PJM Model Review and Recommendations

Eric Fox, Michael Russo, Dr. Frank A. Monforte, Dr. Stuart McMenamin

July 28, 2022

Agenda

» Transitioning to an Hourly Modeling Framework

» Review of the Current PJM Model

» Developing the Long-Term Model Drivers

» Modeling Hourly Loads

» Reshaping Demand – Technology Impacts

» Other Issues
  - Measuring accuracy
  - Weather normalization
  - Capturing energy efficiency
  - Reflecting temperature trends

The PJM Forecast

» 15-Year energy, summer, and winter peak forecasts for 22 Planning Zones

- Mapped into 6 Load Deliverability Areas (LDA)
- Daily energy, own-peak, and coincident peak forecasts
  - Models incorporate
    - Zonal weather conditions
    - Regional population and economic projections
    - Census Division end-use saturation and efficiency
    - Technology and Zone-specific load adjustments

- Forecast generated through zone-specific weather simulations

- The focus of this study is on the underlying Zone-level models.

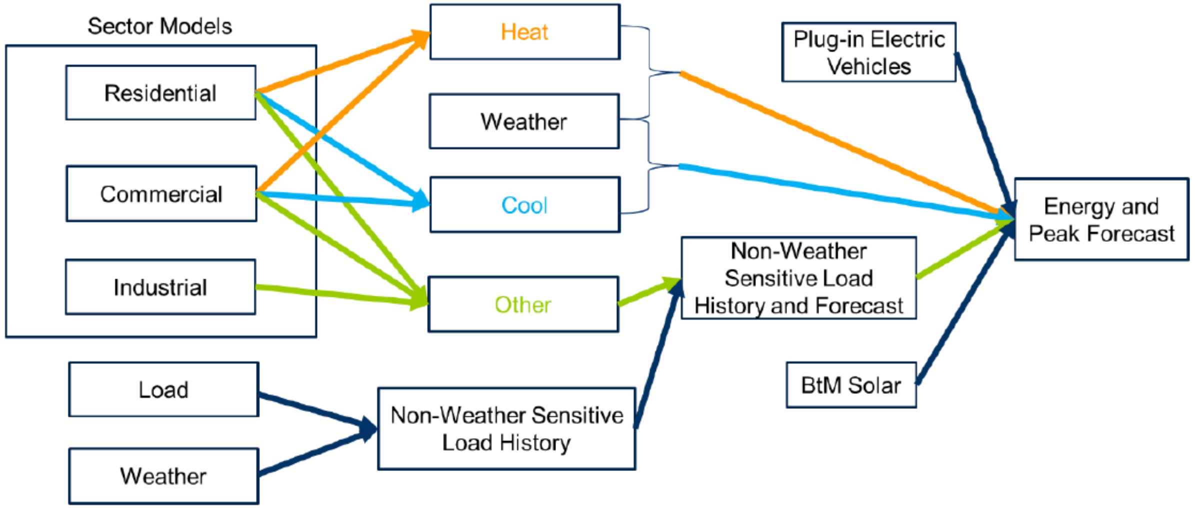
LDAs

PJM Mid-Atlantic
Eastern MAAC
Southern MAAC
Central MAAC
Western MAAC
PJM West

Zones

AE
AEP
APS
ATSI
BGE
COMED
DAYTON
DEOK
DLCO
DOM
DPL
EKPC
JCPL
METED
OVEC
PECO
PENLC
PEPCO
PL
PS
RECO
UGI

The Transition to Hourly Models



» Current daily models has all the right ingredients
- Population and economic drivers
- End-use saturation and efficiency
- Weather

» But it's complex – not a simple problem
- Largely a result of the PJM system size in terms of number of transmission zones, dispersed geography, weather diversity, differences in customer mix, and system planning requirements

» And it's not getting any easier
- Behind the meter solar adoption, expected impacts from electric vehicles, impacts of new electrification programs (such as cold-climate heat pump incentives), and battery storage will reshape loads impacting both the timing and level of zone non-coincident and coincident peak demands

- These load shaping technologies requires transitioning to an hourly load forecasting framework that also addresses the current complexities

The PJM Process – Continuous Model Improvements

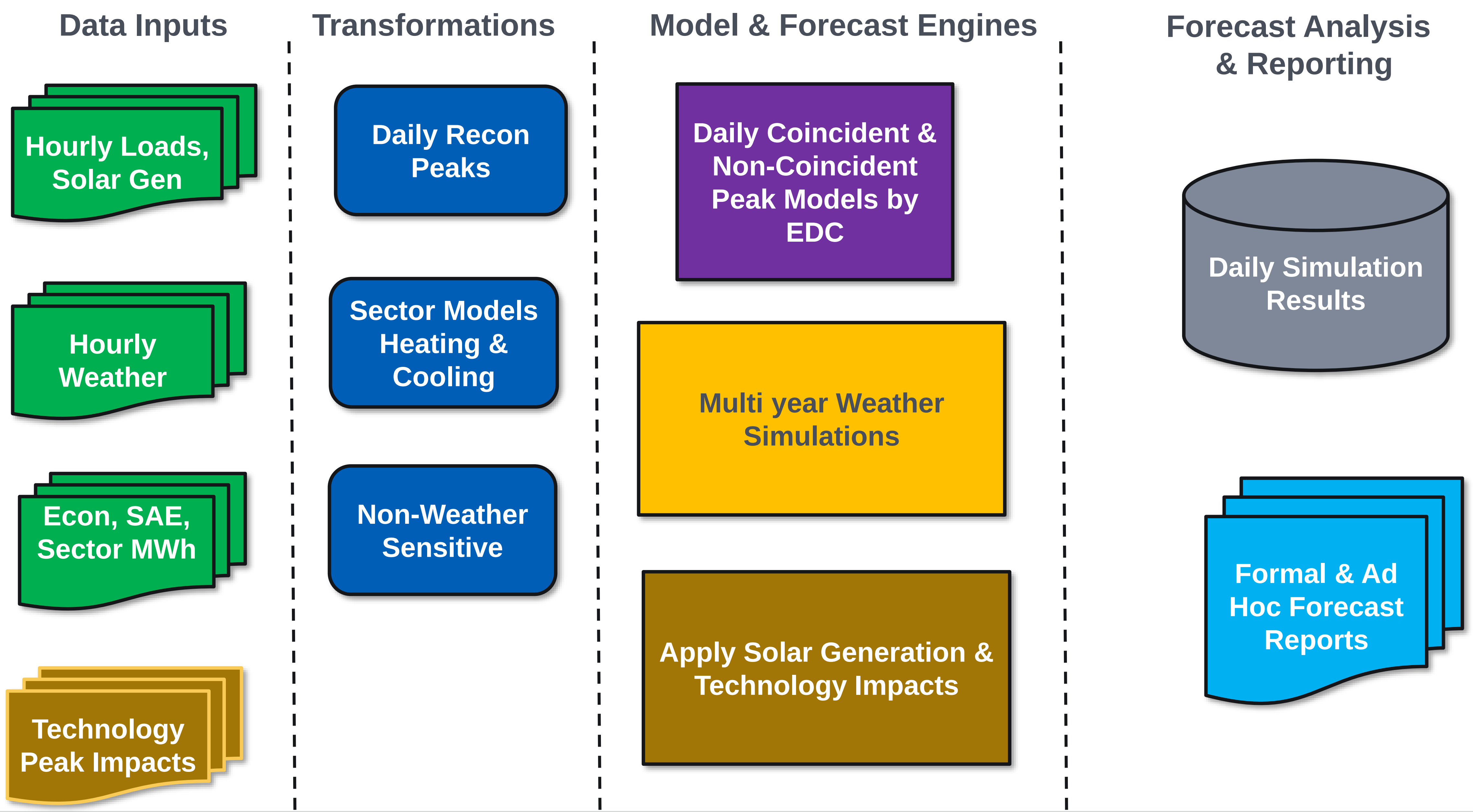
» Past Enhancements

- Developed weather simulation process to address system weather diversity
- Transitioned from GDP to richer economic drivers (blended economic concepts)
- Incorporated end-use saturation and efficiency drivers
- Accounted for relative differences in zonal customer mix – i.e., residential, commercial, and industrial

» Building on the Current Model (Next Phase)

- Transition to zonal hourly load models and forecast
- Develop long-term drivers from regional monthly (vs. annual) sales models
- Construct technology-adjusted hourly load forecasts

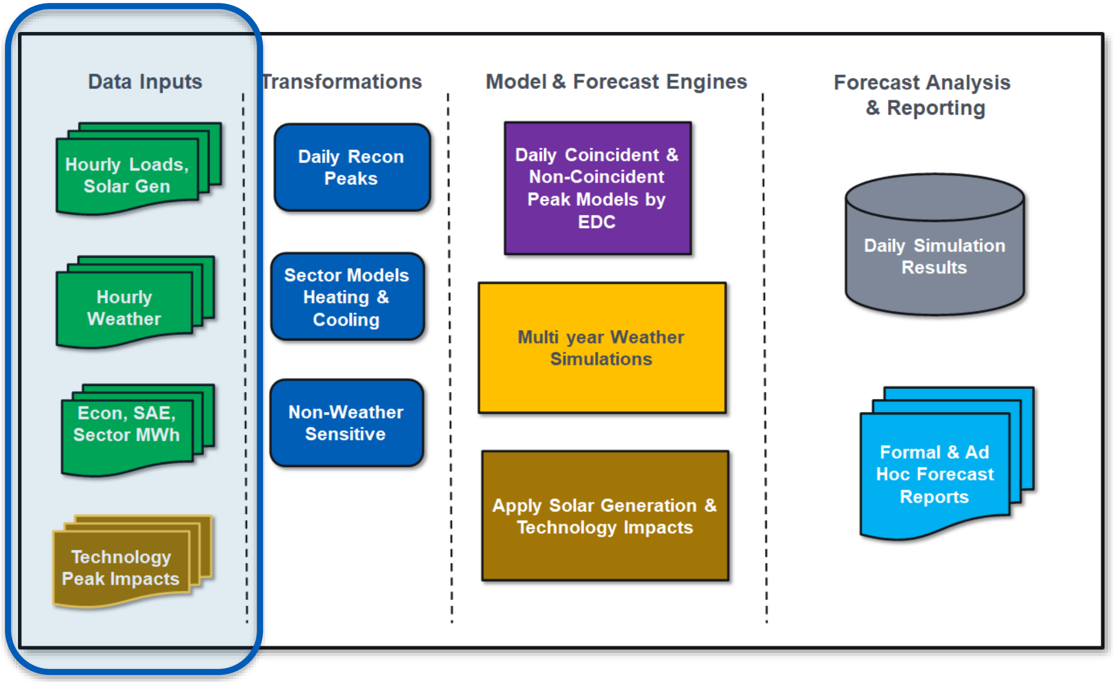
Model Review


Data Inputs
Transformations
Model & Forecast Engines
Forecast Analysis & Reporting
Hourly Loads, Solar Gen
Hourly Weather
Econ, SAE, Sector MWh
Technology Peak Impacts
Daily Recon Peaks
Sector Models Heating & Cooling
Non-Weather Sensitive
Daily Coincident & Non-Coincident Peak Models by EDC
Multi year Weather Simulations
Apply Solar Generation & Technology Impacts
Daily Simulation Results
Formal & Ad Hoc Forecast Reports

PJM Existing Long-Term Load Forecasting Framework (LTFS) Data Inputs Section

Data Inputs Section. Many of the inputs that feed the LFTS come in a variety of electronic formats. This section prepares the raw input data to be ingested by the Data Transformation Section of the LFTS. Data inputs include:

- Economic history and forecasts
- End-use saturation and efficiency trends (SAE)
- Sector (residential, commercial, and industrial) sales (MWh) by EDC
- Hourly Loads by EDC
- Weather data by weather station
- Solar PV, EV Charging, Battery Storage, Demand Response peak demand impacts



PJM Existing Long-Term Load Forecasting Framework
Data Transformations Section

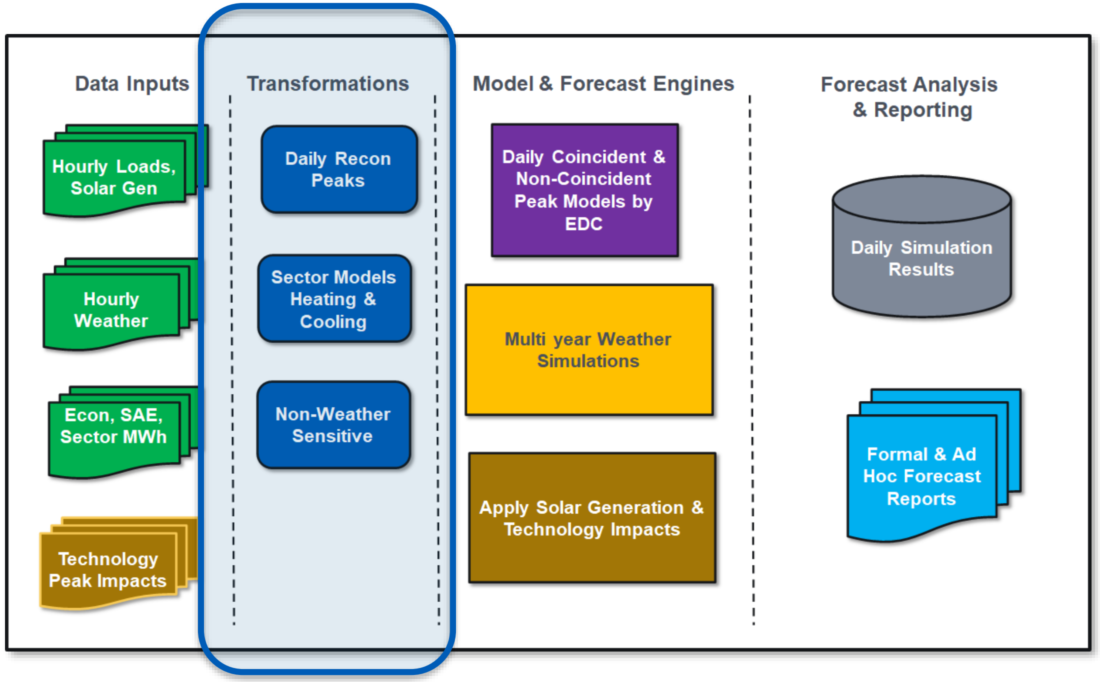
Data Transformation Section. This section converts the raw input into daily coincident and non-coincident model variables.

Reconstituted Coincident and Non-Coincident Peak Loads. To account for load loss due to embedded solar PV generation, historical estimates of aggregate embedded solar PV generation for each hour is added to the historical hourly loads by EDC. Historical coincident and non-coincident peak loads are then computed from the hourly reconstituted load data.

» Note: The coincident and non-coincident peak forecasts are adjusted downward in the forecast period by subtracting off forecasted embedded solar generation values for hour ending 17:00.

Sector Heating & Cooling Models. This series of three modules – Residential, Commercial, Industrial – are used to combine historical and forecasted economic data, and end-use appliance saturation and efficiency trends, with annual sector sales to construct Space Heating, Space Cooling, and Other Non-Weather sensitive load indices.

Non-Weather Sensitive Models. Hourly historical reconstituted load & weather data by EDC and weather station are input into the LFTS. The non-weather sensitive portion of the reconstituted load is extracted from these data and is used to scale the Other Non-Weather sensitive load index to daily MWh for both the historical and forecast period.



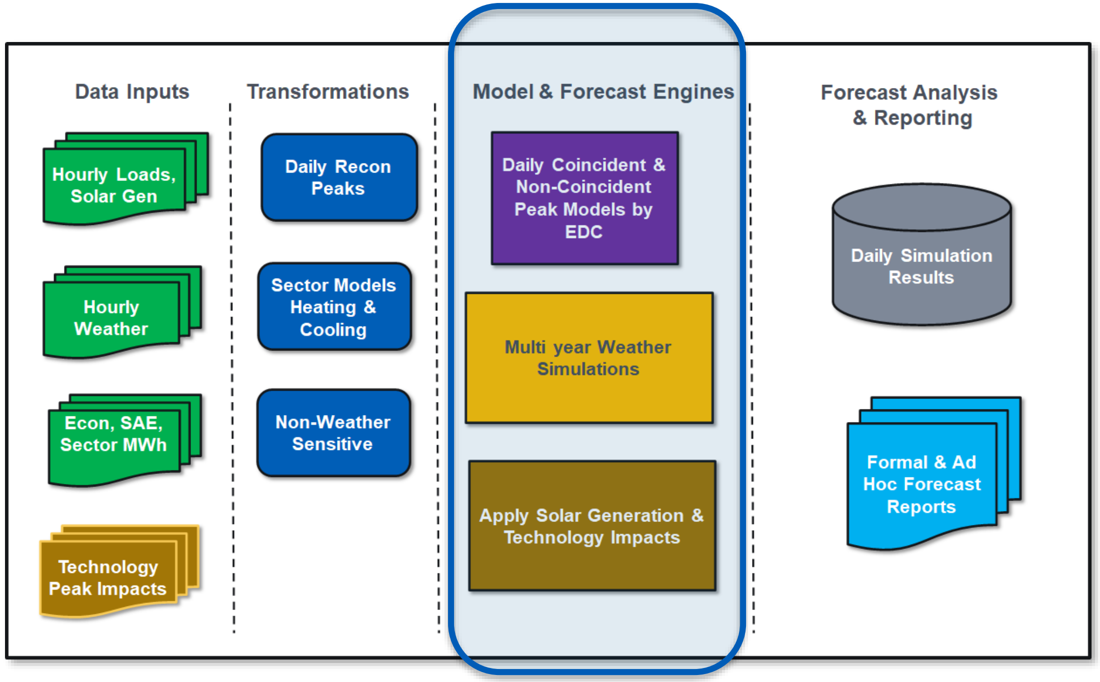
PJM Existing Long-Term Load Forecasting Framework
Model & Forecast Engines

Model and Forecast Engines Section. This section is used to specify and estimate the Daily Coincident and Non-Coincident Peak forecasts by EDC.

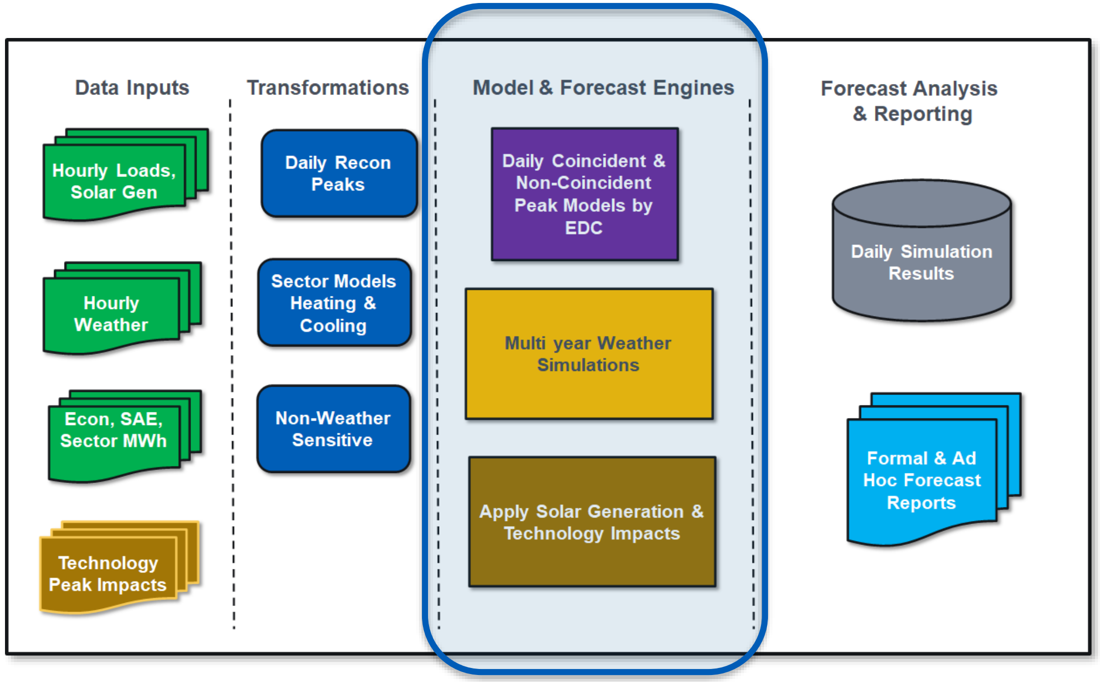
Daily Coincident and Non-Coincident Peak Models. The Space Heating and Space Cooling indices are combined with historical Heating Degree and Cooling Day variables to form MWh estimates of daily space heating and cooling. The daily values are included along with the estimated Other Non-Weather sensitive load data as explanatory variables in a Daily Coincident Peak and Daily Non-Coincident Peak models. A separate set of Coincident and Non-Coincident Peak models is estimated for each EDC using the EDC specific daily Heating MWh, Cooling MWh, and Other MWh values.

Weather Simulations. To account for the geographical and temporal diversity of the large PJM operating footprint weather simulations are used to generate a distribution of daily coincident and non-coincident peak day forecasts for each EDC. The weather simulations use historical hourly weather data by weather station to form forecasts of daily Heating Degree Days and Cooling Degree days by EDC.

Technology Impacts. At the end of each weather simulation embedded Solar PV, EV Charging, and other technology impacts are applied to the coincident and non-peak forecasts.



PJM Existing Long-Term Load Forecasting Framework
Model & Forecast Engines



Multi Year Weather Simulations. The approach is used to produce realistic load forecasts over the PJM footprint.

- Each year of historical weather data is used to construct 13 separate load forecasts for each year in the forecast horizon. With 27 years of historical weather data, each year in the forecast horizon will contain 351 forecasts with which Cumulative Frequency Distributions (CDFs) of coincident and non-coincident peak loads by season and month are constructed. The CDFs are used to select load forecast bands under design condition (e.g., median, 50%, 10% and 90% percentile values).

Actual and Consistent Weather Patterns Across EDCs. Actual weather data for all weather stations are rotated together ensuring consistent and realistic weather patterns are used to drive the load forecasts.

Consistency Across Weather Concepts. All weather concepts are rotated together ensuring consistent and realistic movements in temperatures, humidity, wind speed, wind direction, cloud cover, and solar irradiance. This is critical to align HVAC and solar PV generation interactions.

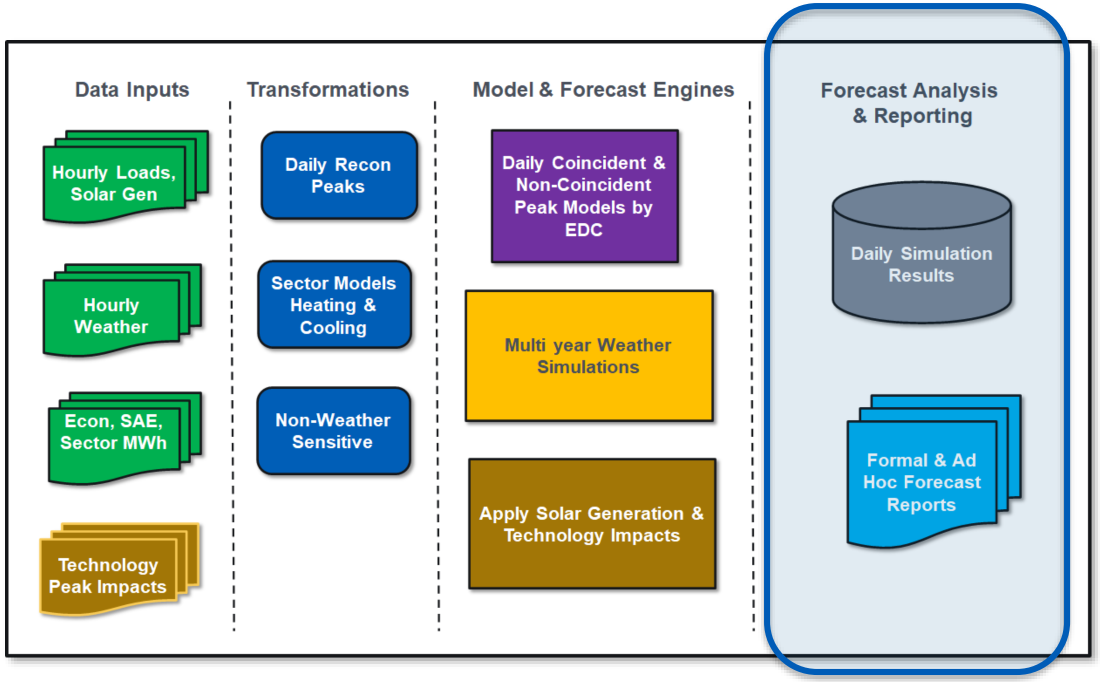
	Weather Scenarios												
		Rotate Forward						Rotate Backward					
Date	A1995	B1995	C1995	D1995	E1995	F1995	G1995	H1995	I1995	J1995	K1995	L1995	M1995
1-Jan	1/1/1995	1/2/1995	1/3/1995	1/4/1995	1/5/1995	1/6/1995	1/7/1995	12/31/1995	12/30/1995	12/29/1995	12/28/1995	12/27/1995	12/26/1995
2-Jan	1/2/1995	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

PJM Existing Long-Term Load Forecasting Framework
Forecast Analysis & Reporting

Forecast Summary Report Module. This section is used to create formal and ad hoc summaries of the long-term forecasts

Non-Coincident Peak Forecasts. Cumulative Frequency Distributions (CDF) of Monthly and Seasonal Non-Coincident Peaks by EDC and Year are constructed from the results of the Model & Forecast Engine section. From these CDFs Median/10%/90% peak values are extracted.

Coincident Peak Forecasts. Cumulative Frequency Distributions (CDF) of Monthly and Seasonal Coincident Peaks by EDC and Year are constructed from the results of the Model & Forecast Engine section. From these CDFs Median/10%/90% peak values are extracted.



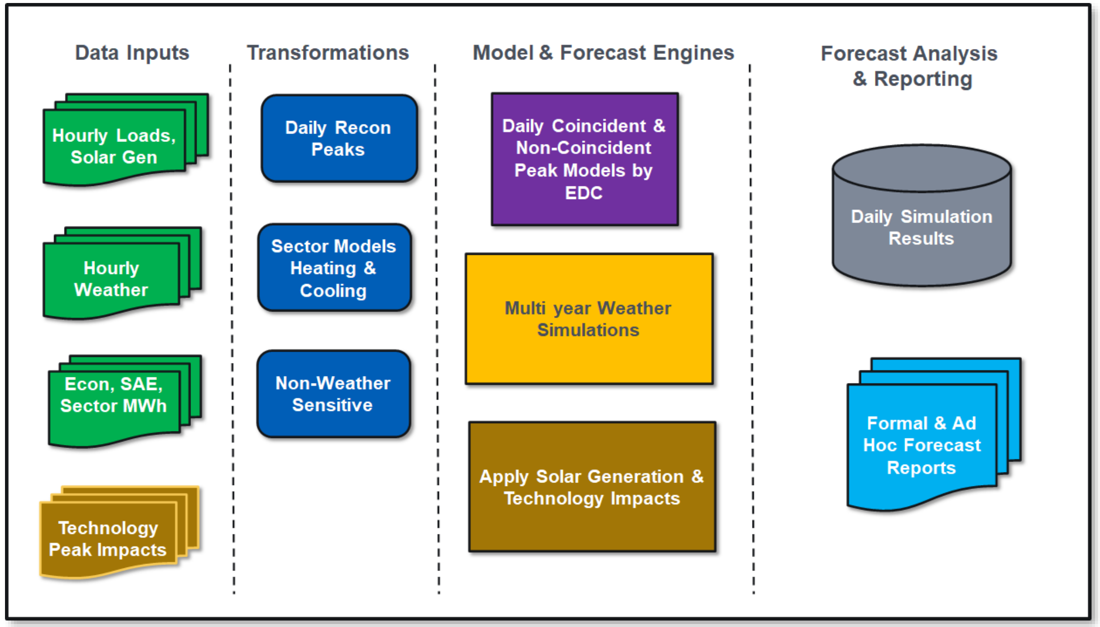
Long-Term Model Drivers

PJM Existing Long-Term Load Forecasting Framework
Market Participant Feedback

Limitations of the Data Used to Construct EDC Specific Sector Models. Currently, PJM utilizes Energy Information Administration (EIA) Form 861 annual EDC sales data by sector to construct the EDC specific space heating, space cooling, and non-weather sensitive models. Because of the limitations of relying on annual sales data, PJM currently fits the sector models to annual data starting in 1998. There is a concern among some Market Participants that the older annual data are not reflective of today's consumption activity implying the resulting space heating, space cooling, and other non-weather sensitive consumption trends lead to forecast errors.

PJM has evaluated shortening the sector models to the most recent ten years and have noted that due to a lack of data it is difficult to develop a stable and realistic set of model coefficients.

Recommendation. PJM should consider using the monthly zonal level sector sales data that is collected on EIA Form 861M. This would allow shortening the estimation range of the sector models to a rolling ten years while expanding the number of observations with the inclusion of monthly data. The monthly data may also help refine the space heating and cooling trends.



Capturing the Primary Forecast Drivers

» Long-term demand is driven by how we use electricity

- For cooling, heating, and other uses

- Residential – air conditioning, resistant heat, cooking, water heating, dryers, refrigeration, ….
- Commercial – lighting, ventilation, air conditioning, cooking, computers, refrigeration, ….
- Industrial – motors, manufacturing equipment, process heating, ….

» Need a set of model variables that capture the factors that drive cooling, heating, and other-use

- Current Approach: Construct **annual** heating, cooling, and base-use estimates and allocate to quarters

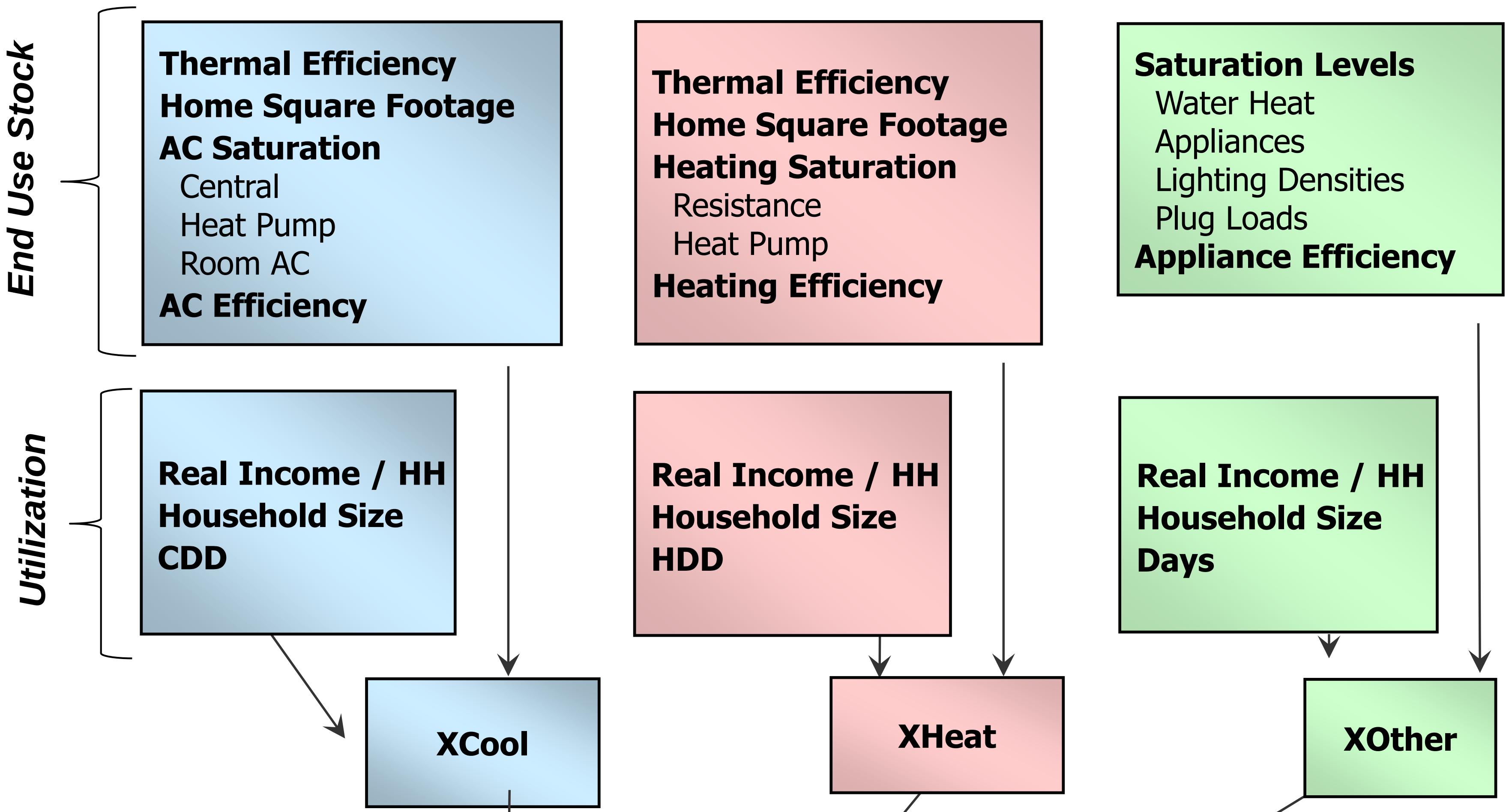
- Recommended Approach: Construct **monthly** heating, cooling, and other-use energy estimates
  - Derive daily heating, cooling, and base-use indices that are integrated with hourly zonal weather, seasonal, and day of the week variables
    - New NREL shapes data could be used to construct hourly other-use model drivers.

Capturing the Primary Forecast Drivers
The sector models

» Estimate monthly customer class models – Residential, Commercial, Industrial

1. Construct initial estimates of cooling, heating, and base-use by customer class
2. Calibrate to actual sales and customers through monthly regression models
3. Use estimated coefficients to isolate heating, cooling, and base use
4. Aggregate heating, cooling, and other-use sales across sectors
5. Calculate heating and cooling indices – interact with hourly weather variables
6. Calculate base-use loads/indices – interact with seasonal, day of the week, and holiday variables

Residential Average Use Model



$$AvgUse_m = a + b_c \times XCool_m + b_h \times XHeat_m + b_o \times XOther_m + e_m$$

Estimate monthly model with historical Use / Customer data

End-Use Variable – Cooling

Technologies
- Room AC
- Central AC
- Heat Pump
- Ground Source Heat Pump

$$\boxed{XCool_{y,m} = CoolEqpy \times CoolUse_{y,m}}$$

End-Use Stock

$$CoolEqp_y = Structural\ Index_y \times \sum_{Type} EI_{2015}^{Type} \times \frac{\left(Sat_y^{Type} \big/ Eff_y^{Type}\right)}{\left(Sat_{2015}^{Type} \big/ Eff_{2015}^{Type}\right)}$$

kWh per household

Utilization is a monthly index

$$CoolUse_{y,m} = \left(\frac{HHSize_{y,m}}{HHSize_{2015}}\right)^a \times \left(\frac{HHIncome_{y,m}}{HHIncome_{2015}}\right)^b \times \left(\frac{CDD_{y,m}}{CDD_{2015}}\right)$$

End-Use Variable – Heating

$$XHeat_{y,m} = HeatIndex_y \times HeatUse_{y,m}$$

Captures change in square footage and thermal shell efficiency

Technologies

- Resistant Heat
- Heat Pumps
- Room Heaters
- Ground Source Heat Pump

End-Use Stock

$$HeatEqp_y = Structural\ Index_y \times \sum_{Type} EI_{2015}^{Type} \times \frac{\left(Sat_y^{Type} / Eff_y^{Type} \right)}{\left(Sat_{2015}^{Type} / Eff_{2015}^{Type} \right)}$$

Utilization

$$HeatUse_{y,m} = \left(\frac{HHSize_{y,m}}{HHSize_{2015}} \right)^a \times \left(\frac{HHIncome_{y,m}}{HHIncome_{2015}} \right)^b \times \left(\frac{HDD_{y,m}}{HDD_{2015}} \right)$$

XOther Variable

$$XOther_{y,m} = OtherEqpIndex_{y,m} \times OtherUse_{y,m}$$

End-Use Stock

Technologies
- Refrigeration
- Cooking
- Dryers and Washers
- Water Heating
- Lighting, ….

$$OtherEqpIndex_{y,m} = \sum_{Type} EI_{2015}^{Type} \times \frac{\left(Sat_y^{Type} \Big/ Eff_y^{Type} \right)}{\left(Sat_{2015}^{Type} \Big/ Eff_{2015}^{Type} \right)} \times MFrac_m^{Type}$$

kWh per household

Monthly Allocation Factor (base on end-use shapes)

Utilization

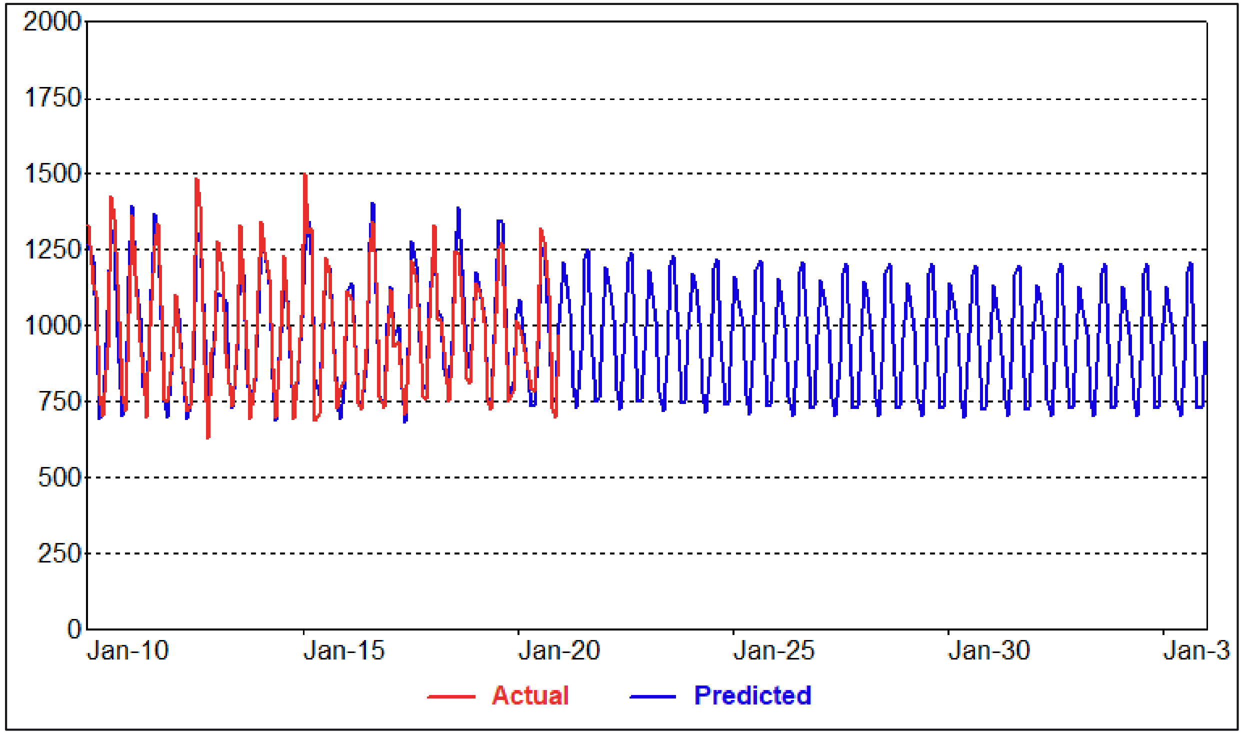
$$OtherUse_{y,m} = \left(\frac{HHSize_{y,m}}{HHSize_{2015}} \right)^a \times \left(\frac{HHIncome_{y,m}}{HHIncome_{2015}} \right)^b \times Days_{y,m}$$

Estimate Model Coefficients

Variable	Coefficient	StdErr	T-Stat	P-Value
mStructResVar.XHeat	0.611	0.041	14.962	0.00%
mStructResVar.lag_XHeat	0.385	0.040	9.596	0.00%
mStructResVar.XCool	0.579	0.029	20.275	0.00%
mStructResVar.lag_XCool	0.271	0.028	9.580	0.00%
mStructResVar.XOther	0.694	0.016	42.501	0.00%

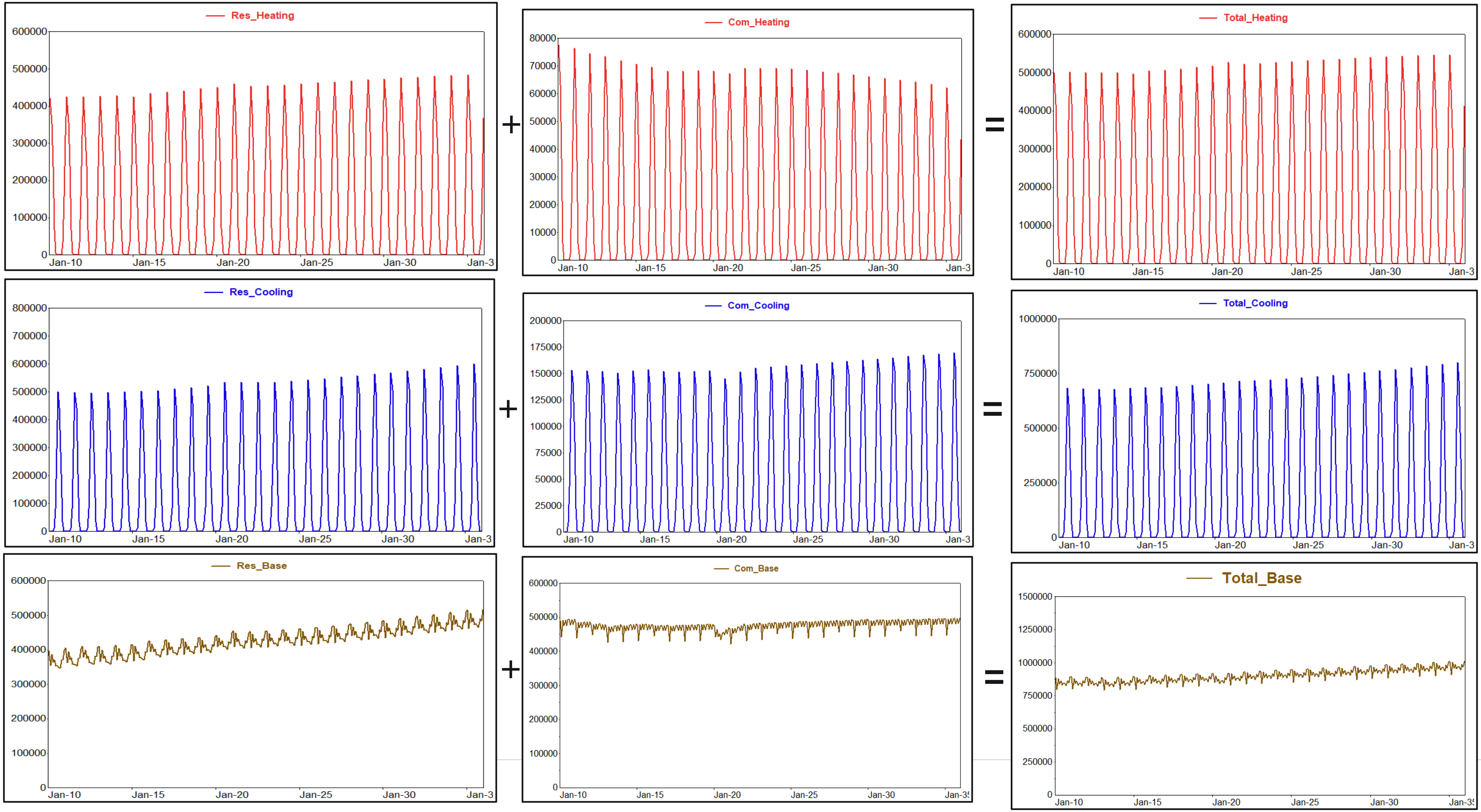
Model Statistics	
Iterations	1
Adjusted Observations	132
Deg. of Freedom for Error	127
R-Squared	0.899
Adjusted R-Squared	0.896
AIC	8.507
BIC	8.616
F-Statistic	#NA
Prob (F-Statistic)	#NA
Log-Likelihood	-743.74
Model Sum of Squares	5,411,707.85
Sum of Squared Errors	605,458.10
Mean Squared Error	4,767.39
Std. Error of Regression	69.05
Mean Abs. Dev. (MAD)	50.32
Mean Abs. % Err. (MAPE)	5.02%
Durbin-Watson Statistic	1.739
Durbin-H Statistic	#NA
Ljung-Box Statistic	21.44
Prob (Ljung-Box)	0.6124

Estimation Period: January 2010 to December 2020 (132 observations)

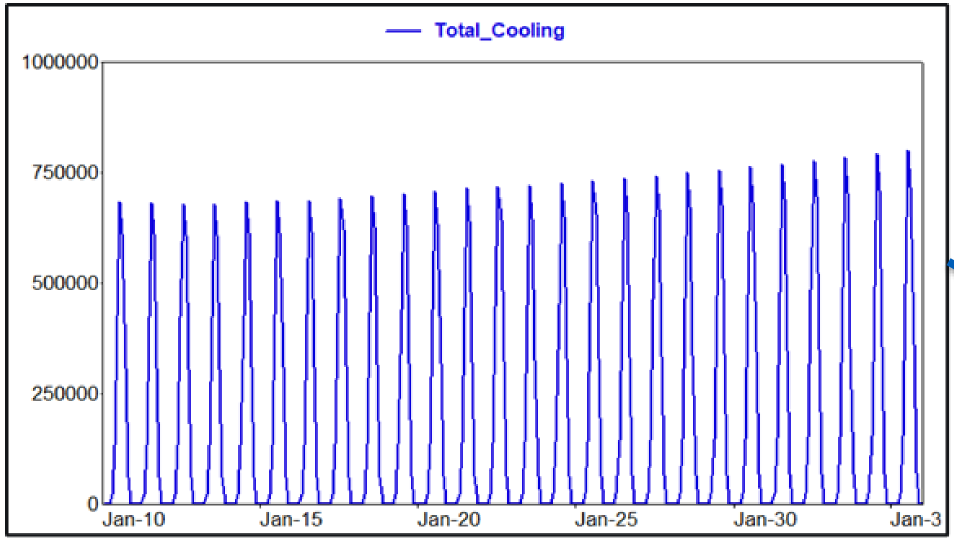


Estimate similar models for the commercial sector and more generalized econometric models for the industrial sector.

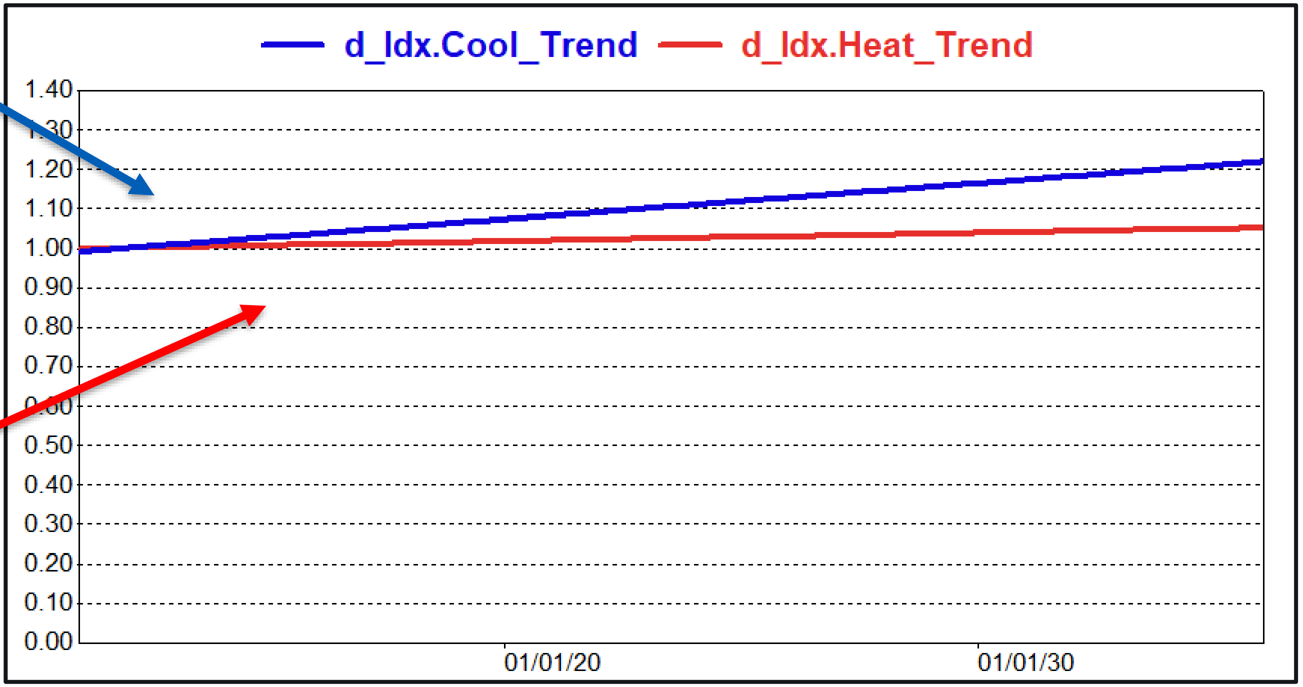
Use Coefficients to Isolate Heating, Cooling, and Base Use Requirements

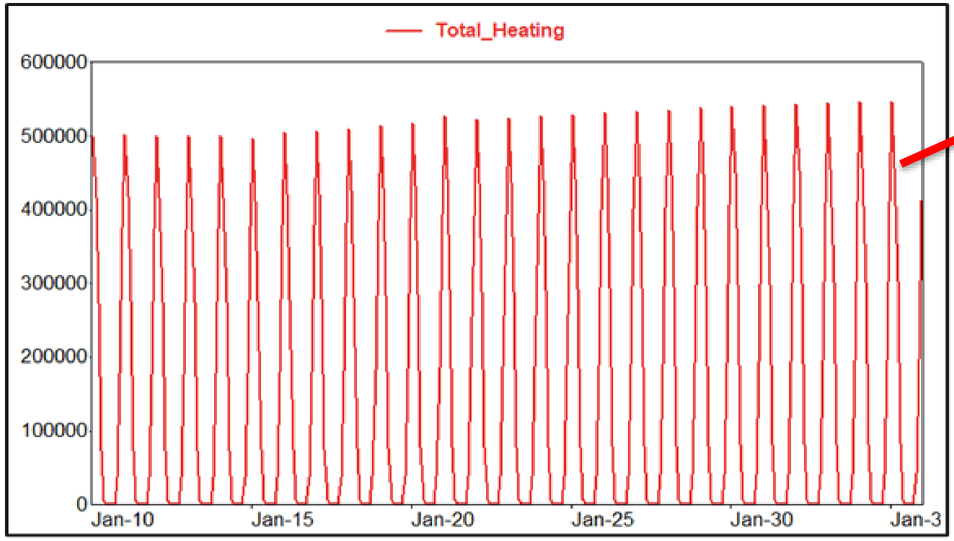


Inputs into Hourly Zonal Models



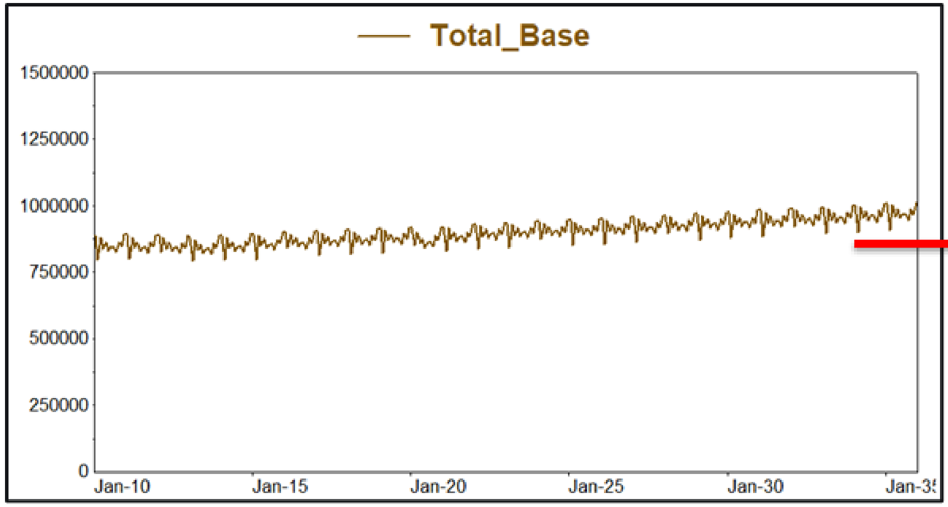
Cooling index interacts with hourly cooling weather variables

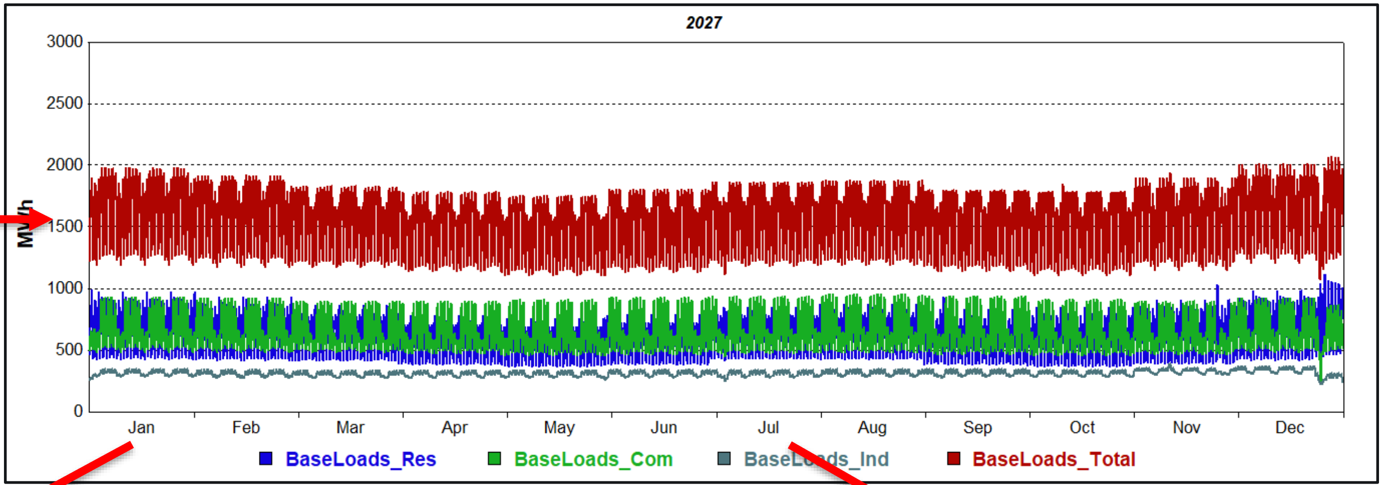


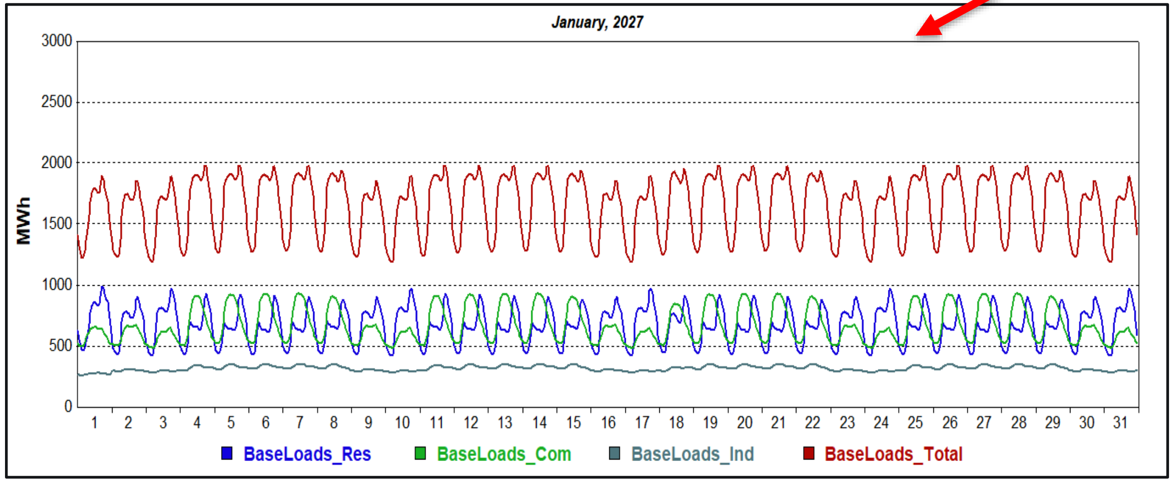


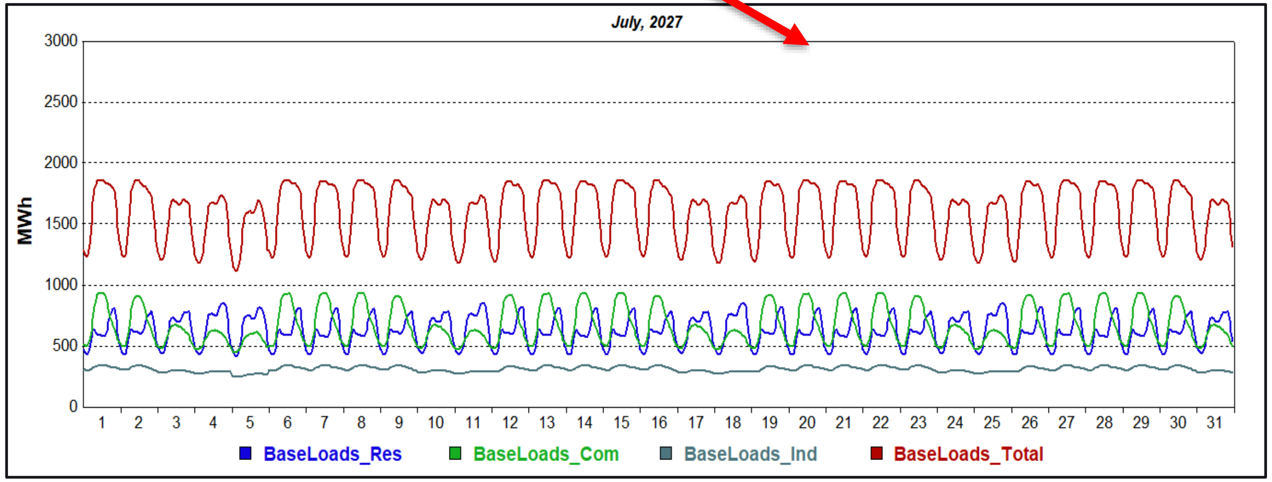
Heating interacts with hourly heating weather variables

Allocate Base Loads to Hours, Input into Hourly Zonal Models









Source: NREL ResStock and ComStock, non-weather sensitive end-use load profiles

Data Sources

- Monthly reported EIA revenue class sales and customers (residential, commercial, industrial, other)
  - Alternatively, can use utility historical billed sales and customers

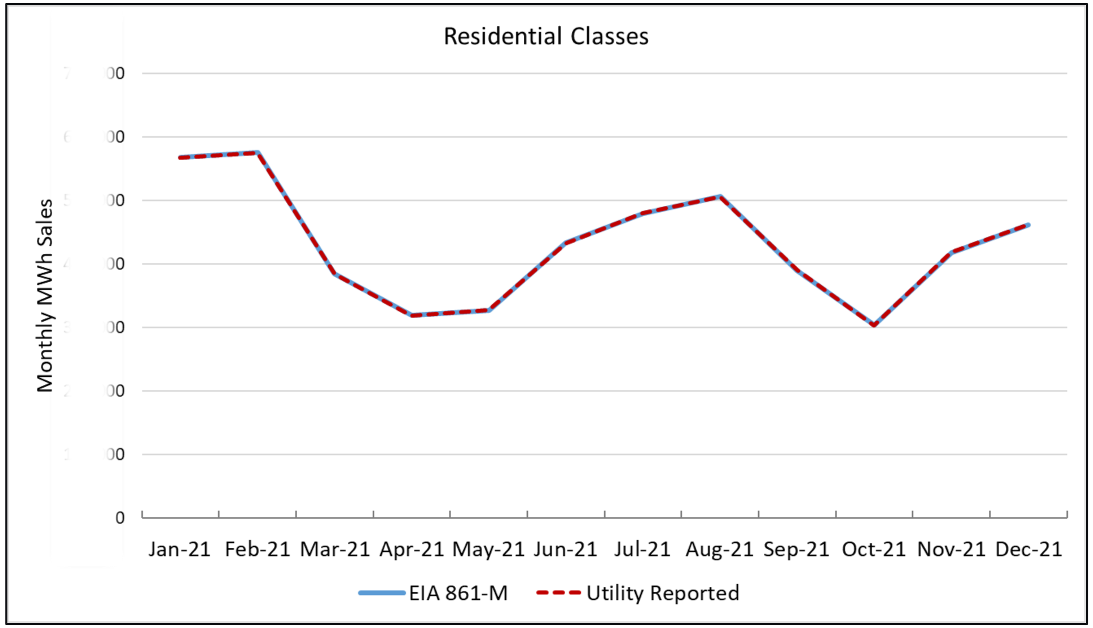
- Moody Analytics demographic and economic forecasts

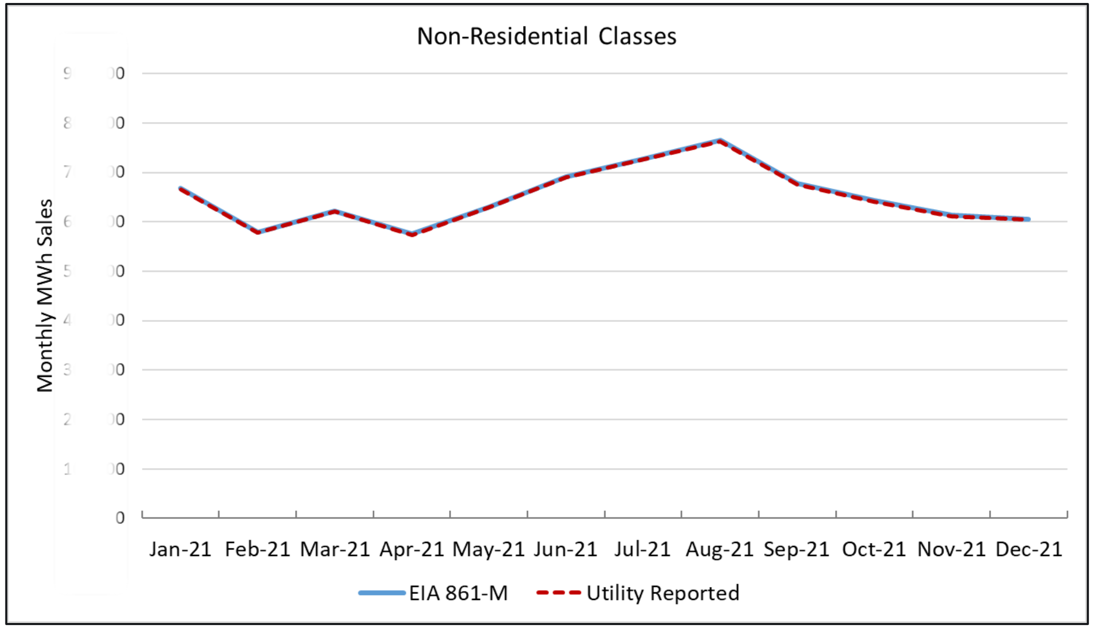
- EIA residential and commercial end-use saturation, stock efficiency, consumption data
  - Calibrate saturations to zones with NREL ResStock and ComStock data
    - Or utility saturation surveys where available
  - Calculate end-use monthly allocation factor from NREL end-use hourly load profiles

- NREL residential and commercial end-use hourly load profiles
  - Or utility AMI data – disaggregated into heating, cooling, and other use (if available)

EIA Data Compared to Utility Data

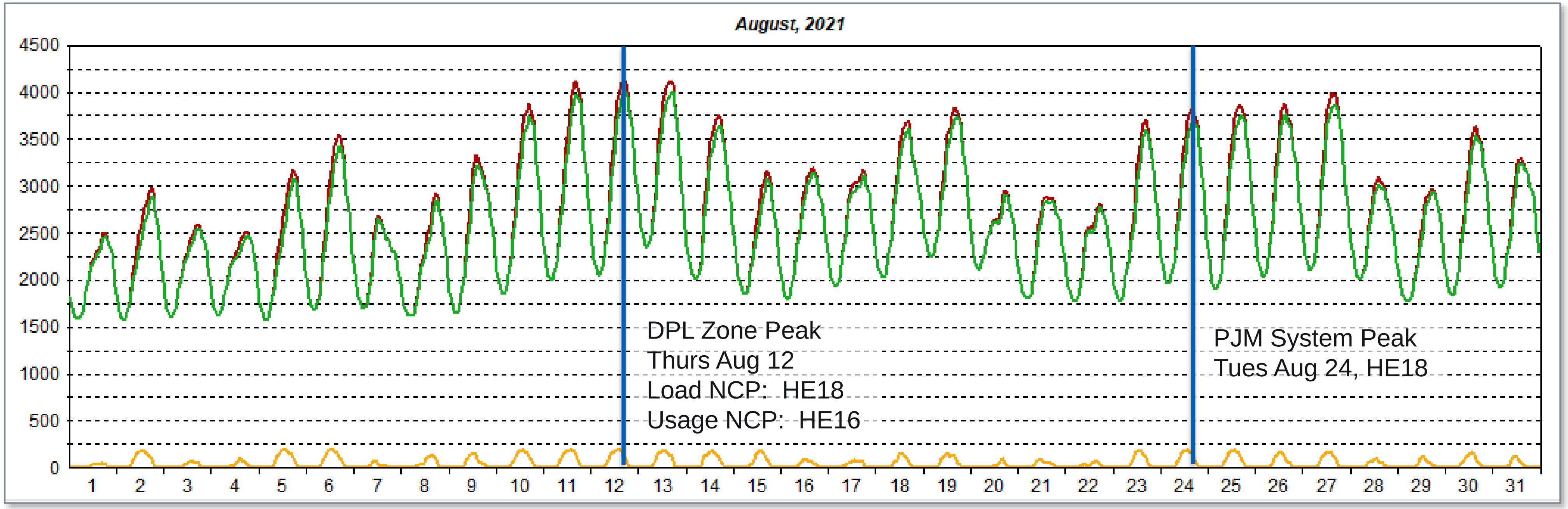
- EIA form 861-M reports historical monthly, class-level, sales and customer counts by utility
- Comparison of 861-M data to actual utility reports sales confirms reasonableness of the 861-M data





Modeling Hourly Loads – Weather Models

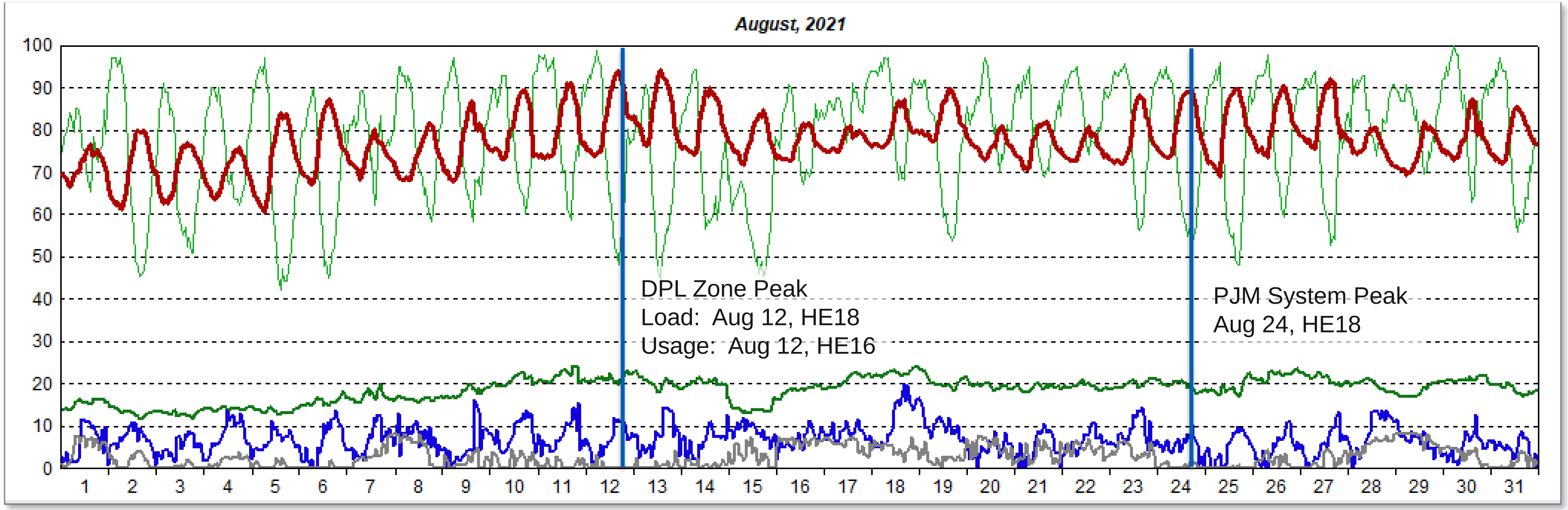
Understanding the Hourly Load Data



- » Load (green) is the net load met by DPL zone resources
- » Solar_DR (orange) is estimated BTM generation + DR impacts
- » Reconstituted load (red) is estimated end-use consumption (usage)
  - • Hourly Usage = Hourly Load + Hourly Solar&DR
- » Zone NCP is the maximum DPL net load
- » Zone CP is the DPL net load at time of PJM peak

	DPL NCP 8/12 HE18	DPL CP 8/24 HE18	Coincidence Factor
Usage	4,100.3	3,753.6	91.5%
Solar_DR	94.3	83.6	
DPL Load	4,006.0	3,670.0	91.6%

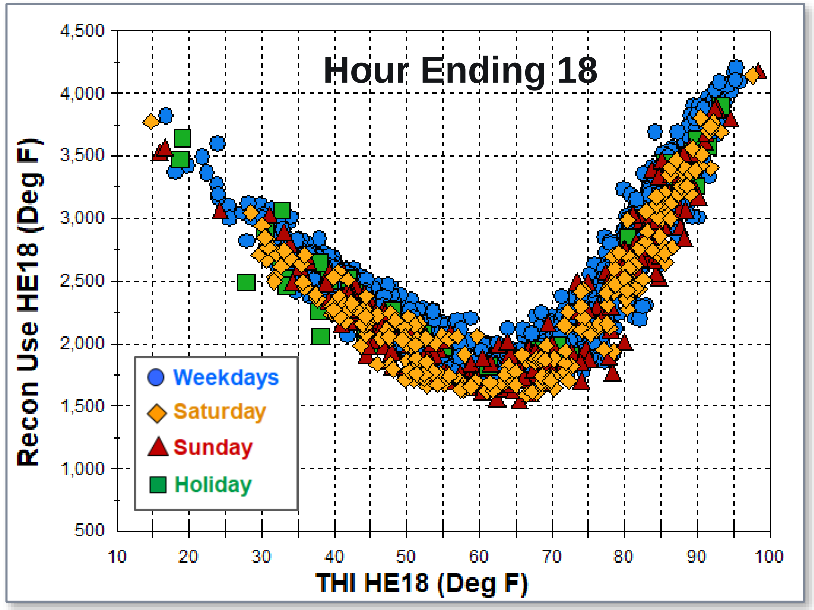
Understanding the Hourly Weather Data



- » Hourly Dry Bulb Temperature in Degrees F (Red)
- » Relative Humidity in Percent (Light Green)
- » Moisture content of air in mmHg (Dark Green)
- » Wind Speed in MPH (Blue)
- » Cloud Cover in octas (0 to 8) (Gray)

	DPL 2021 NCP 8/12 HE18	PJM 2021 Peak 8/24 HE18
AvgDB Temperature (Deg F)	92.7	88.8
Relative Humidity (%)	54.3	53.3
Moisture Content (mmHg)	21.3	19.1
Temp Hum Index (Deg F)	95.4	91.1
Wind Speed (MPH)	11.4	8.4
Cloud Cover (Octas)	0.12	1.6

Proposed Hourly Zone Load Models



» Hourly Load Model
  - Estimate 24 models for each zone
  - Use reconstituted load data as the Y variable
  - Modify weather variable construction
    - Rolling hourly weather variables
  - Integrate class level SAE end-use drivers
  - Maintain weather simulation process
    - Generate 8,760 for each weather rotation and forecast year
  - Post process weather and solar simulation results to get zone NCP values
  - Post process weather and solar simulation results to get RTO and LDA CP values

» Hourly models are a good fit for the weather simulation approach
  - Hourly models can accurately predict loads with alternative hourly weather patterns
  - Impacts of geographic weather diversity can be quantified
  - Avoids the need for a “normal” hourly weather scenario with “normal” weather diversity

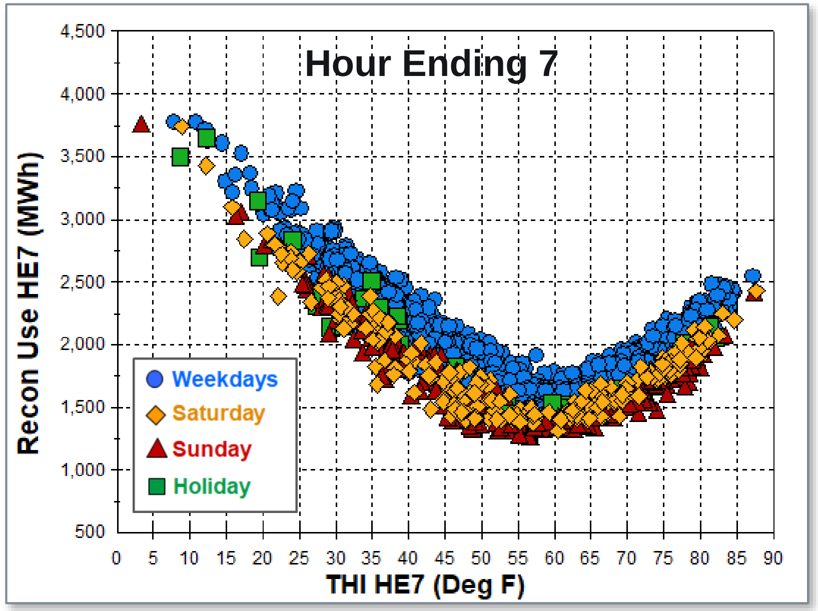
Weather Variable Specification

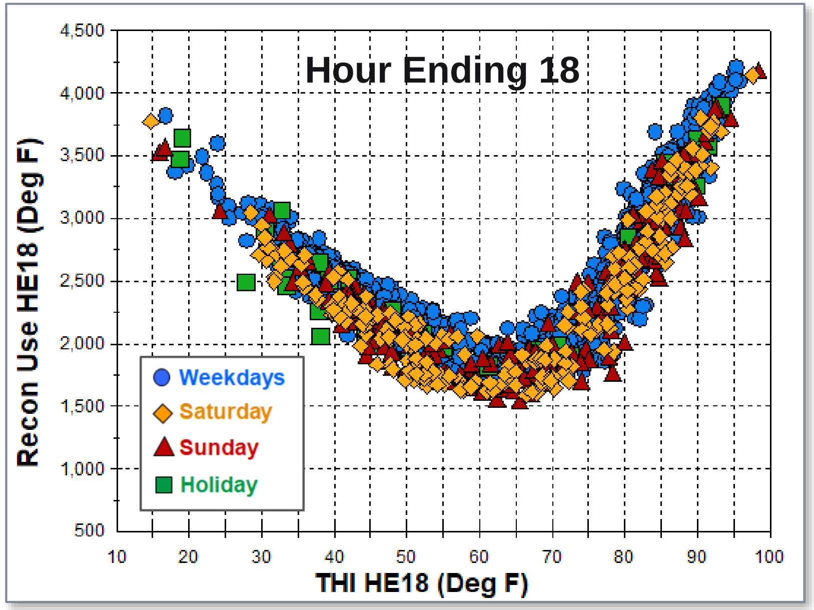
» Alternative ways to use hourly data
- Daily average, Time-of-day average, Rolling
- Recommend rolling hourly approach with two-part TD and HD variables, lags and interactions

» Humidity is important when it is hot
- Currently used to construct temperature-humidity index (THI)
- Recommend alternative THI formula

» Wind has impacts on hot and cold side
- Currently using wind adjusted temperature on cold side only
- Recommend alternative interacting wind this HD1 and TD1

» Clouds have hot and cold side impacts
- Currently not using clouds
- Recommend introducing could variables that interact with HD1 and TD1





Four Approaches to Using Hourly Weather Data

» Y variable is hourly DPL load

1. Daily Average Weather
   - Daily average temperature, humidity, wind, daytime clouds
   - Compute daily average THI, then TD1, TD2, HD1, HD2

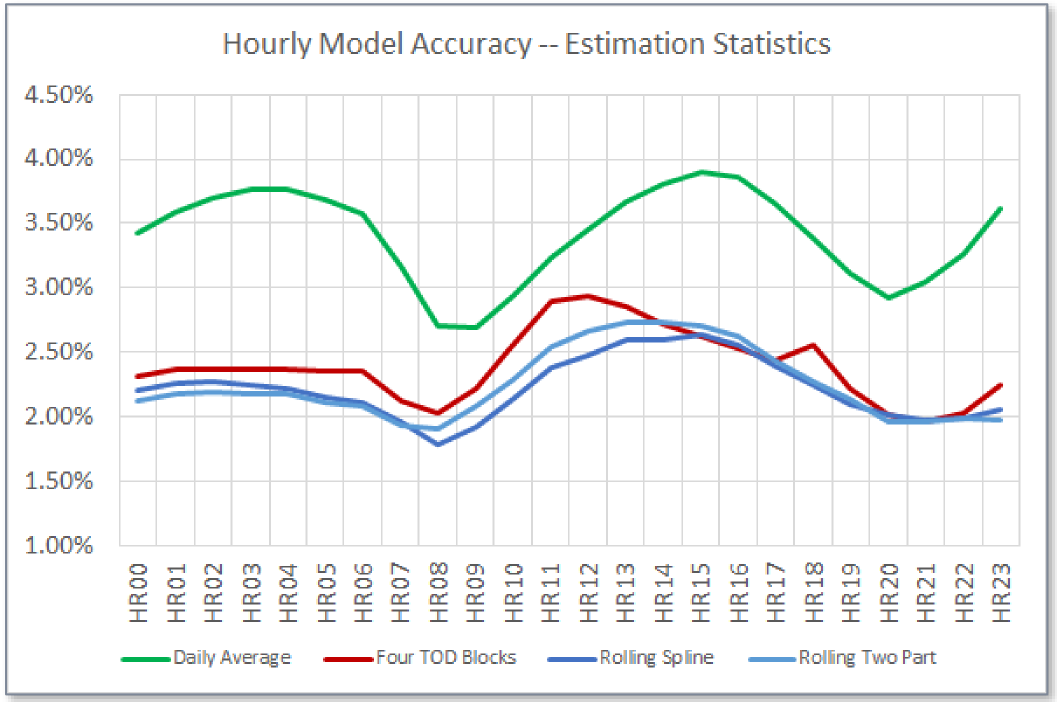
2. Time-of-Day Weather Blocks
   - Four blocks – Night, Morning, Afternoon, Evening
     - Compute AvgDB, Hum, Wind, Clouds for each block
   - TD1, TD2, HD1, HD2 with block specific base values

3. Rolling Hourly Weather – Complex and Simple Models
   - Weather variables for each hour (3 hour centered moving averages)
     - AvgDB, Hum, THI, Wind, Clouds
   - 3a. Complex – HD Spline, TD Spline using 5-degree buckets, weights from neural networks
   - 3b. Simple – TD1, TD2, HD1, HD2 variables

» Common daily variables
- Month binary variables
- Within month trends
- Holiday variables
- Covid Phase 1 to Phase 4 variables
- Long weather lags
  - 10-day HD, TD lags
  - 28-day HD, TD lags

Estimation Results

- » Rolling hourly variables work the best
- » Complex spline and simple two-part degree day methods are similar
- » Recommend rolling two-part approach
  - Two-part method is simpler than detailed spline
  - Accuracy is about the same (2.25% vs 2.22% MAPE)



Hour	DailyAvg Wthr	Four TOD Blocks	Rolling Spline	Rolling Two Part
HR00	3.42%	2.31%	2.20%	2.12%
HR01	3.59%	2.36%	2.26%	2.17%
HR02	3.69%	2.36%	2.27%	2.19%
HR03	3.77%	2.37%	2.24%	2.18%
HR04	3.76%	2.37%	2.22%	2.17%
HR05	3.68%	2.35%	2.15%	2.11%
HR06	3.57%	2.35%	2.11%	2.08%
HR07	3.16%	2.12%	1.96%	1.93%
HR08	2.71%	2.03%	1.78%	1.91%
HR09	2.69%	2.22%	1.92%	2.08%
HR10	2.94%	2.55%	2.14%	2.28%
HR11	3.24%	2.89%	2.38%	2.54%
HR12	3.45%	2.94%	2.48%	2.66%
HR13	3.67%	2.86%	2.59%	2.73%
HR14	3.81%	2.72%	2.59%	2.73%
HR15	3.90%	2.62%	2.64%	2.71%
HR16	3.86%	2.53%	2.56%	2.62%
HR17	3.65%	2.43%	2.39%	2.44%
HR18	3.38%	2.56%	2.25%	2.27%
HR19	3.11%	2.21%	2.10%	2.14%
HR20	2.92%	2.01%	2.01%	1.96%
HR21	3.04%	1.96%	1.97%	1.96%
HR22	3.26%	2.02%	1.99%	1.99%
HR23	3.62%	2.24%	2.05%	1.97%
Avg	3.41%	2.39%	2.22%	2.25%
#Coef	78	75	79-81	82

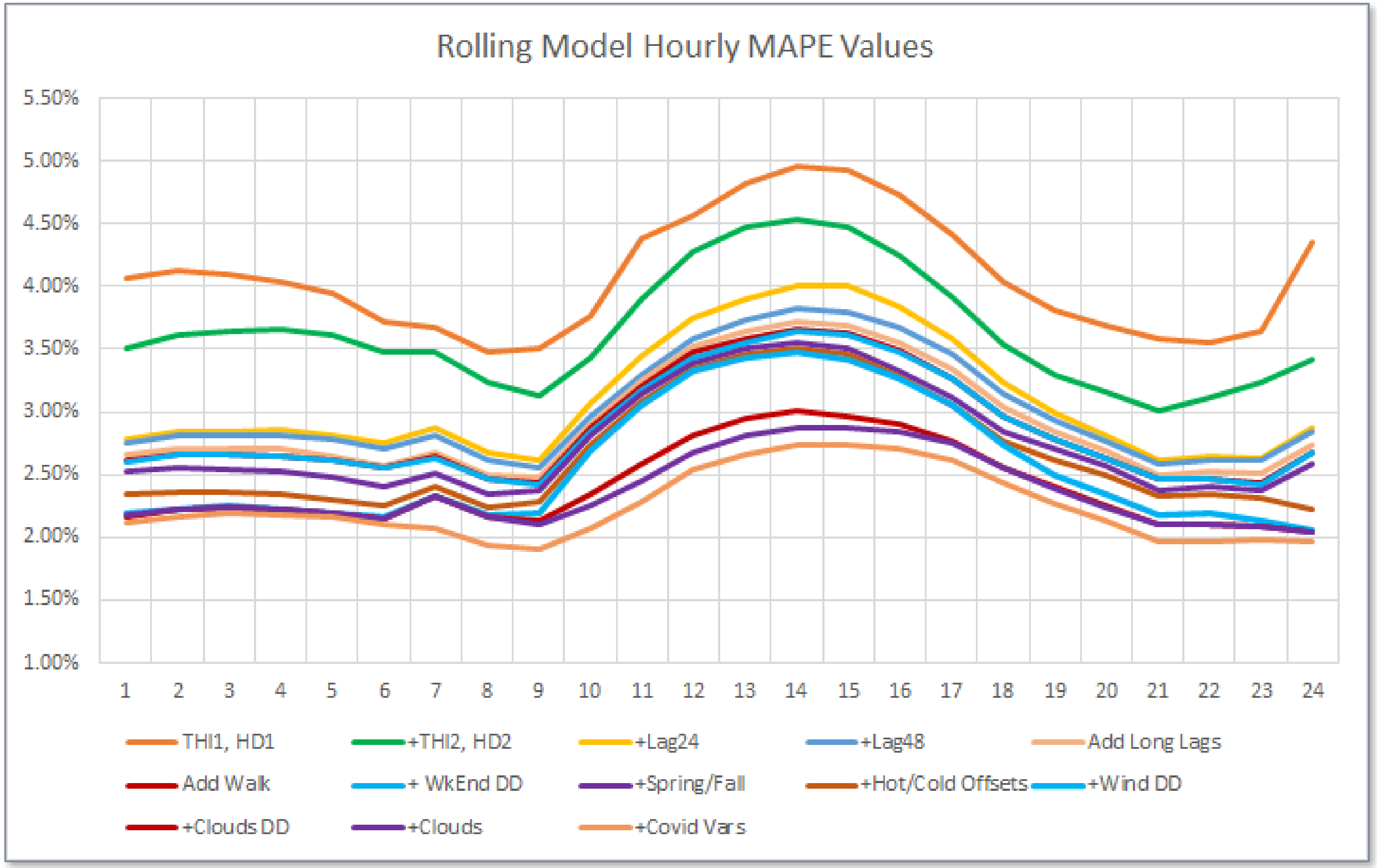
Two-part TD & HD base temperatures

Hour	TD1	TD2	HD1	HD2
HE01	65	75	55	45
HE02	65	75	55	45
HE03	65	75	55	45
HE04	65	75	55	45
HE05	65	75	55	45
HE06	65	75	55	45
HE07	65	75	55	45
HE08	65	75	55	45
HE09	65	75	55	45
HE10	65	75	55	45
HE11	70	80	60	50
HE12	70	80	60	50
HE13	70	80	60	50
HE14	70	80	60	50
HE15	70	80	60	50
HE16	70	80	60	50
HE17	70	80	60	50
HE18	70	80	60	50
HE19	70	80	60	50
HE20	65	75	55	45
HE21	65	75	55	45
HE22	65	75	55	45
HE23	65	75	55	45
HE24	65	75	55	45

Two Part Model Cascade – Estimation MAPE Values

- Using rolling hourly approach
- Estimation 2017 to 2021
- Add groups of variables to the model sequentially
- See the accuracy impact of each variable group

Rolling Model Components	Average MAPE
Month, DayType, Hols	10.31%
Add THI1, HD1	4.08%
Add THI2, HD2	3.64%
Add Lag24 DD	3.10%
Add Lag48 DD	3.02%
Add Long Lags	2.92%
Add Month Trends (Walk)	2.87%
Add WkEnd Shift	2.86%
Add Spring/Fall Shifts	2.77%
Add Hot/Cold Lag Offsets	2.66%
Add Wind * HD,CD	2.58%
Add Clouds * HD,CD	2.41%
Add Clouds	2.37%
Add Covid Vars	2.25%
Add AR1	2.04%

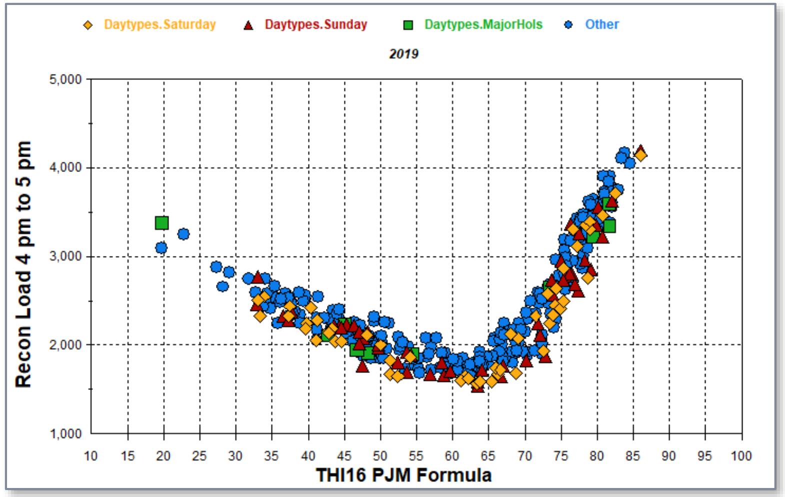


Recommendation for THI Model Variable

» Remaining analysis uses different data
- Switched to reconstituted load
- Estimation 2017 to August 2021

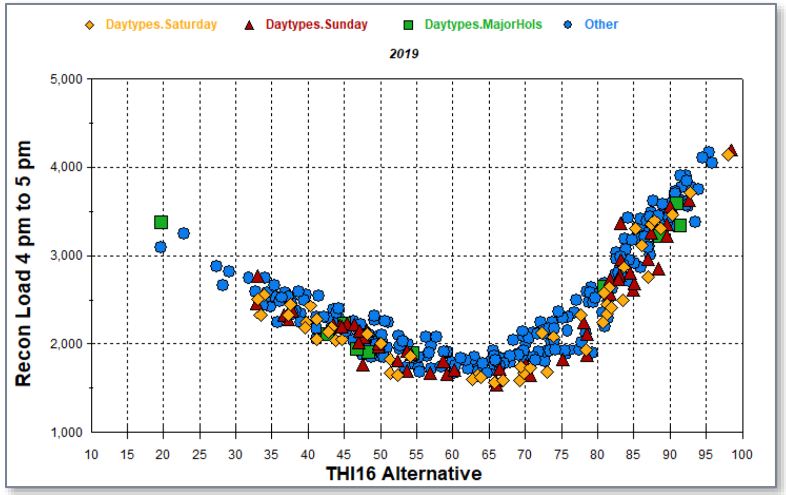
» Current THI variable
- THI = T - .55 (1-H/100) * CD58
- Compresses temperatures above 58
- 90 and humid "feels" cooler than 90

	PJM Formula		
	T - .55 * (1-Hum/100) * CD58		
	Case 1	Case 2	Case 3
T (°F)	90.0	90.0	90.0
Hum (%)	40.0	65.0	90.0
THI	79.4	83.8	88.2



» Recommended THI variable
- THI = T + .55 (H/100 - .4) * CD58
- Expands temperatures above 58
- 90 and humid "feels" hotter than 90

	Recommended Formula		
	T + .55 * (Hum/100 - .4) * CD58		
	Case 1	Case 2	Case 3
T (°F)	90.0	90.0	90.0
Hum (%)	40.0	65.0	90.0
THI	90.0	94.4	98.8



» Not a critical recommendation
- Models will work about the same with either formula
- But new formula is easier to work with using 5 and 10 degree buckets

» Either way, we recommend research to optimize the THI parameters (.55 and 58)

Recommendation for Wind Variables

» Current Model uses Wind Adjusted Temperature (WWP)

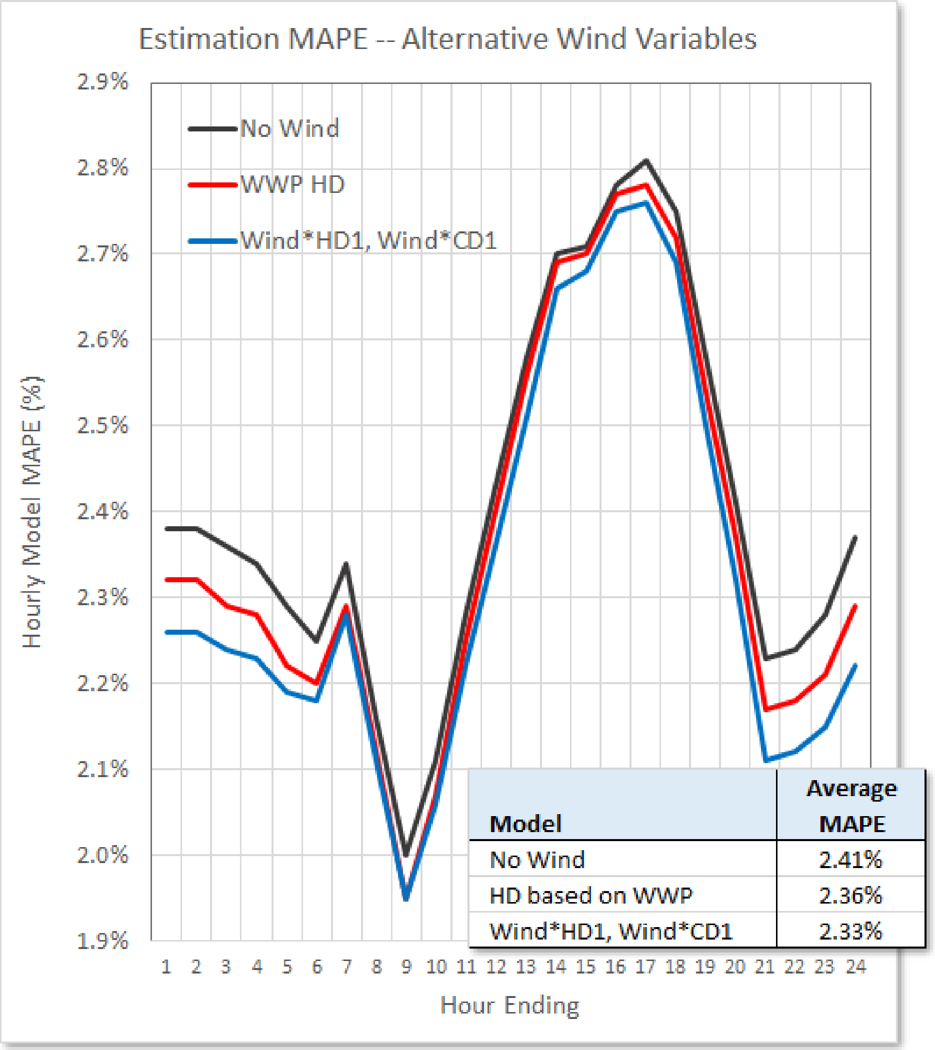
- WWP = T - .5 * Max (Wind – 10, 0)
  - HD values are computed from WWP
  - Wind does not matter until 10 MPh
  - Beyond 10 MPH, + 1 MPH wind = -.5 Degrees
  - Cold side only

» Wind has two effects on building systems

- Insulation: Wind removes thin insulating layer of surrounding air
  - This effect starts at low wind speeds
  - Interacts with solar irradiation, which heats building surfaces
- Infiltration: Causes air changes to increase
- These effects work on both the hot side and cold side

» Recommend alternative specification

- Wind * CD1 on the hot side, Wind * HD1 on the cold side
- With this change:
  - Wind matters at all wind speeds
  - Wind effects are bigger at more extreme temperatures



Model	Average MAPE
No Wind	2.41%
HD based on WWP	2.36%
Wind*HD1, Wind*CD1	2.33%

Estimates from model using reconstituted loads

Recommended Cloud Variables

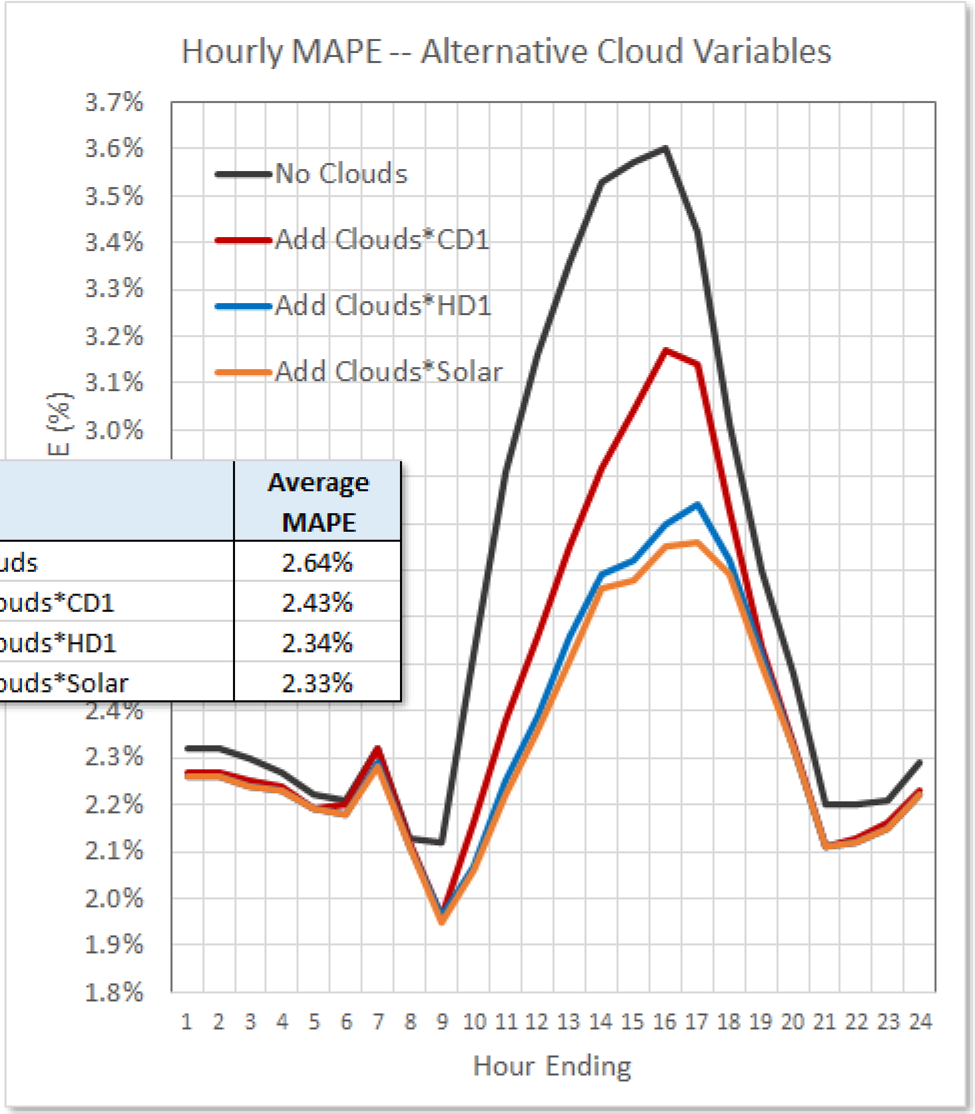
» Current models have no cloud variables

» Clouds have three impacts
- Clouds during the day reduce load when it is hot
- Clouds during the day increase load when it is cold
- Clouds during the day reduce BTM solar production

» Model Y variable is usage (reconstituted load)
- BTM solar generation does not reduce end-use consumption

» Cloud Variables:
- Cloud data ranges from 0 (clear) to 8 (very cloudy)
- Clouds * HD1 to capture cold side cloud impacts
- Clouds * CD1 to capture hot side cloud impacts
- Estimated coefficients are strongly significant
  - T statistics average 6.5 on cold side, 10.7 on hot side

» Recommend adding Cloud variables to the model
- Clouds * HD1, Clouds * CD1
- With reconstituted load, we don't need Clouds*Solar



Model	Average MAPE
No Clouds	2.64%
Add Clouds*CD1	2.43%
Add Clouds*HD1	2.34%
Add Clouds*Solar	2.33%

Estimates from model using reconstituted loads

Incorporate SAE Drivers

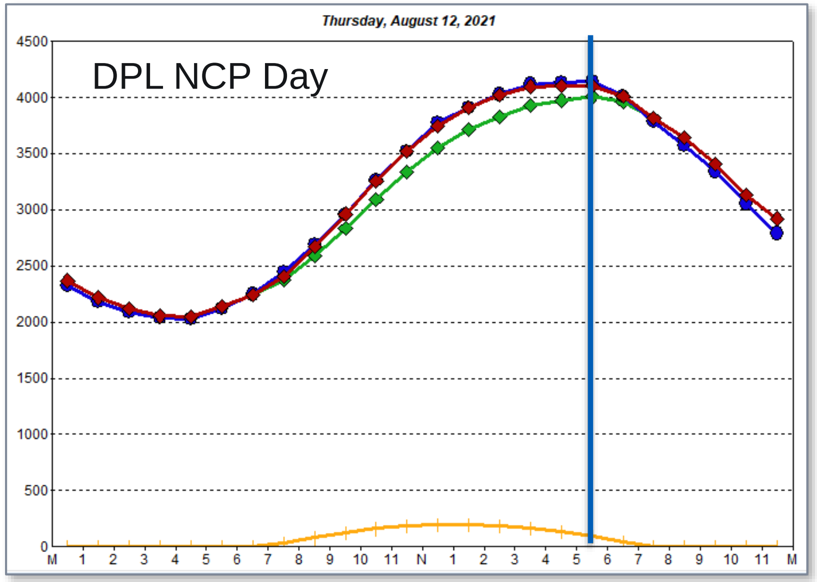
» Incorporate SAE Indexes
  • Interact XOther with constants and shifts
    - Monthly binary, day-of-week, holiday, …
  • Interact XCool with TD variables
    - TD1, TD2, LagTD, Wind*TD1, Clouds*TD1.
  • Interact XHeat with HD variables
    - HD1, HD2, LagTD, Wind*HD1, Clouds*HD1

» End-use assumptions about saturation levels and efficiency are driving the forecast

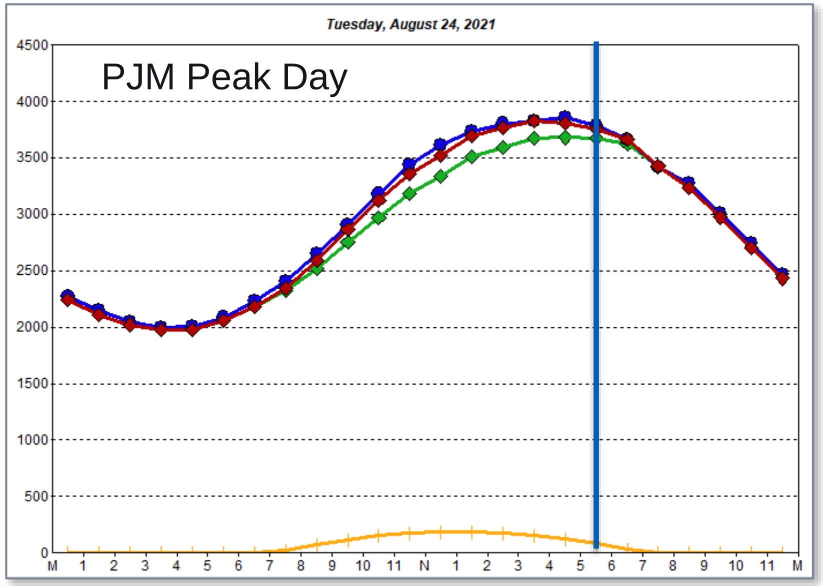
» Results for DPL peak day and PJM peak day in 2021 are shown at right

» Results for DPL coincidence factor (CP/NCP) in 2021 are shown below

	NCP	CP	Coincidence CP/NCP
Actual Load	4,006.0	3,670.0	91.6%
Predicted Load	4,052.4	3,697.5	91.2%
Error	-46.4	-27.5	0.4%
Error %	-1.2%	-0.7%	



	DPL NCP 8/12/HE18
Actual Usage	4,100.3
Predicted Usage	4,146.7
Usage Error	-46.4
Solar_DR	94.3
Actual Load	4,006.0
Predicted Load	4,052.4
Load Error	-46.4



	DPL CP 8/24 HE18
Actual Usage	3,753.6
Predicted Usage	3,781.0
Usage Error	-27.5
Solar_DR	83.6
Actual Load	3,670.0
Predicted Load	3,697.5
Load Error	-27.5

Estimated Model: HE 18 Rolling Two-Part with SAE Interactions

Variable	Coefficient	StdErr	T-Stat
DayTypes.Intercept_XOther	689.463	20.493	33.643
MonthVars.Jan	-5.379	10.614	-0.507
MonthVars.Feb	-29.739	8.883	-3.348
MonthVars.Mar	-77.148	8.422	-9.161
MonthVars.MarDST	-4.426	2.792	-1.585
MonthVars.Apr	-103.989	10.321	-10.075
MonthVars.May	-88.767	10.386	-8.547
MonthVars.Jun	-67.343	11.844	-5.686
MonthVars.Jul	-16.406	14.683	-1.117
MonthVars.Aug	-21.176	16.810	-1.260
MonthVars.Sep	-53.893	15.671	-3.439
MonthVars.Oct	-81.147	12.993	-6.245
MonthVars.Nov	-32.847	12.326	-2.665
MonthVars.NovDST	-16.495	3.604	-4.577
MonthVars.JanWalk	-1.844	1.361	-1.355
MonthVars.FebWalk	-3.773	1.100	-3.429
MonthVars.MarWalk	-0.718	1.583	-0.454
MonthVars.AprWalk	0.915	1.050	0.872
MonthVars.MayWalk	4.216	1.070	3.940
MonthVars.JunWalk	5.269	1.122	4.695
MonthVars.JulWalk	-0.608	1.049	-0.580
MonthVars.AugWalk	-2.870	0.855	-3.357
MonthVars.SepWalk	-4.592	1.119	-4.104
MonthVars.OctWalk	0.475	1.103	0.431
MonthVars.NovWalk	3.100	1.454	2.132
MonthVars.DecWalk	0.283	1.311	0.216
DayTypes.Monday	39.083	3.532	11.066
DayTypes.Tuesday	34.006	3.565	9.540
DayTypes.Wednesday	33.286	3.572	9.320
DayTypes.Thursday	33.936	3.558	9.537
DayTypes.Friday	25.642	3.552	7.219
DayTypes.Saturday	-12.776	2.736	-4.669

Interacts with XOther

Variable	Coefficient	StdErr	T-Stat
Calendar.MLK	37.283	50.244	0.742
Calendar.PresDay	-19.129	47.730	-0.401
Calendar.GoodFri	-8.112	54.883	-0.148
Calendar.MemDay	-114.256	49.459	-2.310
Calendar.July4th	-134.103	47.159	-2.844
Calendar.LaborDay	-39.744	50.150	-0.793
Calendar.Thanks	-378.458	49.853	-7.591
Calendar.FriAThanks	-127.570	49.667	-2.569
DayTypes.WkBeforeXMas	-7.743	15.597	-0.496
Calendar.XMasEve	-145.364	81.256	-1.789
Calendar.XMasDay	-359.132	50.854	-7.062

Variable	Coefficient	StdErr	T-Stat
Calendar.NYEve	-40.271	66.697	-0.604
Calendar.NYDay	-183.177	46.315	-3.955
DayTypes.WkAfterNewYear	20.083	10.327	1.945
DayTypes.WkDayBeforeHol	22.668	26.707	0.849
DayTypes.WkDayAfterHol	-10.909	17.254	-0.632
DayTypes.Phase1	-78.117	14.309	-5.459
DayTypes.Phase2	90.195	11.040	8.170
DayTypes.Phase3	93.289	12.771	7.305
DayTypes.Phase4	31.587	11.900	2.654
DayTypes.Trend2015	-23.962	4.451	-5.384

Heating Vars

Interacts with XHeat

Variable	Coefficient	StdErr	T-Stat
HD1.HD1_17	12.198	4.552	2.680
HD2.HD2_17	21.169	2.171	9.751
Lag6.Lag6HD_17	7.148	2.182	3.276
Lag24.Lag24HD_17	4.650	0.729	6.381
Lag24HC.Lag24CD_HD17	-1.440	7.269	-0.198
WkEndDD.WkEndHD17	0.475	0.628	0.757
SeasHD.SpringHD17	-0.543	0.636	-0.854
SeasHD.FallHD17	-0.977	0.700	-1.397
ColdWind.WindHD17	18.443	2.617	7.047
ColdClouds.CloudHD17	36.500	4.051	9.011
Daily.MA10_HDD	1.248	0.841	1.484
Daily.MA28_HDD	0.289	1.299	0.222
TrendDD.Trend_HD17	-0.670	0.314	-2.137

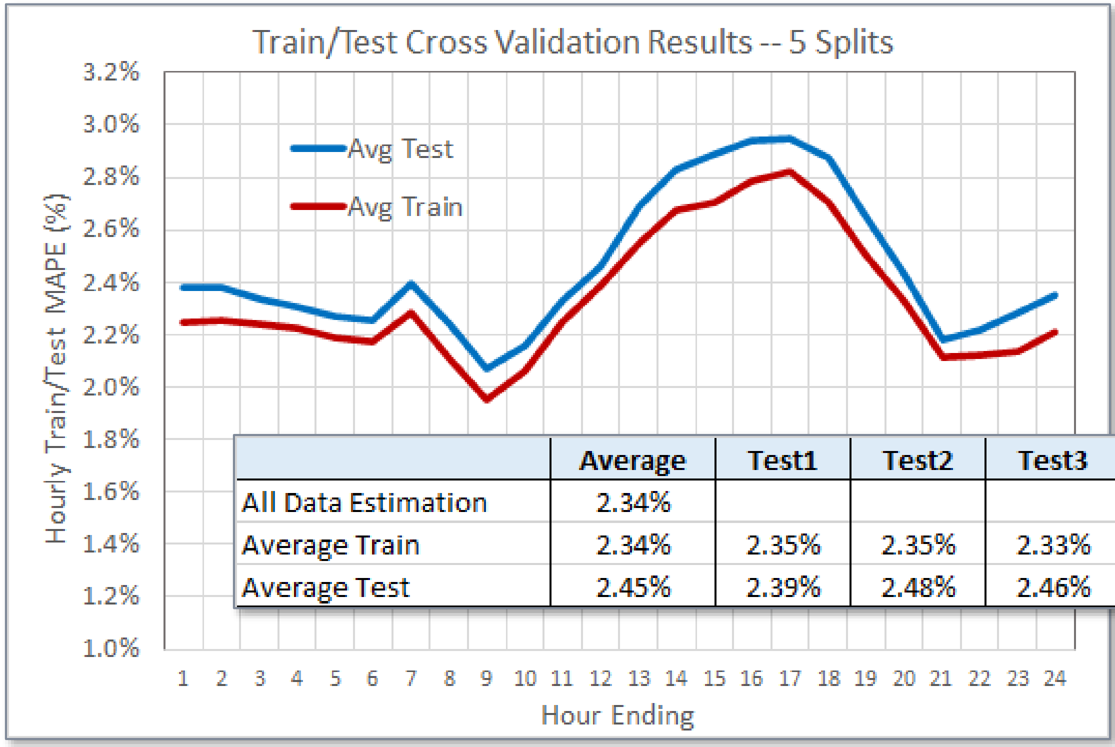
Cooling Vars

Interacts with XCool

Variable	Coefficient	StdErr	T-Stat
CD1.TD1_17	26.219	4.517	5.804
CD2.TD2_17	8.522	1.778	4.792
Lag6.Lag6CD_17	32.113	2.969	10.817
Lag24.Lag24CD_17	6.928	1.307	5.299
Lag24HC.Lag24HD_CD17	-25.499	9.861	-2.586
WkEndDD.WkEndCD17	-0.237	0.870	-0.272
SeasCD.SpringCD17	-19.997	1.493	-13.391
SeasCD.FallCD17	-15.404	2.052	-7.509
HotWind.WindCD17	-2.640	6.026	-0.438
HotClouds.CloudCD17	-137.085	10.435	-13.138
Daily.MA10_CDD	5.254	1.645	3.195
Daily.MA28_CDD	1.096	3.090	0.355
TrendDD.Trend_CD17	1.298	0.330	3.930

Train/Test Statistics for SAE Model

- » Five test runs withholding about 20% of the cases.
- » Average the five train results and the five test results.
- » All hourly generalize very well with about .11% accuracy loss averaged across hours.
- » This is consistent with a robust model that is not over specified.



	Average	Test1	Test2	Test3	Test4	Test5
All Data Estimation	2.34%					
Average Train	2.34%	2.35%	2.35%	2.33%	2.31%	2.34%
Average Test	2.45%	2.39%	2.48%	2.46%	2.52%	2.41%

Hour	Average Train	Average Test
HR00	2.25%	2.38%
HR01	2.26%	2.38%
HR02	2.24%	2.34%
HR03	2.23%	2.30%
HR04	2.19%	2.27%
HR05	2.18%	2.26%
HR06	2.28%	2.39%
HR07	2.11%	2.24%
HR08	1.95%	2.07%
HR09	2.07%	2.16%
HR10	2.25%	2.33%
HR11	2.39%	2.46%
HR12	2.55%	2.69%
HR13	2.67%	2.83%
HR14	2.70%	2.89%
HR15	2.78%	2.94%
HR16	2.83%	2.95%
HR17	2.71%	2.88%
HR18	2.51%	2.66%
HR19	2.33%	2.43%
HR20	2.11%	2.18%
HR21	2.12%	2.22%
HR22	2.13%	2.28%
HR23	2.21%	2.35%
Average	2.34%	2.45%

Reshaping Demand - Technology Impacts

Recommendation for Layering in Future Technology Impacts
Solar and Other Weather-Sensitive Technologies

Extended Weather Simulation.

Step 1. Create Weather Forecast Simulation Traces by EDC and Forecast Year. Continue using the Multi-year Weather Simulation framework.

- Add hourly Solar Generation Capacity factors that are consistent with actual weather patterns.

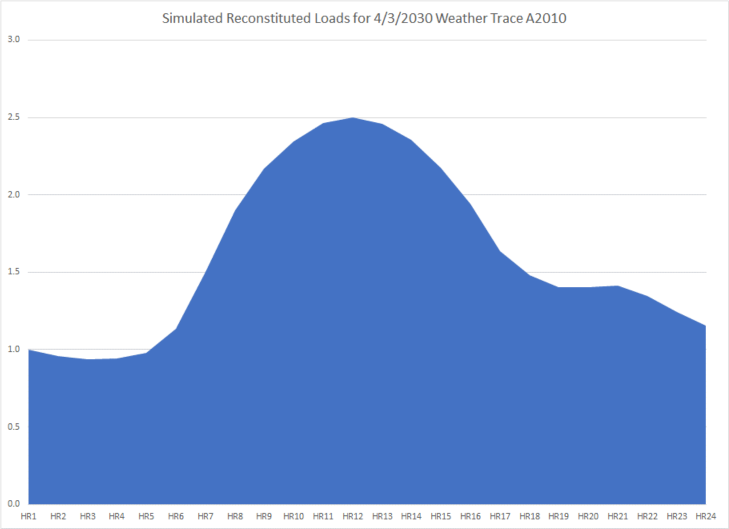
Step 2. For each EDC, Weather Trace and Forecast Year.

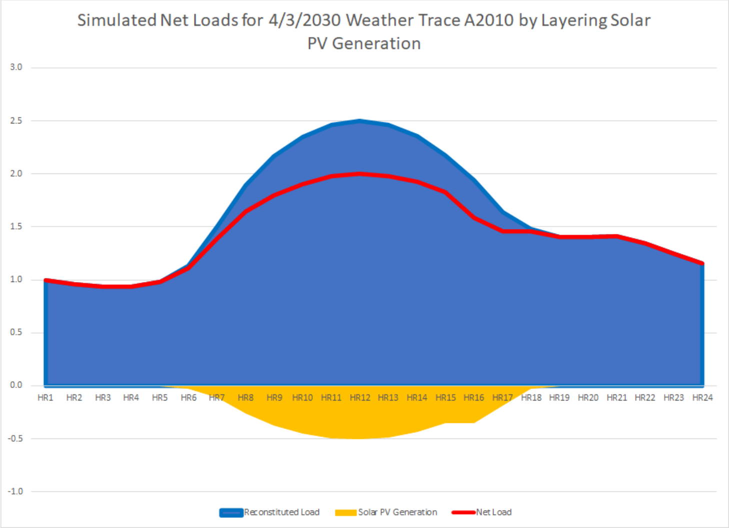
- Simulate EDC-Level Hourly Reconstituted Loads
- Simulate EDC-Level Hourly Solar PV Generation
- Simulate EDC-Level Hourly Other Weather-Sensitive Technology Impact Shapes

Step 3. For each EDC, Layer in Weather Sensitive Impact Shapes by Weather Trace and Forecast Year.

Step 4. For each EDC, Layer in Non-Weather Sensitive Impact Shapes by Forecast Year.

Step 5. For each Weather Trace and Forecast Year construct PJM & LDA Total hourly loads by summing across EDCs by Forecast Year, Weather Trace and EDC.





Recommendation for Layering in Future Technology Impacts
Solar and Other Technologies

Extended Weather Simulation.

Step 1. Create Weather Forecast Simulation Traces by EDC and Forecast Year. Continue using the Multi-year Weather Simulation framework.

- Add hourly Solar Generation Capacity factors that are consistent with actual weather patterns.

Step 2. For each EDC, Weather Trace and Forecast Year.

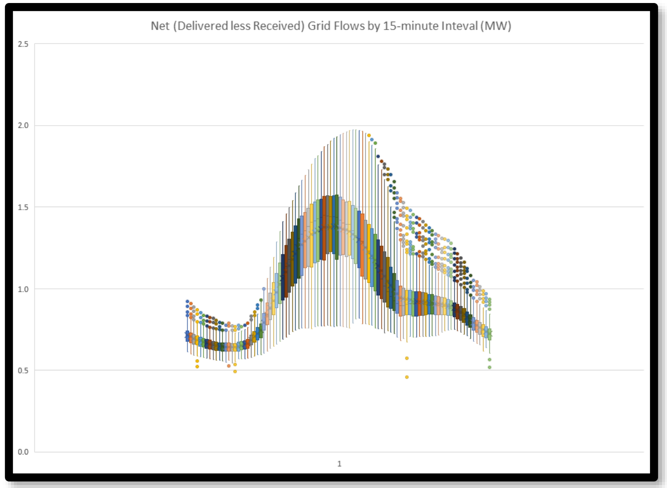
- Simulate EDC-Level Hourly Reconstituted Loads
- Simulate EDC-Level Hourly Solar PV Generation
- Simulate EDC-Level Hourly Other Weather Sensitive Technology Impact Shapes

Step 3. For each EDC, Layer in Weather Sensitive Impact Shapes by Weather Trace and Forecast Year.

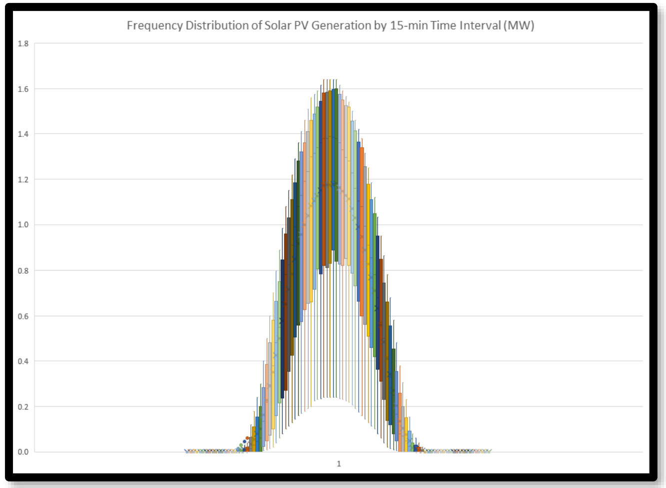
Step 4. For each EDC, Layer in Non-Weather Sensitive Impact Shapes by Forecast Year.

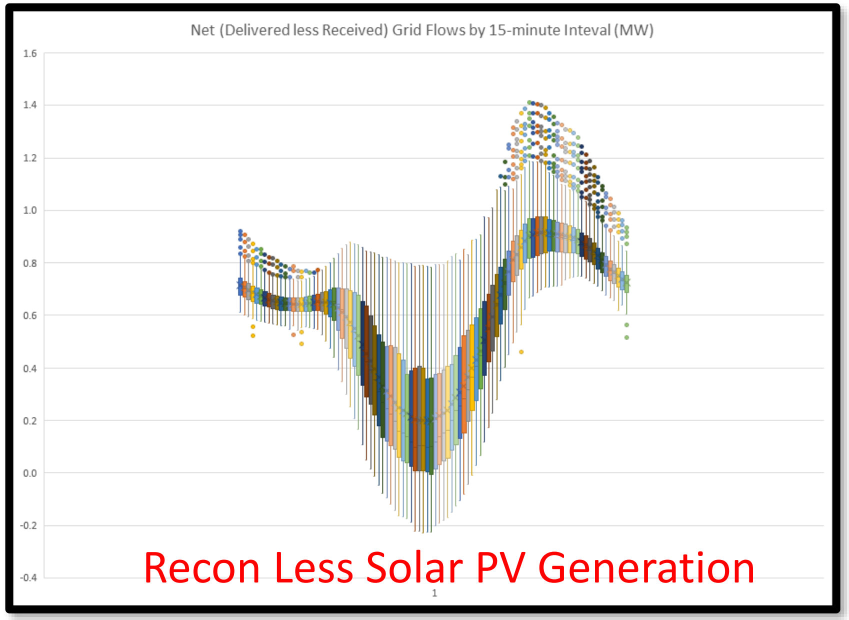
Step 5. For each Weather Trace and Forecast Year Construct PJM & LDA Total hourly loads by summing across EDCs by Forecast Year, Weather Trace and EDC.

Simulated Reconstituted Load



Simulated Solar PV Generation





Recommendation for Layering in Future Technology Impacts
Solar and Other Technologies

Extended Weather Simulation.

Step 1. Create Weather Forecast Simulation Traces by EDC and Forecast Year. Continue using the Multi-year Weather Simulation framework.

- Add hourly Solar Generation Capacity factors that are consistent with actual weather patterns.

Step 2. For each EDC, Weather Trace and Forecast Year.

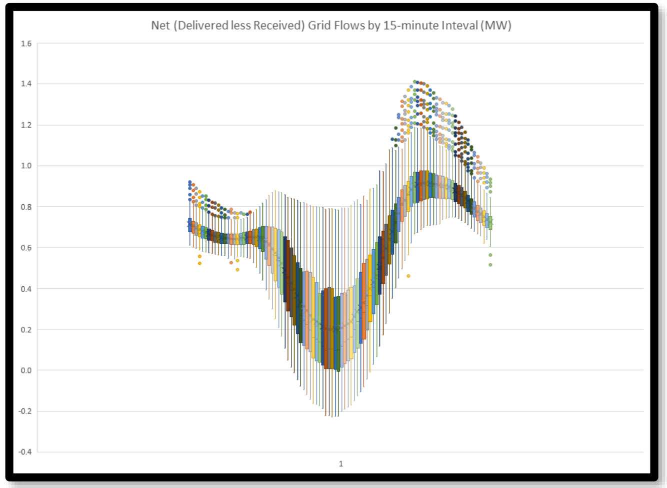
- Simulate EDC-Level Hourly Reconstituted Loads
- Simulate EDC-Level Hourly Solar PV Generation
- Simulate EDC-Level Hourly Other Weather Sensitive Technology Impact Shapes

Step 3. For each EDC, Layer in Weather Sensitive Impact Shapes by Weather Trace and Forecast Year.

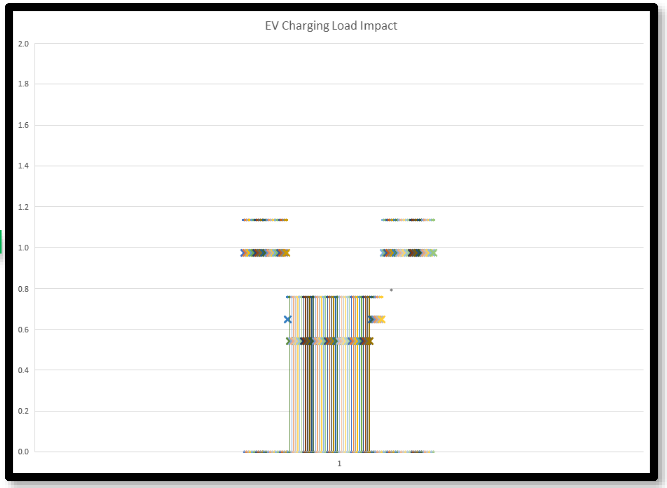
Step 4. For each EDC, Layer in Non-Weather Sensitive Impact Shapes by Weather Trace and Forecast Year.

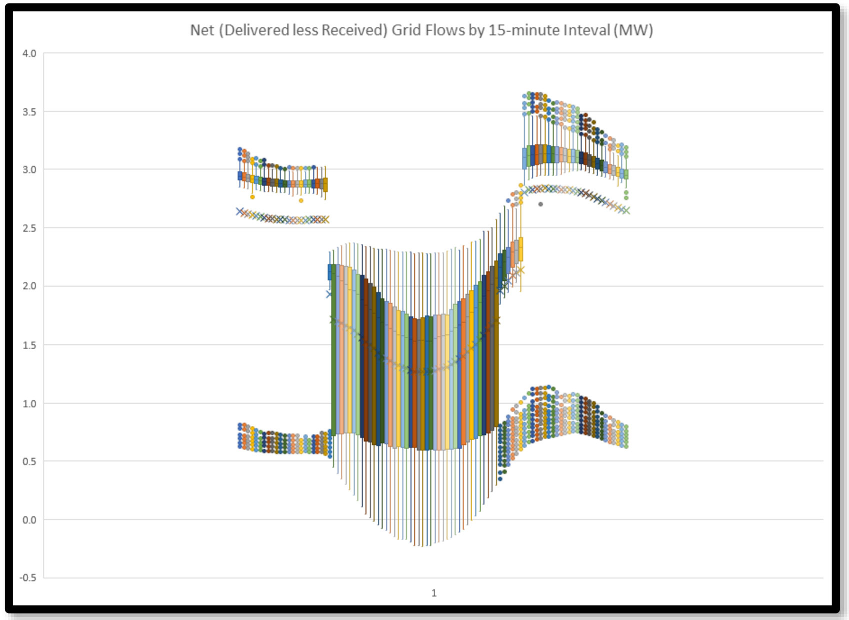
Step 5. For each Weather Trace and Forecast Year Construct PJM Total & PJM Region Hourly Loads by summing across EDCs by Forecast Year, Weather Trace and EDC.

Recon Less Solar PV Generation



EV Charging





Net Load

Other Issues

- Measuring Accuracy
- Weather Normalization
- Capturing Energy Efficiency
- Reflecting Temperature Trends

Measuring Accuracy

- With an SAE model, forecast errors for monthly energy and zone NCP have three components:
  - 1. Actual weather deviations from “normal”
  - 2. Economic growth deviations from forecast
  - 3. Model error, including end-use forecast errors, specification errors, measurement errors, and model error
- We want to track accuracy for energy, zone peaks, and loads coincident with PJM peaks
- The best way to do this is:
  - Calculate predicted values with actual weather and (if possible) actual economics
  - Compute the difference between actual results and predicted results with the actual inputs
    - For monthly energy, this will identify the category 3 or category 2 and 3 errors
    - For zone NCP values, this will identify the category 3 or 2 and 3 errors
    - For zone CP values, there is an additional component related to the specific peculiarities of weather diversity in the weather pattern that actually occurred.
      - This could be a significant source of variance for the more western zones, which tend to exhibit a broader distribution of coincidence factors over time.
      - Compute the model-based coincidence factor using the actual weather data for the year to quantify this source of error.

Weather Normalizing Demand – Changing the Focus

- There are two ways to generate weather normal forecasts and backcasts for a load zone
  - 1: Pre-process weather to create a "normal" weather series
    - a. Create normal weather using: (1) Rank & Average or (2) TMY or (3) Other Averaging Process
    - b. Estimate model with actual weather and forecast/backcast with normal weather
  - 2: Multi-year simulations with post processing
    - a. Run each year of historical weather through the load forecast model
    - b. Post process the distribution of forecasts and select the 50% value as the "normal" load value
- For multiple zones that are geographically dispersed, construction of "normal hourly weather" that yields a consistent set of "normal" NCP and CP values across zones is difficult if not impossible. In contrast, weather simulations using actual historical weather patterns yield realistic distributions of NCP and CP values across load zones.
- Caveat: When using weather simulations, the weather pattern that leads to the 50% or median value for one load zone is likely to be different from the weather patterns that lead to the 50% or median value for the other zones.
- The introduction of Distributed Energy Resources like solar PV adds further complexity to the dimensions of "normal" hourly weather.

Weather Normalizing Demand
Changing the Focus

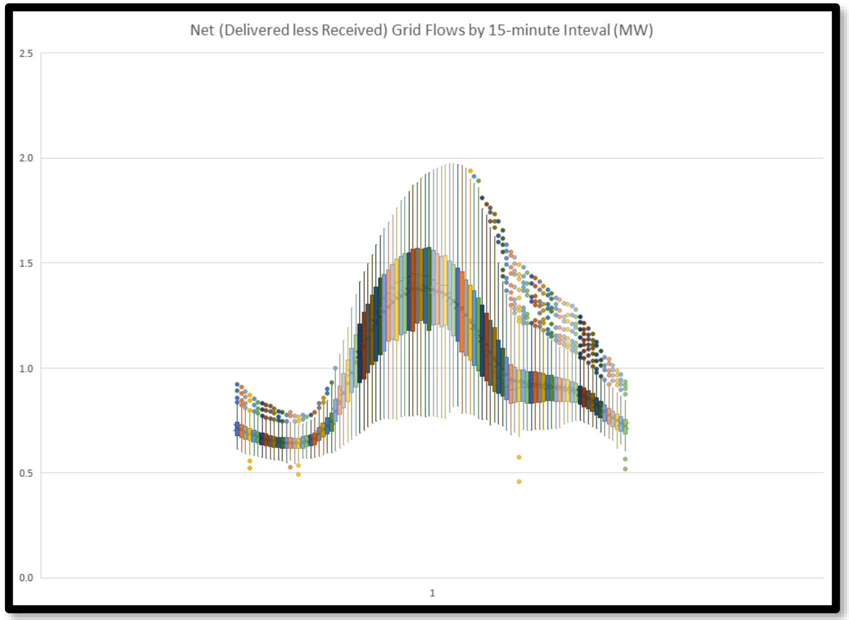
The added load volatility associated with distributed solar PV is driving a growing interest on the part of system operators and planners in quantifying hourly load uncertainty. The weather simulation process that PJM employs is well suited to quantifying hourly load uncertainty.

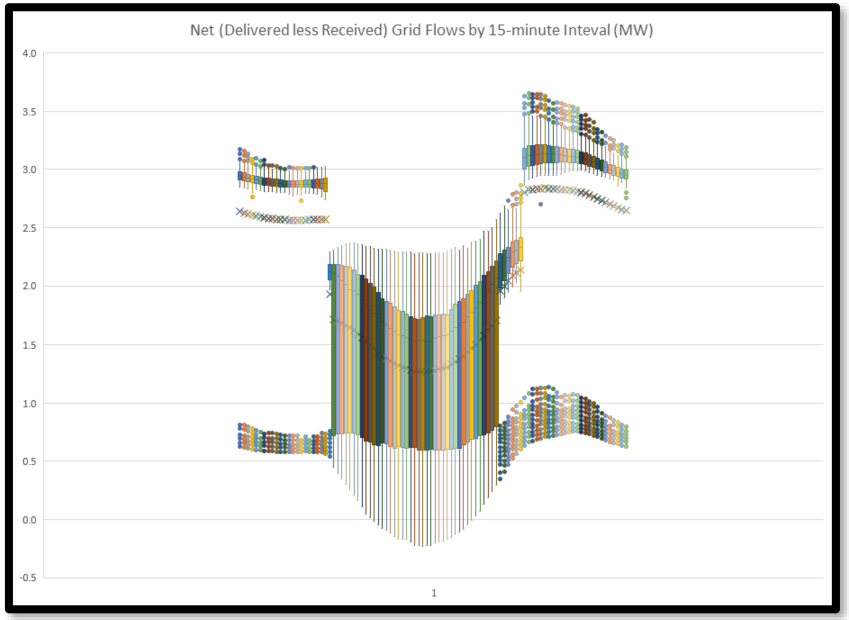
We recommend shifting the focus away from developing “normalized” loads toward using the weather simulation process to quantify historical and future hourly load uncertainty.

The hourly load simulations can also be used to construct distributions of hourly ramp rates that are expected to evolve dramatically with deep penetration of solar PV generation and EV charging.

Hourly load and ramp rate uncertainty can be quantified at key steps in the process. For example,

- Reconstituted Load – quantifies demand for power regardless of how it sourced
- Net Load – quantifies uncertainty around grid-connected supply and ramping requirements

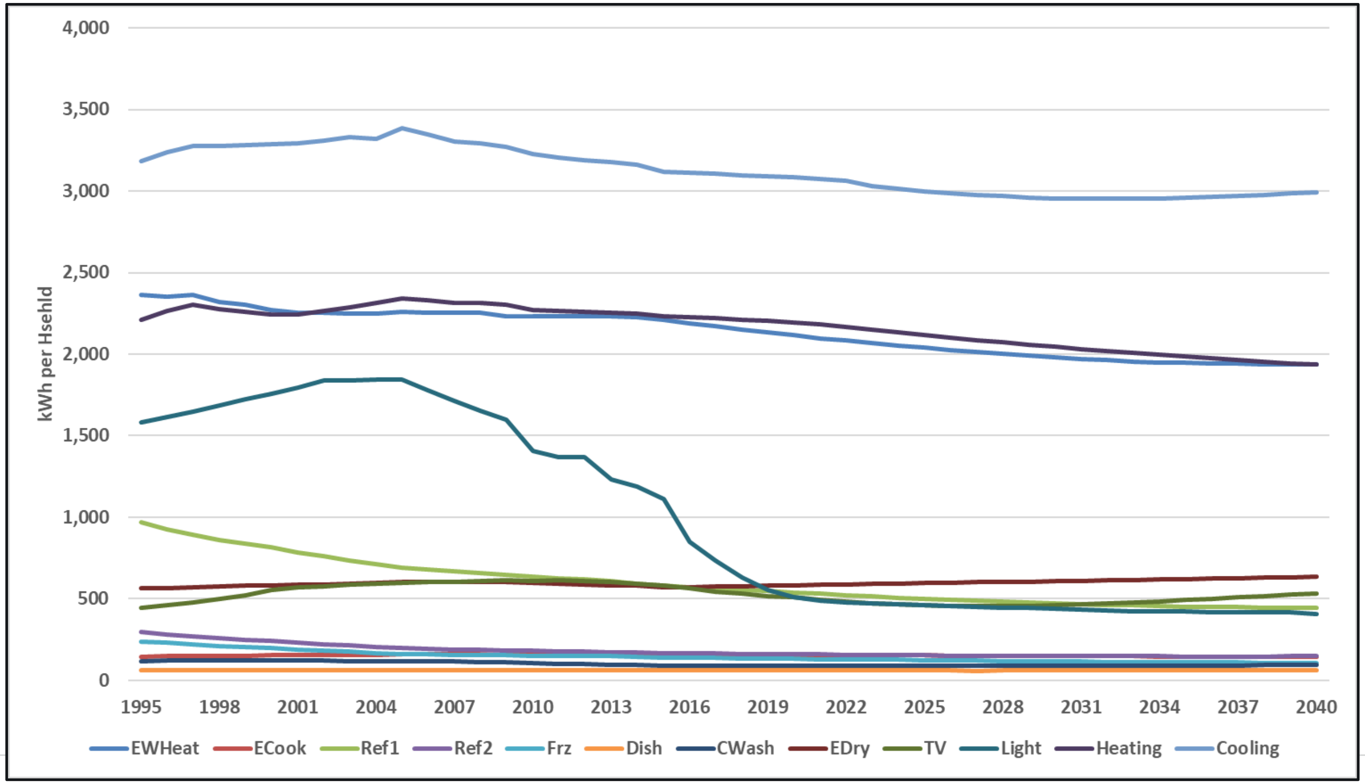




Capturing Efficiency

» Strong energy efficiency trends reflected in the end-use intensities
- Stock efficiency updated annually based on appliance and equipment shipping data
  - End-use choice models captures new standards by limiting the technology efficiency options to those that are greater than the standard
- Accounts for state and utility efficiency programs by “rebating” the more efficient technology options

» Indirectly, the estimation process scales to actual sales data

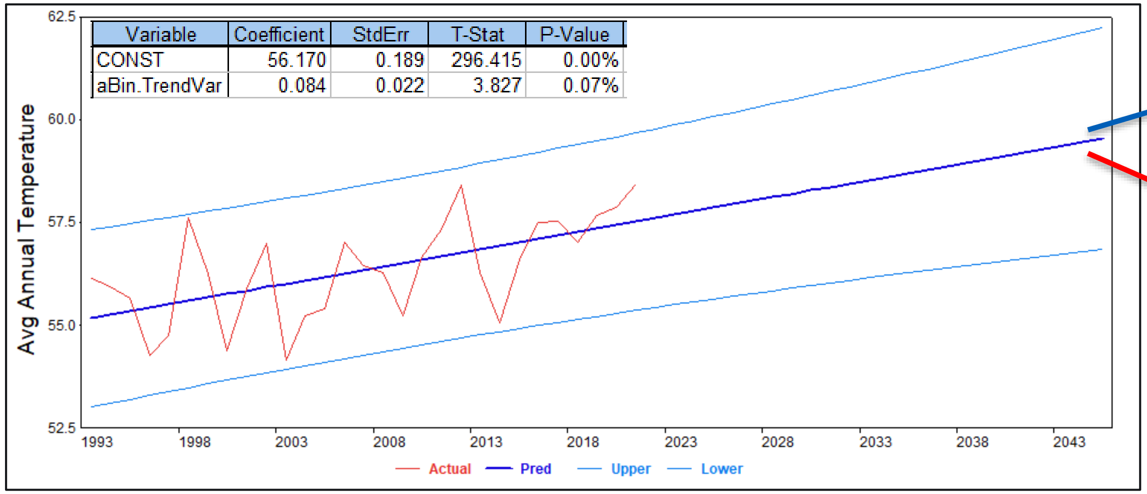


Capturing Temperature Trends

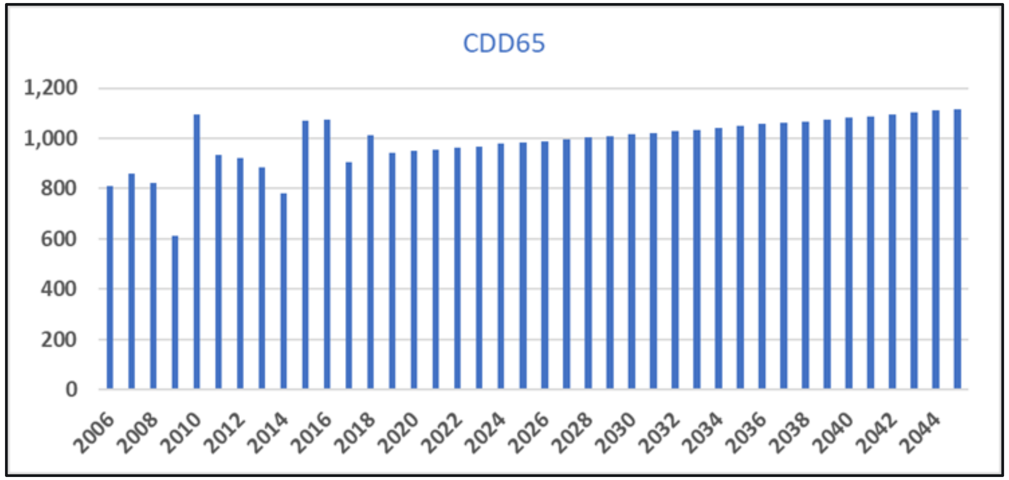
» Average Temperature is Increasing

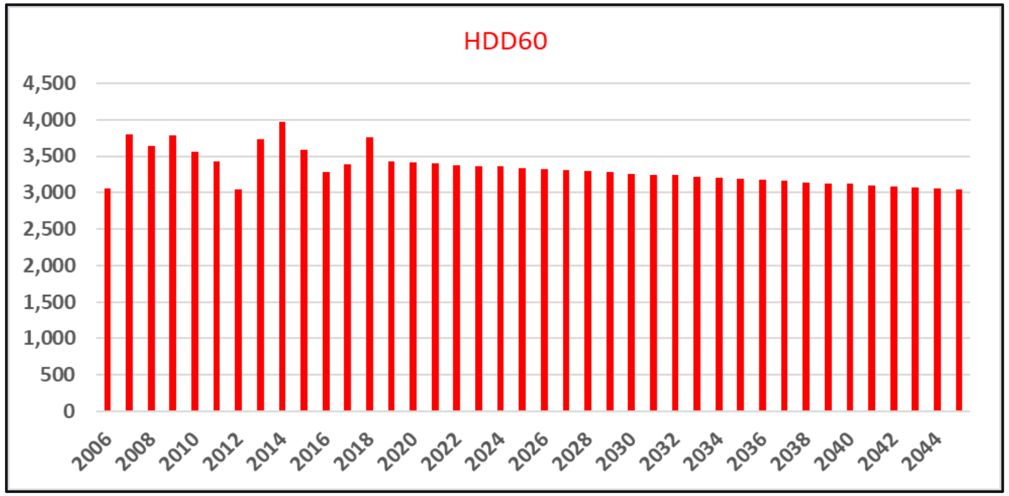
- DPL, 0.8 degrees per decade (1993-2021)

» Translates into increasing CDD and decreasing HDD which can be incorporated into the cooling and heating model indices



Variable	Coefficient	StdErr	T-Stat	P-Value
CONST	56.170	0.189	296.415	0.00%
aBin.TrendVar	0.084	0.022	3.827	0.07%





Summary and Questions

Recommendations

» Replace Annual End-Use Index Calculations with Monthly Calculations

- More observations with monthly data (10 years = 120 observations)
- Capture more variation in heating and cooling loads – stronger weather slopes
- Not as complex – based on well-understood Statistically Adjusted End-Use (SAE) models

» Continue with Weather Simulation Method

- Best approach for accounting for weather diversity
- 20 years of historical data and 7 rotations

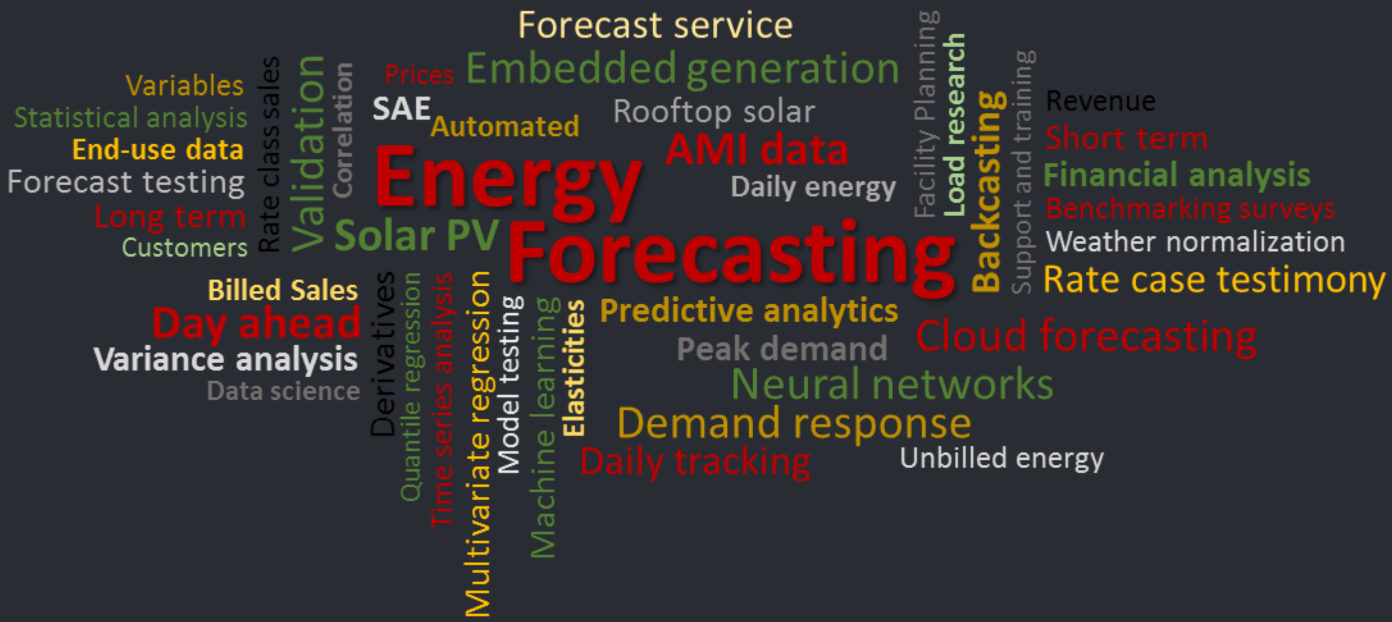
» Replace Daily Peak Models (Independent and Coincident) with Hourly Load Models

- Utilize two-part weather variable construction
- Directly integrate daily or hourly end-use indices

» Adjust Loads using Hourly Technology Load Forecasts

- Carry solar and weather-sensitive end-uses (e.g., cold climate heat pumps) with the weather simulations
- Add non-weather sensitive technologies (e.g. electric vehicles, and data centers) directly to model-generated hourly load forecasts

» Temperatures are increasing. Incorporate temperature trends into the SAE heating and cooling indices



Thank You

http://www.itron.com/forecasting

www.itron.com