Demand Forecasting Methodology

Electricity Distribution Network

Friday, 31 January 2025

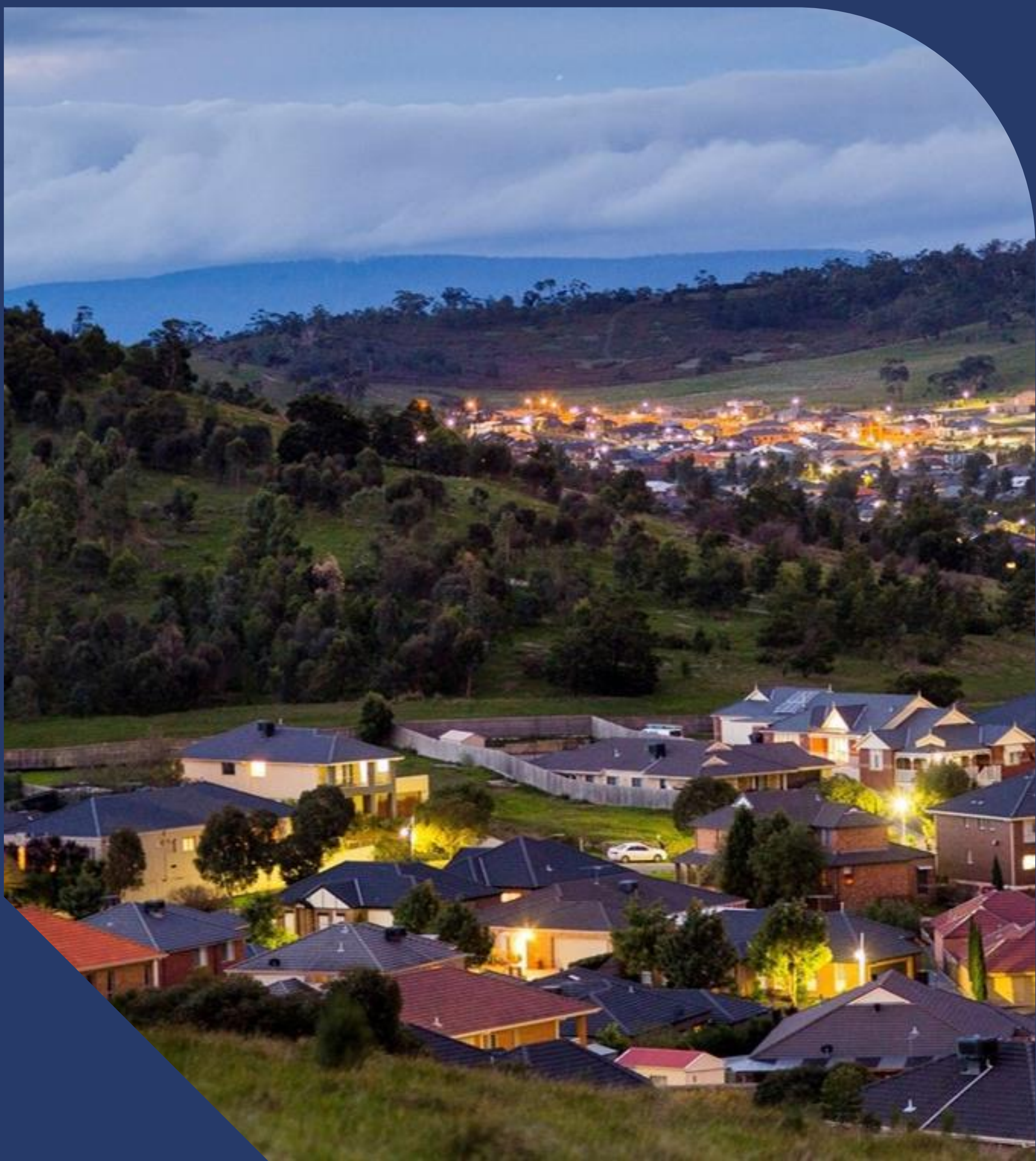


Table of contents

1. Purpose of this document

This document describes AusNet's methodology for forecasting the spatial distribution of maximum and minimum demand on our electricity distribution network. The objective of this forecasting methodology is to produce best practice maximum and minimum demand forecasts that recognise the importance of:

- Accuracy and a lack of bias;
- Transparency and repeatability;
- Incorporation of key drivers; and
- Validation and testing.

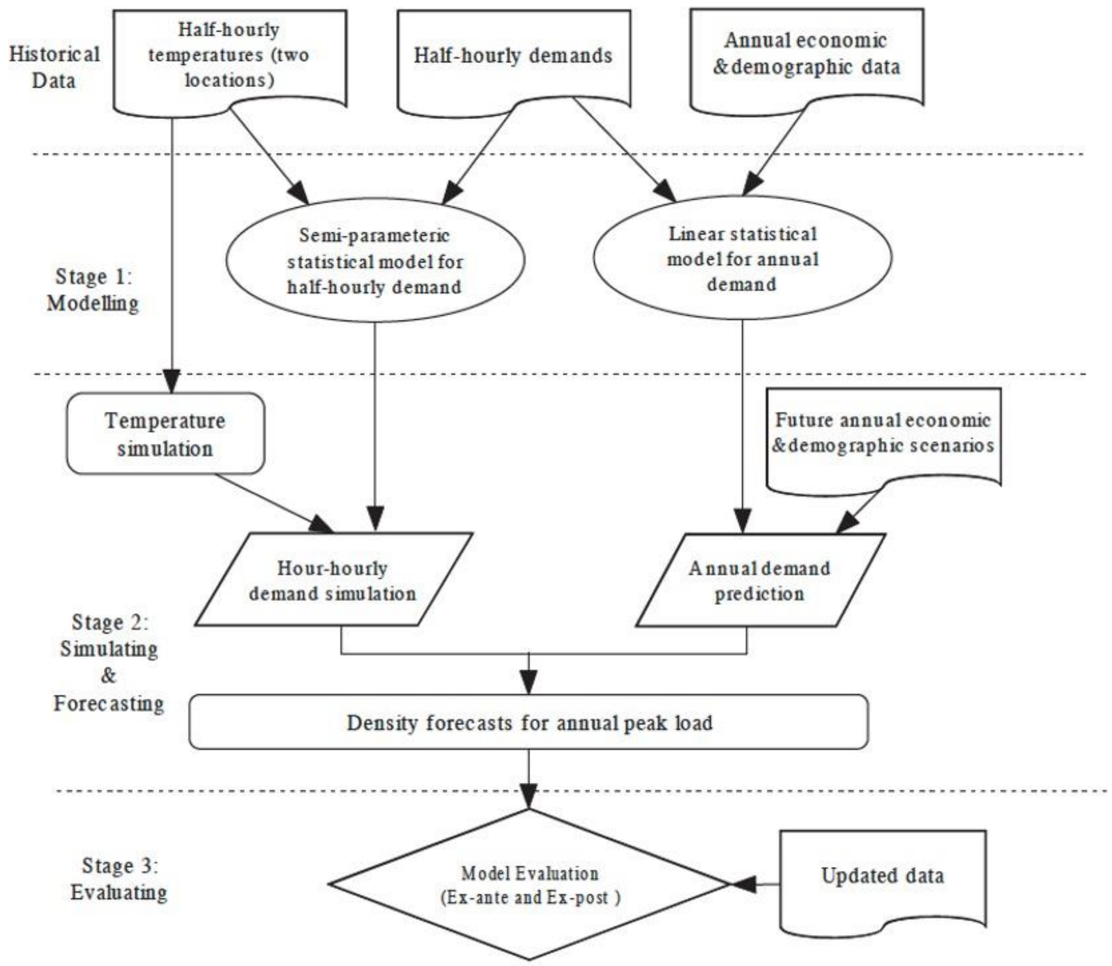
Fifteen-year, forward-looking demand forecasts are prepared annually for AusNet transmission connection assets, zone substations and distribution feeders. The forecasts are prepared for 10%, 50% and 90% Probability of Exceedance (**POE**) conditions, for both maximum and minimum demands, for different seasons during the year.

2. Demand forecasting methodology

2.1. Modelling approach

AusNet's modelling approach is based on a forecasting methodology developed by Monash University in 2015[1], as shown in the figure below. AusNet has adapted the methodology so that it produces maximum and minimum demand forecasts. The modelling approach has also been modified to account for the impact of rooftop PV, Electric Vehicles (EV) and gas electrification, which has become a significant forecasting issue since the model's original development.

Figure 1: Overview of Monash University's demand forecasting methodology



The approach developed by Monash University has two modelling tasks in Stage 1. The first is a model to produce half-hourly forecasts from half-hourly data. The second model produces average demand forecasts from block averages (blocks are defined as two weeks' time windows). The two models are combined in stage 2 to convert the forecast annual demand to half-hourly data, using a number of simulations. From these simulations, the demand forecasts are produced at Stage 3.

The following inputs are used in AusNet's demand forecasting methodology:

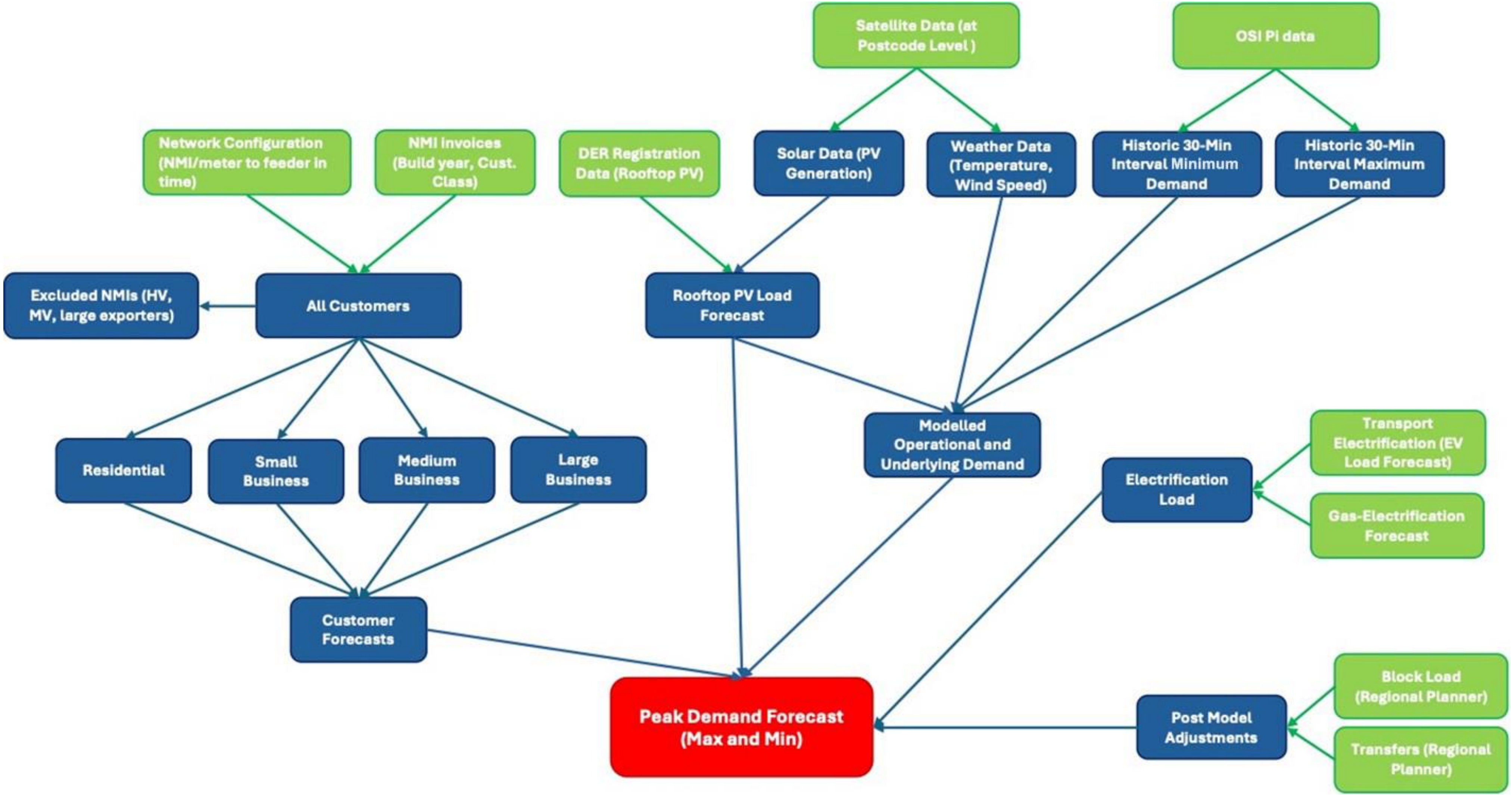
- Assemble historical data, including customer numbers and rooftop PV capacity
- Forecast spatial customer numbers and rooftop PV capacity

[1] Monash University, Electricity Forecasting Model, Professor Rob J Hyndman and Dr Shu Fan, 2015, 15 June 2015, Figure 2.

- Historical and future estimates of electrification trends (e.g. EVs and gas electrification)
- Historical maximum and minimum operational demands at 30-minute intervals
- Historical weather-related and solar variables (e.g., temperature, wind, humidity and solar irradiation) at 30-minute intervals
- Recorded embedded generation including wind as well as large solar generators at 30-minute intervals.

Figure 2 (next page) maps out the various data sources that feed into the demand forecasting methodology (pre-post model adjustments).

Figure 2: Flow chart of data sources used in the demand forecasting methodology



2.2. Process summary

The forecasting process at AusNet is undertaken in R programming language. R is a programming language for statistical computing and data visualisation. It provides practitioners with the tools and techniques they need to complete the forecasting process. The seven key steps in the current spatial and trend analysis forecasting process are:

(1) Assemble historical data. This step includes extracting historical customer numbers, maximum and minimum half hourly operational demand data, rooftop PV capacity, EV numbers, embedded generation, and weather and solar variables.

(2) Forecast spatial customer numbers and rooftop PV capacity, which is informed by government and AEMO's growth estimates.

(3) Model unitised (per customer) underlying maximum (or minimum) half-hourly demand, in two separate but complementary steps, which take account of historical temperature and other factors.

- Average maximum (or minimum) monthly consumption using Cooling Degree Days (**CDDs**) and Heating Degree Days (**HDDs**) temperatures and calendar variables.
- 30-minute fluctuations using weather related variables, time of day and calendar variables.

(4) Forecast the impact of electrification, driven by EVs and gas electrification.

(5) Simulating the future:

- Defining 200 different scenarios for weather and calendar impacts using the bootstrap method.
- Estimating unitised half hourly maximum (or minimum) underlying demand for each scenario and then multiplying by the corresponding customer number forecasts.
- Obtaining half hourly maximum (or minimum) operational demand by subtracting the corresponding rooftop PV generation forecasts.
- Adjusting half hourly maximum (or minimum) operational demand by incorporating the impacts of EV load and gas electrification.
- Determine maximum and minimum operational demands for each scenario by year and season.

(6) Presenting the maximum and minimum demand forecasts in terms of different POEs.

(7) Validate spatial demand forecasts and include post-modelling adjustments.

Each of these seven steps are explained in further detail below.

2.2.1. Assemble historical data – Step 1

The purpose of this section is to describe the historical data that is used in the forecasting process and how that data is obtained. We discuss each input requirement in turn.

Customer numbers and rooftop PV capacity

Customer numbers and growth rates are a major driver of future demand forecasts. Customer counts are used at the start of the modelling process for unitising the underlying demand and then in later steps to aggregate back the forecast unitised underlying demands.

Historical customer numbers are extracted by asset and customer type from the tariff database and spatial asset database to provide both a launch point for the forecasts and a trend on which to inform projections over the forecast period.

The key customer attributes which are relevant for the forecast include:

- The customer's classification, according to their tariff type.
- The feeder and zone substation supplying the customer. Network configurations can change over time, or on certain days (e.g., a customer can be supplied by Feeder X on one day, but Feeder Y on another). We use the most recent assignment between customers (National Meter Identifiers - NMIs) and feeders/zone substations when computing the corresponding past customer counts.
- Whether the customer has rooftop solar PV installed.

Once the customer data is extracted, the customers are grouped into the four classifications (residential, small business, medium business and large business), by the relevant network element (e.g., feeder or zone substation). This serves two purposes:

- Firstly, it establishes the trend of growth (or decline) of the various types of customers on the network element; and
- Secondly, it establishes the starting point for the forecast.

Operational demand

In order to obtain an informed understanding of the existing demand on the network, AusNet extracts demand data from SCADA sensors or a series of sensors at each connection point. Operational demand on various network elements (such as feeders and zone substations) is sourced from OSI-Pi, a data management platform that is responsible for data collection, storage, and management, which records SCADA sensor data. This data allows AusNet to drill into a fine degree of detail, whilst ensuring that the overall demand on the network element is captured and compared to the consumption recorded at the customer level.

After extraction, these data sets are cleansed of any abnormal readings, which can arise from data errors or temporary changes to network configuration. The resultant dataset is used as the basis for calculating the underlying demand, which is a key element of the forecast.

As historical data is used as the basis for forecasting maximum demand, it is imperative that the data excludes any values which are not reflective of the true underlying demand on the network element. This can come about, for example, when the network is temporarily reconfigured and one feeder is supplying load to customers which it would not ordinarily supply, or if embedded generators are offsetting the load on a feeder. Data errors with the telemetry or storing of meter data can also occur.

Data is cleansed in the following ways:

- **Planned temporary feeder reconfiguration**: Where a feeder has been planned to be reconfigured on a peak demand day by Network Planning Engineers, a correction should have already been applied to the data provided to the forecasting team. If this process has been missed, and a step-change in the SCADA data is observed during the selected peak demand day, a correction is applied to the data to remove the step change and create a synthesised data segment for the period of reconfiguration.
- **Unplanned operational feeder reconfigurations**: The network operations control centre may apply unplanned reconfigurations as part of the real time operation of the network. These reconfigurations will show up in the SCADA data as a step change in loading, with subsequent reversal. As above, a correction is applied to the data to remove the step change and create a synthesised data segment for the period of reconfiguration.
- **Large embedded generation**: The impact of large embedded generation (i.e., excluding generation that is installed behind the meter of load customers such as rooftop solar) is removed from the SCADA data trace. Synthesised data is created by adding the generation exports back onto the SCADA values. A register of embedded generation is used to undertake this process, and any generators that are missing from the register are identified through the process of comparing Advanced Metering Infrastructure (AMI) interval data to SCADA data.

Demand data for each network element being forecast is then "unitised" or averaged by dividing the total demand by the corresponding number of total customers. Once the data has been collected, aggregated, cleansed and unitised, it is used as a key input to AusNet's demand forecasting algorithm.

Embedded generation

Each embedded generator's generation data is extracted from the AMI interval database so that it can be added back to our operational demand to calculate the underlying demand.

Weather and solar variables

Weather data including temperature, wind speed and humidity relevant to each feeder and zone substation is extracted from DnA (AusNet's Data and Analytics Database), which is populated by postcode level time series from WeatherZone's satellite data.

The electricity generated by the rooftop PV panels depends on the panel capacity and also the solar variables (e.g., solar irradiance). To estimate the total PV generation, we use the same data source as weather variables.

Electrification trends

The electrification trends including transport electrification (the increasing number of EVs) and gas electrification (people switching from gas to electrical appliances) are two main drivers of electricity demand in the future. In order to forecast the impact of EVs, we require detailed information on EV penetration across AusNet's network. For gas electrification, we require the number of existing gas customers, preferably by location.

2.2.2. Forecast customer numbers and rooftop PV forecasts – Step 2

Customer numbers are a key driver of future electricity demand. Forecasting the growth in customer numbers is therefore an important input to AusNet's demand forecasting methodology. As explained in detail below, customer number forecasts are compiled with reference to both the historical trend in customer growth and the Victorian government's projections of structured private dwellings (**SPD**) in the Victoria in Future (**VIF**) planning publication.[2]

We discuss the forecasting approach for each customer segment below.

Residential customers

AusNet begins with an independent assessment of the projected growth in customers over the forecast horizon. The independent forecast used by AusNet is the Victorian Government's VIF publication. The VIF report provides five-yearly snapshot forecasts of population and dwelling numbers for Statistical Areas (SA) levels from 2 to 4 which are known as SA2, SA3 and SA4 regions. In particular, SA2 regions are comparable to postcodes in terms of area coverage. Since the vast majority of dwellings are connected to the electricity networks, these dwelling projections can be used as a starting point for the growth in residential electricity customers.

The SA2 level forecasts can be approximately mapped to zone substation regions and then up to terminal stations and down to feeders. For example, if Zone Substation A supplies all customers in region 1, 50% of customers in region 2 and 30% of customers in region 3, the projected growth in dwellings in those three regions can be adopted as a forecast for the growth in the number of residential customers connected to the zone substation, with the respective apportionment.

Below the zone substation level, feeders are also apportioned to the nearest region (or multiple regions) in order to derive forecast customer numbers. Particularly in growth corridors it needs to be assumed that feeders are extended over time with a consistent spread of geographic capture of each relevant region. As the feeder is extended over time, decisions by network planning engineers and design engineers on which customers are to be served by which feeder will take into account the forecasts, and the final feeder design will be captured in the following round of forecasting subsequent to the feeder extension project.

This approach is considered adequate where the VIF projections are shown to be accurate. Where there is evidence of VIF projections not reflecting actual growth, the VIF forecasts are adjusted to reflect AusNet's own view of the likely growth, based on recent trends and an assessment of local conditions (for instance, by utilising the knowledge of network planning engineers responsible for particular regions, or other sources such as specific connection inquiries and information made available by housing developers and industry bodies).

Non-residential customers

The above process for residential customers is not able to be replicated for non-residential customers, because the VIF report does not produce forecasts for the growth in the number of commercial or industrial (small, medium and large) businesses.

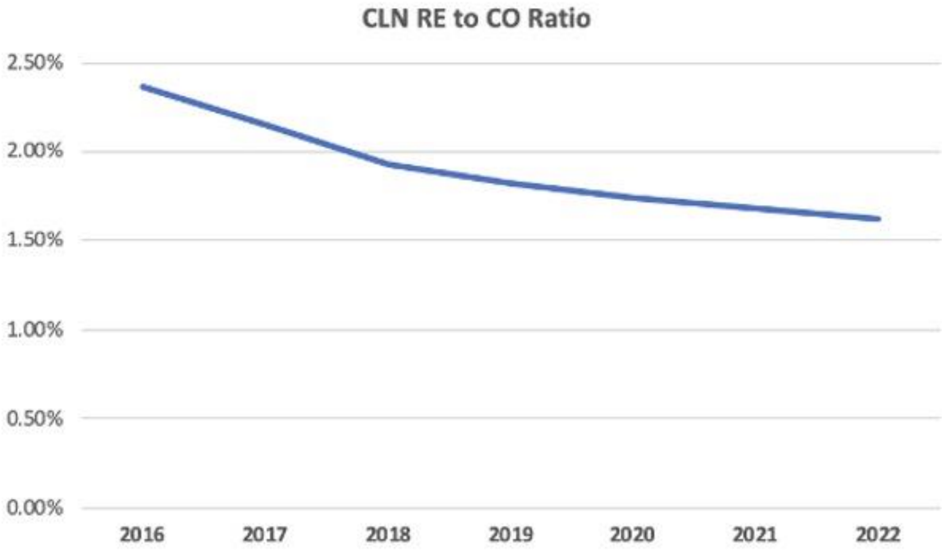
As such, estimating future non-residential customer movements is less certain than residential forecasts, as there is a lack of dependable sources of data on which to base the forecast, and they represent a much smaller proportion of customer numbers. The technique used by AusNet is to apply the relationship between the respective zone substation/feeder residential customer make-up, and the established commercial/industry customers on each feeder, and to linearly project this relationship forward based on the residential customer number forecast.

This approach reflects the typical situation where Local Government town planning parameters are consistent within the geographic coverage of a feeder, such as the land allowed for commercial customers within a residential development area. This results in a relatively stable and smooth outworking of the residential to commercial customer ratio.

As an example, **Error! Reference source not found.**Figure 3 shows the proportion of commercial to residential customers on the Clyde North zone substation. The growth rate in commercial customers in the past two years has dropped and is projected to continue to drop but at a slower rate.

[2] https://www.planning.vic.gov.au/land-use-and-population-research/victoria-in-future

Figure 3: Residential to commercial customer number ratio



A moderator to this is the available general trend data around commercial and industrial movements. Larger commercial and industrial customers require special consideration, and this tends to result in manual interventions. Where local knowledge is available for upcoming commercial and industrial customer connections or disconnections, manual adjustments may also be made.

Finalising the customer number forecasts

Once the customer forecasts for the various network elements have been developed, they are reviewed against historical growth rates as a top-down check that the forecast growth rates are consistent with the trend for the specific network element. For example, if a feeder is showing signs of an initial ramp up in growth rates, the forecast will be reviewed to ensure that the growth over the short to medium term (and into the long term) is reflective of this stage of the feeder's growth cycle.

A simple reconciliation process is also applied across the customer count forecasts of different assets in the hierarchy to ensure that the addition of customer counts in lower levels are equal to the corresponding upper level forecasts.

At the end of the process, AusNet has a customer number forecast for each network element that requires a forecast of demand. Since the demand on the network is a function of the number of customers and the demand per customer, the next step is to establish the existing demand per customer.

Rooftop PV capacity and count

We follow a similar approach to what we explained for customer count to forecast the rooftop PV count and capacity across different network segments and for different customer types. The only major difference is that, instead of the VIF data, we use the latest available AEMO's Electricity Statement of Opportunities (**ESOO**) assumptions to adjust our PV trends. For both residential and non-residential customers if the assumption is that PV penetration rate (PV capacity/Customer count) growth is going to increase/decrease, we apply the same to our distribution regions.

In particular, when obtaining PV count forecasts, we first compute the average PV growth rate in last 0-4 years (g1) and 4-8 years (g2). If g1 is negative and g2 is positive, then net growth rate is considered as the average of g1 and g2. Otherwise, g1 which is the most recent growth rate is considered as the net growth rate. The final growth rate which is used to compute the PV count forecasts is computed as the multiplication between the above net growth rate and the future PV growth rates found from the data published in ESOO. The PV count forecasts are obtained in monthly-basis. Thus, the computed growth rate is sequentially applied to the non-PV customer count corresponding with the previous month to obtain the newly joined PV customers in a given month. The final PV count forecast for a given month is considered as the addition of the newly joined PV customers in that month and the PV customer count of the corresponding previous month. A similar approach is also followed to obtain PV capacity forecasts. As explained, customer count is an input into the rooftop PV forecast process. Hence, the rooftop PV forecast is undertaken after completing the customer count forecast.

We do not consider solar batteries when producing our demand forecasts. Modelling battery load is resource-intensive due to its dependence on numerous variables. While we recognise that battery load can influence various aspects of network planning, its impact on maximum demand which is the specific focus of our analysis is currently considered negligible. Consequently, in developing the current methodology, we have prioritised factors with more material impacts and included the incorporation of batteries on a forecast model improvement plan.

2.2.3. Model underlying half-hourly demand – Step 3

This step estimates the relationship between underlying demand and the driver variables.

The total electricity consumed by customers is denoted as underlying demand. Historically the electricity load on the network (operational demand) and underlying demand were the same. However, over time, with the installation of rooftop solar PV panels and other embedded generators, a significant gap between these two is emerging. The historical underlying demand is calculated as follows:

Underlying Demand = Operational Demand + Embedded Generation + Rooftop PV Generation

Rooftop PV generation is calculated using PV capacity and solar data as follows:

Rooftop PV Generation = PV Capacity * PV Generation per Capacity Based on Solar Factors * Efficiency Factor

Where Efficiency Factor is considered as 0.85. PV generation per capacity is estimated using a Generalised Additive Model (GAM) which is trained considering 3 solar irradiance variables: Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI) as inputs.

The underlying demand varies over time depending on temperature, wind, humidity, time of day, day of week, public and school holidays. Each individual asset (network, terminal stations, zone substations and feeders) is separately modelled using its historical half hourly maximum (or minimum) demand data that belong to the last 14 years after removing outliers by conducting a manual inspection process. In the first step of the modelling process, we divide the calculated maximum (or minimum) half hourly underlying demands by the corresponding historical number of customers and obtain the unitised maximum (or minimum) underlying demand. In following, we continue working on the unitised maximum (or minimum) demand until the later steps where we multiply our estimates by future customer counts.

As part of our modelling process, we develop an approach that decomposes the unitised maximum (or minimum) half hourly underlying demand (y) into unitised average demand and an adjustment coefficient (y*), which reflects weather and calendar effects. Note that here the unitised average demand refers to the average of maximum (or minimum) demand values of each half hour interval that belong to a particular month, where the adjustment coefficient adjusts this average to estimate the unitised maximum (or minimum) half-hourly demand in that month, which fluctuates around the unitised average demand.

Denoting the unitised maximum (or minimum) underlying demand as y, we can write the following equation:

$$y = Average(y)\ X\ Adjustment\ Coefficient$$

Let i denotes month, t denotes half hourly interval and p denotes period (one of morning, afternoon and evening for each half hourly interval) and demand drivers (weather and calendar variables) that are used to estimate the adjustment coefficient as explained in the following sections. Then, unitised maximum (or minimum) underlying demand corresponding with a given half hour interval can be written as:

$$y_{t,p} = \bar{y}_i\ X\ y^*_{t,p}$$

These two components, average and adjustment coefficient, are modelled separately, in parallel, and then put together to obtain the modelled underlying demand.

Modelling average demand ($\bar{y}$)

After decomposing the unitised maximum (or minimum) underlying demand, we estimate the relationship between monthly average maximum and minimum demands and the weather variables. The average of maximum (or minimum) demands is estimated mainly by long term weather variables (e.g., amount of Cooling Degree Days (**CDDs**) and Heating Degree Days (**HDDs**)) and calendar variables (e.g., month of the year). In practice, we estimate the following linear regression model. This is to cover the impact of medium-term factors that impact the average maximum (or minimum) consumption but could be unrelated to half-hourly fluctuations in demand, which are addressed by the adjustment coefficient (y*) in the next step.

Unitised monthly average maximum (or minimum) demand = Constant + Amount of CDDs + Amount of HDDs + Month of the Year + Error

Note that here a CDD is considered as a day where the corresponding average temperature is greater than 18.5°C. For each CDD in a given month, the differences between the corresponding mean temperatures and 18.5°C are computed. The summation of these temperature differences is considered as the amount of CDDs of the corresponding month. An HDD is considered as a day where the corresponding average temperature is less than 18.5°C. The amount of HDDs corresponding with a particular month is also computed in a similar way to the amount of CDDs.

Modelling adjustment coefficient (y*)

This step is the semi-parametric model referred to in Monash Electricity Forecasting Model to estimate the relationship between half-hourly demand and the driver variables.

For each asset, we split the historical data into 3 subsets evenly: morning: 06:00-13:59; afternoon: 14:00-21:59; and evening: 22:00-05:59. The best parameters that should be used to model half hourly maximum (or minimum) underlying demand in each period are separately found. By selecting the model variables for three different periods of the day separately, the model is able to handle both peak and minimum demand forecasting without significantly increasing the computational burden. During variable selection, models are separately trained for each of the 3 periods where to train a model in a given period, all past half hour intervals of the corresponding asset which belong to that period are used. The following variables corresponding with the past half hour intervals are used as inputs when training these three separate models.

- Wind speed
- Humidity
- Current temperature as well as temperatures that were there 1 hour ago, 2 hours ago, 3 hours ago, 1 day ago, 2 days ago and 3 days ago
- Maximum, minimum and average temperatures of the day that the half hour period belongs to
- Day of week
- Whether the day is a school holiday or not
- Whether the day is a public holiday or not

A number of different formulas are then constructed using the above variables. A model is then fitted per each formula and computes cross-validation Mean Absolute Error (MAE) separately for morning, afternoon and evening periods. Then, for each period, the formula with the minimum MAE is selected as the best formula to model the maximum (or minimum) underlying demand in that period.

For a given asset, we then fit a separate GAM per each half hourly interval in the day (48 models) using the corresponding historical half hourly maximum (or minimum) underlying demands of that asset. The demand patterns change throughout the day and better estimates are obtained if each half-hourly period is treated separately. For each half hourly model, the best parameters corresponding with its period (morning, afternoon or evening) are used as inputs. In summary, the formula of a fitted model corresponding with a half-hourly interval will be similar to:

$$\log\left(\frac{y_{t,p}}{\bar{y}_i}\right) = F_p + Constant + Error$$

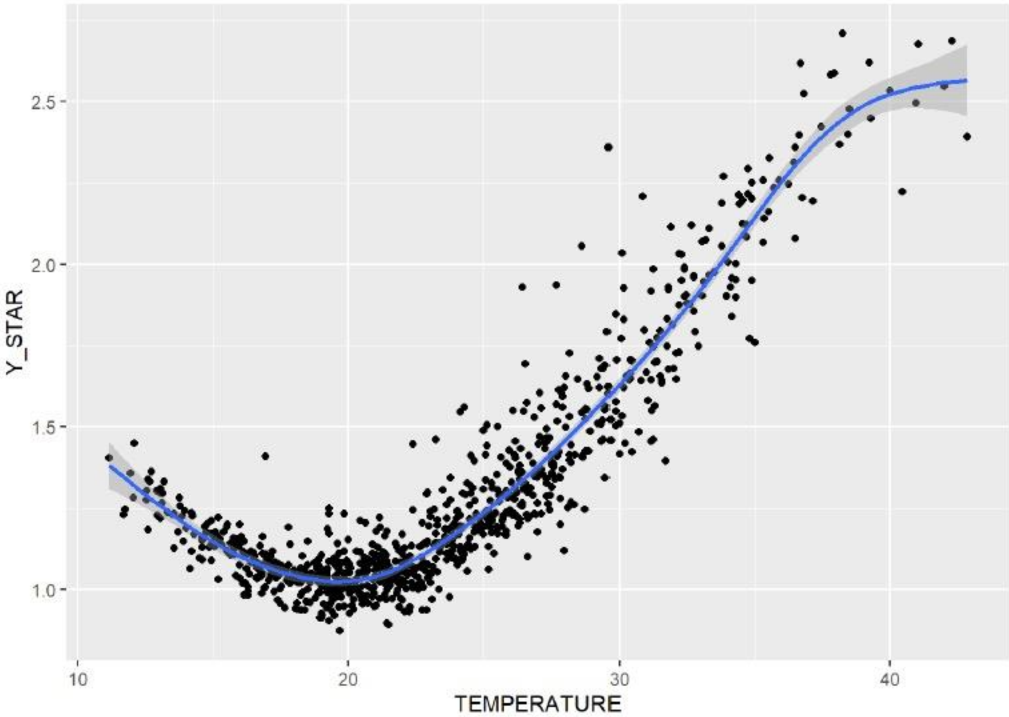
where F_p denotes the best set of input variables corresponding with the period of a particular half hour interval.

Let output of the half hourly model is $O_{t,p}$. Then the adjustment coefficient, $y^*_{t,p}$, can be written as: $y^*_{t,p} = \exp\,(O_{t,p})$

During forecasting, the explanatory (input) variables used in estimating the Adjustment Coefficient can be used to convert our forecast of average maximum (or minimum) demand into half-hourly data using a bootstrapping process. It also enables us to simulate the range of values that demand could take based on historical weather and explanatory variables.

The figure below shows an example relationship between temperature and fitted values for daily maximum demand and maximum temperature. One feature of the data is that demand increases decline as temperatures exceed 40 degrees, as demand reaches saturation.

Figure 4: Relationship between the modelled Adjustment Coefficient and temperature



2.2.4. Forecasting the impact of electrification – Step 4

This step is not explicitly identified in Figure 1, although Monash University recognise that adjustments may be required for electrification. The impact of electrification as a result of increased penetration from EVs and the impact of lower gas customers is modelled separately, as explained below.

Forecasting load from EVs

The forecast number of EVs for each network asset (Feeders, Zone Substations, Transmission Connection Point and the Distribution Network) is produced for each customer type (Residential, Small and Medium Business, and Large Business).

From a forecasting perspective, the EV penetration rate is a key input to forecasting EV numbers. The EV penetration ratio for different geographical areas are estimated by comparing customer numbers with EV sales data obtained from the Victorian Department of Transport and Planning, disaggregated to postcodes to determine the level of EV penetration across our network.

Future penetration ratios are forecast using data from AEMO's ESOO, which contains projected penetration ratios for Victoria. Our existing EV penetration ratios are increased at the same growth rate as AEMO's forecast. For instance, if AEMO's Victorian EV penetration rate is growing from 2% to 10% over 5 years, which is 5 times growth; we grow asset A with initial EV penetration rate of 3% to 15% and asset B with initial EV penetration rate of 1% to 5%. The resulting base forecast penetration ratios across our network are adjusted by applying linear adjustments as required so that:

- At an aggregated level, we adopt to the same EV forecast growth rate as estimated by AEMO for Victoria.
- Our forecasts reconcile to AEMO's base assumption that the EV penetration will be 100% by 2050.

To produce an EV load forecast, the number of EVs are combined with EV charging profiles, obtained from AEMO's ESOO EV workbook, to estimate the EV load for each half-hour interval. AEMO's EV workbook provides different charging profiles for vehicle types and charging method, i.e., convenience charging, daytime charging, highway fast charging; and nighttime charging.

Our forecast method produces the following forecast EV load:

*EV Load (kW) = Number of EVs * Share (%) in each charging profile * Load (kW) per EV in each charging profile*

The resulting EV Load (kw) for each half-hour interval for each year is adopted in the demand forecasting process.

Forecasting gas electrification load

An important driver for increased electricity demand over time is the impact of gas electrification, in response to efficiency considerations and government policy objectives. To estimate the load that results from gas electrification, we use the below formula:

Gas electrification load (kw) =

(1) Gas penetration rate (%, percentage of customers with gas connection) multiplied by

(2) Electrification rate (%, percentage of customers switching from gas to electricity) multiplied by

(3) Impact on electricity consumption (%, percentage for different seasons and for different times per day) multiplied by

(4) Base electricity consumption (kw), for different seasons and for different times per day).

Each input to the formula is obtained as follows.

- **Gas penetration rate**: We obtain gas penetration rates using AGIG maps at postcode level. In metro regions, we have more than 90% gas penetration, with some regional areas varying from 0% to 80%. We also consider the impact of the Victorian Government's policy on new gas connections. From 1 January 2024, new gas connections for new dwellings, apartment buildings, and residential subdivisions requiring planning permits are being phased out.
- **Electrification rate**: We use latest available GSOO gas-customer number forecasts to estimate the electrification rates, which reflects the decline in the expected number of gas customers.
- **Impact on electricity consumption**: We use an internal study of the customers in our network to estimate the impact on electricity consumption. We run a regression model on the data to obtain an estimation of the gas-electrification impact for each season half-hour pair.
- **Base electricity consumption**: To estimate the gas electrification load, it is assumed that gas customers switching to electricity customers have the same underlying demand as existing electricity customers.

The final output of this step is a half-hourly load per season (kw), which will be used in the next step, simulating the future.

2.2.5. Simulating the future – Step 5

To simulate the future conditions that explain the electricity demand, we use a bootstrap method. This process involves using the historical observations of inter-related temperature, solar, and season variables and randomly select 200 sets of explanatory variables. In doing so, we use a block bootstrap with variable blocks method that, first, increase the number of existing scenarios to choose from and, second, maintains the statistical characteristics of the original explanatory variables. The process of bootstrapping was inspired by Monash's methodology document, however, the block lengths that we use are different from the lengths mentioned in Monash paper. For temperature, calendar and solar variables simulation, four varied length blocks are used where blocks 1, 2 3 and 4 are respectively spanned across the periods, November to March, April to May, June to August and September to October. In each simulation iteration, one historical demand year is randomly selected per each block. This means for the above defined 4 groups of months, per each half hour interval, explanatory variables such as temperature, calendar and solar variables are randomly selected from different historical demand years. In this way, we create 200 different sets of explanatory variables, at half hour intervals, which are then used to estimate half hourly maximum (or minimum) underlying demand using the model developed in the previous step.

Residuals are simulated using one-month blocks. In each simulation iteration, one demand year is randomly selected per each of the 12 months. This means for each month, we have half hourly data coming from different historical demand years. Operational demand predictions are obtained for this randomly selected dataset using the model developed in the previous step, and then the corresponding residuals are computed. Centered residuals are then computed per each half hour interval which is the subtraction between the original residual and the median of residuals in the corresponding month. Then, using an Autoregressive Integrated Moving Average (ARIMA) model of order (1,0,0), a set of residual medians are simulated, and these simulated residual medians are added back to the centered residuals, which are considered as the final simulated residuals. These final simulated residuals are added to the corresponding half hourly maximum (or minimum) underlying demand predictions.

The solar variables that are randomly selected along with the other explanatory variables are used in the next step to calculate the operational demand from the estimated underlying demand.

Note that this step provides unitised maximum (or minimum) underlying demand. In order to put the underlying demand in practice, we need to multiply it by the number of customers for each year and then calculate the half hourly maximum (or minimum) operational demand, as follows. Base Operational Demand (MW) =

Underlying Demand ((unitised underlying demand + simulated residuals)* customer count) minus

PV Generation (PV generation per capacity from bootstrapped solar data * PVcapacity * efficiency factor)

As a complementary step, for each half-hourly simulated maximum (or minimum) operational demand, we add the matching EV load and gas-electrification load, to account for these trends in the forecast. So, the above formula will be completed as:

Operational Demand (MW) =

Underlying Demand ((unitised underlying demand + simulated residuals) * customer count) minus

PV Generation (PV generation per capacity from bootstrapped solar data * PV capacity * efficiency factor) plus

Electrification Load (calculated in the previous step, for each half-hour for each season converted into MW)

These half hourly maximum and minimum operational demands are then used to extract seasonal maximum and minimum operational demands.

The output of the simulation step is 200 maximum and 200 minimum values, for each season for each year.

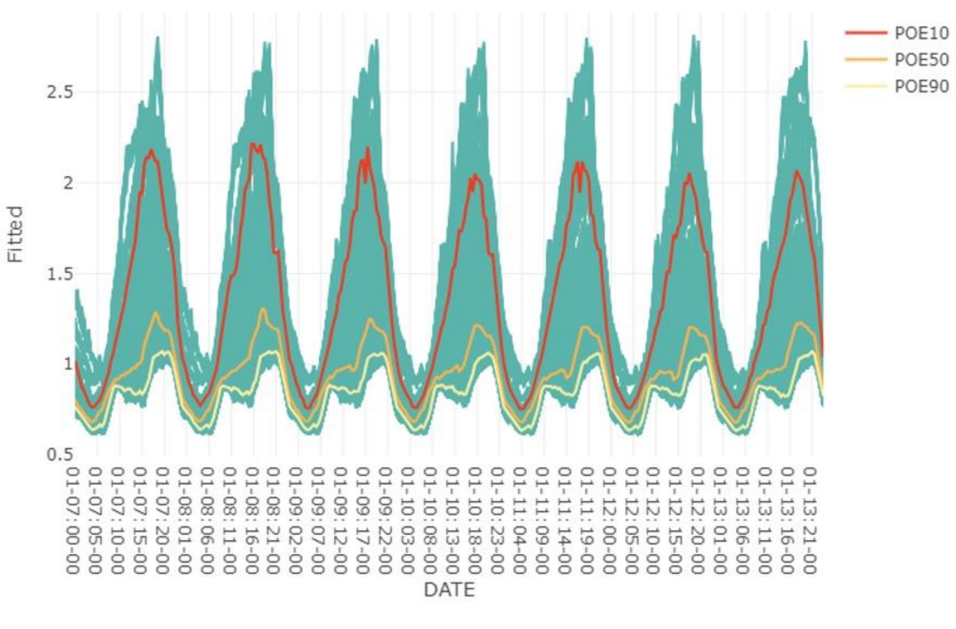
2.2.6. Presenting maximum and minimum demand forecasts as POEs – Step 6

The simulated maximum and minimum demands are used to extract different probabilities of exceedance (POE). POE10, POE50 and POE90 are the three most used for network planning purposes. Starting with the maximum demand POE10 is equal to the 90th percentile of the simulated series, where only 10% of modelled outcomes exceed that value. In an analogous way, POE50 and POE90 are equal to 50th and 10th percentiles of the simulated maximum demand, respectively.

In the case of the minimum demand, however, there is a slight difference in definition. POE here means the probability of operational demand being lower than a specific value. Therefore, POE10, POE50 and POE90 are exactly 10th, 50th and 90th percentiles of the simulated minimum demands, respectively. Whilst on face value this is a contradiction in terms (probability of exceedance reflects the chance of demand being lower than the value), this approach retains the intuition that a POE10 scenario is describing an outcome with a 10% chance of being lower than the forecast (e.g., even lower demand than a minimum demand forecast predicts).

The figure below shows an example of the 200 simulated demand traces for the second week of January, and the extracted POEs for each 30-minute interval. Just note that in practice we extract POEs at annual level, not 30-min intervals.

Figure 5: Simulated demand traces and POEs for each 30-min interval



2.2.7. Validate spatial demand forecasts and post modelling adjustments – Step 7

Regional network planning engineers in conjunction with the sub-transmission planning engineer validate the relevant forecasts. Validation involves magnitude checks and trend line checks informed by knowledge of the loadings and network configuration changes recently completed and pending. Adjustments are undertaken to improve the accuracy of the forecast by addressing factors such as:

- large customers (above 1MVA) that are known to have connected recently or will connect in the near term (i.e. block loads);
- impact of known network projects that have recently been undertaken or are in train such as feeder reconfigurations;

- Inconsistencies in AMI data that lead to offsets in the final forecast.

For block loads, we apply a post model adjustment by assessing actual connection requests for loads over 1MVA that are well progressed and are expected to proceed. When computing a block load from a project that is corresponding with a particular asset for a given year, we consider several facts such as the probability of the project commencing, whether the load is coming from a residential or commercial/industrial customer, expected start year of the project and load uptake in years. We then identify whether those block loads are already captured in the demand forecasts by comparing the demand forecast growth with the loads coming from the list of requested connection projects computed based on the above factors. If the requested block loads are higher than the forecast demand growth, we adjust the demand forecasts for those block loads, only incorporating the difference between the block load and the demand forecast growth from the model (to ensure there is no overlap). Furthermore, the loads coming from the commercial fast chargers are omitted as block loads to avoid double counting them with the EV charging load forecasts.

2.2.8. Bottom up and top down reconciliation

By using the same NMI data in a bottom-up build of forecasts for both feeders and zone substations, the forecasts are inherently consistent between feeders and zone substations. As such, a rigorous reconciliation process is not required.

The purpose of reconciliation is verifying that the sum of the feeder maximum demand forecasts for a particular zone substation align to the forecast maximum demand of the zone substation. A difference between the two values is expected due to the fact that feeder peak demand times will never perfectly coincide with one another. Visualising the degree of coincidence is sufficient to verify the alignment of the forecasts between feeders and zone substations.

2.3. Planned methodology improvements

AusNet adopts a continuous improvement model to the demand forecasting methodology in order to progressively improve the forecasts and respond to changing parameters such as the impact of new technologies and the availability of new data sets.

We will continue to review our forecasting methodology to improve our approach in light of experience, including by reviewing the accuracy of our forecasts. We will also continue to improve our forecasting process where possible, including by increasing automation and taking steps to improve the quality of our data inputs.

2.4. Demand forecast uses

The preparation of the annual maximum and minimum demand forecasts is the starting point for our distribution planning process and expenditure plans. The demand forecasts are employed for a wide variety of tasks, including:

- Network constraint identification
- Generator connections
- Assessing network hosting capacity
- System strength studies
- Forecasting capital expenditure requirements
- Tariff setting
- Operational network planning
- Annual reporting
- Voltage management
- Distribution loss factors
- Transmission connection planning.

The demand forecasts prepared by AusNet and other distributors are also used by AEMO in its roles as market operator and national planner.

AusNet Services

Level 31
2 Southbank Boulevard
Southbank VIC 3006
T +613 9695 6000
F +613 9695 6666
Locked Bag 14051 Melbourne City Mail Centre Melbourne VIC 8001
www.AusNetservices.com.au

Follow us on

@AusNetServices
@AusNetServices
@AusNet.Services.Energy