ACIL ALLEN CONSULTING

REPORT TO

**AUSTRALIAN ENERGY MARKET OPERATOR**

26 JUNE 2013

# CONNECTION POINT FORECASTING

## A NATIONALLY CONSISTENT METHODOLOGY FOR FORECASTING MAXIMUM ELECTRICITY DEMAND

ACIL ALLEN CONSULTING PTY LTD
ABN 68 102 652 148

LEVEL FIFTEEN
127 CREEK STREET
BRISBANE QLD 4000
AUSTRALIA
T+61 7 3009 8700
F+61 7 3009 8799

LEVEL TWO
33 AINSLIE PLACE
CAN BERRA ACT 2600
AUSTRALIA
T+61 2 6103 8200
F+61 2 6103 8233

LEVEL NINE
60 COLLINS STREET
MELBOURNE VIC 3000
AUSTRALIA
T+61 3 8650 6000
F+61 3 9654 6363

LEVEL ONE
50 PITT STREET
SYDNEY NSW 2000
AUSTRALIA
T+61 2 8272 5100
F+61 2 9247 2455

SUITE C2 CENTA BUILDING
118 RAILWAY STREET
WEST PERTH WA 6005
AUSTRALIA
T+61 8 9449 9600
F+61 8 9322 3955

ACILALLEN.COM.AU

**FOR FURTHER INFORMATION PLEASE CONTACT:**

Jeremy Tustin
Principal
Tel: (03) 8650 6027
Mob: 0421 053 240
Email: j.tustin@acilallen.com.au

Jim Diamantopoulos
Principal
Tel: (03) 8650 6060
Mob: 0403 927 398
Email: j.diamantopoulos@acilallen.com.au

# Contents

## List of figures

## List of tables

# Definitions and short forms

| Shortened form | Meaning |
|---|---|
| ABS | Australian Bureau of Statistics |
| AEMO | Australian Energy Market Operator |
| Baseline forecast | An interim step in the methodology described in this report is to produce forecasts that have not been reconciled to an independently prepared system level forecast or had any post model adjustments applied to them. Baseline forecasts are prepared by applying a growth rate to a starting point. They typically do not reflect expectations relating to economic activity or other drivers of electricity demand or applicable policy interventions. |
| Block Load | An identified, step change in demand, either positive or negative, attributable to a specific project or customer. |
| BOM | Bureau of Meteorology |
| Coincidence factor | The ratio of demand at a connection point at the time of system wide maximum demand to demand at the same connection point at its maximum. A coincidence factor can take a value between 0 and 1. |
| Coincidence factor | A ratio calculated between two network elements the coincidence factor is the ratio of demand at a lower element when it peaks to demand at the lower element when the higher element peaks. See Appendix D for a more detailed description and a worked example. |
| CP | A connection point, or CP, is the physical point at which the assets owned by a TNSP meet the assets owned by a DNSP. While there is a physical point at each connection point where the transmission and distribution networks meet the particular asset where this happens is not necessarily the same in all cases therefore we do not describe it physically.<br><br>CPs is sometimes referred to as terminal stations and bulk supply points. |
| Demand | Is the quantity of electricity that flows through a connection point at any given time. It is measured in watts, generally expressed megawatts (MW), though it can be converted to megavolt amperes (MVA) using a power factor.<br>For the purposes of this report demand excludes electricity that need not be transferred through a connection point such as electricity generated by embedded generators. It also excludes electricity that would be used but is not because of demand side participation, the failure of network infrastructure or for some other reason.<br>Demand exceeds the quantity of electricity that is sold to end users, including by distribution network losses.<br>Strictly speaking demand is an instantaneous concept, though it is typically measured in 15 or 30 minute intervals.<br>This definition is similar to 'operational demand' as defined by AEMO, though the treatment of losses may differ depending on the level of the network at which the forecasts are made. |
| Demand forecasts | Forecasts of the maximum demand that would be observed at a CP in a given year if standard conditions are experienced. Those conditions are typically expressed in POE, usually 10, 50 or 90. |
| DEWHA | Department of Environment, Water, Heritage and the Arts (former Commonwealth Government department) |
| DNSP | Distribution Network Service Provider |
| Embedded generation | Electricity generation plant connected to the distribution network. |
| Energy | Is average demand over a period of time longer than 30 minutes, often a year. Energy is measured in Wh (usually GWh at the CP level) though it could also be measured in Joules. |
| kW | Kilo watt, one thousand watts. |
| MVA | One million Volt Amperes. |
| MW | One million Watts. |
| MWh | Megawatt hour, one million watt hours. |
| National | means in all regions of the NEM |
| NEM | National Electricity Market, including Queensland, New South Wales, the Australian Capital Territory, Victoria, Tasmania and South Australia. |
| Normal weather conditions | The weather conditions that would be expected to occur at the POE level being discussed or applied at the time. For example, in the context of 50 POE forecasts, normal weather conditions usually means the temperature that would be met or exceeded only one year in two. |
| NSLP | Net System Load Profile. The aggregate small customer load profile used to settle the wholesale electricity market in jurisdictions without interval data collected from all second tier customers. |
| NSP | Network Service Provider, either a DNSP or TNSP. |
| OLS | Ordinarly Least Squares. A method of estimating regressions. |

| Shortened form | Meaning |
|---|---|
| Organic growth | Growth in electricity demand that is not attributable to a block load. This is attributable to factors including small scale urban infill or to changes in the electricity demand of existing customers over time. |
| POE | Probability Of Exceedence. Typically, actual maximum demand is standardised to either, or both, of 10% and 50% POE levels.<br>The 50 (10) POE demand level is the level of maximum demand that, on average, would be exceeded in 50% (10%) of seasons. It can be thought of as the maximum demand that would be observed or exceeded once every two (ten) years on average.<br>The key driver of variability in demand is usually weather. However this is not always the case and the concept of POE is not tied directly to weather. |
| Power factor | is the ratio of demand expressed in MW to demand expressed in MVA |
| Region | A region of the NEM, South Australia, Victoria, Tasmania, New South Wales (including the Australian Capital Territory) or Queensland. |
| RRN | Regional Reference Node |
| R-squared | The coefficient of determination, a measure of the goodness of fit of a regression. |
| Solar PV | Solar Photovoltaic. A technology commonly used to generate electricity by domestic customers for their own use and for export to other customers. |
| Statistical Local Area | The Statistical Local Area (SLA) is an Australian Standard Geographical Classification (ASGC) defined area. SLAs are Local Government Areas (LGAs) or part thereof. |
| Switching | Temporary changes in network configuration made by a NSP for operational reasons. |
| TNSP | Transmission Network Service Provider |
| Transfers | Permanent (or indefinite) changes in network configuration made by an NSP usually to manage demand growth. |
| ZSS | Zone substation |

# Structure of an electricity network

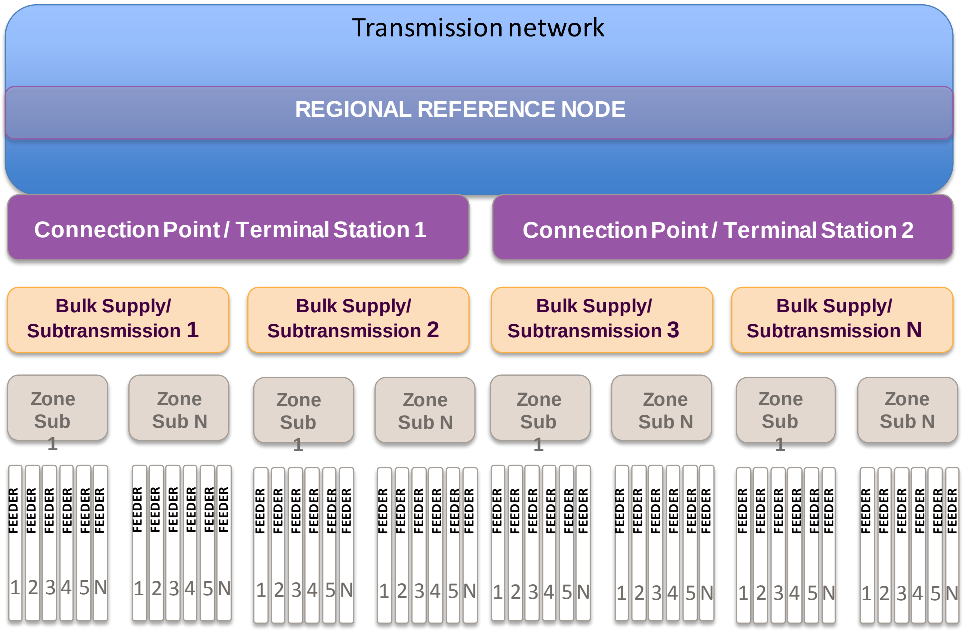
This report refers to an electricity network configured as follows and as illustrated in Figure A

- Electricity is generated and transferred on a transmission network at high voltage
- A transmission network meets a distribution network at a connection point
- A distribution network may transfer electricity at a voltage lower than transmission voltage to a subtransmission station, but not necessarily
- A distribution network transfers electricity from either a subtransmission station (if used) or a connection point to a zone substation (ZSS) at a voltage lower than the transmission (or subtransmission) voltage
- A distribution network transfers electricity to small customers at a further reduced voltage on a feeder and connection assets
- Larger customers may connect to the distribution network at higher voltages than small customers or to the transmission network.

In addition, we take account of the fact that the wholesale electricity market is *notionally*, but not *physically*, settled at the regional reference node (RRN) in each region. As the Australian Energy Market Operator (AEMO) says in "*Treatment of Loss Factors in the National Electricity Market*"[1], there are transmission losses between the RRN and each Connection Point (CP). However, electricity is not *physically* transferred this way, so some of these losses are negative. It is important for reconciliation purposes to account for these losses consistently, though the actual treatment is not important. That is, the CP forecasts and the system forecast to which they are reconciled could either be 'adjusted' to the RRN or applied at the CPs themselves. What is important is that both are the same.

[1] AEMO, "Treatment of Loss Factors in the National Electricity Market", 1 July 2012, available from http://www.aemo.com.au/Electricity/Market-Operations/Loss-Factors-and-Regional-Boundaries/Treatment-of-Loss-Factors

FIGURE A TYPICAL HIERARCHY OF DNSP'S ELECTRICITY DISTRIBUTION NETWORK



# Preface - Connection point maximum demand forecasting methodology

The following is a description of the proposed methodology for forecasting maximum electricity demand at the Connection Point (CP) level. Further detail regarding each step is provided in the body of the report.

The methodology outlined here will produce forecasts of maximum demand under weather normalised conditions. That is, forecasts of what demand would be if certain weather conditions occur. Typically, weather normalised conditions are discussed at the 10%, 50% and/ or 90% Probability of Exceedence (POE) level, though the methodology can be used to produce forecasts at any POE level.

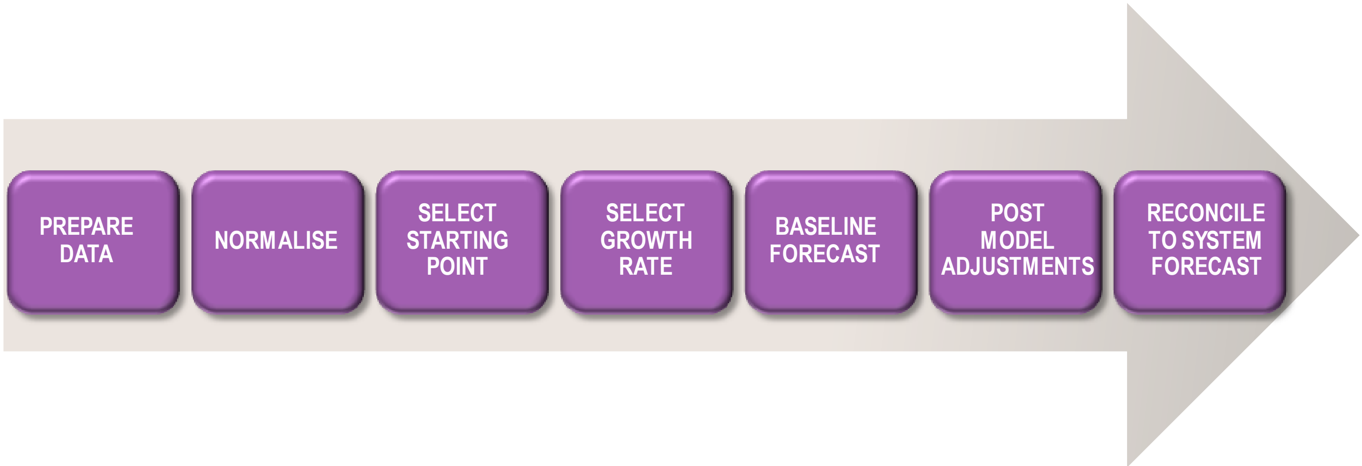
Forecasts of this type are an input into the network planning and related processes of Network Service Providers (NSPs). The typical forecast horizon is ten years.

It is assumed that the forecasts are to be in MW. However, the same methodology could be used to produce forecasts in MVA, though in this case the forecaster would begin with data in MVA.[2]

## Overview of methodology

The methodology consists of seven steps as depicted in Figure B

FIGURE B **FORECASTING METHODOLOGY**




A summary of each step follows below. The post model adjustments referred to in step 6 should be conducted after the forecasts are reconciled to the system forecast. They may be unnecessary if the relevant factors are included in the system forecast. This is discussed in more detail below.

## Data preparation

The first step in the methodology is to collect the necessary data and manage it appropriately.

2 Of course, it is also possible to convert from MW to MVA or vice versa using power factors.

Data relating to historical maximum demand and weather conditions should be collected along with metadata relating to the CPs for which forecasts are to be prepared to assist with judgements that must be made during the forecasting process. Those judgements would also benefit from historical data and projections of likely drivers of electricity demand, which should be collected as well.

More detail regarding data collection and management is provided in section 3, which relates to data requirements and sources; and chapter 4, which relates to data processing and dataset preparation.

### Demand data

The first data requirement is a time series of high frequency data (15 or 30 minute interval) for each CP to be forecast. Ideally this time series should go back for at least 10 years.

We envisage that the forecasts are to be produced in MW, so the data should also be in MW. We also envisage that the data would be collected at the CP level so no adjustment for losses is necessary. If data were collected at a different level of the network (see Figure A) then an adjustment would be needed to account for losses between that level and the CP.

These data should be well understood and should relate closely to what is being forecast. Three factors that may require adjustments to the historical data should be considered:

1. Network configuration
2. Block loads
3. Output of embedded generation

A more detailed discussion of the approach to accounting for changes in network configuration and block loads is provided in the context of data cleansing in sections 4.1 and 4.2.

#### Network configuration

First, the historical data should reflect the *current* network configuration. From time to time, loads may be transferred from one CP to another, either temporarily or permanently.

From the perspective of the demand data, temporary *switches* ‘obscure’ history while permanent *transfers* ‘change’ history.

In both cases the historical data that were observed do not relate to what is to be forecast. Both issues should be addressed in the data set by either adjusting the data to reverse the changes if possible or by deleting the ‘tainted’ observations from the data if not.

Reversing the changes could draw on a detailed bottom-up rebuild of the historical data or could be based on metadata regarding transfers that have been made. That is, historical data at a very granular level, such as feeders, could be built up to the CP level to reflect *current* network configuration. This would ameliorate most of the problems associated with switches and transfers. However, it requires a detailed database and will not always be possible.

‘Tainted’ observations could be disregarded based on their values, with inexplicably high or low values deleted from the dataset (these may also be errors). There is no hard and fast rule to identify ‘inexplicable’ variability. A workable rule would be to examine all observations that lie more than two standard deviations from the mean[3] and try to identify reasons why this is the case. Another approach is to produce a scatterplot of the data and delete data points that appear not to belong to the series.

#### Block loads

The data set should also be adjusted for the impact of substantial block loads in history.

Block loads cause demand data to 'step' either up or down depending on the circumstances. In either case they can obscure the underlying growth pattern in the data.

Drawing on metadata regarding the CP, adjustments should be made to remove the impact of block loads from the historical data.

This should relate to block loads that are introduced or removed from the historical data permanently as well as those which are temporary. For example, if a large industrial facility shut down for maintenance during the summer period one year, maximum demand observed at the CP will be reduced in that year. This should be accounted for in the forecasting process to avoid treating the temporary shutdown as a permanent shift in demand.

Adjustments should generally only be made for block loads that are large relative to the total load supplied by the CP. For example, a load that represents more than five per cent of the load on a CP would typically warrant an adjustment, but a smaller load may not. The smallest load for which an adjustment should be made to CP data will usually be much larger than a load that would warrant an adjustment at the zone substation level.

Another case where an adjustment may be warranted is if a load is unique for the CP in question. For example, the first mine ever to be built in an area may warrant an adjustment even if it is small relative to total load at the CP as it is unlikely to be reflected in the historical growth pattern at that CP.

#### Embedded generation

In some places electricity generators are 'embedded' or connected directly to the distribution network. Generally speaking, these generators cause demand as measured at the CP to be less than demand measured at the customer's meter because some of the latter is supplied from within the network.

The appropriate treatment of embedded generation depends on the purpose of the forecasts. If the objective is to obtain a projection of the actual demand for electricity in an area then the location of the generator that meets that demand is immaterial. In this case, demand as measured at the CP will need to be adjusted (increased) by adding the output of embedded generators to historical data.

On the other hand, the objective may be to determine the appropriate size of a CP. In this case the objective is to determine the maximum demand that will be measured at the CP. In practice this will be reduced by the output of the embedded generator and therefore it *might* be appropriate not to add back that output. However, two further issues need to be considered.

First, for weather normalisation it will be necessary to use total demand, so the output of embedded generators should be added back for weather normalisation purposes.

[3] This should be considered *after* correcting for weather effects.

Second, in planning the appropriate size of the network it will be important to consider the question of certainty of output of the embedded generator.

Generally, if the NSP can be certain that the embedded generator will generate a particular amount at the time of system peak then it might be appropriate to plan a smaller CP asset to account for this. Therefore, it might be appropriate to adjust the forecast downward by the same amount. However, if that certainty is not provided, it would typically be appropriate to add back the output of embedded generators. That certainty could be based on observations of the generator's behaviour in the past or on a load support agreement between the generator and the relevant NSP.

Small embedded generators may require different treatment again. Some of these, notably solar photovoltaic (PV) systems, are passive in the sense that they do not need to be 'turned on' by an operator. Simply put, if it is sunny they will generate, if not, they won't. This makes the output of solar PV systems more predictable, and therefore reliable, than the output of other forms of embedded generation.

Therefore, while solar PV systems are not free of uncertainty, they are not subject to the same type of certainty as a thermally powered embedded generator whose operator has discretion over whether or not to 'turn on'.

A further characteristic of solar PV systems is that, individually, they are small relative to total load and have recently become very numerous. This reduces the uncertainty associated with them.

This may not be true of all forms of small embedded generation in future, though. For example, if fuel cells become common, consideration may need to be given to the incentives on their operators and the certainty that can be assigned to their output.

At present, the dominant form of small embedded generation in the NEM is solar PV and, generally speaking, no adjustment to 'add back' solar output would be necessary.

### Weather data

The next dataset to collect is weather data for normalisation. As discussed in section 3.2, daily maximum and minimum ambient temperatures should be obtained as well as other weather variables, such as rainfall, humidity etc.

The Bureau of Meteorology (BOM) publishes weather data for many weather stations in Australia and one must be chosen and assigned to each CP for normalisation. A discussion on choosing the appropriate weather station is provided in section 3.2.

The purpose of using weather data is to correct for changes in demand that are partly behavioural.[4] Therefore, it may be appropriate to use data from the weather station that is used for local weather reporting, i.e. the station for which forecasts and actuals are reported in local news.

The choice of weather station and variable is empirical. It would be appropriate to test data from several stations to identify which is most closely correlated to demand at each CP, which may reflect microclimatic influences etc.

Once a weather station has been identified as appropriate to use in forecasting demand at each CP, it is unlikely to change, although the analysis should be revisited from time to time to confirm that it is still the best alternative.

4 That is, customers may respond to the weather forecast as well as the actual weather.

The weather data should be collected for a very long time series, thirty years or more. It does not matter that this will be longer than the time series of demand data as a longer time series is required for weather normalisation.

Weather data will typically have some missing observations. These should be imputed as discussed in section 4.3.1. If more than about five per cent of the observations are missing it would be more appropriate to use a different weather station if a suitable alternative exists.

### Other data

A variety of other data may be useful in forecasting demand at CPs. Some of these data will not be used formally, but may assist in making the various judgements that must be made along the way. These are discussed in section 3.4 and include data relating to:

1. Planned transfers between connection points (for either post model adjustments or altering network configuration)
2. Changes in relationship between drivers and demand due to:
   a) changed use of demand management
   b) ongoing uptake of embedded generators, in particular solar PV systems
   c) block loads, historical and future
   d) regional economic and population data
3. Metadata concerning CPs, such as:
   a) industrial/ commercial/ residential mix of customers
   b) nature of industry/ commerce.

### Data sources

It is important that the data used in preparing forecasts are accurate, sourced transparently and cannot be said to have been chosen selectively to the forecaster's benefit.

The best approach is to take historical data from publicly available sources such as the Australian Bureau of Statistics (ABS) and the BOM. Of course some data are only available from NSPs. Those data should be taken from the most accurate source possible according to a consistent process, such as giving preference to data from revenue meters over other sources. The sources should be accurately described. This is discussed further in section 3.

## Normalisation

The objective of demand forecasting is not to forecast what *actual* electricity demand will be in any given year.

Rather, the objective is to forecast what demand *would be* under 'normal' conditions.

This requires that the random element of demand is 'normalised' out of historical data before forecasts can be produced.

In doing this it should be noted that electricity demand varies based on drivers, day type, weather and other factors. Drivers are accounted for econometrically (partly through the reconciliation stage). The remaining factors are addressed at the normalisation stage.

Generally, when drivers and day type have been accounted for, the main remaining source of variability in demand is weather. Therefore, normalisation is commonly referred to as weather normalisation and a link is often drawn between normalised demand and normalised weather conditions.

In some cases the key source of variability may be something other than weather. For example, in some areas the key source of variability in demand may be water pumping load, which may be only loosely connected to temperature. In these cases it may be more appropriate to normalise for this other factor. However, the ability to do so depends on identifying that other factor and appropriate data to measure it.[5]

The weather normalisation procedure comprises four steps, which are discussed in more detail in sections 5.2 and 5.3:

1. prepare the dataset for normalisation (section 5.2)
2. estimate the relationship between temperature and demand at the CP (section 5.2)
3. create a distribution of maximum demands for each CP for each year (section 5.3)
4. identify 'normal' maximum demand from that distribution (section 5.3).

The procedure is performed separately for each season for which demand forecasts are required, typically winter and summer.

### The dataset to use for normalisation

The appropriate dataset to use for normalisation is a subset of the demand data collected at stage 1. Generally, it should:

1. reflect only the season of interest, i.e. summer or winter
2. be one year's data unless conditions were very mild or very extreme
3. be truncated to remove:
   a) demand on 'mild' days
   b) demand on non-working days.

#### The season of interest

Demand for electricity is largely due to heating or cooling and therefore varies between seasons. The first step is to narrow the dataset so that it relates to only one season. This raises the question of how to define the seasons of interest. This is particularly relevant for summer, which is sometimes treated as extending from November to March.

This should be resolved at the local level. The objective is to ensure that the summer data capture summer peaks. If these occur routinely in November, or in March, those months could be added to the summer season. However, in doing so, thought should be given to whether the relationship between demand and temperature in these months is the same as it is in other summer months.

The same issues should be considered in relation to winter demands.[6]

---

[5] In the specific case of water pumping load, it may be most appropriate to treat it as a block load by deleting it from the historical series and adding projections of it to the forecast.

[6] The objective of weather normalisation in winter is to ensure that the data capture winter peaks. However, these tend to occur more reliably in 'calendar' winter so adding months to the season is typically not necessary.

### One season unless very mild

Generally speaking, the data used for normalisation should relate to only one season (that is, only one year's summer or winter). The exception to this is if the season in question was very mild. In this case there may not be enough data to describe the relationship between temperature and demand at extreme conditions accurately.

The solution to this is to pool data. That is, combine the data from more than one consecutive season. Pooling increases the number of observations that are available, but it is not without problems. When data are pooled, an allowance must be made for underlying growth.

However, no allowance is made for changes in the relationship between weather and demand.[7] For this reason, data should only be pooled over two or three years.

### Truncation

The third step is to truncate the dataset to remove days when weather was 'mild' and non-working days.

At mild temperatures, demand is relatively unresponsive to changes in temperature. As temperature increases (in summer) demand begins to rise before reaching a point where it will not increase (much) further regardless of temperature.

'Mild' days should be removed from the dataset. The reason is that at mild temperatures the relationship between temperature and demand is likely to be different than at 'hot' or 'cold' temperatures, and these days play no role in determining maximum demand.

The threshold between mild and 'hot' or 'cold' temperatures is identified by examining a scatterplot of maximum demand data to identify the temperature at which the relationship between temperature and demand begins to increase.

On non-working days such as weekends and public holidays,[8] load is typically less than it would be on working days. Including non-working days in the dataset without accounting for them explicitly will bias the relationship between temperature and demand.

An example of the truncation process is provided in section 5.2.1.

## Estimate the relationship between demand and weather

Weather normalisation begins with data for an entire season in a single year. The necessary data are:

1. observed daily maximum demand
2. weather.

Weather may be daily maximum and minimum temperature, the average of these or another variable. This could be examined empirically to identify which is most closely correlated with demand at the CP in question. It may change between CPs and between seasons. For example winter demand may be correlated with temperature at 6:00PM rather than daily (overnight) minimum.

---

7 Making this type of allowance is possible, but increases the complexity of the model considerably.

8 There may be CPs in strongly tourism related areas at which load on weekends or public holidays is not less than load on other days. This is because, in those areas, these days are not non-working days. In these cases it may be more appropriate to truncate other days or not to truncate.

If Maximum and Minimum demand are used,[9] compute a linear regression of the following form:

$$MD_d = m * MAXtemp_d + n * MINtemp_d + c + \varepsilon$$

where $MD_d$ is maximum demand observed on day d for all days in the dataset MAXtemp and MINtemp are daily maximum and minimum temperature respectively *m, n* and *c* are regression parameters, and $\varepsilon$ is an error term.

Collect the coefficients of the regression model and the standard error of the estimate.

### Create a distribution of demands for each CP for each year

Use the coefficients of the regression describing the weather relationship and *all* of the available weather data to produce estimates of what daily maximum demand would have been in the most recent season under *all* historical weather conditions observed. If there are thirty years of weather data available and 75 working days in the summer[10], this would produce 2,250 fitted daily maximum demand values.

Notionally these suggest what demand would have been in 2012/13 for each weather outcome that has been seen in the last 30 years. 30 of these are seasonal maximum demand values, that is, one maximum for each year of weather data. Those 30 annual maximum demand values are recorded.

Having done this, apply the standard error of the regression randomly to the initial 2,250 notional demand observations. Every time this is done another set of 30 annual maxima is produced. Therefore, if 100 trials are conducted, the result will 225,000 simulated daily demands, of which 3,000 are annual maxima.

### Identify 'normal' maximum demand from that distribution

There are now a large number of simulated demands for the season. The final step of the normalisation process is to identify the 'normal' maximum demand for the season from that distribution of simulated demands.

This is simply a matter of collecting demand at the desired percentile. For example, 50 POE demand is the 50th percentile of the 3000 demands produced in the previous step. 10 POE demand is the 90th percentile etc.

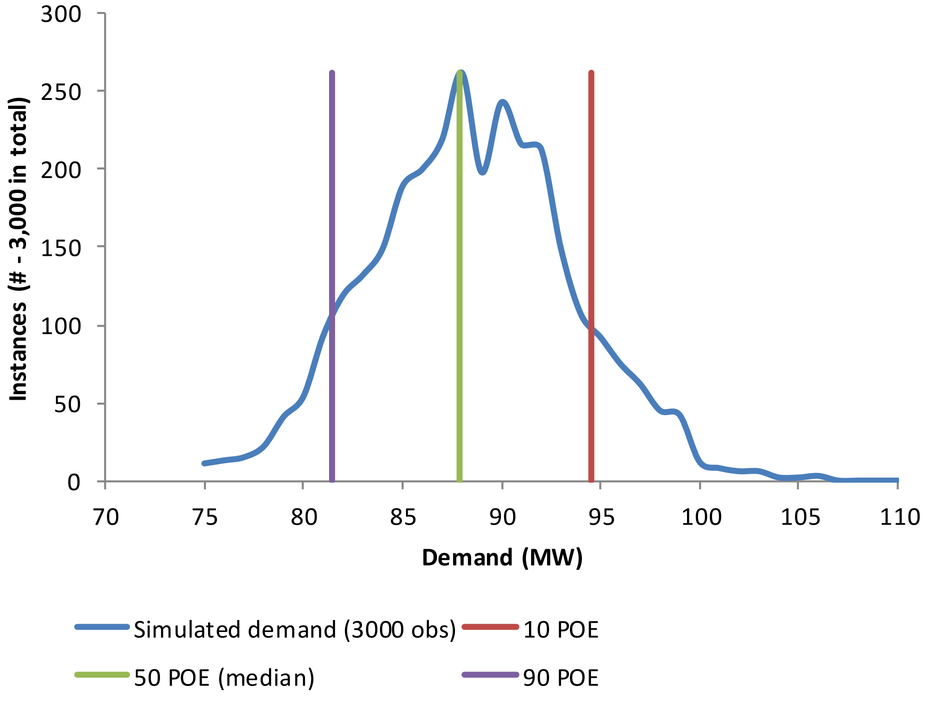
Any POE level can be taken from the distribution of simulated demands.

---

9 In our experience this approach accounts for the majority of weather related variation. More complicated models relating demand to other weather variables can be built, but there is a question of assessing the extra benefit of doing so against the additional cost.

10 Note that weather data are used for all non-working days while demand data are deleted for mild days. Weather data could also be truncated to remove mild days, though this would not alter the analysis as these days are extremely unlikely to generate peak demand.

FIGURE C **DISTRIBUTION OF SIMULATED ANNUAL MAXIMUM DEMANDS**




*SOURCE: ACIL ALLEN CONSULTING*

Figure C illustrates the result of simulating 225,000 daily maximum demands and 3,000 annual maxima with mean 88 MW and standard deviation 5 MW. In this case the 10, 50 and 90 POE weather corrected demands are as shown in Table A. Note that the simulation process is random. If it was repeated the values would be slightly different.

TABLE A **WEATHER CORRECTED DEMAND AT 10, 50 AND 90 POE**

| POE level | Demand (MW) |
|---|---|
| 10 | 94.5 |
| 50 | 87.9 |
| 90 | 81.4 |

## Selecting the starting point

When the historical data have been weather corrected the starting point for the forecasts can be selected. Conceptually, this is the weather normalised demand in the last year for which actual data are available.

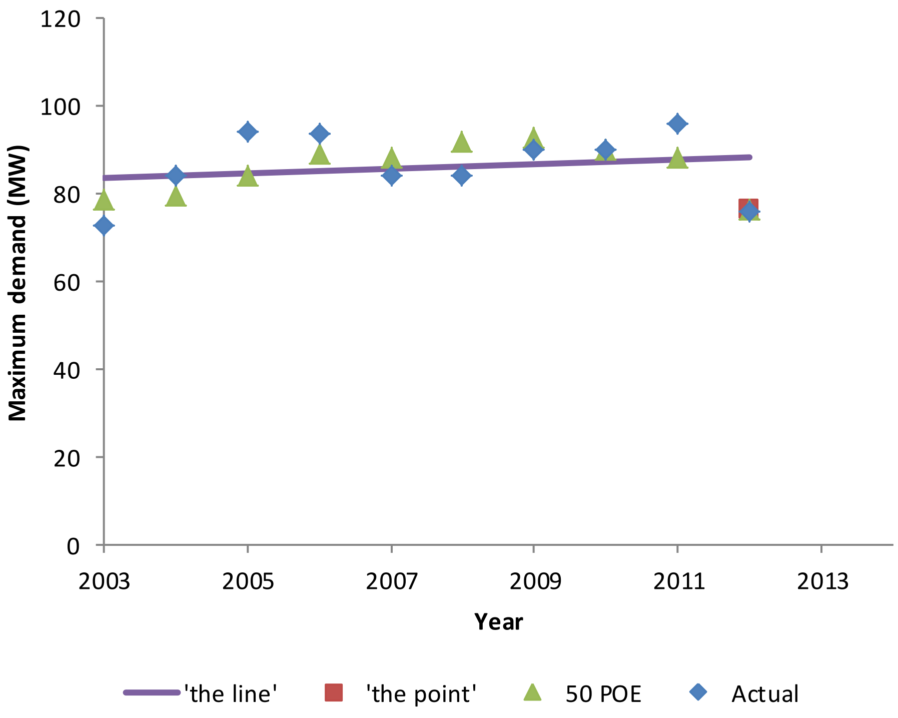
Practically, two options are available and a judgement must be made.

The options are to define the starting point (see Figure D for illustration):

— 'off the point' taking the simulated 50 (or 10) POE value for the last available year

— 'off the line' taking the value off a regression line fitted to the weather normalised history (discussed below).

FIGURE D **WEATHER NORMALISATION OFF THE POINT OR OFF THE LINE**




Starting the forecast 'off the line' may force a step change in demand between the last actual year and the first forecast year. Whether this is appropriate depends on the reason for that step change.

In the example given above, which is based on synthesised data, the observed maximum demand in the last year for which data were available was approximately 76 MW. This was significantly below the maximum demand observed in the previous year and other recent years.

This was very close to a 50 POE year, so the weather normalised 50 POE is within one MW of the actual.

However, the drop in maximum demand from the previous year was substantial so the value on 'the line' for 2012 is approximately 88 MW, substantially above the observed value.

The question to consider is whether the difference between 76MW and 88MW (i.e. from 'the point' to 'the line') is attributable to weather or whether there is another reason for it such as a worsening economic environment or substantial price response.

If no valid reason to support starting 'off the point' can be found, it is reasonable to conclude that the difference between the 'point' and the 'line' is due to randomness in the data. In this case, starting 'off the line' is preferred over assuming that the same random outcome will be repeated in every forecast year.

Therefore, generally speaking if the two options are close to one another the preferred approach is to take the starting point 'off the line'. However, if the line and point are 'far' from one another the preferred option is to start 'off the point' unless no reason can be found to justify doing so.

It is also important to remember that this choice will be rendered largely obsolete by the reconciliation process.

Further detail is provided in section 6.1.

**Starting point – developing 'the line'**

For a starting point to be taken 'off the line' a regression must be developed. This is a regression of the weather normalised historical maximum demands, with block loads removed, on either:

1. customer numbers (or population as a proxy)
2. a time trend.

There is no theoretical basis to prefer one of these regressions over the other. Therefore, both regressions should be estimated and the choice made empirically.

The forecaster should consider whether the coefficient in the population regression makes theoretical sense, in particular whether it has the expected sign.[11] If not, it should be disregarded. There is also a strong likelihood that a model using only population as an explanatory variable will be mis-specified if other drivers are significant. In this case it would be more appropriate to use the time trend.

If the population regression has the expected sign, the goodness of fit of both regressions should be considered. Preference should be given to the variable that produces the best fit.

In some cases, especially areas consisting primarily of industrial customers, the coefficient on either regression may be very small, indicating little or no growth. In these cases, growth may be accounted for through identifiable block loads added to the forecasts later.

## Select the initial growth rate

Growth rates are chosen based on the regression developed in choosing the starting point, though again some judgement is required.

The first step is to use the growth rates implied by the regression used to develop 'the line' and therefore the starting point.[12] That is, apply either a projection of customer numbers or a continuation of the time trend to develop (baseline) forecasts.

The second step is to sense check this projection with local experts. Areas to take care of are:

1. projections of growth faster than four per cent per annum – even if growth has occurred at this speed in the past, it will not necessarily be sustained. A local planner will know whether growth at this rate is plausible
2. loads known to be flat – fluctuations in demand in history may correlate with a time trend or customer numbers but lead to implausible forecasts, especially in areas dominated by commercial and industrial loads. Local planners will know whether the expected changes in these loads are likely to support the modelled growth rates
3. loads where the 50 POE demands grow faster than 10 POE demands – this is unlikely to occur, especially in summer. This would also imply that, at some time, 50 POE demand will exceed 10 POE demand, which is not possible

[11] There is no theoretical basis to predict the sign of the coefficient in the trend regression. That is, demand might increase or decrease over time.

[12] This should be considered even if the starting point is taken 'off the point'.

4. distortions in growth rates arising from undetected load transfers across and within seasons or arising from errors in the data.

In these cases it may be necessary to modify the growth rate suggested by the regressions or to substitute a growth rate selected manually, for example by substituting the growth rate from a nearby CP or from the system forecast. The decision to do this, and the reasons the particular changes were made, should be recorded.

Further detail is provided in section 6.2.

## Baseline forecasts

At this stage baseline forecasts can be computed by applying the growth rate[13] to the starting point and adding anticipated block loads and future network transfers.

It may be appropriate at this point to make adjustments to these forecasts to account for policy changes, though this will depend on the nature of the policy and the available data. More detail on this is provided below.

Until the post model adjustments are made and the forecasts are reconciled to an independently prepared system forecast the forecasting process is not complete. In particular, the forecasts do not yet reflect the likely impact of drivers of electricity demand.

## Post model adjustments

The baseline forecasts are now adjusted to account for changes in demand that have not otherwise been accounted for in the methodology.

This is to ensure that factors that *are* expected in future but *are not* incorporated into the system or baseline forecasts are reflected in the final forecasts.

The most common post model adjustment is accounting for known block loads, either increments or decrements, to be made at given connection points.[14]

Perhaps the next most common cause of post model adjustment is a change in Government policy, or factors caused by such a change, that impact maximum demand. For example, the uptake of solar photovoltaic (PV) systems driven by the various feed-in tariffs and other policies that have been developed in the last five years would have justified a post model adjustment in the past, and may do so going forward. Another example is increased use of demand side management.

Another possible cause is increasing energy efficiency, though energy efficiency is essentially targeted at average demand. Its relationship with *maximum* demand is complex (see section 9.2 for a discussion of this issue).

Failure to make the necessary adjustments would result in forecasting on the wrong basis. That is, the maximum demand forecasts would be forecasts of what would happen if the PV systems were not installed or the demand side management did not occur.

---

[13] Note that the coefficient in the time trend regressions is not a growth *rate* but a linear coefficient. That is, it is not consistent with a constant percentage increase each year.

[14] Adding these block loads to the CPs where they are expected will improve the *allocation* of growth between CPs but, because it is done before reconciliation, will not allow *total* growth to exceed growth in the system forecast. This is important as applying block loads after reconciliation would lead to double counting because the block loads are 'in' the system growth.

The appropriate way to forecast the impact of a policy change will vary with the policy. The key issue is to focus on *changes* in demand attributable to the policy. For example the rise in solar PV systems has now been experienced for several years. Therefore, the starting point should *already* reflect the existing systems. So reducing historical demand to account for existing systems would be double counting. Similarly, it is possible that the chosen growth rate will also reflect the uptake of solar PV systems in recent years and that this uptake will remain constant. In this case, no post model adjustment would be required if no change in uptake is forecast.

On the other hand, if the growth rate is based on a long term trend it may understate the impact of solar PV systems and an adjustment may be appropriate. Similarly, if uptake of PV systems is expected to accelerate (or decelerate) an adjustment may be warranted.

Post model adjustments are inherently difficult to prepare and contentious. By their nature they rely on assumptions which are difficult or impossible to verify. They often rely on expectations of future Government policy, which can change. It is imperative that the adjustments made and the assumptions and methodology used to develop them are stated explicitly and available for review. To the maximum extent possible the forecast impacts should satisfy the same principles as the forecasts themselves.

## Reconciliation to system forecast

The final stage in the demand forecasting process is to reconcile the CP forecasts to an independently prepared system forecast. In this context ‘system’ refers to forecasts of maximum demand at the NEM region level.[15]

The purpose of doing this is to ‘import’ the likely impact of drivers of electricity demand into the CP forecasts. Some drivers, such as economic activity are not well understood or forecast at the CP level and lend themselves more to incorporation into a higher level forecast. Depending on their nature, some policy changes will be more appropriately modelled at the system level and ‘imported’ into the CP forecasts.

This report does not address how a system level forecast would be prepared. Rather, it is assumed that the CP forecasts will be reconciled to the system level forecast AEMO prepares (or commissions) from time to time.

More discussion of the reconciliation process is provided in section 8.2.

### Develop coincident forecasts and account for losses

The first step in the adjustment process is to bring the CP forecasts onto a common basis to the system forecast by accounting for diversity and losses between the CP and the RRN.[16] This reflects the fact that the demand at all CPs will not necessarily peak at the same time.

This is done by observing the ratio of maximum demand observed at each CP to the demand at each CP when system maximum demand occurs (coincident maximum demand).[17] The baseline forecast for each CP is scaled accordingly.

[15] Note that this is a notional forecast. There is no single physical point where this demand will be observed.

[16] We assume that the system forecast is computed at the RRN in each region. If it is not, it may also need to be adjusted for losses between the point where it is computed and the RRN.

[17] A more detailed discussion of this issue is in Appendix D.

The ratio should be considered over several years and thought given to the reason for any change that may be observed (bearing in mind that some change will be attributable to randomness). In many cases the ratio will be stable. However, if a trend is observable over time it may be appropriate to continue this through the forecast period. If not, the average of the ratio in recent years or the value observed in the last year for which actual data are available would be used.

**Identify necessary adjustments and calculate and apply ‘trim’**

When the difference between the system level forecast and the sum of the coincident CP forecasts is calculated, the extent of any necessary adjustment to the CP forecasts can be identified. This can theoretically be positive (an increase in the CP maximum demand forecasts) or negative (a decrease in the CP maximum demand forecasts), though typically it is negative (i.e. system forecast smaller than sum of CPs).

The system and CP forecasts are produced on different bases and can be expected to be different.

Generally speaking, as discussed in section 8.1, the system forecast is likely to be preferred over the CP forecasts. Therefore, the CP forecasts will generally be altered to match the system forecast. This will be particularly so if the CP forecasts are based on trend growth and the system forecast is based on a projection of economic activity and other macroeconomic factors not incorporated into the CP forecasts. Projected economic growth is likely to vary from year to year, causing the forecast rate of growth in demand to change as well, which will not be reflected in the trend based CP forecasts.[18]

If the CP forecasts are thought to be based on better or more up to date information than the system forecast and are therefore thought to be superior to it, the preferred option would be to revisit the system forecast to incorporate the best and most up to date information.

The forecaster should consider the difference between the CP and system forecasts as adjusted to the regional reference node (or other common point) and make a judgement about the appropriate extent of reconciliation or ‘trim’.

As a starting point, ‘trim’ should be allocated proportionally. That is, CPs with higher maximum demand should be ‘trimmed’ more than smaller CPs.

However, there may be exceptions for CPs where the likely future growth is well understood. In the case of CPs where demand is dominated by a small number of large industrial loads, planners may have a highly developed understanding of future maximum demand. This will already be reflected in the baseline forecasts through the choice of starting points and growth rates. These CPs should be quarantined from ‘trimming’ and the reasons for doing so recorded.

[18] A similar issue arises if the CP forecasts are based on a population projection that grows at a different rate(s) to economic growth.

# 1 Introduction

The role of the Australian Energy Market Operator (AEMO) includes operating as the National Transmission Planner. In this role it:

- oversees the strategic development of the national electricity grid
- delivers strategic gas and electricity planning advice
- develops forecasts which guide long-term investment in network infrastructure and resource management.

This includes modelling possible future scenarios, advising how the transmission network might develop under those scenarios, and leading the development of the transmission network to meet forecast energy requirements.

On 29 June 2012 AEMO released the first National Electricity Forecasting Report providing independent electricity demand forecasts developed on a consistent basis for the five National Electricity Market (NEM) regions. AEMO also published its associated input data, assumptions and methodology.

Those forecasts are used in a range of applications, including as inputs to the development of regional and Connection Point (CP) forecasts by Transmission Network Service Providers (TNSPs) and Distribution Network Service Providers (DNSPs), as well as the development and review of revenue reset applications and Regulatory Investment Tests for Transmission.

Looking forward, AEMO has developed a three year plan to develop electricity demand forecasts that enable holistic and coordinated decision-making for infrastructure investment across the electricity markets. This aligns closely with AEMO's mission to plan, develop and operate markets that support long-term investment in Australia.

In turn, this will assist AEMO in supporting the delivery of the National Electricity Market Objective, which is:[19]

> to promote efficient investment in, and efficient operation and use of, electricity services for the long term interests of consumers of electricity with respect to –
>
> a. price, quality, safety, reliability, and security of supply of electricity; and
>
> b. the reliability, safety and security of the national electricity system.

Among other things, AEMO's three year strategy includes:

1. publishing a methodology for CP maximum demand forecasting (at the transmission level) across the NEM and developing common definitions for the different types of gas[20] and electricity demand in 2013
2. developing CP maximum demand forecasts against AEMO's published methodology and publishing a methodology for regional gas forecasting for south-eastern Australia in 2013-14.

In early 2013 AEMO commissioned ACIL Allen to prepare the methodology referred to in point 1 above. ACIL Allen's advice to AEMO in relation to that methodology is

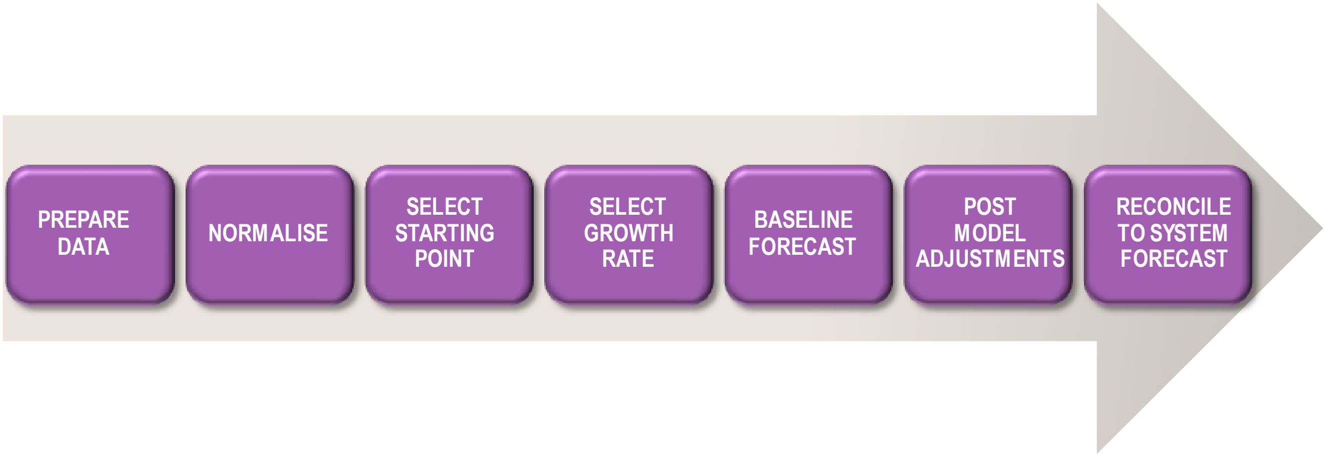
[19] National Electricity Law, s.7

[20] This report relates only to the electricity aspect.

contained in this report. AEMO will manage the implementation of that methodology, including the development of forecasts, in 2013-14.

The CP maximum demand forecasting methodology ACIL Allen recommends AEMO use consists of the steps shown in Figure 1.

FIGURE 1 **OVERVIEW OF CONNECTION POINT MAXIMUM DEMAND FORECASTING PROCESS**




This report contains:

- a list of definitions and shortened forms on page vi
- a summary of the CP maximum demand forecasting methodology in the preface, beginning at page ix
- a description of some broad principles of best practice forecasting which are relevant at the CP level in chapter 2
- a more detailed discussion of the proposed methodology including discussions of:
  - data collection and management in chapter 3
  - data processing and dataset preparation in chapter 4
  - the recommended method of weather normalisation in chapter 5
  - the starting point and growth rates in chapter 6
  - post model adjustments in chapter 7
  - the process for reconciling the baseline forecasts to an independently prepared system forecast in chapter 8.

In addition, there are several appendices.

- Appendix A provides a general discussion of the relationship between electricity demand and its drivers. These relationships are used to 'sense check' forecasts throughout the process
- The following two appendices summarise alternative methods of electricity demand forecasting. Specifically:
  - Appendix B provides a discussion of alternative methods of weather normalisation other than the method proposed in this report
  - Appendix C provides a summary of various approaches that can be taken to forecasting electricity demand
- Appendix D provides a discussion of coincidence factors and diversity between different levels of an electricity network

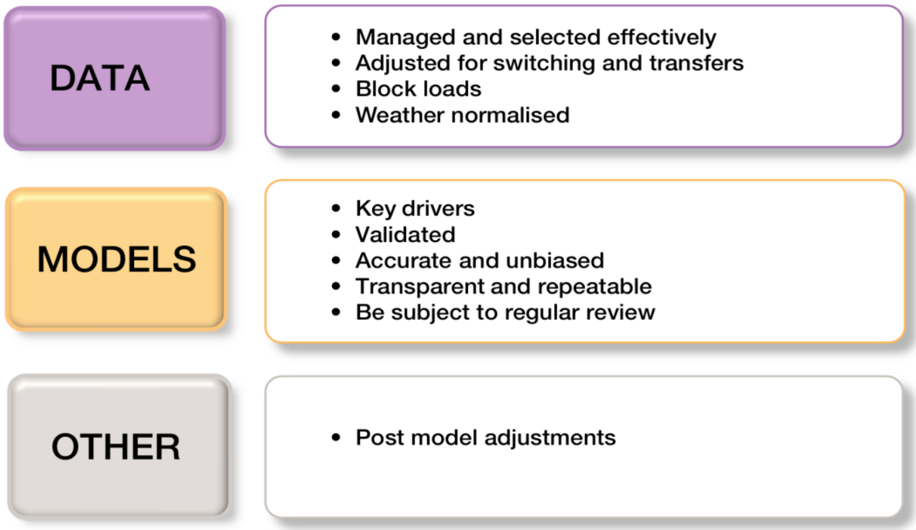
- Appendix E provides a brief discussion of the role of energy forecasts and the reasons that DNSPs typically do not prepare these at the CP or other spatial level.

# 2 Best practice forecasting principles

Electricity demand forecasts may be prepared and used for many reasons including network planning and policy development.

Whatever the final use, electricity demand forecasts should be prepared with regard to a number of key principles, which are displayed in Figure 2 and discussed further below.

FIGURE 2 **BEST PRACTICE FORECASTING PRINCIPLES**




*SOURCE: ACIL ALLEN CONSULTING*

## 2.1 Incorporating key drivers

An electricity demand forecasting methodology should incorporate the key drivers of electricity demand, either directly or indirectly. These may include[21]:

1. Economic growth
2. Electricity price
3. Population growth and/ or growth in the number of households
4. Temperature, humidity and rainfall/wind data
5. Growth in the number of air conditioning systems
6. Growth in the number of heating systems
7. Growth and change in usage of key appliances and other relevant technological changes.

[21] This is a list of drivers that may be applicable, but it does not necessarily follow that the ideal forecasting methodology will automatically incorporate all of these drivers. Whether individual drivers should be used in a given forecasting methodology is partly an empirical question and depends partly on the available data.

In the methodology described in this report the impact of weather is incorporated through the weather normalisation process described in chapter 5 and discussed in section 2.2.

The remaining factors are incorporated when bottom-up forecasts are adjusted to conform with an independently prepared system forecast as described in chapter 8. They should also be considered when:

- selecting appropriate growth rates (section 6.2)
- making post model adjustments (chapter 7).

A discussion of the expected relationship between electricity demand and its drivers is in Appendix A.

## 2.2 Weather normalisation

Electricity demand is well known to be sensitive to weather. The stochastic nature of weather means that any comparison of historical demand is only meaningful if the historical data are adjusted to standardised weather conditions. If this is not done, the analysis becomes, at least partly, an analysis of historical weather rather than electricity demand.

Another issue is that electricity demand forecasts prepared for regulatory purposes are not intended to forecast what electricity demand will be in any given year. Rather, they are intended to forecast what demand *would be* under normal weather conditions.[22] This cannot be estimated without accounting for the impact of weather on historical data appropriately.

For these reasons, an electricity demand forecasting methodology should incorporate weather normalisation as discussed in chapter 5.

## 2.3 Adjustments for temporary switching and permanent transfers

From time to time network operators change the way their networks are configured. Some loads are shifted between existing zone substations or CPs while others are shifted to new assets. These load transfers can be temporary (referred to herein as ‘switching’) or permanent nature (referred to as ‘transfers’).

To compare maximum demands over time, historical data must be adjusted to system normal conditions by removing or accounting for the impact of switching and transfers.

If this is not done, the data will suggest changes in demand at particular locations over time that did not actually occur.[23] This would result in regression models with lower explanatory power and which are subject to biased coefficient estimates.

In the methodology described here, these issues are accounted for at both the data collection (chapter 3) and data cleansing (chapter 4) stages.

[22] In most cases the objective is to forecast demand under ‘normal’ weather conditions or, more particularly, what demand would be under 50 per cent probability of exceedence (POE) conditions. In some cases the objective is different. For example in South Australia the transmission code currently requires that forecasts relate to the maximum demand that would be observed during heatwave conditions.

[23] For example a transfer would appear in the data to suggest that load declined at the transferring network element and grew at the receiving network element. Applying regression techniques to these data would provide growth projections that are biased downwards at the transferring element and upwards at the receiving element.

## 2.4 The likely impact of block loads

Growth in electricity demand can be thought of as comprising two 'parts':

1. Organic growth – referring to ongoing growth in the number of customers supplied by a particular CP and changes in the demand per customer. These typically lead to gradual, incremental changes[24] in demand each year. This is typically the result of growth in the usage per household or incremental population growth.
2. Block loads – referring to significant 'steps' in demand on a particular connection point. To 'qualify' as a block load, the step would need to be large relative to the load supplied at the connection point.[25]

In principle, block load and organic growth should be dealt with separately. However, block loads that occurred in the past are 'included in' the historical data so they obscure the forecaster's view of organic growth in the past. A forecasting methodology based on appropriate econometric techniques will incorporate the contribution block loads may make to growth in maximum demand in future. It is similar to assuming that future growth will include a similar rate of block load growth as has been observed in the past.[26]

The first part of the proposed approach to dealing with block loads is done at the data collection and preparation stages described in chapters 3 and 4 respectively.

There is also an inclination to add, to forecasts any future block loads that are known or expected. If the block loads are large relative to the annual growth expected at a particular CP then failing to add them will cause forecasts to be too low.

However, if block loads are included in the historical data, and expected new block loads are added to the forecasts the n the new block loads may be double-counted resulting in inflated forecasts.

To avoid double counting, past and future data should be prepared on the same basis. The conceptual challenge is to identify block loads that are *additional* to the historical trend.

The approach to doing this is described in section 7.1.

## 2.5 Accuracy and unbiasedness

All forecasting models will include errors by nature of the fact that they are an approximation of the real world. Those errors will limit the model's accuracy. Nonetheless, any credible forecasting methodology must produce forecasts that are reasonably accurate and whose accuracy can be measured objectively.

Assessing a model's accuracy should include both in-sample and out-of-sample tests. Poor performance on these tests could typically be traced to shortcomings in the modelling approach or to deficiencies in the data used. Whichever is the case, these should be addressed until the model performs satisfactorily.

Similarly, models should be free of bias, meaning that they should be no more likely to produce high than low forecasts.

---

24 In some cases demand per customer can fall by enough to offset increases in the number of customers and organic growth can be negative.

25 A block load could arise from population growth if very large subdivisions were built as may soon occur in Western Sydney for example.

26 This may be influenced by the stage of growth of the network.

The methodology proposed here is to prepare forecasts from regression equations using the ordinary least squares method of estimation (OLS).

OLS is a popular, powerful and widely used method of estimation which has been proved, subject to certain assumptions, to be the best linear unbiased estimator.[27] Therefore, subject to those same assumptions, forecasts prepared using the methodology outlined in this report will be free from statistical bias.

## 2.6 Transparency and repeatability

A transparent forecasting process is one that is easily understood and well documented and, if it was repeated by another forecaster, would produce the same result. It is generally incumbent on a forecaster who intends that their forecasts be used for regulatory or similar purposes to be able and willing to explain how they were prepared and the assumptions that were made in preparing them.

Forecasting electricity demand will inherently include subjective elements, exposing it to the judgement of individual forecasters. This is not inappropriate and 'judgement' should not be considered a 'dirty word' in this context.

However, the use of judgement increases the importance of transparency. In cases where judgement is used, those judgements should be documented and reasons explained, either as a process or individually.

To achieve this any documentation needs to set out and describe clearly the data inputs used in the process, the sources from which the data are obtained, the length of time series used, and details of how the data used in the methodology are adjusted and transformed before use.

The functional form of any specified models also need to be clearly described, including:

- the variables used in the model
- the number of years of data used in the estimation process
- the estimated coefficients from the model used to derive the forecasts
- detailed description of any thresholds or cut-offs applied to the data inputs
- details of the forecast assumptions used to generate the forecasts.

The process should clearly describe the methods used to validate and select one model over any others. Any judgements applied throughout the process need to be documented and justified. Adjustments to forecasts that are outside of the formal modelling process that are not documented with a clear rationale justifying that course of action should be avoided.

The methodology should be systematic so that any third party that follows a series of prescribed steps will be able to replicate the results of the forecasting methodology.

This report is the starting point for a transparent CP forecasting process. It outlines a methodology that could be followed by a third party. It also identifies a range of factors that should be documented through the process.

## 2.7 Estimated models should be validated

Models derived and used as part of any forecasting process need to be validated and tested. This is done in a number of ways:

---

27 The assumptions in question are technical, and are widely discussed in Econometrics texts. See, for example, Gujarati, Damodar N. "Basic Econometrics", second edition, 1988, p63.

- assessment of the statistical significance of explanatory variables
- goodness of fit
- in sample forecasting performance of the model against actual data
- diagnostic checking of the model residuals
- out of sample forecast performance.

These should be done after forecasts are prepared and an attitude of continuous improvement should be applied to the forecasting methodology.

## 2.8 Effective management and selection of data

The forecasting methodology requires effective management of data used in the process. This means keeping a central repository of all the data series used in the forecasting methodology in one or more electronic databases. The importance of the data collected implies that these databases need to be developed such that the management and collection of data is auditable and has integrity. These issues are discussed in chapters 3 and 4.

Ideally a number of electronic databases would be constructed which would split the data into categories depending on the type of data involved (for example demographic, economic, demand and temperature data) and the extent to which it has been processed.

Selection of which data series to use will depend on factors such as their:

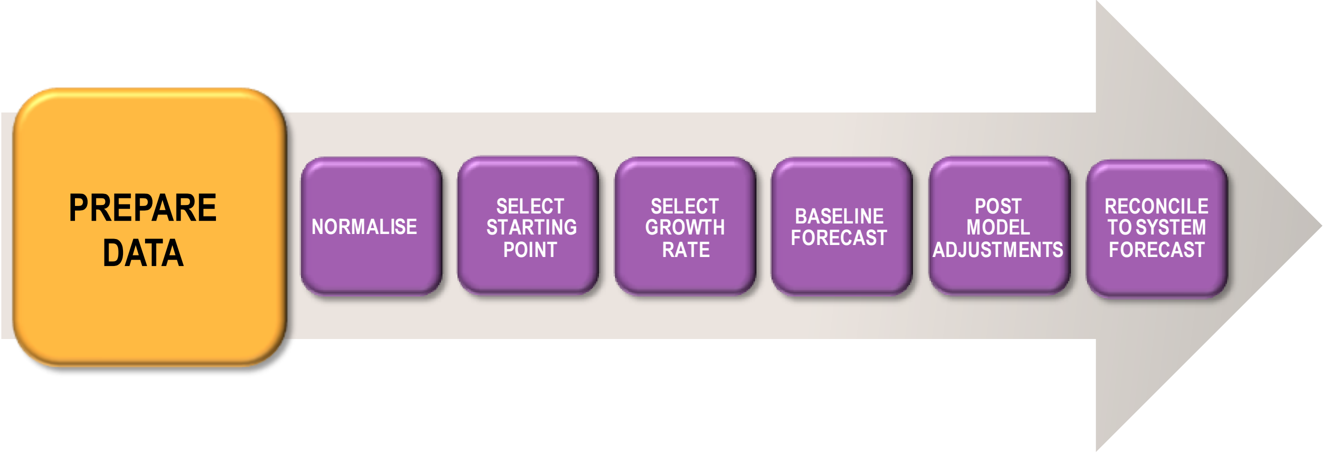
- reliability and accuracy
- the reputation of the data source
- the degree of completeness of the data and the absence of significant gaps
- the consistency of the data series through time
- the extent to which they cover a sufficiently long time series.

## 2.9 Regular review

The forecasting process should be subjected to review on a regular basis (annually or bi-annually) to ensure that the data inputs have been collected and utilised adequately and that the applied methodology meets the above principles.

The review should also focus on forecast performance and consider the possible causes of any divergence of observed maximum demand from the forecasts. The causes of the divergence could relate to factors such as differences between forecasts of the explanatory variables and the actual levels observed, or could be due to structural issues with the way the models are constructed.

# 3 Data collection and management




Preparing maximum demand forecasts begins with collecting data and managing it appropriately. The data that should be collected and their likely sources are summarised in Table 1. In summary, data relating to historical maximum demand and weather conditions should be collected along with historical data and projections of likely drivers of electricity demand and metadata relating to the CPs for which forecasts are to be prepared.

TABLE 1 **DATA REQUIREMENTS**

| Data | Description | Source |
|---|---|---|
| Historical | | |
| Demand data | High frequency, long time series of demand measured at each CP to be forecast | TNSPs and/ or DNSPs |
| Historical switching and transfers | Size of load switched or transferred and timing. | NSPs |
| Block loads | Historical block load timing and size | NSPs, large customers |
| Embedded generation | Historical output for some time and frequency as demand data | AEMO, generators |
| Weather data | Long (30 year) time series of weather data at a weather station assigned to each CP | BOM |
| Economic activity | Economic indicators for the area supplied by each CP, or as close an approximation as can be found | ABS, Local Government, State Dept. of Treasury or Planning |
| Population | Population of area supplied by CP | ABS, Local Government, State Government |
| Policy data | Impact or nature of policies that may have influenced electricity demand | Commonwealth, State and Local Governments |
| Metadata | Descriptive information relating to area supplied by each CP | Various including NSP, local Government, ABS etc. |
| Projection | | |
| Block loads | For each candidate block load identified, size of expected block load, likelihood it will proceed, coincidence with system | NSPs and/ or other local area experts, large customers |
| Large loads | Expected shut down or start up of large loads | NSPs, local area experts, large customers |
| Embedded generation | New plant expectations or shut down of existing plant | Generators, AEMO (market registration if large enough) DNSPs |
| Transfers | Timing and load to be transferred | NSPs |
| Economic activity | Projections of economic drivers | State and Local Governments, Consultancies |
| Population | Projected population in area supplied by CP | ABS, Local Government |

The main criteria used to determine the suitability for modelling of a particular data source are:

- reputability of the data source
- reliability of the data source
- completeness (no or few missing values)
- suitably long time series
- accuracy of the data.

The majority of external[28] data used in the modelling process are sourced from the Australian Bureau of Statistics (ABS) and the Bureau of Meteorology (BOM). Both of these agencies have established very strong reputations. In particular the ABS is known to apply best practice data collection principles and has been rated by 'The Economist' magazine as one of the top few statistical agencies in the world.

Data sourced from the ABS can be regarded as accurate and the probability of subsequent revisions to data is generally low. ABS data can also be considered to be reliable in the sense that the economic and demographic data used in this forecasting methodology are unlikely to be discontinued.

Similarly, the BOM has established a number of high quality temperature sites which have few missing values, are homogeneous over time and have a very long record of unbroken daily recordings.

[28] That is, external to NSPs.

Some information will be required from people with expertise in the area supplied by the CP. For example the likely size and timing of block loads. Local area expertise can usually be found within NSPs or in Local Governments etc.

## 3.1 Demand data

The first data requirement for the CP forecasting methodology is a sufficient time series of half hourly or 15 minute interval data for each CP to be forecast. Ideally this time series should go back at least 10 years, although it is possible to work with shorter time periods.

It is assumed that the forecasts are to be presented in MW's, so the data should also be in MW. If forecasts are to be in MVA the process should begin with data in MVA.[29]

## 3.2 Weather data

The next data set to be collected is weather data. The methodology requires a very long time series of daily maximum and minimum temperature data. It may also be worthwhile to add other weather variables such as rainfall and wind speed. Which variable to use in the forecast is an empirical question to be answered at a later stage.

To allow a thorough understanding of long run weather conditions it is preferable to use at least 30 years of weather data, more if it is available. Data must be accurate with few missing observations. Missing observations should be imputed as discussed in section 4.3.1. If more than about five per cent of data are missing, preference should be given to other weather stations. The weather station should also still be in operation. If it has been discontinued another weather station should be used for the entire process.

The data should be obtained from the BOM. The BOM supplies data for hundreds of weather stations across the country. It will be necessary to choose which is the most suitable for each CP.

The key aspects of assessing the suitability of a weather station include the following, with each discussed below in turn:

1. relationship between demand and weather as measured at the station
2. prevalence of missing data at the station
   a) at least 90% of cells should be populated
3. length of time series available
   a) the time series should cover 30 years if possible, although some compromises may need to be made on this criterion
   b) the weather station needs to be in current operation
4. availability of important weather variables
   a) at the very minimum the weather station will need to collect temperature data
   b) other variables such as humidity, rainfall, sunshine and wind speed may also be required.

It may seem reasonable to expect that the most appropriate weather station to use at each CP will be the one that is closest to the CP and meets the minimum quality

[29] It is possible to convert from one to the other using a power factor, though it is preferable to begin with data measured in the units to be forecast.

requirements listed above. However, due to the vagaries of microclimates this may not necessarily be the case.

Before assigning a weather station to each CP, we recommend a detailed analysis of the daily correlations between daily maximum demand and each potential weather variable to determine which stations provide the best explanatory power for movements in the daily maximum demands. This analysis should be conducted separately for both the summer and winter seasons (as defined in the methodology) and for a range of weather variables.

It is also not out of the question for different weather stations to be used between the summer and winter seasons at the same CP.

Conceptually, it is important to use weather data collected at a place where weather conditions are similar to those experienced by customers in the area supplied by the CP in question. Therefore, the most suitable weather station to use to weather correct a CP will probably be the weather station closest to it.

However, this might not always be the case. For example, the purpose of using weather data is to correct for changes in demand that are partly behavioural. Therefore, it may be appropriate to use data from the weather station that is used for local weather reporting, i.e. the station for which forecasts and actuals are reported in local news, even if other weather stations are closer to the CP.

In some cases weather stations will be moved, which can cause structural breaks in the data. Data from these stations may be unsuitable, and should be treated with care and possibly avoided. The historical data may have been adjusted to account for the break, in which case they should be examined carefully if they are to be used to identify inexplicable changes in the relationship between weather and demand before and after the move.

Once a weather station has been identified as appropriate to use in forecasting demand at each CP it is unlikely to change. However, the analysis should be repeated periodically to confirm that it is still the most suitable weather station to allocate to the CP. This is particularly important in fast growing areas where the area supplied by a CP changes significantly over time.

## 3.3 Block loads

Data should be collected relating to:

- actual block loads installed during the historical period for which maximum demand data were collected
- block loads expected to be installed during the forecast period.

The necessary data would include details of:

- the size of the load (expressed in MW[30])
- the timing of the load (year and season of commencement)
- an assessment (for future loads) of the probability that the load will proceed as planned (see section 7.1).

## 3.4 Metadata concerning connection points

To better understand the nature of the loads at specific CPs, it is helpful to obtain a breakdown of each CP load by customer type (residential/commercial/industrial). If

[30] Or MVA, see section 3.1.

this is not available, it would be helpful to have an estimate of the percentage breakdown of the load at each CP into the five separate categories, Residential, Commercial, Industrial, Agricultural and Mining. If this is not available then classifying each CP into one of the following categories would suffice:

— predominantly residential
— mixed residential/industrial /commercial
— predominantly commercial
— predominantly industrial
— predominantly agricultural
— mining related.

This information would not necessarily be used formally in the forecasting process, though in some cases this will occur. For example, weather normalisation would not be applied to a CP where the load is dominated by industrial customers that are not weather sensitive.

In any case metadata can give the forecaster useful guidance concerning the type of behaviour to expect from each CP. For example, those that are residential would be expected to display the greatest degree of weather sensitivity while those that are mining related might not be significantly sensitive to weather.

This data may be available from the relevant NSPs or it could potentially be assembled independently.

## 3.5 Other data

### 3.5.1 Planned transfers between connection points

Data should be obtained relating to any historical and planned transfers between CPs. The data should include the timing of the transfer.

The data should also include the load (to be) transferred, ideally on the same basis as the forecasts themselves. That is, for 50 POE forecasts, the load to be transferred should be in 50 POE terms. For 10 POE forecasts, it should be in 10 POE terms. In practice, though, this could only be determined if a historical data series is available for the load to be transferred. This would be available if, for example, all the load at a given zone substation was to be transferred. However, if load is ‘split’ and only part of a measured historical load is transferred, the historical data necessary to determine 10 or 50 POE load may not exist. In this case the only thing that can be done is to obtain the best estimate of load to be transferred and apply that same estimate to both the 10 and 50 POE levels.

### 3.5.2 Network demand management

If a DNSP is planning a significant network demand management program with a large reduction in maximum demand over the forecast period, details should be provided to the forecaster before the forecasting process is implemented.

While it may not be possible, details of the size of any demand management initiatives at the CP level should be provided. However, it may only be possible to obtain an estimate of the impact of any network demand management at the system level only. In this case it would be factored into the forecasts through the reconciliation with the system forecast.

### 3.5.3 Solar PV and other embedded generation

In order to capture the impact of increasing solar PV penetration and possibly other forms of embedded generation, it is desirable that the forecaster obtain a historical time series of the capacity of solar PV installed in the area supplied by each CP. In practice, it may only be possible to obtain this data by state (NEM region). The key issue to consider is the increase in the use of PV, not the existing installed capacity, which is already reflected in the historical demand data.

### 3.5.4 Regional population data

Population data for the area supplied by each CP should be collected.

Good historical data is available from the ABS at both the local government area (LGA) level and statistical local area (SLA). Both spatial units are small enough to capture demographic changes around specific CPs. Data is available through the collection, Australian Demographic Statistics and also the most recent Census conducted in 2011.

Regional population forecasts at the Local Government Area level are also produced by many Local Councils or Shires, although these tend to present an optimistic picture of the future growth prospects of their region. An alternative source of regional population growth forecasts might be the various Departments of Planning across State jurisdictions.

### 3.5.5 Large load changes

Data relating to substantial changes in load should be collected, both historical and expected. For example, if there were industrial facilities which shut down for maintenance or similar these should be accounted for. The preferred approach is to modify historical data either to remove the affected observations entirely or to 'add back' the load that would have been observed had the shutdown not occurred.

# 4 Data processing and data set preparation


Before any forecasting models can be constructed and estimated the data that have been collected must be prepared by:

- cleansing and filtering the data
- accounting for permanent shifts in levels
- imputing missing data.

## 4.1 Data cleansing and filtering

Before the data can be used, any anomalies that may be present should be removed or accounted for.

The two main sources of difficulty arise from outliers caused by temporary network switching and from permanent shifts in the level of maximum demand over time caused by block loads, permanent load transfers or other shifts in load (see section 3.5.5).

Failure to account for these can lead to biased coefficient estimates, poor weather normalisation results and poor maximum demand forecasts.

There is a certain degree of overlap between the cleansing and filtering process and data collection. Notionally, if data can be collected in a clean and filtered state, cleansing and filtering is unnecessary. However, in practice it is unlikely that the data that are available will be entirely 'clean'.

There can also be numerous small changes which are not easily detected. The quality of the data set can be substantially improved by removing these though clearly, this is a difficult task without detailed local knowledge and records of what was happening within the network over the period being considered.

We recommend an iterative process of analysing the data to identify outliers. This is essentially a matter of building evidence to test a conclusion that individual observations should be omitted from the dataset. It is important that data are not omitted too readily because the point of the forecasting being done is to understand the behaviour of demand at peak times.

The first step in the process is to consult with the relevant NSP to identify switching activity in its operating records. Data recorded on days when switching was known to have occurred should be deleted or corrected for the known switching by reducing the observed load at the receiving CP and increasing it at the donating CP.[31]

Next, the data should be visualised, using a scatter plot or similar to identify outliers.

This process will identify observations which may be anomalies and should perhaps be omitted from the data set. No hard and fast rule can be applied.

The proposed approach is to apply a series of tests to the data. The more ‘tests’ that an observation ‘fails’, the more likely that it is an anomaly. Before any observations are omitted, though, it would be advisable to consult with the DNSP and explore the possibility that there is an explanation for the anomaly that supports a conclusion that it should be omitted.

Outliers usually show up as irregular movements in the pattern of daily demand and often present as a large spike or dip that is some distance outside the normally observed bounds of behaviour. The first test is to visualise the data using a scatterplot or similar and identify observations that appear to be outside the norm.

The second test is to subject the data to a statistical procedure that captures those values that lie outside some threshold from the mean. For example, identifying all observations that lie more than two standard deviations from their seasonal mean.[32]

It is important to control for weather sensitivity in this process. Therefore, the residuals from a simple regression between daily maximum demand and weather data (usually maximum and minimum temperature) with dummy variables for day of the week effects and public holidays, should be collected. Those residuals should be subjected to the same two tests (i.e. visualised and compared to their seasonal mean).

## 4.2 Account for load transfers - permanent shifts in levels

The discussion above relates to anomalous readings in the historical data due to temporary changes.

There is also the possibility that load is transferred permanently, causing a permanent discrete shift in the level of the series. The same effect can arise from a significant block load being added to or removed from a CP (for example a large industrial load).

Permanent transfers will be less common at the CP level than the zone substation level. However, adjustment may still be required to the level that is consistent with the average level of demand in the proceeding period. It is also likely that the relevant NSP will be aware of historical transfers.

The approach is broadly the same as the testing process discussed in section 4.1 above. That is, examine the data, both in raw form and as residuals from a crude weather normalisation, to identify possible changes. In this case, though, the search is for changes that persist. Graphically these will appear as either upward or

[31] In other words, identify the size of the load that was switched and ‘shift’ that amount back from the CP that *actually* supplied it to the CP that *usually* supplies it.

[32] It is important that the mean value used for reference here covers a relatively short period (two or three years) so that organic growth does not cause routine observations to appear to be outliers.

downward steps in demand. In practice the same testing process will capture both anomalies and permanent shifts in demand.

If such steps are identified and attributed to a discrete event such as a block load, it would be appropriate to adjust the average level of the data up or down relative to the last season from which the forecasts are going to start. If this is not done, the step change will bias the growth rate as observed in the data.

On the other hand it is also possible that the discrete jump arises as a result of a shift in the underlying demand at a CP that is not related to transfers or large loads. In this instance, the recorded jump or decline in the peak demand in a given year is real and should not be adjusted.

This highlights the importance of strong local knowledge, good reliable data capture and storage systems and having a strong understanding of what is happening within the distribution network. While visual inspection of time series data is useful, local intelligence is also important in ascertaining differences between artificial and real movements in peak demand.

It is also important to note the possibility of changes between seasons. That is, the forecaster may be focussed on analysing summer peaks and, in doing so, concentrating on summer demand data. This may conceal the fact that a permanent change occurred in spring (for example) that caused demand to be different in later summers.

## 4.3 Imputation of missing data

### 4.3.1 Imputation of missing weather data

At most weather stations there will be a small number of missing observations. It is necessary to impute the missing values to use these data in a forecast model.

The approach is to impute missing values either by using a value that is representative of the time series or by choosing a value from a day that appears similar. Therefore, where there are missing observations of:

- maximum and minimum temperature – substitute the missing value for another day in the time series when demand was similar and the day was of the same ‘type’ (i.e. weekend, public holiday etc)
- wind speed (if used) – replace with the average daily wind speed observed in the corresponding season over the sample
- rainfall (if used) – replace with zero, which is typically the most commonly observed value.

If the necessary observations are missing on days when peak demand occurred it may be more appropriate to use data from a different weather station, though this should be traded off against the reasons the alternative weather station was not chosen in the first place.[33] It would also be possible to impute missing data at one weather station using a statistical relationship estimated between it and another weather station. This may provide superior results, which could be tested by imputing data that are not actually missing.

This imputation process should only involve a small number of adjustments if the weather station that is used is of a sufficiently high quality.

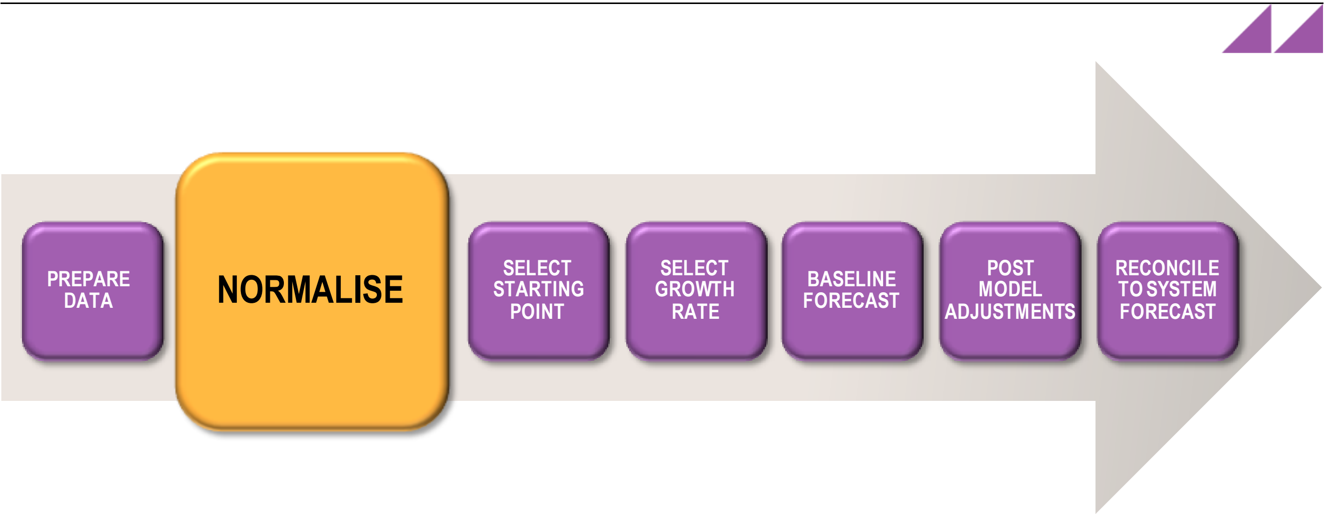
[33] For example, the alternative weather data will probably not be as closely correlated with demand or it would have been used in the first place.

### 4.3.2 Dealing with missing maximum demand data

It is also possible that there will be segments of missing maximum demand data. It is not appropriate to impute these missing data. They should be omitted from the data series.

In other words these data points should be treated as missing values within any calculations. They should not be treated as zero demand, which would bias the results.

# 5 Weather normalisation

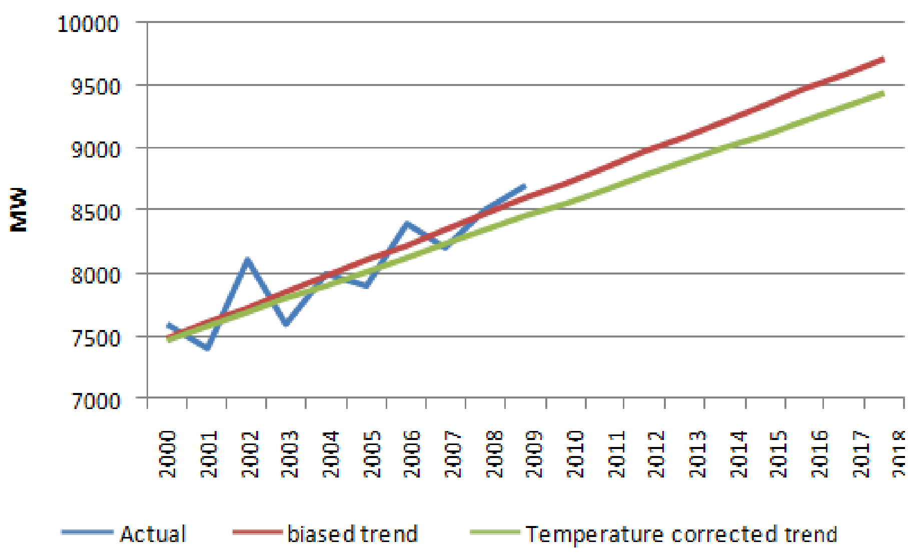



It is well established that weather is a significant driver of demand for electricity. Because weather varies from year to year it is important to normalise historical data to allow a constant basis of comparison over time.

Using data that have not been weather normalised can bias the observed trend when time series data are used as a basis for forecasting future peak demands.

Figure 3 provides a simple illustrative example of biased forecasts arising from fitting basic trends to data that have not been weather normalised. In the example, the weather was warmer than average in the later years, causing demand to be higher than it would have been under 50 POE weather conditions (the example is summer demand on a summer peaking network element).

FIGURE 3 **BIAS ARISING FROM NON WEATHER NORMALISED DATA**




*SOURCE: ACIL ALLEN CONSULTING*

Figure 3 shows:

- a (red) trend line fitted to the raw historical data
- a (green) trend line fitted to the weather normalised historical data.

The example shows that the hotter than normal conditions in the later years cause the red (raw) trend line to have a higher growth rate than the (green) weather normalised trend line. This is the influence of weather, specifically the coincidental fact that the later summers were warmer than earlier.

The raw data give a misleading impression of the underlying growth in demand at the CP.[34] While there may be very little difference in the two trend lines over the estimation period and into the early years of the forecast period, the degree of divergence between the two lines becomes larger the further out the forecasts extend.

This chapter sets out our recommended approach to accounting for the impact of weather in preparing electricity demand forecasts. There are alternative approaches, which are discussed in Appendix B.

Weather normalising historical data begins with estimating the relationship between demand and weather. Our recommended approach to doing this is described in sections 5.1 and 5.2.

When that relationship has been estimated, the remaining task is to produce weather normalised historical data. Section 5.3 describes our recommended procedure for doing this.

## 5.1 Which weather variables to use in the regressions?

To estimate the relationship between demand and weather, it is first necessary to choose the appropriate way to measure weather conditions. As part of the model building exercise we recommend testing the strength of the relationship between alternative weather variables at each CP and in each season to identify which best captures the relationship between weather and demand.[35]

A list of candidate weather variables is in Table 2 below with a description of the likely nature of their correlation with demand.

TABLE 2 **POSSIBLE WEATHER VARIABLES TO BE USED IN ESTABLISHING THE RELATION BETWEEN MAXIMUM DEMAND AND WEATHER**

| Measure | Summer peak | Winter peak |
|---|---|---|
| Air Temperature | Hotter increases load | Colder increases load |
| Humidity | Higher moisture increases load | |
| Sun/sky cover | More sunshine increases load | More sunshine decreases load |
| Wind | Indeterminate/Mixed impacts | More wind increases load |
| Rainfall | Potentially reduces load | Potentially increases load |

*SOURCE: ACIL ALLEN CONSULTING*

The appropriate weather variables should be chosen based on their relationship with demand. All else being equal, the best weather variables are those that are most closely correlated with demand.

[34] The bias will not necessarily be upwards. If the weather had been cooler than average in the example the bias would be downwards.

[35] That is, which variable provides the best fit in a regression against truncated daily maximum demand.

The correlation between demand and candidate variables can be examined informally, using scatter plots, or by estimating correlation coefficients or regression lines.

It is also worthwhile considering using lags of up to several days into the model of the relationship between daily maximum demand and weather. This would account for the possibility that customers may have a higher propensity to use heating or cooling appliances *today* because weather was extreme *yesterday.*[36] It may also reflect physical characteristics of the housing stock in an area or that feelings of discomfort are greater when a heat spell has lasted longer, even though the outside ambient temperature remains the same.

In our experience, temperature is typically the most important of the weather factors in driving daily maximum demand. Together, the daily maximum and night before minimum temperatures tend to capture the majority of weather related variation in the daily maximum demand.

## 5.2 Estimating the relationship between maximum demand and weather

Weather normalisation is based on an estimated (quantitative) relationship between maximum demand at any given CP and the weather variables that drive it.

Given that the objective is to forecast (weather normalised) maximum demand it is only necessary to analyse seasons when demand peaks are likely to occur. Generally speaking this means that only summer and winter need to be analysed, possibly not even both. References in this report to 'season' exclude autumn and spring.

Our approach to weather normalisation is applied to each season individually. The relationship between weather and demand is estimated by fitting a multiple regression to daily maximum demand and weather data for each season for which data are available. The model coefficients from those seasonal regressions are used in implementing the weather normalisation procedure.

Weather normalisation begins with data for an entire season in a single year. The necessary data are:

1. observed daily maximum demand
2. weather.

As discussed in section 5.1, weather may be daily maximum and minimum temperature, the average of these or another variable.

If Maximum and Minimum demand are used,[37] compute a linear regression of the following form:

$$MD_d = m * MAXtemp_d + n * MINtemp_d + c + \varepsilon$$

where $MD_d$ is maximum demand observed on day d for all days in the dataset MAXtemp and MINtemp are daily maximum and minimum temperature respectively *m, n* and *c* are regression parameters, and $\varepsilon$ is an error term.

Collect the coefficients of the regression model and the standard error of the estimate for use in the weather normalisation procedure described in section 5.2.

---

[36] Or the day before etc.

[37] In our experience this approach accounts for the majority of weather related variation. More complicated models relating demand to other weather variables can be built, but there is a question of assessing the extra benefit of doing so against the additional cost.

Four aspects of specifying the weather normalisation model for each CP are discussed below, namely:

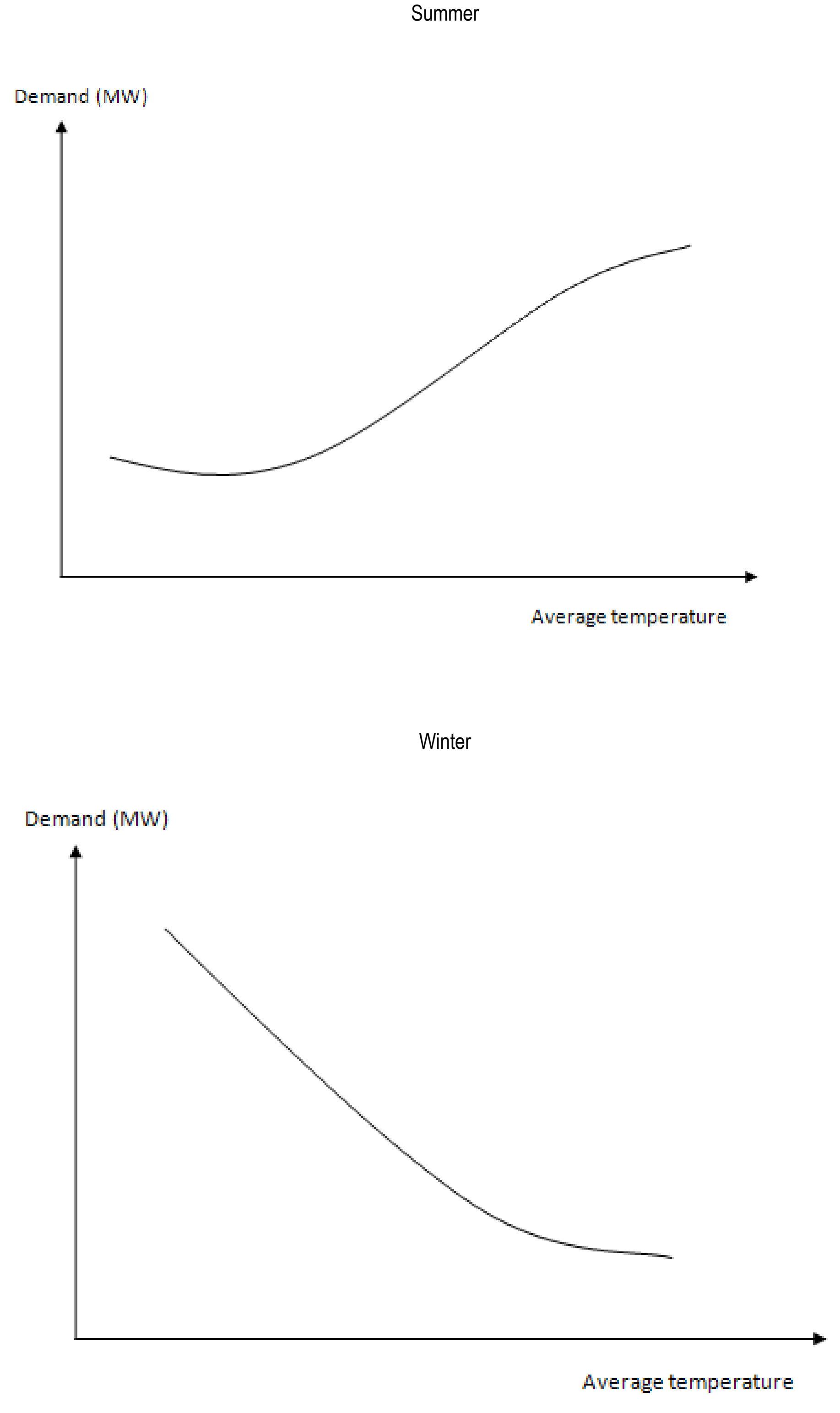
1. the functional form of the relationship, i.e. linear or non-linear, and the maximum demand data to be used for weather normalisation, discussed in section 5.2.1
2. the data used for weather normalisation and how to define the season for weather normalisation purposes, discussed in section 5.2.1
3. how to treat CPs where demand appears not to be weather sensitive, discussed in section 5.2.2
4. the approach to take if the last season for which data are available was very mild, discussed in section 5.2.4.

### 5.2.1 The functional form of the relationship - linear approximation method

The relationship between demand and weather tends to follow the S-shape illustrated in Figure 4. As temperature increases in summer (or decreases in winter):

1. there is a temperature range in which demand is not responsive to temperature (flat part of the curve)
2. then there is a range in which demand increases with temperature (steep part of the curve)
3. then there is a range where the relationship 'flattens out' under extreme conditions. This occurs under extreme conditions only and is only observed during extreme weather conditions (more extreme than 10 POE) (saturation part of the curve).

FIGURE 4 **THEORETICAL RELATIONSHIP BETWEEN DAILY MAXIMUM DEMAND AND TEMPERATURE**




With sufficient data, it would be possible to specify the relationship between demand and temperature as a curve, and to estimate the parameters of that curve using multiple regression. However, as discussed in Appendix B.1.2, in practice there are rarely sufficient data for this. The key reason is that, by definition, observations at the 'top' of the curve, where weather is extreme, are observed only infrequently, no more often than one year in ten.

Our preferred approach is to use a linear approximation of the curve by truncating the dataset to omit observations on the 'flat' part of the S-curve described above and fitting a linear regression to estimate the 'steep' part of the S-curve. If the saturation part of the curve is observed in the data, it would be worthwhile considering the use of an S curve instead of fitting a linear approximation. However, this part of the curve is rarely observed; even low POE years are defined as such by only one observation, as illustrated in the figures on the following pages. By contrast, the flat part of the curve is experienced during mild conditions and is always observed in the data, so it must be dealt with.

To estimate the relationship between weather and demand, the first step is to truncate the dataset to remove non working days, generally weekends, public holidays and a period from shortly before Christmas until early January.

This process is illustrated in Figure 5, which shows maximum demand data from CP in a single summer.

The top pane of Figure 5 shows the maximum demand observed each day in the summer.[38] A regression line fitting daily maximum demand to daily average temperature is shown to give an indication of the relationship that these data suggest exists between demand and weather. For reference the 10 and 50 POE daily average temperatures are also shown.

In the second pane of Figure 5 weekends and public holidays have been deleted. The slope of the fitted line is approximately the same. It is noteworthy that the hottest day has been deleted from the dataset (it was a weekend) and that demand on that day was slightly lower than the hottest remaining day even though that day was notably cooler.

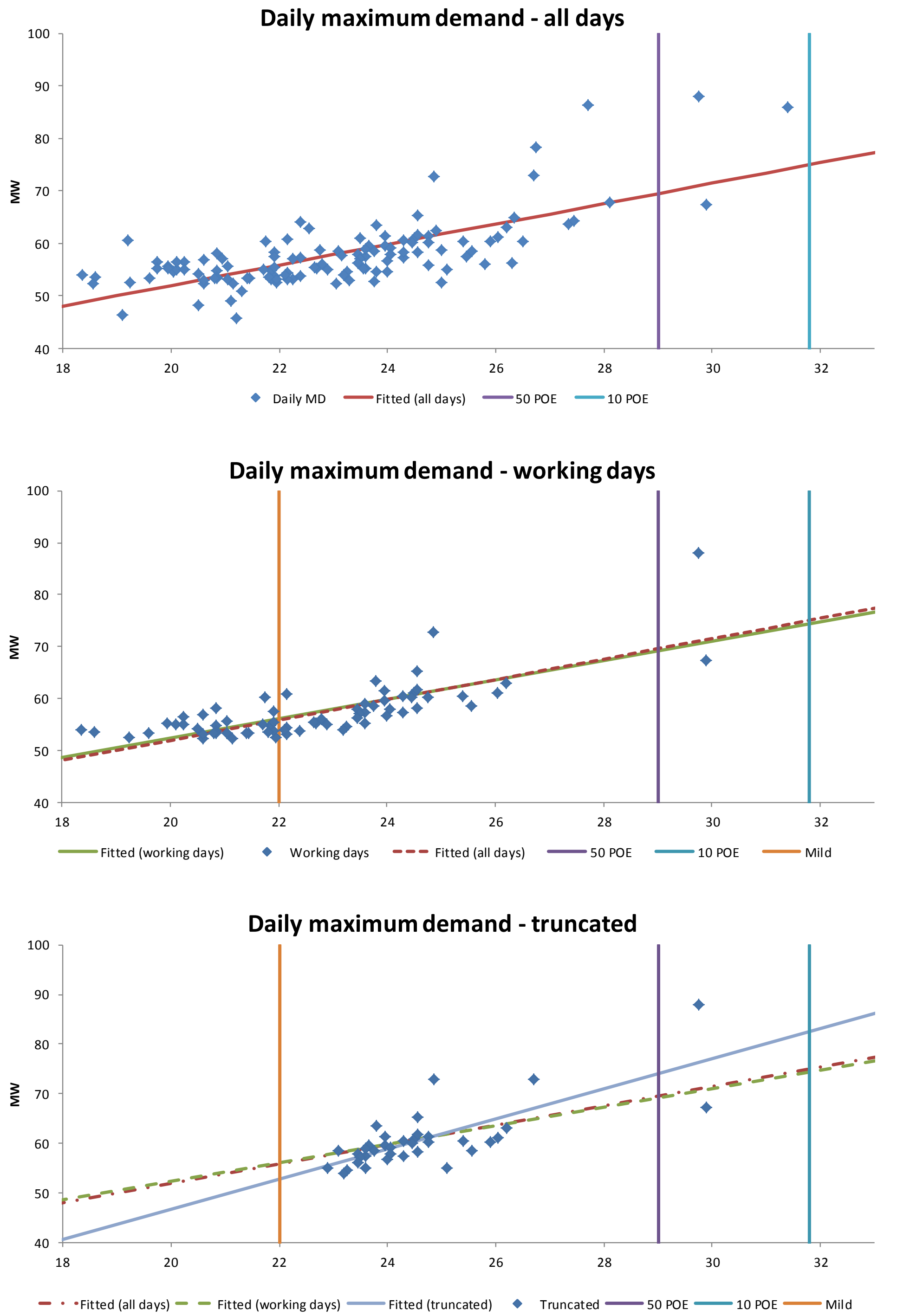
Note that Figure 5 shows that, even though this was close to a 10 POE year, there were only three observations above the 50 POE temperature. Of these, only two occurred on working days. This is insufficient to allow the saturation part of the S-curve to be estimated meaningfully.

The second step is to identify the threshold temperature that defines the beginning of the 'steep' part of the S-curve. That is, the point where weather conditions stop being 'mild'.

The second pane of Figure 5 shows the chosen threshold for 'mild' conditions at 22C. At higher (daily average) temperatures there are a number of observations that are significantly above the cluster of other observations.

[38] In this example summer was defined as 1 December to 31 March. Section 5.2.2 discusses the approach to determining how to define seasons for weather normalisation purposes.

FIGURE 5 **TRUNCATING DATASET FOR WEATHER NORMALISATION**




*SOURCE: ACIL ALLEN CONSULTING*

The third pane of Figure 5 shows only the truncated dataset. In this pane the slope of the weather relationship is significantly higher than in the previous two panes and the regression line passes closer to the observed maximum demand.

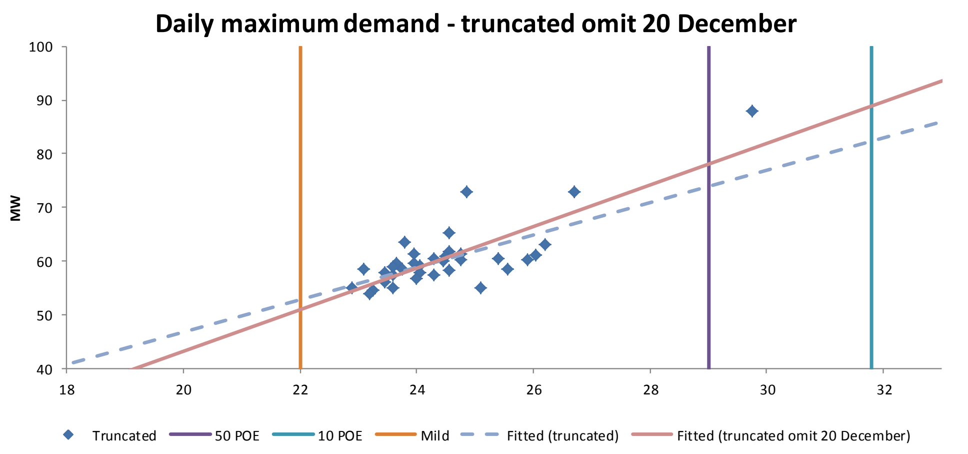
The third pane of Figure 5 also highlights a challenge in weather normalising data. There are two days in the truncated dataset when daily average temperature was approximately 30°C. These were:

— 20 December, when average temperature was 29.9C and maximum demand was 67MW

— 18 January, when average daily temperature was 29.75 (slightly lower) and maximum demand was 88 MW (substantially higher)

The demand observed on 20 December is omitted from the dataset in Figure 6. As Figure 6 shows, this has a significant impact on the estimated relationship between weather and demand. Omitting that observation causes the goodness of fit of the regression line to improve (R-squared increases from approximately 0.56 to approximately 0.64). However, the improved fit alone is no reason to conclude that the observation should be omitted.

FIGURE 6 **TRUNCATED DATASET – 20 DECEMBER OMITTED**




*SOURCE: ACIL ALLEN CONSULTING*

In this situation it would be worthwhile examining potential causes of the lower than expected demand on 20 December that summer. It may turn out that a temporary switch had taken place that day or that demand was reduced for technical problems such as a substantial outage. Alternatively, given its proximity to Christmas, the day may perhaps be properly described as a non-working day (as is likely in this example).

On the other hand, there may be nothing unusual about that day other than that demand was lower than on another day when the weather was similar.

If an explanation is discovered to establish that demand was lower than 'normal' on 20 December, it would be appropriate to omit the observation and use the regression relationship illustrated in Figure 6 for weather normalisation. On the other hand, if no explanation can be found, the conclusion should be that part of the relationship between demand and weather is that demand can sometimes be low on hot days. In this case the regression illustrated in the third pane of Figure 5 should be used for weather normalisation.

### 5.2.2 CPs that are not weather sensitive

The process of estimating regressions between demand and weather data may show that some CPs are not weather sensitive. In practice, CPs for which the regression between weather and demand weather data provides an R-squared value of less than 0.3 should be considered not to be weather sensitive, or the data should be considered too unreliable to allow normalisation, and normalisation should not be used.

Similarly, some CPs may be known not to be weather sensitive, particularly CPs that supply only industrial customers. Normalisation should not be used in these cases.

### 5.2.3 The season of interest

Demand for electricity is largely due to heating or cooling and, therefore, varies between seasons. Therefore, weather normalisation should be conducted on a dataset that relates to only one season. This raises the question of when the seasons occur, which is particularly relevant for summer which is sometimes treated as extending from November to March.

This should be resolved at the local level. The objective is to ensure that the summer data capture summer peaks. If very high demands[39] occur routinely in November, or in March, those months could be added to the summer season. However, in doing so, thought should be given to whether the relationship between demand and temperature in these months is the same as it is in other summer months.

The same issues should be considered in relation to winter demands. As with weather normalising summer demand, the objective of weather normalisation in winter is to ensure that the data capture winter peaks. However, these tend to occur more reliably in 'calendar' winter (June, July and August) so adding months to the season is typically not necessary in our experience. However, it is still possible that there are parts of the NEM where winter peaks can occur outside 'calendar' winter. In these cases, additional months should be included.

### 5.2.4 One season unless very mild

Generally speaking, the data used for normalisation should relate to only one season (that is, only one year's summer or winter). The exception to this is if the season in question was very mild. In this case there may not be enough data to describe the relationship between temperature and demand under unusual conditions accurately.

The solution to this is to pool data, that is, to combine the data from more than one consecutive season and apply the process described in section 5.2.1 to the pooled data.

Pooling increases the number of observations that are available, but it is not without problems. When data are pooled, no allowance is made for changes in the relationship between weather and demand over time.[40] Therefore, data should not be pooled over more than two or three seasons. When pooling data it is necessary to account for underlying growth in demand over time, which can be done by scaling the data by the growth in average demand or using a dummy variable in the regression.

## 5.3 Weather normalisation procedure

Once the relationship between demand and weather has been estimated, that relationship is used to produce a set of weather normalised historical demands. As discussed in Appendix B.1.3, there are several procedures that can be employed in doing this.

We recommend that a regression and simulation approach be applied using the seasonal regressions developed in section 5.2. This approach allows more complex relationships between weather and demand than other methods, which assume a one to one relationship between temperature and demand.

---

[39] 'Very high' means relative to demand observed at the CP in question.

[40] This is possible but it increases the complexity of the model significantly.

In addition to using all available maximum demand data, the regression and simulation approach uses all available weather data. This will usually include weather data collected before maximum demand data are available. This helps ensure that the weather normalisation captures as much of the natural variability in weather as possible.

The procedure is to:

1. Use the estimated regression models for each season with all available weather data to produce estimates of what annual maximum demand *would have been* in each season under all of the weather conditions that have been observed given the relationship estimated in section 5.2.
2. Allow for natural variability in daily maximum demand using the standard error of each fitted regression. This is done by adding an error taken as a draw from a normal distribution with a mean of zero and standard deviation equal to the standard error of the regression to each fitted value from the regression above. [41]

To illustrate, consider a CP for which there are 30 years of summer weather data at the weather station that has been identified as most closely correlated. Assume that the chosen weather variable is temperature, with both maximum and minimum used independently, and that there were 75 working days each summer.[42]

First, use the coefficients of the regression describing the weather relationship (from section 5.2) and the weather data for the 75 working days in all available years (30) to produce estimates of what the daily maximum demand would have been in the most recent season under *all* weather conditions observed.

In this example there would now be 2,250 fitted daily maximum demand values (that is 75×30). Notionally these suggest what daily demand would have been in 2012/13 for each weather outcome that has been seen in the last 30 years. Of these 2,250 demand values, 30 are annual (summer) maximum demand values, one maximum for each year of weather data. Record those 30 annual maximum demand values.

Then a random error is added to each of the initial 2,250 notional demand observations. This is done by taking a draw from a distribution with mean zero and standard deviation equal to the standard error of the estimate of the regression described in section 5.2.

Every time this is done (i.e. every time a trial is conducted) another set of 30 annual maxima is produced. Therefore, if 100 trials are conducted, the result will be 225,000 daily maximum demands, of which 3,000 will be annual maxima.

The 10 and 50 POE demand values are taken from these 3,000 observations. That is, the 50 POE is the median of these 3,000 values and the 10 POE is the value that falls on the ninetieth percentile. This is illustrated in Figure 7. Any POE level can be taken by choosing the corresponding percentile.

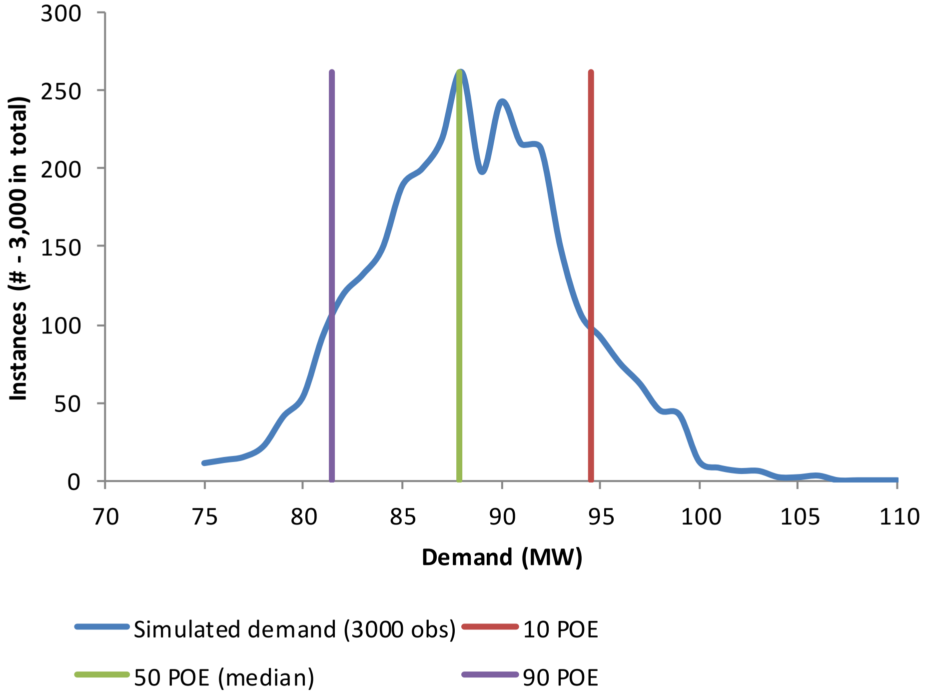
Figure 7 illustrates the result of simulating 225,000 observations from a normal distribution with mean 88 MW and standard deviation 5 MW and taking the 3,000

41 Assuming that the errors are normally distributed is an adequate starting point that should be used unless there is strong evidence to the contrary. If the analysis suggests that the errors follow another distribution, then it is reasonable to use it. It may be the case that the true distribution of the errors has 'fatter tails' than the normal distribution. In this case the simulated errors will be smaller than they should be and consequently the simulated maximum demands will be smaller. Weather corrected 10 and 50 POE demands will be lower than they should be in this case.

42 For simplicity we have assumed that there are four months in summer, December to March (inclusive). We deleted two sevenths of those days (35 days) as weekends and a further 11 days as holidays (one for Australia Day and ten weekdays for the Christmas/ New Year break). In practice, the number of omissions will vary each year depending on when weekends and public holidays fall. In some places it may be appropriate to allow more days for the Christmas break if, for example, factory closures are for three weeks rather than two.

simulated annual maxima. The 10, 50 and 90 POE values from the simulated data are shown in Table 3.

FIGURE 7 **DISTRIBUTION OF SIMULATED MAXIMUM DEMANDS AND POE (PERCENTILE) VALUES - EXAMPLE**




*SOURCE: ACIL ALLEN CONSULTING*

TABLE 3 **WEATHER CORRECTED DEMAND AT 10, 50 AND 90 POE - EXAMPLE**

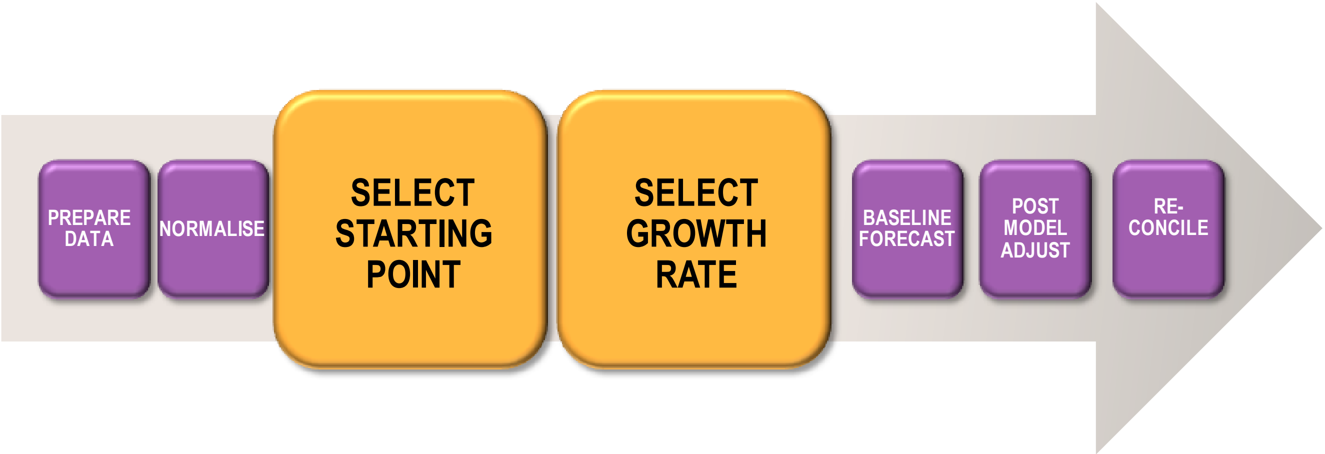
| POE level | Demand (MW) |
|---|---|
| 10 | 94.5 |
| 50 | 87.9 |
| 90 | 81.4 |

This process is repeated for each historical season for which data are available and for which a relationship between maximum demand and weather has been estimated.

After this process has been completed, there will be a set of estimated 10 and 50 POE maximum summer and winter demands, one for each historical year for which data was available.

These historical weather normalised maximum demands can then be used as the basis for estimating historical growth rates at each CP. The most recent weather normalised set of maximum demands can also be used as the starting points to which a set of growth rates can be applied.

# 6 Select the starting point and growth rate




With a set of weather normalised historical maximum demands at each CP the starting point and growth rates can be identified.

## 6.1 Selecting the starting point

The starting point is the value from which the growth rate is applied. It should be used in the last year for which data are available.

Conceptually there are two choices for the starting point. It could be either:

1. the last observed weather normalised maximum demand (off the point)
2. the fitted value for the relevant year from a regression of demand on explanatory variables (off the line).

It is extremely unlikely that the point and line produce the same value in the last year for which data are available. If the two options are close to one another the preferred approach is to take the starting point 'off the line'. However, if the line and point are 'far' from one another the preferred option is to start 'off the point'. Starting 'off the point' does not involve a step change from the last year for which data are available to the first forecast year.

'Close' and 'far' should be defined by comparing the size of the difference to the total demand and considering the implications, for example for future investment, of one starting point rather than the other. The question to consider is whether it is plausible that the step change implied by starting 'off the line' will happen. This should be supported by reasons based on metadata surrounding the CP and documented for transparency.

In any case, 'the line' must be estimated, meaning that a regression should be estimated between the weather normalised demands and an appropriate explanatory variable. Two options should be tested:

1. regression on a time trend

2. regression on the population of the service area of the CP.

As discussed in section 2.4, historical transfers and large block loads should be removed from the historical data before these regressions are fitted. This will prevent any distortion of the true underlying growth rate and will help avoid double counting of block loads when these are added to the forecasts at the end of the process.

There is no theoretical basis to prefer one of these regressions over the other. Therefore, both regressions should be estimated and the choice made empirically.

The forecaster should consider whether the coefficient in the population regression makes theoretical sense, in particular whether it has the expected sign.[43] If not, it should be disregarded. There is also a strong likelihood that a model using only population as an explanatory variable will be mis-specified if other drivers are significant. In this case it would be more appropriate to use the time trend.

If the population regression has the expected sign, the goodness of fit of both regressions should be considered. Preference should be given to the variable that produces the best fit.

In some cases, especially areas consisting primarily of industrial customers, the coefficient on either regression may be very small, indicating little or no growth. In these cases, growth may be accounted for through identifiable block loads added to the forecasts later.

If the coefficient on population is negative, implying that electricity demand *falls* as population *rises*, it should be disregarded.

If not, the goodness of fit should be examined. The regression that exhibits the best fit should be chosen.

In residential areas, it would be expected that the population-based approach would provide a superior fit. The same may be true in primarily commercial areas because population is expected to be correlated with economic activity or, in other words, commercial activities typically service the surrounding population.

This is unlikely to hold in industrial areas. In these areas the trend line seems likely to provide a superior fit, though it would not be surprising if the analysis showed that the load was stationary (i.e. flat over time).

## 6.2 Selecting the growth rate

When starting points have been selected, it remains to select the appropriate growth rate. The appropriate choice depends on the regression used to develop the starting point.

If it was decided to use population for that regression then maximum demand should be projected forwards using a suitable population projection. The population projection should be ‘sense checked’ and projections that appear overly optimistic should be avoided. Population projections are available from the ABS for a range of scenarios.

If no suitable projection is available for the area in question, which is unlikely, then the long term (historical) growth rate from the population data could be substituted.[44]

[43] There is no theoretical basis to predict the sign of the coefficient in the trend regression. That is, demand might increase or decrease over time.

[44] ACIL Allen recommends some caution in applying population forecasts produced by local Governments as these have a tendency to present an unrealistically rosy picture of the regions prospects, often not matching recent behaviour.

If the trend approach was used to selecting the starting point, demand should be projected using a continuation of that trend, though it would not be inappropriate to apply population growth in place of the trend..

If 10 and 50 POE demand projections are both required (or other levels), demand should be projected from the weather normalised historical data for both the 10 and 50 POE maximum demand independently. That is, the preferred approach is not to project the 50 POE series and ‘scale up’ to 10 POE (or vice versa).

Using the estimated relationship between maximum demand and population at each level (10 and 50 POE) separately creates projections of maximum demand at these two levels independently and allows for the possibility that demand grows at different rates at these two levels, which simple scaling does not make possible.

### 6.2.1 Fixed rules to apply

Before the baseline forecasts are produced a number of rules should be applied to ‘sense check’ the baseline growth rates. These will prevent nonsensical outcomes and also to help reduce the impact of unreasonable growth rates generated by the process.

Rules to be considered include:

- CPs known to be dominated by flat loads (perhaps some industrial areas[45]) should have low growth rates or should be constrained to having a zero growth rate in the baseline forecasts, with all growth ultimately coming through block loads.
- It is unusual for (distribution) ZSS to exhibit annual growth exceeding four per cent per annum. Given that CPs are mainly the aggregate of a number of ZSS, it is even more unlikely that maximum demand at a CP would grow as fast as four per cent per annum.[46] Therefore, consideration should be given to setting four per cent as an upper limit to the baseline growth rate.
- The 50 POE growth rate should be prevented from exceeding the 10 POE growth rate to avoid the problem where the 50 POE forecast exceeds the 10 POE forecast.

These are guides for sense checking. They should not be taken as ‘hard and fast’ rules. In practice it would be helpful to discuss the CPs to which these rules may apply with local experts to ensure that their ‘story’ is understood. This is discussed below.

### 6.2.2 Role of judgment and local experts

Selecting the appropriate growth rate to use in developing baseline forecasts is partly a judgement based process. There are a number of reasons that judgement may need to be applied, including that the historical data upon which forecasts are based could be affected by load transfers that have been missed during the data cleansing and processing process.

It is also possible that history is not a good guide to the future in some areas. This may be due to factors such as a greenfield development that is being serviced by a specific CP reaching completion, with all areas of potential development now being exhausted. This would cause growth to be slower in future than in the recent past. On the other hand, substantial greenfield development in an area that was sparsely

45 At the CP level this may apply to transmission connected industrial customers.

46 This would require that the average growth of each ZSS connected to the CP is four per cent per annum or more.

populated historically would mean that future growth may be much faster than history.

In any of these situations, and possibly others, growth rates that are derived using econometric analysis, which is based on historical relationships, may lead to the wrong answers. The forecaster needs to consider the likely future of an area before blindly applying the calculated growth rates to each CP.

The process of determining growth rates at each CP is not a simple mechanical one. It requires the forecaster to exercise their own judgement as well as the expert knowledge of planning engineers with strong local area knowledge. It would be appropriate to discuss baseline forecasts for each CP with appropriately qualified experts. This is not necessarily so that those experts can choose the growth rates, but so that they can identify issues such as those discussed above for the forecaster to take into account.

For CPs that supply predominantly residential areas, calculated growth rates should be compared against historical and forecast population growth rates in and around each CP as a means of validating the calculated results

### 6.2.3 Treatment of new connection points

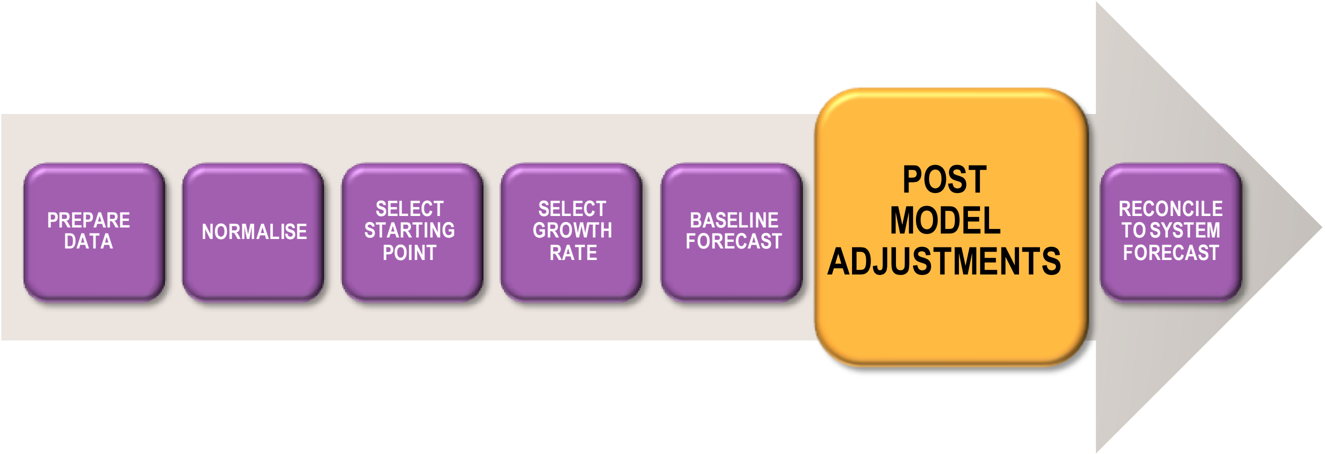
There is also a question about how to treat new CPs for which there are little or no historical data.

Typically the starting point is known. The challenge is to identify the appropriate growth rate. There are a number of ways to proceed:

- choose a similar CP that has been operating for some years for which the new CP is expected to have a similar development cycle or profile - the growth rates applied to the new CP will then follow the same path as that of the established and mature CP
- adopt the forecast rate of population growth for the area (assuming it is substantially residential or commercial)
- apply a judgement based growth rate drawing on the knowledge of a local expert
- in the absence of any of the above options, apply an arbitrary growth rate such as the average for the distribution network in question.[47]

[47] This should happen very infrequently.

# 7 Post model adjustments




At this stage it may be necessary to make a number of post model adjustments to take into account factors that are known, but not yet incorporated into the forecasts.

These factors include:

- new large block loads, load transfers and network demand management
- impact of government policies driving factors such as the uptake of solar PV at each CP
- demand side management initiatives
- other large changes in load, such as mines disconnection etc.

## 7.1 Treatment of forecast block loads and demand management

Forecasting block loads can be difficult and involves the expert knowledge and judgement of local asset managers and planners.

The difficulty in accurately predicting these discrete loads arises from three main areas, relating to:

- the size of the new load
- the timing of the new load
- the likelihood of the new load going ahead.

Generally speaking, project proponents are often optimistic about their projects. Consistent with this, there is a tendency for them to advise NSPs that their loads will be larger and brought on sooner than a more neutral assessment may suggest. Similarly, proponents sometimes overstate the chance that the project will proceed.

The recommended approach to accounting for this begins with identifying ‘candidate’ block loads.

In identifying these loads, regard should be given to the expected size of the load relative to the 50 POE maximum demand on the CP in question. If the expected block load is smaller than five per cent of the 50 POE load, it should be disregarded as making an adjustment may amount to double counting (see section 7.1.1).

'Candidate loads' large enough to warrant an adjustment to forecasts at the CP level will be well known to relevant NSPs and local area experts.

Once a list of loads that are large enough to warrant examination is produced, the forecaster has the option to *weight or wait*. That is:

- in some cases it will be appropriate to apply probabilistic *weights* to the identified block loads and add their *expected value* to a baseline forecast (see section 7.1.1)
- in other cases it will be appropriate to *wait* until some of the uncertainty is resolved before making the adjustment. In particular, it may be appropriate to wait until connection applications are made (and fees paid) before adding block loads to the forecasts.

The best choice will depend on the number of block loads anticipated. The expected value (weighted) approach is well suited to situations where there are numerous loads anticipated and it is likely that some, but not all, proceed or where the uncertainty is not so much about whether the load will proceed, but how soon and how large it will be.

The 'wait' approach is better suited where there are relatively few block loads that can be readily identified. It may not be well suited to situations where regulatory processes 'lock in' an NSP's revenue in advance.[48]

Regardless of the approach taken, the degree of coincidence of the new load with the system peak must be evaluated correctly. Failure to do so will tend to exaggerate the block load's impact on peak demand (either 50 or 10 POE).

Correctly evaluating the likelihood, timing and size of prospective block loads is difficult, but can be dealt with in a number of ways:

- by identifying the size of any historical bias in timing, size and likelihood, and then adjusting future block loads accordingly
- by evaluating the block loads regularly, through close contact between the project sponsor and the forecaster, to quickly incorporate changing circumstances
- by conducting independent analysis of the financial viability of the industry and likelihood of commitment surrounding the prospective block load
  - for example, because a prospective mining project may be rendered uneconomic by significant commodity price falls, careful analysis of the price outlook for a commodity may help inform the estimated probability of proceeding.

The extent to which block loads are added to the forecast should also be considered when selecting the growth rate (see section 6.2). Generally, if a more conservative approach is taken to including block loads (fewer loads added), it would be appropriate to allow a higher growth rate. The reverse is also true.

---

48 In this case it may be more suitable to use a contingent project approach.

### 7.1.1 Probabilistic expected value approach

In many cases the appropriate approach to incorporating block loads into a baseline forecast will be to use the *expected value* approach. To do this, the forecaster should consider the different outcomes that may be expected from a project or group of projects and assign a probability to each. In rare cases it may be possible to do this formally, but in practice it is more likely to be an informal set of weights and scenarios. For example, in adding residential subdivisions to zone substation forecasts some DNSPs routinely add 80 per cent of the anticipated load to account for uncertainty.[49]

During this process, loads that are projected to come online three years or more into the forecast period should be discounted heavily or even set to zero if a conservative approach is adopted. This is to account for the considerable economic uncertainty associated with projects more than a few years into the future. It is very common for projects to be re-evaluated and put on hold in the period leading up to the load's commencement.

Smaller loads may warrant an adjustment if they are unique to the area, such as the first mine to be built in a certain area.

This is an aspect of the forecasts that would be considered during the review with local area experts.

### 7.1.2 Potential double counting of block loads

Where econometric analysis is applied to a historical time series and individual block loads are added to the same forecast, there is a prospect of double counting the impact of block or discrete loads. Because block loads are included in the historical data, any fitted regression line will incorporate their contribution to the growth in the peak demand over time. Hence adding expected new block loads to the forecasts may result in double-counting leading to inflated forecasts.

The potential for double counting is typically reduced by applying a threshold to the size of future discrete loads, with only loads exceeding a certain size being added onto the forecast. Smaller loads are assumed to comprise part of the underlying growth determined by the historical trend. Alternatively, where an accurate record of historical block loads is available, the effect of these loads could be removed from the historical time series prior to undertaking any econometric analysis. However, accurate historical block load data is usually difficult to access.

The preferred approach is to use a formal size threshold of five per cent[50] of the measured load at a given CP for new block loads. Loads that are expected to be smaller than this would not warrant an adjustment. Larger loads would be considered using the expected value approach.

### 7.1.3 Adjust for coincidence between block load and connection point peak

It is also important to recognise that a new block load may not peak coincidentally with the CP to which it is attached. For this reason, its contribution to the CP peak may be less than its own peak.

---

49 This is just an example. A single residential subdivision is unlikely to be large enough to warrant addition to a CP forecast.

50 The size of this threshold is somewhat arbitrary and experience has played a large role in determining it. Its purpose is to keep the time series smooth and not allow discrete shifts to bias results. A threshold that is lower than 5% will mean that many small loads that could be best captured and treated as organic growth will need to be treated as block loads. A load that is about 5% of the measured load on a CP will be reasonably rare and lead to a discrete jump in the time series. Thresholds that are much larger than 5% will then lead to a time series that contains some large discrete jumps for which we haven't made any adjustment. These will prove inadequate.

Care needs to be exercised to make an adjustment for this effect in the forecasts, by obtaining an estimate of the degree of coincidence between the block load and its CP.

## 7.2 Network demand management and transfers

Conceptually, the same approach should be taken to incorporating network demand management (including embedded generation) and transfers as is taken to block loads. That is, likely changes in load should be identified and incorporated into the forecasts if there is sufficient confidence that they will occur.

In practice this is only likely to be possible with a detailed understanding of the relevant NSP's plan for the future operation of the network. However, even then, changes may be unexpected, especially over the longer term.

## 7.3 Small embedded generation

The use of small embedded generation is also likely to have an impact on electricity demand. The most commonly used form of embedded generation in Australia at the moment is solar photovoltaic (PV) systems, which are referred to here, though the same principles would be applied if other forms of embedded generation become more common.

The uptake of small-scale solar PV systems in Australia has increased at an extremely high rate over the last five years. Data regarding the exact number and capacity of systems are mixed due to regional differences, but it is clear that there has been very rapid uptake. For example, in its 2012 Rooftop PV Information Paper, AEMO reported that the installed capacity of solar PV in the NEM rose from approximately 23 MW in 2008 to 1,450 MW in 2012.

In any case, there is no doubt that solar PV systems have reduced maximum demand to some extent, at least in places where maximum demand occurs in summer.[51]

### 7.3.1 The impact of solar PV systems on maximum demand

The amount that demand on the network is reduced is related to the *total* output of the PV system, not only the portion of that output that is exported.

That is, from the CP perspective, it does not matter whether the output of a PV system is used by its owner, their neighbour or another customer nearby. What matters is that the output of the PV system need not flow through the CP.[52]

In turn, the output of a solar PV system is a function of the surface area of the panel, its technical efficiency and the insolation, or amount of light that falls on it. [53] The first two factors are partly at the discretion of the owner (subject to technical and other limits). Insolation depends mainly on latitude, though the extent to which an individual solar PV system 'sees' light depends on other factors such as cloud cover

[51] Strictly speaking, embedded generation systems have no impact on demand for electricity. Rather, they provide an alternative source from which a customer can take electricity. Whether a customer uses electricity from a PV system or from 'the grid' its demand is the same.

However, from a network perspective this does not matter; from that perspective the key question is the demand that will be placed on the network. Measured in this way, demand is reduced by small-scale PV systems.

[52] This could not necessarily be said about the distribution network, which would need to transfer the output of the PV system between customers.

[53] Solar insolation is the amount of solar radiation that reaches the earth's surface. In a sense it is the fuel that a solar PV system uses to generate electricity.

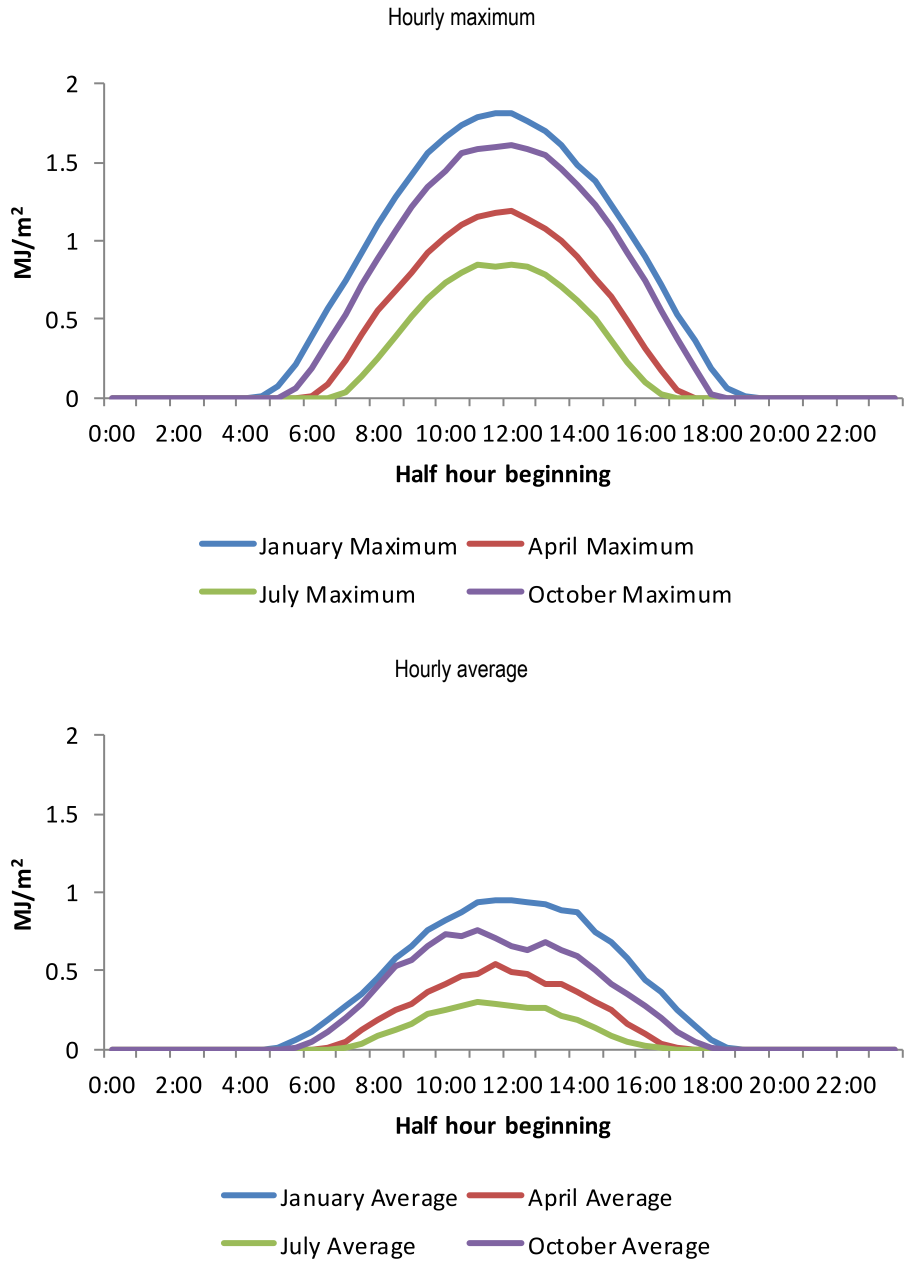
at a particular time and the size and proximity of trees and other objects that may cast shade on the panel.

However, in aggregate terms it would be reasonable to disregard these individual factors and base a projection on insolation data, which can be obtained from the BOM.

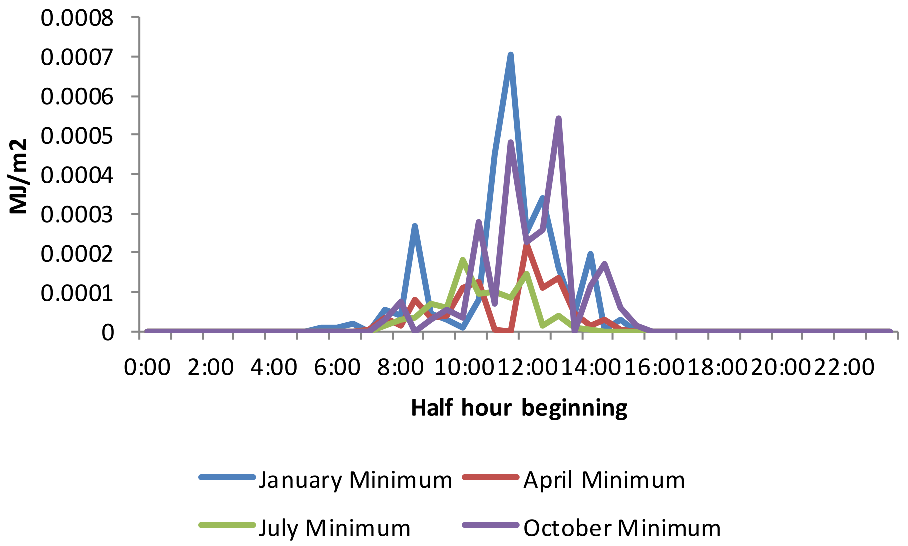
For example, Figure 8 illustrates the solar insolation as observed by the BOM at Melbourne Airport in 2010, 2011 and 2012. To account for daily variability, the figure shows the maximum, minimum and average insolation observed during each half hour of the day on a monthly basis.

Broadly, the pattern illustrated in Figure 8 is as would be expected. In particular, insolation begins later in the day, finishes earlier and is generally lower during the cooler months. It starts earlier, lasts longer and is generally higher during the warmer months. However, the third pane of Figure 8 (which shows the hourly minima) shows that insolation can be very low at any time in the year. This has potentially severe implications for the reliability of photovoltaic DG systems.

FIGURE 8 **SOLAR INSOLATION PROFILE – MELBOURNE AIRPORT**

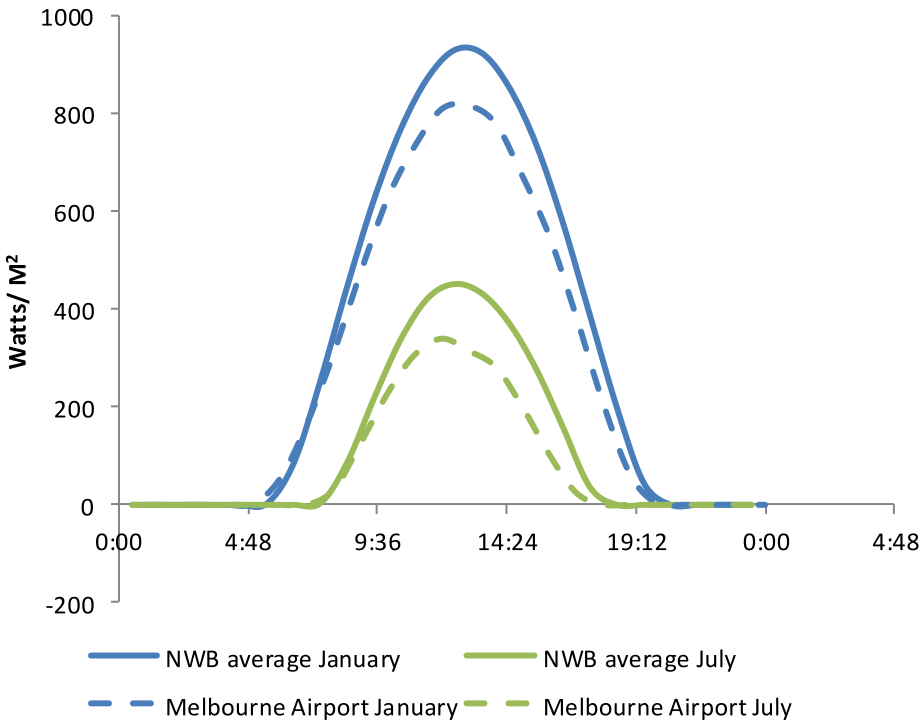


Hourly minimum (note smaller scale)




The level of solar insolation changes with latitude and other climatic conditions, though the shape tends not to change as much. This is illustrated in Figure 9, which shows the Melbourne airport average insolation for January and July alongside corresponding data for North West Bend (NWB) in South Australia.

FIGURE 9 **SOLAR INSOLATION AT MELBOURNE AIRPORT AND NORTH WEST BEND (SA)**




The difference in insolation in various parts of Australia is reflected in the Renewable Energy Target and its predecessor schemes, which have divided Australia into four zones for the purpose of deeming the amount of electricity a typical solar PV system would generate in different places (on a capacity installed basis).[54] Output factors

[54] The total energy output is reflected by the area under the insolation curve.

have been published for each zone which provide an estimate of the energy that a solar PV system will generate, on average, in each zone (in MWh per kW per year).

By combining solar insolation data and the output factors applicable to a given CP's location, it is possible to create an hourly estimate of the output of a small-scale solar PV system in the area supplied by each CP on a per kW installed basis. This can be done by weighting the total energy output of a notional 1 kW system, as implied by the output factor, by hourly insolation data from the BOM. It would be appropriate to average the hourly insolation data to the monthly level to account for daily variation.

This hourly estimate can then be used to estimate the impact of a notional 'fleet' of solar PV systems in the area supplied by a CP by scaling it up to correspond with the estimated capacity of the solar PV systems in the area in question. NSPs may have data relating to applications for new systems for a few months but, in the absence of a preferred alternative, the approach would be to assume that uptake will continue at the rate observed in recent months.[55]

This will produce an estimate of the total output of additional PV systems in the area supplied by the CP in question in each hour of the month. The remaining step is to determine in which hour peak demand is likely to occur and subtract the output of PV systems estimated in that hour.

A complicating factor is that DNSPs have noted that solar PV systems seem to be 'pushing back' peak demand to later in the day as well as reducing it. This is due to the timing of the PV output and demand peak as illustrated in Figure 10.

Figure 10 compares the output of a notional 2.5 kW solar PV system installed at (or near) Melbourne Airport with the electricity demand of an average residential customer at the same location.[56] Both variables are averaged over the period from November to March. Demand is presented in two forms. First, the green unbroken line shows the customers actual use of electricity.[57] Second, the red dashed line shows their imports from the grid or, in other words, the demand they place on the network. The latter is negative during the middle of the day because the customer is generating more electricity than they are using and exporting the surplus.

The impact the customer's solar PV system has on their maximum demand is highlighted in the lower pane of Figure 10, which shows the same data at a higher resolution.

It is apparent that the PV system causes the customer's maximum demand to be smaller and later than it would otherwise have been.

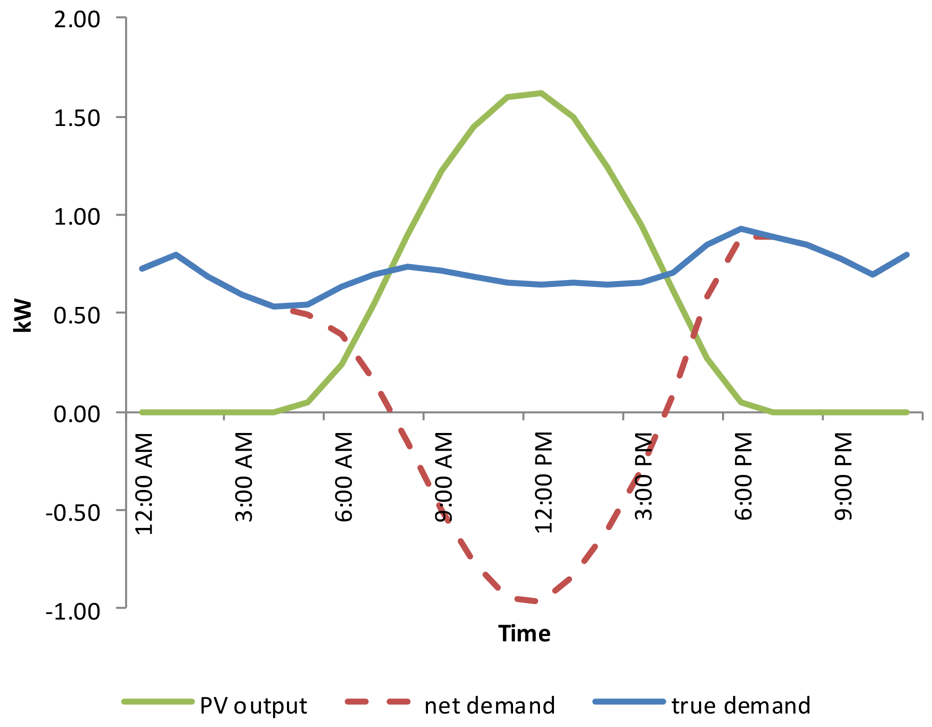
It may be appropriate to moderate the estimated impact of PV systems on peak demand to take account of this effect. This could be done by choosing the estimated output from an hour or two later in the day than the hour thought most likely to exhibit peak demand.

55 In some areas it would be appropriate to use different definitions of 'recent'. The objective would be to observe the rate of uptake since the most recent change (reduction) in feed in tariff or other relevant policy initiative. The timing varies state by state.

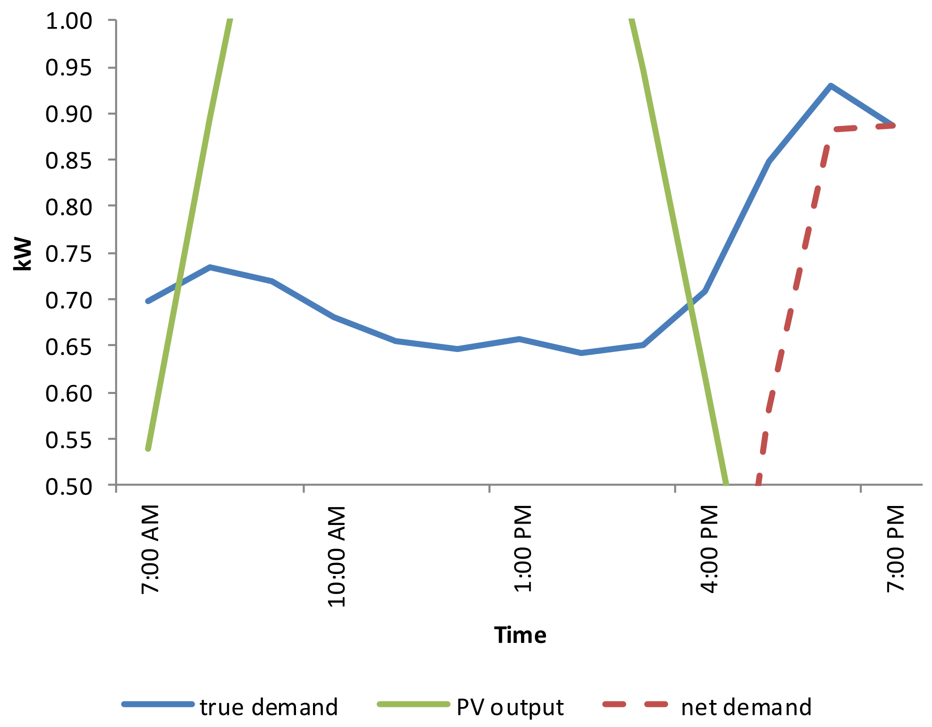
56 In other words, the solar output is based on Melbourne Airport solar insolation data and the consumption is based on the NSLP.

57 This is based on the NSLP so strictly speaking it reflects demand net of the output of solar panels. For the current illustration this is not important as long as the shape of the NSLP is a reasonable approximation of a typical customer's demand profile.

FIGURE 10 **ILLUSTRATIVE RESIDENTIAL CUSTOMER DEMAND AND SOLAR PV OUTPUT**




Higher resolution




*SOURCE:* ACIL ALLEN CONSULTING

# 8 Reconciling with an independently derived system demand forecast


Best practice spatial load forecasting requires that both top-down and bottom-up spatial forecasts are produced independently of one another.

Top-down macro level forecasts have the advantage of allowing the methodology to incorporate the impacts of broader macroeconomic and demographic aggregates. System level data is usually more accurately and regularly recorded, covers a broader area and is more amenable to the fitting of econometric models which can be used to generate forecasts. Furthermore, system level data does not require adjustment for transfers or block load increases with future block loads usually accounted for by variation in the key drivers.

Bottom-up spatial forecasts are important as they allow the capture of underlying characteristics of the areas serviced by individual CPs. For example, variations across the spatial environment based on observed and planned local developments (local population growth, housing developments etc.) lead to different growth requirements across the system space. Spatial level demand forecasts are therefore necessary to capture the growth rates of individual CPs relative to one another.

It is important, also, that the forecasts generated by the methodology should be consistent. That is, allowing for diversity between different levels of the network and losses, the forecasts produced at the CP level should be the same as the system level forecasts.

## 8.1 Reasons for reconciling

### 8.1.1 Ability to better incorporate high level macroeconomic/policy drivers

Detailed disaggregated economic data is very difficult to obtain for regional Australia. For this reason, it is very difficult to construct a model of maximum demand at the CP level which incorporates all of the specific drivers in to the model specification. Therefore, we consider it reasonable to reconcile the lower level CP forecasts to an independently derived system maximum demand forecast that is better able to capture the impact of high level macroeconomic drivers.

The system model is better suited to capturing the key drivers of maximum demand for a number of reasons:

- State level data on which the model is based are readily available and accurate
- State or DNSP-wide maximum demand data exhibits greater regularity and are less affected by non-weather related randomness such as transfers and switching.

For example, CPs which have very flat historical load characteristics, for reasons such as being predominantly mining related loads, can be excluded from any reconciliation process.

The higher level forecasts can also incorporate the impact of system wide policy and other impacts, such as the uptake of solar PV and network demand management by the DNSP.

### 8.1.2 Ability to apply scenarios

With a system model based on economic drivers it is possible to prepare demand projections on different projections of those drivers. For example, it would be possible to project maximum electricity demand under high, medium or low economic growth assumptions.[58]

This type of scenario analysis is more appropriately conducted at the system level. At the CP level it will typically be no more sophisticated than altering growth rates arbitrarily.

However, through the reconciliation process, the high, medium and low scenario is applied to all the individual CPs.

### 8.1.3 Better weather normalisation on smoother data

Weather normalisation at the system maximum demand level usually results in models with significantly higher explanatory power compared to those developed for smaller spatial substations.

The system based weather normalisation procedure will generally result in less erratic weather corrections, which are also more reliable.

While it is still necessary for weather normalisation to be applied to each of the CPs under consideration, the reconciliation with the system level forecasts can reduce the impact of any systematic biases which may be present at the CP level. The smoother weather correction at the system level is therefore transmitted to each of the CPs, while keeping the relative changes between the CPs intact.

### 8.1.4 No transfers and switching and minimal impact of block loads

At the system level, the historical time series data remains unaffected by transfers and switching events. In addition to this, most block loads are also too small to make any impact on the system maximum demand time series and simply blend into the historical trend. For this reason, block loads generally do not need to be taken into account and added onto the system level forecasts as they would be for many CPs or other spatial forecasts.

---

[58] Depending on specifics it may be possible to test more complex scenarios such as exchange rate changes etc.

The exception to this rule, of course, is if the block loads are extremely large. These would need to be added onto the system maximum demand forecast as well as at the CP where it is expected to impact.

It is likely that there will be only a very small number of such loads that are large enough to require an adjustment to the system maximum demand forecasts. One possible example could be some of the very large Liquefied Natural Gas developments currently under construction in Queensland.

Block loads that are added at both the system and CP level, should be added to both sets of forecasts after the reconciliation between the system maximum demand and the CPs has taken place. This is to prevent the very large load from impacting other CPs (through the reconciliation) unrelated to the new load.

### 8.1.5 Ability to incorporate higher level policy factors

While it may be difficult to assess the impact of high level policy factors at the spatial level, it is possible to build their impact into the forecasts by taking account of them within the system level forecasts and then reconciling the CP forecasts to the system forecasts.

### 8.1.6 Spatial forecasts limited to simplistic models

Inputs to spatial or regional forecasting methodologies are typically limited to historic growth adjusted to account for the judgement of network planners. This lacks the transparency and repeatability that is possible at the system level.

Bottom-up maximum demand forecasts tend to rely heavily on trend analysis which, by its nature, assumes a continuation of historical conditions into the future. If there is an expectation that these conditions will be different in the future, then such forecasts will perform poorly. Reconciling them with an independent system level forecast which takes these expected changes into account helps to reduce this problem.

### 8.1.7 Check against independently produced maximum demand

Another reason to reconcile the forecasts against an independently derived set of system level forecasts is that it serves as a means of checking the validity of the forecasts against an independent alternative.

Before the formal reconciliation is applied, the sum of the coincident CP forecasts can be compared with the independent system level forecasts for each jurisdiction, and the divergence between the two sets of forecasts analysed.

## 8.2 Approach to reconciliation

### 8.2.1 Calculation of system coincident maximum demands and coincidence factors

The first step in the process of reconciling the system level forecast with the CP forecasts is to determine the historical system coincident maximum demands at each CP. This is the maximum demand that occurred at each CP at the time of the system peak (either DNSP or state wide peak). The Coincidence factor is the ratio, calculated at each CP, of :

- demand at the CP when system demand peaks coincident peak to
- maximum demand experienced at the CP.

### 8.2.2 Apply coincidence factors to non-coincident connection point forecasts to obtain system coincident forecasts

ACIL Allen recommends the relatively simple approach of using coincidence factors to move from the CP level to the system total.

The historical coincidence factors calculated at each CP form the basis of the conversion from non-coincident to coincident forecasts. The system coincident forecast is obtained by multiplying the non-coincident forecast by the coincidence factor.

**Best approach to calculation of coincidence factors**

There is a question as to which coincidence factor to apply into the forecast period. Several approaches are possible:

- the coincidence factor in the most recent year
- the average coincidence factor over a number of years
- a coincidence factor that trends over time.

The forecast period coincidence factors should be based on the behaviour of a set of historical coincidence factors. If there is no clear trend evident in the coincidence factors, then an average value based on say the last 3 or 5 years should be applied into the forecast period.

Using the average is superior to applying the most recent year, particularly in the case of industrial CPs, as the peak is determined by industrial processes rather than time of day behaviours or temperature sensitive load. As a result, the historical coincidence factors can display significant volatility from year to year. A single coincidence factor may therefore not be representative.

Where a trend is evident and there appear to be clear theoretical grounds for the continuation of the trend, a linear time trend could be applied to obtain the coincidence factors to be applied during the forecast period.

### 8.2.3 Apply proportional adjustment to reconcile system level forecasts to the connection point level

Once the coincident level demands are obtained for each CP, and these are aggregated up to the system level, a proportional adjustment should be applied to each of the CP forecasts to formally reconcile the spatial forecasts with those at the system level. The reconciliation will be applied and will differ for each year in the forecast period, so as to make the coincident demand forecasts aggregate exactly up to the forecasts produced at the system level.

ACIL Allen believes if the CP forecasts are credible and reasonable, then the adjustment should be relatively small in percentage terms, at least in the early years of the forecast period.[59]

An alternative approach to the adjustment might exclude CPs that are not expected to generate any growth during the forecast period so as to prevent them from participating in the adjustment.

The key point is that reconciliation does not adjust the relative forecasts between the CPs in any substantial way. Its main purpose is to make the lower level forecasts logically consistent with those at the system level.

---

[59] The exception to this is if there is a significant change in the projected drivers of demand compared to history such as an economic recession of recovery. This would cause the system forecast to diverge from the (trend based) CP forecasts.

### 8.2.4 Conversion back to non-coincident forecasts

Once the reconciled coincident forecasts are calculated for each CP, then these are converted back to non-coincident maximum demand forecasts by multiplying them by the inverse of the relevant coincidence factor.

# 9 Issues for further analysis

## 9.1 Price elasticity of demand

Typically, the price of a product is a key driver of demand for that product. Generally, the relationship between price and demand is summarised in the price elasticity of demand, often referred to simply as elasticity.

The price elasticity of demand for different products can vary widely so a price elasticity of demand is not fully defined unless the product being demanded is identified. When the product in question is electricity the word 'demand' has a very particular meaning, which complicates the analysis.

In this context, it is important to distinguish between *maximum demand* and *energy consumption.* Therefore, it is important to distinguish between:

1. the price elasticity of demand for energy
2. the price elasticity of demand for megawatts.[60]

The first of these two concepts, the price elasticity of demand for energy, is readily understandable. This is the relationship between the price of energy and the quantity, in watt-hours, of energy demanded over time. Broadly, if price increases, customers might be expected to switch fuels from electricity to gas and, thereby, reduce their consumption of electricity. Alternatively, an increase in price might motivate them to improve their energy efficiency and achieve the same outcomes by using less electricity.[61]

The second concept is less widely discussed. It is the relationship between the price of electricity and the quantity that the customer will demand when their demand is at its maximum. This elasticity also deals with reduced consumption in response to increased price, but the reduction must occur at very specific times. Depending on tariff structures, this elasticity might also deal with the possibility that an increase in electricity price might cause a customer to engage in load shifting from times of high price (and demand) to times of lower price (and demand).

These elasticities are not conceptually the same. They would take different values.

Our review of the literature in this area indicates that the majority of studies of 'elasticity of electricity demand' focus on the price elasticity of demand for energy. Previous studies have typically given little or no consideration to the price elasticity of demand for megawatts.

Elasticity estimates for energy consumption were summarised by Fan and Hyndman in 2008.[62] They identified that estimates of the elasticity of demand (for energy consumption) ranges from -0.1 to -0.7 as shown in Table 4.

---

60 Given our earlier definition of demand, this would be more accurately named the price elasticity of demand for demand rather than identifying it by the units of measure. However, this is a cumbersome name so we refer to the price elasticity of demand for megawatts.

61 The outcomes in question might range from the level of comfort provided (for a residential customer) to the quantity of aluminium produced (for a commercial customer).

62 Fan, S and Hyndman, R J "The price elasticity of electricity demand in South Australia and Victoria", 2010. available online, www.buseco.monash.edu.au/ebs/pubs/wpapers/2010/wp16-10.pdf

TABLE 4 **PRICE ELASTICITY OF ENERGY DEMAND ACROSS JURISDICTIONS**

| Researcher | Year | Region | Sector | Elasticity |
|---|---|---|---|---|
| Bohi & Zimmerman | 1984 | U.S (various utilities) | Residential, industrial and commercial | Residential sector<br>Short-run: -0.2<br>Long-run: -0.7 |
| Filippini | 1999 | Swiss<br>(40 cities) | Aggregation | -0.3 |
| Beenstock et al. | 1999 | Israel | Residential and industrial | Residential<br>-0.21 to -0.58<br>Industrial<br>-0.002 to -0.44 |
| NIEIR | 2007 | Australia | Residential, industrial and commercial | Residential: -0.25<br>industrial: -0.38<br>commercial:- 0.35 |
| King & Shatrawka | 1994 | England | Residential and industrial | Substitution elasticity<br>Inter-day: -0.1 to -0.2<br>Intra-day: -0.01 to -0.02 |
| Patrick & Wolak | 1997 | England and Wales | Industrial and commercial | Water supply industry:<br>-0.142 to -0.27 |
| King | 2003 | California | Residential | -0.1 to -0.4. |
| Reiss | 2005 | California | Residential | -0.39 |
| Faruqui & George | 2005 | California | Residential, industrial and commercial | Substitution elasticity:<br>-0.09 |
| Taylor et al. | 2005 | U.K. | Industrial | -0.05 to -0.26 |

*SOURCE:* PRICE ELASTICITY OF ELECTRICITY DEMAND IN SOUTH AUSTRALIA, SHU FAN AND ROB HYNDMAN, DEPARTMENT OF ECONOMETRICS AND BUSINESS STATISTICS WORKING PAPER 16/10, MONASH UNIVERSITY, AUGUST 2010

The elasticities presented in Table 4 are estimates of the elasticity of demand for energy consumption, not the elasticity of demand for megawatts.

We are not aware of estimates of the latter, though the reasonable expectation is that their absolute value would be lower, not higher, than the estimates shown in Table 4.

In theory, the elasticity of demand for megawatts can be determined from a properly specified econometric model of demand. The model used to forecast demand at the system level would presumably be able to estimate it. In our view this is a more appropriate way to introduce the impact of price into the forecasting methodology than making adjustments to the CP forecasts directly.

In our view it is worthwhile examining the possibility that electricity demand is sensitive to price, but it would not be surprising if no reliable relationship was found and a forecasting model of electricity demand excluded a price variable.[63]

## 9.2 Energy efficiency

One factor that should be considered in forecasting electricity demand but that is not discussed in great detail elsewhere in this report is energy efficiency.

In recent years significant effort has gone into improving Australians' energy efficiency. This is likely to have been due partly to rising energy prices and partly to various policy efforts made by Australian Governments in recent years.

63 To be clear, this statement is limited to maximum demand. It does not extend to models of demand for energy.

Energy efficiency improvements can be thought of as comprising two parts:

- behavioural changes – customers may improve their energy efficiency by being more sparing in their use, for example by turning lights off in office buildings overnight
- technical changes – customers may choose appliances that use less energy to perform the same task, for example LCD televisions use less energy than plasma televisions.

While it may be possible to estimate the change in energy use that is expected as a result of these two factors, incorporating the effect of energy efficiency into forecasts is not as simple as this. There are two further complications to consider:

1. the extent of double counting between reductions in demand forecasts due to energy efficiency adjustments and reductions due to price projections (through price elasticity)
2. the extent to which energy efficiency changes will reduce *maximum* demand.

The first issue arises because, as electricity price increases, the incentive on customers to be more energy efficient also increases. This is likely to be incorporated in the system level forecast to which the CP forecasts are reconciled. That is, the system level forecast would typically incorporate a variable for historical prices, use this to estimate the price electricity of demand[64] and price, and include a projection of price in producing forecasts.

The risk of double counting arises if changes in electricity use in the past have caused customers to become more energy efficient. That is, all else being equal as energy price increases, customers have a greater incentive to be energy efficient. This may manifest itself as increased behavioural energy efficiency or by making the 'business case' for more efficient appliances more attractive. The price elasticity of demand inherently 'captures' improvements in energy efficiency.

Therefore, the challenge is to identify improvements in energy efficiency additional to the improvements that would be driven by projected changes in price. In practice, this might be done by estimating the number of technical changes that a given policy is expected to make and then making a further adjustment to account for the impact of rising price. For example, the Victorian Energy Efficiency Target might be estimated to cause a certain number of globes to be replaced with energy efficient options, but perhaps half of these might be assigned to price increases. Only the remaining half would need to be considered in making an adjustment to demand forecasts.

The example of energy efficiency improvements being achieved by replacing light globes highlights the second issue, which is that energy efficiency typically targets *average* demand (energy use) rather than *maximum* demand.

Consider an energy efficiency policy that focusses solely on replacing residential light globes with more efficient options. In practice this policy may drive little energy efficiency improvement today because the efficiency of lighting has already been improved substantially. Regardless of this, though, any change that was made to energy efficiency would be made almost entirely at night, when lights are used. In most parts of Australia this would have no impact on maximum demand at all because maximum demand tends to occur during daylight hours (on summer afternoons). Figure 11 illustrates this point by providing a stylised model of the structure of electricity demand. The data were synthesised to illustrate the typical

[64] One way of expressing this relationship is as an elasticity, though it would not necessarily be structured this way in the model.

'shape' of demand on a summer's day when annual maximum demand might be observed.

In Figure 11, industrial loads are flat from 7:00AM until 9:00PM. Outside these times they are at 80 per cent of their day time level. Commercial loads are also flat between these hours, and drop to one third of their daytime level overnight.

Residential loads are in three parts:

- residential flat loads
- residential 'other loads', accounting for all residential loads not accounted for elsewhere
- residential lighting
- residential air conditioning.

Residential flat loads are at the same level throughout the day, accounting for 'always on' appliances such as refrigerators.

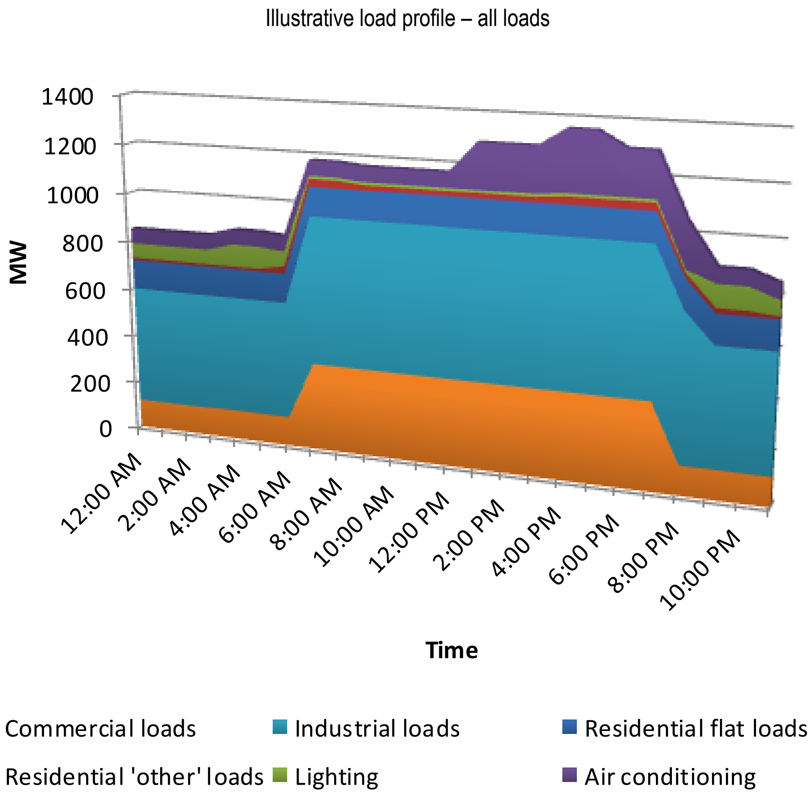
Residential 'other loads' account for appliances such as televisions and other entertainment appliances and cooking appliances. They increase from a low overnight base at 6:00AM and then drop back somewhat during the working day. They increase again from 4:00 PM before returning to overnight levels at 11:00PM.

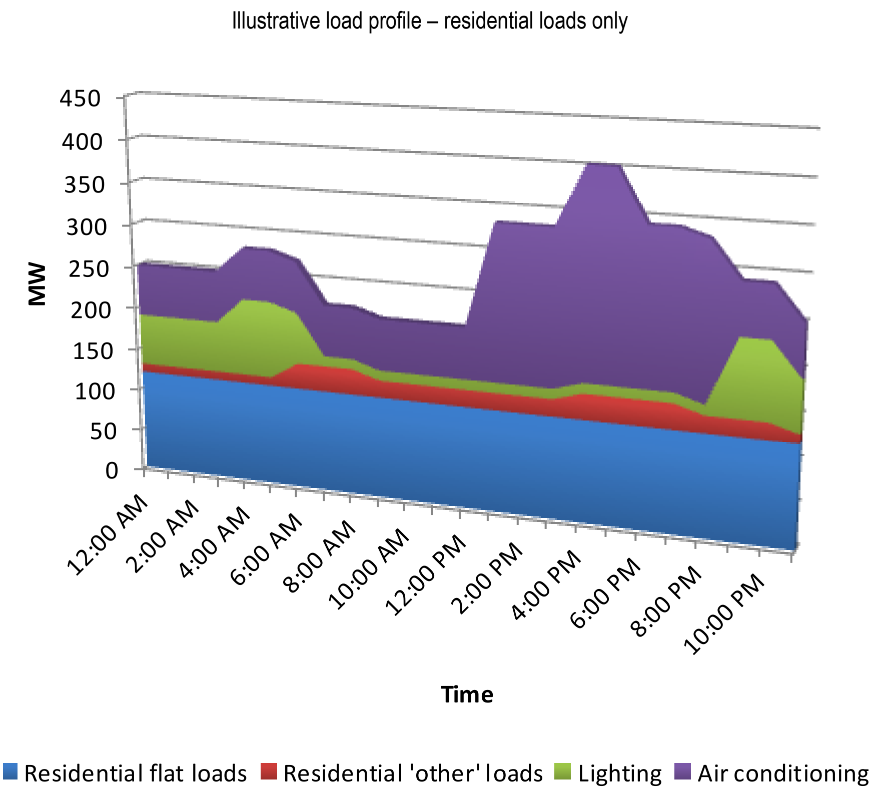
Residential lighting is at a low level before 4:00AM when it increases until 6:00AM to coincide approximately with sunrise in summer. Lighting is then at a low level, though not zero, until 9:00PM, to coincide approximately with sunset. It falls away again at 11:00 PM.

Residential air conditioning is at a low level until 1:00 PM when it triples. It then increases to four times its original level at 4:00 PM and remains at that level until 6:00 PM.

On the day illustrated here, maximum demand occurs between 4:00 PM and 6:00 PM. Residential air conditioning is only approximately one fifth of total demand at this time, but it is the increase in this portion of demand that causes demand to peak.

FIGURE 11 **ILLUSTRATIVE DAILY LOAD RPOFILE BY COMPONENT**




*SOURCE: ACIL ALLEN CONSULTING*

While this example is only illustrative, it shows that the impact that increased energy efficiency has on maximum demand varies depending on the specifics of the way energy efficiency is improved.

Energy efficiency could only reduce maximum demand by improving the energy efficiency of appliances in use between 4:00 PM and 6:00PM on the hypothetical day shown. Further, it could not reduce demand by more than the amount that each 'portion' of demand contributes at that time.[65]

Table 5 shows the contribution that each 'portion' of demand makes to the peak.

TABLE 5 **CONTRIBUTION TO MAXIMUM DEMAND OF EACH 'PORTION' OF DEMAND – HYPOTHETICAL SUMMER'S DAY**

| Portion | Contribution to demand |
| --- | --- |
| Residential flat loads | 8.8% |
| Residential 'other' loads | 2.2% |
| Lighting | 0.9% |
| Air conditioning | 17.6% |
| Industrial loads | 44.1% |
| Commercial loads | 26.4% |

It is also possible that an improvement in energy efficiency may have no impact on maximum demand at all if it is achieved through appliances that are not used when the peak occurs. For example, the example discussed here assumes that there is no heating load in use when the peak occurs. Therefore, it would be impossible to reduce maximum demand by improving the energy efficiency of heating appliances.[66]

It is also impossible that efforts to improve energy efficiency could *increase* maximum demand. For example, energy efficient building design principles encouraged building designers to endeavour to reduce the heating of a building (whether home or other building). Designers may achieve this by designing the building to capture as much warmth from the sun as possible. The reduced heating requirement could lead to a substantial reduction in energy used for heating. However, on a hot summer's day this design may 'work against' the occupant creating a larger air conditioning requirement and thereby *increasing* demand on peak days. In this way it is possible that the *energy* requirement of a building reduces while the *demand* requirement increases.

This is not to say that energy efficiency cannot cause reductions in maximum demand. There is no doubt that it could do so. However, it is insufficient to assume that a general improvement in energy efficiency will lead to the same change in maximum demand.

[65] It could only reduce it by that amount if it was reduced to zero.

[66] This would not be the case in a winter peaking region, though in that case the same could be said of space coolers.

# Appendix A Drivers of electricity demand

This appendix provides an overview of the likely drivers of electricity demand. Two aspects of it should be noted up front.

First, the discussion here is presented on an 'all else being equal' basis. That is, the discussion of the effect each driver has on electricity demand assumes that every other driver is unchanged. In practice this may not hold true, in which case multiple drivers would be included in a forecasting model.

Second, the discussion here recounts theoretical expectations. It has not been empirically tested and indeed the results of empirical testing may vary across the NEM.

In addition to the discussion of likely drivers, the nature and potential source of the available data, both historical and forecast, are discussed.

## A.1 Population, household formation and customer numbers

The more people that live and work in an area, all else being equal, the more electricity they will use. Therefore, the population of an area is a likely driver of electricity demand.

To some extent, electricity is used by households rather than individuals. For example, a household with one resident may use a certain amount of electricity for cooling. Demand may increase if a second person occupied the household, but it would not necessarily double. In fact, if the house is already being cooled, adding a second occupant may have little or no impact on demand.

This goes to the question of the best measure of population for forecasting purposes. The two main options are population or the number of electricity customers.

The preferred variable is an empirical question. That is, it cannot be resolved without testing which measure performs better in an historical analysis of demand. Nor will the same measure necessarily perform better in all cases.

Historical population data are readily available from the ABS as are household formation data. Further, DNSPs have access to historical customer numbers data through internal records.

Regarding projections, the ABS and most Governments publishes a suite of population projections and household formation data that can be used to create customer numbers projections (assuming that the entire population has access to the electricity network). These projections are typically published at a granular level that a DNSP could build up to the necessary level of its network.

Alternatively, projections are available commercially from a range of economic forecasters.

Population is tied closely to economic outcomes so it is important to consider the timing of projections and therefore the economic outlook that they reflect. It is also worthwhile considering obtaining population and economic projections from the same source to ensure consistency.

## A.2 Household income and economic growth

Economic activity and income are key drivers of electricity demand, potentially operating through several mechanisms. Broadly, they both measure electricity customers’ capacity to consume. At the household level rising incomes are likely to be associated with purchases of more, new and larger appliances. At the commercial and industrial level they are associated with rising activity in a wide range of ways, most of which use electricity.

However, incomes and economic activity are closely related to one another and likely to be correlated. Therefore, it is unlikely to be appropriate to use both variables in a forecasting model. As with the best population variable (population or customer numbers) the best economic variable is an empirical question which should be answered by testing alternative model specifications.

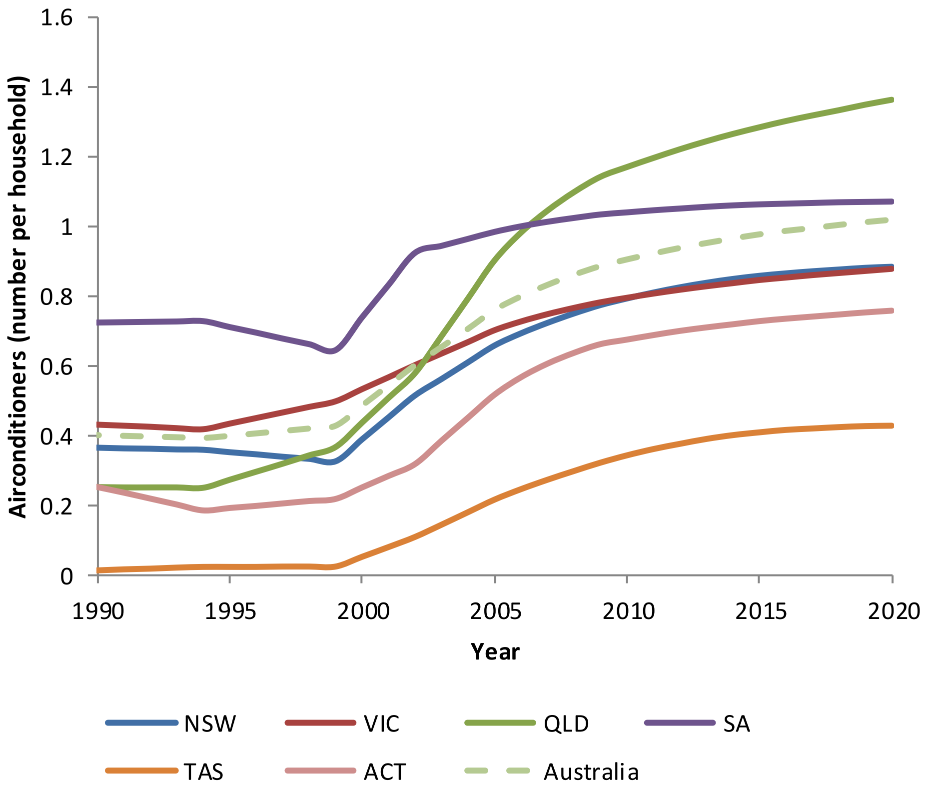
## A.3 Air conditioning system numbers and market penetration

It is well understood that air conditioner load is a large contributor to the variability in electricity demand. That is, a large part of the reason that demand is higher on some days than others is air conditioner use, particularly in private homes.

It follows that a demand forecasting methodology should take account of likely growth in the use of air conditioners at peak times over the forecast period. A key factor to consider in this projecting air conditioner take up is saturation. That is, leaving aside new household formation, there are a finite number of dwellings in a service area and thus an upper limit to the number of air conditioners that will be installed in that area. The limit is not as simple as ‘one per household’, but nor is it sufficient to allow a projection of air conditioner usage to continue along a linear trend into the future.

In 2008 the (then) Department for Environment, Water, Heritage and the Arts released a report which found that, between 1998 and 2008, the share of dwellings with space cooling installed more than doubled to about 65%. Looking forward, DEWHA projected that air conditioning penetration would decelerate over time, with different states peaking at different times as shown in Figure A1. Broadly, the figure shows that air conditioning penetration was stable in most states until the late 1990s when it began to increase. It was projected to increase into the future, though at a decreasing rate.

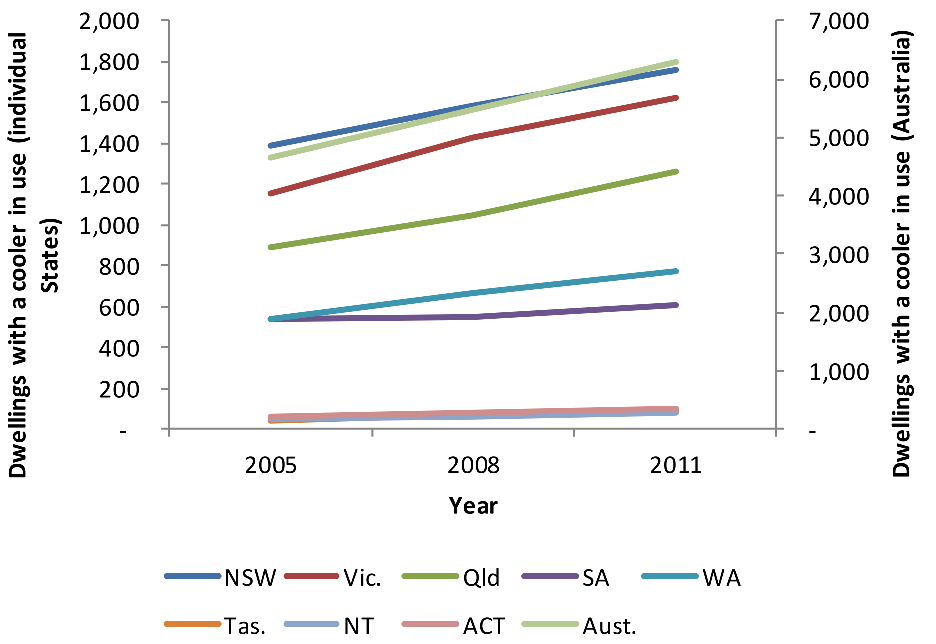
FIGURE A1 **AIR CONDITIONING PENETRATION BY STATE – DEWHA DATA (PROJECTION FROM 2008)**




*DATA SOURCE:* ENERGY USE IN THE AUSTRALIAN RESIDENTIAL SECTOR, 1986-2020, *DEPARTMENT OF THE ENVIRONMENT, WATER, HERITAGE AND THE ARTS, 2008.*

The ABS has also published data which support these conclusions, shown in Figure A2.

FIGURE A2 **DWELLINGS WITH A COOLER IN USE (,000) – 2005, 2008 AND 2011 - BY STATE – ABS DATA**




*DATA SOURCE:* ABS ENVIRONMENTAL ISSUES: ENERGY USE AND CONSERVATION, MAR 2011

The ABS data are more limited than the DEWHA data, with only three data points for each State.

Our discussions with DNSPs and our experience more generally suggests that air conditioner uptake in existing homes has already slowed in many parts of the country.

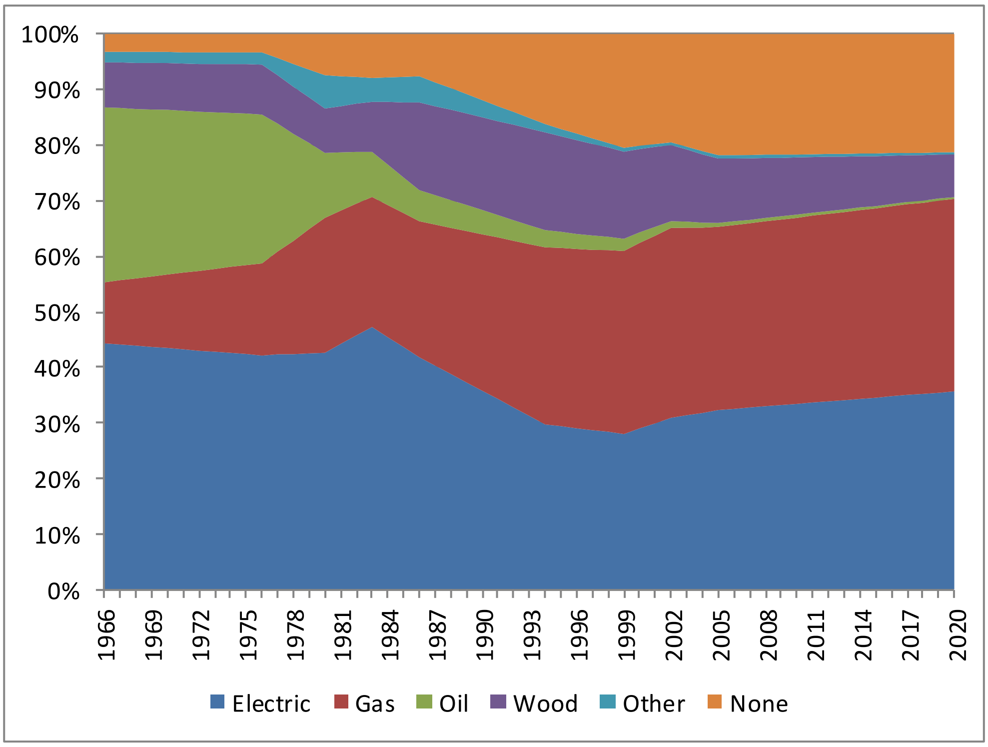
## A.4 Heating loads

Customers' use of space heating, and their choice of fuel, influences electricity demand similarly to the way that air conditioning does. In this case, they key issue is the substitution between electricity and gas as space heating fuels and the fact that there have been conflicting trends.

Where customers have access to gas, there has been a tendency to replace electric space heating systems with gas alternatives. At the same time, it appears that new homes are increasingly using reverse cycle (electric) systems for both heating and cooling.

DEWHA's survey results in respect of heater type are shown in Figure A3. They show a rise in the proportion of gas heating beginning in the 1980s and continuing until approximately 2000. This was mainly at the expense of electricity. Since then, though, the relative proportion of gas use for heating has remained constant, while the use of electric space heating has grown at the expense of other fuels, particularly oil.

FIGURE A3 **MAIN FORM OF SPACE HEATING**




*DATA SOURCE:* ENERGY USE IN THE AUSTRALIAN RESIDENTIAL SECTOR, 1986-2020, *DEPARTMENT OF THE ENVIRONMENT, WATER, HERITAGE AND THE ARTS, 2008.*

## A.5 Electricity prices

Typically, the price of a product is a key driver of demand for it. The relationship between price and demand is summarised in the price elasticity of demand, often referred to simply as elasticity. As discussed in section 9.1 of the report, the relationship between maximum demand and price is complex and different than the relationship between energy demand and price.

## A.6 The carbon price

No discussion of electricity price in the current environment is complete without a discussion of the carbon price.

From 1 July 2012, an explicit carbon price has been in place in Australia through the *Clean Energy Future* legislation. The stationary energy sector accounts for a substantial proportion of Australia's greenhouse gas emissions and much of this is attributable to electricity generation. Therefore, the carbon price is expected to have significant implications for electricity retail prices in the future.

The impact of an explicit carbon pricing mechanism on electricity prices is moderated by the fact that less than half of the price an end user pays for electricity is attributable to generation and therefore subject to carbon pricing.

Further, from 1 July 2015, the carbon price will be determined by the market. Until mid 2012, the Commonwealth Government's intention was that the market would include a 'floor price' preventing the carbon price from falling below $15 per tonne $CO_2$e in 2015-16, with the floor increasing slowly thereafter.

However, on 28 August 2012, the Commonwealth Government announced its intention to remove the price floor to facilitate linking between the Australian and European carbon pricing mechanisms.

The objective, and likely effect, of this change is that the carbon price in the two schemes will be similar to one another after 1 July 2015. Current projections for the European Carbon market indicate that future Australian Carbon prices will be much lower than their current levels.

There is also the possibility that the Commonwealth Government will change at the September 2013 election and that the carbon price may then be repealed.

For these reasons, we consider it reasonable to assume that the forward carbon price in Australia will be low.

The appropriate methodology for incorporating the carbon price into electricity demand forecasts is to ensure that the price projection used includes the impact of the carbon price. However, in practice, the appropriate adjustment may be small enough that it falls within the error margin of the price projection.

# Appendix B Weather normalisation

As discussed in section 2.2, the relationship between demand and weather is strong in many parts of Australia. This is usually due to heating and, in some places, cooling loads. Electricity demand forecasts must take this inherent randomness into account. The most common approach is to weather normalise the data upon which they are based.

While there are several ways of weather normalising data, the broad approach comprises the same two steps, namely:

1. estimating the relationship between demand and weather
2. adjusting the historical data to weather normal conditions using that relationship.

These two steps are discussed in turn below.

## B.1 Estimating the relationship between demand and weather

All approaches to weather normalisation begin by estimating a relationship between maximum demand and weather. To do so requires the forecaster to:

1. select a variable to measure 'weather'
2. choose the shape of the curve that describes the relationship
3. estimate the weather sensitivity
4. calculate the weather normalised historical data.

### B.1.1 Measuring weather

The weather is often represented by:

- average daily temperature
- weighted average temperature with a large weighting in favour of the daily maximum temperature relative to the daily minimum
- daily maximum and daily overnight minimum as separate variables within aregression
- weather indices including temperature and additional variables like humidity, rainfall and average wind speed.

Ideally, the choice of weather variable would be empirical rather than theoretical. That is, the variable that provides the best fit in a model would be chosen. Lags of the specified variables may also be included for several days. Again, this is an empirical question.

### B.1.2 Estimate the weather sensitivity

The weather sensitivity, or the relationship between weather and demand, is estimated by fitting a regression between daily maximum demand and the chosen weather measure. It may be estimated for summer only (in networks where annual maximum demand occurs in summer) or separately for summer and winter.

Weather sensitivity is estimated for each CP separately.

The shape of the curve to fit is discussed in section B.2 below.

In some cases the fit of the regression will be poor, indicating that the network element in question is not weather sensitive. This is to be expected in the case of network elements that mainly supply industrial customers (whose demand is relatively 'flat'). It is also to be expected in areas where water pumping load is the main driver of variation in demand. In these cases the regression results should be discarded and weather normalisation would not be used.

The estimated regressions produce a set of coefficients which measure the temperature sensitivity of maximum demand. These coefficients then form the basis of the weather normalisation.

### B.1.3 Adjustment of maximum demand to weather corrected level

Once the relationship between maximum demand and weather has been estimated (for each season), the next step is to adjust observed demand to standard weather conditions. There are four ways this can be done:

1. Taking weather normalised demand 'off the point'

   In this approach, the relationship between demand and temperature is expressed in MW per degree above standard. The forecaster identifies the actual temperature and demand on the peak day and adjusts demand accordingly.

2. Taking weather normalised demand 'off the line'

   In this approach, weather normalised demand is calculated using the weather sensitivity regression discussed above at the value of the weather variable that corresponds to standard weather conditions. This would typically be calculated from a suitably long time series of weather data (more than 30 years). For example, the 50 POE temperature would be the median value of 30 years of temperature data.

3. Taking weather normalised demand 'above the line'

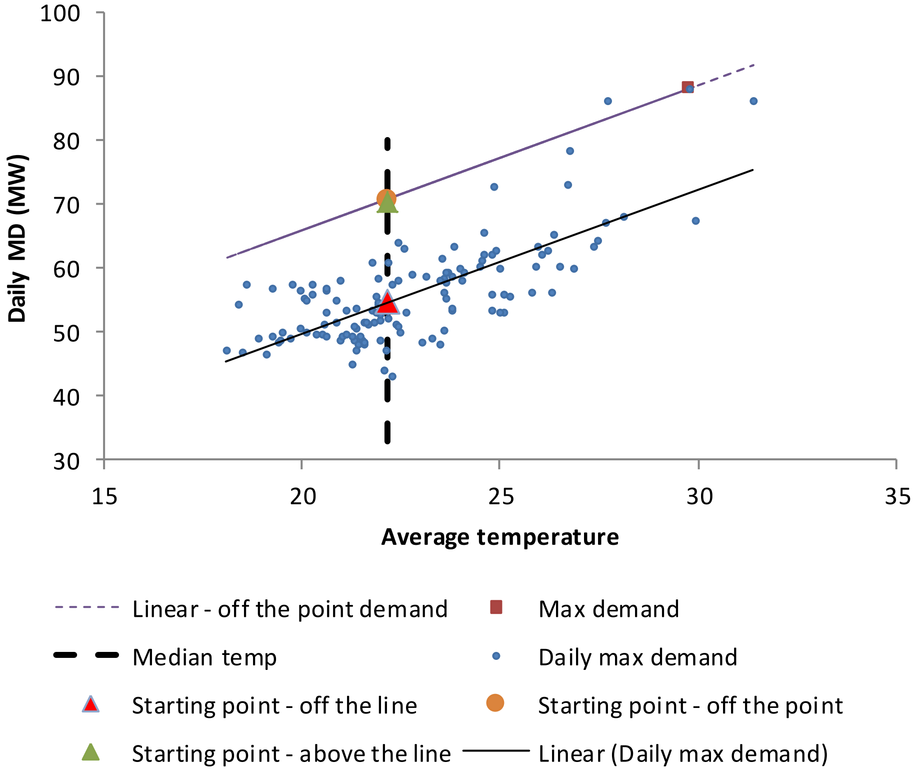
   This approach is similar to approach 2 above, but weather normalised demand is defined as a point above the regression line, often 2 standard deviations above it.

4. Applying a simulation approach

   In this approach the complete long run time series of weather data are fitted to the estimated model to produce an estimated maximum demand for each day over a long weather history, often dating back 50 years. The standard error of the regression is used to simulate results in a large number of trials to develop a long run distribution of maximum demands from which the median, $10^{th}$ and $90^{th}$ percentiles can be extracted. This approach is described in more detail in section 5.3 of the report.

The first three approaches are illustrated in Figure B1.

FIGURE B1 **WEATHER NORMALISATION**




Most DNSPs have traditionally used one or another of these approaches. These approaches are common in that each is based on the notion that 'normal' demand is calculated by using 'normal' weather conditions. They rely on a one-to-one relationship between demand and weather conditions.

Some DNSPs are beginning to use the simulation based approach described in the report (section 5.3). This allows for more complex relationships between maximum demand and weather. With this approach, the notion of standard weather conditions is less important. Rather, the underlying concept of standard demand is estimated more directly. That is, this approach estimates the demand that would be observed on a 50 POE (or other) basis given the demands that have been observed, and weather normalised, in history.

The traditional approaches are inherently less robust than the simulation approach because they entail weather normalising only a single day each season, namely the day on which maximum demand was observed.

The main advantage of the simulation approach is that it tends to produce weather normalised maximum demands that are not biased, as would be the case if only a single peak day was adjusted to some weather corrected value. Also, there is now no longer a requirement for there to be a 1 for 1 relationship between the level of maximum demand and some measure of weather, giving the forecaster additional flexibility to specify better models relating maximum demand to weather.

Simulation methods on the other hand construct an entire distribution of maximum demands over a large number of years from which the 50% and 10% POE maximum demand are obtained.

## B.2 Shape of fitted curve

The relationship between daily maximum demand and weather is often said to be S-shaped. This reflects the understanding that the weather sensitivity of demand is driven by cooling load (in summer) and heating load (in winter). It also reflects the fact that heaters and coolers are typically not used when conditions are 'mild'.

Therefore, for summer, daily maximum demand becomes unresponsive to changes in temperature at cooler (summer) temperatures, often around an average temperature of 22 to 23 degrees. There is also a slight flattening of the curve at extremely high temperatures where maximum demand reaches a point of saturation and cannot rise any further in response to higher temperatures. At these levels, all or most air conditioning systems within a network are turned on.

The opposite relationship applies for winter, with lower temperatures associated with higher levels of maximum demand. As temperatures increase the maximum demand response to changes in temperature flattens also.

In winter a saturation point could theoretically be expected at very low temperatures when the majority of the available heaters are in use, though this is rarely observed in Australia.

The S-shaped relationship between weather and demand can be used in weather normalisation. Some DNSPs specify the relationship as a polynomial line of best fit to the data and estimate the parameters of that function using linear regression.

Other DNSPs take the view that there is usually insufficient data to estimate the shape of the S-curve with sufficient accuracy, particularly the top of the curve. To illustrate, consider Figure B2, which shows daily maximum demand observed at a CP in the 2012/13 summer along with the long term 50 and 10 POE (daily average) temperatures for the area.

FIGURE B2 **WEATHER SENSITIVITY**

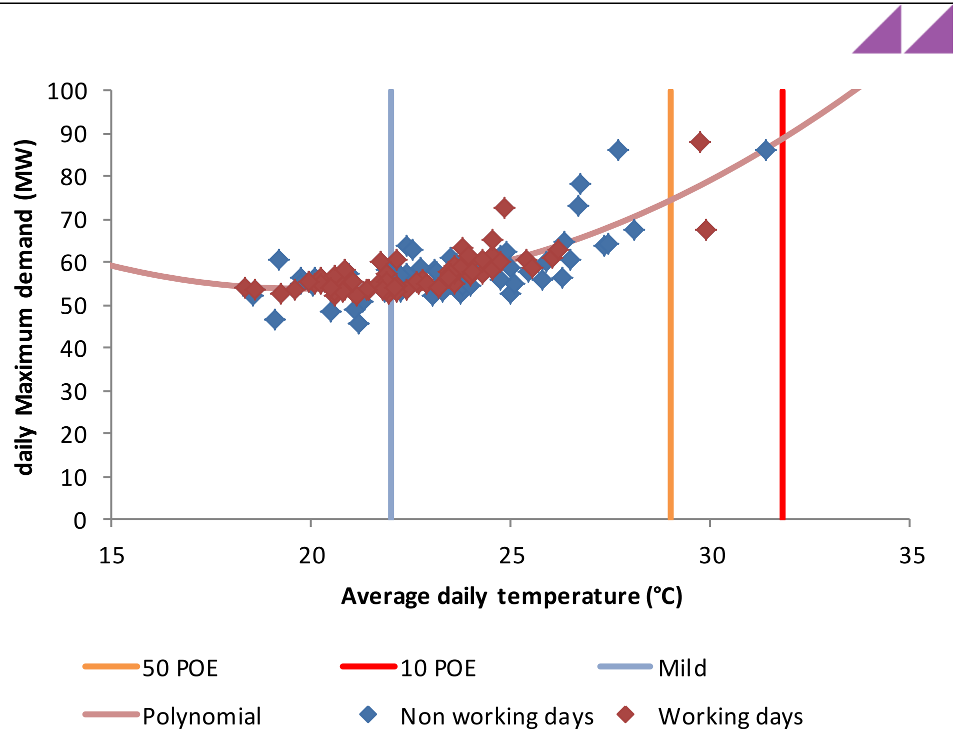
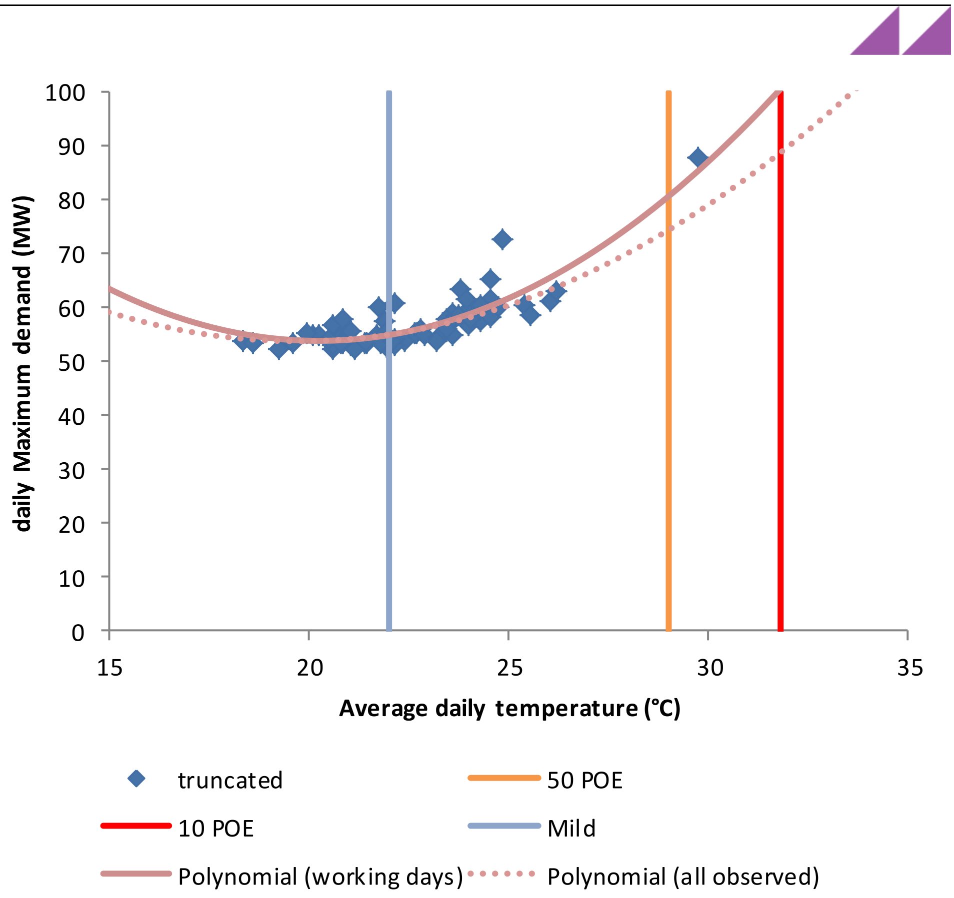



Figure B2 shows that, this year, the temperature rose above the 50 POE level three times, reaching almost, but not quite, to the 10 POE level. Therefore, by comparison with the long term average, this was a very hot year. However, there are only three observations above the 50 POE and the hottest of these three days was a weekend.

The polynomial line of best fit passes through the demand and temperature observed on the hot weekend day. This is unsurprising given the arrangement of the data. However, given that it was a Saturday, demand was likely to be less on that day than it would have been if the same temperature had occurred on a weekday. Therefore, the polynomial curve fitted here is not likely to capture the true relationship between temperature and maximum demand.

This can be addressed by truncating the dataset to omit non-working days as shown in Figure B3.

FIGURE B3 **WEATHER SENSITIVITY – WORKING DAYS**




The figure shows that removing the non-working days from the dataset caused the line of best fit to shift upwards at higher temperatures, which is the expected result (the polynomial curve through the full data set is shown as a dashed line for comparison).

However, this figure also shows that the polynomial line is dependent on demand at mild temperatures. The line of best fit now suggests that, as daily average temperature falls below approximately 20°C, demand will increase. This is not consistent with reasonable expectations. That is, this is still quite a warm temperature, so it is unlikely that any significant amount of space heating load would be introduced at this temperature. Rather, this is the result of the inherently symmetrical nature of a second order polynomial.[67]

[67] This could be avoided by using other functions that produce a natural S or sigmoid shape such as a Gompertz function.

This problem can be removed by truncating the dataset to remove observations at temperatures when demand is expected not to be temperature sensitive. This is shown in Figure B4.

FIGURE B4 **WEATHER SENSITIVITY – TRUNCATED DATASET**

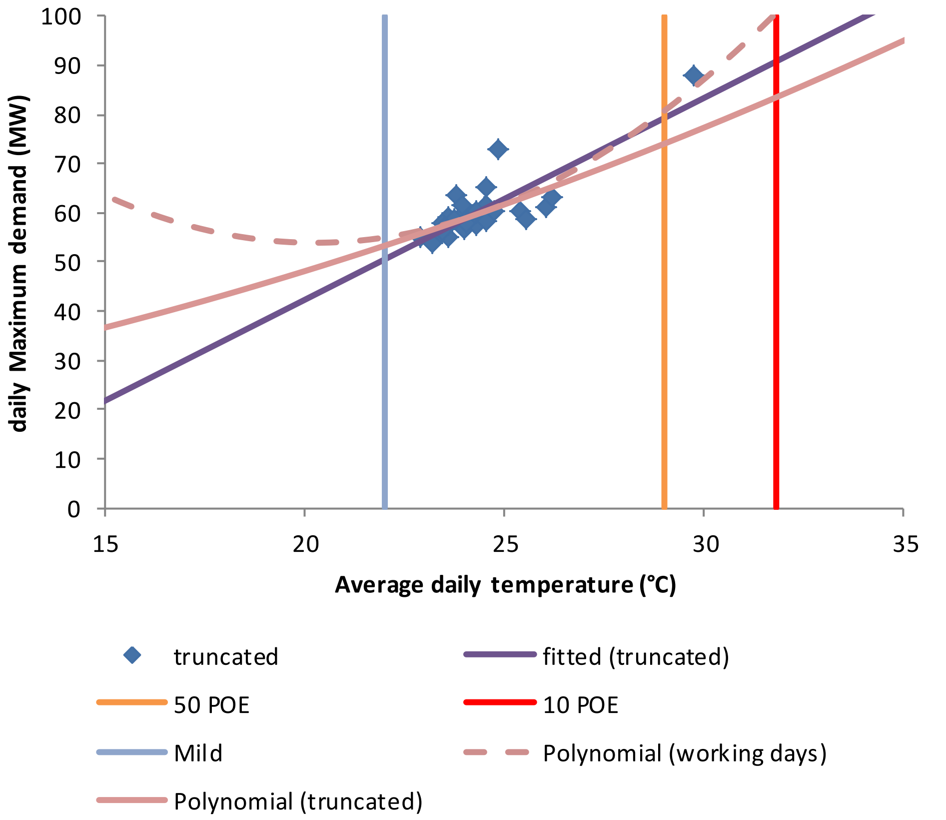



Figure B4 shows that, if the dataset is truncated so that only hot days are taken into account, the polynomial of best fit between temperature and demand is 'flatter' than in the 'all working days' data set (in this figure the dashed line is the polynomial curve from the 'all working days' data shown in Figure B3). This confirms that the relationship between demand and temperature under cooler conditions was influencing the estimated relationship.

However, there are now very few observations on which to base an analysis of the relationship between temperature and demand at this CP. For this reason, it is preferable to estimate a linear relationship.[68]

Figure B4 also shows the result of fitting a linear regression to the truncated data set. The result is close to the polynomial. The figure also shows that there are only a few observations to use in specifying this relationship, but this is unavoidable.

[68] The solution cannot be to add other days from the same year because, as discussed above, they have already been eliminated for valid reasons.

# Appendix C Forecasting approaches

## C.1 Top-down v bottom-up forecasting

Most forecasting methodologies can be classified as top-down or bottom-up in nature.

Top-down approaches apply at a high or macro level and typically incorporate economic drivers.

Bottom-up forecasts apply at a much lower level such as a single zone substation or feeder.

The data used in top-down approaches often display more stable and systematic characteristics. This makes them more amenable to the application of statistical or econometric techniques and approaches. Further, top-down approaches can draw on a rich array of macroeconomic data produced by national statistical agencies such as the ABS. This allows them to incorporate causal factors that cannot be incorporated into bottom-up forecasts.

However, electricity networks are planned and built at a very local level. A network could not be planned or managed with a system level forecast alone because planners need to know the likely demand on key elements of the network. The main disadvantage of top-down approaches is that they do not produce disaggregated forecasts.

The disaggregated forecasts needed for network planning and operation are produced using bottom-up approaches. Detailed and reliable *disaggregated* data tend not to be available. Therefore, the econometric methods in bottom-up methodologies tend to be less sophisticated, often involving the simple linear (or non-linear) extrapolation of trends. A further complication is that the highly disaggregated data used in bottom-up forecasts is often affected by errors or irregularities that must be corrected.

The contrasting strengths of top-down and bottom-up approaches can be obtained by combining them.

Armstrong notes that forecasting accuracy can be improved by combining different forecasts, particularly when the different forecasts are obtained from substantially different methods and draw from different sources of information.[69] This is generally true of bottom-up versus top-down methodologies in maximum demand forecasting as they employ very different methods and also use quite different information.

Combining different sets of forecasts is also useful when it is unclear which of the competing methodologies would be the most accurate and where the implications of large errors are severe.

## C.2 Forecasting approaches

This section provides an overview of the following four ‘families’ of approach to forecasting electricity demand:

1. causal models
2. trend models
3. end use models

[69] Armstrong, J. S. “Principles of Forecasting: A Handbook for Researchers and Practitioners”, (2001)

4. judgemental approaches.

### C.2.1 Econometric/Causal models

Causal models are based on the prior knowledge of theoretical relationships that exist between a dependent variable and those variables (or drivers) that affect it. For example, a causal model of electricity demand may be based on the knowledge that, all else being equal, demand will increase as the number of customers in a service area increases.

By incorporating the major factors affecting maximum demand, causal models improve the forecaster's ability to explain changes in the structure of demand.

Causal models are calibrated (estimated) using a range of econometric techniques, commonly multiple linear regression and ordinary least squares estimation.[70] These approaches attempt to quantify the relationship between demand and those factors that influence it.

By quantifying these relationships explicitly causal models allow the forecaster to incorporate their view (or the views of other experts) on the future course of these drivers into the demand forecasts. This is not possible with other techniques and is the main advantage of this approach.

According to Armstrong (2001), causal models are most useful when:

— strong relationships are expected between the dependent variable (electricity demand in this case) and the independent variable(s) (the drivers of electricity demand)

— the direction of the causal relationships are well known

— large changes are expected in the causal variables.

As discussed in Appendix A, the theoretical relationships between electricity demand and a range of drivers is well understood, though it may be complex. Similarly, while there are often contradictory pressures on demand due to competing relationships, the direction of each causal relationship is typically well known.

It is also well known that many of the drivers of electricity demand have changed substantially in recent years and many are expected to change in future.

The main disadvantages of this approach lie in the requirement for detailed information (both historical data and forecasts of the drivers) and that the drivers, while well understood theoretically, may be difficult to measure. The absence of a suitably long period of historical data may make this approach impractical. Further, applying causal models is also more complex than some of the alternatives.

### C.2.2 Trend based approaches

An alternative approach to applying a causal model is to estimate future demand from historical demand. The notion underpinning this approach is that demand in future will reflect the experience in the past. For example, growth may continue, on average, at the same rate that has been observed historically. The advantage of this approach is that it does not require large amounts of data other than historical maximum demand data. A trend based approach is appropriate during periods of stability when the underlying drivers of maximum demand are expected to be similar during the projection period to the historical period for which data are available.

---

[70] Causal models could also be described as econometric models, though the latter term may imply that the causal relationship must be economic. For example, it may be argued that a model of electricity demand based on temperature alone is causal, but not econometric. In this report we use the terms interchangeably to avoid confusion.

However, a trend based approach encounters significant problems when the underlying drivers are expected to change.

Because of its simplicity, the extrapolation of past trends is an approach that is commonly applied at the feeder or zone substation level. At these levels, a trend based approach is likely to be the best available as there are unlikely to be meaningful measures of economic or demographic data at the highly disaggregated spatial level.

### C.2.3 End use models

An alternative bottom-up approach involves forecasting maximum demand based on assumptions of energy usage patterns of individual end users. These 'end use' models are usually informed by detailed surveys of those users, in particular their appliance usage and demographic characteristics. They work by building up profiles of the individual energy users and then applying these to projections of the total number of connected customers or appliances in use.

The main advantage of this approach is that it offers the most detailed understanding of the way energy is used. This is particularly useful when assessing the impact of demand conservation measures, take up of new technologies and appliances, or energy efficiency through technological progress.

However, the heavy survey requirement makes them a more complex approach to obtaining forecasts of maximum demand. It also makes this the most data intensive and costly of the approaches discussed here.

The end use model approach is unlikely to be a cost-effective option to forecast maximum demand at the CP level, as it requires the forecasts be built up from the level of an individual household. This has implications for the sample size required in the survey. The more disaggregated the forecasts to be prepared, the larger the sample that is required. For example, if reasonably accurate results are required at the system level, then a sample size of several thousand is likely to produce a relative standard error within 5%. However, if the results are required to 10 or 20 regions across a single jurisdiction, then the required sample size will expand significantly (by many multiples) to achieve a similar level of relative standard error.[71]

The application of end use models at the regional level will therefore be highly prone to error.

### C.2.4 Judgmental approaches

Maximum demand forecasting at the spatial level often involves a judgemental component. Judgement based forecasts aim to exploit the detailed local area knowledge of planning engineers and other subject matter experts.

Forecasting electricity demand will inherently include subjective elements, exposing it to the judgement of individual forecasters. This is not inappropriate and 'judgement' should not be considered a 'dirty word' in this context. Armstrong (2001) considers that it is good procedure to use judgment when:

- experts are unbiased
- large changes are unlikely

[71] Another complication is that the sample must also provide reasonable coverage of different groups in society. The most cost effective way to conduct surveys is via the internet using pre-registered panels. However, while very useful for some applications, these panels are typically underrepresented from certain key groups such as low income and elderly consumers. This can have serious implications for the accuracy of the sample.

- relationships are well understood by experts
- experts possess privileged information
- experts receive accurate and well summarised feedback about their forecasts.

The judgment of experts is often called upon to assess the calculated trend growth rates that are obtained from any trend analysis that might be applied at a zone substation. In this case, the use of objective judgement is useful in validating the growth rates that are obtained from the automatic fitting of trends, especially since the historical trends can be significantly affected by permanent transfers and temporary switching which can bias the fitted trend line significantly. Also, the local asset managers are well placed to understand what is happening in and around an area serviced by a particular zone substation. They often possess knowledge of recent developments that may not be reflected in the historical maximum demand numbers, and are able to amend the growth rates to reflect this additional local area knowledge.

There may be, however, a concern about the impartiality of the network managers when forecasts are prepared for regulatory purposes because the planners are employed by the company whose revenue is being determined. While this may be more an issue of perception than reality, it is a strong argument for detailed and transparent recording of the judgements that were made and the reasons for them.

# Appendix D Accounting for diversity using coincidence factors

FIGURE D1 **ILLUSTRATIVE ELECTRICITY NETWORK**

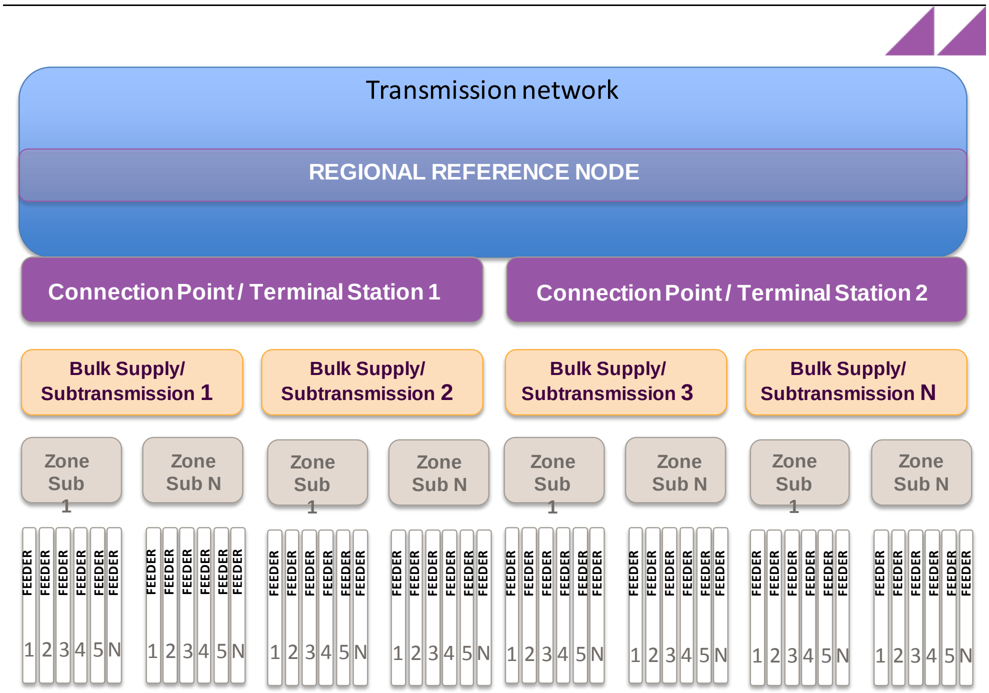



Figure D1 shows the structure of the electricity network upon which the forecasting methodology described here is based. In this appendix, the issue of diversity between different levels of the network is discussed and the method of calculating coincidence factors is described.

The network illustrated in Figure D1 consists of 62 elements, namely:

1. Two CPs
2. N subtransmission stations, each associated with a single CP
3. N ZSS, each associated with a single subtransmission station
4. N feeders, each associated with a single ZSS.

To illustrate the concept of diversity, it is necessary to think of a long time such as a season or a year that comprises a number of shorter periods, often five or 30 minutes.

In any given period, each feeder in the network supplies a certain quantity of electricity. Therefore, there is a certain demand at each feeder every period. [72,73]

Given the physical structure of the network, in any given period, demand at a higher level is the sum of demand at a lower level. Therefore, in period '*t*', the *same* demand is experienced simultaneously at:

- CP1

[72] Demand may not be measured at every element, particularly feeders, but it exists nonetheless.

[73] Demand is conceptually an instantaneous concept, though it is measured over 'periods'. Different measurement technologies provide data over different periods, with 5 minutes and 30 minutes both common. The length of the period does not matter for the current example.

- subtransmission stations 1 and 2
- the first four ZSS
- the first 24 feeders.

In any given season (or year), every element in the network will experience one period when demand is higher than in any other period. That is, each element experiences maximum demand in one period.[74] However, there is no particular reason to believe that maximum demand will happen at different elements in the same period.

The fact that different elements can experience maximum demand at different times has implications for the maximum demand at 'higher' levels of the network. This is illustrated in Figure D2, which takes a very simple network consisting of four feeders (F1, F2, F3 and F4), two ZSS (Z1 and Z2) and one CP. For simplicity losses are disregarded in this example.

FIGURE D2 **SIMPLIFIED ELECTRICITY NETWORK**


SOURCE: ACIL ALLEN CONSULTING

74 Theoretically the same demand could be observed in more than one period. This is unlikely, but does not matter. What matters is the level of maximum demand, not the number of times it is observed.

In the first pane, demand is:

- 14 MW at F1
- 8 MW at F2
- 9 MW at F3
- 6 MW at F4

Therefore, demand is 22 MW at Z1, 15 MW at Z2 and 37 MW at CP

In the second pane, demand is

- 10 MW at F1
- 11 MW at Z2
- 9 MW at F3
- 9 MW at F4

Therefore, demand is 21 MW at Z1, 18 MW at Z2 and 39 MW at CP.

The coincidence factor is the ratio of maximum demand at an element to demand at that element when demand at a higher element was at maximum. Therefore, a coincidence factor is computed between two elements. Each element has a coincidence factor between itself and each higher network element to which it is connected.

Coincidence factors can be calculated based on the two periods shown in Figure D2. Consider F1, which is connected to Z1 and CP. Maximum demand at F1 for the year was experienced in period 1. It was 14MW. However, maximum demand for Z1 was experienced in period 2, when demand at F1 was 10 MW. Therefore, the coincidence factor for F1 with Z1 is 10/14. CP also peaked in period 2, so the coincidence factor for F1 with CP is also 10/14. If CP had peaked in a third period, the coincidence factor between F1 and CP would be the ratio of F1's demand in period 3 and 14 MW.

In this example, the coincidence factor of both Z1 and Z2 is 1, because both peak simultaneously with CP. In a more complex example this would not necessarily be the case.

# Appendix E Energy Forecasts

AEMO's original terms of reference also sought advice relating to energy forecasts at the CP level.

As discussed in the supplement to this report, DNSPs in the NEM typically do not produce energy forecasts at this level or any other level of spatial disaggregation. Most indicated these forecasts would be of no business benefit to them. That is, DNSPs prepare forecasts of energy sales for their whole system, but do not disaggregate these forecasts in any spatial way.

The reasons for this relates to the reason that DNSPs (and TNSPs) are regulated and the role that forecasts of maximum demand and energy play in that regulation.

DNSPs are natural monopoly businesses. This means that, unlike the vast majority of Australian businesses, competition cannot be relied upon to give DNSPs an incentive to provide consumers with a service that represents an appropriate combination of price and quality. Therefore, DNSPs are subject to economic regulation.

Under the National Electricity Rules, DNSPs are subject to a *control mechanism* which may consist of one or more of the following:[75]

1. a schedule of fixed prices
2. caps on the prices of individual services
3. caps on the revenue to be derived from a particular combination of services
4. tariff basket price control
5. revenue yield control
6. a combination of any of the above.

The specifics of each control mechanism vary but, broadly, each requires that two things are determined:

1. the total revenue that the DNSP should be entitled to earn, which is based largely on the capital and operating expenditure it will need to make
2. the price that the DNSP should be entitled to charge its customers to recover that revenue.

The forecasts a DNSP prepares relate to these two parameters directly.

The maximum demand that will be experienced on a DNSP's network relates directly to the total cost a DNSP will need to incur. The higher the demand is expected to be, the more infrastructure will be needed to supply it and, therefore, the more the DNSP will need to spend and the more revenue it will be entitled to earn.

DNSPs require maximum demand forecasts to be prepared on a spatial basis because they need to know not only how large demand will be, but which of their assets will be required to deliver it and, therefore, *where* demand will occur. If DNSPs only had system level demand forecasts they would know that their system may need to be upgraded but would not know which assets to upgrade.

On the other hand, DNSPs require energy forecasts to be disaggregated by tariff class. In a simple model, where there is only one tariff class, a DNSP's revenue requirement would be divided by the amount of energy it will deliver to reveal the price it would be permitted to charge. In practice, this is more complex because tariffs typically consist of fixed charges and usage charges, and usage charges may

[75] *National Electricity Rules* clause 6.2.5

have more than one block. Further, different tariffs apply to different classes of customer such as industrial, commercial, residential etc.

However, while DNSPs typically have more than one tariff class, they invariably *do not* distinguish tariff class on geographic grounds. Therefore, a residential customer in one part of a DNSP's network is invariably charged the same (distribution) tariff as a residential customer in any other part of that DNSP's network.

For this reason it simply does not matter to a DNSP *where* they deliver their energy. Accordingly, DNSPs apply significant effort to forecasting energy sales by tariff class, but not by location or network element.