

Demand Response Baseline Methodologies

**Version 20.0**

Table of Contents

[1. Introduction 1](#_Toc149225423)

[2. Baselines 2](#_Toc149225424)

[2.1. Statistical Regression Model 2](#_Toc149225425)

[2.1.1. Model Decomposition 2](#_Toc149225426)

[2.1.2. Daily Energy and Hourly Fraction Models 3](#_Toc149225427)

[2.1.3. Interval Multipliers 3](#_Toc149225428)

[2.1.4. Discussion of Model Variables 3](#_Toc149225429)

[2.1.5. Hourly and Interval Load Variables 4](#_Toc149225430)

[2.1.6. Calendar Variables 4](#_Toc149225431)

[2.1.7. Day of the Week Variables 4](#_Toc149225432)

[2.1.8. Holiday Variables 5](#_Toc149225433)

[2.1.9. Season Variables 6](#_Toc149225434)

[2.1.10. Weather Variables 7](#_Toc149225435)

[2.1.11. Daylight Variables 13](#_Toc149225436)

[2.2. Meter-Before/Meter-After Model 15](#_Toc149225437)

[2.3. Middle 8-of-10 Like Days Model 16](#_Toc149225438)

[2.4. Nearest-20 Like Days Model 16](#_Toc149225439)

[2.5. Matching Day Pair Model 17](#_Toc149225440)

[3. Control Group Baseline Methodology 17](#_Toc149225441)

[4. Matching Sites Methodology 18](#_Toc149225442)

[5. Event Day Adjustment Methodology 19](#_Toc149225443)

# Introduction

As defined in North American Energy Standards Board (NAESB) standards, a baseline is an estimate of the electricity that would have been consumed by a demand response resource in the absence of a demand response event. The baseline is then compared to the resource’s actual metered load during the event to determine the demand reduction value. Baselines are thus the foundation of performance evaluation for demand response.

ERCOT, following exhaustive analysis, has developed and adopted seven baseline types for use in demand response measurement and verification. These baseline types are currently in use in ERCOT’s administration of Emergency Response Service (ERS), where they are referred to as “default baseline types.” They also may be eligible for use in administering demand response participation in other services.

The seven baseline types are: Statistical Regression Model (Reg), Meter-Before/Meter-After Model (MBMA), Middle 8-of-10 Like Days Model (M810), Nearest-20 Like Days Model (Near20), Matching Day Pair Model (MDP), Control Group, Matching Sites. Details for each type are described in sections below.

ERCOT’s analysis has also determined that, for baseline types other than MBMA an event-day adjustment to the model estimates improves the accuracy of the estimates. The same event-day adjustment methodology is applied to each of the four baseline types and is described in the section below titled “Event-Day Adjustment Methodology”.

For each metered load in a demand response resource, whether an individual resource or part of an aggregated resource, ERCOT will determine whether one or more of the baseline types are applicable. If the adjusted load estimates produced by a baseline model, in ERCOT’s judgment, are deemed to be sufficiently accurate and reliable, the adjusted load estimates generated for demand response event days shall be deemed to be the baseline loads for the participating metered loads. These baseline loads, individually or in aggregate as applicable, shall then be compared against the actual loads recorded on those days to assess performance by the resource during the demand response event.

ERCOT also has found that using a Control Group Baseline Methodology for sufficiently large aggregations of residential weather sensitive loads provides load estimates that are accurate and reliable for event performance evaluation. The Control Group Baseline Methodology is described in a separate section of this document.

Some demand response resources’ load patterns are too unpredictable to allow ERCOT to develop a sufficiently accurate baseline model. In these cases, the demand response resource may be disallowed from participation in the applicable ERCOT service. In the case of ERS, such a resource has the option of providing the service under the “Alternate Baseline,” which is known in NAESB standards as the “Maximum Base Load” performance evaluation methodology. Under this methodology, the demand response resource is required to reduce its load to a specified level, regardless of its usage at the time of dispatch, and no attempt is made to estimate the resource’s usage in the absence of the deployment.[[1]](#footnote-1)

It is important to note that while all of the performance evaluation baselines, except Matching Sites, described here are in use for ERS, some of them may not be allowed in the administration of other services.

# Baselines

## Statistical Regression Model

The generalized form of the Statistical Regression Model that will be used for an ESI ID can be written as follows:



Where **e** is the ESI ID,

 **d** is a specific day,

 **h** is an hour on day **d**,

 **int** is a 15-minute interval during hour **h**,

 **kW** is the average load for an ESI ID in a specific 15-minute interval,

 Weather represents weather conditions on the day and preceding days,

 Calendar represents the type of day involved, and

 Daylight represents solar data, such as the time of sunrise and sunset.

Within this general specification, there are an unlimited number of detailed specifications that involve different types of data (such as hourly versus daily weather variables) and different functional specifications that can be used to capture specific nonlinear relationships and variable interactions.

Note that interval load data values recorded during Energy Emergency Alert (EEA) events, during periods of scheduled unavailability of load for curtailment and apparent outlier load values will be excluded from the baseline model building process.

### Model Decomposition

The model to be used is based on the following definitional decomposition.



This decomposition allows analysis of three separate problems. The first is a model of daily energy (kWhd). The second is a model of the fraction of daily energy that occurs in a specific hour (Fracd,h). The third is a model of the load in an interval relative to the average load in the hour to which that interval belongs (Multd,h,int). This breakdown allows development of a robust and relatively rich daily energy model that relies primarily on daily weather and calendar information. The hourly fraction models can then focus more on things that effect the distribution of loads through the day. The interval models can then be designed to distribute the loads within an hour to the 15-minute intervals in that hour.

As an example of how this works, suppose that the following conditions occur:

* + Estimated energy for the day is 36.0 kWh.
	+ The fraction of daily energy that occurs in hour 17 is estimated to be 5.0%.
	+ The load in the first interval of this hour relative to the hourly load is 1.020.

Then the estimated load in kW for the interval from 4 p.m. to 4:15 p.m. is 1.836, computed as follows:



### Daily Energy and Hourly Fraction Models

All three parts of each baseline model are estimated using multivariate regression. In their basic form, the daily energy and hourly fraction models are structured as follows:

 

Where **Y** is the variable to be explained, the **X**’s are the explanatory variables, the **b**’s are the model parameters, and **e** is the statistical error term. For a baseline model, there is one equation of this form for daily energy and 24 equations for the hourly fractions. Although each equation is linear in the parameters, the equations may be highly nonlinear in the underlying variables, such as temperature. These nonlinearities are introduced in the definition of the **X** variables from the underlying weather and calendar factors.

Later sections provide discussions of weather variables, construction of model variables from the weather variables, and interactions between weather and calendar variables.

Business loads vary considerably across ESI IDs in terms of their weather sensitivity and, in general, are less weather sensitive than Residential loads. As a result, some of the baseline models will use a limited set of weather variables. Some ESI IDs will not have significant weather sensitivity on a daily basis, and, as a result, the models for such ESI IDs will be estimated using a simplified season/day-type specification that does not consider the influence of daily and hourly weather patterns.

### Interval Multipliers

The translation from hourly results to 15-minute interval results is performed using multivariate regression of the following form:



and



Thus, the load in a particular 15-minute interval is treated as a function of the hourly load estimate for the hour containing the interval and the hours immediately preceding and following that hour.

### Discussion of Model Variables

The groups of variables that appear in these models are:

* + Hourly and Interval Load Variables
	+ Calendar Variables
* Day of the Week Variables
* Holiday Variables
* Weekday and Weekend Variables
* Season Variables
* Season/Day-Type Interaction Variables
	+ Weather Variables
* Hourly Weather Data
* Weather Zones
* Temperature Variables
* Temperature Slopes
* Weekend Slope Release Variables
* Weather-Based Day-Types
* Heat Buildup Variables
* Temperature Gain Variables
* Time-of-Day Temperature Variables
	+ Daylight Variables
* Daylight Saving
* Time of Sunrise and Sunset
* Fraction of dawn and dusk hours that is dark.

In what follows, each of these groups of variables is discussed separately.

### Hourly and Interval Load Variables

The load data that are used as the dependent variable in the baseline models are developed from 15-minute interval load data in kWh for the individual ESI IDs. Hourly interval load values are created by summing the corresponding 15-minute interval load values.

### Calendar Variables

The main calendar variables include the day of the week, indicators of season, and holiday schedules.

### Day of the Week Variables

The variables used in the models are:

* + **Monday** = 1 on Mondays, 0 otherwise.
	+ **TWT** = 1 on Tuesdays, Wednesdays, and Thursdays, 0 otherwise.
	+ **Friday** = 1 on Fridays, 0 otherwise.
	+ **Saturday** = 1 on Saturdays, 0 otherwise.
	+ **Sunday** = 1 on Sundays, 0 otherwise.

These variables are used in the daily energy and hourly fractional models. The following provides a discussion of the importance of these variables.

* + **Saturday**-Commercial loads tend to be lower on Saturday than on weekdays, reflecting low levels of activity in office buildings and businesses that operate five days per week.
	+ **Sunday**-Commercial loads tend to be lower on Sunday than on weekdays, reflecting low levels of activity in office buildings and small retail and services businesses that are closed or that have abbreviated hours on Sunday.
	+ **Monday**-Monday loads tend to be slightly different than days in the middle of the week. This is especially true for manufacturing operations, where there is often no third shift on Sunday night and Monday morning.
	+ **Tuesday, Wednesday, and Thursday (TWT)**-These days in the middle of the week tend to be highly similar for business loads.
	+ **Friday**-Friday loads tend to be slightly different than days in the middle of the week. Many businesses ramp down earlier on Friday.

### Holiday Variables

In the daily energy models and the hourly fraction models, specific variables are introduced for each individual holiday. Weekday holidays have higher residential loads than typical weekdays and lower business loads. The exact effect on business loads depends on the holiday. For example, on Thanksgiving, most commercial operations are closed. However, on the day after Thanksgiving, office-type operations are usually closed but retail operations are open. All major national holidays fall on fixed days of the week with the exception of Christmas, July 4th, and New Year’s Day, making these three holidays the most difficult to model. The following is a list of all specific holidays that are included in the ERCOT models.

* + **NewYearsHoliday** = Binary variable for New Year’s Day holiday
	+ **MemorialDay** = Binary variable for Memorial Day
	+ **July4thHol** = Binary variable for Independence Day
	+ **LaborDay** = Binary variable for Labor Day
	+ **Thanksgiving** = Binary variable for Thanksgiving
	+ **FridayAfterThanks** = Binary variable for the Friday after Thanksgiving
	+ **ChristmasHoliday** = Binary variable for the Christmas Holiday
	+ **XMasWkB4** = Binary variable for week before Christmas Holiday
	+ **XMasAft** = Binary variable for the week after Christmas Holiday

For NewYearsHoliday, ChristmasHoliday, and July4thHol, the holiday variables are set to 1 for the preceding Friday if the holiday date falls on a Saturday, and on the following Monday if the actual holiday date falls on a Sunday.

#### Major Holidays

In addition, to the individual holidays, a binary variable is constructed for major holidays (MajorHols). The MajorHols variable is defined as the sum of NewYearsHoliday, MemorialDay, LaborDay, Thanksgiving, FridayAfterThanks, and ChristmasHoliday. This variable is used in the definition of the WkDay and WkEnd variables.

#### Weekday and Weekend Variables

The WkDay variable is set to 1 on any weekdays that are not major holidays, and it is set to 0 on any Saturdays, Sundays, or days that are major holidays. The WkEnd variable is defined to be the complement of the WkDay variable. It is 1 on any Saturday, Sunday, or day that is a major holiday, and is 0 otherwise. Formally,

* + **WkDay** = Monday + Tuesday + Wednesday + Thursday + Friday - MajorHols
	+ **WkEnd** = 1 – WkDay

The WkDay and WkEnd variables are interacted with weather slope variables to allow weather slopes to be different on weekdays than they are on weekend days and holidays. For example, to allow the slope on average dry bulb temperature (AveDB) to differ between weekdays and weekend days, the following specification can be used:



Where **KWhd** = the estimated kWh for day d,

 **a** = constant term,

 **b** = slope on average temperature on a weekday,

 **AveDBd** = average dry bulb temperature on day d,

 **c** = slope release for weekend days,

 **WkEndd** = weekend day d.

In this way, the slope on average temperature is given by the value b on a weekday and by the value (b+c) on a Saturday, Sunday, or Major Holiday. If c is positive, then the weather sensitivity on weekends is larger than on weekdays. If c is negative, then the weather sensitivity on weekends is smaller. As a result, the coefficient c is often called a “slope release,” since it releases the weather slope to be different on specific days.

### Season Variables

Two season variables are defined, one for summer months and one for winter months. Effects for remaining months are included in constant terms in the models. The variables are defined as follows:

**Summer** = 1 for days in June, July, August, and September and 0 otherwise.

**Winter** =1 for days in December, January, and February and 0 otherwise.

#### Season/Day-Type Interactions Variables

Several interaction variables are defined to be used in the hourly fraction models. Each of these variables interacts a season variable with a day-type variable. The variables are:

* + **SummerMon** = Summer x Monday
	+ **SummerTWT** = Summer x TWT
	+ **SummerFri** = Summer x Friday
	+ **SummerSat** = Summer x Saturday
	+ **SummerSun** = Summer x Sunday
	+ **WinterMon** = Winter x Monday
	+ **WinterTWT** = Winter x TWT
	+ **WinterFri** = Winter x Friday
	+ **WinterSat** = Winter x Saturday
	+ **WinterSun** = Winter x Sunday

### Weather Variables

#### Hourly Weather Data

Weather variables that are used in the Statistical Regression Baseline Models are:

* + Dry Bulb Temperature
	+ Dew Point

Since different weather providers use different methods to access and download data from the automated stations, the hourly values will show minor variations from one commercial weather data provider to the next. These values are maintained on standard time throughout the year.

#### Weather Zones

Eight weather zones are defined as indicated in Figure 1. This figure also indicates the location of hourly weather stations used to represent each zone.



 Figure 1: Weather Stations Used in ERCOT Systems

#### Computing Weather Zone Variables

Weather variables are defined for each zone based on multiple stations in that zone. The stations that are used and the weights that are applied are presented in Table 1. The weights in this table are in percent and sum to 100 for each zone.



 Table 1: Weather Stations and Zone Weights

At least two weather stations are used to represent weather conditions in each zone. The main advantage of using multiple stations is that the weather variables are less liable to reflect local conditions that are impacting a specific measurement station at a point in time but that are not impacting the larger geographical area.

The weather variables used in the models are calculated from weighted hourly data which were aggregated within a zone. In the calculations, the values for each station are weighted first and then aggregated across stations. For example, the MornDB variable represents the minimum morning dry bulb temperature. In defining this variable for each zone, the order of calculation is:

* Compute the weighted average dry bulb temperature for each hour for stations in the zone.
* Determine the minimum morning dry bulb temperature of the aggregated values for the zone.

The same approach is used for calculating the afternoon and evening maximum values. The daily average values are also computed this way, although the order of the calculations does not matter for computing the daily average values.

#### Temperature Variables

Dry bulb temperature is the temperature of the air as measured by any standard thermometer. As a result, the terms dry bulb temperature and temperature are used interchangeably.

As mentioned above, in the ERCOT models, these variables are transformed by computing aggregates and by computing weighted averages of these aggregate measures across weather stations in a zone. The aggregate concepts are:

* + **AveDB.** This is the Average Dry Bulb Temperature. It is computed as the arithmetic average of the 24 values for the day.
	+ **MornDB.** This is the Minimum Dry Bulb Temperature in Morning Hours. In terms of WeatherBank variables, which are labeled from 0 to 23, the minimum is computed over values labeled Hour4 to Hour8. When the 24 values for a day are renumbered from 1 to 24, the minimum is computed over values 5 to 9.
	+ **AftDB.** This is the Maximum Dry Bulb Temperature in Afternoon Hours. In terms of WeatherBank variables, which are labeled from 0 to 23, the maximum is computed over values labeled Hour11 to Hour16. When the 24 values for a day are renumbered from 1 to 24, the maximum is computed over values 12 to 17.
	+ **EveDB.** This is the Maximum Dry Bulb Temperature in Evening Hours. In terms of WeatherBank variables, which are labeled from 0 to 23, the maximum is computed over values labeled Hour18 to Hour21. When the 24 values for a day are renumbered from 1 to 24, the maximum is computed over values from 19 to 22.

Once these aggregate values are computed for each station in a zone, the weighted average of the values is computed. For example, for average temperature:

 

Where AveDBz = the Average Dry Bulb Temperature in zone z,

 s∈z = the list of stations that are used to represent zone z,

 Wgts = the weight assigned to station s in zone z,

 AveDBs = the average dry bulb temperature for station s.

#### Temperature Slopes

Figure 2 below shows an example of the relationship between daily average temperature and daily energy use (kWh per customer) for the residential sector. This plot shows a strong nonlinear relationship and provides motivation for the temperature variables that are used in the ERCOT models. Specifically, the plot suggests a relatively flat relationship between 60 and 70 degrees, with cooling effects showing on the hot side of the curve (average temperatures above 70) and heating effects showing on the cold side of the curve (average temperatures below 60). To allow further nonlinearities, a second set of cut points are introduced at 50 and 80 degrees. Finally, to allow for “capping” effects that occur when cooling equipment reaches capacity, a final break point is introduced at 85 degrees. The final sets of dry bulb variables included in the models are as follows:

* + XColdSlopez = Max(50 – AveDBz, 0)
	+ ColdSlopez = Max(60 – AveDBz, 0)
	+ MidSlopez = Min(Max(AveDBz – 60, 0), 10)
	+ HotSlopez = Max(AveDBz – 70, 0)
	+ XHotSlopez = Max(AveDBz – 80, 0)
	+ XXHotSlopez = Max(AveDBz – 85, 0)

 Figure 2: Daily Energy vs. Average Temperature

#### Weekend Slope Release Variables

The two main weather slopes are interacted with indicators of day-type, allowing the temperature sensitivity levels to be different on weekend days than they are on weekdays.

* + HotSlopeWkEndz = HotSlopez × WkEnd
	+ ColdSlopeWkEndz = ColdSlopez × WkEnd

#### Weather-Based Day-Types

In addition to the slope variables, a set of weather-based day-type variables are constructed for each weather zone z. These variables are used as interaction variables to allow remaining weather concepts to have different effects when temperatures are warm than they do when temperatures are cold.

* + HotDayz = 1 if AveDBz >70
	+ ColdDayz = 1 if AveDBz <60
	+ MildDayz = 1 – HotDayz - ColdDayz

#### Heat Buildup Variables

In addition to the current day temperature, temperature on preceding days impacts loads through heat buildup effects. The buildup variable is defined as follows.



Where BuildUpz,d = Weighted average lagged temperature for zone z on day d,

 s∈z = the list of stations that are used to represent zone z,

 Wgts = the weight assigned to station s in zone z,

 AveDBs,d-1 = the average temperature for station s on day d-1,

 AveDBs,d-2 = the average temperature for station s on day d-2.

Figure 3 below shows a scatter plot of average temperature versus the buildup variable. In modeling loads, we expect a positive sign on the buildup variable on warm days, since heat buildup will increase cooling requirements for a given temperature level. We expect a negative coefficient on cold days, since higher temperatures on preceding days will reduce heating requirements. As a result, three variables are introduced to allow impact of buildup in a zone be different on hot and cold days. In constructing these variables, the mean value of the Buildup variable across all areas (68.4 degrees) is subtracted out, giving the following:

* + HotBuildUpz = HotDayz × (BuildUpz – 68.4)
	+ ColdBuildUpz = ColdDayz × (BuildUpz – 68.4)
	+ MildBuildUpz = MildDayz × (BuildUpz – 68.4)

By converting this type of variable to deviation-from-the-mean form, it is possible to include the buildup variable interacted with temperature-based day-type variables without including constant releases for the day-type variables, which simplifies the specification.

 Figure 3: Temperature Buildup vs. Temperature



#### Temperature Gain Variables

In addition to the average variables, a temperature gain variable provides an indication of the temperature range. It is computed as the afternoon high temperature minus the morning low temperature. On most days this value is positive, and the average temperature gain is about 17.7 degrees. On a small number of days, the gain is negative, indicating that afternoon temperatures are below morning temperatures.

On each day, the average temperature gain for a zone is computed from the station data as follows:



Where TempGainz = Average temperature gain for a zone,

 s∈z = the list of stations that are used to represent zone z,

 Wgts = the weight assigned to station s in zone z,

 AftDBs = the maximum afternoon temperature for station s,

 MornDBs = the minimum morning temperature for station s.

When modeling daily energy, a bigger value for the range (given the average temperature) will typically imply a larger value for daily energy. This occurs since a bigger range implies larger extreme values, which imply more heating in the winter and more cooling in the summer. To measure these effects, the temperature gain variable is first reduced by its overall mean value of 17.7 degrees, and the result is then interacted with a set of day-type variables, as follows.

* + HotTempGainz = HotDayz × (TempGainz – 17.7)
	+ ColdTempGainz = ColdDayz × (TempGainz – 17.7)
	+ MildTempGainz = MildDayz × (TempGainz – 17.7)

#### Time-of-Day Temperature Variables

The hourly fraction models discussed above utilize the time-of-day temperature variables (MornDB, AftDB, and EveDB). By including all of these variables in each equation, it is possible to model fractions that reflect the full weather pattern for each day. For example, for two days with the same average temperature, the baseline on days that have cool mornings and hot afternoons will be different from the baseline on days that have warm mornings and cool afternoons. In the models, these variables are also interacted with the day-type (hot days and cold days) and with the weekend variable, allowing slopes to differ between weekdays and weekend days. The full set of temperature variables used in the hourly fraction equations is as follows.

HotMornDBz = Hotz × MornDBz

HotAftDBz = Hotz × AftDBz

HotEveDBz = Hotz × EveDBz

WkEndHotMornDBz = WkEnd × HotMornDBz

WkEndHotAftDBz = WkEnd × HotAftDBz

WkEndHotEveDBz = WkEnd × HotEveDBz

MildMornDBz = Mildz × MornDBz

MildAftDBz = Mildz × AftDBz

MildEveDBz = Mildz × EveDBz

ColdMornDBz = Coldz × MornDBz

ColdAftDBz = Coldz × AftDBz

ColdEveDBz = Coldz × EveDBz

WkEndColdMornDBz = WkEnd × ColdMornDBz

WkEndColdAftDBz = WkEnd × ColdAftDBz

WkEndColdEveDBz = WkEnd × ColdEveDBz

### Daylight Variables

#### Daylight Saving/Time of Sunrise and Sunset

Lighting loads have a significant impact on load shapes in the dawn and dusk hours. The timing of these loads is impacted by changes in the time of sunrise and sunset. Although the change is gradual through the annual cycle, there is a one-hour jump at the transitions into and out of Daylight-Saving Time.

Data is stored in one of two ways, clock time and standard time. In either case, adjustments can be made for the changes in the solar cycle and for the changes in human behavior associated with Daylight Saving. The baseline models are estimated with data that are on clock time rather than standard time, implying that sunrise and sunset shift one hour later in March and one hour earlier in November.

 Figure 4: Residential Profiles Before and After Daylight Saving



To track the combination of solar cycles and the incidence of daylight saving, the following variables are included in the daily energy models.

* + DLSav = 1 for days on Daylight Saving Time, 0 otherwise.
	+ HLight = Sunset – Sunrise (both measured in fractions of hours)

#### Fraction of dawn and dusk hours

In addition, in the hourly fraction models, a series of variables is defined to quantify the fraction of each of the dawn and dusk hours that is dark.

* + FracDark7 = Fraction of the hour from 6 a.m. to 7 a.m. that is dark (before sunrise)
	+ FracDark8 = Fraction of the hour from 7 a.m. to 8 a.m. that is dark (before sunrise)
	+ FracDark18 = Fraction of the hour from 5 p.m. to 6 p.m. that is dark (after sunset)
	+ FracDark19 = Fraction of the hour from 6 p.m. to 7 p.m. that is dark (after sunset)
	+ FracDark20 = Fraction of the hour from 7 p.m. to 8 p.m. that is dark (after sunset)
	+ FracDark21 = Fraction of the hour from 8 p.m. to 9 p.m. that is dark (after sunset)

Figure 5 to Figure 9 show examples of these variables plotted over the 1999 calendar. In each chart, the heavy red line is the Fraction Dark variable, and the thin green line represents the number of hours of sunlight. The weighted average values for each zone have been computed and are included in the baseline model spreadsheets.

Figure 5: Hours of Light and Fraction Dark in Hour 7 (6 a.m. to 7 a.m.)



Figure 6: Hours of Light and Fraction Dark in Hour 18 (5 p.m. to 6 p.m.)



Figure 7: Hours of Light and Fraction Dark in Hour 19 (6 p.m. to 7 p.m.)



Figure 8: Hours of Light and Fraction Dark in Hour 20 (7 p.m. to 8 p.m.)



Figure 9: Hours of Light and Fraction Dark in Hour 21 (8 p.m. to 9 p.m.)



## Meter-Before/Meter-After Model

The underlying rationale for the Meter-Before/Meter-After Model is that, for consistently flat loads, the load for a Load site at any point in time can be accurately predicted based on the load immediately preceding that point in time. For this model the kWh consumption during the full fifteen-minute interval that ends immediately preceding the issuance of the dispatch instruction is used as the baseline for all subsequent intervals. If the interval before is affected by other demand response-related events, including Energy Emergency Alert (EEA) events, self-reported deployment events (unrelated to ERCOT actions), periods of notified unavailability of load for curtailment, and other apparent outlier load values, ERCOT may use a different interval that is more representative of the normal load.

## Middle 8-of-10 Like Days Model

The underlying rationale for the Middle 8-of-10 Like Days Model is that the load for a Load site on days of the same day-type that occur close to a demand response event are likely to be similar to “business as usual” load for the event day. In most cases days prior to the event day are used. In some cases, where improved baseline accuracy is achieved, days following the event day, or a combination of days before and after the event day are used.

This concept is used widely in many jurisdictions, and with numerous variations especially on the number of days to evaluate. ERCOT has concluded using Middle 8-of-10 produces the best results. This approach consists of identifying the 10 days having the same day-type as the event day.

Day-types are defined as follows:

* + Weekdays (Monday – Friday excluding holidays)
	+ Weekend / Holidays (Saturday, Sunday and ERCOT Holidays)

The 10 initially selected like days are examined for other demand response-related events, including Energy Emergency Alert (EEA) events, self-reported deployment events (unrelated to ERCOT actions), periods of notified unavailability of load for curtailment, and other apparent outlier load values. If any such days are found among the 10 initially selected like days, they are excluded from the baseline calculation. In this situation additional like days will be incorporated in the baseline calculation unless they also should be similarly excluded. This process continues until ten acceptable days are identified.

The second step consists of calculating kWh consumption for each of the ten days and eliminating the days with the highest and lowest consumption. The third step consists of averaging the interval loads for the eight remaining days together for each interval. The result of this is the unadjusted baseline.

## Nearest-20 Like Days Model

The underlying rationale for the Nearest-20 Like Days Model is the same as that for the Middle 8-of-10 Like Days Model: that the load for a Load site on days of the same day-type that occur close to a demand response event are likely to be similar to “business as usual” load for the event day. For this Model, the twenty like days, both before and after the event day, are used.

As with the Middle 8-of-10 Like Days Model Day-types are defined as follows:

* + Weekdays (Monday – Friday excluding holidays)
	+ Weekend / Holidays (Saturday, Sunday and ERCOT Holidays)

The 20 initially selected like days are examined for other demand response-related events, including Energy Emergency Alert (EEA) events, self-reported deployment events (unrelated to ERCOT actions), periods of notified unavailability of load for curtailment, and other apparent outlier load values. If any such days are found among the 20 initially selected like days, they are excluded from the baseline calculation. In this situation additional like days will be incorporated in the baseline calculation unless they also should be similarly excluded. This process continues until twenty acceptable days are identified.

The second step consists of averaging the interval loads for the twenty days together for each interval. The result of this is the unadjusted baseline.

## Matching Day Pair Model

The underlying rationale for the Matching Day Pair Model is that historical matching day-pairs can be found that have load similar to the actual load on the day preceding the event and the “business as usual” load on the day of the event.

The Matching Day Pair approach consists of matching intervals for the entire day before and the day of the event up to one hour before the start of the event with the corresponding intervals for all day-pairs of the same day-pair type for the preceding year.

Day-pair types are classified by the event day and are defined as follows:

* + Weekdays (Monday – Friday excluding holidays)
	+ Weekend / Holidays (Saturday, Sunday and ERCOT Holidays)

Day-pair types are examined for Energy Emergency Alert (EEA) events, periods of notified unavailability of load for curtailment, and other apparent outlier load values. If any such days are found, they are excluded from the baseline determination.

The similarity of matching day-pairs to the event day-pair is evaluated by calculating the sum of squared differences across the matching intervals between a matching day-pair and the event day-pair as shown in the following formula:

****

The ten-matching day-pairs with the lowest sum of squared differences are identified, deemed to be the best available matches and are averaged together on an interval-by-interval basis to calculate an unadjusted baseline.

The ten-matching day-pairs with the lowest sum of squared differences are identified, deemed to be the best available matches and are averaged together on an interval-by-interval basis to calculate an unadjusted baseline.

If the day before is affected by other demand response-related events, including Energy Emergency Alert (EEA) events, self-reported deployment events (unrelated to ERCOT actions), periods of notified unavailability of load for curtailment, and other apparent outlier load values, ERCOT may use a different baseline methodology that is more representative of the normal load.

# Control Group Baseline Methodology

Note: use of this baseline methodology is limited to Residential sites.

ERCOT will develop a sample design based on the set of sites being evaluated with the Control Group Baseline Methodology to establish the parameters for a sample that will accurately represent the aggregate load of the sites. The sample design may be a simple random sample or a stratified random sample. If a stratified sample is chosen, the stratification variable will be defined based on its ability to improve the accuracy of sample estimates. Stratum boundaries and sample allocation to strata will be defined to optimize the over-all sample accuracy.

ERCOT will randomly select sites to comprise the sample (the control group) and communicate those sites to the QSE representing the resource within the time constraints applicable to the service.

During a test or actual deployment, the QSE representing the resource will deploy all sites except those in the ERCOT sample (control group). Sites in the control group will have their interval-by-interval loads during the event averaged together, and these averages with constitute the per-site baseline for the event. This per-site baseline will be compared to the interval-by-interval per-site load for the deployed group to determine the per-site load reduction.

Periodically ERCOT will select a new control group and communicate those sites to the QSE representing the resource. The new control group will take effect based on the time constraints applicable to the service and be used to develop baselines for subsequent actual or test deployments. The selection of a new control group will be driven by the frequency of actual and test events (more events will increase the frequency).

Whenever significant changes to the set of sites in the resource are made, ERCOT will review the control group sample design, and if deemed necessary, modify the design and select a new control group.

# Matching Sites Methodology

Note: use of this baseline methodology is limited to Residential sites.

For all sites being analyzed for modifications in Load (Load increases or decreases), ERCOT will identify one or more matching sites whose load was similar over a historical time-period, and whose load is expected to approximate what the load of the analysis site otherwise would have been during the period of Load modification.

Initially, candidate matching sites may be identified by pertinent attribute data, such as:

1. Profile type,
2. TDSP code,
3. Zip code,
4. Substation code,
5. KWH stratum (historical period)
6. Date site was first energized,
7. Presence of onsite registered or unregistered generation.

Candidate matching sites are then ranked by the similarity of their historical usage level to usage level of the analysis site during that historical period. The specification of the historical time-period used for comparison will vary based on the nature of the load modification being analyzed. Candidate matching sites will be classified by their use over the historical as higher or lower than the analysis site use for the historical time-period.

The ranking determinant will be applied separately to sites classified as in the lower or higher use category. In most cases the determinant be the mean absolute percent difference between the candidate site’s load and the analysis site’s load. The ‘n’ candidate sites in each of the lower and higher use categories with the lowest mean absolute percent differences will be used as the matching sites and their interval-by-interval loads will be averaged to calculate the baseline for the analysis site. For this purpose, the value of ‘n’ will be based on achieving sufficient baseline accuracy while managing the computational effort involved.

It should be noted that applying the Day of Adjustment methodology described in paragraph below may be considered at the site level to improve the accuracy of the baseline estimates.

# Event Day Adjustment Methodology

Except for the Meter-Before/Meter-After model, an event-day adjustment may be applied to the model estimates (unadjusted baselines) to improve the accuracy of the baseline.

The event-day adjustment is based on kWh usage for an adjustment period preceding the event start time.

In most cases the adjustment period will be the eight fifteen-minute intervals beginning three hours before the event start time. The eight intervals may be moved earlier if necessary, so that none of them occur during an EEA or to offset the effect of outlier usage prior to the event. This prevents atypical load changes occurring before the event, from affecting the baseline adjustment.

In some cases, if more accuracy is realized, ERCOT may elect to use a different adjustment period prior to or following an event. For events that occur early in the day, the adjustment period may be for the prior day.

Actual kWh for the adjustment period is determined by summing the ESI ID’s actual kWh across the intervals in the adjustment period. Baseline kWh is determined by summing the baseline estimates across the same intervals. The adjustment factor is then calculated by dividing the actual kWh by the baseline kWh. The unadjusted baseline kWh for each interval during the event is multiplied by the adjustment factor to calculate the adjusted baseline kWh.

1. For details on the MBL performance evaluation methodology (Alternate), consult the ERCOT Protocols Sections 8.1.3.1.1 and 8.1.3.1.4, and the prevailing ERS Technical Requirements and Scope of Work. [↑](#footnote-ref-1)