F

Demand Forecasting Models

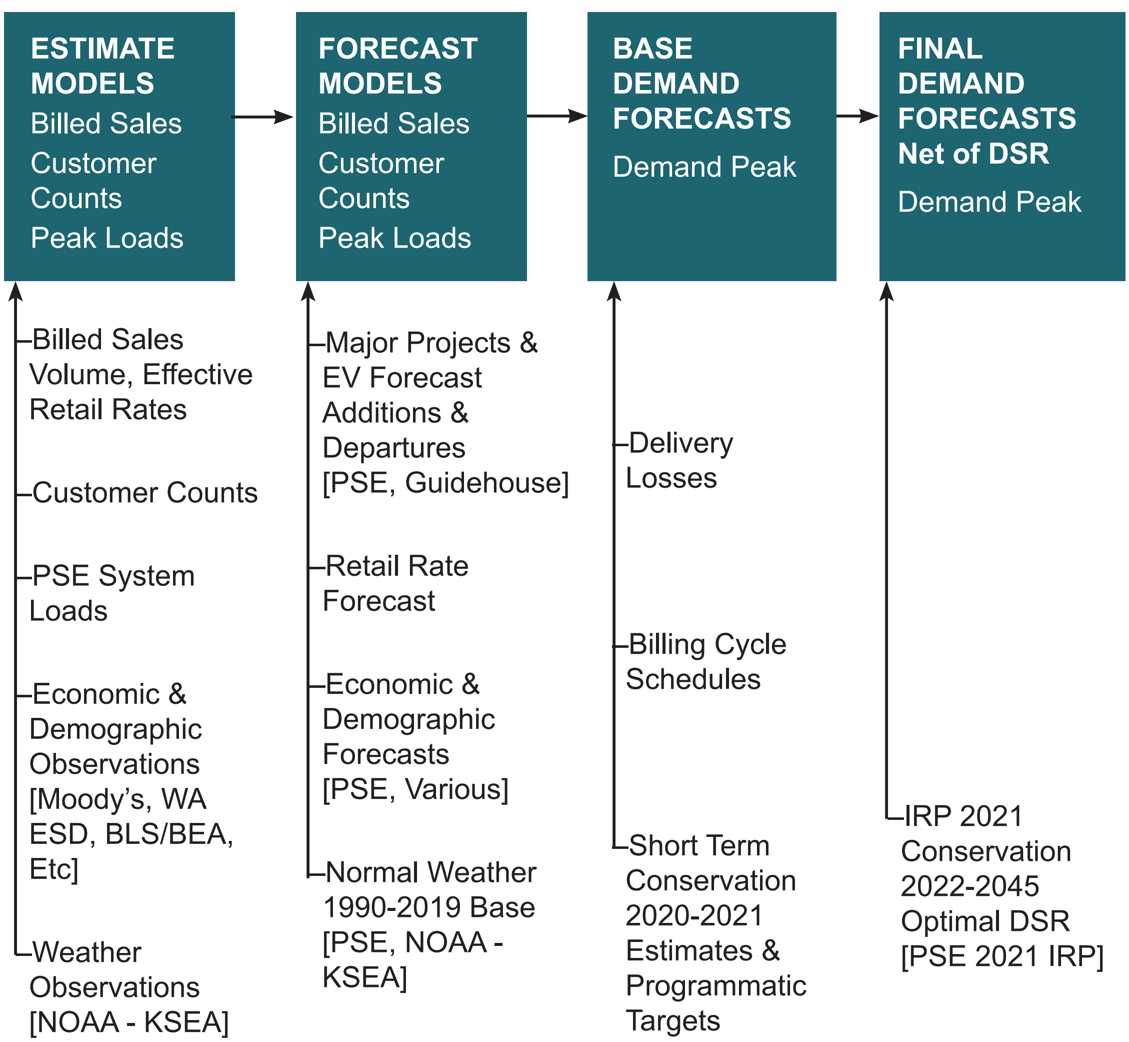
This appendix describes the econometric models used in creating the demand forecasts for PSE's 2021 IRP analysis.

Contents

1. THE DEMAND FORECAST

PSE employs time series econometric methods to forecast monthly energy demand and peaks for PSE's electric and natural gas service territories. PSE gathers observations of sales, customer counts, demand, weather and economic/demographic variables to estimate models of use per customer (UPC), customer counts and peaks. Once model estimation is complete, PSE utilizes internal and external forecasts of new major demand (block sales), retail rates, economic/ demographic drivers, normal weather and programmatic conservation to create a 20-year projection of monthly demand and peaks. The 2021 IRP Base Demand Forecast for energy reflects committed, short-term programmatic conservation targets; the 2021 IRP Base Demand net of demand-side resources (DSR) additionally reflects the optimal DSR chosen in the 2021 IRP analysis. The following diagram depicts the demand forecast development process:

Figure F-1: Demand Forecast Development Process Flow



Model Estimation

To capture incremental customer growth and temperature/economic sensitivities, PSE forecasts billed sales by estimating use per customer (UPC) and customer count models. The models are disaggregated into the following major classes and sub-classes (or sectors, as determined by tariff rate schedule) in order to best estimate the specific driving forces underlying each class.

- Electric: residential, commercial (high-voltage interruptible, large, small/medium, lighting), industrial (high-voltage interruptible, large, small/medium), streetlights and resale
- Natural gas: firm classes (residential, commercial, industrial, commercial large volume and industrial large volume), interruptible classes (commercial and industrial) and transport classes (commercial firm, commercial interruptible, industrial firm and industrial interruptible).

Each class's historical sample period ranges from, at earliest, January 2003 to December 2019.

> > > See Chapter 6, Demand Forecasts, *for discussion of the development of economic/demographic input variables.*

Customer Counts

PSE estimates monthly customer counts by class and sub-class. These models use explanatory variables such as population, employment (both total and sector specific), and unemployment. Larger customer classes are estimated via first differences, with economic and demographic variables implemented in a lagged or polynomial distributed lag form to allow delayed variable impacts. Some smaller customer classes are not estimated, and instead held constant. ARMA(p,q) error structures are also imposed, subject to model fit.

The estimating equations for **customer counts** are specified as follows:*

$$CC_{C,t} = \boldsymbol{\beta_C}\left[\propto_C \quad \boldsymbol{D}_{M,t} \quad T_{C,t} \quad \boldsymbol{ED}_{C,t}\right] + u_{C,t},$$

where:

Customer Count ("$CC_{C,t}$")	=	Count of customers in Class/sub-class "C" and month "t"
Class ("C")	=	Service and class/sub-class, as determined by tariff rate
Time ("t")	=	Estimation time period
Regression Coefficients ("$\boldsymbol{\beta_C}$")	=	Vector of CC_C regression coefficients estimated using Conditional Least Squares/ARMA methods
Constant ("$\propto_C$")	=	Indicator variable for class constant (if applicable)
Date Indicator ("$\boldsymbol{D}_{M,t}$")	=	Vector of month/date specific indicator variables
Trend ("$T_{C,t}$")	=	Trend variable (not included in most classes)
Economic/Demographic Variables ("$\boldsymbol{ED}_{C,t}$")	=	Vector of economic and/or demographic variables
Error term ("$u_{C,t}$")	=	ARMA error term (ARMA terms chosen in model selection process)

** The term vector or boldface type denotes one or more variables in the matrix.*

Use Per Customer

Monthly use per customer (UPC) is estimated at class and sub-class levels using explanatory variables including degree days, seasonal effects, retail rates, average billing cycle length, and various economic and demographic variables such as income and employment levels. Some of the variables, such as retail rates and/or economic variables, are modelled in a lagged form to account for both short-term and long-term effects on energy consumption. Finally, depending on the equation, an ARMA(p,q) error structure is employed to address issues of autocorrelation. The estimating equations for **use per customer** are as follows:*

$$\frac{UPC_{C,t}}{D_{c,t}} = \boldsymbol{\beta_C}\left[\propto_C \quad \frac{\boldsymbol{DD}_{C,t}}{D_{C,t}} \quad \boldsymbol{D}_{M,t} \quad T_{C,t} \quad \boldsymbol{RR}_{C,t} \quad \boldsymbol{ED}_{C,t}\right] + u_{C,t}$$

where:

Use Per Customer ("$UPC_{C,t}$")	=	Billed Sales ("$Billed\ Sales_{C,t}$") divided by Customer Count ("$CC_{C,t}$"), in class "C", month "t"
Cycle Days ("$D_{C,t}$")	=	Average number of billed cycle days for billing month "t" in class "C"
Regression Coefficients ("$\boldsymbol{\beta_C}$")	=	Vector of UPC_C regression coefficients estimated using Conditional Least Squares/ARMA methods
Constant ("$\propto_C$")	=	Indicator variable for class constant (if applicable)
Degree Days ("$\boldsymbol{DD}_{C,t}$")	=	Vector of weather variables. Calculated value that drives monthly heating and/or cooling demand.

$$\boldsymbol{HDD}_{C,Base,t} = \sum_{d=1}^{Cycle_t} |max(0, Base\ Temp - Daily\ Avg\ Temp_d)| * BillingCycleWeight_{C,d,t}$$

$$\boldsymbol{CDD}_{C,Base,t} = \sum_{d=1}^{Cycle_t} |max(0, Daily\ Avg\ Temp_d - Base\ Temp)| * BillingCycleWeight_{C,d,t}$$

Date Indicator ("$\boldsymbol{D}_{M,t}$")	=	Vector of month/date specific indicator variables
Trend ("$T_{C,t}$")	=	Trend variable (not included in most classes)
Effective Retail Rates ("$\boldsymbol{RR}_{C,t}$")	=	The effective retail rate. The rate is smoothed, deflated by a Consumer Price Index, and interacted with macroeconomic variables and/or further transformed.
Economic and Demographic Variables ("$\boldsymbol{ED}_{C,t}$")	=	Vector of economic and/or demographic variables
Error term ("$u_{C,t}$")	=	ARMA error term

** The term vector or boldface type denotes one or more variables in the matrix.*

Peak Electric Hour and Natural Gas Day

The electric and natural gas peak demand models relate observed monthly peak system demand to monthly weather-normalized delivered demand. The models also control for other factors, such as observed temperature, exceptional weather events, day of week, or time of day.

The primary driver of a peak demand event is temperature. In winter, colder temperatures yield higher demand during peak hours, especially on evenings and weekdays. The peak demand model uses the difference of observed peak temperatures from normal monthly peak temperature and month specific variables, scaled by normalized average *monthly* delivered demand, to model the weather sensitive and non-weather sensitive components of monthly peak demand. In the long-term forecast, growth in monthly weather-normalized delivered demand will drive growth in forecasted peak demand, given the relationships established by the estimated regression coefficients.

The **electric peak hour** regression estimation equation is:

$$\max\left(Hour_{1,t} \dots Hour_{H_{t},t}\right) = \boldsymbol{\beta}\left[\frac{Demand_{N,t}}{H_t}\boldsymbol{D}_{M,t} \quad \Delta Temperature_{N,t}\frac{Demand_{N,t}}{H_t}\boldsymbol{D}_{S,t}\boldsymbol{D}_{PeakType,t} \quad \boldsymbol{D}_{DoW,t} \quad D_{LtHr,t} \quad D_{Hol,t} \quad T_{Hot,t}\right] + \varepsilon_t$$

where:

Hourly Demand ("$Hour_{J,t}$")	=	Hourly PSE system demand (MWs) for hour j=1 to $H_{t,}$
Total Hours ("H_t")	=	Total number of hours in a month at time "t"
Regression Coefficients ("$\boldsymbol{\beta}$")	=	Vector of electric peak hour regression coefficients
Normalized Demand ("$Demand_{N,t}$")	=	Normalized total demand in month at time "t"
Temperature Deviation ("$\Delta Temperature_{N,t}$")	=	Deviation of actual peak hour temperature from *hourly* normal minimum peak temperature
Month Indicator ("$\boldsymbol{D}_{M,t}$")	=	Vector of monthly date indicator variables
Month Indicator ("$\boldsymbol{D}_{S,t}$")	=	Vector of seasonal date indicator variables
Peak Type ("$\boldsymbol{D}_{PeakType,t}$")	=	Vector of heating or cooling peak indicators
Day of Week Indicator ("$\boldsymbol{D}_{DoW,t}$")	=	Vector of Monday, Friday, and Mid-Week indicators
Evening Peak ("$D_{LtHr,t}$")	=	Indicator variable for evening winter peak
Winter Holiday ("$D_{Hol,t}$")	=	Indicator variable for holiday effects
Cooling Trend ("$T_{Hot,t}$")	=	Trend to account for summer air conditioning saturation
Error term ("ε_t")	=	Error term

Similar to the electric peaks, the natural gas peak day is assumed to be a function of weather and non-weather-sensitive delivered demand, the deviation of actual peak day average temperature from normal daily average temperature in a month, and type of days.

The **natural gas peak day** estimation equation is:

$$\max\left(Day_{1,t} \ldots Day_{Days_{t},t}\right) = \boldsymbol{\beta}[BDemand_{N,t} \quad \Delta Temperature_{N,t} HDemand_{N,t} \quad \boldsymbol{D}_{M,t} \quad \boldsymbol{D}_{WE,t}] + \varepsilon_t$$

where:

Daily Demand ("$Day_{i,t}$") = Firm delivered dekatherms for day "i"

Total Days ("$Days_{t}$,") = Total number of days in a month at time "t"

Regression Coefficients ("$\boldsymbol{\beta}$") = Vector of gas peak day regression coefficients

Normalized Firm Heating Demand ("$HDemand_{N,t}$") = Normalized monthly firm delivered heating demand

Normalized Firm Base load Demand ("$BDemand_{N,t}$") = Normalized monthly firm delivered base load demand

Temperature Deviation ("$\Delta\, Temperature_{N,t}$") = Deviation of observed daily average temperature from the normal minimum temperature for that month

Month Indicator ("$\boldsymbol{D}_{M,t}$") = Vector of monthly date indicator variables

Weekend Indicator ("$\boldsymbol{D}_{WE,t}$") = Vector of date specific indicator variables

Error term (or "ε_t") = Error term

The natural gas peak day equation uses monthly normalized firm delivered demand as an explanatory variable, and the estimated model weighs this variable heavily in terms of significance. Therefore, the peak day equation will follow a similar trend as that of the monthly firm demand forecast with minor deviations based on the impact of other explanatory variables. An advantage of this process is that it uses demand of distinct natural gas customer classes to help estimate gas peak demand.

Billed Sales Forecast

To forecast billed sales, PSE uses the UPC and customer count models derived above with external and internally derived forecast drivers. Economic, demographic and retail rate forecasts, as well as “normal” monthly degree days, are fitted with model estimates to create the 20-year use per customer and customer count projections by class. The class total billed sales forecasts are formed by multiplying forecasted use per customer and customers ($\widehat{UPC}_{C,t} * D_{C,t} * \widehat{CC}_{C,t}$), then adjusting for known future discrete additions and subtractions (“$Block\ Sales_{C,t}$”).

Major block sales changes are incorporated as additions or departures to the sales forecast as they are not reflected in historical trends covered in the estimation sample period. Examples of such items include emerging electric vehicle (EV) demand, large greenfield developments, changes in usage patterns by large customers, fuel and schedule switching by large customers, or other infrastructure projects. Finally, for the IRP Base Demand Scenario, the forecast of billed sales is reduced by new programmatic conservation (“$Conservation_{C,t}$”) by class, using established conservation targets in 2020-2021.

The total **billed sales forecast** equation by class and service is:

$$Billed\ Sales_{C,t} = \widehat{UPC}_{C,t} * D_{C,t} * \widehat{CC}_{C,t} + Block\ Sales_{C,t} - Conservation_{C,t}$$

Where:

Time (“t”)	=	Forecast time horizon
Use Per Customer (“$\widehat{UPC}_{C,t}$”)	=	Forecast use per customer
Cycle Days (“$D_{C,t}$”)	=	Average number of scheduled billed cycle days for billing month “t” in class “C”
Customer Count (“$\widehat{CC}_{C,t}$”)	=	Forecast count of customers
Conservation (“$Conservation_{C,t}$”)	=	Base Scenario: Ramped/shaped programmatic conservation targets
Major New Sales (“$Block\ Sales_{C,t}$”)	=	Ramped/shaped expected entering or exiting sales not captured as part of the customer count or UPC forecast.

Total billed sales in a given month are calculated as the sum of the billed sales across all customer classes:

$$Total\ Billed\ Sales_t = \sum_c Billed\ Sales_{c,t}$$

Base Demand and Final Demand Net of DSR Forecasts

Demand

Total system demand is formed by distributing monthly billed sales into calendar sales, then adjusting for company own use and losses from distribution, and for electric only, transmission. The electric and natural gas demand forecasts ("$\widehat{Demand}_{N,t}$") form the 2021 IRP Electric and Natural Gas Base Demand Forecasts. For the IRP Final Demand scenario, the optimal conservation bundle is found in the 2021 IRP.

Peak Demand

PSE forecasts peak demand using the peak models estimated above, plus assumption of normal design temperatures, forecasted total system normal demand less conservation ("$\widehat{Demand}_t - Conservation_t$"), and short-term forecasted peak conservation targets. Peak conservation and demand conservation are distinct: they are related, however, different conservation measures may have larger or small impacts on peak when compared with energy. Thus, the peak models seek to reflect exact peak conservation assumption from programmatic activities and the previous Conservation Potential Assessment, as opposed to simple downstream calculations from demand reduction. These calculations yield system hourly peak demand each month based on normal design temperatures.

$$Peak\ Demand_t = F\left(\widehat{Demand}_{t,}, \Delta Temperature_{N,Design,t}\right) - Conservation_{Peak,t}$$

Where:

$Peak\ Demand_t$	=	Forecasted maximum system demand for month "t"
Time ("t")	=	Forecast time horizon
Delivered Demand Forecast ("$\widehat{Demand}_t$")	=	Forecast of delivered demand for month "t"
Temperature Deviation ("$\Delta Temperature_{Normal,Design,t}$")	=	Deviation of peak hour/day design temperature from monthly normal peak temperature
Conservation ("$Conservation_{Peak,t}$")	=	Ramped/shaped peak conservation resulting from programmatic conservation targets; IRP Optimal DSR

For the electric peak forecast, the normal design peak hour temperature is based on the median ("1 in 2" or 50th percentile) of the last of seasonal minimum temperatures for years 1988 to 2017 during peak hours (HE8 to HE20) observed at Sea-Tac (KSEA), as reported by NOAA. For winters spanning 1988 to 2017, the median observed peak temperature is 23 degrees Fahrenheit. The annual winter peak forecast is set at the maximum normal peak observed in a year, which is currently a December weekday evening.

For the natural gas peak day forecast, the design peak day is a 52 heating degree day (13 degrees Fahrenheit average temperature for the day). This standard was adopted in 2005 after a detailed cost-benefit analysis requested by the WUTC. The analysis considered both the value customers place on reliability of service and the incremental costs of the resources necessary to provide that reliability at various temperatures. We use projected delivered demand by class with this design temperature to estimate natural gas peak day demand. PSE’s natural gas planning standard covers 98 percent of historical peak events, and it is unique to our customer base, our service territory and the chosen form of energy.

For the 2021 IRP Base Peak Demand Scenario, the effects of the 2020 and 2021 DSR targets are netted from the peak demand forecast to account for programmatic conservation already underway. This enables the choice of optimal resources and conservation to meet peak demand. Once the optimal DSR is derived from the IRP, the peak demand forecast is further adjusted for the peak contribution of future conservation.

2. STOCHASTIC DEMAND FORECASTS

Demand forecasts are inherently uncertain, and to acknowledge this uncertainty, the IRP considers stochastic forecast scenarios. Examples of drivers of forecast uncertainty include future temperatures, customer growth, usage levels and electric vehicle growth. To model these uncertainties, multiple types of stochastic forecast scenarios are created for different IRP Analyses. These demand and peak forecast permutations include:

- Monthly demand and peak forecasts
  - 250 gas and 310 electric stochastic monthly demand and peak forecasts
  - high/low forecast monthly demand and peak forecasts
- Hourly demand forecasts
  - A typical hourly load shape
  - 88 stochastic hourly forecasts for years 2027-2028 and 2031-2032.

Monthly Demand and Peak Demand

To create the set of stochastic electric and natural gas demand forecasts, the demand forecasts assume economic/demographic, temperature, electric vehicle and model uncertainties. The high and low demand forecasts are derived from the distribution of these stochastic forecasts at the monthly and annual levels.

Economic and Demographic Assumptions

The econometric demand forecast equations depend on certain types of economic and demographic variables; these may vary depending on whether the equation is for customer counts or use per customer, and whether the equation is for a residential or non-residential customer class. In PSE's demand forecast models, the key service area economic and demographic inputs are population, employment, unemployment rate, personal income, manufacturing employment and US gross domestic product (GDP). These variables are inputs into one or more demand forecast equations.

To develop the stochastic simulations of demand, a stochastic simulation of PSE's economic and demographic model was performed to produce the distribution of PSE's economic and demographic forecast variables. Since these variables are a function of key U.S. macroeconomic variables such as population, employment, unemployment rate, personal income, personal consumption expenditure index and long-term mortgage rates, we utilized the stochastic simulation functions in EViews[1] by providing the standard errors for the quarterly growth of key U.S. macroeconomic inputs into PSE's economic and demographic models. These standard errors were based on historical actuals from the last 30 years, ending 2019. This created 1,000 stochastic simulation draws of PSE's economic and demographic models, which provided the basis for developing the distribution of the relevant economic and demographic inputs for the demand forecast models over the forecast period. Outliers were removed from the 1,000 economic and demographic draws. Then 250 draws were run through the electric and natural gas demand forecasts to create the 250 stochastic simulations of PSE's demand forecasts.

Temperature

Uncertainty in the levels of heating and cooling load is modeled by considering varying historical years' degree days and temperatures. Randomly assigned annual "normal" weather scenarios are sourced from actual observations of degree days for electric and natural gas demand and seasonal minimum/maximum on-peak hourly temperatures for electric peak. The years considered for stochastic energy demand and peak range between 1990 and 2019.

Electric Vehicles

PSE's high and low EV energy consumption scenarios are based on PSE's base case EV forecast. The high and low scenarios were developed by calibrating data from the Pacific Northwest National Laboratory's "Electric Vehicles at Scale – Phase I; Analysis: High EV Adoption Impacts on the Western U.S. Power Grid" (July 2020) to PSE's EV forecast. To determine EV energy consumption and peak loads, the ratios of kWh/vehicle and kW/vehicle for residential charging and commercial charging were calculated based on PSE's load forecast data in the year 2028. The ratios were applied to the high and low scenarios of incremental EVs in the PSE balancing area.

Model Uncertainty

The stochastic demand forecasts consider model uncertainty by adjusting customer growth and usage by normal random errors, consistent with the statistical properties of each class/sub-class regression model. Model adjustments such as these are consistent with Monte-Carlo methods of assessing uncertainty in regression models.

The high and low demand forecasts are defined in the IRP as the monthly 90th and 10th percentile, respectively, of the 250 stochastic simulations of demand based on uncertainties in the economic and demographic inputs and the weather inputs.

1 / EViews is a popular econometric forecasting and simulation tool.

Hourly Demand

Resource Adequacy Modelling

For the resource adequacy model, 88 stochastic hourly forecasts for year 2027-2028 and 2031-2032 were developed. For the period April 1, 2013 to December 31, 2019, PSE used the statistical hourly regression equation to estimate hourly demand relationships:

$$Demand_{h,d,s,t} = \boldsymbol{\zeta_h}\left[(1 - D_{h=1})Demand_{h-1,d,t} \quad \boldsymbol{D_{M,t}} \quad D_{Hol,d,t} \quad \boldsymbol{D}_{DoW,d,t} \quad \boldsymbol{T}_{h,d,t}\right] + u_{i,d,t}$$

where:

$$\boldsymbol{T}_{h,d,t} = \left[\max(55 - T_{h,d,t}, 0) \quad \max(T_{h,d,t} - 55,0) \quad \max(55 - T_{h,d,t}, 0)^2 \quad D_{h=1}\max(40 - DAvg_{t-1}, 0) \quad D_{h=1}\max(DAvg_{t-1} - 70,0)\right]$$

Hourly Demand ("$Demand_{h,d,t}$")	=	PSE hourly demand
Hour "h"	=	Hour of day {1…24}
Day "d"	=	Day grouping {Weekday, Weekend/Holiday}
Date "t"	=	Date
Daily temperature shape "s"	=	Indicator of daily average temperature type
Regression Coefficients ("$\boldsymbol{\zeta_h}$")	=	Vector regression coefficients
Hourly Temperature ("$T_{h,d,t}$")	=	Hourly temperature at Sea-Tac ("KSEA")
Lag Daily Average Temp ("$DAvg_{t-1}$")	=	Previous daily average temperature
Monthly Indicator ("$\boldsymbol{D}_{M,t}$")	=	Vector of monthly date indicator variables
Day of Week Indicator ("$\boldsymbol{D}_{DoW,d,t}$")	=	Vector day indicators {Monday, Friday, Sunday}
Holiday Indicator ("$D_{Hol,d,t}$")	=	Holiday indicator
Hour Ending 1 Indicator ("$D_{h=1}$")	=	Indicator Variable for hour ending 1
Error term (or "$u_{i,d,t}$")	=	ARMA(1,1) error term

Demand is estimated for each hour, day of week type and daily average temperature type, yielding 24x2x4 sets of regression coefficients. An annual hourly demand profile is forecasted by fitting an annual 8,760-hour temperature profile and calendar. After creating this fitted value, the forecast is further calibrated by additional hourly demand from an annual EV profile, an AC saturation adjustment for future peak hours with temperatures greater than 72 degrees, the monthly delivered demand ("$\widehat{Demand}_{N,t}$") forecasted for the 2021 Base Demand Forecast, and various stochastic temperature and demand scenarios.

AURORA Modeling Process

An hourly profile of PSE electric demand was produced to support the IRP portfolio analyses. We use our hourly (8,760 hours + 10 days) profile of electric demand for the IRP as an input into the AURORA portfolio analysis. One full year of hourly data is created and then the monthly demand forecast is shaped to the hourly data when running the portfolio analysis. Day one of the hourly shape is a Monday, day two is a Tuesday and so on, so the AURORA model adjusts the first day to line up January 1 with the correct day of the week. The estimated hourly distribution is built using statistical models relating actual observed temperatures, recent demand data and the latest customer counts.