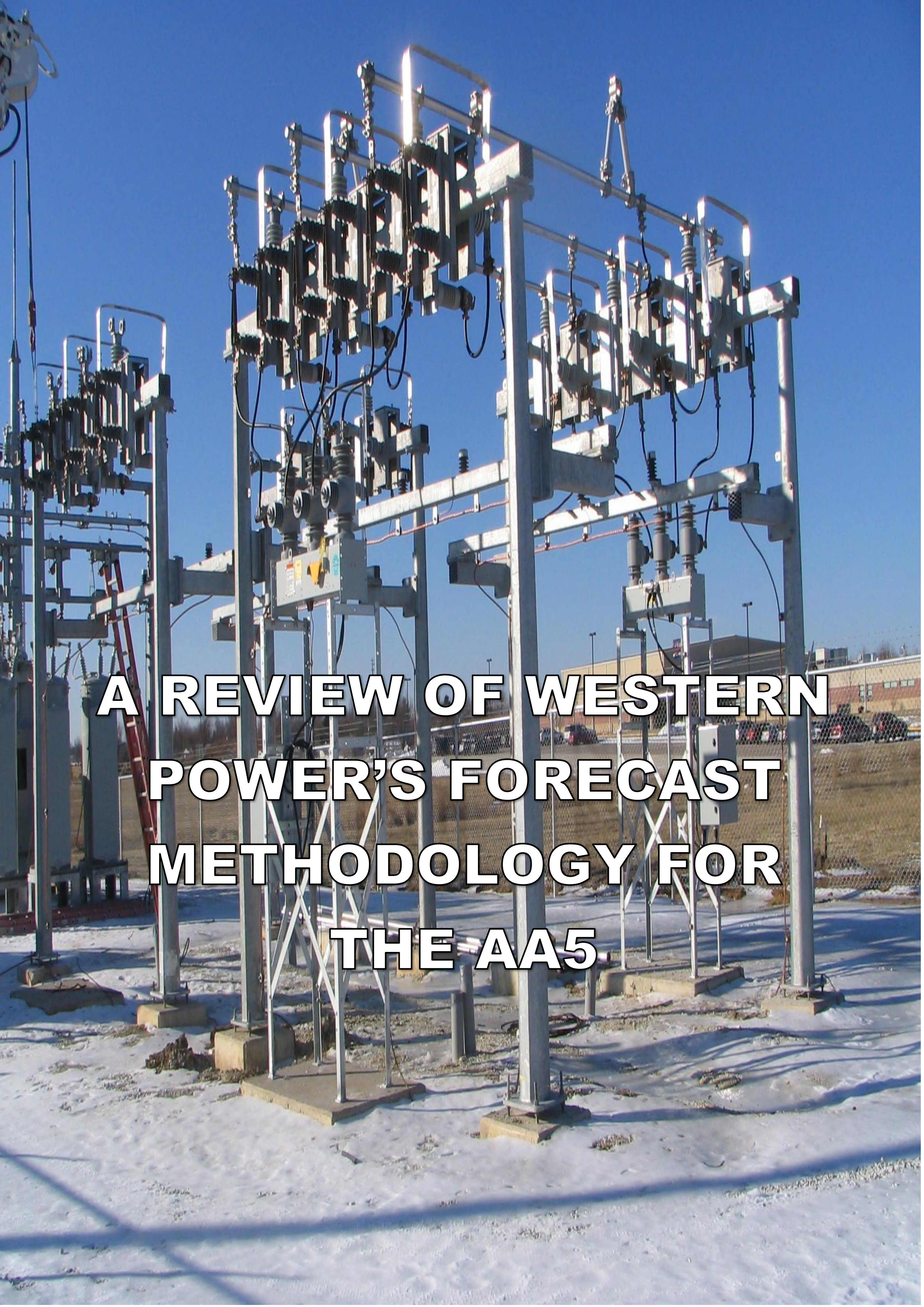
Attachment 7.7

Report on Western Power's Forecasting Methodology

Access Arrangement Information

1 February 2022

EDM 57144170



A REVIEW OF WESTERN POWER'S FORECAST METHODOLOGY FOR THE AA5

A report for
WESTERN POWER

June 2021

While the National Institute endeavours to provide reliable forecasts and believes the material is accurate it will not be liable for any claim by any party acting on such information.

Prepared by the
National Institute of Economic and Industry Research
ABN: 72 006 234 626
Lower Ground, Unit 1A, 663 Victoria Street, Abbotsford, Victoria, 3067
Telephone: (03) 9488 8444; Email: admin@nieir.com.au

TON905 – Western Power-Methodology (report 1)/Western Power/2021

Contents

		Page no.
Executive summary		**i**
1.	**Introduction**	**1**
1.1	Project background	1
1.2	Project scope and deliverables	1
2.	**Energy and customer number forecasting methodologies**	**2**
2.1	Introduction	2
2.2	Methodological approach	2
3.	**Maximum and minimum demand forecasting methodologies**	**7**
3.1	Overview of forecasting approach	7
3.2	Data validation and preparation	7
3.3	Methodological approach	8
3.4	Methodology and forecast validation	10
3.5	Comparison to other methodologies	11
3.6	Discussion and recommendations	12
4.	**Summary of recommendations**	**13**
Glossary		**14**
Appendix A:	**Documentation reviewed by NIEIR**	**17**
Appendix B:	**NIEIR's relevant experience**	**18**

List of tables

		Page no.
2.1	Reference tariffs	2
2.2	Western Power – Energy and customers by class and major tariff groups	3
2.3	Description of the independent variables used by Western Power	4
2.4	PV capacity forecasts and independent variables	5
3.1	Candidate point process models used to fit the substation data	9
4.1	Summary of recommendations/suggestions for improvement in Western Power forecasts	13

List of figures

		Page no.
2.1	A schematic of the SAS Forecast Studio process	3
2.2	The forecast hierarchy for energy and customers	4
3.1	Arkana substation – Drivers of maximum demand	9
3.2	Muchea substation – Summer maximum demand forecast (2018)	10

Executive summary

Western Power is currently preparing for its AA5 submission to the Economic Regulation Authority (ERA) for the control period 1 July 2022 to 30 June 2027. Western Power prepares forecasts of customer connections, energy and peak demands for the AA5 period.

In June 2021, NIEIR was commissioned by Western Power to review their forecasting methodologies.

Western Power produces forecasts of customer connections and energy at the substation levels. Maximum and minimum demands are also generated at the substation level. As part of this process, Western Power uses independent variables to drive the forecasts, as well as internally generated PV customer and capacity forecasts.

NIEIR reviewed a number of documents relating to the forecast methodologies employed. In addition, NIEIR and Western Power held a number of video conferencing sessions regarding the forecast process at Western Power.

Western Power forecast methodologies

Overall, NIEIR's review of the forecast methodology was positive. Notwithstanding some potential methodological improvements, Western Power's forecasting models appear reasonable, robust and fit for purpose.

Some other key points in relation to Western Power's forecasting methodology are as follows.

- The use of SAS Forecast Studio for energy and customers, implies the forecasting process is transparent and repeatable. The models produced should, therefore, be accurate and unbiased.
- Western Power monitors the accuracy of its forecasts concentrating on its latest set of forecasts. Evidence of this process is in Western Power's energy forecasting worksheets in Excel files provided to NIEIR.
- Western Power has documented its process for developing the energy and peak demand forecasts. The process steps are:
  - NTWK.1.2.11 Validate forecasting data;
  - NTWK.1.2.1.2 Prepare block load forecasts;
  - NTWK.1.2.1.3 Develop underlying load growth energy and customer numbers forecasts;
  - NTWK.1.2.1.4 Develop underlying load growth forecasts;
  - NTWK.1.2.1.5 Develop annual energy and customer forecasts;
  - NTWK.1.2.1.6 Develop demand forecast;
  - NTWK.1.2.1.7 Develop adjusted demand forecast;
  - NTWK.1.2.1.8 Prepare the Network Capacity Mapping Tool data; and
  - NTWK.1.2.1.9 Compile the peak demand trend report.
- Western Power also has a documented process of quality controlling its input data which is industry best practice.
- Peak demand forecasts are generated using Extreme Value Theory (EVT). This approach is based on modelling extreme deviations from the median of probability distributions.
- Quantile regression is also used by Western Power in modelling peak demands in some instances.
- The customer and energy forecasts are generated using SAS, a publically available software package. The peak demand forecasts are generated by scripts written in R. R is an open source and free software environment for statistical computing and graphics. It runs on a variety of UNIX platforms.

A checklist of recommendations regarding the forecasting methodologies is presented in Section 6 of this report. These recommendations cover both the customer and energy forecasts and the maximum and minimum forecasts for the Western Power distribution area. Not all recommendations will be adopted by Western Power as there are trade-offs between different objectives of the forecast process. These include detail and complexity, transparency and repeatability. Battery and EV models need to be developed as a matter of priority.

1. Introduction

Western Power commissioned the National Institute of Economic and Industry Research (NIEIR) to review its forecasting methodologies for energy, customers and maximum demands.

In 2016, NIEIR was commissioned to review Western Power's forecasting methodology for energy, customers and peak demands. New forecasting methods and processes were developed in 2014-2016. In 2010, NIEIR was commissioned to review the Western Power's zone substation forecasting model.[1] Post-2016, new forecasting methodologies have been developed for maximum demand forecasts. These are based on extreme value theory.

Western Power requires a suitably experienced consultant to review the demand forecasts being prepared for the next Access Arrangement submission ("the AA5 submission").

1.1 Project background

Western Power is a publicly owned Government Trading Enterprise (GTE) governed by an independent Board that reports through the Minister for Energy to Parliament.

It is responsible for building, maintaining, and operating the electricity network in the South West corner of Western Australia. The network forms the vast majority of the South West Interconnected Network, which together with the electricity generators comprises the South West Interconnected System.

Western Power is currently preparing its AA5 for submission to the ERA which is due on 1 September 2022. The AA5 regulatory control period will span 1 July 2022 to 30 June 2027.

1.2 Project scope and deliverables

As part of its preparations, Western Power prepares forecasts of demand of its core services, namely connection, capacity, and volumetric services. The forecasts are used, among other things, to determine the efficient level of capital expenditure.

Ahead of the actual submission, Western Power is seeking assurance that the prepared forecasts are an appropriate basis for developing its capital expenditure plans.

Western Power will make available the following supporting materials for the forecast review:

- Energy and Customer Numbers Forecast 2020.xlsx;
- Energy & Customer Numbers Report 2020;
- CUSTED2020 Forecast Method Report;
- 2020 Day minimum demand forecasts by zone substation report;
- 2020 Maximum demand forecasts (SUMMER) by zone substation report;
- 2020 Maximum demand forecasts (WINTER) by zone substation; and
- Other materials, such as the forecasting input and output data sets, computer code can also be provided on request.

This report is structured as follows.

- Sections 2 and 3 describe the forecasting methodologies for energy and customer numbers, and minimum and maximum demands for Western Power.
- Section 4 provides a checklist of recommendations for possible future works for both energy customers and maximum demands.

[1] "*A review of issues raised regarding the Western Power zone sub-station forecasting model*", a report for Western Power, by NIEIR, January 2010.

In 2005, NIEIR also reviewed Western Power's forecasting methodologies in "*Verification of Western Power Corporation Forecasts of Demand and Energy for the Access Arrangement for the SWIN*", a report for Western Power Corporation, by NIEIR, March 2005.

2. Energy and customer number forecasting methodologies

2.1 Introduction

Western Power prepared forecasts of energy and customer numbers from 2020-21 to 2024-25 in October 2020. They were subsequently extended out to 2026-27 in April 2021.

This section outlines the methodological approach adopted by Western Power in developing these forecasts.

2.2 Methodological approach

Western Power's energy and customer number forecasts are developed on a network tariff basis. The reference tariffs modelled by Western Power are shown in Table 2.1. The total energy for 2019-20 and customers at 30 June 2020 are also provided in Table 2.1. The tariffs include small single and multi-rate for residential and business, high voltage demand tariffs and bi-directional tariffs for solar customers.

Western Power's network tariff models are based on monthly data extending back to 2007-08 (up to 156 observations). In order to smooth out the effects of tariff switching over time, Western Power also prepares summary forecasts on a class basis and major tariff group. These are shown in Table 2.2.

Table 2.1 Reference tariffs

Tariff group	Description	Energy – 2019-20 (GWh)	Customers 30 June 2020 (number)	Average use (MWh)
RT1	Anytime Energy Residential	3502	725972	4.8
RT2	Anytime Energy Business	611	68796	8.9
RT3	Time of Use Energy Residential	32	5155	6.3
RT4	Time of Use Energy Business	318	4015	79.2
RT5	High Voltage Metered Demand	687	308	2230.8
RT6	Low Voltage Metered Demand	1818	3667	495.9
RT7	High Voltage Contract Maximum Demand-1	2786	278	10022.2
RT7Z	High Voltage Contract Maximum Demand-2	461	19	24242.5
RT8	Low Voltage Contract Maximum Demand	167	54	3101.0
RT11	Distribution Entry	5	25	194.7
RT13	Anytime Energy Residential Bi-directional Service	1566	316420	4.9
RT14	Anytime Energy Business l Bi-directional Service	28	2144	13.0
RT15	Time of Use Energy Residential l Bi-directional Service	57	10264	5.5
RT16	Time of Use Energy Business l Bi-directional Service	103	747	138.2
RT17	Part Time use of Energy Residential	109	18306	5.9
RT18	Part Time use of Energy Business	386	5852	66.0
RT19	Part Time use of Demand Residential	10	150	66.6
RT20	Part Time use of Demand Business	824	6390	129.0
RT21	Multi Part Time use of Energy Residential	0	1	0.4
RT22	Multi Part Time use of Energy Business	1	38	22.4
Other		0	14	33.0
RT9	Public Lighting	135	275857	0.5
RT10	Unmetered Supplies	51	19273	2.6
Distribution Total		**13658**	**1463745**	**9.3**
TRT1	Transmission Exit	4200	40	104996.2
Total		**17858**	**1463785**	**12.2**
Total excluding public lighting and unmetered		17672	1168655	15.1

Source: Prices 2020-21, Reference Tariffs Western Power.

Table 2.2	Western Power – Energy and customers by class and major tariff groups				
	Energy – 2019-20	Customers – Average use – 30 June 2020		Per cent share – 2020	
	GWh	Number	MWh	Energy	Customers
Generators	42	63	669.3	0.2	0.0
Large business	3934	605	6502.3	22.2	0.1
Medium business	1986	3721	533.7	11.2	0.3
Residential	4870	1023432	4.8	27.5	87.6
Small business	2678	140832	19.0	15.1	12.1
Transmission	4200	40	104996.2	23.7	0.0
Grand total	**17710**	**1168693**	**15.2**	**100.0**	**100.0**

Data preparation and validation

Western Power, in its forecasting process, undertakes a data validation procedure. The data validation procedure involves:

- testing the data against established benchmarks (e.g. comparison against previously validated data sets);
- comparative benchmark testing;
- graphical presentations of the data; and
- tops down sense-checks on sales, distribution losses, customers, PV and demand at the substation level.

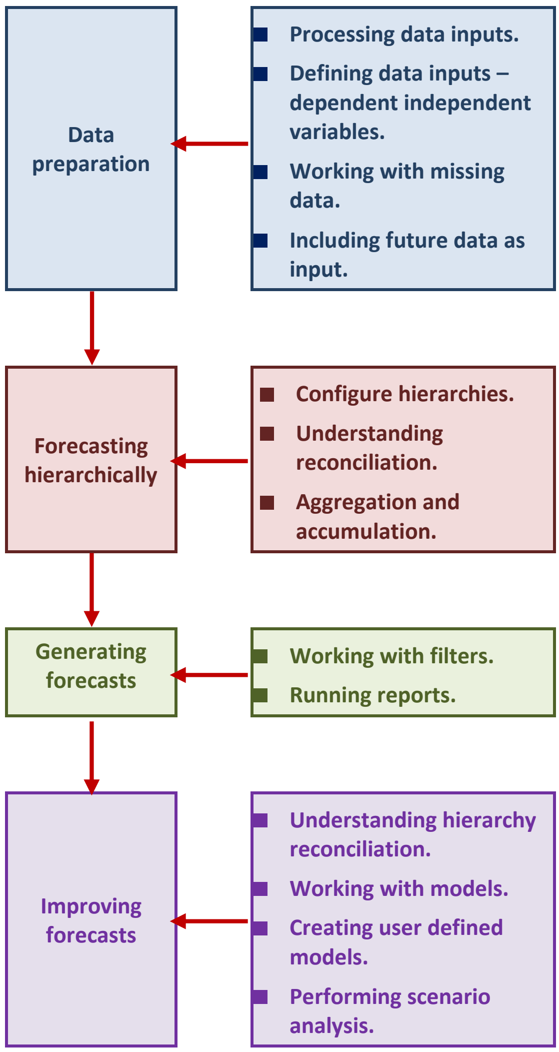
Forecast process

Western Power's forecasts of energy, customers and PV are produced using SAS Forecast Studio.[2] SAS Forecast Studio is an application that seeks to speed up the forecasting process through automation. Using this application allows the user to:

- generate forecasts automatically using a model selection list;
- create own forecast models;
- perform top down, bottom up and middle-out hierarchical forecasting;
- visually analyse and diagnostic checking; and
- perform simulations and generate reports.

Figure 2.1 shows some key elements of the SAS Forecast Studio process.

Figure 2.1: A schematic of the SAS Forecast Studio process



Source: www.sas.coms>software>support.

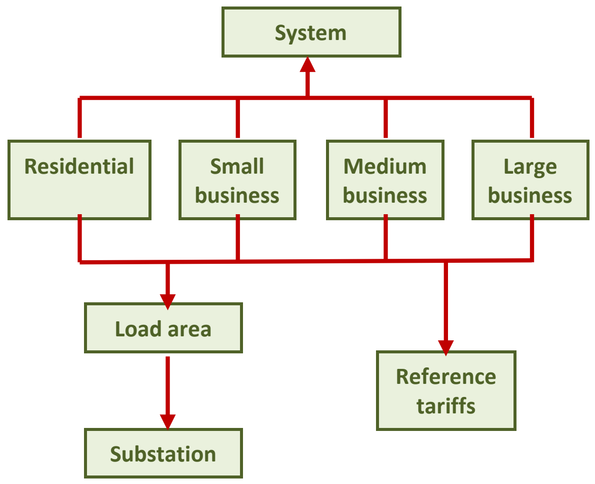
2 SAS stands for Statistical Analysis System, a software system for data analysis and report writing.

Western Power elects to utilise user defined models in SAS Forecast Studio. Three types of time series models are employed, including:

- autoregressive integrated moving average (ARIMA);
- unobserved component models; and
- multivariate regression models.

By specifying an appropriate hierarchical structure, the Western Power forecasts by tariff, class and region add up to the total. The forecast hierarchy is reproduced below.

Figure 2.2: The forecast hierarchy for energy and customers



Source: Medium-term Forecasting Methodology Information Paper, pp. 2, 9 October 2020.

The SAS Forecast Studio automatically generates an array of diagnostic test results for forecast models.

Forecast drivers

The independent variables or forecast drivers used in SAS Forecast Studio include the following.

- economic activity;
- electricity prices; and
- substitution factors.

Table 2.3 provides a description of the independent variables used by Western Power in its forecasting process.

The data in Table 2.3 are sourced from the Western Australian Government, the Australian Bureau of Statistics, the Housing Industry Association and BIS Oxford Economics (engaged by Western Power). The HIA data is based at the regional level (SA3 and SA4).

Table 2.3 Description of the independent variables used by Western Power

Category	Variable	Description
Economic	New buildings	Housing Industry Association forecast of new buildings by location.
Economic	Regional final demand	SWIS regional final demand forecasts (in million $) (annual, monthly and percentage change) prepared by BIS Oxford Economics.
Economic	Gross regional product	SWIS gross regional product forecasts (in million $) (annual, monthly and percentage change) prepared by BIS Oxford Economics.
Price	Tariff A1	Synergy residential retail tariff, variable (in $/KWh) and fix (in $/day) parts of the tariff. Forecast tariffs sourced from budget papers. Tariffs beyond the budget paper forecast horizon were assumed to remain constant in real terms.
Price	Tariff L1	Synergy business retail tariff, variable (in $/KWh) and fix (in $/day) parts of the tariff. Forecast tariffs sourced from budget papers. Tariffs beyond the budget paper forecast horizon were assumed to remain constant in real terms.
Substitution	PV count	Count of customers with a bidirectional network tariff.
Substitution	PV capacity	Sum of PV inverter capacity (in MW).

Source: Table 1, see Figure 2.2.

Western Power's medium-term forecast models do not take into account all new and emerging technologies, such as electric vehicles and batteries. Up to 2024-25 this is not likely to be that significant. It may be more significant post 2024-25.

Connections forecasts

Forecasts of connections in SAS Forecast Studio are by customer type and supply area. Information specified in the hierarchical structure ensures the customer number forecasts are reconciled at each level of the hierarchy, the top, middle and bottom.

SAS produces connection forecasts using ARIMA and Vector Auto-Regressive (VAR) models, and unobserved component models. The SAS user then selects the most appropriate model specifications based on diagnostic checks and other model selection criterion.

Customer connections by tariff are generated using an Unobserved Components Model. Some tariff groups are aggregated for this component of the forecast process. The Tariff Model also adjusts for tariff migrating where customers shift from one reference tariff to another reference tariff.

Energy forecasts

Western Power's forecasts of energy are developed in SAS Forecasting Studio. The model produces monthly energy forecasts by tariff type, customer and substation levels reconciled across the different hierarchy levels specified in Figure 2.2 above.

Different models are used across alternative customer segments, although autoregressive models are predominately utilised. The model for large business for Western Power also includes adjustment for Block Loads.

Forecasts for street lighting are obtained from a simple linear regression model for connections and an ARIMA model for energy.

Photovoltaic capacity forecasts

Western Power produces forecasts of solar PV capacity by zone substation. Forecasts are produced on a monthly basis for four classes:

- residential;
- small business;
- medium business; and
- large business.

A linear regression model was fitted for each class above. The independent variables in the regression model were the number of connections and the fixed and variable electricity tariffs. Table 2.4 shows the variables in Western Power's PV capacity model.

Block loads

Large increments to energy and demand, as well as large industry closures, are also factored into Western Power's energy forecasts.

Western Power assesses connection applications to the transmission and distribution network. Most applications are regarded as part of natural load growth.

A small number of applications could result in a significant new load on the Western Power network. The likelihood of these block loads to proceed is assessed via a consultative process, both internally and with external parties such as AEMO.

Table 2.4 PV capacity forecasts and independent variables

Variable	Description	Unit
$Res\_PVC_{i,t+1}$	Monthly residential PV capacity for substation i in month $t+1$	KVA
$\Delta Res\_PVC_{i,t}$	Monthly increment of residential PV capacity for substation i in month t	KVA
$Bus\_PVC_{i,t+1}$	Monthly business PV capacity for substation i in month $t+1$ month	KVA
$\Delta Bus\_PVC_{i,t}$	Monthly increment of PV capacity for substation i in month t	KVA
$Res\_NMI_{i,t}$	Monthly residential connections (including both historical and forecasted values) for substation i in t month from January 2008 to June 2025	Count
$Bus\_NMI_{i,t}$	Monthly (small, medium and larger) business connections (including both historical and forecasted values) for substation i in t month from January 2008 to June 2025	Count
$TariffA1_t$	Residential tariff in t month from January 2008 to June 2025, source from Energy Operators (Electricity Generation and Retail Corporation) (Charges) By-laws 2006	Cent/KWh
$TariffA1\_SC_t$	Residential service charge in t month from January 2008 to June 2025, source from Energy Operators (Electricity Generation and Retail Corporation) (Charges) By-laws 2006	Cent/day
$TariffL1_t$	Business tariff in t month from January 2008 to June 2025, source from Energy Operators (Electricity Generation and Retail Corporation) (Charges) By-laws 2006	Cent/KWh
$TariffA1\_SC_t$	Business service charge in t month from January 2008 to June 2025, source from Energy Operators (Electricity Generation and Retail Corporation) (Charges) By-laws 2006	Cent/day

Source: Medium-term Forecasting Methodology Information Paper, pp. 18, 9 October 2020.

Comments on Western Power's energy and customer number methodology

NIEIR has considered the forecasting methodologies employed in preparing forecasts of customer and energy by tariff. Overall, the approach is sound, robust and is fit for purpose. Western Power has a process manual for developing energy, connection and peak demand forecasts. This manual has been followed.

Section 6 outlines some key recommendations in respect to the forecasting methodologies for energy and customers. These include:

- increasing the range of explanatory variables or drivers for energy and customers at the substation level in SAS Forecast Studio; and
- adding battery and electric vehicles to the energy forecasts, even if these factors are considered as a post-modelling adjustment.

3. Maximum and minimum demand forecasting methodologies

3.1 Overview of forecasting approach

Western Power prepares forecasts of maximum and minimum demand for annual planning purposes and for Western Powers fifth Access Arrangement (AA5), which is submitted to the Economic Regulation Authority.

The forecasts are prepared for both maximum demand, which typically occurs during periods of extremely hot weather during summer, and for a daytime (6:00 am to 6:00 pm) minimum demand, which is becoming increasingly important as solar photovoltaics drive down daytime network demand. Forecasts of maximum demand are prepared for both summer and winter seasons. While minimum demand usually occurs during mild, sunny days in either Autumn or Spring where temperature load is minor but solar photovoltaic generation is high. Western Power measures maximum (or minimum demand) as the highest (or lowest) five-minute interval of electricity demand for a given season.

Western Power employs models at the network total level (top-down) and at the zone substation level (bottom-up). A combination of top-down and bottom up methodologies allows Western Power to reconcile the macro drivers to the local drivers of electricity demand. Forecasts are either coincident to the time of system maximum/minimum demand or non-coincident, where a non-coincident peak for a zone substation is the local maximum/minimum and does not necessarily occur at the same time as the system peak.

Western Power have adopted a new methodology to forecast electricity network demand that uses Extreme Value Theory. The Extreme Value Theory approach has been in development since 2017 and replaces the previous deterministic load factor approach to modelling maximum demand.

Forecasts and historical peak demand are presented as a probability of exceedance distribution with:

- 10^{th} Probability of exceedance level (POE10), which can be expected to be exceeded every 1 in 10 years;
- 50^{th} Probability of exceedance level (POE50), which can be expected to be exceeded every 5 in 10 years; and
- 90^{th} Probability of exceedance level (POE90), which can be expected to be exceeded every 9 in 10 years.

A robust methodology will generate historical probability of exceedance distributions for maximum demand that align with the expected behaviours. This means, for example, if ten summer seasons of historical data were available, then half would be expected to be above the POE 50 level of demand and half would be below the POE 50 level of demand. A longer historical series is therefore preferred to ensure that any model is providing a good fit and displaying the desired probability characteristics.

3.2 Data validation and preparation

NIEIR finds that Western Power's approach to data validation and preparation is thorough and rigorous. This starts by ensuring separation between the source data, the intermediate calculations and the final outputs.

To ensure that the data used is valid, Western Power treats the data through numerous steps including:

1. check the raw data against previously verified data for an overlapping period;
2. compare the new raw data against alternative credible data sources;
3. make an assessment of the patterns and trends exhibited by the new data against the old data;
4. use reviewer judgement as to whether the data is valid or problematic;
5. scrub data for errors or switching events, which could lead to false positives for maximum/minimum demand.

3.3 Methodological approach

Since 2017, Western Power have developed a new method for forecasting maximum demand using Extreme Value Theory (EVT). Previously, Western Power used a load factor approach to forecast maximum demand.

This EVT approach is based on a modelling extreme deviations from the median of probability distributions. The new methodology has been peer reviewed and published in an international journal *IEEE Transactions on Power Systems*.

Secondly, Western Power also use Quantile Regression instead of the EVT approach in some cases.

The main drivers behind the forecasts are:

- customer numbers
- energy consumption; and
- PV capacity installed.

Each of these series are originally available on a monthly basis and are allocated into a daily series using various smoothing techniques. Daily maximum demands are extracted from the relevant 5-minute interval series. Daily maximum demands are further analysed for outliers, which involves identifying observations that deviate outside a certain threshold, and then investigating the reason behind potential outliers and removing if necessary.

The forecasts for each of these three series indirectly take into consideration other drivers of maximum demand, such as the economic outlook, electricity prices, and substitution effects.

Each of the models is briefly outlined below.

Model 1: Extreme Value Theory (EVT)

In Extreme Value Theory, the series of annual maximum demand can be fitted by a Generalised Extreme Value (GEV) distribution. Western Power does not directly estimate a GEV distribution with annual maximum demand, but the parameters of the GEV distribution are estimated for a potential 12 candidate models fitted to daily data. Each of the parameters within the candidate models is specified by a functional form that is a combination of customer numbers (Z_1), energy consumption (Z_2) and PV capacity installed (Z_3). Table 3.1 shows the candidate models used to fit the substation data. For example, Model B is specified as a function of customer counts only. Model D contains both customer numbers and energy consumption, while Model G contains all three of the drivers, including PV capacity installed.

The daily demands that are used for the candidate models are censored below a certain threshold so that they take into account the more extreme peaks only.

Each candidate model is estimate for each zone substation or network total demand. The best candidate model is selected through an evaluation process.

(1) Checking to see how well the candidate model fits the data. The Akaike information criteria (AIC) is used for this purpose.

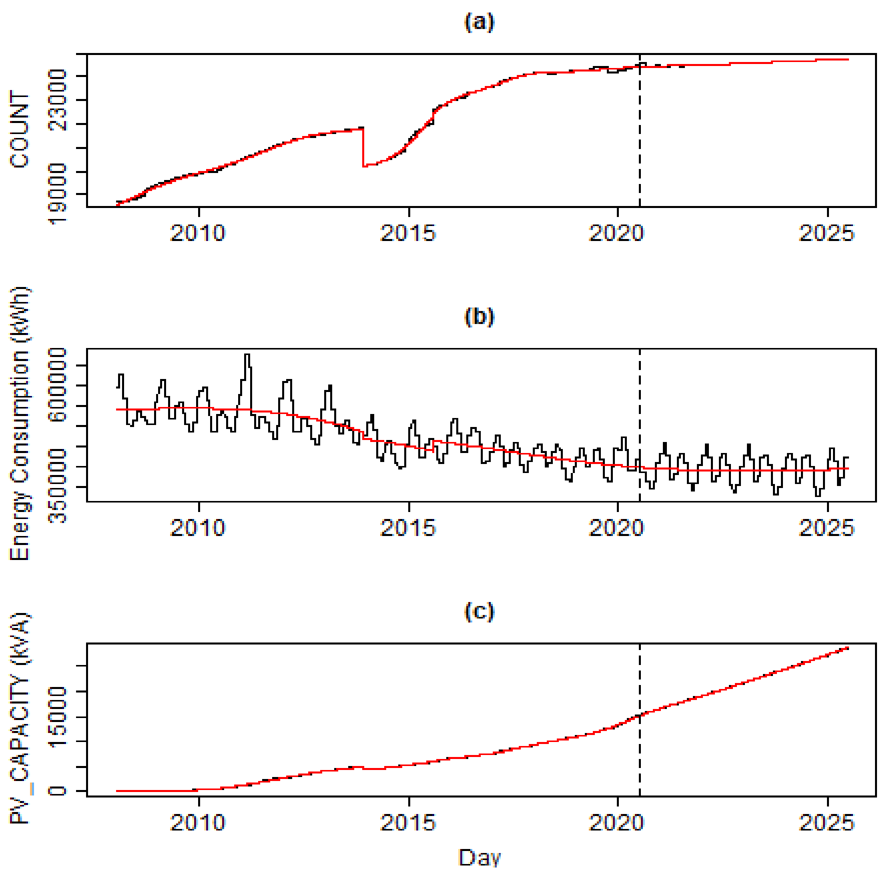
(2) Testing the significance of the difference between the candidate non-stationary model and the stationary model.

(3) Check to see whether the chosen candidate model is valid. The validity of the model is assessed by checking the signs on customer numbers, energy consumption and PV capacity installed. One would expect customer numbers and energy consumption to increase demand, while PV capacity to decrease demand.

Once the GEV distribution has been parameterised using the best fit models, the probability of exceedance levels (0.1, 0.5 and 0.9) enters the equation as a separate parameter. The historical and forecast POE levels are then generated using the three historical and forecast drivers of maximum demand.

Figure 3.1 shows an example of the drivers that are used to estimate the parameters for the GEV distribution and push forward the forecasts at the Arkana zone substation. While Figure 3.2 shows an example of the fitted POE distributions of demand for the Muchea zone substation as well as the historical demand observations and the censoring threshold.

Figure 3.1: Arkana substation – Drivers of maximum demand



Note: Black line shows the original historical and forecast data. The red line is the smoothed series that is used in the maximum demand EVT modelling.

Source: Western Power, *Medium-term forecasting methodology information paper: Connections, energy, demand and reactive power*.

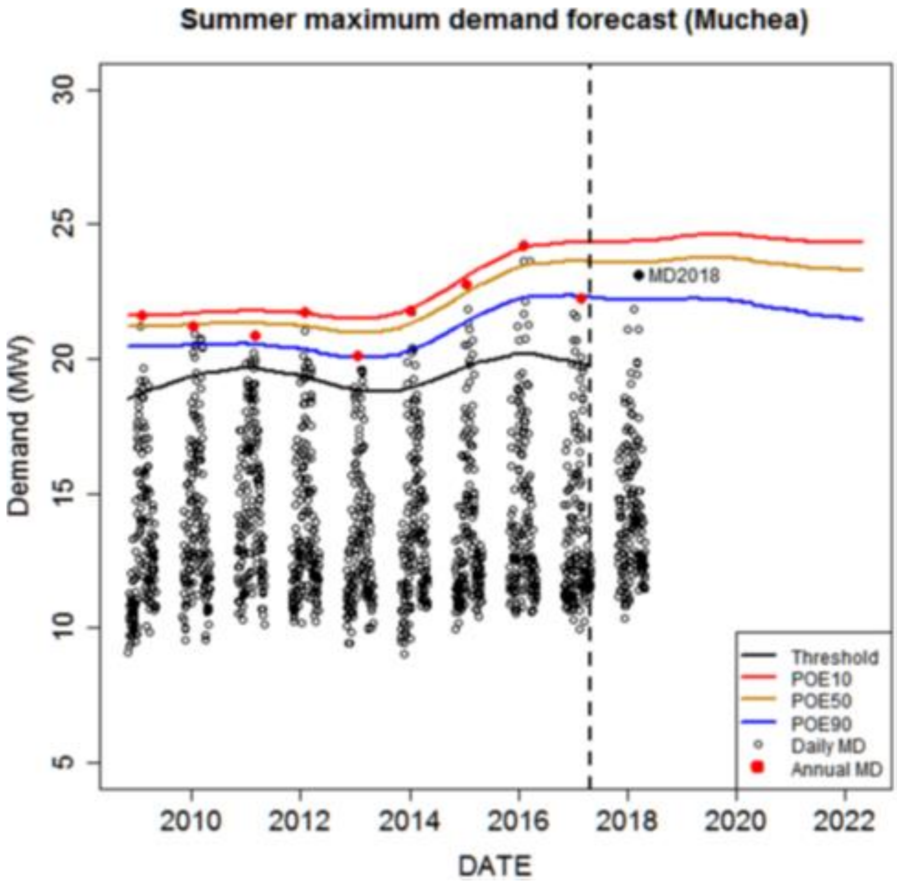
Table 3.1 Candidate point process models used to fit the substation data

Model	*Notation*	*Location*	*Scale*	*Shape*	*Covariate*
A	mu0	$\mu(t)=\mu$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Stationary
B	mu1	$\mu(t)=\mu_0+\mu_1 Z_{1t}$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Z_{1t}
C	mu2	$\mu(t)=\mu_0+\mu_1 Z_{2t}$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Z_{2t}
D	mu12	$\mu(t)=\mu_0+\mu_1 Z_{1t}+\mu_2 Z_{2t}$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Z_{1t} & Z_{2t}
E	mu13	$\mu(t)=\mu_0+\mu_1 Z_{1t}+\mu_3 Z_{3t}$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Z_{1t} & Z_{3t}
F	mu23	$\mu(t)=\mu_0+\mu_2 Z_{2t}+\mu_3 Z_{3t}$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Z_{2t} & Z_{3t}
G	mu123	$\mu(t)=\mu_0+\mu_1 Z_{1t}+\mu_2 Z_{2t}+\mu_3 Z_{3t}$	$\sigma(t)=\sigma$	$\xi(t)=\xi$	Z_{1t}, Z_{2t} & Z_{3t}
H	mu2_sig3	$\mu(t)=\mu_0+\mu_2 Z_{2t}$	$\sigma(t)=\exp(\sigma_0+\sigma_1 Z_{3t})$	$\xi(t)=\xi$	Z_{2t} & Z_{3t}
I	mu12_sig3	$\mu(t)=\mu_0+\mu_1 Z_{1t}+\mu_2 Z_{2t}$	$\sigma(t)=\exp(\sigma_0+\sigma_1 Z_{3t})$	$\xi(t)=\xi$	Z_{1t}, Z_{2t} & Z_{3t}
J	mu13_sig3	$\mu(t)=\mu_0+\mu_1 Z_{1t}+\mu_3 Z_{3t}$	$\sigma(t)=\exp(\sigma_0+\sigma_1 Z_{3t})$	$\xi(t)=\xi$	Z_{1t} & Z_{3t}
K	mu23_sig3	$\mu(t)=\mu_0+\mu_2 Z_{2t}+\mu_3 Z_{3t}$	$\sigma(t)=\exp(\sigma_0+\sigma_1 Z_{3t})$	$\xi(t)=\xi$	Z_{2t} & Z_{3t}
L	mu123_sig3	$\mu(t)=\mu_0+\mu_1 Z_{1t}+\mu_2 Z_{2t}+\mu_3 Z_{3t}$	$\sigma(t)=\exp(\sigma_0+\sigma_1 Z_{3t})$	$\xi(t)=\xi$	Z_{1t}, Z_{2t} & Z_{3t}

Notes: Z_{1t} = Customer count.
Z_{2t} = Energy consumption.
Z_{3t} = PV capacity.

Source: Western Power, *Medium-term forecasting methodology information paper*.

Figure 3.2: Muchea substation – Summer maximum demand forecast (2018)



Source: Western Power, *The Use of Extreme Value Theory for Forecasting Long-Term Substation Maximum Electricity Demand*.

Model 2: Quantile Regression Model (QRM)

Alternatively, Western Power have also employed a Quantile Regression Model (QRM) to model maximum and minimum demand in some cases. The QRM model is computational quicker than the EVT and is used when implementation speed is a priority. The QRM model directly estimates linear equations based on the three drivers Western Power uses to drive demand; customer numbers, energy consumption and solar PV capacity. The linear equations are fitted to daily load data, and are used to estimate the quantile for a given probability of exceedance level. The Quantiles are adjusted so that the POE demands fit the annual peaks, rather than daily peaks, which the equations are modelled on.

Block load forecasts

Major expansions of existing customers and new large customer connections have the potential introduce a step change in forecast demand that is largely outside the capability of statistical techniques to forecast. Western Power adds forecasts of block loads to demand if the potential customer has a high likelihood to proceed over the next five years. The likelihood to proceed is based on internal and external stakeholder consultation. This includes consultation with AEMO.

3.4 Methodology and forecast validation

Western Power employs a number of techniques to validate their forecasts, this includes:

(i) using out of sample observations to test the model;

(ii) comparing forecasts with AEMO; and

(iii) model choice using statistical test to get the best fit.

NIEIR finds that overall, the Western Power validation process is sound.

3.5 Comparison to other methodologies

The most common methodologies used to forecast maximum demand for Australian electricity networks can be broadly categorised into:

(i) deterministic models of maximum demand that may take a load factor approach and use long-term weather standards to generated probability of exceedance levels; and

(ii) half-hourly models that model each half hour during the day by making use of simulation techniques.

The use of General Extreme Value type models has become more prominent over the past few years with AEMO and other Australian electricity networks (such as Evoenergy in the ACT) estimating GEV distributions.

AEMO

AEMO use both a half-hourly simulation model and GEV models to forecast maximum and minimum demand. AEMO use a GEV model to forecast short term maximum demand (one year ahead forecasts) and use the half-hourly simulation model to forecast medium to long term maximum demand over 20 to 30 years ahead[3].

AEMO's GEV model was fitted for each region using variables for month, rooftop PV capacity, non-scheduled solar generators, NMI count, average temperature and solar irradiance.

A key difference between the GEV and half-hourly model is the ability to analyse the changing patterns of demand across the day as the result of PV generation, other new technologies, and government programs and policies. This is increasingly important as PV continues to increase and electric vehicles are expected to have a significant impact, especially in the 2030's decade.

Evoenergy

Evoenergy have adopted similar concepts to forecast zone substation maximum demand starting with their 2020 Annual planning report[4]. Extreme temperatures are fitted with a Gumbel distribution to forecast maximum demand. Previously, Evoenergy had used a half-hourly demand model (Monash Electricity Forecasting Model) for annual planning.

Evoenergy estimate their model using season/time of day, temperature, and 'baseline' demand. The forecasts are mainly driven by time trends and trends in weather.

NIEIR

NIEIR conducted preliminary modelling of Western Power system total network summer maximum demand using PeakSim. NIEIR has previously used this model to forecast the SWIS network. The PeakSim model is a half-hourly simulation model that segments demand into base load and temperature sensitive load for each half hour and season. POE distributions of demand are estimated by bootstrapping historical weather.

Western Power provided NIEIR with 5-minute interval demand data for the system total network from 2008 to 2021. The 5-minute interval data was averaged into half-hourly interval data so that it could be run through NIEIR's PeakSim model.

The following observations were made based on comparing the historical POE series for summer maximum demand:

- The spread between POE bands is similar between the methodologies. That is the difference between POE 10 and POE 50; and between POE 50 and POE90 levels;
- Both 2020 and 2021 years have very similar POE levels between the methodologies. 2021 was the first forecast year for Western Power, while NIEIR's POE demand distributions were historically estimated;
- The historical POE distribution from PeakSim shows more historical variation than Western Powers methodology. This is partly explained by the PeakSim model tracking trends in weather impacts (air conditioner use); and
- However, the average annual difference over 2010 to 2021 at the POE 50 level is relatively small. Western Powers historical POE 50 summer demand is 0.13 per cent higher than the PeakSim model on average over 2010 to 2021.

[3] Electricity Demand Forecasting Methodology Information Paper August 2020, *AEMO.*

[4] Annual Planning Report 2020, December 2020, *Evoenergy.*

3.6 Discussion and recommendations

Since NIEIR last reviewed Western Powers demand forecasts in 2017, there have been several beneficial improvements to the methodology. These improvements have led to greater internal consistency between the forecasts produced by Western Power.

This includes:

- modelling electricity demand at the network level (top-down) and later reconciling these forecasts to the zone substation (bottom-up) forecasts;
- ensuring greater consistency between forecasts of energy consumption and peak demand;
- forecasting minimum demand at the network level and zone substation level;
- improvements to fitting maximum demand through using GEV distributions;
- historical POE distributions appear to show more stability (smoother profiles) across the years than what was estimated from the previous methodology. That is there appears to be less volatility between years.

There are certain trade-offs that need to be made when considering model choice.

Timing

The models are specified using the daily maximum or minimum for the network or zone substation. They do not specifically take into account peak timing. For example, under an increasing stock of PV capacity, daytime load will experience downwards pressure and this has the potential to push the timing of the peaks into the evening to low lighting conditions. This implies that the historical impact of PV capacity on peak demand (as estimated by a coefficient) will decrease into the forecasting period as the timing shifts backwards. There may be timing shifts when Electric Vehicles become more common place, or battery storage penetration increases.

A separate, complimentary half-hourly model to estimate medium term to long term trends in new technologies may improve the forecasts.

Other shifts in electricity usage

Over the next ten years, especially closer to 2030, both Electric Vehicles and Battery storage have the potential to have a large impact on electricity demand. Even if they now only have a small impact on electricity networks, inclusion within the methodology and forecasts would prepare Western Power for the future.

NIEIR recognises that Western Power is actively monitoring these technologies but has not included them in the forecasts.

Weather

Western Power do not directly use weather in the EVT model, but weather impacts are indirectly considered through the energy consumption driver. Depending on the network NIEIR estimates that around 40 to 60 per cent of electricity demand at system level summer maximum demand is due to temperature load. This is driven by the use of temperature sensitive equipment, such as air conditioners during summer peak demand. This is a much higher proportion of electricity usage than for annual energy consumption (or average demand). Changes in air conditioner use, such as increased stock or energy efficiency improvements, may not be directly accounted for within Western Powers current methodology when driven by average demand alone.

4. Summary of recommendations

This section summarises the recommendations following the review of Western Power forecasting methodologies for energy forecasting and peak demand forecasting.

Table 4.1 shows these recommendations. NIEIR acknowledges that not all of these recommendations would necessarily be pursued by Western Power. Some recommendations could compromise other objectives of the forecasting process. Confidential survey information collected from major customers for example, should not be released to third parties compromising the 'transparency' objective.

Table 4.1 Summary of recommendations/suggestions for improvement in Western Power forecasts

Proposal	Variables
PART A: Energy	
1. Modify extrapolated forecast inputs so that energy customers and energy forecasts extend out 8-10 years (rather than only 5 years).	Modify SAS Forecast Studio inputs for forecasting.
2. Integrate zone sub-station specific driver variables via mapping ABS LGAs or ABS Statistical Areas (SA1 – SA4).	Population, dwelling stock, real income, gross product by sector or industry (NIEIR quarterly regional data sets).
3. Industry coding of Western Power NMIs to facilitate proposal 2 (above).	NMIs by network tariff and industry – large, medium and small business.
4. Major customer survey – large business.	Anticipated energy and load changes over next five years.
5. Modelling old and new customers (residential initially) separately.	Collating data by class and year of connection, weather normalising and conducting statistical analysis.
6. Applying post-modelling adjustments to energy projections if required.	Possible State and Commonwealth energy and environmental programs and new technologies, such as batteries and electric vehicles.
7. Procure or develop regional specific battery and EV forecasts covering installations, capacity and energy.	These models should combine economic and socio-demographic drivers of battery use and EV. Sociodemographic includes age structure, dwelling type, income and geographical location.
PART B: Maximum demand	
1. Directly estimate the impact of weather on maximum demand by including weather as an additional variable in the EVT/QRM models.	Dry temperature at nearest weather station for each zone substation. Air conditioner stock/use.
2. Estimate long term weather standards for Western Australian weather stations for use in 1.	Calculate temperature sensitivities and temperature POEs.
3. Development of long-term drivers of peak demand models by zone sub-station (post 2024-25).	These data could possibly be developed from proposal number No. 1 under Part A: Energy.
4. Include forecasts demand impacts of electric vehicles and battery storage.	
5. Construct half-hourly (or 15 minute) interval models of new technologies to assess the changes in demand across the day and the influence on maximum demand forecasts.	PV, electric vehicles and battery storage network impacts should be considered for each interval. Either estimated as a part of an integrated half-hourly model, or as a post-modelling adjustment to current EVT/QRM models.
6. Further testing at the zone substation level to detect divergence of forecast POE compared to historical POE distributions.	Affects a small number of zone substations.

Glossary

Adjusted *R-Squared*	A goodness-of-fit measure in multiple regression analysis that penalises additional explanatory variables by using a degrees of freedom adjustment in estimating the error variance.
Bias	The difference between the expected value of an estimator and the population value that the estimator is supposed to be estimating.
Bootstrap	A resampling method that draws random samples, with replacement, from the original data set.
Classical Linear Model	The multiple linear regression model under the full set of classical linear model assumptions.
Combined cycle generator	An electric generating unit in which electricity is produced from waste heat exiting from one or more combustion turbines. This process increases the efficiency of the electric generating unit.
Confidence interval (CI)	A rule used to construct a random interval so that a certain percentage of all data sets, determined by the confidence level, yields an interval that contains the population value.
Confidence level	The percentage of samples in which we want our confidence interval to contain the population value; 95 per cent is the most common confidence level, but 90 and 99 per cent are also used.
Correlation coefficient	A measure of linear dependence between two random variables that does not depend on units of measurement and is bounded between 1 and 1.
Covariance	A measure of linear dependent between two random variables.
Cross-sectional data set	A data set collected by sampling a population at a given point in time.
Data frequency	The interval at which time series data are collected. Yearly, quarterly and monthly are the most common data frequencies.
Degrees of freedom (*df*)	In multiple regression analysis, the number of observations minus the number of estimated parameters.
Demand	The amount of electricity being consumed by customers at any given time.
Demand side participation (DSP)	Refers to a voluntary reduction in net demand, typically in response to a pool price signal, achieved either by switching off a load or running an embedded generator to reduce the net level of demand. Although the NEC provides scope for participants to bid their loads directly into the pool, DSP is more usually an off market arrangement negotiated directly between a participant and an end-user of electricity or the owner of an embedded generator. The extent and terms of these bilateral contracts are not publicly available, making it difficult to predict future levels of DSP. DSP may also be referred to as demand side response, load curtailment contracts or demand side management.
Dependent variable	The variable to be explained in a multiple regression model (and a variety of other models).
Descriptive statistic	A statistic used to summarise a set of numbers; the sample average, sample median and sample standard deviation are the most common.
Distributed Lag Model	A time series model that relates the dependent variable to current and past values of an explanatory variable.
Distribution network	The system of wires, switches, transformers and other related infrastructure that delivers electricity from the high-voltage transmission network to the end-use consumer for use in homes and businesses.
Econometric Model	An equation relating the dependent variable to a set of explanatory variables and unobserved disturbances, where unknown population parameters determine the ceteris paribus effect of each explanatory variable.

Elasticity	The percentage change in one variable given a 1 per cent ceteris paribus increase in another variable.
Embedded generators	A generator whose generating units are connected to a distribution system (rather than the high voltage transmission system). Such generators are usually small and often cogeneration facilities.
Error Correction Model	A time series model in first differences that also contains an error correction term, which works to bring two I(1) series back into long-run equilibrium.
Estimator	A rule for combining data to produce a numerical value for a population parameter; the form of the rule does not depend on the particular sample obtained.
Exogenous variable	Any variable that is uncorrelated with the error term in the model of interest.
Forecast error	The difference between the actual outcome and the forecast of the outcome.
Gigawatt (GW)	One gigawatt equals 1 billion watts, 1 million kilowatts or 1 thousand megawatts.
Gigawatt-hour (GWh)	One gigawatt-hour equals one billion watt-hours.
Kilowatt-hour (KWh)	The basic unit of electric energy equal to one kilowatt of power supplied to or taken from an electric circuit steadily for one hour. One kilowatt-hour equals 1,000 watt-hours. One kilowatt-hour can power ten 100 watt light bulbs for one hour.
Least squares estimator	An estimator that minimises a sum of squared residuals.
Limited Dependent Variable (LDV)	A dependent or response variable whose range is restricted in some important way.
Linear function	A function where the change in the dependent variable, given a one-unit change in an independent variable, is constant.
Load	The amount of electricity needed to meet demand at any given time.
Loss factor	A multiplier used to describe the additional electrical energy loss for each increment of electricity used or transmitted.
Losses	For the end-user to obtain electricity at the desired location and at the desired voltage level, electricity must be transmitted through wires and transformed. Throughout this process some electricity-energy is converted into heat-energy resulting in the 'loss' of end-use electricity.
Mean absolute error (MAE)	A performance measure in forecasting, computed as the average of the absolute values of the forecast errors.
Mean squared error (MSE)	The expected squared distance that an estimator is from the population value; it equals the variance plus the square of any bias.
Median	In a probability distribution, it is the value where there is a 50 per cent chance of being below the value and a 50 per cent chance of being above it. In a sample of numbers, it is the middle value after the numbers have been ordered.
Megawatt (MW)	One megawatt equals one million watts.
Megawatt-hour (MWh)	One megawatt-hour equals one million watt-hours. One MWh of electricity can power ten thousand 100 watt light bulbs for one hour.
Missing data	A data problem that occurs when we do not observe values on some variables for certain observations (individuals, cities, time periods, and so on) in the sample.
Normal distribution	A probability distribution commonly used in statistics and econometrics for modelling a population. Its probability distribution function has a bell shape.
Off-peak	A time span of lower electricity usage, which would typically include public holidays, weekends and 9:00 p.m. to 7:00 a.m. on weekdays.
Ordinary least squares (OLS)	A method for estimating the parameters of a multiple linear regression model. The ordinary least squares estimates are obtained by minimising the sum of squared residuals.
Peak	A time span of higher electricity usage, which would normally include non-holiday weekdays from 7:00 a.m. to 9:00 p.m.
***R*-Square**	In a multiple regression model, the proportion of the total sample variation in the dependent variable that is explained by the independent variable.
Random sample	A sample obtained by sampling randomly from the specified population.

Residual	The difference between the actual value and the fitted (or predicted) value; there is a residual for each observation in the sample used to obtain an OLS regression line.
Seasonality	A feature of monthly or quarterly time series where the average value differs systematically by season of the year.
Sent out energy	The amount of electricity supplied by a scheduled generator to the transmission or distribution network at its connection point.
Standard deviation	A common measure of spread in the distribution of a random variable.
Standard error of the regression (SER)	In multiple regression analysis, the estimate of the standard deviation of the population error, obtained as the square root of the sum of squared residuals over the degrees of freedom.
Stationary process	A time series process where the margins and all joint distributions are invariant across time.
Statistical significance	The importance of an estimate as measured by the size of a test statistic, usually a '*t*' statistic.
Sub-station	An assemblage of electrical equipment for the purposes of changing and/or regulating the voltage of electrical circuits.
Time series data	Data collected over time on one or more variables.
Transformer	A device that converts electricity from one voltage level to another.
Transmission network	Electricity power lines and associated infrastructure that convey electricity between certain generators and distribution systems.
Transmission Network Service Provider (TNSP)	An organisation which engages in the activity of owning, controlling or operating a regulated transmission system.
Unbiased estimator	An estimator whose expected value (or mean of its sampling distribution) equals the population value (regardless of the population value).
Watt-hour	The total amount of energy used in one hour by a device that uses one watt of power for continuous operation. Electric energy is commonly sold by the kilowatt-hour.

Appendix A: Documentation reviewed by NIEIR

Western Power provided the following supporting materials for the forecast review:

- Energy and Customer Numbers Forecast 2020.xlsx;
- Energy & Customer Numbers Report 2020;
- CUSTED2020 Forecast Method Report;
- 2020 Day minimum demand forecasts by zone substation report;
- 2020 Maximum demand forecasts (SUMMER) by zone substation report; and
- 2020 Maximum demand forecasts (WINTER) by zone substation.
- Upon request, Western Power also provided the following information:
- Substation_demand_Forecast_data.xlxs; and
- BIDA_NMI_PV_Forecasts_2021_05_12 Final.

Appendix B: NIEIR's relevant experience

The National Institute of Economic and Industry Research was founded in 1984. NIEIR's current client base includes:

- energy retailers;
- distribution and transmission network businesses;
- industry planning and energy markets operators;
- government departments and agencies; and
- business and community groups.

NIEIR was employed by Western Power to review its forecasting methodologies for energy, customers and peak demand in 2005, 2010 and again in 2016. NIEIR also has experience in preparing energy and peak demand forecasts for the SWIS for over 10 years for AEMO (formerly the Independent Market Operator (IMO)). Gas demand forecasts for Western Australian pipelines were also prepared by NIEIR for AEMO up to 2015. NIEIR currently produces commodity forecasts to support AEMO's gas forecasting process.

NIEIR is one of the premier energy demand forecasters in Australia; NIEIR regularly prepares energy demand forecasts for:

- medium- and long-term network planning;
- budgeting and capital expenditure planning;
- regulatory price reviews, including new tariff designs (time-of-use, cost reflective);
- uptake of disruptive technologies such as plug-in electric vehicles, photovoltaic systems, and electricity storage systems;
- strategic and investment planning decisions;
- energy efficiency and climate change policy studies;
- in-house research and analysis; and
- due diligence process in asset acquisition process.

As part of these services, NIEIR prepares:

- energy forecasts;
- demand (peak-day issue or maximum daily/hourly quantity) forecasts; and
- customer number forecasts.

These metrics are disaggregated into various segments including:

- sectoral and industry classifications;
- geographical dimensions (including national, state, regional and sub-regional level);
- tariff and customer type; and
- technology type (hot water, space conditioning, etc.)

The following table lists organisations that NIEIR has successfully delivered the above services for over the past five years. Many of these companies remain long term NIEIR clients. This demonstrates that NIEIR is a professional and trusted source of expertise within the Australian energy industry. Selected case studies for energy sector projects are presented in Appendix B and NIEIR's recently commissioned reports are listed in Appendix C. The complete list of commissioned reports is available upon request.

■ Agility Management (AGL Electricity)	■ AEMO
■ United Energy	■ Western Power
■ Envestra (Australia Gas Networks)	■ ENERGEX
■ CitiPower	■ Multinet
■ Powercor Australia	■ Powerlink Queensland
■ Endeavour Energy	■ Essential Energy
■ Ergon Energy	■ Ausgrid
■ Jemena	

NIEIR also issues reports and reviews on such issues as macroeconomic forecasts, energy efficiency improvement (EEI) potential and programs, the take-up of photovoltaic systems, plug-in electric vehicles, energy storage systems and substitution of electrical technology such as heat pumps in the place of gas heating technology.