

Appendix B: Residential SAE Modeling Framework

The traditional approach to forecasting monthly sales for a customer class is to develop an econometric model that relates monthly sales to weather, seasonal variables, and economic conditions. From a forecasting perspective, econometric models are well suited to identifying historical trends and to projecting these trends into the future. In contrast, end-use models can incorporate the end-use factors driving energy use. By including end-use structure in an econometric model, the statistically adjusted end-use (SAE) modeling framework exploits the strengths of both approaches.

There are several advantages to this approach.

- The equipment efficiency and saturation trends, dwelling square footage, and thermal integrity changes embodied in the long-run end-use forecasts are introduced explicitly into the short-term monthly sales forecast. This provides a strong bridge between the two forecasts.
- By explicitly incorporating trends in equipment saturations, equipment efficiency, dwelling square footage, and thermal integrity levels, it is easier to explain changes in usage levels and changes in weather-sensitivity over time.
- Data for short-term models are often not sufficiently robust to support estimation of a full set of price, economic, and demographic effects. By bundling these factors with equipment-oriented drivers, a rich set of elasticities can be incorporated into the final model.

This section describes this approach, the associated supporting SAE spreadsheets, and the MetrixND project files that are used in the implementation. The main source of the residential SAE spreadsheets is the 2020 Annual Energy Outlook (AEO) database provided by the Energy Information Administration (EIA).

Statistically Adjusted End-Use Modeling Framework

The statistically adjusted end-use modeling framework begins by defining energy use ($USE_{y,m}$) in year (y) and month (m) as the sum of energy used by heating equipment ($Heat_{y,m}$), cooling equipment ($Cool_{y,m}$), and other equipment ($Other_{y,m}$). Formally,

$$USE_{y,m} = Heat_{y,m} + Cool_{y,m} + Other_{y,m} \quad (1)$$

Although monthly sales are measured for individual customers, the end-use components are not. Substituting estimates for the end-use elements gives the following econometric equation.

$$USE_m = a + b_1 \times XHeat_m + b_2 \times XCool_m + b_3 \times XOther_m + \varepsilon_m \quad (2)$$

$XHeat_m$, $XCool_m$, and $XOther_m$ are explanatory variables constructed from end-use information, dwelling data, weather data, and market data. As will be shown below, the equations used to construct these X-variables are simplified end-use models, and the X-variables are the estimated usage levels for each of the major end uses based on these models. The estimated model can then be thought of as a statistically adjusted end-use model, where the estimated slopes are the adjustment factors.

Constructing XHeat

As represented in the SAE spreadsheets, energy use by space heating systems depends on the following types of variables.

- Heating degree days
- Heating equipment saturation levels
- Heating equipment operating efficiencies
- Average number of days in the billing cycle for each month
- Thermal integrity and footage of homes
- Average household size, household income, and energy prices

The heating variable is represented as the product of an annual equipment index and a monthly usage multiplier. That is:

$$XHeat_{y,m} = HeatIndex_{y,m} \times HeatUse_{y,m} \quad (3)$$

Where:

- $XHeat_{y,m}$ is estimated heating energy use in year (y) and month (m)
- $HeatIndex_{y,m}$ is the monthly index of heating equipment
- $HeatUse_{y,m}$ is the monthly usage multiplier

The heating equipment index is defined as a weighted average across equipment types of equipment saturation levels normalized by operating efficiency levels. Given a set of fixed weights, the index will change over time with changes in equipment saturations (Sat), operating efficiencies (Eff), building structural index ($StructuralIndex$), and energy prices. Formally, the equipment index is defined as:

$$HeatIndex_y = StructuralIndex_y \times \sum_{Type} Weight^{Type} \times \frac{\left(\frac{Sat_y^{Type}}{Eff_y^{Type}} \right)}{\left(\frac{Sat_{15}^{Type}}{Eff_{15}^{Type}} \right)} \quad (4)$$

The $StructuralIndex$ is constructed by combining the EIA's building shell efficiency index trends with surface area estimates, and then it is indexed to the 2015 value:

$$StructuralIndex_y = \frac{BuildingShellEfficiencyIndex_y \times SurfaceArea_y}{BuildingShellEfficiencyIndex_{15} \times SurfaceArea_{15}} \quad (5)$$

The $StructuralIndex$ is defined on the $StructuralVars$ tab of the SAE spreadsheets. Surface area is derived to account for roof and wall area of a standard dwelling based on the regional average square footage data obtained from EIA. The relationship between the square footage and surface area is constructed assuming an aspect ratio of 0.75 and an average of 25% two-story and 75% single-story. Given these assumptions, the approximate linear relationship for surface area is:

$$SurfaceArea_y = 892 + 1.44 \times Footage_y \quad (6)$$

In Equation 4, 2015 is used as a base year for normalizing the index. As a result, the ratio on the right is equal to 1.0 in 2015. In other years, it will be greater than 1.0 if equipment saturation levels are above

their 2015 level. This will be counteracted by higher efficiency levels, which will drive the index downward. The weights are defined as follows.

$$Weight^{Type} = \frac{Energy_{15}^{Type}}{HH_{15}} \times HeatShare_{15}^{Type} \quad (7)$$

In the SAE spreadsheets, these weights are referred to as Intensities and are defined on the *EIADData* tab. With these weights, the *HeatIndex* value in 2015 will be equal to estimated annual heating intensity per household in that year. Variations from this value in other years will be proportional to saturation and efficiency variations around their base values.

For electric heating equipment, the SAE spreadsheets contain two equipment types: electric resistance furnaces/room units and electric space heating heat pumps. Examples of weights for these two equipment types for the U.S. are given in **Table 1**.

Table 1: Electric Space Heating Equipment Weights

Equipment Type	Weight (kWh)
Electric Resistance Furnace/Room units	916
Electric Space Heating Heat Pump	346

Data for the equipment saturation and efficiency trends are presented on the *Shares* and *Efficiencies* tabs of the SAE spreadsheets. The efficiency for electric space heating heat pumps are given in terms of Heating Seasonal Performance Factor [BTU/Wh], and the efficiencies for electric furnaces and room units are estimated as 100%, which is equivalent to 3.41 BTU/Wh.

Price Impacts. In the 2007 version of the SAE models and thereafter, the Heat Index has been extended to account for the long-run impact of electric and natural gas prices. Since the Heat Index represents changes in the stock of space heating equipment, the price impacts are modeled to play themselves out over a 10-year horizon. To introduce price effects, the Heat Index as defined by Equation 4 above is multiplied by a 10-year moving-average of electric and gas prices. The level of the price impact is guided by the long-term price elasticities:

$$HeatIndex_y = StructuralIndex_y \times \sum_{Type} Weight^{Type} \times \frac{\left(\frac{Sat_y^{Type}}{Eff_y^{Type}} \right)}{\left(\frac{Sat_{15}^{Type}}{Eff_{15}^{Type}} \right)} \times (TenYearMovingAverageElectric Price_{y,m})^\phi \times (TenYearMovingAverageGas Price_{y,m})^\gamma \quad (8)$$

Since the trends in the Structural index (the equipment saturations and efficiency levels) are provided exogenously by the EIA, the price impacts are introduced in a multiplicative form. As a result, the long-run change in the Heat Index represents a combination of adjustments to the structural integrity of new

homes, saturations in equipment and efficiency levels relative to what was contained in the base EIA long-term forecast.

Heating system usage levels are impacted on a monthly basis by several factors, including weather, household size, income levels, prices, and billing days. The estimates for space heating equipment usage levels are computed as follows:

$$HeatUse_{y,m} = \left(\frac{WgtHDD_{y,m}}{HDD_{15}} \right) \times \left(\frac{HHSize_y}{HHSize_{15}} \right)^{0.25} \times \left(\frac{Income_y}{Income_{15}} \right)^{0.20} \times \left(\frac{ElecPrice_{y,m}}{ElecPrice_{15,7}} \right)^\lambda \times \left(\frac{GasPrice_{y,m}}{GasPrice_{15,7}} \right)^\kappa \quad (9)$$

Where:

- *WgtHDD* is the weighted number of heating degree days in year (*y*) and month (*m*). This is constructed as the weighted sum of the current month's HDD and the prior month's HDD. The weights are 75% on the current month and 25% on the prior month.
- *HDD* is the annual heating degree days for 2015
- *HHSize* is average household size in a year (*y*)
- *Income* is average real income per household in year (*y*)
- *ElecPrice* is the average real price of electricity in month (*m*) and year (*y*)
- *GasPrice* is the average real price of natural gas in month (*m*) and year (*y*)

By construction, the *HeatUse_{y,m}* variable has an annual sum that is close to 1.0 in the base year (2015). The first two terms, which involve billing days and heating degree days, serve to allocate annual values to months of the year. The remaining terms average to 1.0 in the base year. In other years, the values will reflect changes in the economic drivers, as transformed through the end-use elasticity parameters. The price impacts captured by the Usage equation represent short-term price response.

Constructing XCool

The explanatory variable for cooling loads is constructed in a similar manner. The amount of energy used by cooling systems depends on the following types of variables.

- Cooling degree days
- Cooling equipment saturation levels
- Cooling equipment operating efficiencies
- Average number of days in the billing cycle for each month
- Thermal integrity and footage of homes
- Average household size, household income, and energy prices

The cooling variable is represented as the product of an equipment-based index and monthly usage multiplier. That is,

$$XCool_{y,m} = CoolIndex_y \times CoolUse_{y,m} \quad (10)$$

Where

- $XCool_{y,m}$ is estimated cooling energy use in year (y) and month (m)
- $CoolIndex_y$ is an index of cooling equipment
- $CoolUse_{y,m}$ is the monthly usage multiplier

As with heating, the cooling equipment index is defined as a weighted average across equipment types of equipment saturation levels normalized by operating efficiency levels. Formally, the cooling equipment index is defined as:

$$CoolIndex_y = StructuralIndex_y \times \sum_{Type} Weight^{Type} \times \frac{\left(\frac{Sat_y^{Type}}{Eff_y^{Type}} \right)}{\left(\frac{Sat_{15}^{Type}}{Eff_{15}^{Type}} \right)} \quad (11)$$

Data values in 2015 are used as a base year for normalizing the index, and the ratio on the right is equal to 1.0 in 2015. In other years, it will be greater than 1.0 if equipment saturation levels are above their 2015 level. This will be counteracted by higher efficiency levels, which will drive the index downward. The weights are defined as follows.

$$Weight^{Type} = \frac{Energy_{15}^{Type}}{HH_{15}} \times CoolShare_{15}^{Type} \quad (12)$$

In the SAE spreadsheets, these weights are referred to as Intensities and are defined on the *EIADData* tab. With these weights, the *CoolIndex* value in 2015 will be equal to estimated annual cooling intensity per household in that year. Variations from this value in other years will be proportional to saturation and efficiency variations around their base values.

For cooling equipment, the SAE spreadsheets contain three equipment types: central air conditioning, space cooling heat pump, and room air conditioning. Examples of weights for these three equipment types for the U.S. are given in **Table 2**.

Table 2: Space Cooling Equipment Weights

Equipment Type	Weight (kWh)
Central Air Conditioning	1,012
Space Cooling Heat Pump	306
Room Air Conditioning	277

The equipment saturation and efficiency trends data are presented on the *Shares* and *Efficiencies* tabs of the SAE spreadsheets. The efficiency for space cooling heat pumps and central air conditioning (A/C) units are given in terms of Seasonal Energy Efficiency Ratio [BTU/Wh], and room A/C units efficiencies are given in terms of Energy Efficiency Ratio [BTU/Wh].

Price Impacts. In the 2007 SAE models and thereafter, the Cool Index has been extended to account for changes in electric and natural gas prices. Since the Cool Index represents changes in the stock of space heating equipment, it is anticipated that the impact of prices will be long-term in nature. The Cool Index

as defined Equation 11 above is then multiplied by a 10-year moving average of electric and gas prices. The level of the price impact is guided by the long-term price elasticities.

$$CoolIndex_y = StructuralIndex_y \times \sum_{Type} Weight^{Type} \times \frac{\left(\frac{Sat_y^{Type}}{Eff_y^{Type}} \right)}{\left(\frac{Sat_{15}^{Type}}{Eff_{15}^{Type}} \right)} \times (TenYearMovingAverageElectric Price_{y,m})^\phi \times (TenYearMovingAverageGas Price_{y,m})^\psi \quad (13)$$

Since the trends in the Structural index, equipment saturations and efficiency levels are provided exogenously by the EIA, price impacts are introduced in a multiplicative form. The long-run change in the Cool Index represents a combination of adjustments to the structural integrity of new homes, saturations in equipment and efficiency levels. Without a detailed end-use model, it is not possible to isolate the price impact on any one of these concepts.

Cooling system usage levels are impacted on a monthly basis by several factors, including weather, household size, income levels, and prices. The estimates of cooling equipment usage levels are computed as follows:

$$CoolUse_{y,m} = \left(\frac{WgtCDD_{y,m}}{CDD_{15}} \right) \times \left(\frac{HHSize_y}{HHSize_{15}} \right)^{0.25} \times \left(\frac{Income_y}{Income_{15}} \right)^{0.20} \times \left(\frac{Elec Price_{y,m}}{Elec Price_{15}} \right)^\lambda \times \left(\frac{Gas Price_{y,m}}{Gas Price_{15}} \right)^\kappa \quad (14)$$

Where:

- *WgtCDD* is the weighted number of cooling degree days in year (y) and month (m). This is constructed as the weighted sum of the current month's CDD and the prior month's CDD. The weights are 75% on the current month and 25% on the prior month.
- *CDD* is the annual cooling degree days for 2015.

By construction, the *CoolUse* variable has an annual sum that is close to 1.0 in the base year (2015). The first two terms, which involve billing days and cooling degree days, serve to allocate annual values to months of the year. The remaining terms average to 1.0 in the base year. In other years, the values will change to reflect changes in the economic driver changes.

Constructing XOther

Monthly estimates of non-weather sensitive sales can be derived in a similar fashion to space heating and cooling. Based on end-use concepts, other sales are driven by:

- Appliance and equipment saturation levels
- Appliance efficiency levels
- Average number of days in the billing cycle for each month
- Average household size, real income, and real prices

The explanatory variable for other uses is defined as follows:

$$X_{Other,y,m} = OtherEqpIndex_{y,m} \times OtherUse_{y,m} \quad (15)$$

The first term on the right-hand side of this expression (*OtherEqpIndex_y*) embodies information about appliance saturation and efficiency levels and monthly usage multipliers. The second term (*OtherUse*) captures the impact of changes in prices, income, household size, and number of billing-days on appliance utilization.

End-use indices are constructed in the SAE models. A separate end-use index is constructed for each end-use equipment type using the following function form.

$$ApplianceIndex_{y,m} = Weight^{Type} \times \frac{\left(Sat_y^{Type} / \frac{1}{UEC_y^{Type}} \right)}{\left(Sat_{15}^{Type} / \frac{1}{UEC_{15}^{Type}} \right)} \times MoMult_m^{Type} \times (TenYearMovingAverageElectric Price)^\lambda \times (TenYearMovingAverageGas Price)^\kappa \quad (16)$$

Where:

- *Weight* is the weight for each appliance type
- *Sat* represents the fraction of households, who own an appliance type
- *MoMult_m* is a monthly multiplier for the appliance type in month (m)
- *Eff* is the average operating efficiency the appliance
- *UEC* is the unit energy consumption for appliances

This index combines information about trends in saturation levels and efficiency levels for the main appliance categories with monthly multipliers for lighting, water heating, and refrigeration.

The appliance saturation and efficiency trends data are presented on the Shares and Efficiencies tabs of the SAE spreadsheets.

Further monthly variation is introduced by multiplying by usage factors that cut across all end uses, constructed as follows:

$$ApplianceUse_{y,m} = \left(\frac{BDays_{y,m}}{30.44} \right) \times \left(\frac{HHSize_y}{HHSize_{15}} \right)^{0.46} \times \left(\frac{Income_y}{Income_{15}} \right)^{0.10} \times \left(\frac{Elec Price_{y,m}}{Elec Price_{15}} \right)^\phi \times \left(\frac{Gas Price_{y,m}}{Gas Price_{15}} \right)^\lambda \quad (17)$$

The index for other uses is derived then by summing across the appliances:

$$OtherEqIndex_{y,m} = \sum_k ApplianceIndex_{y,m} \times ApplianceUse_{y,m} \quad (18)$$

Supporting Spreadsheets and MetrixND Project Files

The SAE approach described above has been implemented for each of the nine Census Divisions. A mapping of states to Census Divisions is presented in **Figure 16**. This section describes the contents of each file and a procedure for customizing the files for specific utility data. A total of 18 files are provided. These files are listed in Table 3.

Figure 16: Mapping of States to Census Divisions

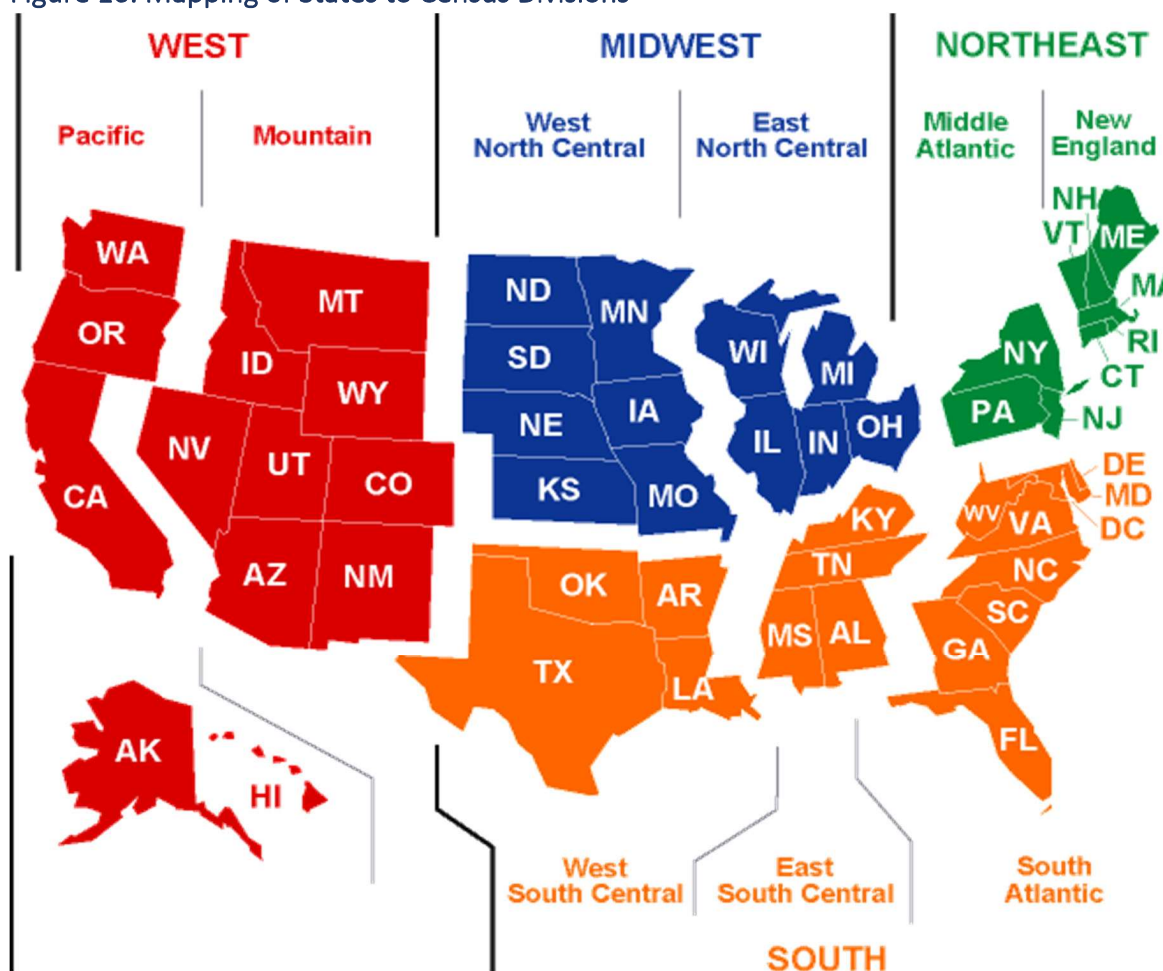


Table 3: List of SAE Files

Spreadsheet	MetrixND Project File
NewEngland.xlsx	SAE_NewEngland.ndm
MiddleAtlantic.xlsx	SAE_MiddleAtlantic.ndm
EastNorthCentral.xlsx	SAE_EastNorthCentral.ndm
WestNorthCentral.xlsx	SAE_WestNorthCentral.ndm
SouthAtlantic.xlsx	SAE_SouthAltantic.ndm
EastSouthCentral.xlsx	SAE_EastSouthCentral.ndm
WestSouthCentral.xlsx	SAE_WestSouthCentral.ndm
Mountain.xlsx	SAE_Mountain.ndm
Pacific.xlsx	SAE_Pacific.ndm

As defaults, the SAE spreadsheets include regional data, but utility data can be entered to generate the *Heat*, *Cool*, and *Other* equipment indices used in the SAE approach. The MetrixND project files link to the data in these spreadsheets. These project files calculate the end-use Usage variables are constructed and the estimated SAE models.

Each of the nine SAE spreadsheets contains the following tabs:

- **Definitions** contains equipment, end use, worksheet, and Census Division definitions.
- **Intensities** calculates the annual equipment indices.
- **Shares** contains historical and forecasted equipment shares. The default forecasted values are provided by the EIA. The raw EIA projections are provided on the *EIAData* tab.
- **Efficiencies** contains historical and forecasted equipment efficiency trends. The forecasted values are based on projections provided by the EIA. The raw EIA projections are provided on the *EIAData* tab.
- **StructuralVars** contains historical and forecasted square footage, number of households, building shell efficiency index, and calculation of structural variable. The forecasted values are based on projections provided by the EIA.
- **Calibration** contains calculations of the base year Intensity values used to weight the equipment indices.
- **EIAData** contains the raw forecasted data provided by the EIA.
- **MonthlyMults** contains monthly multipliers that are used to spread the annual equipment indices across the months.
- **EV** contains a worksheet for incorporating electric vehicle (EV) impacts.
- **PV** contains a worksheet for incorporating photovoltaic battery (PV) impacts.

The MetrixND Project files are linked to the *AnnualIndices*, *ShareUEC*, and *MonthlyMults* tabs in the spreadsheets. Sales, economic, price and weather information for the Census Division is provided in the linkless data table *UtilityData*. In this way, utility specific data and the equipment indices are brought into the project file. The MetrixND project files contain the objects described below.

Parameter Tables

- **Elas.** This parameter table includes the values of the elasticities used to calculate the Usage variables for each end-use. There are five types of elasticities included on this table.
 - Economic variable elasticities
 - Short-term own price elasticities
 - Short-term cross price elasticities
 - Long-term own price elasticities
 - Long-term cross price elasticities

The short-term price elasticities drive the end-use usage equations. The long-term price elasticities drive the Heat, Cool and other appliance indices. The combined price impact is an aggregation of the short and long-term price elasticities. As such, the long-term price elasticities are input as incremental price impact. That is, the long-term price elasticity is the difference between the overall price impact and the short-term price elasticity.

Data Tables

- **AnnualEquipmentIndices** links to the *AnnualIndices* tab for heating and cooling indices, and *ShareUEC* tab for water heating, lighting, and appliances in the SAE spreadsheet.
- **UtilityData** is a linkless data table that contains sales, price, economic and weather data specific to a given Census Division.
- **MonthlyMults** links to the corresponding tab in the SAE spreadsheet.

Transformation Tables

- **EconTrans** computes the average usage, and household size, household income, and price indices used in the usage equations.
- **WeatherTrans** computes the HDD and CDD indices used in the usage equations.
- **ResidentialVars** computes the *Heat*, *Cool* and *Other Usage* variables, as well as the *XHeat*, *XCool* and *XOther* variables that are used in the regression model.
- **BinaryVars** computes the calendar binary variables that could be required in the regression model.
- **AnnualFcst** computes the annual historical and forecast sales and annual change in sales.
- **EndUseFcst** computes the monthly sales forecasts by end uses.

Models

- **ResModel** is the Statistically Adjusted End-Use Model.

Steps to Customize the Files for Your Service Territory

The files that are distributed along with this document contain regional data. If you have more accurate data for your service territory, you are encouraged to tailor the spreadsheets with that information. This section describes the steps needed to customize the files.

Minimum Customization

- Save the MetrixND project file and the spreadsheet into the same folder
- Select the spreadsheet and MetrixND project file from the appropriate Census Division
- Open the spreadsheet and navigate to the *Calibration* tab
- In cell “B9”, replace base year Census Division use-per-customer with observed use-per-customer for your service territory
- Save the spreadsheet and open the MetrixND project file
- Click on the *Update All Links* button on the *Menu* bar
- Review the model results

Further Customization of Starting Usage Levels

In addition to the minimum steps listed above, you can also utilize model-based calibration process described previously to further fine-tune starting year usage estimates to your service territory.

Customizing the End-use Share Paths

You can also install your own share history and forecasts. To do this, navigate to the *Share* tab in the spreadsheet and paste in the values for your region. Make sure that base year shares on the *Calibration* tab reflect changes on the *Shares* tab.

Customizing the End-use Efficiency Paths

Finally, you can override the end-use efficiency paths that are contained on the *Efficiencies* tab of the spreadsheet.