1/3/1995	1/4/1995	1/5/1995	12/29/1995	12/28/1995	12/27/1995	12/26/1995	12/25/1995	12/24/1995
31-Dec	12/31/1995	1/1/1995	1/2/1995	1/3/1995	1/4/1995	1/5/1995	1/6/1995	12/30/1995	12/29/1995	12/28/1995	12/27/1995	12/26/1995	12/25/1995

This approach has two key advantages. One, by rotating the data on the calendar, peak-producing weather will be applied to peak producing days. Two, by producing scenarios over a wide range of weather conditions, the weather rotation method is able to identify both the probable and possible levels of future peak load.

The process is repeated for the remaining years of historical weather data. For example, using twenty years of weather history, this approach will result in 260 (20 weather years x 13 days) separate forecast simulations for each year in the forecast horizon. These simulations produce a frequency distribution of NCP demands by zone, as illustrated in Figure III-1 below.

Figure III-1 : Summer NCP Forecast Distribution



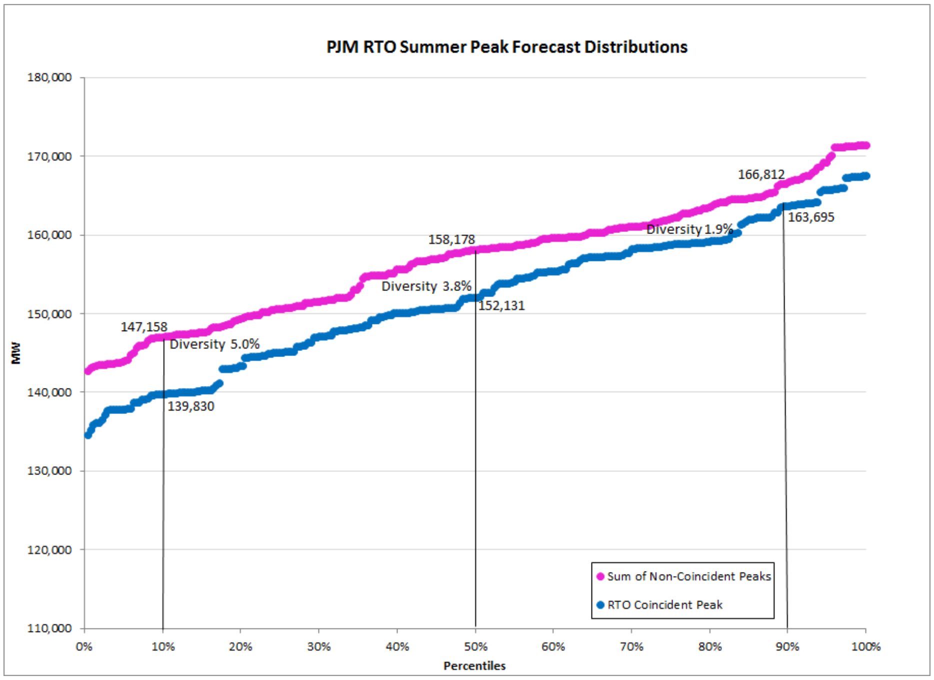
For each weather scenario, monthly NCPs are determined by obtaining the maximum NCP for the month. Seasonal NCPs are determined as the maximum over the summer/winter/spring/fall months. At this point, only monthly and seasonal zonal NCP forecasts are retained. For each season, the ratio of each month's peak to the highest monthly peak is taken, and then each month's ratio is multiplied by the seasonal peak. In this way, one of the month's peaks is the seasonal peak while still preserving the relationship between monthly peaks.

For purposes of system planning, only a couple of the values in the forecast distribution are used. After ranking the scenario forecasts by MW value, the median value is selected as the base (or 50/50) forecast. This is the value used for most system planning studies. The 90th percentile (or 90/10) result is used for studies where the system is assumed to be at system emergency conditions.

Forecast Simulation – Coincident Peaks

To obtain peak forecasts for the entire PJM RTO and Locational Deliverabity Areas (LDA), the forecast simulation and weather rotation method described above is applied to the results of the coincident peak (CP) model equations. Each zone has an RTO CP model; additionally, zones will have CP models for each LDA of which they are a member (e.g., PJM Western Region, EMAAC LDA, ReliabilityFirst region, etc.). In addition to those listed above for NCP forecasting, rotating the weather data on the calendar provides an additional benefit for identifying coincident peaks - the natural diversity of weather patterns that impact the PJM footprint is simulated. This is a more plausible approach to peak forecasting than traditional methods which tend to have all weather stations having peak producing weather on the same day. Using the latter approach would overstate the PJM or LDA peak forecast by minimizing diversity. The impact of modeling diversity is illustrated in Figure III-2 below, which shows the RTO CP forecast compared to the sum of all zones' NCP forecasts.

Figure III-2 : PJM RTO Summer Peak Forecast Distribution



The general decline in expected diversity as forecasts increase is primarily related to weather; lower forecasts are consistent with milder weather conditions which lead to other factors driving the peak which results in greater diversity, while the highest forecasts are consistent with RTO-wide extreme weather which results in less diversity.

To obtain the RTO/LDA peak forecast, the solution for each of the zonal coincident peak (CP) models are summed by day and weather scenario to obtain the RTO/LDA peak for the day. Monthly RTO/LDA peaks are determined by obtaining the maximum of the summed zonal values for the month. Seasonal RTO/LDA peaks are determined as the maximum over the summer/winter/spring/fall months. At this point, the values of the overall RTO/LDA peaks are set, but not the contribution of each zone.

To determine the final zonal RTO/LDA-coincident peak forecasts, a methodology similar to the process for deriving zonal NCPs is applied. By weather scenario, the maximum daily CP load contribution for each zone over the month/season is identified. For each zone a distribution of zonal CP versus weather scenario is developed. From this distribution the median value is selected. The median zonal CPs are summed and this sum is then used to apportion the forecasted RTO/LDA peak to produce the final zonal CP forecasts.

For the RTO/LDA, a distribution of the seasonal RTO/LDA peak vs. weather scenario is developed. For each season, the ratio of each month's peak to the highest monthly peak is taken, and then each month's ratio is multiplied by the seasonal peak. From this distribution, the median result is used as the base (50/50) forecast; the values at the 10th percentile and 90th percentile are assigned to the 90/10 weather bands.

Forecast Simulation – Energy

To obtain forecasts of net energy for load for zones, LDAs, or the entire RTO, the forecast simulation and weather rotation method described above is applied to the results of the energy model equations. The simulation process produces a distribution of monthly forecast results by summing the daily values per forecast year for each weather scenario.

Forecast Adjustments

After completion of the weather rotation process to formulate the distribution of forecast peaks, adjustments are applied to account for events outside of the forecast model and for distributed solar generation[7]. Events outside of the forecast model are future load additions or subtractions that are deemed to not be captured by the forecast process. The procedure to evaluate these forecast adjustments is described in Manual 19 Attachment C.

Distributed solar adjustments are netted off the forecasted load. Recall from Section III - *Dependent Variables* that the dependent variable in the load forecast model is load with an addback for historical distributed solar generation. Thus the resulting load forecast must be reduced by expected future distributed solar generation. This process is explored in depth in Section IV - Distributed Solar Generation.

In summary, the final load forecast by zone and LDA is a function of the forecasted load without distributed solar generation adjusted for existing and future distributed solar generation and other exogenous forecast events as is laid out in Equation III-6.

Equation III-6 : Application of Solar and Forecast Adjustments

Final_Load = Model_Load – Dist_Solar_Forecast + Forecast_Adjustment

Where
Final_Load = Forecasted load (peak or energy) used by PJM Planning and Market functions
Model_Load = Forecasted load (peak or energy) resulting from the econometric model and weather rotation process
Dist_Solar_Forecast = Expected future distributed solar generation (peak or energy)
Forecast_Adjustment = Exogenous plus/minus adjustment for an individual zone

[7] Distributed solar generation are resources that are not interconnected to the PJM markets. These resources do not go through the full interconnection queue process and do not offer as capacity or as energy resources. Furthermore, the output of these resources is netted directly with the load. PJM does not receive metered production data from any of these resources.

IV. 2015 CHANGES

Recognizing the need for better accuracy, PJM sought to make improvements to the load forecast model. Throughout 2015, PJM examined the load forecast model and discussed revisions with the stakeholder community. The most important changes to the model framework were:

- New variables to account for equipment and appliance saturation and efficiency
- A refined weather treatment
- Shorter weather simulation period
- Distributed solar energy resource adjustments

The changes were thoroughly vetted through the PJM stakeholder process and ultimately endorsed by the Markets and Reliability Committee in December, 2015. The changes are documented in the PJM Load Forecasting Manual (M-19) and were implemented for the 2016 Load Forecast. In the following sections, the mechanics of each area of change is described.

Equipment/Appliance Saturation and Efficiency

The most important methodological change was adding variables that capture trends in equipment/appliance saturation and efficiency. Significant customer behavior changes post-2010, including more efficient technology adoption, contributed to the breakdown in the relationship between economics and electricity demand. These factors help explain why electricity demand languished even as the economy began to recover. PJM compiled three indexes to explain these developing trends in different usage segments: Cooling, Heating, and Other. The following subsections explain this process.

Base Data Overview

The Energy Information Administration (EIA) is the originator of data used by PJM to model the impacts of equipment/appliance saturation and efficiency. Once a year, the EIA publishes the Annual Energy Outlook (AEO)[8]. The AEO has a Reference Case and several Scenarios based on changing various model assumptions. The Reference Case is the basis for the data used in the PJM Load Forecast. From the AEO2015 Preface:

> *The AEO2015 Reference case projection is a business-as-usual trend estimate, given known technology and technological and demographic trends…The main cases in AEO2015 generally assume that current laws and regulations are maintained throughout the projections. Thus, the projections provide policy-neutral baselines that can be used to analyze policy initiatives.*

[8] Link to AEO: http://www.eia.gov/forecasts/aeo/

The projections for the AEO are a product of the National Energy Modeling System (NEMS)[9]. NEMS contains modules to represent demand at the Census Division[10] level across different sectors: residential, commercial, industrial and transportation[11].

Within each demand module, consumers are modeled over time. Equipment/appliance stock evolves as existing stock becomes obsolete and is replaced or as new demand is warranted through new homes or buildings. Increases or replacements of equipment/appliance stock consider expected technology choices available[12]. National efficiency standards influence the technology choices available, but are not a ceiling. Adopted efficiency can and does exceed standard.

Itron compiles data from NEMS[13] on equipment/appliance saturation and their associated efficiency, and combines it with historical data that they maintain and update that is consistent with the EIA data set. This is a necessary step as EIA does not provide a full historical data set with each annual release, and long-term load forecasting requires more data than what EIA makes available. PJM receives Residential and Commercial sector data for each Census Division from Itron through membership in its Energy Forecasting Group[14]. Residential and Commercial data have the following equipment/appliance detail.

Residential

- Heating: Electric Furnaces and Resistant Room Space Heaters, Heat Pumps, Ground-source Heat Pumps, Secondary Heating, Furnace Fans
- Cooling: Central Air Conditioning, Heat Pumps, Ground-source Heat Pumps, Room Air Conditioners
- Other: Electric Water Heating, Electric Cooking, Refrigerator, Second Refrigerator, Freezer, Dishwasher, Electric Clothes Washer, Electric Clothes Dryer, TV Sets, Lighting, Miscellaneous Electric Appliances[15]

Commercial

- Heating: Single Heating Type
- Cooling: Single Cooling Type

[9] Link to NEMS Overview: http://www.eia.gov/forecasts/aeo/nems/overview/pdf/0581(2009).pdf
[10] Census Divisions are groupings of states. There are nine Census Divisions.
[11] NEMS Documentation: http://www.eia.gov/reports/index.cfm#/KNEMS%20Documentation,T144
[12] More information on the EIA Technology Forecasts can be found here: http://www.eia.gov/analysis/studies/buildings/equipcosts/
[13] Information on independently retrieving NEMS data can be found here: http://www.eia.gov/forecasts/aeo/info_nems_archive.cfm
[14] Link to information on Itron's Energy Forecasting Group: https://www.itron.com/na/productsAndServices/Pages/Energy%20Forecasting%20Group.aspx?market=electricity
[15] Residential Miscellaneous Electric Appliances is intended as a catch-all of other appliances not captured under the other listed types. Examples would be Laptops, Desktop PCs, Rechargeables, Security Systems, and Pool Heaters.

- Other: Ventilation, Water Heating, Cooking, Refrigeration, Outdoor Lighting, Indoor Lighting, Office Equipment (PCs), Miscellaneous[16]

Saturation

The saturation (or share) term is a percentage value to indicate the pervasiveness of a certain equipment type. Considering saturation is important, because these values will eventually be applied to transmission zones of many different sizes. Saturation rates, rather than counts, are size neutral.

In the case of the Residential sector, saturation is the percent of households that use a certain equipment type. This can be calculated by considering the total stock of a certain equipment type divided by the number of households. Certain equipment types (Refrigerators, TVs, Lighting and Miscellaneous) are considered ubiquitous, and are assigned a 100% saturation rate. The usage of these types over time will be handled solely by their associated efficiency term.

In the case of the Commercial sector, saturation is the percent of floor space that uses a certain equipment type. The EIA considers commercial energy consumption over 11 building types, and saturation for each type is calculated as the weighted percent of floor space. Only ventilation is considered universal, and assigned a 100% saturation rate.

Efficiency

The efficiency term measures the relative energy use of equipment/appliance types over time. In each case, positive movements in the efficiency term indicate an improvement and negative movements indicate deterioration.

The energy efficiency of Residential equipment/appliances is typically measured by some type of energy efficiency metric that measures energy out relative to energy in. Common examples of these would be the Energy Efficiency Ratio (EER) used to measure Central AC or the Efficiency Factor (EF) used to measure water heaters. In these instances, increases in the efficiency term indicate efficiency improvements.

However, many types (such as refrigerators or televisions) in the Residential sector are measured by a Unit Energy Consumption (UEC) metric measured in expected kilowatt-hours per year. Increases in a UEC-based metric would indicate deterioration; as a result PJM takes the reciprocal of these efficiency terms to maintain consistency.

In the case of the Commercial sector, all efficiency measures are in the form of an efficiency metric, therefore no transformations to the data are required.

[16] Commercial Miscellaneous is intended as a catch-all of other appliances not captured under the other listed types. Examples would be Security Systems, Medical Imaging Equipment, Elevators, and Escalators.

Equipment Index Construction

Three equipment indexes are constructed for each zone for use in the Load Forecast Model: Heating, Cooling, and Other. Constructing the indexes is a two-step process:

1) Develop Residential and Commercial sector indexes for each usage category;

2) Weight sector indexes and combine into a single index for each usage category.

Step 1: The sector index for each usage category is a weighted average across equipment types of saturation normalized for efficiency. Heating and Cooling Equipment Indexes are handled in the same manner as they will both be linked to weather parameters, which can be seen in the next section. Indexes are created for each Census Division[17] defined by the equation:

Equation IV-1 : Sector Equipment Index Construction (Heating and Cooling)

$$\mathbf{EquipIndex_{Sector,Category,y} = \sum_{type} W_{Type} \times (Sat_{Type,y}/Eff_{Type,y})/(Sat_{Type,1998}/Eff_{Type,1998})}$$

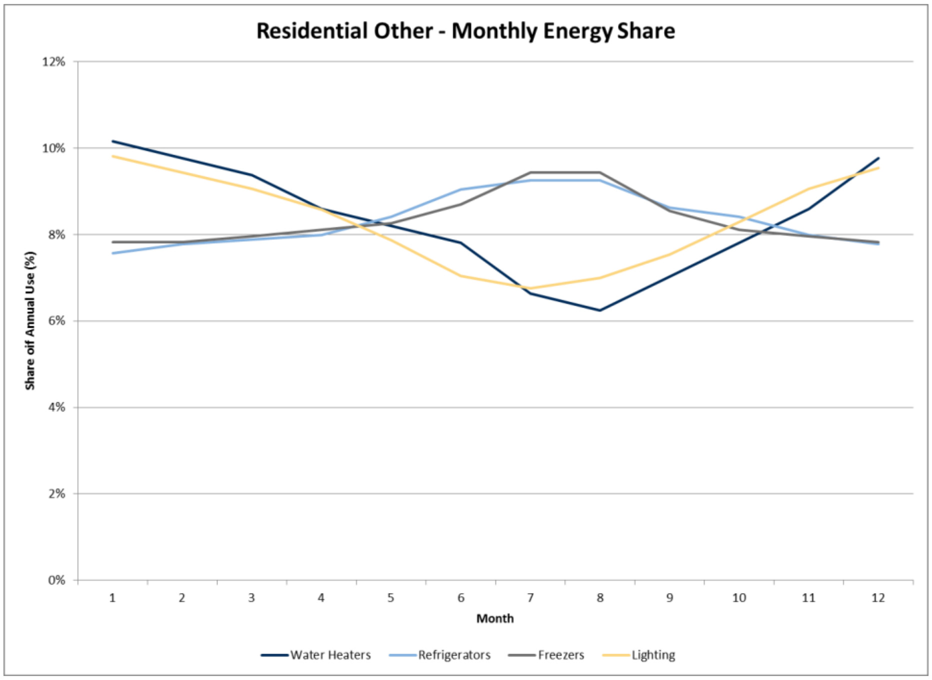
Where
Sector = Residential or Commercial
Category = Heating or Cooling
y = Year
Type = Equipment/Appliance type (see types listed earlier in this document)
W = Weight, defined here as the 1998 share of annual usage used by the specified equipment type in the specified category
Sat = Saturation, defined by year and equipment type
Eff = Efficiency, defined by year and equipment type

As can be seen in the equation, indexes are anchored to 1998 both by the weight and by the denominator of the equation. However, the index will evolve with time as saturations of various equipment types change and efficiencies improve.

Unlike the Heating and Cooling measures, the Other Equipment index is not linked with weather parameters. It instead is linked with each equipment type's expected use at different points of the year. Residential appliances in the Other category have different monthly energy usage profiles. Some appliance types are assumed to use the same amount of energy year-round (Electric Cooking, Dishwashers, Clothes Washers, Clothes Dryers, Televisions, and Miscellaneous). In these cases, the monthly weight is 8.33% (or 1/12). The remaining four Residential appliance types (Water Heaters, Refrigerators, Freezers and Lighting) have monthly weights derived from 2001 Itron/EIA data (see Figure IV-1). All Commercial equipment types in the Other category are given the 8.33% monthly weight.

[17] Additionally, some zones have modified indexes based on data they supplied. This is discussed in APPENDIX D.

Figure IV-1: Residential Other Equipment Monthly Energy Usage Share



The monthly weights are combined with the equipment saturation and efficiency data to provide a slight variation on the equations seen for the Heating and Cooling Indexes.

Equation IV-2 : Sector Equipment Index Construction (Other)

$$\mathbf{EquipIndex_{Sector,Other,t} = \sum_{type} Mw_{Type,m} \times W_{Type} \times (Sat_{Type,y}/Eff_{Type,y})/(Sat_{Type,1998}/Eff_{Type,1998})}$$

Where
Sector = Residential or Commercial
t = Time, year and/or month
Mw = Monthly weight
m = Month (1-12)
Type = Equipment/Appliance type (see types listed earlier in this document)
W = Weight, defined here as the 1998 share of annual usage used by the specified equipment type in the specified category
Sat = Saturation, defined by year and equipment type
Eff = Efficiency, defined by year and equipment type

Step 2: Total indexes for Heating, Cooling, and Other are now created by weighting and summing the sectoral components determined in Step 1. The weighting mechanism takes into account both the importance of that category within the sector (for instance, cooling usage as a share of a residential consumer's total usage) and the importance of that sector within the zone (for instance, residential sales as a share of a zone's residential and commercial sales).

Equation IV-3 : Overall Equipment Index Construction

$$\mathbf{EquipIndex_{Category,t} = \sum\nolimits_{Sector} Weight_{Sector,Category} \times EquipIndex_{Sector,Category,t}}$$

$$Weight_{Residential,Category} = AdjSales_{Residential,Category} / (AdjSales_{Residential,Category} + AdjSales_{Commercial,Category})$$
$$Weight_{Commercial,Category} = AdjSales_{Commercial,Category} / (AdjSales_{Residential,Category} + AdjSales_{Commercial,Category})$$
$$AdjSales_{Residential,Category} = TotalSales_{Residential} \times (Intensity_{Residential,Category} / Intensity_{Residential,All})$$
$$AdjSales_{Commercial,Category} = TotalSales_{Commercial} \times (Intensity_{Commercial,Category} / Intensity_{Commercial,All})$$

Where
Sector = Residential or Commercial
Category = Heating, Cooling, or Other
t = Time, year and/or month
Intensity = 5-year Sector Average Annual Energy Usage (by Category or All) from Itron/EIA data set
TotalSales = 5-year Sector Average Annual Energy Sales from FERC Form 1

Equipment Index Application

Each index—Heating, Cooling, and Other—is used according to the load behavior type it is meant to explain. The Heating and Cooling Equipment Indexes are intended to help explain weather sensitive load and therefore are tied to weather variables in the model. This technique allows weather condition magnitude as well as climate control equipment prevalence and efficiency to drive weather sensitive load. The Other Equipment Index is more centered on year-round load behavior (i.e. base load conditions), and thus does not have weather variable linkages.

Heat

The Heating Equipment Index is utilized primarily in winter months (January, February, and December), though it is also used in parts of the shoulder months (March, April, October, November) when conditions can be cold and windy.

Equation IV-4 : Heating Equipment Index Utilization

Winter (January, February, December)

$$\mathbf{EquipIndex_{Heat} \times S\_WWP_N}$$

Where
S_WWP = Variable based on values of the Winter-Weather Parameter or Wind-Adjusted Temperature at Hour Ending 19

N = Spline segment (1, 2, 3, or 4[18])

Shoulder (March, April, October, November)

$EquipIndex_{Heat}$ x (WWP – 50)
When WWP < 50

All Months

$EquipIndex_{Heat}$ x EconIndex x HDD
$EquipIndex_{Heat}$ x EconIndex x HDD(-1)

Where
EconIndex = Economic Index
HDD = Heating Degree Days
HDD(-1) = Heating Degree Days, lagged one day

Cool

The Cooling Equipment Index is utilized primarily in summer months (May through September), though it is also used in parts of the shoulder months (March, April, October, November) when conditions can be hot and humid.

Equation IV-5 : Cooling Equipment Index Application

Summer (May - September)

$EquipIndex_{Cool}$ x S_THI_N

Where
S_THI = Variable based on values of the Maximum Temperature Humidity Index
N = 2, 3, or 4[19]

Shoulder (March, April, October, November)

$EquipIndex_{Cool}$ x MaxTHI
When WWP > 70

All Months

$EquipIndex_{Cool}$ x EconIndex x CDD
$EquipIndex_{Cool}$ x EconIndex x CDD(-1)

Where

[18] The construction of winter seasonal weather variables is discussed in *Refined Weather Treatment*
[19] The construction of summer seasonal weather variables is discussed in *Refined Weather Treatment*

MaxTHI = Maximum daily temperature-humidity index
EconIndex = Economic Index
CDD = Cooling Degree Days
CDD(-1) = Cooling Degree Days, lagged one day

Other

As was discussed earlier, the Other Equipment Index does not have weather variable linkages. It is more intended to help explain base load conditions, which are nevertheless an important determinant for peak load demand.

Equation IV-6 : Other Equipment Index Application

All Months

EconIndex x $\mathbf{EquipIndex_{Other}}$

Where
EconIndex = Economic Index

Refined Weather Treatment

Summer and Winter

In the PJM load forecast model, different weather variables and relationships are used to represent the seasons. In the summer months, load tends to increase with hotter and more humid the weather conditions. However, in the winter months, load tends to increase with colder and windier weather conditions. While the forecast model has always allowed for varied treatment over the seasons, it now permits more granular treatment within the seasons.

Previously, the forecast model had a single independent variable in each season: the Temperature-Humidity Index (THI) and Winter-Weather Parameter (WWP) in summer and winter, respectively. The model also had additional time-of-day weather variables that allowed for a dynamic relationship of load with weather. However, because of multicollinearity (a situation where two or more independent variables are highly correlated), the coefficients on these variables were sometimes counterintuitive and obscured model transparency[20]. The goal of re-examining the seasonal weather treatment was to clean up this section of the model and improve transparency. In addition to re-visiting how weather is represented in the model, the definition of WWP was revised to be its value at hour ending 19 (7:00 PM) instead of the daily minimum. Daily minimums of WWP often occur in the morning or overnight hours, and thus this switch was considered to be more coincident with actual PJM winter peaks.

[20] See Slide 16 of the April 30, 2015 Load Analysis Subcommittee presentation: http://www.pjm.com/~/media/committees-groups/subcommittees/las/20150430/20150430-item-03-load-model-enhancements.ashx

Assuming a linear relationship with a constant slope between THI (or WWP) and load may lead to misrepresenting the underlying dynamic relationship of load to weather (see Figure IV-2 and Figure IV-3).

Figure IV-2 : THI Example (No Granular Treatment)

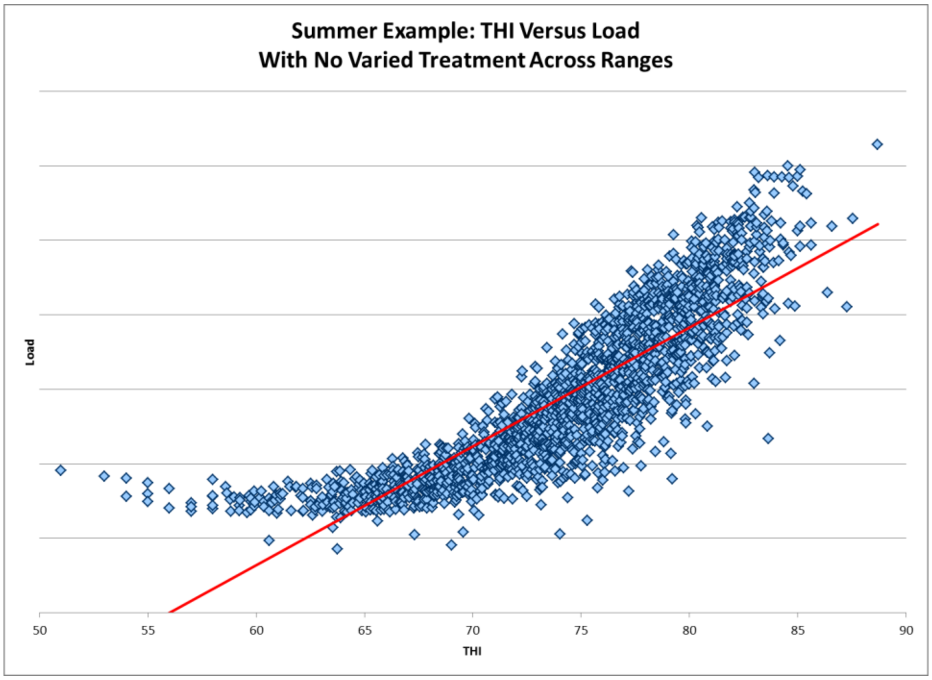
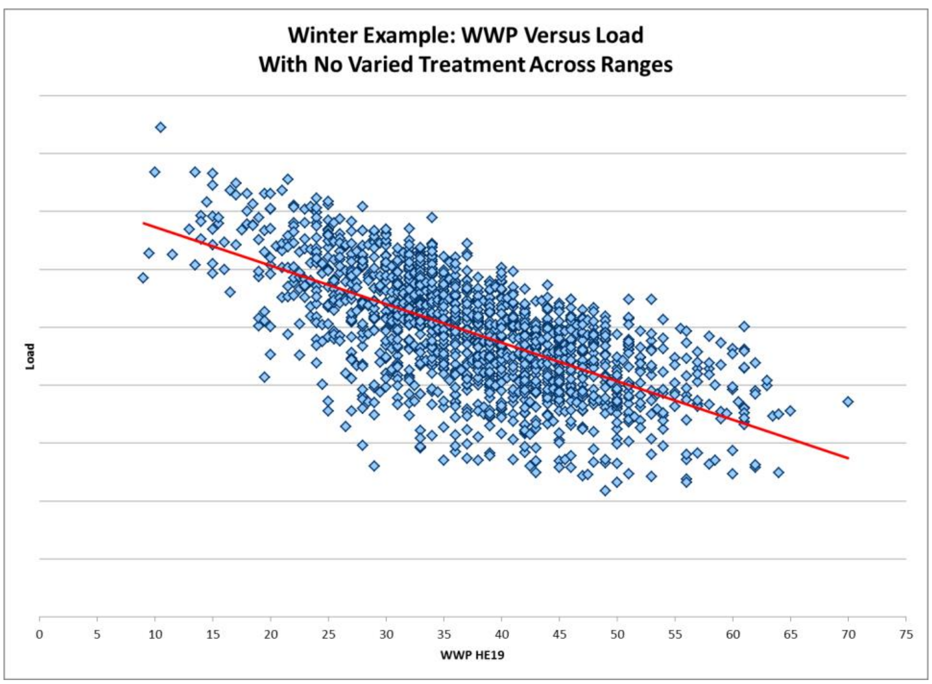


Figure IV-3 : WWP Example (No Granular Treatment)



A better representation of the load to weather relationship is to allow for a different load response at different ranges of the given weather variable. There are mild days that elicit minimal load-weather response. Also, there is some degree of HVAC saturation in extreme weather (i.e. the load to weather response moderates as equipment operates constantly).

The impact of mild days can be seen in both the summer and winter seasons. At THI values less than 65 or WWP values greater than 40, there appears to be minimal load response to weather conditions. At THI values around the high 70s and higher, there is often some moderation in load response from mid-range THI values. This reflects some degree of AC saturation. Modeling this intricacy should be explicit (see Figure IV-4 and Figure IV-5).

Figure IV-4 : THI Example (Granular Treatment)

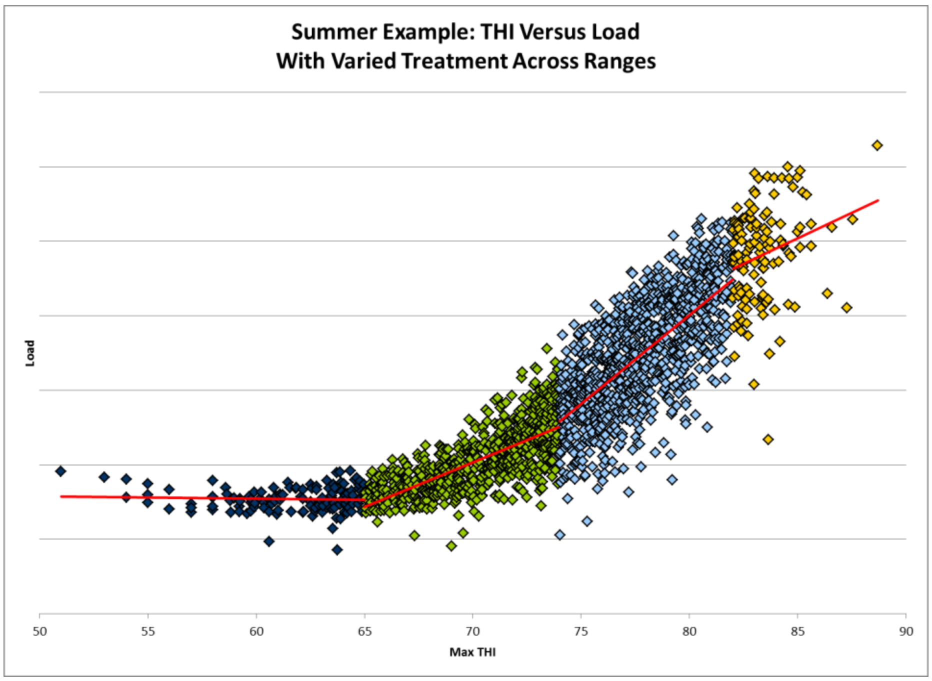
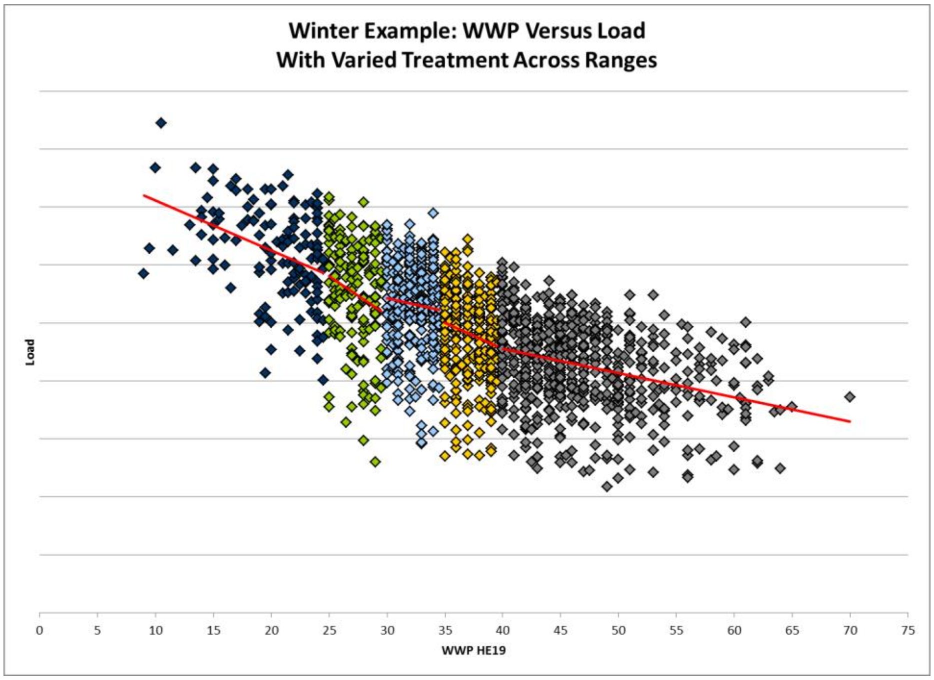


Figure IV-5 : WWP Example (Granular Treatment)



Weather conditions vary across the PJM footprint so it is best to treat zones individually, recognizing that what characterizes extreme conditions can vary zone by zone. Spline cut-off points were selected with the goal of best representing the behavior of load under extreme weather conditions when seasonal peaks will occur. The process to accomplish this was as follows:

1. Using time series analytical techniques, de-trend load history. This removes the impact of economics and other factors that impact load on a year-to-year basis, and puts all load on a level playing field;

2. Using regression analysis, model de-trended load against the weather parameter over a subset of weather values (for instance THI greater than 80 or WWP less than 25). Each regression is conducted from 1998 through 2014, and has additional variables to control for weekends and holidays;

3. Subset the load data to only show the top 10 zonal peaks per year. Compare the model-generated predicted values to the de-trended loads determined in step 1;

4. Choose the range cut-point that minimizes average percent error.

Once the extreme cut-points are determined for each zone and for each season, the remaining cut-points are determined. Mild conditions for all zones are defined as less than or equal to 65 THI and greater than or equal to 40 WWP for summer and winter, respectively. The two intermediate ranges are then the remaining range split roughly in half. The resulting THI and WWP spline range definitions are below (see Table IV-1 and Table IV-2). These are the ranges used in the 2016 Load Forecast; they will be re-visited periodically.

Table IV-1 : Thresholds for THI Splines

	Spline 2	Spline 3	Spline 4
AE	65	74	82
AEP	65	73	81
APS	65	73	81
ATSI	65	73	81
BGE	65	74	83
COMED	65	73	81
DAYTON	65	73	81
DPL	65	74	82
DQE	65	73	80
DUKE	65	73	81
EKPC	65	74	82
JCPL	65	73	81
METED	65	73	81
PECO	65	74	82
PENLC	65	72	78
PEPCO	65	74	83
PL	65	72	79
PS	65	73	81
RECO	65	73	81
UGI	65	72	79
VEPCO	65	74	82

Table IV-2 : Thresholds for WWP Splines

	Spline 2	Spline 3	Spline 4
AE	40	32	24
AEP	40	32	24
APS	40	31	22
ATSI	40	29	18
BGE	40	34	28
COMED	40	28	17
DAYTON	40	30	21
DPL	40	34	29
DQE	40	30	21
DUKE	40	32	25
EKPC	40	32	25
JCPL	40	33	27
METED	40	31	23
PECO	40	33	27
PENLC	40	30	20
PEPCO	40	35	30
PL	40	31	23
PS	40	34	28
RECO	40	33	27
UGI	40	29	18
VEPCO	40	36	32

These ranges are then used to create a series of weather variables, which build off one another. When THI or WWP is in the range of Spline 4 (S4), the load impact is calculated by taking into consideration the three splines that precede it (S1-S3).

Equation IV-7 : Summer and Winter Weather Spline Definitions

Summer (May – September)

S1_THI = MaxTHI

IF MaxTHI > $S2_{\text{Lower Bound}}$ THEN S2_THI = MaxTHI – $S2_{\text{Lower Bound}}$; ELSE S2_THI = 0

IF MaxTHI > $S3_{\text{Lower Bound}}$ THEN S3_THI = MaxTHI – $S3_{\text{Lower Bound}}$; ELSE S3_THI = 0

IF MaxTHI > $S4_{\text{Lower Bound}}$ THEN S4_THI = MaxTHI – $S4_{\text{Lower Bound}}$; ELSE S4_THI = 0

Winter (January, February, and December)

S1_WWP = WWP

IF WWP < $S2_{\text{Upper Bound}}$ THEN S2_WWP = WWP – $S2_{\text{Upper Bound}}$; ELSE S2_WWP = 0

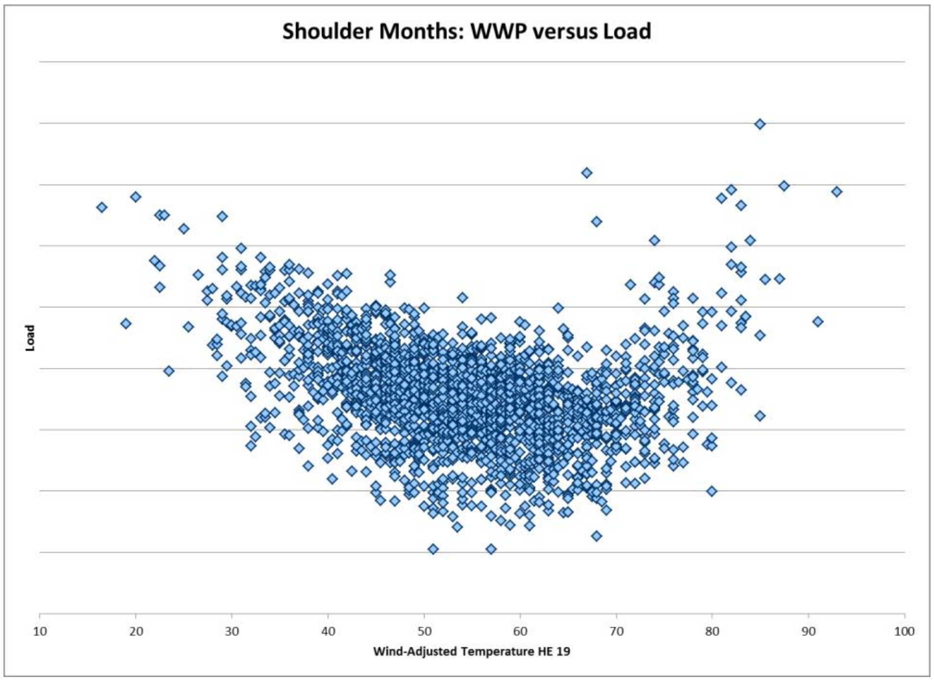
IF WWP < $S3_{\text{Upper Bound}}$ THEN S3_WWP = WWP – $S3_{\text{Upper Bound}}$; ELSE S3_WWP = 0

IF WWP < $S4_{\text{Upper Bound}}$ THEN S4_WWP = WWP – $S4_{\text{Upper Bound}}$; ELSE S4_WWP = 0

Shoulder Months

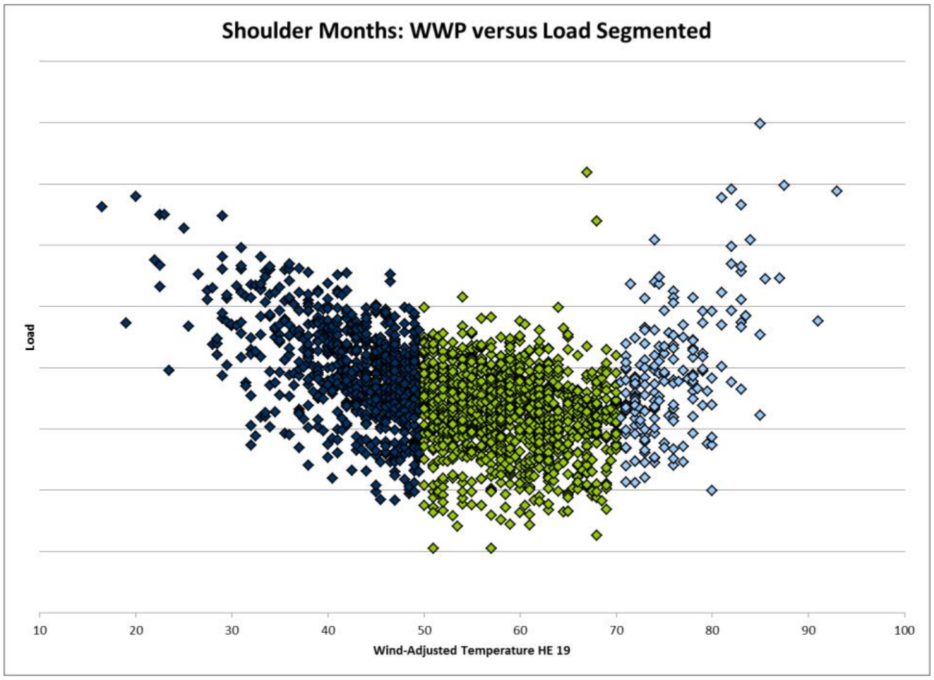
Unlike the summer and winter seasons, the shoulder months have more diverse weather with some days being summer-like and others winter-like. This leads to a U-shaped relationship of load to weather (see Figure IV-6).

Figure IV-6 : Shoulder Months - Load vs Weather Example



To accommodate this situation, the load to weather relationship is permitted to vary over three segments (see Figure IV-7). In doing so, the model can account for conditions that mimic winter or summer, as well as mild conditions in which there is little weather sensitivity.

Figure IV-7 : Shoulder Months – Load vs Weather Segmented



The left-most and center segments will be explained using the same WWP that is used in the winter, which accounts for both temperature and wind (see Figure IV-8). The right-most segment will be explained using THI, as in the summer months (see Figure IV-9).

Figure IV-8 : Shoulder Months – Load vs Weather (Less than 70 WWP)

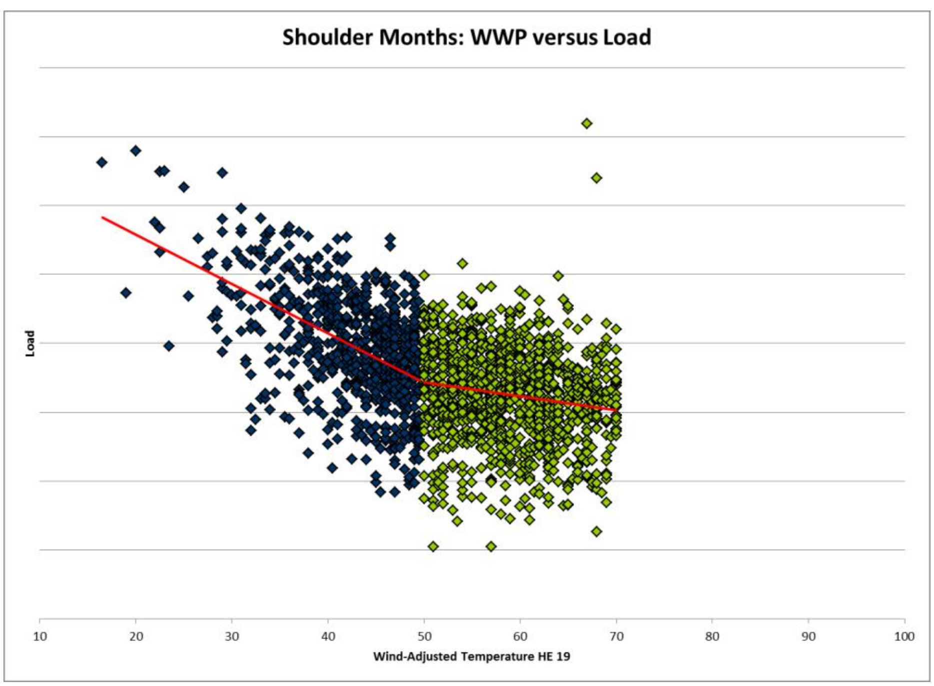
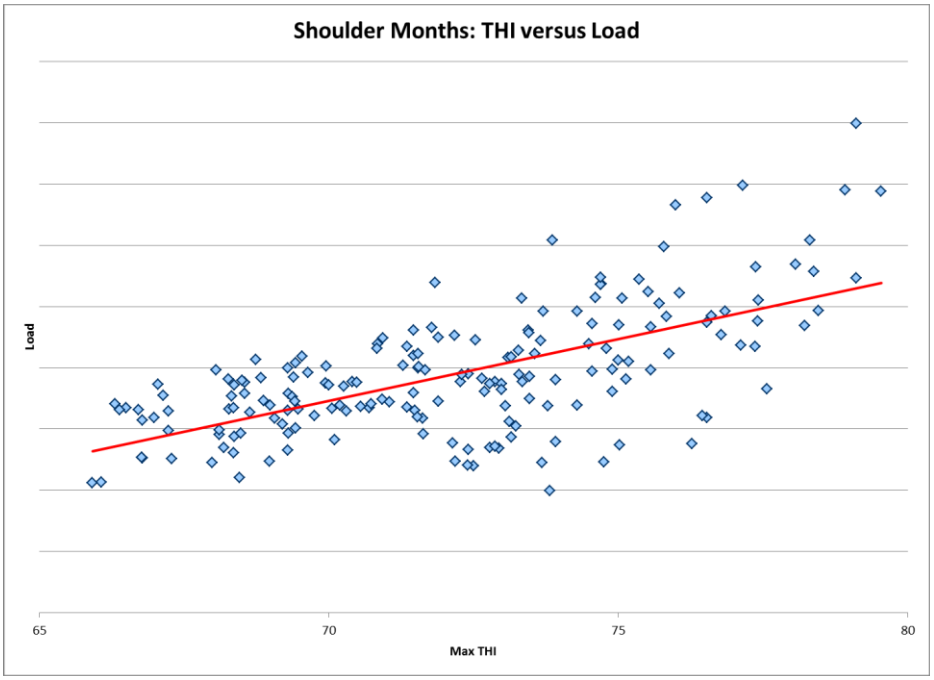


Figure IV-9 : Shoulder Months – Load vs Weather (Greater than 70 WWP)



Equation IV-8 : Shoulder Month Weather Definitions

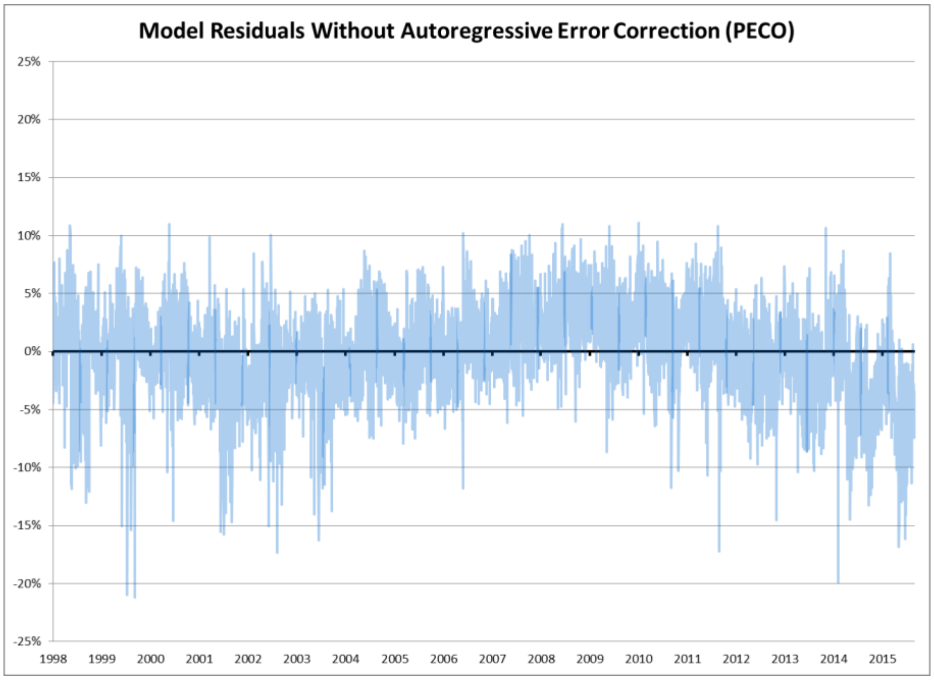
Shoulder Months (March, April, October, November)
IF WWP < 50 THEN Shldr_WWP_50LT = WWP - 50; ELSE THEN Shldr_WWP_50LT = 0 IF 50 <= WWP <= 70 THEN Shldr_WWP_BASE = WWP; ELSE THEN Shldr_WWP_BASE = 0 IF Shldr_WWP_50LT = 0 AND Shldr_WWP_BASE = 0 THEN Shldr_THI = MaxTHI; ELSE Shldr_THI = 0

Autoregressive Error Term

As part of its comprehensive model redevelopment, PJM examined whether there was cause to address the pervasive autocorrelation in the model's error terms. Autocorrelation is a common occurrence with time series modeling, and is an issue because it violates the assumption that the model errors are independent of one another (i.e. white noise). Consequently, model coefficient estimators are inconsistent and standard errors of the coefficient estimates are both biased and inconsistent, resulting in test statistics that are no longer valid. However, the presence of autocorrelation does not necessarily indicate that a forecast is biased.

Visual inspection and statistical tests can determine whether autocorrelation is present. In time series modeling, the goal is to have the residuals resemble white noise, meaning that they have no discernible pattern. This was clearly not the case for most zones (see Figure IV-10). Positive residuals indicate the model is under-predicting and negative residuals vice versa. The residuals seem to show a pattern, wherein residuals are correlated with the prior period(s). Moreover, the pattern of negative residuals in the most recent years may contribute to over-forecasting in the early years of the forecast period as present errors with autocorrelation are tied to past errors.

Figure IV-10 : Residuals Without AR Error Correction Example

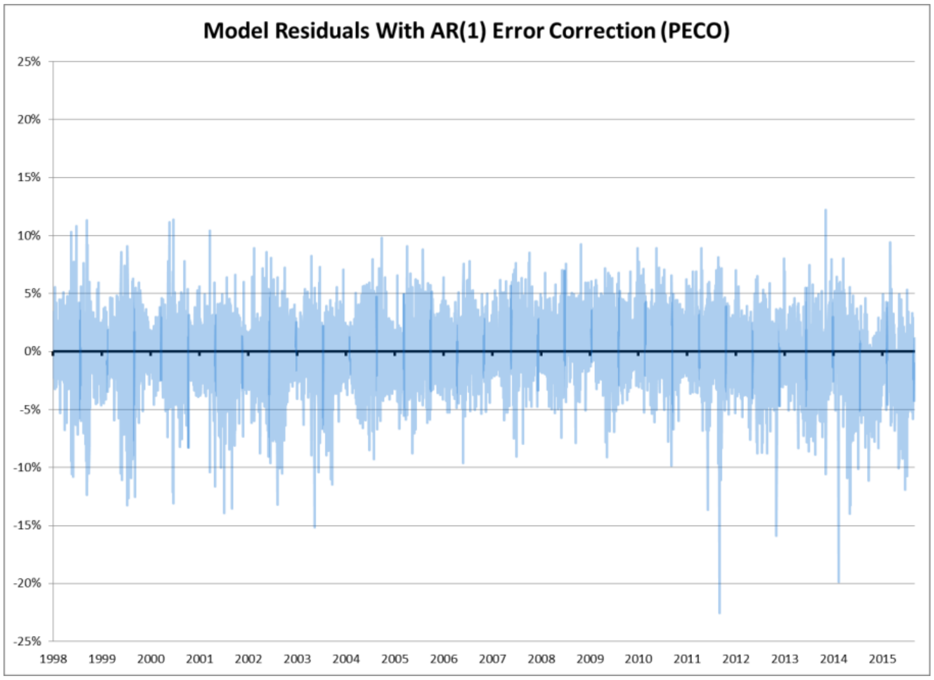


To confirm or refute the presence of autocorrelation in the residuals, Durbin-Watson test statistics were computed for each zonal model. This is a common method in this type of analysis. In each instance, the results indicated the presence of positive autocorrelation (a positive residual in one period is likely to be followed by another positive residual and vice versa). The appropriate fix can then be determined through observing plots of both the autocorrelation function (ACF) and partial autocorrelation function (PACF) produced via statistical software (PJM uses SAS). These plots pointed to the need for an autoregressive error structure with 1 lag, alternatively referred to as AR(1). This means that residuals can be modeled using a one period lag of itself.

Upon fitting the models with an AR(1) error structure, the residuals appeared more like white noise (see Figure IV-11). To confirm that there was no longer autocorrelation in the residuals, Durbin-Watson test statistics were computed on the modified zonal models. The results confirmed that this was the case and that the issue had been adequately resolved.

The impact of adopting the autoregressive error term into PJM's model was to reduce the starting point of forecasts, with little impact on growth rates. This was deemed a worthwhile change as it contributed to addressing the problem of the first year of the forecast often exhibiting unrealistically high growth rates.

Figure IV-11 : Residuals With AR Error Correction Example



Weather Simulation Period

The preceding sections cover model structure changes, specifically revisions to the regression model that determine the parameter estimates to generate the load forecast. There is a distinction between these revisions and a change to assumptions that are applied to derive the forecast: economic growth, equipment saturation and efficiency projections, and weather.

Weather is an important electricity demand driver; however, in order to arrive at a peak forecast, PJM does not forecast weather. Instead, future weather is assumed to mimic past weather. Past practice had been to use all available weather data back to the early 1970s. Weather is then sampled to arrive at the distribution of potential peaks. From this distribution, PJM can determine a range of values that are used in RTEP such as the 50/50 (50% chance the peak will be less than and 50% the peak will be greater than) or the 90/10 (90% chance the peak will be less than and 10% the peak will be greater than).

The past weather used to generate the distribution is an important input in determining the forecast values used. In using data back to the early 1970s, this implicitly makes the assumption that all weather is equally likely in the future. In other words, weather in 2020 is just as likely to resemble the 1970s as it is to resemble the 2010s. An investigation was conducted to determine whether this is a prudent assumption.

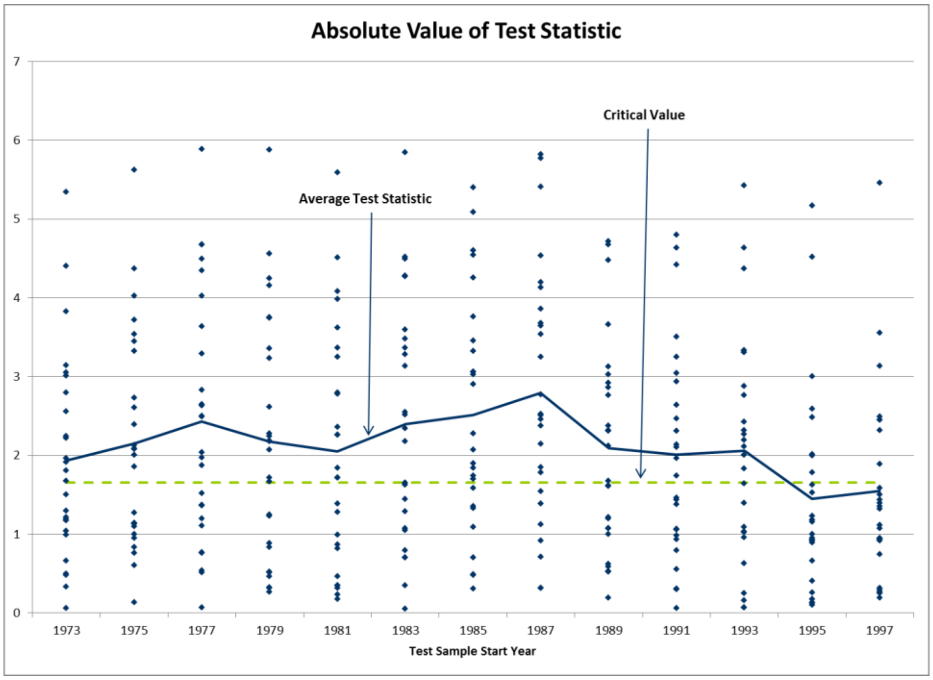
At the time of the investigation in Summer 2015, the forecast model was using data from 26 weather stations across the footprint. The data was broken down in a number of ways, comparing median and 90th percentile values of maximum THI from 1973-1993 to those from 1994-2013. The preponderance of evidence indicated that the majority of weather stations had experienced higher THI values in the more recent 20 years than in the prior 20 years[21]. The concern was that using the full amount of weather history may lead to understating the extremeness of weather and thus potentially also understating peak load.

Once it was determined that using weather history back to the early 1970s was imprudent, a new time period needed to be selected. To do so, weather data from 2005-2014 was considered a control sample, since it was a sample of sufficient size to have a mean and standard deviation to compare to other samples. This sample was then compared with samples from historical years to evaluate consistency. For each weather station for each year, the 13 highest THI values were selected. Multiple test samples were then constructed with an end point of 2004 and a starting point that varied incrementally from 1973 by two years (i.e. 1973-2004, 1975-2004, 1977-2004, etc.).

Test samples were then compared with the control sample using a two-sample t-Test to determine consistency. In this statistical test, the null hypothesis is that the two samples are equivalent and the null hypothesis is rejected when the test statistic exceeds a critical value (in this case, evaluated at a 5% significance level). On average, across weather stations, the test statistic did not lie below the critical value until the test sample had a start year of 1995 (see Figure IV-12). At this point, 18 of the 26 weather stations had test statistics that lie below the critical value (each dot in the figure represents a weather station's test statistic).

[21] Analysis was shared at Load Analysis Subcommittee on September 2, 2015: http://www.pjm.com/~/media/committees-groups/subcommittees/las/20150902/20150902-item-04-forecast-update.ashx

Figure IV-12 : Weather Simulation Period Evaluation – Test Statistic Values



Based on these results, PJM decided to use only weather from 1994/1995 forward as the assumption for weather in the forecast model. This assumption will be revisited periodically as more data becomes available.

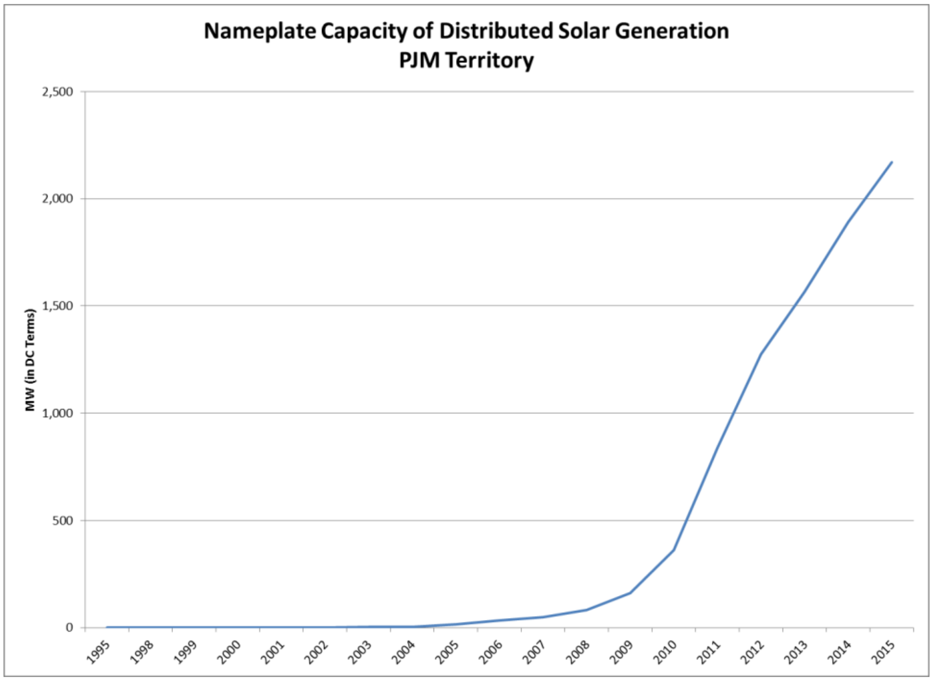
Distributed Solar Generation

In early 2015, recognizing the growing market of solar installations, PJM began to investigate and develop a plan to incorporate distributed solar generation into the long-term load forecast. For the purposes of the long-term load forecast, PJM defines distributed solar generation as any solar resource which is not interconnected to the PJM markets. These resources do not go through the full interconnection queue process and do not offer as capacity or as energy resources. Furthermore, the output of these resources is netted directly with the load. PJM does not receive metered production data from any of these resources.

PJM EIS and GATS

PJM Environmental Information Services (EIS), a wholly owned subsidiary of PJM Technologies, Inc. which is a subsidiary of PJM Interconnection, operates the Generation Attribute Tracking System (GATS). The functional design of the GATS has been developed through considerable deliberation of a stakeholder group that included representatives from various state agencies (state public utility commissions, state environmental protection offices, state Energy offices and consumer advocates), market participants, environmental advocates, and PJM staff. The design of the GATS is an "unbundled," certificates-based tracking system. This means that the attributes or characteristics of the generation are separated from the megawatt hour (MWh) of Energy and recorded onto a Certificate after the MWh of Energy is produced. There is one Certificate, with a unique serial number representing the attributes of the generation for each MWh produced. The Certificate's value is that it can be traded separately from the actual MWh of Energy in a voluntary bilateral market. The generation data which GATS collects includes distributed solar generation that are behind the meter. Utilizing this collection of data, PJM estimates the amount of distributed solar generation in terms of direct current (DC) nameplate capacity. Figure IV-13 below shows the amount of DC nameplate capacity (MW) of distributed solar generation as recorded in GATS for the PJM footprint. In the last five years, there has been over a 1000% increase of installations in the PJM region, and the number of installations is expected to continue to grow.

Figure IV-13 : Nameplate Capacity of Distributed Solar Generation in the PJM Footprint



Approach

PJM recognizes the need to isolate distributed solar generation from the historical load. Not addressing these resources would mean that the reductions in the load attributable to distributed solar generation would be explained by the other independent variables in the model, thus misrepresenting the actual values of the model parameters and biasing the forecast. PJM is using a two-step approach to incorporate distributed solar generation into the long term load forecast.

Step 1: Back-casting Historical Distributed Solar Generation

In this first step, historical hourly impacts of distributed solar generation are estimated in order to add them back to the historical load values. This allows the load forecast model to accurately estimate the relationship between load and the various explanatory variables without undue interference from distributed solar energy resource impacts.

GATS and Solar Insolation

To account for the historical impacts of distributed solar generation, PJM used the distributed solar generation installation data from the Generation Attribute Tracking System (GATS) as the basis of the amount of installed MW[22]. The first step to estimate the historical back-casted values is to convert the nameplate capacity amount to a value which incorporates the maximum possible insolation. Solar insolation is the amount of solar energy that would reach the Earth's surface at a given time of day based on a cloudless sky. PJM referenced National Oceanic and Atmospheric Administration (NOAA) data via an ITRON application to determine the hourly value of how much energy would be produced from a particular solar panel on any given day[23]. The distributed solar generation were mapped to the local weather stations that are listed in Section 3 of Manual 19 and the applicable latitude and longitude definitions were used as the point on the Earth's surface.

Cloud Cover

A cloud cover variable based on the weather stations listed in Section 3 of Manual 19 was also applied to the nameplate MW amount. The cloud cover variable is important to include because the varying degrees of cloud cover can greatly impact the amount of energy produced by a solar panel. Metar, a weather measurement, data includes five categories for cloud cover as well as a separate code for weather phenomena (rain, thunderstorm, blowing snow, etc.). In the weather data PJM receives from its weather vendor, these two measurements have been merged into one column. PJM translates the combination cloud cover/weather phenomena into a 0-8 index; a value of 8 is considered overcast, whereas a value of zero is considered to be a weather condition where there are no clouds. Each MW value was multiplied by:

Equation IV-9 : Cloud Cover Adjustment

$$1 - \frac{\text{cloud cover index}}{8}, \text{ where cloud cover index is less than 8}$$

Research showed that even when there is cloud cover, a variable amount of energy is still produced, so for hours when the cloud cover index was equal to eight, instead of applying Equation IV-9, the MW nameplate capacity amount was multiplied by ten percent.

Temperature Impacts

In addition to cloud cover and solar insolation, higher temperatures have also been found to result in degradation of energy output. For every degree above 55°F, the energy output of the solar panel is reduced by 0.27%[24]. The equation applied is:

Equation IV-10 : Temperature Degradation Adjustment

$$1-(\text{Maximum}(\text{Temperature}-55,0)*0.0027)$$

---

[22]The values from GATS are available publicly through the PJM EIS website
[23]A more complete description and list of equations is provided in this ITRON Forecast Practitioner's Handbook.
[24] Page 6 of the ITRON Forecast Practitioner's Handbook.

Tilt of Solar Panel and DC to AC Conversion

Finally, two last variables were considered to determine the historical back-casted values: the DC to AC conversion; and the 27 degree tilt factor[25]. PJM also assumed that the panels were southern facing. Southern facing panels are the ideal position so it was assumed that all installations strived to be as southern facing as possible[26]. Since the GATS data is submitted in direct current (DC) terms, a conversion to alternating current (AC) terms needed to be applied. Most homes utilize AC, thus the need for the DC/AC inverter losses. Also, on average, panels are not installed completely flat; rather they are installed based on a particular tilt. This tilt generally corresponds to the angle of the roof on which it is installed. On average, this tilt is approximately 27 degrees. To determine the impact of these two adjustments, PJM conducted simulations using the PVWATTS calculator[27]. PJM calculated the hourly average of the ratio of a zero degree tilt to a 27 degree tilt by month for each weather station as listed in Section 3 of Manual 19. PJM used a 96% DC to AC conversion factor which was based on analysis of the PVWATTS simulations. Since PVWATTs uses a typical meteorological year these values represent a good estimate of their impact on the energy production of the solar panels. Table IV-3 shows the Annual GWh of back-casted distributed solar generation by zone. Please note that all years except for 2015 are based on calendar year, however, 2015 is only through August 31, 2015.

Table IV-3 : Annual Energy of Back-casted Distributed Solar

Annual GWh of Back-casted Distributed Solar Generation

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
AE					0.0	0.1	0.5	1.2	4.8	7.9	11.6	26.2	60.8	105.5	181.6	214.8	270.8	211.9
AEP					0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	10.7	13.4	19.0	23.9	26.0	22.3
APS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	1.1	6.0	16.6	23.1	31.1	43.3	28.1
ATSI					0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.4	4.4	15.9	31.7	43.1	49.9	41.8
BGE							0.0	0.0	0.0	0.1	0.2	2.0	5.9	15.2	32.8	46.0	62.8	68.0
COMED			0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.5	0.8	14.0	22.2	23.6	21.6
DAYTON										0.0	0.0	0.1	0.3	1.0	3.9	9.8	10.1	8.0
DEOK		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.4	0.8	3.5	8.5	11.0	11.2	8.7
DLCO							0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.6	1.4	3.0	3.2	2.5
DOM		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.6	3.2	6.5	9.9	18.4	35.8	94.3
DPL		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.7	2.1	5.2	9.2	32.2	55.2	99.8	129.1	113.4
EKPC										0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.2	0.3
JCPL			0.0	0.0	0.1	0.6	1.5	3.8	9.8	16.6	24.0	38.1	66.1	121.7	250.1	315.5	395.3	328.2
METED			0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.8	6.8	27.4	38.0	37.2	38.2	28.2
PECO	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.3	0.4	0.6	0.7	1.3	9.1	27.1	42.2	47.4	49.2	35.9
PENLC			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.8	2.6	4.8	5.3	6.0	4.9
PEPCO						0.0	0.0	0.0	0.1	0.2	0.5	4.1	9.4	20.9	36.1	57.2	83.1	46.6
PL								0.0	0.0	0.1	0.2	0.8	17.9	66.3	93.9	94.0	97.3	75.3
PS			0.0	0.3	0.4	0.5	0.9	2.5	8.2	17.2	24.9	48.1	114.2	196.7	365.8	481.9	518.2	403.1
RECO			0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.7	0.9	1.0	1.8	3.2	8.2	11.0	11.5	8.6
UGI											0.0	0.0	0.2	0.4	0.4	0.5	0.5	0.4
PJM RTO	**0.0**	**0.0**	**0.1**	**0.4**	**0.7**	**1.5**	**3.4**	**8.4**	**24.2**	**44.6**	**66.2**	**130.7**	**328.3**	**677.7**	**1,220.8**	**1,573.4**	**1,865.4**	**1,552.3**

Application in Load Forecast Model

Once all of the adjustments were applied, PJM summed the estimated energy impacts on an hourly basis by zone and these estimates were added to the unrestricted load used in PJM load models to generate a forecast that essentially removes distributed solar generation impacts from the load.

[25] Common roof pitches range from 4:12 to 9:12 which translates to a roof angle of as little as 18.4 degrees to as high as 36.9 degrees. 27 degrees was chosen as it lies firmly in the middle of this range.

[26] http://rredc.nrel.gov/solar/calculators/pvwatts/system.html

[27] The PVWATTS calculator is available via this web link: http://pvwatts.nrel.gov/

Step 2: Ex-Post Adjustment of Distributed Solar Generation

The second step in the process to incorporate distributed solar generation into the long term load forecast was to utilize a long term forecast of distributed solar generation.

IHS Forecast

PJM procured a distributed solar generation forecast from IHS. This entailed working with IHS to define inputs which represented the power demand growth rate specific to PJM, as well as the policies which were in effect at the time the forecast was conducted. Some key assumptions included that the 30% Investment Tax Credit (ITC) would expire for any resources installed after 2016 and that it would drop to a 10% ITC thereafter. Additionally, PJM assumed that current net metering policies will remain in effect through the duration of the forecast.[28] Table IV-4 below shows the forecast of the distributed solar generation annual additions of nameplate capacity, by state within the PJM territory.

Table IV-4 : Distributed Solar Energy Resource Forecast by State - PJM Territory Only -Annual Additions of Nameplate Capacity

PJM Territory of the State	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
DC	3.9	3.6	3.8	4.6	4.1	3.4	3.3	3.3	3.2	3.2	3.2	3.2	3.1	3.1	3.1	3.1
DE	14.6	19.4	24.1	25.4	16.8	18.3	30.9	34.2	24.3	23.9	25.2	28.0	34.7	45.4	60.3	81.2
IL	37.2	22.3	15.1	15.0	14.9	24.2	30.0	29.7	32.0	32.8	32.5	32.2	32.0	31.7	33.7	36.4
IN	2.6	2.0	1.3	1.2	1.2	1.2	1.2	1.2	1.2	1.3	2.0	2.9	4.1	5.7	5.9	6.2
KY	0.9	0.4	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2	2.8	3.7
MD	131.3	117.1	119.4	88.4	39.2	14.5	10.3	15.6	25.7	37.6	51.3	68.0	95.9	129.6	142.5	154.7
MI	0.6	0.4	0.4	0.6	0.7	0.9	1.2	1.6	2.0	2.6	3.4	4.5	5.8	6.7	7.6	8.7
NC	85.4	70.1	37.1	37.3	37.5	37.8	38.3	39.0	40.0	42.1	44.8	48.0	52.9	59.6	68.0	77.2
NJ	209.3	116.9	56.4	43.8	52.5	56.2	70.1	92.0	118.6	176.8	245.6	293.6	316.1	354.6	367.3	372.4
OH	10.9	35.6	39.6	39.2	38.9	43.1	44.3	44.8	48.3	49.2	48.8	17.7	8.9	9.9	12.7	16.0
PA	68.3	45.7	47.8	51.5	53.7	38.3	32.4	32.1	31.9	31.6	31.4	32.2	35.2	39.1	53.5	61.5
TN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
VA	35.3	33.5	29.9	51.7	72.5	81.3	107.2	121.0	126.5	129.8	132.9	138.2	145.3	153.9	164.4	175.3
WV	1.1	3.9	5.0	4.9	4.9	4.8	4.8	4.8	4.7	4.7	4.6	4.6	4.6	4.5	8.3	14.3
Total	601.4	470.8	379.9	363.9	337.0	324.3	374.2	419.3	458.6	535.7	625.8	673.2	738.8	844.1	930.1	1,010.6

Converting State to Zone

The IHS report presented nameplate capacity additions by state, for the PJM portion of the state. To derive additions by zone, PJM converted the state values to zone using EIA 826 data to estimate each utility's share of sales in the state. Once the zone's share of the state was calculated, that ratio was applied to the state level annual additions of installed capacity. For those transmission zones that span multiple states, their state shares were aggregated to obtain a total for the zone. Table IV-5 below shows the annual additions of installed capacity of distributed solar generation by zone.

[28] The full report is available via this link: http://pjm.com/~/media/committees-groups/subcommittees/las/20151130/20151130-item-04-ihs-pjm-pv-forecast-report.ashx.

Table IV-5 : Distributed Solar Energy Resource Forecast by Zone Annual Additions of Nameplate Capacity

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
AE	27.6	15.3	7.4	5.7	6.8	7.3	9.1	11.9	15.3	22.8	31.6	37.8	40.6	45.4	47.0	47.5
AEP	13.6	23.1	23.9	27.1	30.3	33.3	37.8	40.3	42.9	44.3	46.1	37.2	37.5	41.6	48.4	56.5
APS	27.3	23.6	24.5	22.2	17.3	12.5	12.2	13.4	14.9	16.5	18.3	20.7	25.0	30.2	35.9	41.4
ATSI	6.7	14.9	16.4	16.4	16.4	17.3	17.5	17.6	18.8	19.1	18.9	7.6	4.6	5.1	6.6	8.0
BGE	64.3	57.2	58.3	43.1	19.1	7.0	5.0	7.6	12.5	18.2	24.9	32.9	46.4	62.7	68.9	74.7
COMED	37.2	22.3	15.1	15.0	14.9	24.2	30.0	29.7	32.0	32.8	32.5	32.2	32.0	31.7	33.7	36.4
DAYTON	1.2	3.8	4.2	4.2	4.1	4.6	4.7	4.8	5.1	5.2	5.2	1.9	1.0	1.1	1.4	1.7
DEOK	1.8	5.6	6.2	6.1	6.1	6.8	7.0	7.0	7.6	7.7	7.7	2.8	1.4	1.6	2.2	2.8
DLCO	6.2	4.2	4.3	4.7	4.9	3.5	2.9	2.9	2.9	2.8	2.8	2.9	3.1	3.5	4.8	5.4
DOM	113.4	96.7	60.9	78.5	95.3	102.7	123.9	135.8	141.3	146.2	151.3	158.9	169.5	183.1	200.1	218.0
DPL	23.7	27.5	32.4	31.8	20.1	20.1	32.7	36.4	27.2	27.6	29.9	33.8	42.4	55.4	71.3	93.2
EKPC	0.5	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1.6	2.1
JCPL	60.1	33.7	16.3	12.7	15.2	16.2	20.3	26.6	34.4	51.2	71.2	85.3	92.0	103.4	107.2	108.9
METED	6.8	4.5	4.8	5.2	5.4	3.8	3.2	3.2	3.2	3.2	3.2	3.3	3.6	4.0	5.5	6.3
PECO	17.7	11.8	12.4	13.4	13.9	9.9	8.4	8.4	8.3	8.2	8.2	8.4	9.2	10.2	14.0	16.1
PENLC	7.6	5.0	5.2	5.6	5.8	4.2	3.5	3.5	3.4	3.4	3.3	3.4	3.7	4.1	5.5	6.3
PEPCO	46.3	41.4	42.3	33.2	16.8	8.1	6.6	8.3	11.5	15.3	19.7	25.1	34.1	44.9	49.1	53.0
PL	17.5	11.7	12.2	13.2	13.8	9.8	8.3	8.2	8.2	8.1	8.0	8.2	9.0	10.0	13.7	15.7
PS	117.6	65.7	31.7	24.6	29.5	31.6	39.5	51.8	66.7	99.4	138.2	165.1	177.7	199.3	206.3	209.2
RECO	4.0	2.2	1.1	0.8	1.0	1.1	1.3	1.7	2.2	3.3	4.6	5.5	5.9	6.6	6.8	6.9
UGI	0.4	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.4
PJM RTO	**601.4**	**470.8**	**379.9**	**363.9**	**337.0**	**324.3**	**374.2**	**419.3**	**458.6**	**535.7**	**625.8**	**673.2**	**738.8**	**844.1**	**930.1**	**1,010.6**

The annual additions shown in the table above include an assumption that there is a degradation factor of 0.8% for every year installed[29]. While the median degradation is 0.5%, PJM chose to use the more conservative average value of 0.8%. Table IV-6 below shows the total cumulative nameplate capacity by zone.

Table IV-6 : Cumulative Nameplate Capacity of Distributed Solar Generation Forecast by Zone

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
AE	27.6	42.9	50.3	56.0	62.8	70.1	79.2	91.1	106.4	129.2	160.8	198.6	239.2	284.6	331.6	379.1
AEP	13.6	36.7	60.6	87.7	118.0	151.3	189.1	229.4	272.3	316.6	362.7	399.9	437.4	479.0	527.4	583.9
APS	27.3	50.9	75.4	97.6	115.0	127.5	139.7	153.1	168.0	184.5	202.7	223.5	248.5	278.7	314.5	355.9
ATSI	6.7	21.6	38.1	54.5	70.8	88.1	105.6	123.1	141.9	161.0	179.9	187.5	192.1	197.1	203.7	211.8
BGE	64.3	121.5	179.7	222.9	242.0	249.0	254.0	261.6	274.1	292.3	317.2	350.1	396.6	459.3	528.1	602.9
COMED	37.2	59.5	74.6	89.7	104.6	128.8	158.8	188.5	220.5	253.3	285.8	318.0	350.0	381.7	415.3	451.7
DAYTON	1.2	4.9	9.1	13.3	17.4	22.0	26.7	31.5	36.6	41.8	47.0	48.9	49.9	50.9	52.3	54.0
DEOK	1.8	7.3	13.5	19.7	25.8	32.5	39.5	46.5	54.1	61.9	69.6	72.4	73.8	75.4	77.6	80.4
DLCO	6.2	10.4	14.7	19.4	24.2	27.7	30.6	33.5	36.4	39.2	42.0	44.9	48.1	51.5	56.3	61.7
DOM	113.4	210.1	271.0	349.5	444.8	547.5	671.4	807.2	948.5	1,094.7	1,246.0	1,404.9	1,574.5	1,757.6	1,957.6	2,175.7
DPL	23.7	51.2	83.5	115.3	135.4	155.4	188.1	224.4	251.6	279.2	309.1	342.9	385.3	440.8	512.1	605.3
EKPC	0.5	0.7	0.8	0.9	1.0	1.1	1.1	1.2	1.3	1.4	1.4	1.5	1.6	1.7	3.3	5.5
JCPL	60.1	93.8	110.2	122.9	138.0	154.3	174.5	201.1	235.5	286.7	357.9	443.2	535.1	638.5	745.8	854.6
METED	6.8	11.3	16.1	21.2	26.6	30.4	33.7	36.9	40.1	43.3	46.5	49.7	53.3	57.3	62.7	69.0
PECO	17.7	29.5	41.9	55.3	69.3	79.2	87.6	96.0	104.3	112.5	120.7	129.1	138.3	148.5	162.5	178.6
PENLC	7.6	12.6	17.9	23.5	29.3	33.5	37.0	40.4	43.8	47.2	50.5	53.9	57.6	61.6	67.1	73.5
PEPCO	46.3	87.7	130.0	163.2	180.0	188.1	194.8	203.1	214.6	229.9	249.7	274.8	308.8	353.8	402.8	455.9
PL	17.5	29.2	41.4	54.6	68.4	78.2	86.5	94.7	102.9	111.0	119.0	127.2	136.2	146.2	159.9	175.6
PS	117.6	183.3	214.9	239.5	269.1	300.7	340.2	391.9	458.6	558.1	696.3	861.4	1,039.1	1,238.3	1,444.6	1,653.8
RECO	4.0	6.2	7.3	8.1	9.1	10.1	11.5	13.2	15.4	18.7	23.3	28.8	34.7	41.2	48.0	54.9
UGI	0.4	0.7	1.0	1.4	1.7	1.9	2.1	2.3	2.5	2.7	2.9	3.1	3.3	3.6	3.9	4.3
PJM RTO	**601.4**	**1,072.2**	**1,452.1**	**1,816.1**	**2,153.1**	**2,477.4**	**2,851.6**	**3,270.9**	**3,729.6**	**4,265.3**	**4,891.1**	**5,564.3**	**6,303.1**	**7,147.2**	**8,077.4**	**9,088.0**

Historical Distributed Solar Generation Installations

Since PJM added the estimated historical back-casted solar impacts to the unrestricted loads, those must be considered as well. Using the GATS data of currently installed distributed solar generation, along with a 96% DC to AC conversion factor, and a 0.8% degradation assumption, PJM calculated the amount of

---

[29] Beginning on Page 5 of this NREL study: http://www.nrel.gov/docs/fy12osti/51664.pdf

nameplate capacity of distributed solar generation that have already been installed. Table IV-7 below shows how much of the currently installed resources are assumed to operate in the future.

Table IV-7 : Cumulative Nameplate Capacity of Existing Distributed Solar Generation by Zone

ZONE	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
AE	192.0	190.4	188.9	187.4	185.9	184.4	182.9	181.5	180.0	178.6	177.1	175.7	174.3	172.9	171.5	170.2
AEP	36.3	36.0	35.8	35.5	35.2	34.9	34.6	34.4	34.1	33.8	33.5	33.3	33.0	32.7	32.5	32.2
APS	41.7	41.4	41.1	40.7	40.4	40.1	39.8	39.4	39.1	38.8	38.5	38.2	37.9	37.6	37.3	37.0
ATSI	50.4	50.0	49.6	49.2	48.8	48.4	48.0	47.6	47.3	46.9	46.5	46.1	45.8	45.4	45.0	44.7
BGE	105.0	104.2	103.4	102.5	101.7	100.9	100.1	99.3	98.5	97.7	96.9	96.2	95.4	94.6	93.9	93.1
COMED	33.9	33.7	33.4	33.1	32.9	32.6	32.3	32.1	31.8	31.6	31.3	31.1	30.8	30.6	30.3	30.1
DAYTON	12.6	12.5	12.4	12.3	12.2	12.1	12.0	11.9	11.8	11.7	11.6	11.6	11.5	11.4	11.3	11.2
DEOK	13.5	13.4	13.3	13.2	13.1	12.9	12.8	12.7	12.6	12.5	12.4	12.3	12.2	12.1	12.0	11.9
DLCO	4.4	4.4	4.3	4.3	4.3	4.2	4.2	4.2	4.1	4.1	4.1	4.0	4.0	4.0	3.9	3.9
DOM	109.6	108.7	107.9	107.0	106.2	105.3	104.5	103.6	102.8	102.0	101.2	100.4	99.6	98.8	98.0	97.2
DPL	112.6	111.7	110.8	109.9	109.0	108.2	107.3	106.5	105.6	104.8	103.9	103.1	102.3	101.4	100.6	99.8
EKPC	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
JCPL	381.3	378.2	375.2	372.2	369.2	366.3	363.4	360.4	357.6	354.7	351.9	349.1	346.3	343.5	340.7	338.0
METED	37.2	36.9	36.6	36.3	36.0	35.7	35.4	35.1	34.9	34.6	34.3	34.0	33.8	33.5	33.2	32.9
PECO	49.8	49.4	49.0	48.6	48.2	47.8	47.5	47.1	46.7	46.3	46.0	45.6	45.2	44.9	44.5	44.1
PENLC	5.5	5.5	5.4	5.4	5.3	5.3	5.2	5.2	5.2	5.1	5.1	5.0	5.0	5.0	4.9	4.9
PEPCO	82.9	82.3	81.6	81.0	80.3	79.7	79.0	78.4	77.8	77.2	76.5	75.9	75.3	74.7	74.1	73.5
PL	86.7	86.0	85.3	84.6	83.9	83.3	82.6	81.9	81.3	80.6	80.0	79.3	78.7	78.1	77.5	76.8
PS	585.1	580.5	575.8	571.2	566.6	562.1	557.6	553.1	548.7	544.3	540.0	535.7	531.4	527.1	522.9	518.7
RECO	12.4	12.3	12.2	12.1	12.0	11.9	11.8	11.7	11.6	11.5	11.4	11.3	11.2	11.1	11.1	11.0
UGI	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
PJM RTO	**1,953.8**	**1,938.2**	**1,922.7**	**1,907.3**	**1,892.0**	**1,876.9**	**1,861.9**	**1,847.0**	**1,832.2**	**1,817.5**	**1,803.0**	**1,788.6**	**1,774.3**	**1,760.1**	**1,746.0**	**1,732.0**

Determination of Total Nameplate Capacity of Distributed Solar Generation in the Forecast

To determine the net amount of nameplate capacity of distributed solar generation for each of the forecast years, PJM added the forecasted values and the amount of existing resources. Table IV-8 below shows the total cumulative impact of distributed solar generation for each of the forecast years.

Table IV-8 : Cumulative Nameplate Capacity of Existing and Forecasted Distributed Solar Generation by Zone

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
AE	219.6	233.3	239.2	243.4	248.7	254.5	262.1	272.6	286.4	307.8	337.9	374.3	413.5	457.5	503.1	549.3
AEP	49.9	72.7	96.4	123.2	153.2	186.2	223.7	263.8	306.4	350.4	396.2	433.2	470.4	511.7	559.9	616.1
APS	69.0	92.3	116.5	138.3	155.4	167.6	179.5	192.5	207.1	223.3	241.2	261.7	286.4	316.3	351.8	392.9
ATSI	57.1	71.6	87.7	103.7	119.6	136.5	153.6	170.7	189.2	207.9	226.4	233.6	237.9	242.5	248.7	256.5
BGE	169.3	225.7	283.1	325.4	343.7	349.9	354.1	360.9	372.6	390.0	414.1	446.3	492.0	553.9	622.0	696.0
COMED	71.1	93.2	108.0	122.8	137.5	161.4	191.1	220.6	252.3	284.9	317.1	349.1	380.8	412.3	445.6	481.8
DAYTON	13.8	17.4	21.5	25.6	29.6	34.1	38.7	43.4	48.4	53.5	58.6	60.5	61.4	62.3	63.6	65.2
DEOK	15.3	20.7	26.8	32.9	38.9	45.4	52.3	59.2	66.7	74.4	82.0	84.7	86.0	87.5	89.6	92.3
DLCO	10.6	14.8	19.0	23.7	28.5	31.9	34.8	37.7	40.5	43.3	46.1	48.9	52.1	55.5	60.2	65.6
DOM	223.0	318.8	378.9	456.5	551.0	652.8	775.9	910.8	1,051.3	1,196.7	1,347.2	1,505.3	1,674.1	1,856.4	2,055.6	2,272.9
DPL	136.3	162.9	194.3	225.2	244.4	263.6	295.4	330.9	357.2	384.0	413.0	446.0	487.6	542.2	612.7	705.1
EKPC	0.8	1.0	1.1	1.2	1.3	1.4	1.4	1.5	1.6	1.7	1.7	1.8	1.9	2.0	3.6	5.8
JCPL	441.4	472.0	485.4	495.1	507.2	520.6	537.9	561.5	593.1	641.4	709.8	792.3	881.4	982.0	1,086.5	1,192.6
METED	44.0	48.2	52.7	57.5	62.6	66.1	69.1	72.0	75.0	77.9	80.8	83.7	87.1	90.8	95.9	101.9
PECO	67.5	78.9	90.9	103.9	117.5	127.0	135.1	143.1	151.0	158.8	166.7	174.7	183.5	193.4	207.0	222.7
PENLC	13.1	18.1	23.3	28.9	34.6	38.8	42.2	45.6	49.0	52.3	55.6	58.9	62.6	66.6	72.0	78.4
PEPCO	129.2	170.0	211.6	244.2	260.3	267.8	273.8	281.5	292.4	307.1	326.2	350.7	384.1	428.5	476.9	529.4
PL	104.2	115.2	126.7	139.2	152.3	161.5	169.1	176.6	184.2	191.6	199.0	206.5	214.9	224.3	237.4	252.4
PS	702.7	763.8	790.7	810.7	835.7	862.8	897.8	945.0	1,007.3	1,102.4	1,236.3	1,397.1	1,570.5	1,765.4	1,967.5	2,172.5
RECO	16.4	18.5	19.5	20.2	21.1	22.0	23.3	24.9	27.0	30.2	34.7	40.1	45.9	52.3	59.1	65.9
UGI	0.8	1.1	1.4	1.8	2.1	2.3	2.5	2.7	2.9	3.1	3.3	3.5	3.7	4.0	4.3	4.7
PJM RTO	**2,555.2**	**3,010.4**	**3,374.8**	**3,723.4**	**4,045.1**	**4,354.3**	**4,713.5**	**5,117.9**	**5,561.8**	**6,082.8**	**6,694.1**	**7,352.9**	**8,077.4**	**8,907.3**	**9,823.4**	**10,820.0**

Capacity Factors

Having obtained the nameplate capacity of the distributed solar generation, PJM must apply this to both the energy and peak forecasts. In order to do so, PJM calculated a capacity value at peak, as well as hourly capacity factors for the energy forecasts.

To calculate the capacity value at peak, PJM derived a capacity factor based on the back-casted values. This capacity factor at summer peak represents the ratio of estimated output divided by the nameplate amount taken as an average of hour ending 17:00 over the months of June, July, and August. The capacity factor at fall peaks represents the ratio of estimated output divided by the nameplate amount taken as an average of hour ending 17:00 over the months of September, October, and November. The capacity factor at spring peaks represents the ratio of estimated output divided by the installed amount taken as an average of hour ending 17:00 over the months of March, April, and May. The capacity factor at winter peaks represents the ratio of estimated output divided by the installed amount taken as an average of hour ending 19:00 over the months of December, January, and February. The specific hour selected for each season corresponds to the typical peak hour of that season. The capacity factors applied to the installed capacity for energy purposes represent the hourly average by month and zone. The capacity factor for energy purposes is the estimated output divided by the installed amount at that particular point in time. Table IV-9 below shows the capacity factors for the seasonal peaks that are applied to the installed capacity amount.

Table IV-9 : Capacity Factors at Seasonal Peak by Zone[30]

	Capacity Factors by Season			
	Fall	Winter	Spring	Summer
AE	14%	0%	26%	32%
AEP	14%	0%	18%	22%
APS	11%	0%	17%	20%
ATSI	15%	0%	25%	31%
BGE	11%	0%	18%	21%
COMED	13%	0%	19%	24%
DAYTON	15%	0%	19%	25%
DEOK	14%	0%	19%	24%
DLCO	10%	0%	16%	22%
DOM	13%	0%	20%	23%
DPL	14%	0%	25%	29%
EKPC	21%	0%	32%	32%
JCPL	10%	0%	18%	23%
METED	10%	0%	19%	23%
PECO	9%	0%	16%	19%
PENLC	13%	0%	25%	33%
PEPCO	11%	0%	16%	19%
PL	12%	0%	24%	27%
PS	8%	0%	15%	18%
RECO	8%	0%	15%	18%
UGI	12%	0%	23%	25%

[30] Fall capacity factors are calculated as an average of HE 17:00 during the months of September, October, and November. Winter capacity factors are calculated as an average of HE 19:00 during the months of December, January, and February. Spring capacity factors are calculated as an average of HE 17:00 during the months of March, April, and May. Summer capacity factors are calculated as an average of HE 17:00 during the months of June, July, and August.

Ex-Post Bias Adjustment of Distributed Solar Generation in the Long Term Peak Load Forecast

Table IV-10 below shows the result of multiplying the cumulative installed capacity of the net forecasted and historical distributed solar generation by the derived capacity factor at summer peak.

Table IV-10 : Cumulative Capacity at Summer Peak of Forecasted and Historical Distributed Solar Generation by Zone

	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031
AE	69.3	73.6	75.5	76.8	78.5	80.3	82.7	86.0	90.4	97.1	106.6	118.1	130.5	144.4	158.8	173.3
AEP	11.1	16.2	21.4	27.4	34.1	41.4	49.8	58.7	68.2	78.0	88.2	96.4	104.7	113.9	124.6	137.2
APS	13.7	18.4	23.2	27.5	30.9	33.4	35.7	38.3	41.2	44.5	48.0	52.1	57.0	63.0	70.0	78.2
ATSI	17.8	22.3	27.3	32.3	37.3	42.5	47.8	53.2	58.9	64.8	70.5	72.8	74.1	75.6	77.5	79.9
BGE	35.6	47.4	59.5	68.3	72.2	73.5	74.4	75.8	78.2	81.9	87.0	93.7	103.3	116.3	130.6	146.2
COMED	17.2	22.5	26.1	29.7	33.2	39.0	46.2	53.3	60.9	68.8	76.6	84.3	92.0	99.6	107.6	116.3
DAY	3.4	4.4	5.4	6.4	7.4	8.5	9.7	10.9	12.1	13.4	14.7	15.1	15.4	15.6	15.9	16.3
DEOK	3.7	5.0	6.5	8.0	9.4	11.1	12.7	14.4	16.2	18.1	19.9	20.6	20.9	21.3	21.8	22.4
DLCO	2.3	3.3	4.2	5.2	6.3	7.0	7.7	8.3	8.9	9.6	10.2	10.8	11.5	12.2	13.3	14.5
DOM	50.9	72.7	86.4	104.1	125.6	148.8	176.9	207.7	239.7	272.8	307.1	343.2	381.7	423.2	468.7	518.2
DPL	39.7	47.4	56.6	65.6	71.2	76.8	86.0	96.4	104.0	111.8	120.3	129.9	142.0	157.9	178.5	205.4
EKPC	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	1.1	1.8
JCPL	100.1	107.0	110.0	112.2	115.0	118.0	121.9	127.3	134.5	145.4	160.9	179.6	199.8	222.6	246.3	270.4
METED	10.0	11.0	12.0	13.1	14.2	15.0	15.7	16.4	17.0	17.7	18.4	19.0	19.8	20.6	21.8	23.2
PECO	13.0	15.3	17.6	20.1	22.7	24.6	26.1	27.7	29.2	30.7	32.2	33.8	35.5	37.4	40.0	43.1
PENLC	4.3	5.9	7.6	9.4	11.3	12.6	13.7	14.8	15.9	17.0	18.1	19.2	20.4	21.7	23.4	25.5
PEPCO	24.3	32.0	39.8	45.9	49.0	50.4	51.5	53.0	55.0	57.8	61.4	66.0	72.3	80.6	89.7	99.6
PL	28.5	31.6	34.7	38.2	41.8	44.3	46.3	48.4	50.5	52.5	54.5	56.6	58.9	61.5	65.1	69.2
PS	125.3	136.2	141.0	144.6	149.1	153.9	160.1	168.6	179.7	196.6	220.5	249.2	280.1	314.9	350.9	387.5
RECO	2.9	3.3	3.5	3.6	3.8	3.9	4.1	4.4	4.8	5.4	6.2	7.2	8.2	9.3	10.5	11.7
UGI	0.2	0.3	0.4	0.5	0.5	0.6	0.6	0.7	0.7	0.8	0.8	0.9	0.9	1.0	1.1	1.2
PJM RTO	**573.7**	**676.0**	**759.0**	**839.3**	**913.8**	**986.0**	**1,070.4**	**1,164.7**	**1,266.8**	**1,385.2**	**1,522.7**	**1,669.0**	**1,829.4**	**2,013.2**	**2,217.4**	**2,441.0**

The values shown in this table were then subtracted from the summer peak forecast, in order to adjust the forecast for the impact of distributed solar generation. Table IV-10 corresponds to Table B-8 in the 2016 Load Forecast Report.

V. WEATHER NORMALIZATION

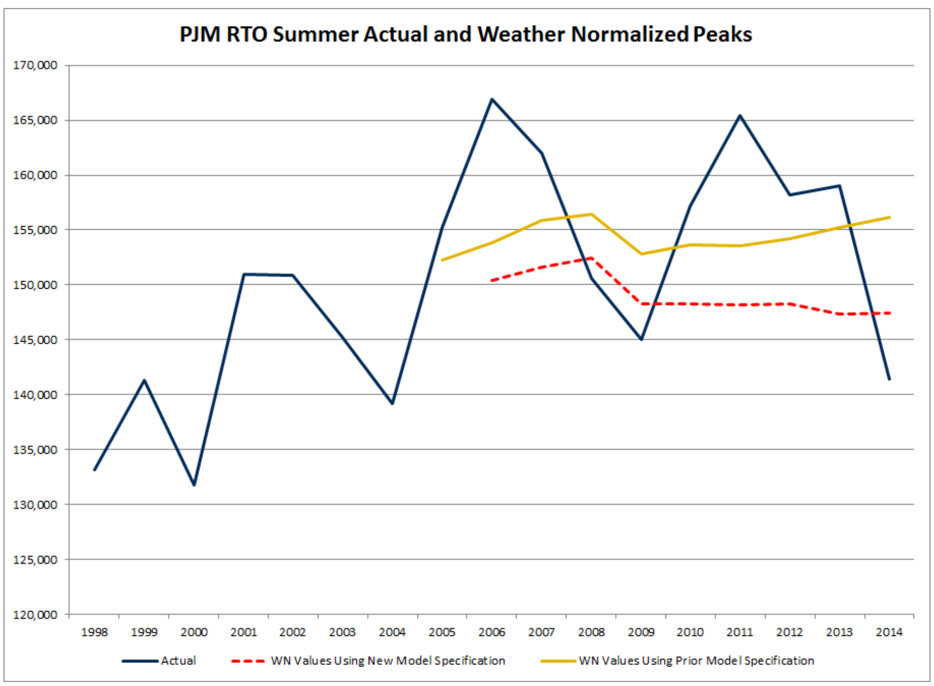
PJM produces Weather Normalized (WN) values for each zone's summer and winter non-coincident and coincident peaks. Weather normalized loads are not explicitly used in PJM's planning or load forecast modeling processes. The intent of WN loads is to indicate a long-term trend in each zone's seasonal peak loads and to help understand the forecast path.

Background of Change

Previously, PJM used its zonal peak models to derive the official weather normalized peak for each season. After each season, each zonal NCP and CP model was re-estimated, adding the most recent historical load, weather, and economic data. The weather simulation process was then run, including historical weather through the just-completed season. For the RTO WN value, the zonal results were aggregated and from the resulting distribution of results the median value was selected as the weather normalized seasonal peak.

With the significant model changes adopted in 2015, PJM's method of weather normalizing peak loads using the forecast models resulted in a considerably different weather-normalized history.

Figure V-1 : PJM RTO Model-Driven Weather Normal Comparison



As a result, PJM investigated a weather normalization approach that was not tied to the model but would produce meaningful historical load trends for each zone.

New Process for Summer & Winter

PJM separately weather normalizes the summer and winter seasons, though the process is identical for each season. PJM began the investigation for a new weather normalization process by revisiting the former bottom-up weather normalization procedure developed by the Load Analysis Subcommittee and used until 2007. In that method, the zone's daily peak unrestricted loads for a season were regressed against a weather parameter, and the resulting equation was solved at the zone's weather standard (the long-term average of weather conditions on peak days). The months included in the summer regression were June, July and August. The weather parameter used in the summer was a two-day weighted temperature humidity index (THI) with a weighting of 4 for the current day (t) and 1 for the prior day (t-1).

Equation V-1 : Summer Weather Parameter – Weighted THI

$$\mathbf{WTHI = (4*maxTHI_t + 1*maxTHI_{(t-1)})\ /5}$$

If DB ≥ 58,

$$THI = DB - 0.55 * (1 - HUM) * (DB - 58)$$

If DB < 58,

$$THI = DB$$

Where: THI = Temperature humidity index;

DB = Dry bulb temperature (°F);

HUM = Relative Humidity (where 100% = 1)

The months included in the winter regression were December, January and February. The weather parameter used in the winter was a two-day weighted wind-adjusted temperature called winter weather parameter (WWP) at hour ending 19:00 with a weighting of 4 for the current day (t) and 1 for the prior day (t-1).

Equation V-2 : Winter Weather Parameter – Weighted WWP

$$\mathbf{WWWP = (4*WWP_{hr19t} + 1*WWP_{hr19(t-1)})\ /5}$$

If WIND > 10 mph,

$$WWP = DB - (0.5 * (WIND - 10))$$

If WIND ≤ 10 mph,

$$WWP = DB$$

Where: WIND = Wind velocity, in miles per hour;

WWP = Wind speed adjusted dry bulb temperature;

DB = Dry bulb temperature (°F)

PJM chose to retain the weather parameters from the prior method and then investigated using a single year or three year history for the regression. The single year analysis resulted in slightly more volatile weather normalized series as sudden shifts in load growth and extreme weather in a single year dominated the results. The three year history, which controls for inter-year load growth, smoothed the weather normalized series and allowed for more weather history in the regression.

Weather Standard

PJM considered a number of methods to produce the weather standards at which the three year regression equations were solved. The time periods reviewed were 1998 to 2014 to be consistent with the historical load period used and 1994 to 2014 to be consistent with the historical years used in the weather simulation. In order to keep the weather standard in line with the forecast weather distribution, PJM decided to use the time period of 1994 to the most recent year available.

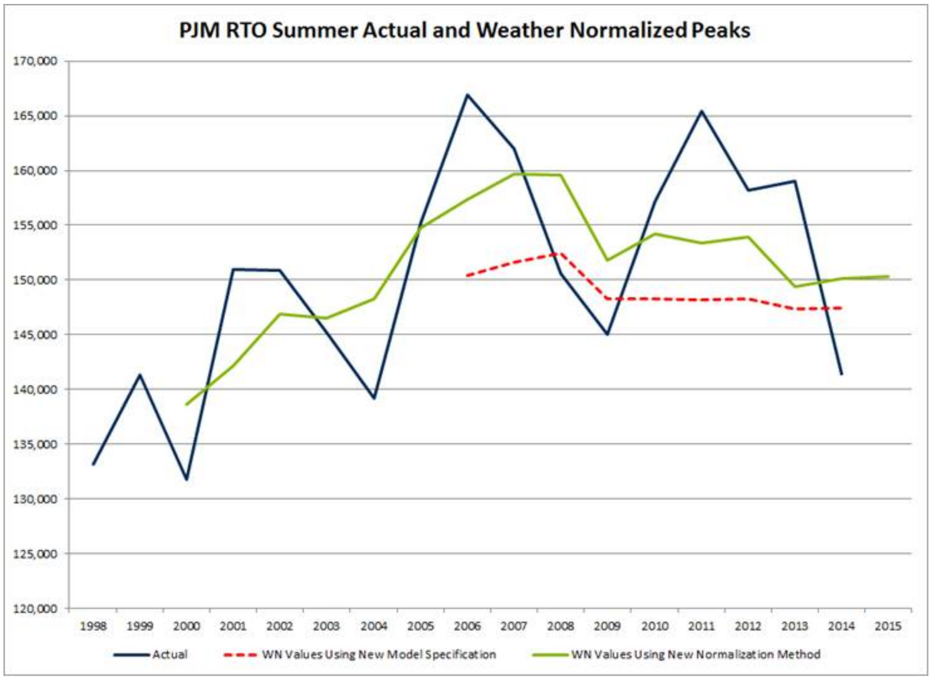
Next, the pool of days to pick the maximum weather standard was examined. Three categories of weather days were considered. First, only peak day weather; second, seasonal extreme weather; and finally, weather that occurred on non-holiday weekdays were investigated. Using only weather that occurred on peak days produces a weather standard that is lower in the summer and higher in the winter since extreme weather on non-peaking days is not captured. By using seasonal extremes for the weather standard the weather normalized series may be inflated and thus overstate the long term trend of historical load and first year load growth. To balance possible and probable weather, PJM chose to use weather from non-holiday weekdays.

PJM had the option of using a rolling number of years in the weather standard or a static number of years. If a rolling number of years is used then trends in weather would be picked up but the weather standard would be slightly different each year. By using a static number of years in the weather standard the weather normalized loads indicate load changes but not weather trends. It was decided to use a static number of years with the intent to occasionally revisit and, if necessary, revise the number of years used in the weather standard. This method will have the effect of producing updated historical weather normalized loads each year as every three year regression will be solved at the updated weather standard for all historical years. Using a static number of years has the added benefit of being similar to the load forecast and thus helps in comparing the WN values to the forecast.

Finally, the method of using either an average or median to compute the weather standard was reviewed. Using the average or median reflects central tendency and will only differ significantly if the distribution is not normal. PJM decided to use the average as it is an industry standard and using either average or median produced similar results.

The resulting weather normalization procedure (illustrated in Figure V-2) produces a series of values which successfully meets the goal of indicating the long-term trend in each zone's seasonal peak. However, its ability to accurately portray first year growth in the forecast will be compromised following a year in which a zone experienced significantly milder or harsher weather than average.

Figure V-2: Comparison of New and Model-Driven Weather Normalization Procedures



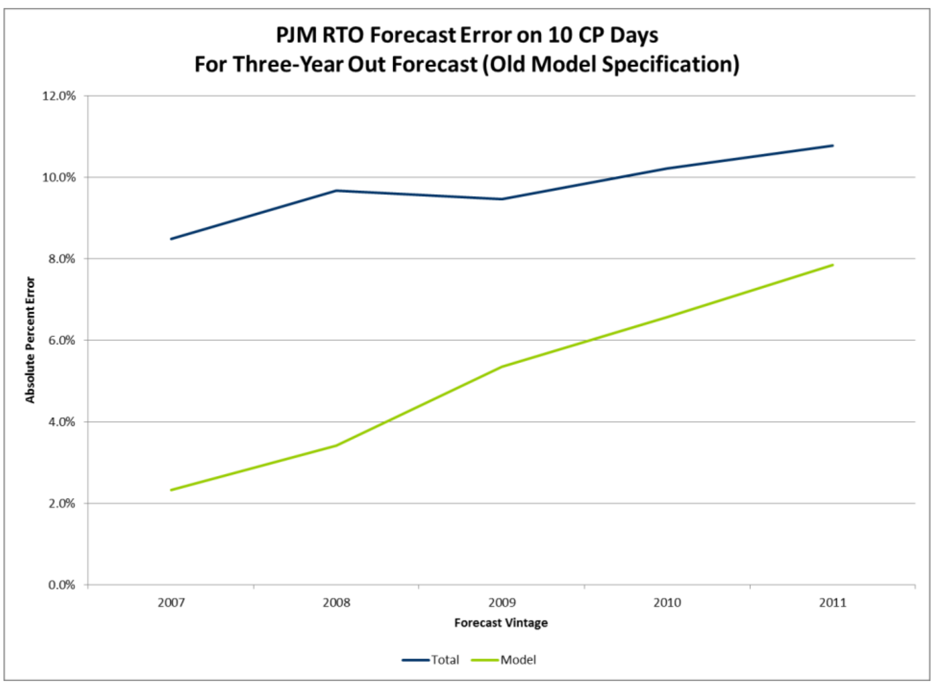
APPENDIX A. ACCURACY ANALYSIS

The catalyst for extensive model review was the growing concern of increasing forecast error. PJM has been producing an independent forecast since 2006. This short period of time limits the number of data points from which to draw firm conclusions about accuracy, especially when perhaps the most critical points of the long-term forecast are 3 and 5 years out (the RPM and RTEP years, respectively). However, the trend seemed clear that error was high and increasing.

When considering accuracy, the total effective error comprises the economic error and the model error. PJM does not forecast economics, and instead uses a series of economic drivers forecast by Moody's Analytics. As a result, there is some error that is outside of the modeler's control. Model error is the remaining error, once some control is put in place for economic error.

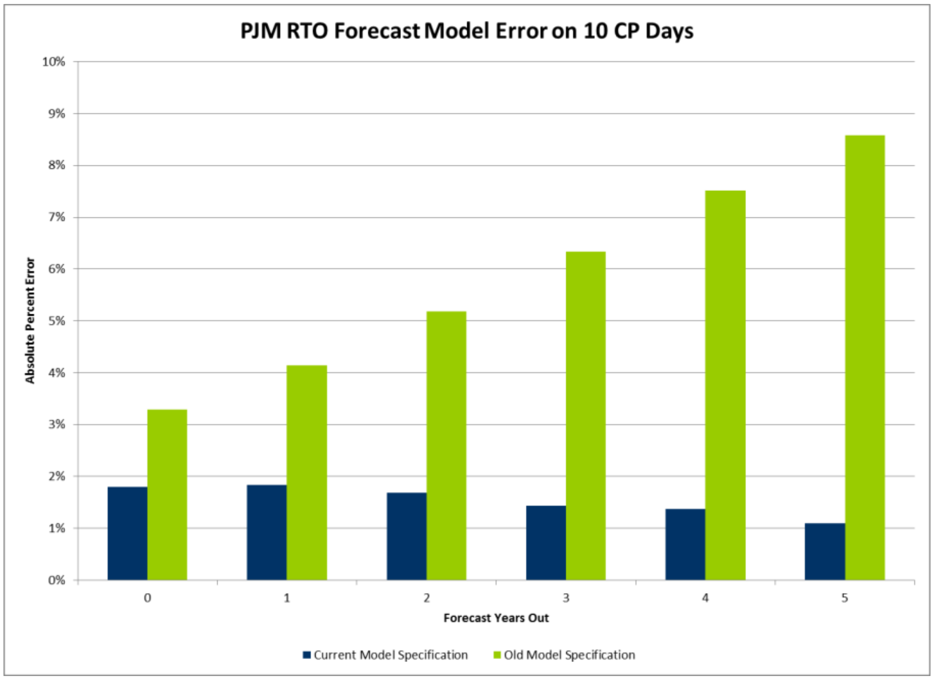
In Figure A-1, total error and model error for the three-year out forecast are plotted. Total error is computed using only the information that would have been available at the time the forecast was generated. Model error is calculated in a similar way, but instead of using economics available at the time, a more recent snapshot of economics is used. In each case, the models are evaluated as to how they would perform on the 10 highest load days of each year.

Figure A-1 : Three-Year Out Forecast Error (Old Model Specification)



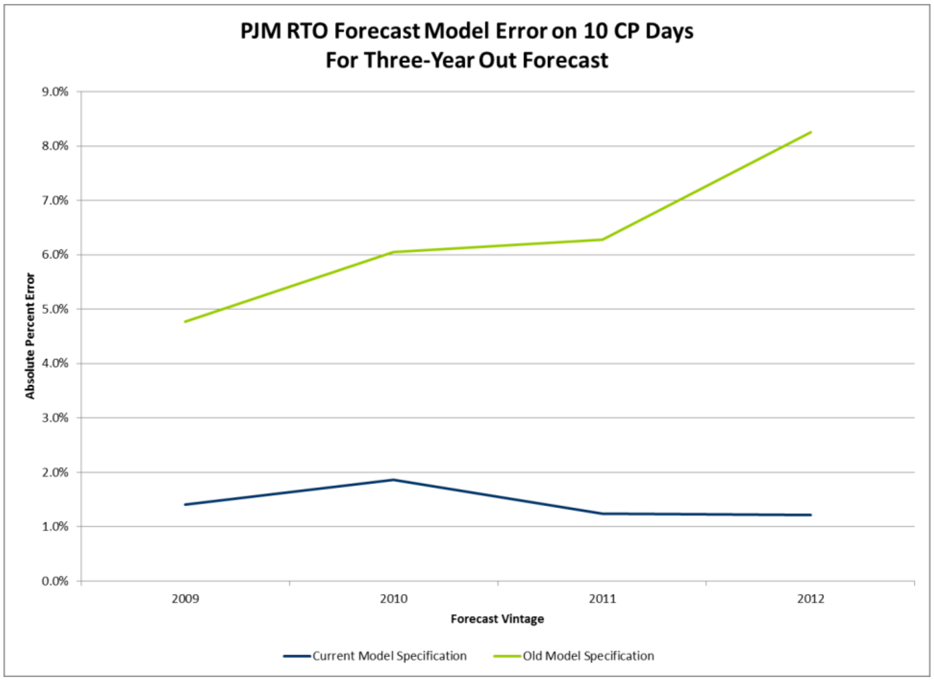
Model error increased with successive forecast updates. In the early years of PJM producing a load forecast (which coincided with the Great Recession in 2007 to 2009), much of the forecast error was attributable to inaccuracy in the economic forecast. However, more recent years were showing this to not be the case and error was increasingly attributable to the model. As model development progressed throughout 2015, improvement of model error was the identified metric for comparison purposes. The current model specification showed significant improvement in model error (see Figure A-2 and Figure A-3)[31].

Figure A-2 : Forecast Model Error (Current versus Old)



[31] In the Figures, forecast performance is observed over 10 days for each forecast for each year out. Thus, there are a different number of points used in each calculation. The performance at 0 years out is calculated over 70 points, whereas the performance at three years out is calculated over 40 points, and so forth.

Figure A-3: Three-Year Out Forecast Model Error (Current versus Old)



APPENDIX B. DECOMPOSITION AND SENSITIVITY ANALYSIS

Decomposition Analysis

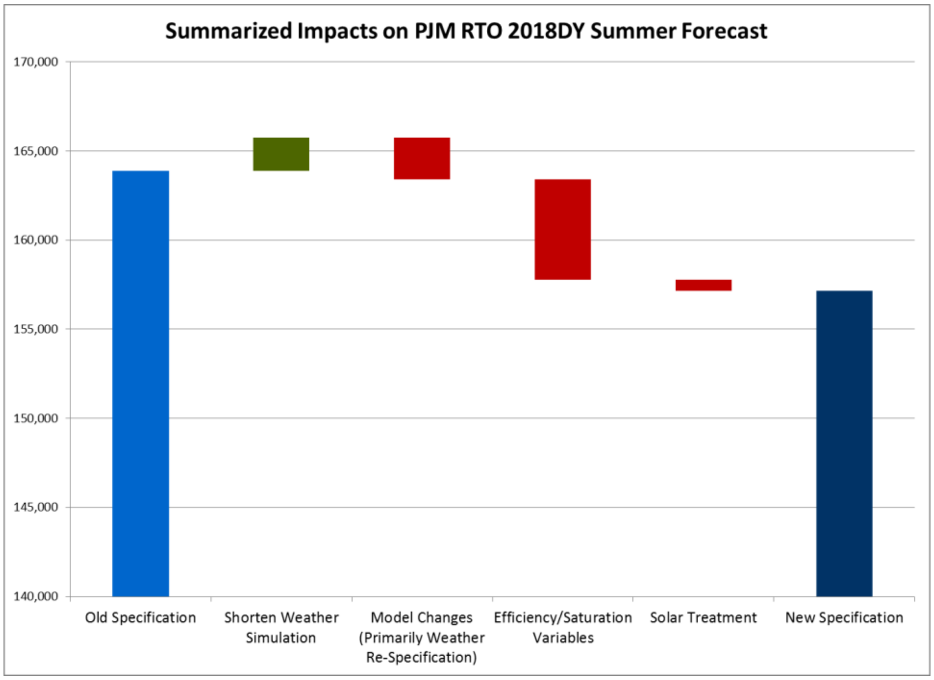
Each change made throughout the load forecast model investigation had varying degrees of significance. The analysis that follows breaks down the model changes into distinct categories to isolate the forecast impact of each improvement. Analysis was conducted using the 2015 forecast model as the base case, with some slight differences from the official forecast released in January 2015. As a result, the starting point of the analysis is not identical to what was presented in that report. The differences are:

- Economics were updated to the Moody's Analytics July 2015 release;
- No binary variable adjustment was made;
- No forecast adjustments were applied.

The breakdown of the impacts on the summer forecast can be seen in Figure B-1. The new specification has a forecast that is 4.1% below the forecast that would have been produced using existing practices at the time. The introduction of the new equipment index variables that capture trends in saturation and efficiency were the single largest contributor to the forecast reduction (-3.4%). Various model changes, which include refinements to the weather specification as well as the introduction of an autoregressive (AR(1)) error term were the next largest contributor to the forecast reduction (-1.4%). Also acting to lower the forecast was the treatment of distributed solar generation (-0.4%)[32]. Partially offsetting these reductions is the shorter weather simulation period (+1.1%).

[32] The forecast reduction attributed to distributed solar generation is calculated using the solar forecast that was in place for the 2016 Load Forecast, as no prior PJM distributed solar energy resource forecast is available.

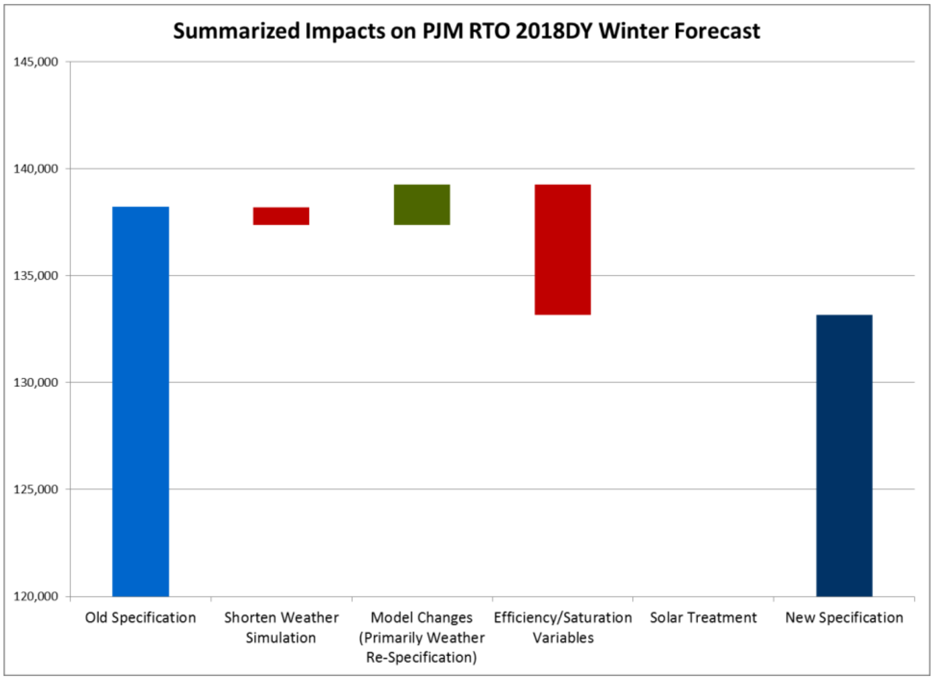
Figure B-1 : Component Decomposition – Summer Forecast



It is important to keep in mind that these contributions are calculated for a particular point in time (three years out). Earlier years in the forecast would show less of a reduction as there would be less time for equipment efficiency gains and distributed solar generation to influence the forecast. Similarly, further years would show a larger reduction as these types of trends become more significant.

The breakdown of the impacts on the winter forecast can be seen in Figure B-2. The new specification has a forecast that is 3.7% below the forecast that would have been produced using existing practices at the time. The introduction of the new equipment index variables that capture trends in saturation and efficiency were the single largest contributor to the forecast reduction (-4.4%). A shorter weather simulation also acted to reduce the forecast (-0.6%) as winter peak weather conditions have tended to be milder in the most recent 20 years than in the 20 years prior. Various model changes–including refinements to the weather specification and the AR(1) error term—offset some of these changes (+1.4%). This indicates that the old weather specification understated the elasticity of load to weather in severe conditions. There is no impact from solar treatment as the output from distributed solar generation at the time of the winter peak (typically 7:00 PM) is forecast to be 0 MW.

Figure B-2 : Component Decomposition – Winter Forecast



Sensitivity Analysis

With the inclusion of end-use variables in the load forecast model, the forecast is now influenced by input from another outside information source in addition to the economic forecast already provided by Moody's Analytics. The end-use variables obtained from EIA data include projections of both equipment saturation and equipment efficiency. These variables affect the load forecast in two major ways. First, they better align the forecast with historical trends which lowers the starting point beyond what it otherwise would be. Second, they augment the forecast with EIA's projections for future end-use behavior.

Better alignment with historical trends is an important step forward. The previous load forecast model was arguably biased high in recent years as it was failing to adequately capture the rapid evolution of end-use behavior. More accurately capturing the forecast starting point increases the model's ability to project the loads in the out-years more accurately.

The EIA end-use projections also impact the shape of the forecast going forward. Three sensitivities were run to capture these impacts:

- Hold efficiency constant on all equipment at 2015 levels;
- Hold saturation constant on all equipment at 2015 levels;
- Hold saturation and efficiency constant on all equipment at 2015 levels.

Figure B-3 below shows these sensitivity results for the summer peak forecast. The Cooling and Other Equipment Indexes are the measures of consequence in the summer period. The equipment efficiency projections pull the forecast downward on average by 0.6% per year over the next five years. This is somewhat offset by increasing equipment saturation overall[33] that adds on average 0.2% per year over the next five years. Overall, the total impact of the saturation and efficiency projections is 0.4% per year over the next five years compared with these two projections being held constant.

Figure B-3 : End-Use Variable Sensitivities – Summer Forecast

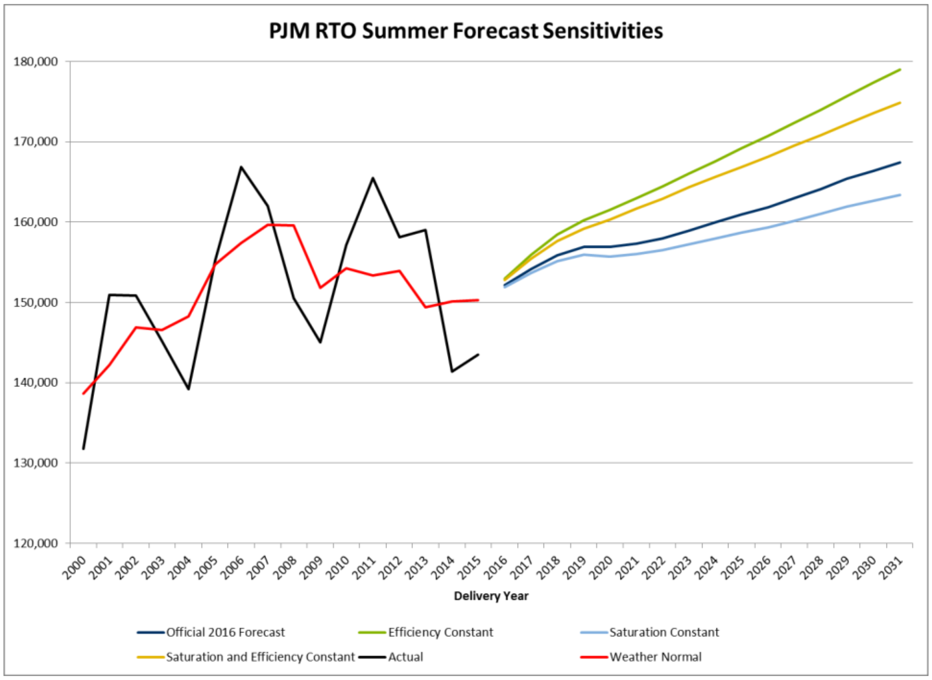
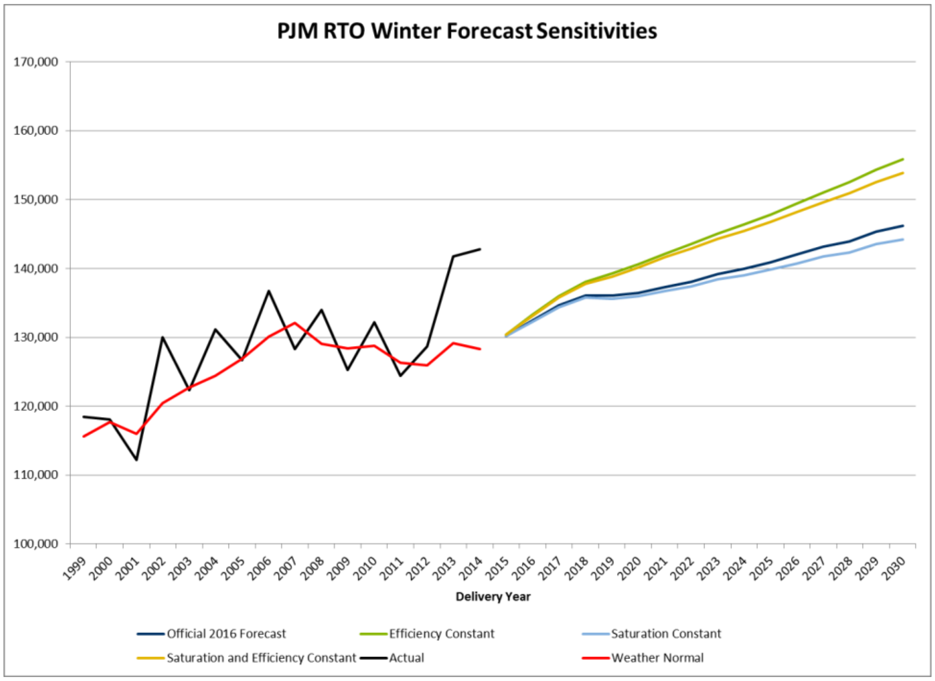


Figure B-4 below shows these sensitivity results for the winter peak forecast. The Heating and Other Equipment Indexes are the measures of consequence in the winter period. The equipment efficiency projections pull the forecast downward on average 0.5% per year over the next five years. This is somewhat offset by increasing equipment saturation overall that add on average 0.1% per year over the next five years. Overall, the total impact of the saturation and efficiency projections is 0.5% per year over the next five years compared with these two projections being held constant.

[33] Saturation of some equipment goes up while others go down, but the net impact is positive. For instance, most commercial equipment types have declining saturation, whereas most residential equipment/appliance types have increasing saturation. An interesting exception is the future declines in room air conditioners, though this is more than offset by increases in other residential cooling types.

Figure B-4 : End-Use Variable Sensitivities – Winter Forecast



The notable difference between the summer and winter seasons is that the winter has less upward pressure from increasing saturation. This is because the prevalence of natural gas heating alternatives impacts heating saturation rates. There is not a significant non-electric equivalent in the cooling space.

APPENDIX C. SEER TO EER CONVERSION

The end-use data supplied by Itron to PJM defines efficiency for residential central air conditioning units and heat pumps by their Seasonal Energy Efficiency Ratio (SEER). Throughout the 2015 model development process, some stakeholders contracted with a consultant, Navigant, who raised the concern that the use of SEER could lead to overstating efficiency improvements and that the Energy Efficiency Ratio (EER) was a more appropriate metric[34]. Navigant conducted analysis to convert SEER to EER and provided it to PJM, and PJM verified their analysis.

This is the only time that SEER ratings for central air conditioners and heat pumps will need to be manually converted to EER ratings. Going forward, Itron will be making EER ratings a part of the standard annual release through the Energy Forecasting Group.

Rationale for EER instead of SEER

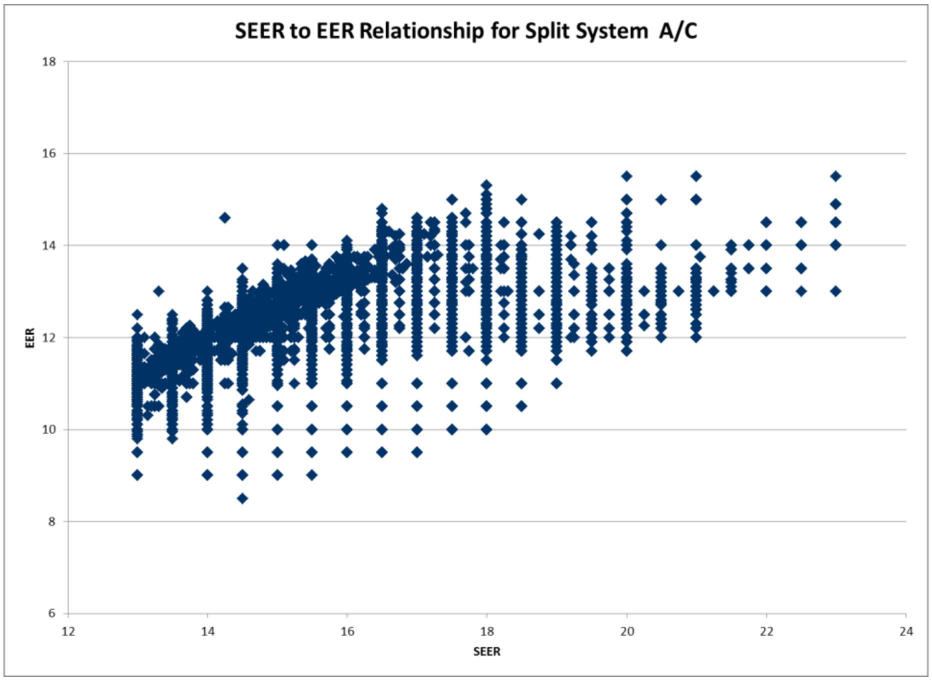
The rationale for using EER instead of SEER can be traced to the intention of each of these metrics. SEER is meant to capture the efficiency of a cooling unit over the course of the entire cooling season. This measure is calculated as a ratio of cooling output during a typical year to the energy consumed. SEER would best be described as measuring seasonal average performance.

The primary goal of the PJM load forecast is to anticipate load under system peak conditions, not daily peaks under average conditions. EER measures unit performance at a single outside temperature of 95 degrees, more in line with this goal. This would not be an issue if the relationship between SEER and EER ratings were static, but the data does not support this. Figure C-1 shows split system air-conditioning unit records data (roughly 1 million records) from the Air-Conditioning, Heating, and Refrigeration Institute (AHRI)[35].

[34] Navigant presented information on their findings at the November 30, 2015 Load Analysis Subcommittee meeting: http://www.pjm.com/~/media/committees-groups/subcommittees/las/20151130/20151130-item-05-navigant-presentation.ashx

[35] AHRI website: https://www.ahridirectory.org/ahridirectory/pages/home.aspx

Figure C-1 : SEER versus EER– Central A/C Units

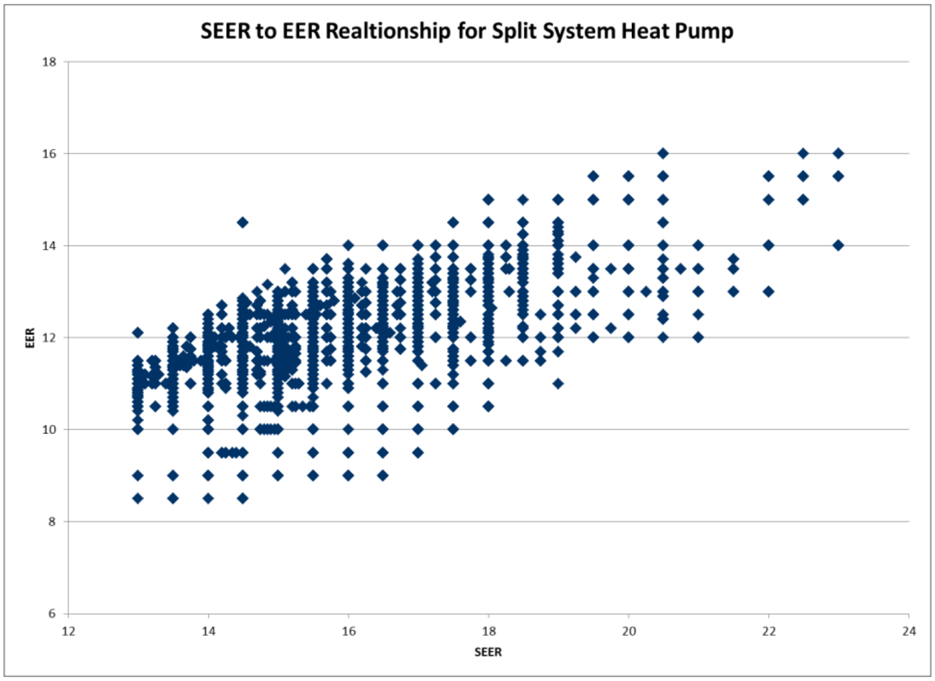


While there has been some movement towards EER-based standards[36], SEER ratings remain the predominant metric used to gauge air conditioner efficiency. This is sensible as SEER ratings are the metric more indicative of a consumer's annual expenses. As a result, there is strong incentive for manufacturers to increase SEER without necessarily increasing EER.

The figure above shows that for a given SEER rating, there is a wide range of potential EER ratings. Also, the EER to SEER ratio decreases with higher SEER ratings. A typical 13 SEER unit has an associated EER rating of 11, a ratio of 0.85. Whereas a typical 18 SEER unit has an associated EER rating of 13, a ratio of 0.72. The data on split system heat pumps tells a similar story (see Figure C-2).

[36] Arizona, California, Nevada, and New Mexico have corresponding EER requirements as a standard.

Figure C-2 : SEER versus EER – Heat Pump Units



Converting SEER to EER

Average SEER ratings provided by Itron need to be converted to EER ratings to reflect the relationship described above. This conversion requires the use of two sets of equations: one set for before 2006 and one after 2006. Two sets of equations are required because the historical timeframe of the data spans two sets of standards. From 1992 to 2005, the minimum SEER for central air conditioners and heat pumps was 10. Starting in 2006, a new standard pushed the minimum SEER to 13. This standard was eventually amended such that heat pumps as of 2015 need to have a minimum SEER of 14.

For SEER ratings less than 13, data is gathered from technical documentation associated with the rule change that took effect in 2006[37]. For SEER ratings 13 and higher, the aforementioned unit data from the AHRI is used. In both cases, the model considers the median unit and fits a second order polynomial equation to the data. A second order polynomial is used rather than a linear trend to account for the declining EER to SEER ratio discussed above. This can be observed in Figure C-3 and Figure C-4, and the equations in Equation C-1.

[37] U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (2002). Technical Support Document to Energy Conservation Standards for Central Air Conditioners and Heat Pumps (Docket No. EERE-2006-STD-0089). Chapter 4 Engineering Analysis, Table 4.27. http://www.regulations.gov/#!documentDetail;D=EERE-2006-STD-0089-0371

Figure C-3 : SEER to EER Relationship – Central A/C

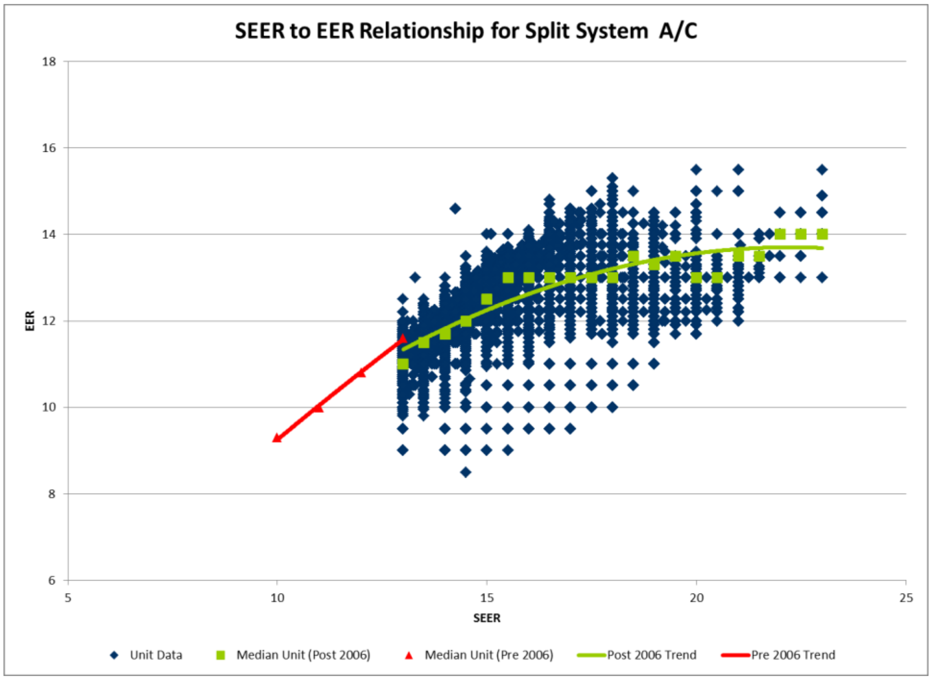
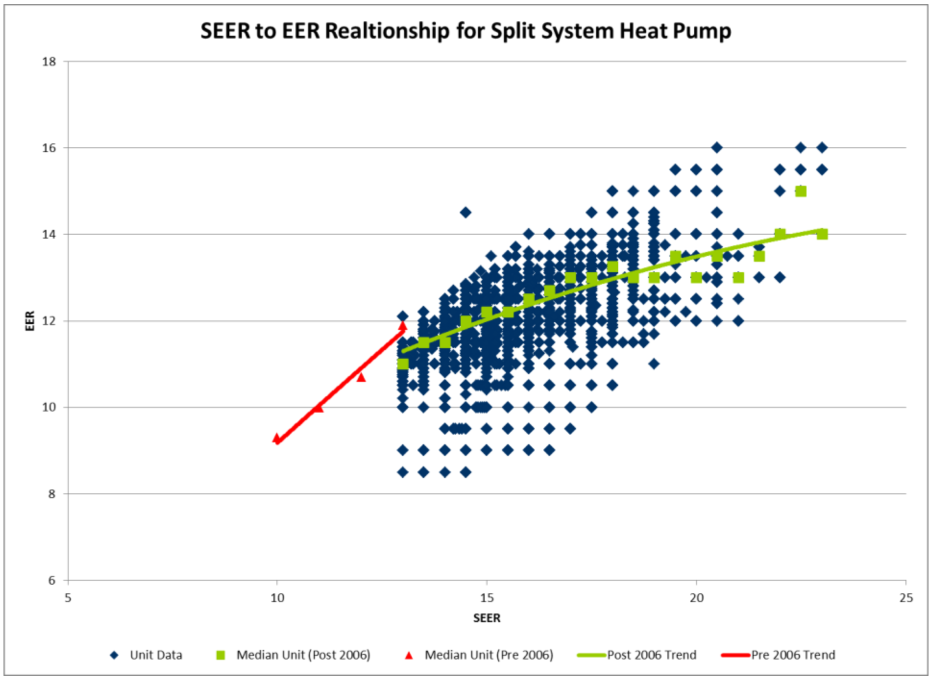


Figure C-4 : SEER to EER Relationship – Heat Pumps



Equation C-1 : SEER to EER Conversion Equations

Central Air Conditioners

Before 2006: EER = -0.0118 x SEER2 + 1.043 x SEER
2006 and after: -0.0274 x SEER2 + 1.227 x SEER

Heat Pumps

Before 2006: EER = -0.0045 x SEER2 + 0.9635 x SEER
2006 and after: -0.0267 x SEER2 + 1.1954 x SEER

The Itron data spreadsheets, arranged by Census Division, contain information on average SEER, the saturation rate, and the number of households. The combination of the latter two provides the number of central air conditioners or heat pumps. With this information, PJM derives an estimate of average EER by equipment type using a multi-step process.

Step 1: Estimate the number of cooling units that retire each year, and the number of cooling units added each year. Cooling units are assumed to have a lifetime of 19 years.

Equation C-2: Calculating New and Retired Cooling Units (Step 1)

$$\text{Retired\_Units}_t = \text{Total\_Units}_t / 19$$
$$\text{New\_Units}_t = \text{Retired\_Units}_t + (\text{Total\_Units}_{t+1} - \text{Total\_Units}_t)$$

Step 2: Calculate the Average EER in the first historical year (1995) using SEER-to-EER conversion equations for before 2006.

Step 3: Estimate the average SEER of new cooling units. This is done by resolving algebraically the identity that the average SEER rating in one period is related to the average SEER rating in the prior period after adjusting for new and retired cooling units.

Equation C-3 : Calculate Average SEER of New Units (Step 3)

$$\text{Avg\_SEER}_t \times \text{Total\_Units}_t = (\text{Avg\_SEER}_{t-1} \times \text{Total\_Units}_{t-1}) - (\text{Avg\_SEER}_{t-1} \times \text{Retired\_Units}_{t-1}) + (\text{Avg\_New\_SEER}_{t-1} \times \text{New\_Units}_{t-1})$$

which algebraically can resolve to:

$$\text{Avg\_New\_SEER}_{t-1} = ((\text{Avg\_SEER}_t \times \text{Total\_Units}_t) - (\text{Avg\_SEER}_{t-1} \times \text{Total\_Units}_{t-1}) + (\text{Avg\_SEER}_{t-1} \times \text{Retired\_Units}_{t-1})) / \text{New\_Units}_{t-1}$$

Step 4: Translate the average SEER of new cooling units (Avg_New_SEER) into the average EER of new cooling units (Avg_New_EER) using SEER-to-EER conversion equations.

Step 5: Calculate the average EER of total cooling units. This is done by resolving algebraically the identity that the average EER rating in one period is related to the average EER rating in the prior period after adjusting for new and retired cooling units.

Equation C-4 : Calculate Average EER of Total Cooling Units (Step 5)

$$Avg\_EER_t \times Total\_Units_t = (Avg\_EER_{t-1} \times Total\_Units_{t-1}) - (Avg\_EER_{t-1} \times Retired\_Units_{t-1}) + (Avg\_New\_EER_{t-1} \times New\_Units_{t-1})$$

which algebraically can resolve to:

$$Avg\_EER_t = ((Avg\_EER_{t-1} \times Total\_Units_{t-1}) - (Avg\_EER_{t-1} \times Retired\_Units_{t-1}) + (Avg\_New\_EER_{t-1} \times New\_Units_{t-1})) / Total\_Units_t$$

The average EER calculated using these described methods is then used to replace the average SEER metric and the resulting Cooling Equipment Index is calculated.

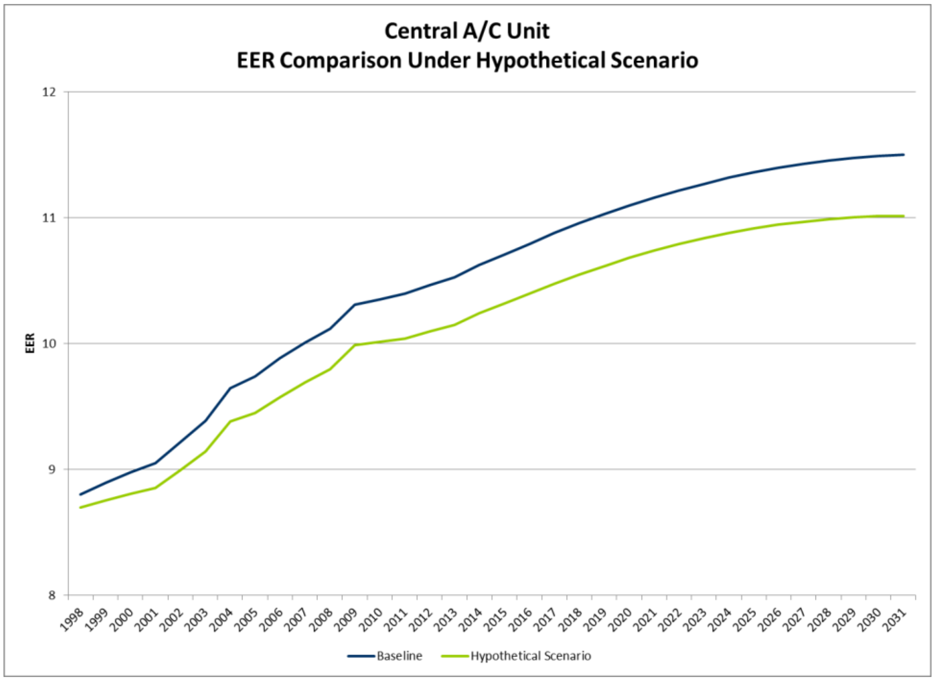
Assumptions Discussion

Aside from the equations derived to map SEER to EER in Equation C-1, there are two additional assumptions made that should be discussed. First, the process assumes that the retiring cooling units have average efficiency. Second, the assumption is made that a cooling unit's lifetime is 19 years. Neither of these assumptions has a significant impact on the analytical results.

Consider the assumption that the retiring cooling units have average efficiency. This is probably not the case as it is intuitive to think that a cooling unit being replaced is likely to be less efficient than average in the year it is replaced. To gauge the impact of this assumption, consider the hypothetical scenario that a retiring cooling unit has 95% the efficiency rating of the average unit in the year it is retired.

Assuming a lower efficiency of retiring cooling units has a couple implications. First, the remaining cooling units implicitly must be more efficient than the average in that year. And consequently, new cooling units are less efficient under this assumption than they otherwise would be and the resulting average EER shifts downward. The results of this hypothetical case can be seen in Figure C-5.

Figure C-5 : EER Comparison – Hypothetical Scenario



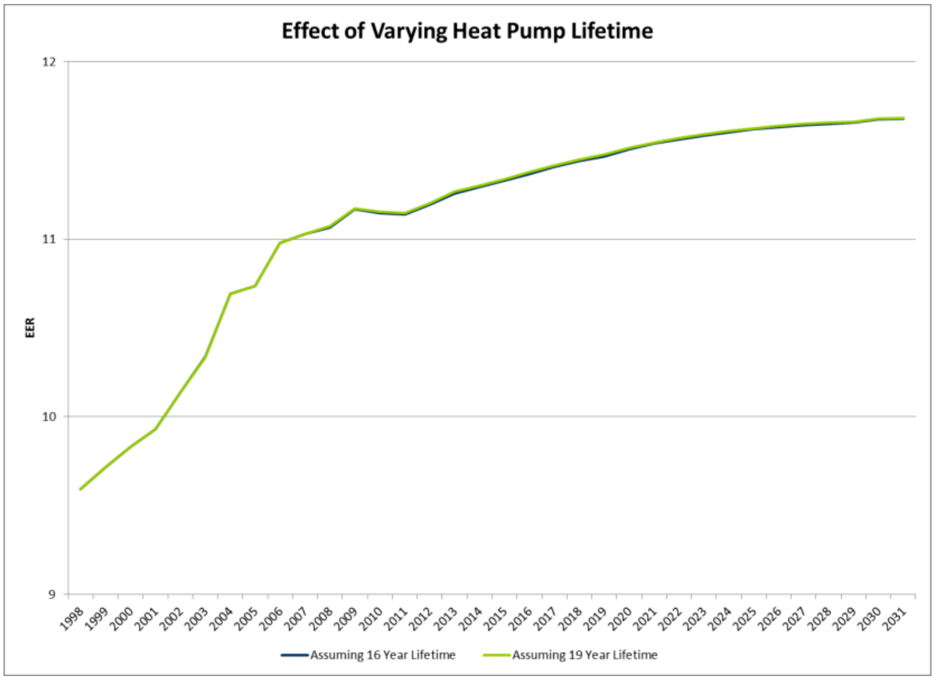
In the construction of the Cooling Equipment Index, it is really the trend that matters as efficiency is captured in how it moves relative to the base year (1998). EER improves at an annualized rate of 1.1% from 1998 to 2015 under the baseline compared with 1% in the hypothetical scenario. However in the forecast period, there is no noticeable difference, as both grow at an annualized rate of 0.6% from 2015 to 2020.

To test the impact on the forecast, these hypothetical EER figures for central air conditioning units and heat pumps were plugged into the equipment index. The results indicate that the forecast would have been 0.1% higher in all years, a negligible impact.

The second assumption to consider is the 19 year average lifetime of cooling units. This assumption is supported by technical documentation[38] produced by the Department of Energy as a part of the amended standards process that took place in 2011, which states that "the mean lifetime for central air conditioners is 19.01 years, for heat pumps is 16.24 years". For simplicity, 19 years was used as the lifetime for both central air conditioners and heat pumps. Assuming a 19 year lifespan for heat pumps rather than 16 years does not make a material difference, as the historical and forecast estimated EER values under either assumption are indistinguishable from one another (see Figure C-6).

[38] U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (2011). Technical Support Document: Energy Efficiency Program for Consumer Products: Residential Central Air Conditioners, Heat Pumps, and Furnaces (Docket No. EERE-2011-BT-STD-0011-0012). Chapter 8 Life-Cycle Cost and Payback Period Analysis, pp 59-67. http://www.regulations.gov/#!documentDetail;D=EERE-2011-BT-STD-0011-0012

Figure C-6 : Effect on Average Efficiency of Varying Unit Lifetime



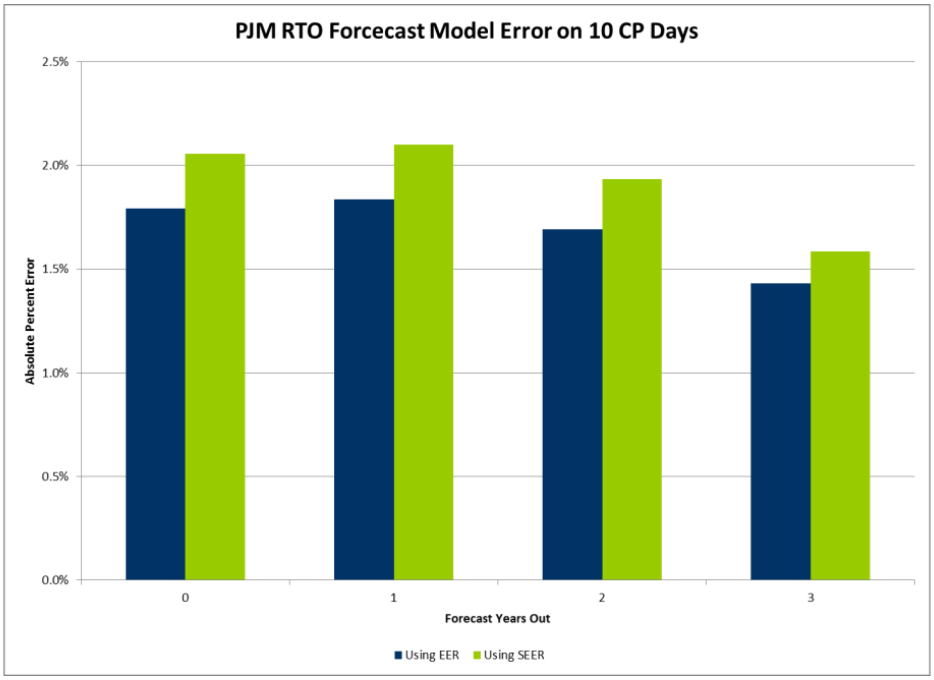
Verifying Accuracy

Adopting EER as the efficiency metric over SEER made intuitive sense, but PJM wanted to confirm that making the switch did not compromise the accuracy of the forecast model. The forecast model was run under two sets of conditions and evaluated for accuracy on the 10 highest load days of the year. The two conditions are:

- Keeping SEER as the efficiency measure for Central A/C and Heat Pumps
- Converting SEER to EER and using EER as the efficiency measure for Central A/C and Heat Pumps

On average, the EER-based approach has been about 10% more accurate than the SEER-based approach in the zero to three year-out forecasts (see Figure C-7). This confirmed that, not only was EER logically correct for use in peak load forecasting, but it was also more accurate.

Figure C-7 : Accuracy Comparison – Using EER Instead of SEER



APPENDIX D. INCORPORATING ZONAL SATURATION DATA

Unless otherwise supplied, the current practice is to assign each zone to a Census Division for its equipment saturation and equipment efficiency data. However, many zones still maintain a load research function and periodically survey their customers regarding their appliances. If this data can improve model results, it was deemed worthwhile to use. Zones were solicited to supply end-use saturation data, and 10 supplied information: AEP, APS, ATSI, ComEd, Dominion, Duke, EKPC, JCPL, MetEd, and PENLC.

The zonal data was then pulled into the existing framework as it was important that the data be compatible with the Itron/EIA data that would remain the basis for the forecast. In this manner, the forecast data for the zones that supplied information would remain comparable to those that did not in order to ensure impartiality. Forecasted saturation data from the allocated Census Division are applied proportionately to the zones as is described in Equation D-1.

The lone exception to the proportional treatment is in the case where a zone has a higher saturation of a particular equipment type than the Census Division, and that equipment type is growing. In this case, the zonal saturation will increase at the same level rate as the Census Division. This is to guard against saturation rates growing out of control. A zone having a higher concentration of a particular equipment type is a snapshot in time, and is not necessarily indicative that its saturation rate should grow at a faster rate in perpetuity. In addition, new historical annual average usages or intensities (used as weights in the Equipment Index calculations) are calculated by type based on the zonal saturation relative to the Census Division saturation.

Equation D-1 : Forecasted Saturation Rates with Zonal Supplied Data

$$\mathbf{Sat_{Zone,Type,t} = Sat_{Census,Type,t} \times (Sat_{Zone,Type,b} / Sat_{Census,Type,b})}$$

$$\textit{Except when } \mathbf{(Sat_{Zone,Type,b} > Sat_{Census,Type,b})} \textit{ and } \mathbf{(Sat_{Census,Type,t} - Sat_{Census,Type,t-1}) > 0}$$

$$\textit{Then } \mathbf{Sat_{Zone,Type,t} = Sat_{Zone,Type,t-1} + (Sat_{Census,Type,t} - Sat_{Census,Type,t-1})}$$

Where
Sat = Saturation, defined by year and equipment type
Type = Equipment/Appliance type (see types listed earlier in this document)
Zone = Transmission Zone
Census = Census Division

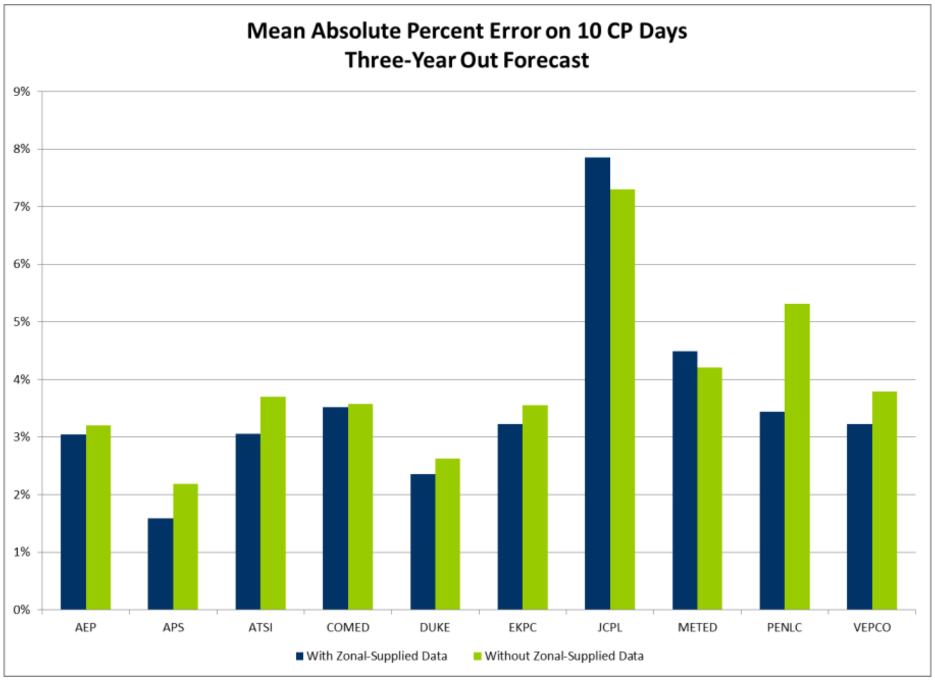
$$\mathbf{Intensity_{Zone,Type,t} = Intensity_{Census,Type,t} \times (Sat_{Zone,Type,t} / Sat_{Census,Type,t})}$$

Where
Intensity = Average Annual Use

The resulting saturation rates are then used to calculate the equipment indexes as described earlier in the document. In the event that data on a particular equipment type is not supplied, then Census Division data is used.

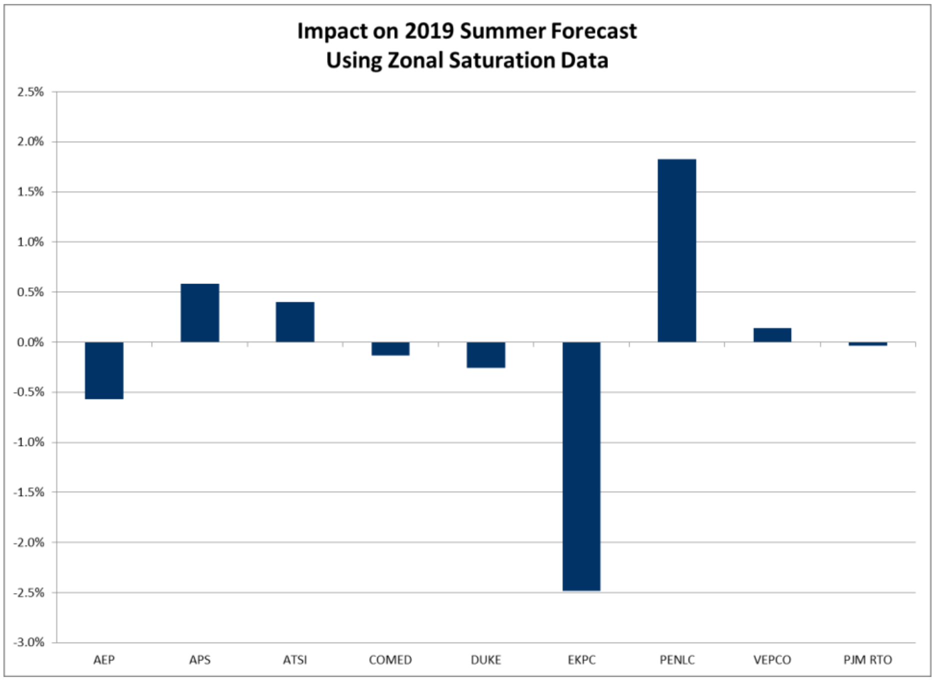
Some zones have two possible sets of equipment indexes. One set that incorporates zonal-supplied data, and one that does not. The inclination is to use zonal-supplied data, but it should not be used at the expense of accuracy. Accuracy was tested by computing three-year out forecast model error, and to then only use the equipment index that leverages zonal-supplied data if it results in improved accuracy. The results of this analysis can be seen in Figure D-1.

Figure D-1 : Accuracy Comparison – With and Without Zonal Data



Of the 10 zones that supplied data, only JCPL and METED produced forecast results that were less accurate. As a result, they were not used in the official forecast. Figure D-2 shows the results of using zonal saturation data. There is negligible impact at the RTO level. For six out of the eight zones affected, the change is in the realm of plus or minus 0.5%. The exceptions are EKPC and PENLC, the latter of which had the most noticeable improvement in accuracy due to inclusion of their zonal saturation data. In EKPC, the lower forecast is due to a higher proportion of relatively more efficient heat pumps than central air conditioners than is reflected in the East South Central Census Division data. In PENLC, the higher forecast is due to a sharper historic rise in cooling equipment saturation than in the Mid-Atlantic Census Division. Cooling equipment was comparatively sparse in PENLC in the late 1990s followed by a sharp rise in the 2000s.

Figure D-2 : Forecast Impact of Using Zonal Data



APPENDIX E. 2013 INVESTIGATION

In 2013 PJM conducted a reexamination of the load forecast model with the intent to identify and implement model changes that would address the persistent over-forecasting. Areas of investigation focused on the variable types then used in the model: day type, weather and economics, as well as revisions to model processes. Tests were conducted on each zone's non-coincident peak model and potential changes were scored based on their contribution to lowering the mean absolute percent error (MAPE) compared to the official 2013 model, with priority given to MAPE calculated on a daily basis as well as on the five highest and the single highest peaks of each summer. The results for each category studied were:

• Day type variables - Deleted weekday and month variables from the model and replaced with Day/month variables (e.g., MonJan, MonFeb, etc.). Result: insignificant improvement, at the cost of significantly more variables in the model;

• Weather variables - Replaced Temperature-Humidity Index with Summer Simmer Index, a related measure that results in a broader range than THI. Result: insignificant changes and in some cases a worse model fit;

• Weather variables - Replaced Temperature-Humidity Index (THI) with Summer Simmer Index, a related measure that results in a broader range than THI. Result: insignificant changes and in some cases worse model fit;

• Economic variables - Replaced the indexed variable with its six individual components: Gross Domestic Product, Gross Metropolitan Product, Population, Households, Non-Manufacturing Employment, and Real Personal Income. Result: improved fit in all zones but produced unexpected coefficient signs and wildly unstable forecast results;

• Economic variables - Replaced the indexed variable with its six individual components and added Manufacturing Employment. Result: improved fit in all zones but produced unexpected coefficient signs and wildly unstable forecast results;

• Economic variables - Replaced the indexed variable with a simpler one, using only Households, Non-Manufacturing Employment and Manufacturing Employment. Result: worse fit in all but a few zones;

• Weekday-Only - Removed weekend days and holidays from model. Result: no significant change;

• Price Term - Added a price shift dummy variable to capture the impact of significant historical changes in retail electricity prices, as recommended by Itron. Result: notable improvement in all cases where zones had significant price shifts;

• Recession Bands - Added a dummy variable for the recession bands as defined by the National Bureau of Economic Research. Result: no significant changes;

• Shortened the estimation period - PJM currently uses an estimation period that spans 1998 to the preceding August. Testing was done by incrementally trimming two years from the front end of the estimation period for the forecast model. Result: no significant improvements in fit until estimation period was shortened by at least eight years, at which point zonal forecasts became unstable – some much higher and others much lower than the baseline forecast.

PJM's results were reviewed with the Load Analysis Subcommittee and PJM management. None of the changes received sufficient support to be adopted and no changes to the model were made at that time.

APPENDIX F. 2014 INVESTIGATION

In a continued effort to investigate the cause of accuracy degradation, PJM began to seek other avenues for improvement. Analysis pointed to a breakdown in the relationship between load and economics, which was being driven in part by the acceleration of energy efficiency activities. PJM's goal was to incorporate these trends into an alternative energy model, and to pursue incorporation into the peak forecast model in the following year. Since the energy forecast is primarily used internally, it could be used as a starting point of testing concepts and laying the groundwork for future stakeholder discussion.

To begin understanding efficiency trends, PJM contracted with Itron to obtain Residential and Commercial equipment/appliance saturation and efficiency data[39]. Historical and forecast data were available at the Census Division level, and each zone was assigned to a Census Division. Three equipment indexes were then created: Heating, Cooling, and Other. Each index is constructed using a Residential and a Commercial component, that are then combined based on FERC Form 1 sectoral electricity sales. Within each sector component the indexes are a weighted average across equipment types of saturation (the share of households or businesses using an equipment type) normalized for efficiency.

Separately, usage indexes were created to indicate the need for each equipment category (Heating, Cooling or Other). Heating and Cooling usage are determined using Heating Degree Days (HDD) and Cooling Degree Days, respectively. Other usage is determined by using monthly weights from EIA data, which indicate the share of each appliance's annual usage in a given month.

Equipment indexes were then interacted with usage indexes to create four variables:

- ***XHeat***: Heating Equipment Index interacted with Heating Use;
- ***XCool***: Cooling Equipment Index interacted with Cooling Use;
- ***XCoolHum***: Cooling Equipment Index interacted with Cooling Use and Humidity;
- ***XOther***: Other Equipment Index interacted with Other Use.

The model was transitioned from a daily frequency to a monthly frequency. Due to this change, several variables that existed in the daily model were no longer appropriate and were removed. Variables removed included day of week variables, holiday variables, and various time-of-day weather variables. PJM adopted this new monthly method as its alternative energy forecast for the 2015 Load Forecast (Tables E-1a, E-2a, and E-3a), which reflected a near 2% reduction in the 5-year out forecast from the in-place model. PJM indicated to stakeholders that this was the energy forecast it would use for its own internal planning.

.

[39] These data are explained in detail in section IV - Equipment/Appliance Saturation and Efficiency