Transmission Connection Point Forecasting Methodology

July 2021

Maximum and minimum demand

Important notice

PURPOSE

AEMO has prepared this document to provide information about its transmission connection point forecasting methodology for annual maximum and minimum demand.

This publication has been prepared by AEMO using information available at 30 June 2021. Information made available after this date may have been included in this publication where practical.

DISCLAIMER

The data presented in figures or tables in this report are for demonstration purposes only, and do not represent the data used in forecasts.

This document or the information in it may be subsequently updated or amended. This document does not constitute legal or business advice and should not be relied on as a substitute for obtaining detailed advice about the Wholesale Electricity Market Rules or any other applicable laws, procedures or policies. AEMO has made every effort to ensure the quality of the information in this document but cannot guarantee its accuracy or completeness.

Accordingly, to the maximum extent permitted by law, AEMO and its officers, employees and consultants involved in the preparation of this document:

- make no representation or warranty, express or implied, as to the currency, accuracy, reliability or completeness of the information in this document; and
- are not liable (whether by reason of negligence or otherwise) for any statements or representations in this document, or any omissions from it, or for any use or reliance on the information in it.

VERSION CONTROL

Version	Release date	Changes
3.0	30/7/2021	Update to methodology to allow forecast of minimum demands. Changes also affected how maximum demand is forecast, allowing the forecast to reflect change in timing of maximum and minimum within the forecast horizon in response to distributed photovoltaics and emerging technologies such as electrical vehicles. This version documents the updated methodology.
2.0	11/10/2019	General changes to text to reflect new AEMO publication names. Updates made to treatment of historical and forecast installed capacity of photovoltaics reflecting improved methods for including their impacts. Updates made to reconciliation reflecting reconciliation of non-coincident forecasts to the regional trend.
1.0	29/7/2016	The original connection point forecasting methodology was produced by ACIL Allen, and published in 2013. A number of subsequent improvements have been made to the methodology, and have been documented in each connection point forecasting report. This version is the first version of a consolidated AEMO Connection Point Forecasting Methodology, including elements of the original ACIL Allen methodology and subsequent improvements.

Contents

Tables

Figures

1. Introduction

In its role as independent market and system operator, AEMO develops maximum and minimum forecasts for each transmission connection point (TCP) to provide a higher level of detail about local trends in demand, relative to the forecasts prepared for AEMO's Electricity Statement of Opportunities (ESOO) for the National Electricity Market (NEM).

The recent trend in uptake of distributed photovoltaic installations, embedded wind and solar generation as well as emerging technologies such as electrical vehicles is changing the load profile observed at TCPs. This is expected to affect the magnitude of minimum demand across a year as well as its timing, including leading many TCPs that have historically recorded minimums in the early morning to record minimums during the afternoon (and for these potentially to occur in a different season).

Together with the regional level maximum and minimum demand forecasts published in the ESOO, the TCP forecasts provide an independent and transparent view of electricity demand in the NEM, supporting efficient network investment and policy decisions for the long-term benefit of consumers. TCP forecasts align with the "Central" scenario in ESOO. The Central scenario reflects the current transition of the energy industry under current policy settings and technology trajectories, where the transition from fossil fuels to renewable generation is generally led by market forces and supported by current federal and state government policies[1].

1.1 Key definitions

- **Connection point**

AEMO's definition of a transmission connection point is the physical point at which the assets owned by a transmission network service provider (TNSP) meet the assets owned by a distribution network service provider (DNSP), as illustrated in Figure 1, or transmission connected industrial loads.

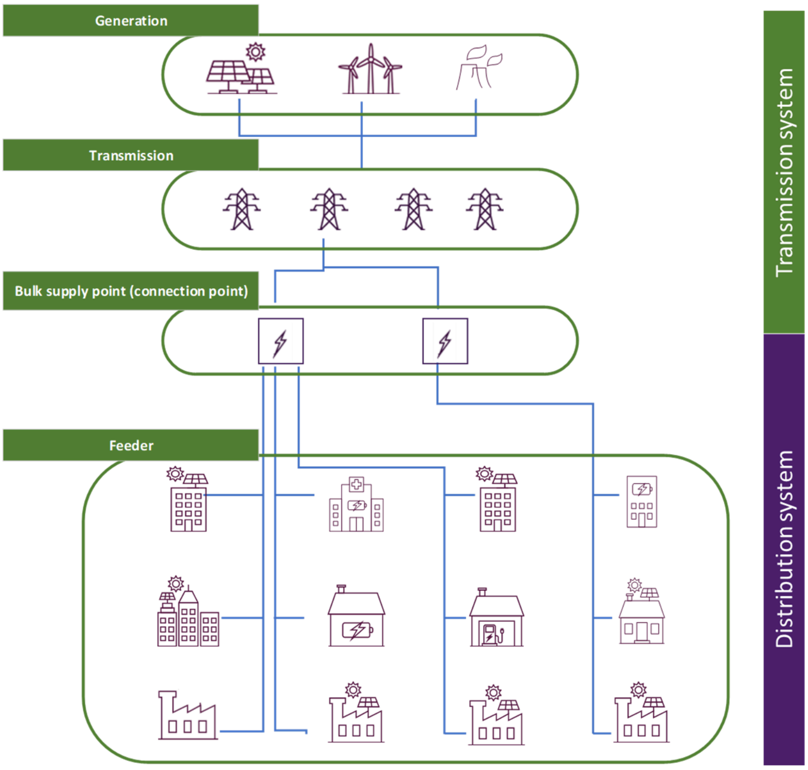
These may also be known as bulk supply points (BSPs), terminal stations, or exit points, and in the NEM's market metering and settlements processes they are called transmission node identities (TNIs).[2]

TNIs may be connected to one another via distribution networks. In situations where the interconnectivity is extensive, AEMO develops a forecast for the aggregated load instead of the individual TNIs. For example, the eastern suburbs of Sydney are connected to the transmission network at multiple TNIs. This ensures there are alternative feeder lines to supply load. The Sydney network is heavily meshed and it is impractical to produce forecasts for each individual TNI while the customer loads might be transferred often from one terminal station to another within the meshed network, so AEMO produces demand forecasts for the aggregated load.

[1] AEMO, 2020 Inputs, Assumptions and Scenarios Report, at https://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Planning-and-forecasting/Inputs-Assumptions-and-Methodologies.

[2] For a complete list of TNIs, refer to the most recent list of regional boundaries and Marginal Loss Factors, at http://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Security-and-reliability/Loss-factor-and-regional-boundaries.

Figure 1 Transmission – Distribution Connection Point diagram



- **Operational demand**

The reported demand and energy on the network depend on where they are being measured. AEMO has elected to forecast operational demand, which refers to the electricity used by residential, commercial, and large industrial consumers, as supplied by operational generators. The detailed composition of operational demand is defined in AEMO's Electricity Market Management Systems (EMMS) Data Model[3], and some adjustments are required to account for the change of focus from regional to connection point level.

At the connection point resolution, operational demand includes:

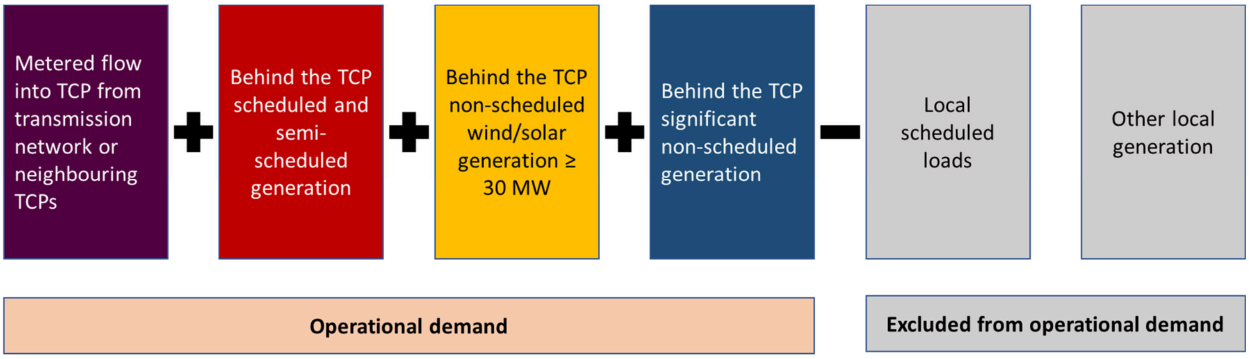
- Metered flow from the transmission network into the connection point, plus
- Metered flow from neighbouring connection points within the same region (if any) into the connection point (negative value will mean a flow out of the connection point), plus
- Generation from operational generators embedded in the connection point (if any), including local scheduled generation, semi-scheduled generation, non-scheduled wind/solar generation of aggregate capacity ≥ 30 MW, and generators that are exempt from registration under market rules, minus
- Local scheduled loads (if any).

When this operational demand is aggregated over all connection points in a region, it is equivalent to the "as consumed" operational demand at regional scale, which excludes transmission losses and auxiliary loads associated with transmission-connected generators.

Figure 2 illustrates the definition of operational demand at connection points, expressed in terms of the supply sources that meet operational demand and others that are excluded.

[3] AEMO, Demand Terms in EMMS Data Model, 2021, at https://www.aemo.com.au/-/media/Files/Electricity/NEM/Security_and_Reliability/Dispatch/Policy_and_Process/Demand-terms-in-EMMS-Data-Model.pdf

Figure 2 Operational demand definition at connection points



- **Probability of Exceedance (POE):** POE is the likelihood a maximum or minimum demand forecast will be met or exceeded. A 10% POE maximum demand forecast, for example, is expected to be exceeded, on average, one year in 10, while a 90% POE maximum demand forecast is expected to be exceeded nine years in 10.
- **Rooftop PV:** Rooftop PV is defined as a system comprising one or more photovoltaic (PV) panels, installed on a residential or commercial building rooftop to convert sunlight into electricity. The capacity of these systems is less than 100 kilowatts (kW).
- **PV Non-Scheduled Generators (PVNSG):** PVNSG is defined as PV systems larger than 100 kW but smaller than 30 MW non-scheduled generators.
- **Distributed PV:** Term for rooftop PV and PVNSG installations combined.
- **Embedded:** Generator or load which is connected to the network considered; in this case a distribution network or transmission connected industrial load associated with a TCP. Sometimes also referred to as "behind-the-TCP" generation/load.
- **Energy Storage Systems (ESS):** ESS are defined as small distributed battery storage for residential and commercial consumers.
- **Electric Vehicle (EV):** EVs are residential and business battery powered vehicles, ranging from small residential vehicles such as motor bikes or cars to large commercial trucks

1.2 Purpose and application

AEMO publishes TCP forecasts as part of its National Transmission Planner (NTP) functions and in accordance with clause 5.22.18(b) of the National Electricity Rules.

TCP forecasts provide a finer level of spatial detail than regional forecasts, such as those published in the Integrated System Plan (ISP) and ESOO. TCP forecasts are intended to support DNSPs in making efficient network investments and to guide regulators in their review of related proposals.

AEMO does not currently have information that allows projections of load growth below the spatial resolution of TCPs. AEMO acknowledges that application of TCP forecasts to zone substations and feeders requires detailed information about the distribution networks to be accurate. For example, it is inappropriate to apply AEMO's TCP forecast to lower level networks by evenly assigning demand across zone substations without sufficient information to support this as a valid approximation.

2. Methodology

2.1 Overview of the methodology

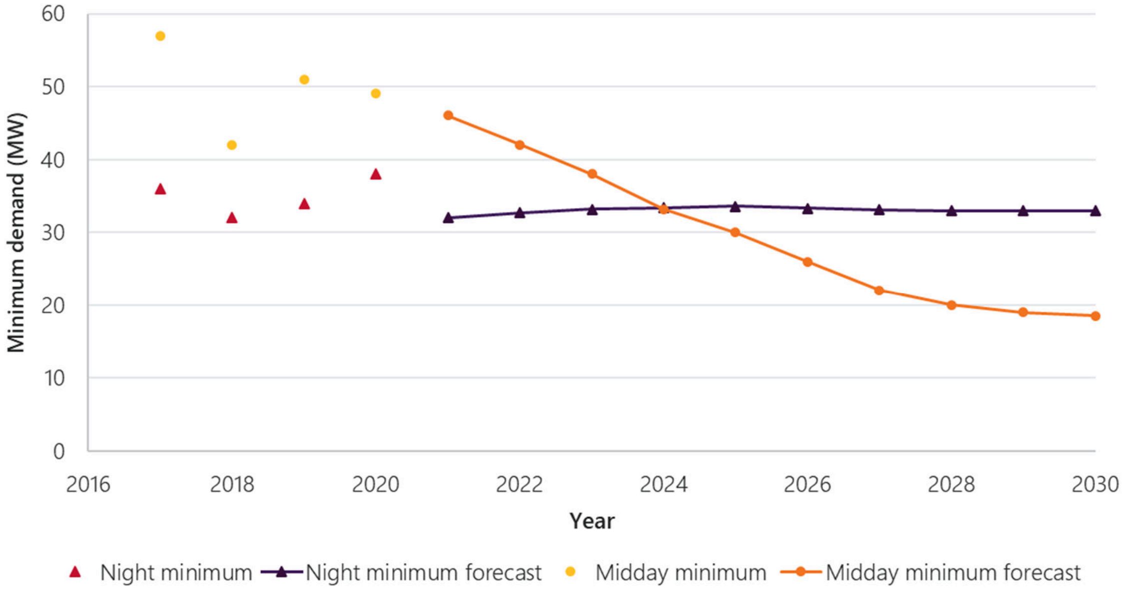
The AEMO connection point forecast methodology is designed to achieve the following objectives:

- To be nationally[4] applicable and consistent with the regional forecast.
- To produce both minimum and maximum demand.
- To model the potential transition of the timing of minimum demand in a computationally efficient way.
- To model the impact of behind-the-meter technologies on demand at different parts of the day.
- To explicitly model spatial demand drivers and reduce the reliance on reconciliation.

To meet the above objectives, the methodology relies on three broad steps:

1. Capturing the connection point specific relations between demand and weather.
2. Determining the weather-corrected starting point for maximum and minimum operational demands for a number of "defined periods" at each TCP.
   - Each "defined period" reflects a different combination of a) period of day (such as morning, afternoon) and b) season (summer, winter, shoulder). Each half-hourly period of a year is to be categorized as being one of a small number of pre-determined "defined periods", for example, afternoon-summer.
   - Multiple "defined periods" are projected forward in the TCP modelling to allow for the potential change in the timing of minimum demands (for example, early morning to mid-afternoon) to be explicitly assessed to ascertain when/if this cross-over occurs. Figure 3 shows an example of minimum demand transitioning from consistently occurring at night to the middle of the day due to an increase in PV generation during sunshine hours.

Figure 3 Example of minimum demand transition time from night to day

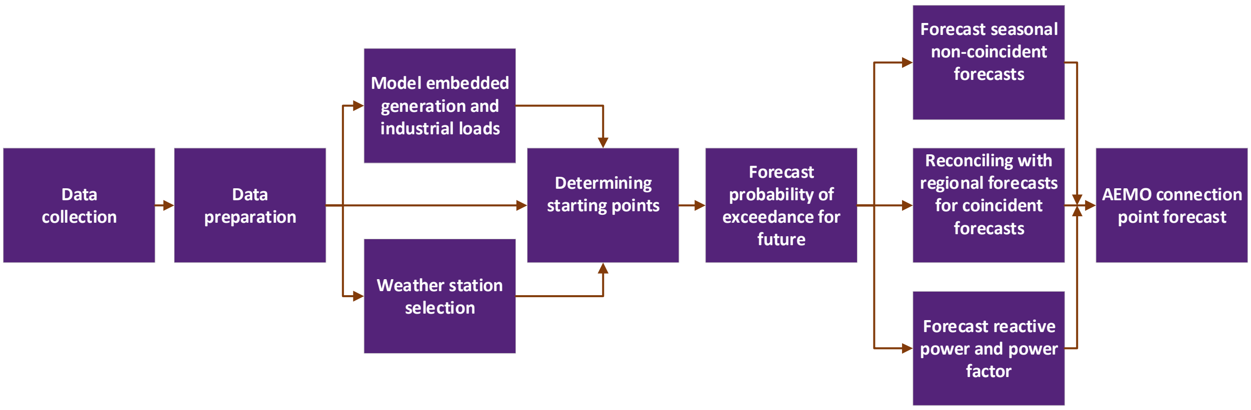


4 Currently, AEMO only forecast the demand at the transmission connection points in the NEM.

3. Forecasting how each of those starting points changes over the forecast horizon (10 years), given certain factors that have a material impact upon the maximum and minimum demands. These factors include, but are not limited to:
   – Growth in number of customers at each TCP.
   – The forecast effect of behind-the-meter technologies such as battery storage, electric vehicles (EVs), PV, and other new embedded generators.
   – Probable loads not otherwise captured by customer growth or behind-the-meter technologies (block loads).
   – Decommissioning of existing embedded generators.

The rest of this report provides more details for the main blocks of the AEMO connection point forecast methodology, as shown in Figure 4 below and summarised as follows:

Figure 4 Methodology blocks

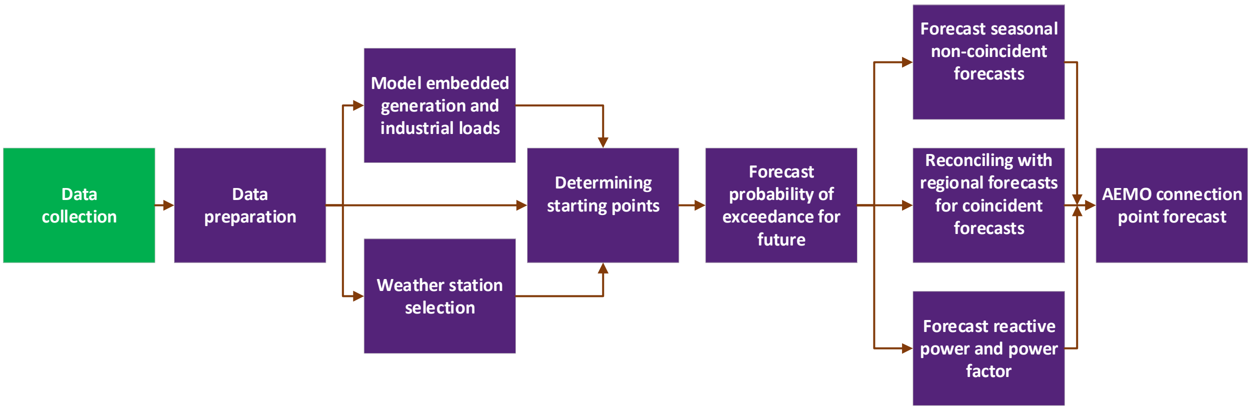


- **Data collection** – collecting the various types of data required to produce connection point forecasts.
- **Data preparation** – involves data cleansing, separating out the impact of large industrial loads and behind-the-meter generation. Removing the impact of the embedded generation and industrial load from the TNI load trace helps with:
  – Getting a better statistical fit for residential/commercial underlying demand and weather models.
  – Growing each component (industrial loads, PV, EV, ESS, etc.) of operational demand individually depending on its specific growth drivers[5].
  – Explicitly accounting for the contribution of each technology in the final operational demand forecast.
- **Modelling embedded generation and industrial loads** – developing models to estimate the demand or generation of the loads and technologies which impact demand at a TCP level for each half-hour of a year.
- **Weather station selection** – selecting the most appropriate weather stations to the connection point being examined.
- **Determining starting point probability of exceedance** – determining the weather-corrected starting point for operational maximum and minimum demands for the defined periods being modelled at each TCP, as it was summarised in steps 1 and 2 above.

[5] Although the impact of behind-the-meter technologies is individually mapped out in the operational demand forecasts, the projected impact accounts for the potential correlation between those technologies; for example, the forecasts of residential battery energy storage are in line with the expected growth of rooftop PV.

- **Forecast probability of exceedance for future** – growing the starting point POEs based on demand growth drivers as explained in step 3 above.
- **Forecasting seasonal non-coincident forecasts** – forecasting seasonal operational maximum and minimum demands at each TCP independent of regional demand.
- **Reconciling with regional forecasts for coincident forecasts** – forecasting the operational demand of TCPs at the time of regional minimum and maximum demand forecasts and reconciling the forecasts to ESOO regional forecasts.
- **Forecasting the reactive power and power factor** – forecasting future reactive power and power factor at TCPs.

2.2 Data collection



Data used in the connection point forecasting process can be grouped into three categories:

- **Measured data** – including electricity demand and hourly temperatures.
- **Modelled data** – obtained from modelling processes (undertaken by AEMO or external parties).
- **Descriptive data** – provided to AEMO by DNSPs via an annual data collection exercise before the start of the forecasting process.

The following tables outline the key sources of data used to develop the connection point forecasts and the methodology step associated with the data.

For further details on the application of the data, see the relevant section of the methodology document.

Table 1 Measured data

Category	Description	Source	Used in
TCP demand	Half-hourly MW demand, metered at the transmission side of each TNI mapped to the corresponding TCP	AEMO	Data preparation
National Metering Identifier (NMI) demand	Half-hourly MW demand and generation of large industrial loads and embedded generators	AEMO	Data preparation
Reactive power	MVAr data, generally from SCADA feeds from transmission network transformers	AEMO	Reactive power
Historical installed capacity of rooftop PV*	Capacity and installation date of rooftop PV systems, by postcode, TCP or TNI	Clean Energy Regulator (CER) Distributed Energy Resources (DER) Register DNSPs	Data preparation

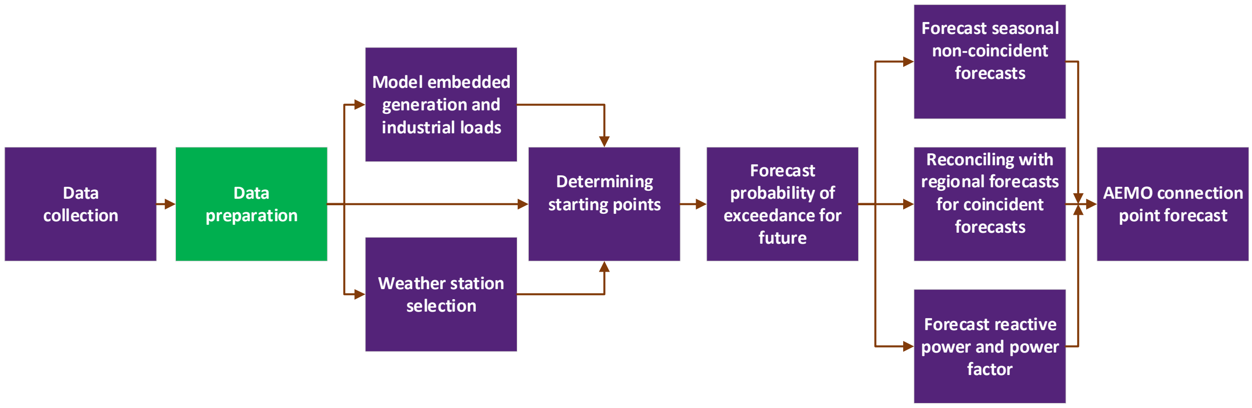
Category	Description	Source	Used in
Historical installed capacity of PV non-scheduled generation (PVNSG)*	Capacity and installation date of PVNSG systems with associated postcode, TCP or TNI	Australian Photovoltaic Institute (APVI) DER register DNSPs	Data preparation
Number of meters	Number of NMIs registered at each TCP	AEMO	Data preparation
Weather	Half-hourly temperatures at selected weather stations	Bureau of Meteorology (BOM)	Data preparation, determining starting points

* The choice between rooftop PV and PVNSG data sources varies between NEM regions where the analysis indicates the more appropriate source. In the future, AEMO may use other data sources, either as supplementary sources or for validation if the analysis suggests they would provide quality information.

Table 2 Modelled data

Category	Description	Source	Step
TCP boundaries	TCP boundaries to estimate spatial demand drivers	Australian Renewable Energy Mapping Infrastructure Project (AREMI)	Data preparation
Local Government Areas (LGAs) boundaries	LGA boundaries to estimate population of TCPs	Australian Bureau of Statistics	Population estimation
DNSP Forecast	Maximum demand forecast, grouped by connection point	DNSPs	Verification of results
Forecast of Rooftop PV Installed Capacity	Forecast of installed capacity of rooftop PV by postcode	External consultants	Embedded technologies adjustment
Forecast of PVNSG Installed Capacity	Forecast of installed capacity of PVNSG by postcode	External consultants	Embedded technologies adjustment
Forecast of installed capacity of ESS	Forecast of installed capacity of ESS by postcode	External consultants	Embedded technologies adjustment
Forecast of EV numbers	Forecast of EV numbers by NEM region	External consultants	Embedded technologies adjustment
Population	Population projections by sub-region	State governments	Growing starting point forecasts
PV Traces	Normalized PV output at connection point locations	External consultants	Embedded technologies adjustment
Profiles of ESS and EV operation	Forecast of ESS and EV operation profile by NEM region	External consultants	Embedded technologies adjustment

2.3 Data preparation



Data preparation includes three main stages, with a few sub-steps as outlined below:

1. Embedded technologies adjustment of the connection point load trace by removing the effect of:
   - embedded generation.
   - industrial loads.
   - rooftop PV and PVNSG.
2. Anomaly detection and correction.
3. Making historical adjustments to the data to account for block loads, transfers, and outages.

2.3.1 Developing underlying residential/commercial historical demand trace

Historical load data at each connection point is used as the starting point for determining underlying demand less large industrial load (or underlying residential/commercial demand), which is the basis of AEMO's demand-weather modelling. This is generally collected from AEMO's databases, including TNI, National Meter Identifier (NMI), and/or Supervisory Control and Data Acquisition (SCADA) data.

Once historical load data has been collected for each connection point, AEMO undertakes the following steps to remove the effect of embedded generators, large industrial loads, and rooftop PV and PVNSG from the traces.

Embedded large loads

Embedded Large Loads (ELLs) are industrial loads that are connected to the distribution network and can materially increase the load as measured at a connection point, given the relative size of a load to the transmission connection point's demand.

To determine the load at the connection point without the effect of the ELLs, the load is removed from the measured connection point load to improve the statistical fit. Each ELL is modelled separately and added back to the connection point forecast.

Historical distributed rooftop PV and PVNSG generation

Historical rooftop PV/PVNSG generation at half-hourly intervals is estimated using the historical installed capacity of rooftop PV/PVNSG systems at each connection point and normalised PV generation traces.

The historical installed capacity of rooftop PV is produced from the following sources:

- Data provided by the relevant DNSP.
- Data from the Clean Energy Regulator (CER).
- The Distributed Energy Resources (DER) Register.

The historical installed capacity of PVNSG is produced from the three following sources:

- Data provided by the relevant DNSP.
- Data from Australian Photovoltaic Institute (APVI).
- The DER Register.

Using these data sources, the amount of installed rooftop PV/PVNSG at each connection point is determined on a monthly basis.

Once the installed rooftop PV/PVNSG capacity for each connection point has been determined, a half-hourly rooftop PV/PVNSG generation trace is developed where historical data is required.

The half-hourly rooftop PV/PVNSG generation trace is developed as follows:

1. Preparing a half-hourly normalised generation trace[6], using the estimated traces provided by an independent consultant for each connection point location.
2. Multiplying the normalised generation trace by the installed capacity for the relevant month and connection point.

To determine the load at the connection point without the effect of the rooftop PV/PVNSG generation, the generation is added back to the measured connection point load, as demonstrated in Figure 5.

Embedded generators

Embedded generators are generators that are connected behind wholesale connection points and reduce the load as measured at a connection point. Half-hourly data for embedded generators is sourced from AEMO's databases.

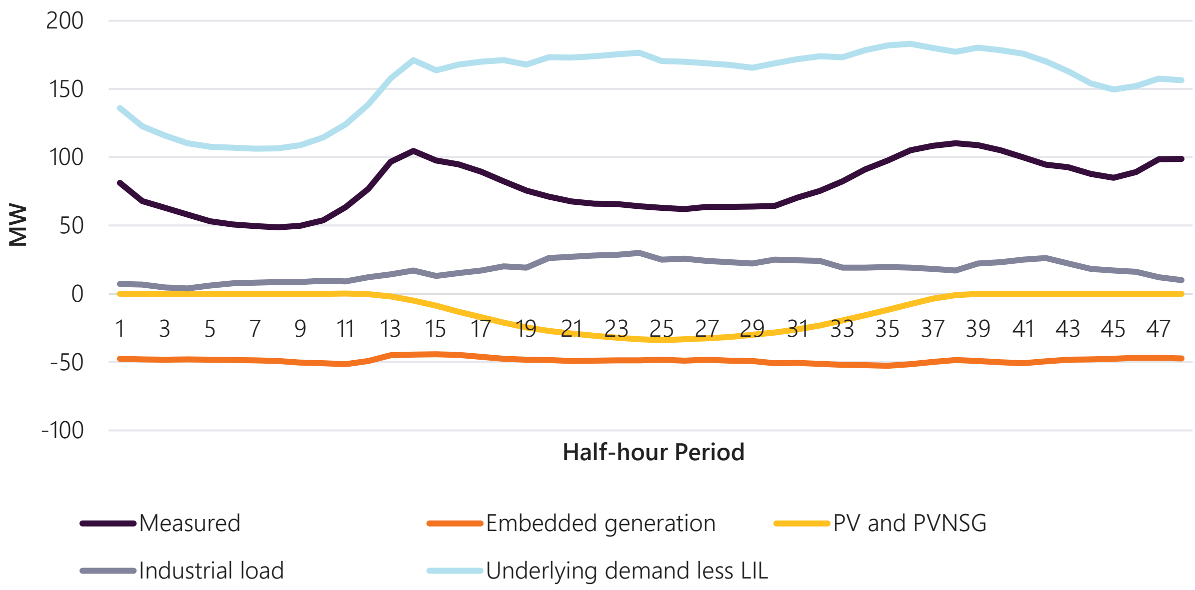
To determine the load at the connection point without the effect of the embedded generator, the generation is added back to the measured connection point load.

Connection point forecasting aims to identify and model the non-scheduled generators which have a significant impact on the demand forecast (particularly minimum demand) of a connection point. This includes generators that have a dedicated meter, separate from customer load.

It is to be noted that some of the embedded solar generation might lie within the definition of PVNSG as well. Depending on the size and other features of the generation units, AEMO categorizes those units as either PVNSG or embedded solar farms. For example, if a PVNSG features tracking solar panels, the normalized solar generation modelled for the TCP might not represent the expected normalized generation of that PVNSG, and therefore, if the unit has a dedicated NMI measuring its generation, then the NMI data rather than estimated generation is used to remove the impact of the unit's generation from TCP load.

[6] The normalised trace has values between 0 and 1 for each half hourly interval. A value of 1 indicates the distributed rooftop PV/PVNSG systems at that connection point are generating at their rated capacity, and 0 indicates no generation from rooftop PV/PVNSG systems.

Figure 5 Example of a load trace by component used to calculate underlying demand less ELL



2.3.2 Setting categorical variables

Each half-hourly datapoint is categorised as being part of one of a small number of pre-determined "defined periods". Each defined period should reflect both a:

- Seasonal classification (summer, winter, shoulder), and
- A time of day classification (for example, early morning – 2.00 am to 6.00 am; morning – 6.00 am to 11.00 am; early afternoon – 11.00 am to 3.00 pm)[7].

The key factors influencing AEMO's specification of the defined periods are:

- Drivers of consumption should be reasonably similar across a defined period, and
- Level of PV and wind output (and any other material source of embedded generation) should be reasonably similar across a defined period.

The defined periods may differ depending on jurisdiction (for example, to reflect daylight hours), or change over time to reflect the increased penetration of new technology (such as battery charging at a certain time of day).

Additional categorical and interaction variables are set up to flag holidays, school holidays, weekends, Christmas periods and days of the week.

2.3.3 Historical adjustments

In the historical data, there will often be evidence of load transfers and block loads (step changes in demand caused either by large customers connecting or changing demand patterns) which have a significant impact on the maximum and/or minimum demand trends at the connection point level. Information on some historical block loads and transfers is provided by the relevant DNSP, where available. Additional block loads and transfers are detected by the anomaly detection algorithms developed in AEMO's connection point forecasting system as well as through visualisation of the historical data.

[7] AEMO retains the flexibility to develop and refine the "defined periods" that are utilised as part of the methodology – both across TCPs and over time.

Figure 6 shows an example of detecting load transfers through the anomaly detection algorithm in AEMO's forecasting software. In this figure, the red dots flag potential events worth investigating.

Figure 6 Example of potential outliers detected by the anomaly detection algorithm

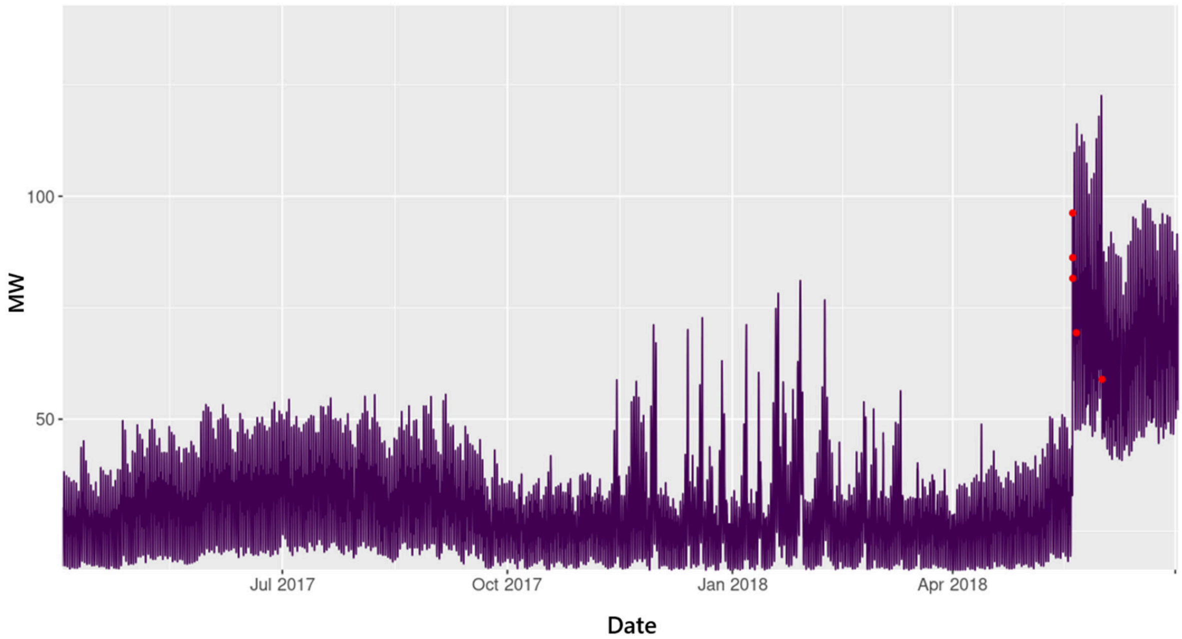
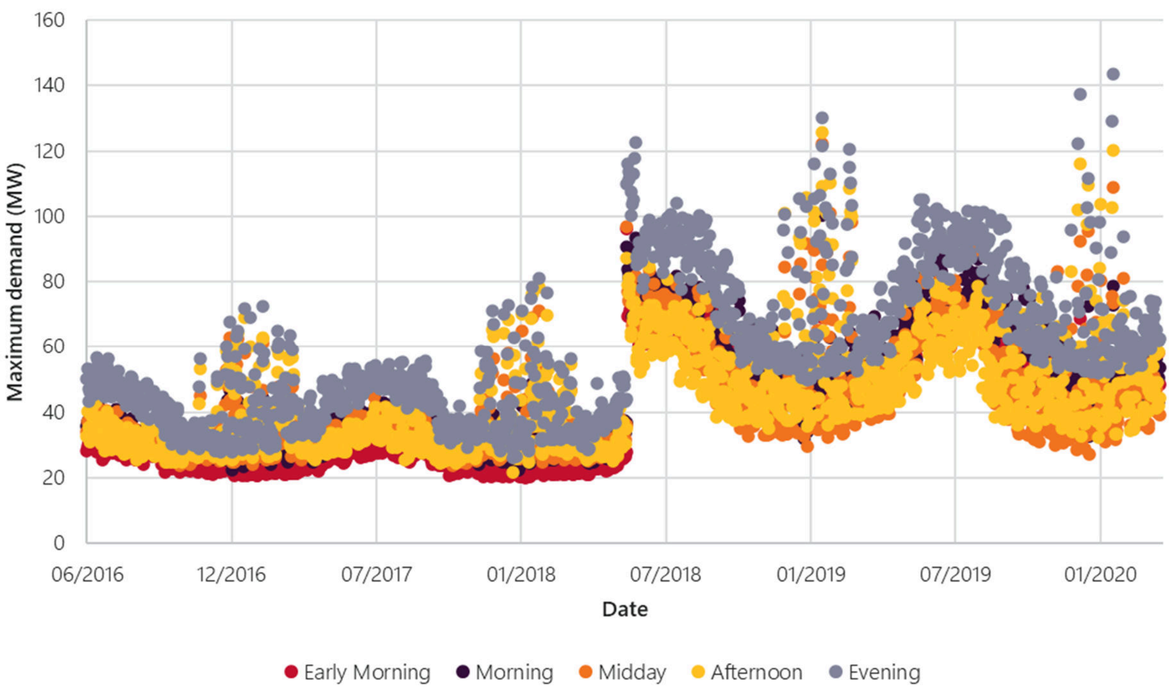


Figure 7 gives an example of detecting the presence of a block load visually. In this figure, a step up has occurred in the maximum demand in 2018. Upon observing an event such as this and verifying that it is atypical using other information (DNSP advice, weather data, and network knowledge), an adjustment is made in the historical data to remove the impact.

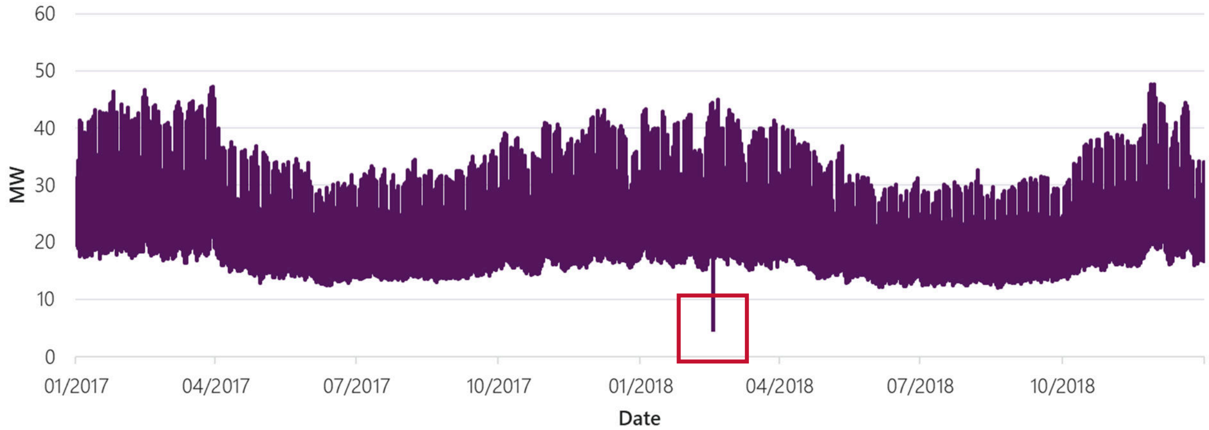
Figure 7 Example of visual detection of block load



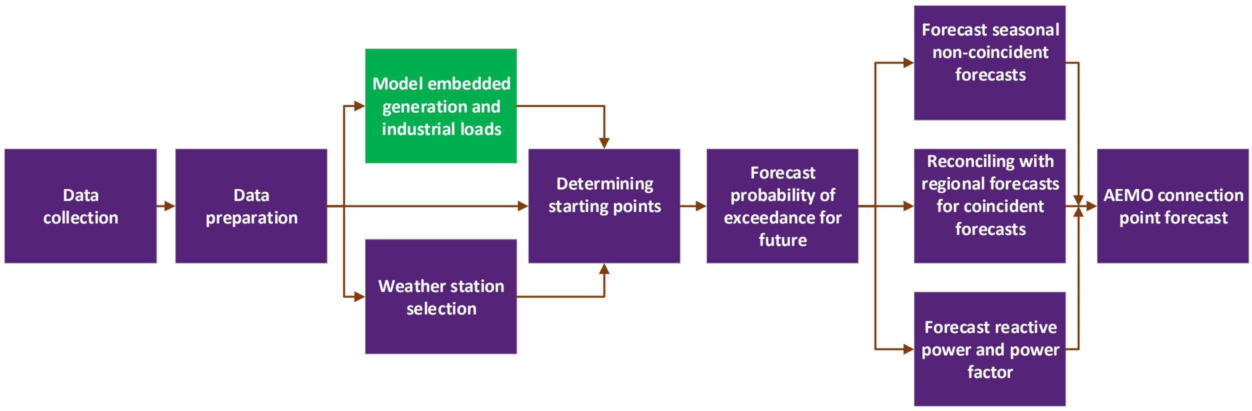
Outages, load transfers, and outliers can have a considerable impact particularly on minimum demand and can potentially distort the goodness of fit of the regression models. Therefore, data cleansing is important to ensure historical demand data best represents the demand profile including its extreme events of maximum and minimum.

Figure 8 shows an example of a load profile with low demand outliers. As mentioned, AEMO uses different tools to detect the outliers and pursue proper adjustments to cleanse the data. Data cleansing is only implemented where the half-hourly periods in question can be verified as being unrepresentative of the true population of the connection point[8]. The true population here refers to the demand under normal operating conditions and in the absence of temporary transfers, outages (including non-credible contingencies) or metering issues.

Figure 8 Example of outliers impacting the historical minimum demand



2.4 Model embedded generation and large industrial loads



The objective is to estimate the demand or generation of the embedded loads and generators which impact the demand at a TCP level for each half-hour of a year.

AEMO aggregates the existing non-operational embedded generators whose outputs are not related to wind speed or solar irradiance at each connection point by fuel type and develops models to estimate their normalised generation. The explanatory variables are time/calendar data such as month of the year, day of the week, and half-hour of the day. The normalised generation values can then be used to estimate the impact on demand of new generators that have no historical load traces.

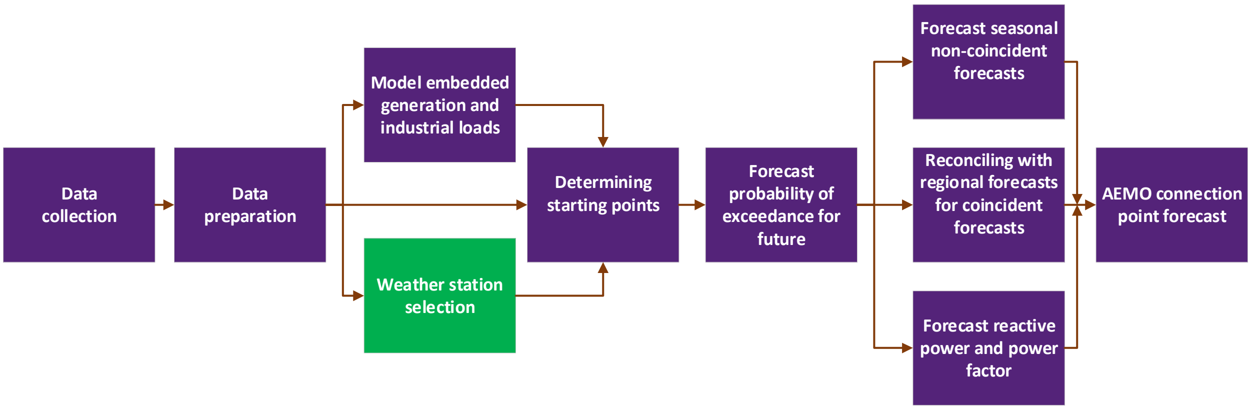
[8] It is to be noted that the "true" population of a transmission connection point can fundamentally change with significant permanent load transfers.

The MW demand of each embedded large industrial load is modelled separately given the explanatory variables such as month of the year, day of the week, and half-hour of the day.

Models are developed to mimic the relationship between reanalysis wind speed and wind power generation for each TCP. AEMO uses reanalysis wind data because current analysis suggests that the actual recorded wind speed data from weather stations do not correlate highly enough with actual wind generation of embedded wind farms.

Embedded solar generation is modelled similar to PVNSG, using the modelled estimated normalized generation provided by AEMO's consultants.

2.5 Weather station selection



It is necessary to select half-hourly temperature data from a weather station that is most appropriate to the connection point being examined.

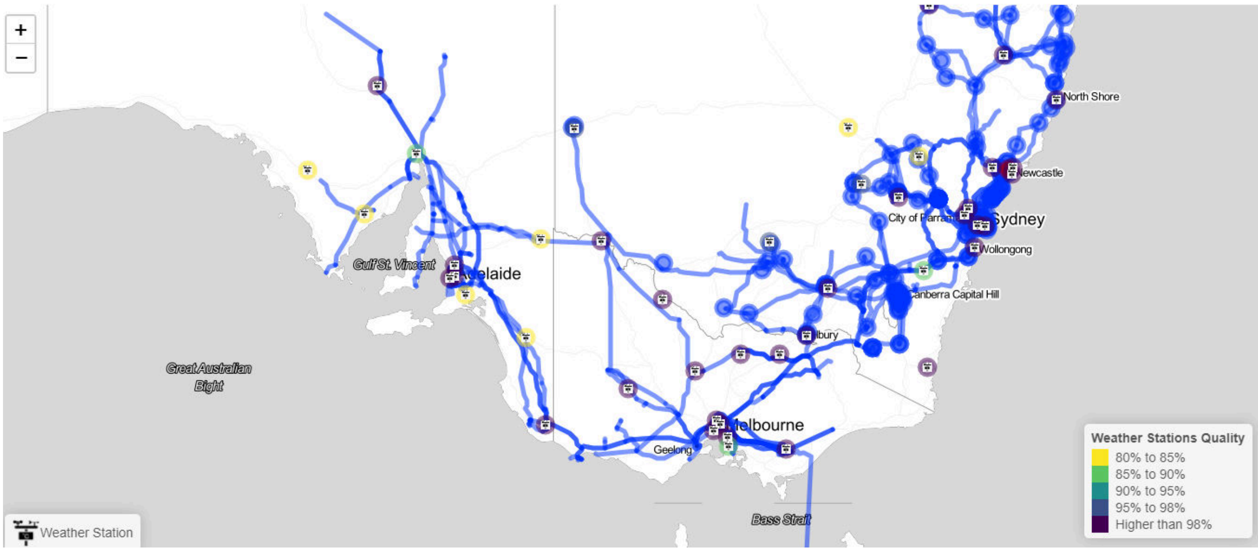
The aspects that identify the suitability of a weather station include:

- Demand–weather relationship.
- Limited gaps in the data.
- The required historical weather data is available, and the station is currently operating.
- Distance from the weather station to connection point.

These factors, together with the forecaster's qualitative assessment, are used to determine the most suitable weather station for a given connection point. Figure 9 shows a snapshot of the weather stations and connection points where the weather stations are colour coded based on the percentage of their half-hourly data available. Weather data is sourced from the Bureau of Meteorology (BOM), with missing data occurring occasionally for a number of reasons including station maintenance, outages and transmission errors. AEMO uses an ensemble machine learning algorithm to gap-fill the selected weather stations' data.

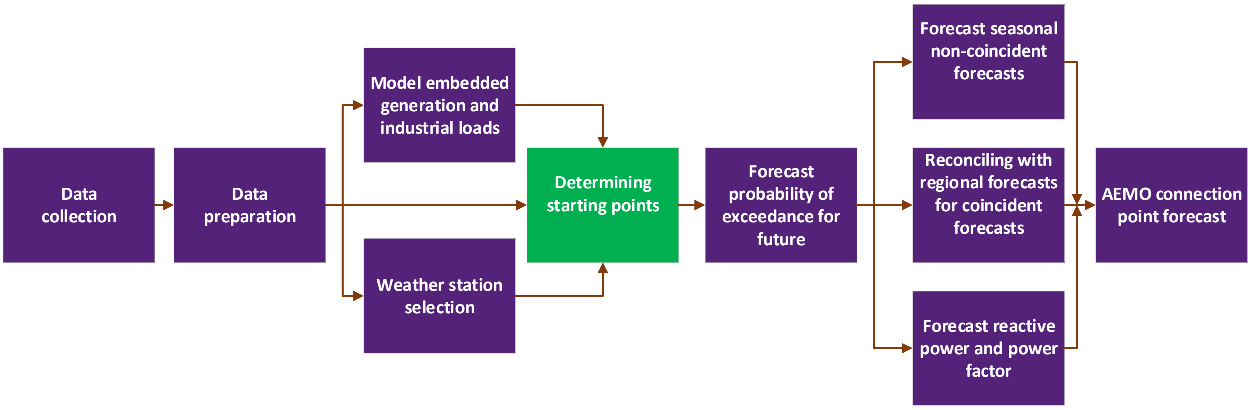
Weather selection is an iterative process. If the selected weather station does not provide the required performance, an alternative station will be tested. In addition, in cases where a single weather station cannot represent a TCP, for example if a TCP covers a large diverse geographic area, a hybrid weather station which is a weighted average of multiple weather stations is created.

Figure 9 Map of connection points and weather stations



2.6 Determining starting points

The idea is to produce the starting point POEs of the maximum and minimum demand for the base year for each connection point and each defined period, and grow the starting point POEs based on the operational demand drivers, with the "official" minimum/maximum demand for forecast years being the lowest/highest one in each year[9]. The base year is one full calendar year before the cut-off date of historical data used to produce forecasts. This section explains how starting point POEs are derived.



2.6.1 Model development and selection

A regression model is built for each defined period and each connection point.[10] AEMO fits regression models to describe the relationship between underlying demand less ELL and the key explanatory variables outlined in Table 3 for all the half-hours within a defined period.

[9] In the current methodology, only the base year is simulated given all the available weather data, and the impacts of the various technologies on different periods of day and seasons are modelled through the defined periods. This is a computationally efficient alternative to simulating all the half-hours in the future years. AEMO is exploring the potential improvements provided by simulating all the future years versus the computational expenses to determine future development plans.

[10] It is possible to train a single regression model for all the half-hours of the day and seasons (all defined periods) and form starting point POEs for the defined periods at a later stage, after weather normalising the base year. This can be done by grouping the simulations by defined periods and producing the starting point POEs for the grouped simulations. Currently, AEMO develops separate regression models for each defined period to be able to focus more on the model performance for the defined periods, which are more important from the maximum or minimum demand perspective, and to have more concentrated residual distribution for each defined period.

AEMO uses a Machine Learning algorithm to derive a model with good fit and strong predictive power. The Least Absolute Shrinkage and Selection Operator (LASSO) regularisation algorithm, a special case of Elastic Net, selects the best model from the range of variables available and all the interactions between the variables. The model is developed trading off the model bias[11] and model variance[12] to derive a parsimonious model with strong explanatory power. AEMO then performs additional in-sample and out-of-sample model diagnostic checks on the best model selected by LASSO. Where the best model fails these checks, AEMO adjusts the LASSO algorithm iteratively.

AEMO:

- Performs k-folds out-of-sample cross-validation[13] to find the optimal model that trades off between bias and variance.
- Inspects the QQ-plots, the residual diagnostics over time, and against the explanatory variables.
- Inspects residuals to ensure that the assumptions made for residuals when simulating minimum and maximum demand are relevant.
- Compares actuals against predictions.

Variables included in the model selection are listed in Table 3. The selected significant variables which eventually are used in the final models can be very different across the connection points and the defined periods. AEMO uses several years of weather–demand data to build the regression models. The length of data used is decided based on a trade-off made between the effort required for data cleansing, especially around minimum demand events, and the required performance.

Table 3 List of variables included in model selection

Variable	Description
Public holiday	Dummy flag for public holidays
School holiday	Dummy flag for school holidays
Christmas period	Dummy flag for Christmas period
Half-hour factor	A factor variable with values for each half-hour of the day
Weekend dummy	Dummy flag for weekends
Month factor	A factor variable with values for each month of the year
Temperature	Linear, quadratic or cubic transformation of temperature, 3-hour and 6-hour rolling averages of temperature depending on model fit.
Time index	To detrend the trace

2.6.2 Weather normalise base year

The weather normalisation process is primarily used to remove the influence of year-to-year variability in weather and provide a distribution of possible maximum and minimum demands. Demand levels at the lower end of the distribution are more likely to be exceeded than demand levels at the higher end.

[11] Under-fitting the model results in a model with high bias.

[12] Over-fitting the model results in a model with high variance.

[13] A k-fold cross-validation is performed by breaking the data set randomly into k smaller sample sets (folds). The model is trained on (k-1) of the folds and validated against the remaining fold. The model is trained and validated k times until each fold is used in the training sample and the validation sample. The forecast accuracy for each fold is calculated and compared between models.

Weather data

To apply weather normalisation for the base year, the available historical records of the half-hourly temperature[14] from the weather station selected for the connection point[15], as well as wind and PV normalised generation data, are collected. Wind and PV normalised values and historical temperature data are joined by date and half-hours in the base year in the same-day order, to preserve aspects of the spatio-temporal correlation of weather[16]. The data for the defined period of interest is retained. It is to be noted that the weather is simulated without climate adjustments, which are expected to have a relatively minimal impact on demand over the 10-year forecast period.

Simulating maximum/minimum demand

Simulations of operational demand are calculated by deploying the model developed in Section 2.6.1. Each year of historical weather data is combined with other explanatory variables in the model to predict operational demand, as described in the following steps. For each simulated half-hour[17] of the day in each defined period:

- The underlying demand less ELL is calculated by the model built in Section 2.6.1 along with the simulated explanatory variables (for example, temperature, the day of week flag assigned to that simulated day, the public holiday flag assigned to that simulated day).
- The simulated PV and wind normalised generation that aligns with the simulated temperature are multiplied by the installed capacity of each solar and wind generator to estimate its generation.
- The megawatt (MW) forecast of ELL is determined.
- The estimated generation of non-operational embedded generators whose outputs are not related to wind speed or solar irradiance are calculated by multiplying the predicted contribution factors by the installed capacities of the associated installations.
- Kernel density estimation is used to generate the empirical distribution of model residuals (developed in Section 2.6.1). n random samples[18] are drawn from the empirical distribution and added to the predictions. This accounts for demand variability not captured in the weather–demand relationship.
- The simulated operational demand is determined as:

$$\begin{aligned} Operational\ Demand \\ &= Underlying\ demand\ less\ ELL - Simulated\ PV\ output - Simulated\ PVNSG\ output \\ &- other\ embedded\ generation + Estimated\ ELL \\ &+ sampled\ regression\ model\ residual \end{aligned}$$

Since n samples are drawn from the empirical distribution of residuals, the highest/lowest operational demand in each set can be retained as the maximum/minimum demand, such that there are n maximum/minimum demands for that year.

The next year of historical weather data is used in the same process, and so on until all the weather data has been processed.

[14] To produce 2020 AEMO connection point forecasts, 12 years of half-hourly temperature records were used. The bottleneck for the length of weather data available for simulations was the reanalysis wind data plus the limited number of weather stations with quality data and close to connection points. AEMO is seeking options available to use more weather data in the following years.

[15] If a single weather station cannot represent the weather at a connection point, AEMO develops hybrid weather stations as a weighted average of several weather stations near a connection point.

[16] AEMO will continue exploring alternative approaches that serve to preserve the spatio-temporal correlation of weather at the connection points and are computationally efficient for connection point forecasts.

[17] Simulated time or simulated half-hour refers to each half-hour in the base year whose demand is to be simulated.

[18] The number of samples drawn from the distribution is determined by making a compromise between accuracy of POEs and available computational resources.

2.6.3 Determine probability of exceedance for base year

If Y years of weather data are used in the simulation process, and n residual sampling events are undertaken each time, then $Y \times n$ maximum and $Y \times n$ minimum demands are produced for each defined period. Each of the $Y \times n$ full years of half-hourly traces is referred to as one simulated year in the rest of this methodology document. These form the maximum and minimum demand distributions for each defined period. Using the distributions, the 10%, 50% and 90% POE levels are obtained from the 90th, 50th, and 10th percentiles, respectively. This becomes the starting point upon which forecasts for maximum and minimum demand for that defined period are made.

2.6.4 Estimate the contribution factors of embedded generation at the time of maximum and minimum demand

For each defined period, the following steps are adopted for determining the contribution factors to be applied in order to reinstate the impact of PV, wind, and other embedded generation facilities:

1. As described in the section above, for each of the simulated years, the simulated PV, wind and embedded generation (where relevant) outputs that underpin each of the simulated maximum/minimum demands in that year will be recorded.
2. Through that process, the normalized generation of wind and PV related technologies that align with the simulated maximum/minimum demand are identified as the contribution factors of those technologies.
3. The contribution factor for PV/wind technologies at the time of maximum/minimum demand is considered to be the median of the values identified in the previous step.
4. The typical half-hour and day of the week underpinning maximum/minimum demands are identified as the mode[19] of the half-hours and days of the week present in all the simulated maximum/minimum demands. The identified typical half-hour and day of the week then are used in the models developed in Section 2.4 to estimate the contribution factor of other embedded generation facilities (such as biogas, fossil generators) at the time of maximum/minimum demand for each defined period.
5. The values from the previous steps are retained in order to reinstate the impact of the embedded generation in forecasts in section 2.7.

It should be noted that the estimated contribution factors do not result in an output for that technology (for example, PV, wind) necessarily being of a statistically pre-defined level (such as 10% POE); rather, it simply reflects the contribution factor at the time when AEMO's modelled POE level maximum/minimum demand occurs in the simulations.

2.6.5 Estimate the contribution of embedded large loads at the time of maximum/minimum demand

The ELLs might be large enough to significantly affect the maximum/minimum demand at the TCP. In addition, the demand profiles of ELLs might change depending on the demand at TCP or system level. Therefore, AEMO aims to predict the demand of each ELL at the time of each seasonal maximum/minimum POE for the corresponding TCP.

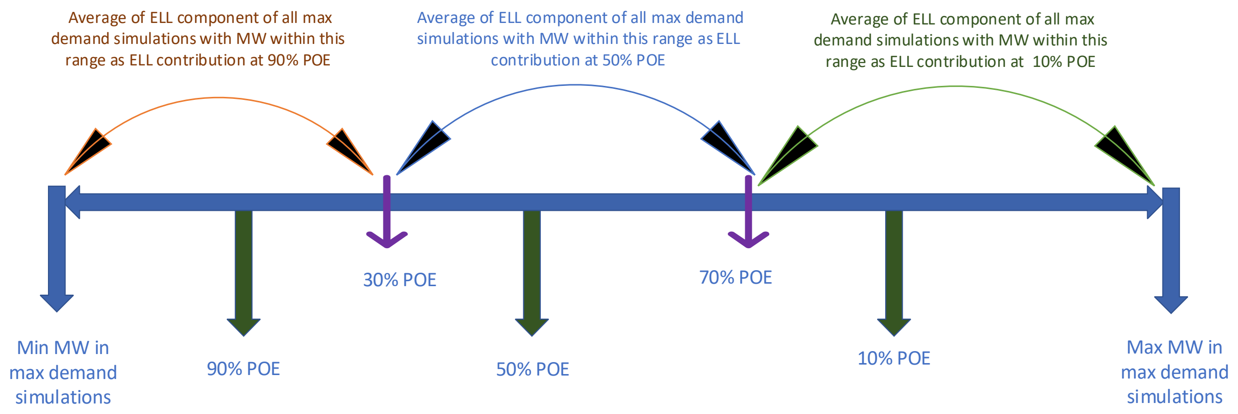
To do so, for each defined period and each of the simulated years, ELL demand that coincides with each of the simulated maximum/minimum demands in that year is recorded. This set of simulated ELL contributions is partitioned into three sets based on the value of simulated minimum/maximum operational demand to which it corresponds – from the minimum simulated operational demand value to the 30th percentile, from the 30th percentile to the 70th percentile, and from the 70th percentile to the maximum simulated operational demand value. The average of the ELL simulations that correspond to the simulation instances in each of these sets is calculated. The value derived from the first set is assigned to the ELL contribution of the 90% POE forecast,

[19] In statistics, the mode is the most observed value in a set of data.

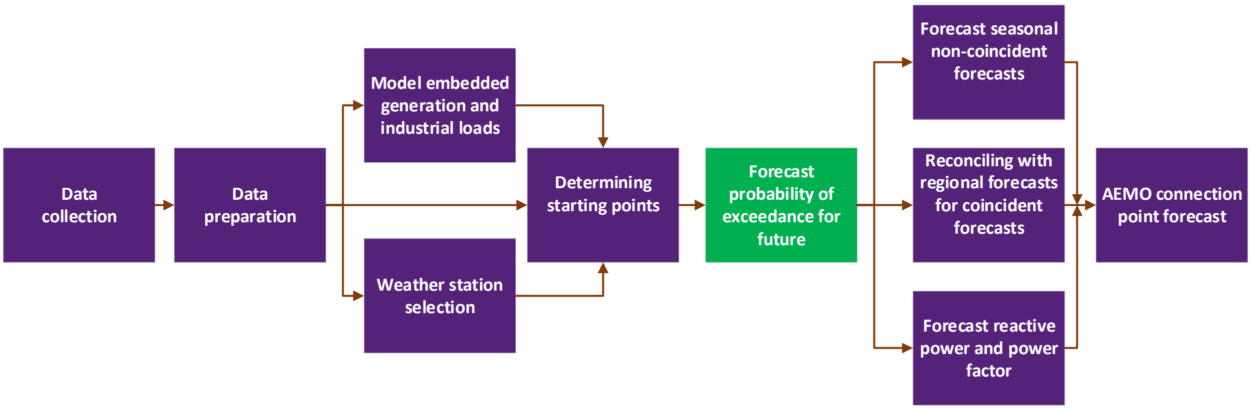
the value from the second set assigned to the ELL contribution of the 50% POE forecast and the value derived from the final set is assigned to the ELL contribution of the 10% POE forecast.

A representation of this process for maximum demand is seen in Figure 10.

Figure 10 Estimate the contribution of ELL at the desired POE levels



2.7 Forecast probability of exceedance for future



AEMO produces component level forecasts to account for the share of each technology in the operational demand to be supplied. More explicit modelling of spatial demand drivers reduces the reliance on reconciliation, and accounts for diversity in the growth of demand drivers and behind the meter technologies across connection points. In the following, it is explained how each component of operational demand is forecast for future years and how the operational demand forecast is calculated based on these components.

2.7.1 Grow underlying demand less embedded large loads

The residential/commercial demand, defined as underlying demand less ELL, associated with the starting point POEs for each defined period are grown proportional to population forecasts for the respective TCP. The estimated historical and forecast population for each TCP is determined by allocating the historical and forecast population of Local Government Areas (LGAs) to connection points based on an estimation of the number of customers in each LGA supplied by each TCP.

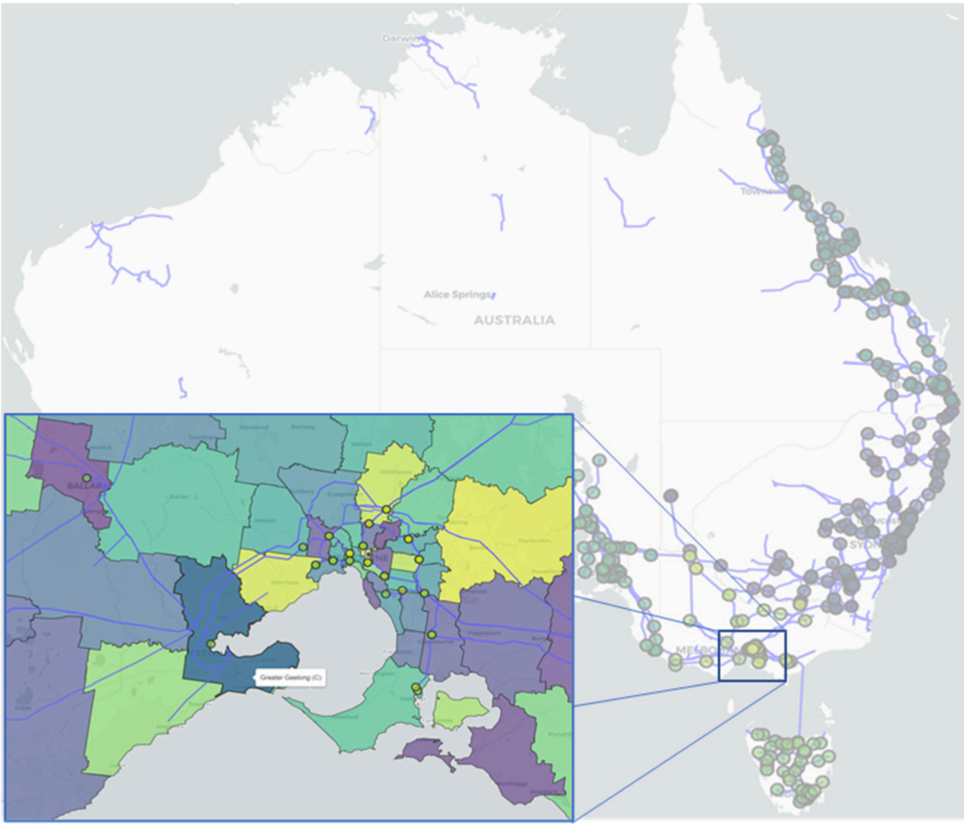
The historical population and population forecasts of LGAs are converted to a connection point basis by:

1. Mapping NMIs to LGAs and connection points based on the coordinates of the NMIs, mapping NMIs to TNIs, and TNIs to TCPs, using AEMO databases and the boundaries of LGAs from the Australian Bureau of Statistics (ABS).
2. Estimating the percentage of retail NMIs in each LGA associated with each connection point.

3. Allocating LGA population to connection points, based on percentages calculated in step 1.
4. Adjusting the forecasts, if required, given the historical data and information available to AEMO.

Figure 11 shows a schematic overlay of TCPs and LGAs. As the methodology matures and more spatial data becomes available, AEMO aims to pave the way for factoring in customer composition (residential, business, industrial, agricultural), and benefit from more determinative and descriptive demand drivers to grow underlying demand less ELL.

Figure 11 Connection points with LGA boundaries



2.7.2 Forecast the impact of embedded technologies and Large Industrial Loads

PV and PV Non-scheduled Generation

AEMO's consultant(s) provide yearly installed capacity forecasts for rooftop PV and PVNSG on a postcode basis. The forecasts are converted to a connection point basis using AEMO data by:

1. Estimating the percentage of retail NMIs in each postcode associated with each connection point.
2. Allocating the rooftop PV/PVNSG installed capacity forecasts to connection points, based on percentages calculated in step 1.
3. Adjusting the forecasts, if required, given the historical data and information available to AEMO.

Given the installed capacity forecasts for each connection point as outlined above and the contribution factors of PV/PVNSG estimated in Section 2.6.4, the total demand offset by PV/PVNSG can be predicted for each defined period's maximum and minimum demand forecasts.

Electric vehicles

AEMO's consultant(s) provide an EV grid-connected charge and discharge profile forecast at the NEM region level, consisting of a per vehicle kilowatt (kW) value for each combination of month, day of week, and half-hour of the day.

As outlined in Section 2.6.4, the typical half-hour and day of the week underpinning maximum and minimum demands for each defined period is extracted from the simulations. By joining EV profile forecasts and the typical time of maximum and minimum demands for each defined period, AEMO estimates the kW per vehicle contribution of EV.

Forecasts of vehicle numbers also are provided by AEMO's consultants per NEM region. Disaggregation is performed to split the regional numbers to connection points using an approach similar to that described earlier for PV installed capacity in which the number of customers per connection point is used as a proxy for disaggregation.

Given the kW per vehicle forecast for each maximum and minimum demand forecast and the vehicle numbers forecast for each future year and each connection point, AEMO can estimate the EV component of maximum and minimum demand forecasts for each defined period.

Battery Energy storage systems

Similar to EVs, ESS charging/discharging profile forecast is provided by AEMO's consultant(s) for each NEM region, and the typical time of maximum and minimum demands for each defined period is used with this profile to estimate the capacity factor contribution of batteries to connection points in the region.

ESS installed capacity forecasts are available at the postcode level, for which AEMO adopts a similar approach as described above for PV/PVNSG to map postcode level battery installed forecasts to connection points. This is performed in relation to an estimation of the number of customers each connection point supplies in each postcode.

Other embedded generation

According to AEMO's definition, operational demand does not include the impact of operational generators, (scheduled generation, semi-scheduled generation, and non-scheduled wind/solar generation of aggregate capacity ≥ 30 MW and exempt generation)[20]. However, the operational demand forecasts are reduced by the impact of other small-scale wind power, hydropower, gas or biomass-based cogeneration, generation from landfill gas or wastewater treatment plants, and smaller peaking plants or emergency backup generators.

To reinstate the impact of the small-scale solar and wind generation, the contribution factors extracted from simulations (as explained in Section 2.6.4) are multiplied by the installed capacity of the respective units to estimate the generation contribution for each defined period's maximum and minimum demand forecasts.

For the rest of the embedded generation, as explained in Section 2.6.4, the typical half-hour and day of the week underpinning maximum and minimum demands for each defined period are extracted from the simulations. These times are input to the models developed in Section 2.4 to estimate the normalised generation for each fuel type at each connection point. The normalised generation is then multiplied by the future installed capacity of the respective units to estimate the generation at the time of the maximum and minimum demand for each defined period.

The future installed capacity includes the committed retirement of the generators and might include committed future generation if enough information regarding their network connections and generation commitments is available at the time of producing connection point forecasts.

[20] Operational generators include local scheduled generation, semi-scheduled generation, and non-scheduled wind/solar generation of aggregate capacity ≥ 30 MW, and generation imports to the region and exempt generation. For more definitions, see AEMO, Demand Terms in EMMS Data Model, 2021, at https://www.aemo.com.au/-/media/Files/Electricity/NEM/Security_and_Reliability/Dispatch/Policy_and_Process/Demand-terms-in-EMMS-Data-Model.pdf

Embedded Large loads

The demand of ELLs at the time of each defined period's maximum or minimum demand and at the POE levels of interest are determined in Section 2.6.5. The ELLs' demand is assumed to be constant over the forecast horizon unless there is information available to AEMO regarding future development plans or change of profile.

2.7.3 Forecast operational demand for each defined period

Operational demand maximum and minimum forecasts for each defined period are determined by reinstating the impact of behind the meter technologies and industrial loads as follows:

$$Operational\ Demand\ = Underlying\ Demand\ less\ ELL\ Forecast - PV\ Generation\ Forecast\ - PVNSG\ Generation\ Forecast\ - Other\ Embedded\ generation\ Forecast\ + ELL\ Forecast\ + EV\ Forecast - ESS\ charge/discharge\ Forecast$$

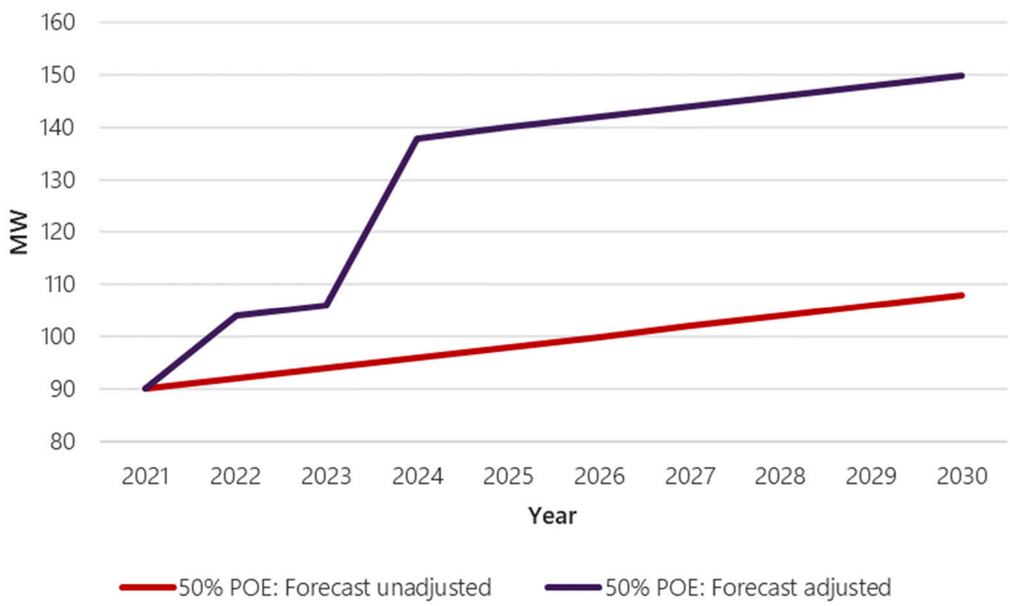
It is to be noted that because AEMO does not remove the impact of existing residential and commercial ESS and EVs from the metered TCP demand, in the above equation, $Underlying\ Demand\ less\ ELL\ Forecast$ is expected to account for the existing EV and ESS in the base year. Therefore, $EV\ Forecast$ and $ESS\ charge/discharge\ Forecast$ include the growth in each future year relative to the base year.

2.7.4 Adjust for future block loads and transfers

Information regarding new block loads and transfers between terminal stations is provided to AEMO by the relevant DNSPs. This is reviewed and applied to the forecast as an increase or decrease in the load at that connection point. The adjustments can be different for maximum and minimum demand forecasts[21].

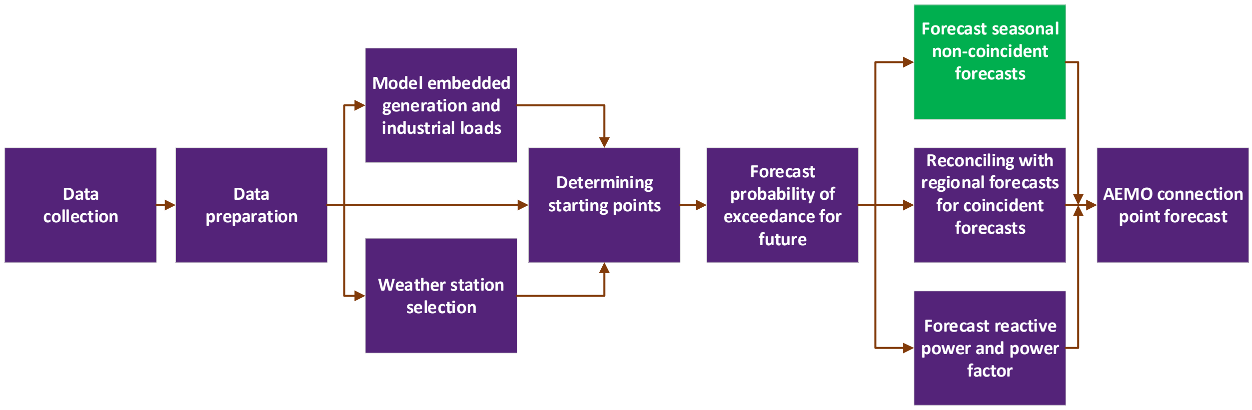
Figure 12 shows an example where a forecast is adjusted for one block load in 2022 and one block load expected in 2024.

Figure 12 Adjusted forecasts for future block loads



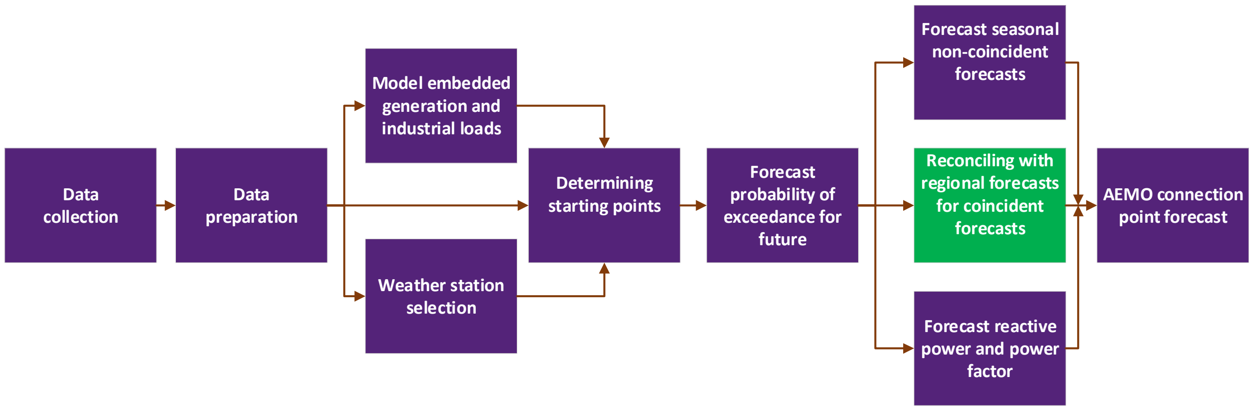
[21] AEMO seeks information from DNSPs regarding the impact of the new loads and transfers on the maximum and minimum demand at connection points. In the absence of any further information regarding the different impact a load can have on the maximum and minimum demand, AEMO makes assumptions by studying the existing loads with similar demand characteristics.

2.7.5 Determine seasonal non-coincident forecasts



The non-coincident seasonal maximum demand POE forecast for each TCP and each future year is derived as the highest forecast for that year among all the defined periods associated with the POE level and season of interest. Similarly, the non-coincident minimum demand POE forecast for each TCP, future year and season is the lowest forecast for that year among all the defined periods associated with the same POE level and season.

2.8 Coincident forecasts and reconciling with regional forecasts



The methodology described in the previous sections focuses on developing non-coincident forecasts for each connection point. They are non-coincident because they do not necessarily coincide with the time of the region's aggregated maximum and minimum demand. Coincident forecasts, on the other hand, represent the demand of the connection point coinciding with the time of the region's aggregated maximum and minimum demand.

Coincident connection point forecasts are reconciled to ESOO regional forecasts. Along with regional demand forecasts, AEMO publishes maximum and minimum demand time for each season, region, and future year. The time is presented in the form of typical half-hours from the distribution of maximum and minimum demand forecast time. The published typical times are considered as the coincident time for each season in the connection point forecast reconciliation process.

The reconciliation approach should

- Lead to robust results in the context of both maximum and minimum demands.
- Account for the diversity in the timing of maximum/minimum demand across connection points.
- Produce robust results when the demand forecast at connection point and/or regional level is negative.

AEMO's approach involves the following steps:

- Identify the "defined period" that aligns with the timing of when regional maximum and minimum demands are forecast to occur (e.g., mid-afternoon in summer). These identified defined periods will be referred to as "coincident defined periods".
- Using the approach outlined in Section 2.7, forecast the maximum and minimum demands for the "coincident defined period" at each TCP in a region.
- Calculate diversity adjustment factors using historical data at each TCP.
- Calculate diversified coincident forecast by applying diversity adjustment factors to non-coincident forecasts of "coincident defined periods" at each TCP.
- Adjust the diversified coincident forecasts to ensure that the sum of TCP forecasts matches the coincident regional forecast.

The above steps are explained in detail in the following sections.

2.8.1 Calculate diversity adjustment factor

To estimate each connection point's demand forecast at the time of the regional maximum and minimum demand, historical diversity adjustment factors are calculated. A diversity adjustment factor for a defined period is formulated as:

$$Diversity\ adjustment\ factor = \frac{TCP\ demand\ at\ time\ of\ regional\ maximum/minimum - TCP\ maximum/minimum\ demand}{|TCP\ maximum/minimum\ demand|}$$

and can be calculated for each historical year and defined period. Diversity adjustment factors are calculated by involving the absolute values of maximum and minimum demand to make the equation robust where minimum demand at a connection point is negative. Diversity adjustment factors are negative for maximum demand and positive for minimum demand.

Even within a defined period, diverse levels of demand are expected at connection points, especially for the case of minimum demand. While maximum demand in most connection points is temperature-driven, minimum demand is highly affected by the contribution of embedded generation and PV. As an example, a connection point may have a minimum demand of -100 MW in February, but a demand of +20 MW at the time of the regional minimum demand in January, both occurring in the middle of the day. In this example, the diversity adjustment factor of minimum demand for the defined period midday-summer is +1.2 for that connection point.

Diversity factors are calculated from historical demand data, taking the average of the most recent three years where possible.

2.8.2 Calculate diversified coincident forecast

After diversity adjustment factors are calculated for each TCP and each coincident defined period, they are applied to non-coincident forecasts produced for the coincident defined periods.

$$\begin{aligned} Unreconciled\ coincident\ forecast \\ &= NonCoincident\ forecast\ for\ the\ coincident\ defined\ period \\ &+ Diversity\ adjustment\ factor\ of\ coincident\ defined\ period \\ &\times |NonCoincident\ forecast\ for\ the\ coincident\ defined\ period| \end{aligned}$$

Table 4 provides a few examples for calculation of coincident forecasts.

Table 4 Example of diversified minimum demand forecasts

Year	Non-coincident connection point minimum demand forecast (MW) for midday	Diversity adjustment factor	Coincident (unreconciled) minimum demand connection point forecast (MW)
2021	-110	0.2	-88

Year	Non-coincident connection point minimum demand forecast (MW) for midday	Diversity adjustment factor	Coincident (unreconciled) minimum demand connection point forecast (MW)
2022	-120	0.2	-96

Similarly, diversified coincident maximum demand forecasts can be calculated as in the examples of Table 5 .

Table 5 Example of diversified maximum demand forecasts

Year	Non-coincident connection point maximum demand forecast (MW) for midday	Diversity adjustment factor	Coincident (unreconciled) maximum demand connection point forecast (MW)
2021	100	-0.2	80
2022	120	-0.2	96

2.8.3 Calculate reconciled coincident forecast

The next step is to determine the adjustment factor that needs to be applied to ensure that the sum of coincident connection point forecasts matches the coincident regional forecast. The reconciliation adjustment factor for each year is calculated as

$$Coincident\ adjustment\ factor = \frac{Regional\ forecast - \sum_{i \in N}(Unreconciled\ coincident\ forecast\ of\ TCP^i)}{\sum_{i \in N}(|Unreconciled\ coincident\ forecastof\ TCP^i|)}$$

with N as the number of connection points in a NEM region.

Coincident adjustment factors are calculated based on the absolute value of unreconciled coincident forecasts to make the equation robust where minimum demand at some connection points is negative. Coincident adjustment factors are calculated for each season, POE level, and maximum and minimum demand, separately. One common coincident adjustment factor is used for each year and region. An example is shown in Table 6.

Table 6 Example of coincident adjustment factors

Year	Sum of coincident (un-reconciled) connection point forecast (MW)	Sum of absolute values of coincident (un-reconciled) connection point forecast (MW)	Regional forecast (MW)	Coincident adjustment factor
2021	400	500	550	0.300
2022	300	550	450	0.273
2023	200	650	350	0.231
2024	150	680	185	0.051
2025	50	700	95	0.064
2026	20	740	60	0.054
2027	-4	780	-20	-0.021
2028	-21	820	-40	-0.023
2029	-40	840	-90	-0.060

Year	Sum of coincident (un-reconciled) connection point forecast (MW)	Sum of absolute values of coincident (un-reconciled) connection point forecast (MW)	Regional forecast (MW)	Coincident adjustment factor
2030	-50	900	-120	-0.078

Each reconciled coincident connection point forecast is then calculated by adjusting the unreconciled coincident connection point forecast, as follows:

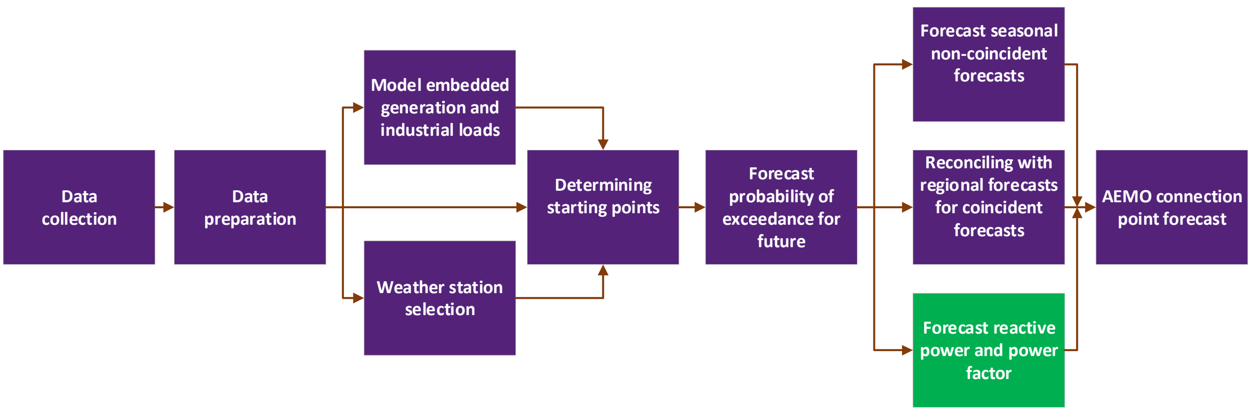
$$\begin{aligned} \mathit{Coincident\ forecast} \\ &= \mathit{Unreconciled\ coincident\ forecast} \\ &+ \mathit{Coincident\ adjustment\ factor} \times |\mathit{Unreconciled\ coincident\ forecast}| \end{aligned}$$

Due to the common adjustment factor, demand at all TCPs in a region is adjusted in the same direction, up or down, depending on the sign of the adjustment factor. The adjustment is larger for TCPs with large absolute values of unreconciled coincident demand. Examples of the adjustment process are given in Table 7.

Table 7 Example of coincident forecast

Year	Coincident adjustment factor	Coincident (unreconciled) connection point forecast (MW)	Coincident (reconciled) connection point forecast (MW)
2021	0.300	16	20.80
2022	0.273	5	6.36
2023	0.231	-1	-0.77
2024	0.051	-5	-4.74
2025	0.064	-8	-7.49
2026	0.054	-12	-11.35
2027	-0.021	-15	-15.31
2028	-0.023	-17	-17.39
2029	-0.060	-19	-20.13
2030	-0.078	-19	-20.48

2.9 Determining reactive power requirements



Production and management of reactive power is an essential element of the planning and operation of an alternating current power system and is important for the operation of many items of end-user equipment. System voltages are also closely linked to the level of reactive flow and must be managed to ensure network assets and consumer equipment are not subject to voltages outside their safe and efficient operating range.

Active demand at a TCP is the result of the accumulation of active power demands at customer connection points plus assessed losses between TCP and customers. Reactive demand at a TCP is more complex to evaluate. In particular, TCP reactive demand is the result of the accumulation of reactive demands at customer connection points *plus* the impact of losses *plus* the injection and extraction of reactive power between customers and TCPs (that is, in the distribution network).

Precise forecasts can only realistically be prepared through substantial network modelling with the help of distribution network providers, which have the knowledge or good estimates of the capability and technical characteristics of lines and other installed equipment. For example, factors such as distances between customers, distribution substations and TCPs can impact losses and voltage.

However, it is possible to approximate the *range* within which any change in power factor at a TCP level might sit. This range – which in effect reflects two book-end scenarios - would reflect assumptions about the degree to which changes in customers' power factor is allowed to (by the actions and equipment of distribution businesses) impact upon the TCP.

At one end, it could be assumed that despite changes to their customers' power factor, distribution businesses manage their networks such that the power factor *at the TCP* does not change at all. If this were the case, megavolt amperes reactive (MVAr) demand would change in line with changes in MW demand at the TCP. At the other end, it could be assumed that changes in customers' power factor flows through to changes in the power factor at the TCP. The actual change in power factor at a TCP will likely sit somewhere between the two ends of this range. A broad approximation of the latter outcome could potentially be made based on the forecasts of MW demand supplied by each TCP and two assumptions[22] as follows:

- Assuming that inverter-based demand has a unity power factor[23].
- The non-inverter-based operational demand power factor will stay constant over the forecast horizon. The non-inverter-based operational demand here refers to the operational demand with the impact of PV, PVNSG, and potentially other inverter-based technologies removed.

Estimate power factor of non-inverter based operational demand

After removing the impact of PV, PVNSG, and operational generators from the historical demand traces, power factors are calculated using active and reactive power values from the top 1% of half-hour demand periods in each defined period and year, ranked by active power. The average value is then adopted as the typical power factor for each year and defined period.

Giving consideration to the trend of calculated average power factors, after sub-setting the historical data to defined periods, a reasonable estimate of future power factors for non-inverter based operational demand is determined by (in order of preference):

- Averaging the power factor over the previous two years *if power factors are within a narrow band of tolerance, 0.03 of each other* and there are two years of data available for that connection point.
- Averaging the power factor over the previous three years *if power factors are within a broader band of tolerance, 0.07 of each other* and there are only three years of data available for that connection point.

[22] The key assumption is that grid-connected and behind-the-meter sources of generation connected through inverters will have effectively unity power factor at the customer connection point.

[23] AEMO acknowledges that recent changes to inverter standards mean that power factor may change, especially at times of minimum demand when many PV systems are exporting power into distribution networks that are near their operating capacity. AEMO will consider revising the connection point forecasting methodology as more data becomes available from inverters meeting the new standard.

- Averaging the power factor over all the previous years of available data and using this long-term average as the estimated future power factor if the most recent power factor estimate is within a band of tolerance, 0.1 from the long-term average.
- Should none of the methods listed above apply to the set of average power factor data, the estimated future power factor is set to the average value from the top 1% of MW demand periods in the previous year.

For each of the criteria listed above, a lagging/leading assessment is also undertaken, whereby a "leading" label is determined if the reactive power is negative for most of these periods, otherwise a "lagging" label is applied.

Forecast reactive power

For each future season and year, the power factors calculated as above for the defined periods associated with the coincident and non-coincident maximum/minimum forecasts at each POE level are used to estimate future reactive power as

$$Reactive\ Power = \ active\ power \times \tan(\cos^{-1}(power\ factor))$$

where active power refers to non-inverter based operational demand forecasts associated with the power factor. Table 8 provides a few examples of the reactive power calculation.

Forecast power factor of operational demand

Given the reactive power forecasts from the previous step and coincident and non-coincident maximum/minimum operational demand forecasts at each POE level for each season and year, the corresponding future power factors can be calculated. As it is assumed that PV and PVNSG do not contribute to reactive power generation or consumption, but they reduce active power, one should expect a declining power factor as PV/PVNSG generation increases. This would be more significant at the time of minimum demand because generally, the contribution of PV/PVNSG at the time of maximum demand is negligible. A few examples are provided in Table 8.

Table 8 Example of power factor and reactive power forecasts

Year	Maximum/minimum	Non-inverter based operational demand	Operational demand	Lead/lag	Non-inverter based operational demand power factor	Reactive power forecast	Power factor
2021	Maximum demand	93.001	87.264	Lag	0.989	13.774	0.988
2021	Maximum demand	93.167	87.089	Lag	0.989	13.799	0.988
2023	Maximum demand	93.292	86.790	Lag	0.989	13.817	0.988
2021	Minimum demand	42.993	7.759	Lead	0.994	-4.627	0.859
2022	Minimum demand	43.298	7.441	Lead	0.994	-4.660	0.848
2023	Minimum demand	43.462	5.470	Lead	0.994	-4.677	0.760

Measures and abbreviations

Abbreviation	Full name
ABS	Australian Bureau of Statistics
AEMC	Australian Energy Market Commission
AER	Australian Energy Regulator
BoM	Bureau of Meteorology
CDF	Cumulative Density Function
CER	Clean Energy Regulator
DER	Distributed Energy Resource
ESS	Energy Storage Systems
ESOO	Electricity Statement of Opportunities
EV	Electric Vehicle
GWh	Gigawatt hours
ISP	Integrated System Plan
KW	Kilowatts
ELL	Embedded Large Load
LNG	Liquefied Natural Gas
MW	Megawatts
NEM	National Electricity Market
NMI	National Meter Identifier
NSG	Non-Scheduled Generation
OLS	Ordinary Least Squares
ONSG	Other Non-Scheduled Generators
POE	Probability of Exceedance
PVNSG	PV Non-Scheduled Generators
PV	Rooftop PV
TCP	Transmission Connection Points
VPP	Virtual Power Plant

Glossary

Term	Definition
Active energy	A measure of the energy that can be converted into useful work, generally expressed in kilowatt hours (kWh).
Active power	The rate at which active energy is transferred.
Apparent power	The square root of the sum of the squares of the active power and the reactive power.
Average annual rate of change	The compound average growth rate, which is the year-over-year growth rate over a specified number of years.
Block loads	Large loads that are connected or disconnected from the network.
Bulk supply point	A substation at which electricity is typically transformed from the higher transmission network voltage to a lower one.
Connection point	A point at which the transmission and distribution network meet.
Coincident forecasts	Maximum/minimum demand forecasts of a connection point at the time of regional maximum/minimum.
Distribution network	The downstream part of the energy network that distributes energy directly to customers. This is generally at lower voltages than the transmission network.
Distribution system	A distribution network, together with the connection assets associated with the distribution network (such as a transformer), which is connected to another transmission or distribution system. Connection assets on their own do not constitute a distribution system.
Electrical energy	The average electrical power over a time period, multiplied by the length of the time period.
Electrical power	The instantaneous rate at which electrical energy is consumed, generated or transmitted.
Electricity demand	The electrical power requirement met by generating units.
Embedded Large Load	Industrial load that is connected to the distribution network and can materially increase the load as measured at a connection point, given the relative size of the load to the total demand at the transmission connection point.
Energy efficiency	Potential annual energy or maximum demand that is mitigated by the introduction of energy efficiency measures.
Generating unit	The actual generator of electricity and all the related equipment essential to its functioning as a single entity.
Generation	The production of electrical power by converting another form of energy in a generating unit.
Installed capacity	The generating capacity in megawatts of the following (for example): • A single generating unit. • A number of generating units of a particular type or in a particular area. • All of the generating units in a region. Rooftop PV installed capacity is the total amount of rooftop PV capacity installed at any given time.
Load	A connection point or defined set of connection points at which electrical power is delivered to a person or to another network or the amount of electrical power delivered at a defined instant at a connection point, or aggregated over a defined set of connection points.
Load transfer	A deliberate shift of electricity demand from one point to another.

Term	Definition
Maximum demand	The highest amount of electrical power delivered, or forecast to be delivered, over a defined period (day, week, month, season or year) either at a connection point, or simultaneously at a defined set of connection points.
Minimum demand	The lowest amount of electrical power delivered, or forecast to be delivered, over a defined period (day, week, month, season or year) either at a connection point, or simultaneously at a defined set of connection points.
National Electricity Market (NEM)	The wholesale exchange of electricity operated by AEMO under the National Electricity Rules.
Network service provider (transmission – TNSP; distribution – DNSP)	A person who engages in the activity of owning, controlling or operating a transmission or distribution system and who is registered by AEMO as a Network Service Provider.
Network Meter Identifier (NMI)	A unique identifier for connection points and associated metering points used for customer registration and transfer, change control and data transfer.
Non-coincident forecasts	The maximum or minimum demand forecasts of a connection point, irrespective of when the system peak occurs.
Operational demand	Operational demand is all demand met by local scheduled generation, semi-scheduled generation, and non-scheduled wind/solar generation of aggregate capacity ≥ 30 MW, and by generation imports to the region, excluding the demand of local scheduled loads. For more definitions, see AEMO, Demand Terms in EMMS Data Model, 2021, at https://www.aemo.com.au/-/media/Files/Electricity/NEM/Security_and_Reliability/Dispatch/Policy_and_Process/Demand-terms-in-EMMS-Data-Model.pdf.
Probability of exceedance (POE) of maximum or minimum demand	The probability, as a percentage, that a maximum or minimum demand (MD) level will be met or exceeded (for example, due to weather conditions) in a particular period of time. For example, for a 10% POE maximum demand in any given season, there is a 10% probability that the corresponding 10% POE projected maximum demand level will be met or exceeded. This means that 10% POE projected maximum demand levels for a given season are expected to be met or exceeded, on average, one year in 10.
Power factor	The ratio of the active power to the apparent power at a metering point.
Reactive energy	A measure, in volt-ampere reactive (var), of the alternating exchange of stored energy in inductors and capacitors, which is the time-integral of the product of voltage and the out-of-phase component of current flow across a connection point.
Reactive power	The rate at which reactive energy is transferred. Reactive power is a necessary component of alternating current electricity which is separate from active power and is predominantly consumed in the creation of magnetic fields in motors and transformers and produced by plant such as: • Alternating current generators • Capacitors, including the capacitive effect of parallel transmission wires • Synchronous condensers.
Region	An area determined by the Australian Energy Market Commission (AEMC) in accordance with Chapter 2A of the National Electricity Rules.
Residential and commercial load	The annual energy or demand relating to all consumers except large industrial load. Mass market load is the load on the network, after savings from energy efficiency and rooftop PV output have been taken into account. Includes light industrial load.
Rooftop photovoltaic (PV) systems	A system comprising one or more photovoltaic panels with the total nominal capacity of up to 100 kilowatts (kW), installed on a residential or commercial building rooftop to convert sunlight into electricity.
PVNSG systems	Commercial PV "non-scheduled" generation systems with the nominal capacity between 100 kW and 30 megawatts (MW).

Term	Definition
Shoulder	Unless otherwise specified, refers to the period 1 April – 31 May and 1 September – 31 October (for all regions except Tasmania), and 1 March – 31 May and 1 September – 30 November (for Tasmania only).
Summer	Unless otherwise specified, refers to the period 1 November – 31 March (for all regions except Tasmania), and 1 December – End of February (for Tasmania only).
Transmission network	A network within any National Electricity Market (NEM) participating jurisdiction operating at nominal voltages of 220 kV and above plus: (a) any part of a network operating at nominal voltages between 66 kV and 220 kV that operates in parallel to and provides support to the higher voltage transmission network (b) any part of a network operating at nominal voltages between 66 kV and 220 kV that is not referred to in paragraph (a) but is deemed by the Australian Energy Regulator (AER) to be part of the transmission network.
Transmission Node Identity (TNI)	Identifier of connection points across the NEM.
Transmission system	A transmission network, together with the connection assets associated with the transmission network (such as transformers), which is connected to another transmission or distribution system.
Winter	Unless otherwise specified, refers to the period 1 June – 31 August (for all regions).