Attachment 6.2

Fitting Distributions for AA4 Service Standard KPIs-Setting the Service Standard Benchmark (SSB) and Service Standard Target (SST)

Access Arrangement Information

2 October 2017

Access Arrangement Information (AAI) for the period
1 July 2017 to 30 June 2022

westernpower

Fitting Distributions for AA4 Service Standard KPIs - SSB and SST

Public

2 October 2017

westernpower

EDM#43111920

An appropriate citation for this paper is:

Fitting Distributions for AA4 Service Standard KPIs - SSB and SST

Western Power

363 Wellington Street
Perth WA 6000
GPO Box L921 Perth WA 6842

T: 13 10 87 | Fax: 08 9225 2660
TTY 1800 13 13 51 | TIS 13 14 50

Electricity Networks Corporation
ABN 18 540 492 861

enquiry@westernpower.com.au
westernpower.com.au

Enquiries about this report should be directed to:

AA4 Project Team

Email: AA4@westernpower.com.au

Contents

1. Summary and Purpose

The purpose of this report is to document the method and outcome for developing performance metrics for the fourth Access Arrangement (**AA4**).

The performance metrics take two forms:

- Service Standard Targets (**SSTs**) under the Service Standard Adjustment Mechanism (**SSAM**) that dictate financial incentives/disincentives
- Service Standard Benchmarks (**SSBs**) that dictate minimum service standards

Minor changes have been applied to an existing method, which involves fitting statistical distributions; allowing a set of proposed SSTs and SSBs that align closely with the AA4 proposed investment program.

1.1 Reason for benchmarking

Section 11.1 of the Electricity Networks Access Code 2004 states that:

A service provider must provide reference services at a service standard at least equivalent to the service standard benchmarks set out in the access arrangement and must provide non-reference services to a service standard at least equivalent to the service standard in the access contract.

By establishing a benchmark, Western Power's customers can assess the level of service performance that they can reasonably expect and that performance can be measured over time. Implicitly, there is a trade-off between the cost of providing reference services and the likely benefit in the form of achievable performance levels.

1.2 Benefits of benchmarking

Measuring service standard performance provides the basis on which to identify deterioration in service as well as help quantify opportunities for improvement. By establishing an objective yardstick, Western Power can assess its performance and take corrective action as required. Moreover, Western Power is well positioned to determine the appropriate level of expenditure to ensure service remains within the required tolerances.

1.3 Principles guiding the benchmarking process

The challenge is establishing a reasonable expectation given the inherent random variation or 'noisiness' in the data. The approach taken in this report is to formally apply a statistical distribution, which explicitly recognises the observed level of randomness.

However, it is important to note that in conducting the statistical analysis, close attention is given to ensuring that the benchmarking process is consistent with the following principles:

- The resulting benchmarks are practical to administer.
- They are adaptive, allowing for changes in the operating environment that are outside Western Power's control.
- While adaptive, the benchmarks should be relatively stable and with minor movement over time being predictable.

- The performance measures should lead to meaningful benefit for customers.

Western Power believes that adopting these principles helps ensure a systematic approach to managing service standards over time. In practice, this means that the text book statistical procedures have been modified to better serve these principles.

The primary outcome is a performance measurement system that conveys clear signals indicating when remedial action is required.

2. Method

Key messages:

- Western Power has based its method firmly on statistics literature. This section provides a detailed description of the standard approach identified in the literature review.
- The standard approach has two parts: visual inspection; and calculation of goodness-of-fit statistics. The visual inspection process, based on P-P and Q-Q plots, serves to visually confirm the statistical results.
- The goodness-of-fit statistics chosen by Western Power for evaluating candidate statistical distributions are the Anderson-Darling statistic for testing continuous distributions; and the Chi-squared statistic for discrete distributions. These statistics are used to reject statistical distributions that are poor representations of the underlying Data Generating Process.
- The Akaike Information Criterion (**AIC**) is used to rank the remaining candidate statistical distributions in terms of how well each distribution fits the data.[1]
- Where the AIC of more than one statistical distribution is close to the lowest AIC score, Western Power has averaged the relevant percentiles across these distributions to derive the corresponding service standard benchmarks and targets. This approach better serves the principle of stability in benchmarking than simply selecting the statistical distribution with the lowest AIC score.

2.1 Introduction

A process for setting SSTs and SSBs was developed in consultation with the Economic Regulation Authority (**ERA**) for the third access arrangement (**AA3**) which is similar to that used by other jurisdictions for the equivalent of SST setting. The process involves fitting Probability Distribution Functions (**PDF**) to historic 12 month rolling average performance data. It is proposed that the AA4 method build on this approach.

2.2 Discussion

Probability distribution fitting describes the process of determining the PDF that best represents the historic data. The objective of probability distribution fitting is to determine which PDF most accurately predicts the observed frequency of observations (i.e. performance measurements).

Before explaining the process for identifying the best fitting PDF, it is necessary to discuss two broad types of distributions. Namely, PDFs that model:

- Continuous variables (e.g. age, weight, height etc) that can be measured at any interval along a number line; and
- Discrete variables that measure the probability of countable outcomes (e.g. coin flips, customer complaints etc).

Irrespective of the type of random variable being investigated, the general distribution fitting process involves the following four steps:

1. Model each candidate PDF

[1] The lower the AIC score, the better the fit of the distribution to the Data Generating Process.

2. Estimate PDF parameters
3. Reject poor fitting PDFs using goodness-of-fit statistical test
4. Select the best PDF based on the quality of fit statistic.

2.3 Fitting continuous PDFs

Two methods for fitting a distribution to continuous variables are:

- Visual inspection of diagnostic plots
- Maximum Likelihood Estimation (**MLE**) and selection based on statistical tests; in this case the p-value of the Anderson-Darling statistic and the Akaike Information Criterion.

Western Power has applied both the visual and the MLE method to determining the best fitting distribution for each performance indicator.

Note that in fitting PDFs, there is an assumption that the chosen PDF is an accurate predictive model of the underlying Data Generating Process. That is, the chosen PDF is deemed to provide a reasonable reflection of the process that causes measured observations to deviate from the centre of the distribution of observations. Particularly important issues are:

- The stability of the measures of central tendency of the observations, such as the mean, median and mode. For example, there is often a strong assumption of a stationary mean. Time series data can exhibit trends, which implies an unstable centre of the distribution of observations. A secondary issue is whether any movement in the mean is due to systematic or random influences.
- The strength of the central tendency, which is measured in terms of the average distance from the centre of the PDF. The higher the average distance, the less strength in central tendency. This implies that the measures of central tendency could be poor predictors of future observations.
- The degree of any asymmetry in the collection of observations. Asymmetry implies that the probability of observing values below the measured centre of the distribution is not the same as the probability of observing values above the measured centre.

2.3.1 Visual inspection

The main visual diagnostic tool is the quantile-quantile (**Q-Q**) scatter plot of observations (i.e. data points) drawn from a specific PDF (e.g. the Normal distribution) against the sampled observations. Note that quantile refers to the division of a sample of observations into equal-sized subgroups. A plot showing quantiles provides a visual guide to the proportion of observations that are below a given value, for example half of observations lie below the median.

The Q-Q plot provides a visual guide of how well observations map to a chosen PDF. A Q-Q plot typically shows a scatter plot of observations against a reference line rising from left to right. A perfect match between observed values and theoretical values drawn from a chosen PDF would be indicated by all the plotted observations lying directly over the reference line.

When dealing with real world data, it is unlikely that all observations would lie perfectly along the reference line, even if the correct PDF is used as the benchmark in the Q-Q plot. This can make deciding which PDF is best to model differences within families of PDFs difficult. Choosing the best distribution is typically more difficult when there are few observations, e.g. less than 100.

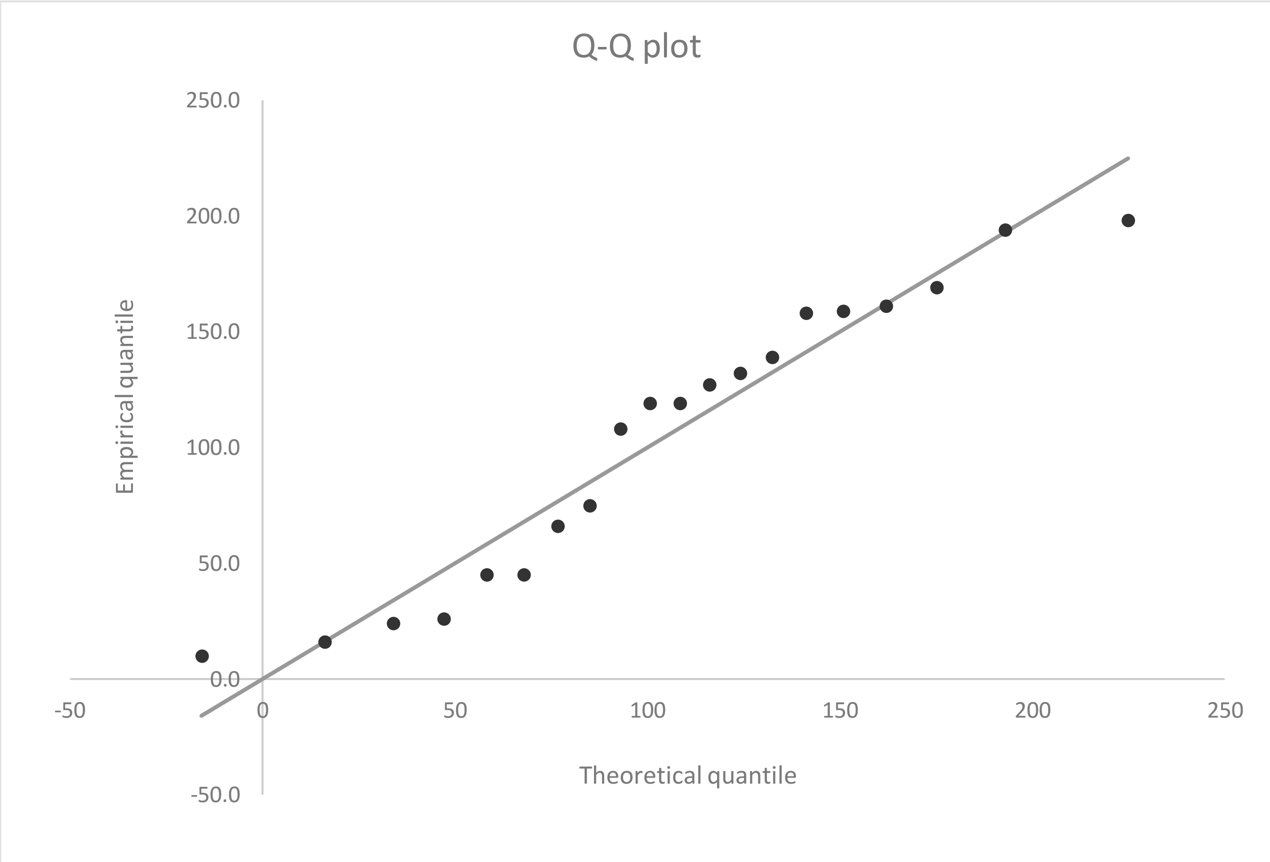

Figure 1: An example of a Q-Q plot

When inspecting a Q-Q plot such as the one shown in Figure 1 , the main features to look for are:

- The distance of low and high value observations from the reference line compared to the mid-value observations. In Figure 1, the observations at the low and high value ends are approximately the same distance from the reference line as the mid-value observations.
- Differences in the distance from the 45-degree line for low-value observations compared to high-value observations. In Figure 1, the two highest value observations lie further from the reference line than the lowest two value observations. This indicates asymmetry in the distribution of observations, but is insufficient to determine that the PDF of the Data Generating Process is asymmetric.
- Clustering of observations, which could indicate the most likely weighted centre of the distribution. There are no signs of clustering in Figure 1, which indicates an even spread of observations along the reference line.

The Q-Q plot is most informative when comparing plots with different theoretical PDFs. Typically, one of the PDFs will exhibit a better match to the observed values than the others. The best matching PDF is then chosen as the PDF to use in predictive modelling or benchmarking.

2.3.2 Q-Q plot versus P-P plot

In addition to the Q-Q plot, it can be helpful to also consider the probability-probability (**P-P**) plot. These plots are similar in intent, with the P-P plot comparing theoretical probabilities drawn from a benchmark

PDF and the empirical probabilities calculated for each observation using the benchmark PDF adjusted for the sample mean and standard deviation.

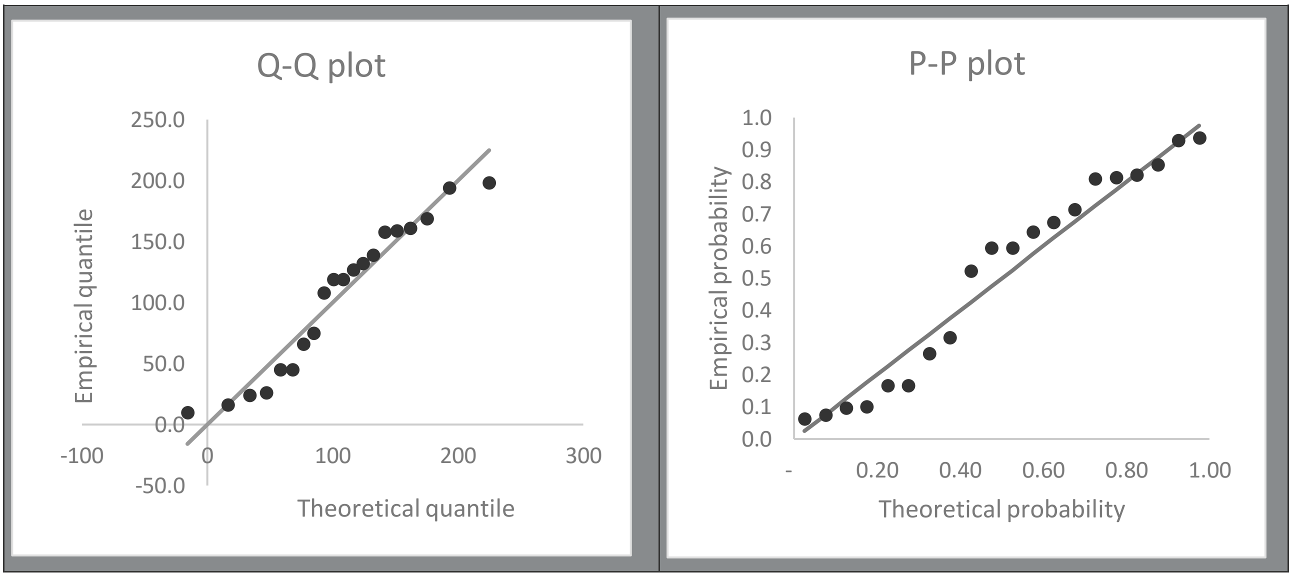

Figure 2: An example of a Q-Q plot and a P-P plot for the same set of observations

The diagnostic information provided by Q-Q and P-P plots are slightly different. Figure 2 presents the Q-Q plot and the P-P plot for the same observations. Comparing the two, it is clearer that the Q-Q plot presents a comparison between the benchmark PDF and the observed data in the same scale as the original data. By contrast, the P-P plot compares the probability of the reference PDF.

The practical difference comes down to the Q-Q plot providing better visibility of the extremes of the distribution whereas the P-P plot provides better visibility of the centre of the distribution.

2.3.2.1 Concluding remarks on visual inspection

The visual inspection of the data using Q-Q and P-P plots were used to confirm that the statistics identified the best fitting plots.

2.3.3 Statistical measures – goodness of fit statistics

In addition to visual methods, there are statistical methods of determining the best fitting PDF. The three main statistics are the:

- Kolmogorov-Smirnov statistic
- Cramer-von Mises statistic
- Anderson-Darling statistic.

These test statistics are the mathematical equivalent of the visual tests.[2] That is, the test statistics compare the empirical distribution with a chosen theoretical distribution, determine the degree of difference and test whether the magnitude of the difference is large enough to reject the chosen theoretical distribution. That is, given $X_1, \cdots, X_n$ identically and independently drawn samples (i.e. observations) from an unknown distribution F, a formal hypothesis test can be established:

[2] This section draws on information provided by Rui Castro 2013: Lectures 2 and 3 – Goodness-of-Fit (GoF) Tests.

$$H_0: F = F_0 \ vs. \ H_1: F \neq F_0$$

Where H_0 represents the null hypothesis, H_1 is the alternative hypothesis, F represents the empirical distribution and F_0 is the hypothesised distribution. According to Castro (2013), these tests are based on the Glivenko-Cantelli theorem, which states

$$\sup_t \left| \hat{F}_n(t) - F_o(t) \right| \xrightarrow{a.s} 0 \text{ as } n \to \infty$$

Where sup indicates the supremum (i.e. the least upper bound), $a.s.$ indicates asymptotic, $\hat{F}_n$ is piece-wise constant and F_0 is ordered so that it is a non-decreasing function. In words, this theorem states that the supremum of the difference between the empirical and theoretical distributions asymptotically approaches zero as the sample size of observations approaches infinity.

This theorem and the tests that are based on it allow for some non-negative differences between the empirical and theoretical distributions. Hence, a threshold critical value needs to be determined which, if exceeded, indicates that the null hypothesis is rejected.

An important feature of all three tests is that the test statistics either do not depend at all on the theoretical distribution (e.g. the Kolmogorov-Smirnov test) or only partly on the theoretical distribution (i.e. the Anderson-Darling test) being tested against the sampled observations. In addition, these tests are consistent under any alternative hypothesis.

2.3.3.1 Kolmogorov-Smirnov statistic and test

The Kolmogorov-Smirnov test is calculated in a piece-wise method by:

- Arranging the observations in ascending order; and then
- Calculating the Empirical Cumulative Density Function (**ECDF**), which is a one-dimensional array containing a series of values derived by dividing the number of observations that are less than each observed value by the total number of observations; and then
- Calculating the maximum distance between the ECDF and the theoretical Cumulative Density Function of a chosen reference PDF.

The Kolmogorov-Smirnov statistic can be calculated for each PDF under consideration. The test is constructed as a hypothesis test:

- The null hypothesis is that the data is generated by the theoretical PDF being tested.
- The alternative hypothesis is that the data is not generated by the theoretical distribution being tested.

The null hypothesis is rejected if the test statistic exceeds a prescribed significance level.

2.3.3.2 Cramer-von Mises statistic and test

The Cramer-von Mises test is like the Kolmogorov-Smirnov test except it replaces the supremum with the product of the squared differences between the empirical and theoretical distributions.

2.3.3.3 Anderson-Darling statistic and test

The Anderson-Darling statistic is another variation of the Kolmogorov-Smirnov test where the difference between the empirical and theoretical distributions on the tails of the distribution receives more importance. Note that the critical values of the Anderson-Darling test are dependent on the theoretical distribution being tested.

According to the literature reviewed, the Anderson-Darling test has higher power than the other two tests. This means that it is better able to reject a false null hypothesis. For this reason, the goodness-of-fit results in this report focus exclusively on the Anderson-Darling test results. Specifically, the p-value of the Anderson-Darling is reported. To accept the candidate distribution as a good fit, the p-value needs to be greater than a critical threshold value.

2.3.3.4 Akaike Information Criterion

The Akaike Information Criterion (**AIC**) is a statistic that provides information about the relative quality of goodness-of-fit. A key difference between the AIC and the other goodness-of-fit statistics is that there is no null hypothesis to be rejected.

The statistic is calculated as follows:

$$AIC_c = 2k - 2ln(\hat{L})$$

Where k represents the number of model parameters and $ln(\hat{L})$ is the natural logarithm of the likelihood score. Note that the likelihood score is derived from maximising the likelihood function. So, there is a clear trade-off between adding more parameters and increasing the likelihood score. That is, for a more complex distribution to be accepted as best fitting, the likelihood score needs to increase (representing a benefit) by more than the cost of adding additional parameters.

This score guards against relying on unnecessarily complex models to determine the best fitting distribution.

2.3.3.5 Using the Anderson-Darling and AIC statistics together

By using both the Anderson-Darling and the AIC statistics, a clear rank can be formed among the candidate distributions for each service standard measure. This is achieved in two steps:

1. Fit each candidate distribution, calculate the Anderson-Darling p-value and reject all distributions registering a p-value less than 0.05.
2. Of those candidate distributions left, choose the distribution that registers the lowest AIC score. In cases where several candidate statistical distributions have AIC scores close to the lowest AIC score, take the average of the resulting performance metrics across statistical distributions.

2.3.3.6 Consideration of the underlying assumptions

Fitting statistical distributions requires adherence to the assumptions underpinning the validity of the chosen statistics. The most important assumptions are:

- The mean and variance of the underlying Data Generating Process are constant;
- That each observation is independent of the other observations contained in the sample data; and
- For a given service standard measure, that each observation is drawn from the same univariate distribution.

These are strong assumptions that could be inadvertently violated. In using real world data, generally accepted practice is to relax these assumptions somewhat. That is, the chosen distribution is regarded as the best approximation of the unknown Data Generating Process.

In some cases (such as a strongly trending mean) relaxing the underlying assumptions could result in gross errors. In such cases, it may be reasonable to adjust the distribution fitting process by:

- Segmenting the sample data and independently fitting the distributions to each segment.
- Pre-whitening the data.

Choosing the appropriate action depends on several factors:

- The overall sample size for each service standard measure – reducing sample size may reduce precision.
- Identifying any underlying causes of a non-stationary mean and/or variance.

Fitting distributions to subsets of the data can be informative. If dividing the data in different ways leads to radically different results, then it can be reasonably inferred that the underlying assumptions are grossly violated. In such cases, additional information needs to be obtained to better understand the Data Generating Process. Once this is obtained, an agreed mitigation strategy to adjust the data can be implemented. Note that no agreed mitigation strategy to adjust the data has been implemented.

2.4 Fitting discrete PDFs

Much of the process for fitting discrete probability distributions is similar to continuous distributions except that the goodness-of-fit statistic is assessed using the chi-square test[3]:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

Where O_i is the observed frequency for bin i and E_i is the estimated frequency. This is distributed with $k - p - 1$ degrees of freedom where p is the number outcomes estimated. Extreme outcomes can lead to unstable p-values; so in our application this process is simulated 2,000 times to compute a stable p-value.

The null hypothesis is that the observations are generated by the chosen distribution. This hypothesis is accepted if the chi-squared statistic is less than the prescribed threshold. This translates to a p-value higher than the p-value of the prescribed threshold.

2.5 Concluding remarks

For the reasons outlined in this section, Western Power has chosen the Anderson-Darling (**AD**) to explicitly reject poor fitting continuous statistical distribution from further consideration while the Chi-squared statistic is used to explicitly reject poor fitting candidate discrete statistical distributions. For the remaining candidate statistical distributions, the AIC is used to determine the best fitting statistical distribution for each performance measure.

[3] https://cran.r-project.org/doc/contrib/Ricci-distributions-en.pdf

3. Application

Key messages:

- The performance data is constructed according to the definitions as agreed for the AA3 period.
- SSBs and SSTs are calculated according to the method described in Section 3.6.
- Consideration is given to determining how much of the historical data should be used to develop appropriate and stable performance measures.
- The main outcome of this benchmarking method are performance benchmarks that are more stringent for AA4 than those applied in the AA3 period.
- Mechanistically adjusting the performance measures risks seriously underestimating the true variance of the Data Generating Process. Setting SSBs using a percentile that is too low could lead to substantially increased expenditure on network reliability to ensure compliance that does not provide a materially improved performance benefit for customers.
- Given this, Western Power has determined that customers would be better served by adopting the 99th percentile as the SSB for all performance measures.

3.1 Introduction

The method used by Western Power and the ERA to determine the SSBs and SSTs for AA3 can broadly be described as follows:

1. Establish the data series relating to historic service standard performance (typically 5 years of 12 month rolling average data).
2. Determine the statistical distribution of best fit.
3. Sample the distribution to obtain the 50th and 97.5th percentiles (or 2.5th percentile where lower value reflects poorer performance), resulting in the SST and SSB accordingly.
4. Where no statistical distribution was found to fit the data appropriately, the historic data was sampled directly.

The purpose of fitting the statistical distribution as opposed to simply sampling the history is to simulate the result should the dataset be larger, and to better model the probabilities in the tail end of the distribution. A similar method is commonly used in other jurisdictions for the equivalent of SST setting.

3.2 Time period for fitting

As described above in section 2.2, this process assumes the history is relatively random and trend free, and that the future will be similar to the past. While true for some metrics, this does not always apply.

In the past, the flaw in this assumption was addressed by using a shorter time series of data than available to fit the distribution, and some use of manual adjustments. Reducing the time period is effective at ensuring the mean of the distribution is more reflective (impacting the SST), however risks underestimating the variance or spread (impacting the SSB).

3.3 Selecting percentiles

Assuming stable performance, the sampling of the 97.5th percentile should indicate a 2.5% probability of exceedance per metric. If the current 17 AA3 SSBs were fully independent, this would result in a 34.98% chance of exceeding at least one per year; effectively necessitating performance improvement to ensure compliance.[4] While the metrics are not fully independent, the impact is still valid.

In AA4, Western Power is proposing network investment to maintain service performance. The proposed network investment aligns closely with customer satisfaction analysis, indicating that customers are satisfied with the current level of performance. As such, Western Power proposes the use of the 99th percentile for setting SSBs. With a 1% probability of exceeding each metric, the total result is a 15.7% probability of exceeding at least one per year. The reduced probability better aligns with the goal of maintaining performance and the proposed investment.

3.4 Distribution selection

In this section, we give details about processing proposed SST and SSB for each of 15 KPIs using the method in section 2. A total of 15 different distributions were chosen (11 continuous and 4 discrete) based on their mathematical simplicity and use in the AA3 approach. The statistical distributions are listed below:

Continuous Distributions		Discrete Distributions
Exponential	3-Parameter Log logistic	Poisson
Gamma	Lognormal	Negative Binomial
3 Parameter Gamma	Normal	Binomial
Generalised Extreme Value	Weibull	Geometric
Logistic	3 Parameter Weibull	
Log Logistic		

Table 1: Candidate statistical distributions

The final statistical distribution for each service standard measure is selected based on the following **criteria**:

- Elimination of poorly fitting distributions. Each candidate distribution is tested to determine whether it should be rejected on the basis of the goodness-of-fit test, i.e., the Anderson-Darling (**AD**) test for continuous distribution and Chi-square test for integer distribution with p-value greater than 0.05. Note that the AD goodness-of-fit test can only be used for continuous distributions such as Normal and Weibull distributions, while the Chi-square test can be used for discrete distributions like Poisson and negative binomial distributions.
- Quality of fit. The distributions that have not been rejected by the relevant goodness-of-fit test are ranked according to the Akaike Information criterion (**AIC**). That is, a good fit is determined as

[4] This is calculated using the formula $P(A_1 \cup A_2 \dots \cup A_n) = 1 - \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_i \cap A_i) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k) - \cdots$,where it is assumed to be no intersection among events, that is, $\sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = 0$

westernpower It's ON

having the lowest AIC. In cases where several statistical distributions have AIC scores close to the lowest, they have also been used to calculate the service standard benchmark and target.

- Stability of the resulting benchmark. Given that reliability data is continuously evolving, there is the risk that small changes in the underlying data would result in the AIC indicating a different distribution. The result can be a radical change in SSB/SST benchmark from one year to the next, which would be contrary to the principles of benchmark setting. To help ensure robustness of the resulting benchmark introduced by small changes to the data, Western Power proposes averaging all distributions considered to be a good fit for a measure based on AIC scores close to the lowest score.

The AIC is a measure of the relative quality of statistical distributions for the data. Given a collection of distributions for the same KPI data, AIC estimates the quality of each model, relative to each of the other distributions. There are two benefits by employing the AIC criterion: (1) discouraging overfit, and (2) overcoming the theoretical difficulty of comparing the AD or Chi-square statistics computed for several distributions fitted on the same KPI data. Hence, AIC provides a means for distribution ranking from a collection of fitted distributions (the smaller the AIC, the better the fit of the distribution).

All statistical distributions are implemented in R[5] with package **fitdistplus**.[6] The selection of distribution(s) are automatically selected by the first two of the above two. A comparison of fitted distribution parameters and associated quantiles are calculated and listed in Tables, and diagnostic plots for the final selected distributions are given per metric.

3.5 Averaging statistical distributions

Often, several distributions fit each metric well. The practical implication of having numerous distributions that fit well, is that a small change in the underlying data can change the selected distribution; sometimes with a large impact on the SSB. In some cases, shifting from the best to the second best distribution of fit can result in a 10-20% change in the proposed SSB.

To demonstrate, the following figure shows the AIC and 99th percentile values for Rural Short SAIDI using Normal and Weibull distributions on a rolling five year-basis. The figure is generated using data back to 2012. The x-axis represents the end-date of the five-year rolling time window. Whichever line has the lowest AIC value would be selected as the distribution of choice.

[5] A copy of the r code can be found at EDM # 43704994

[6] For more information about the fitdistplus package, see: https://cran.r-project.org/web/packages/fitdistrplus/index.html. Note that R packages on Cran are peer reviewed by experts.

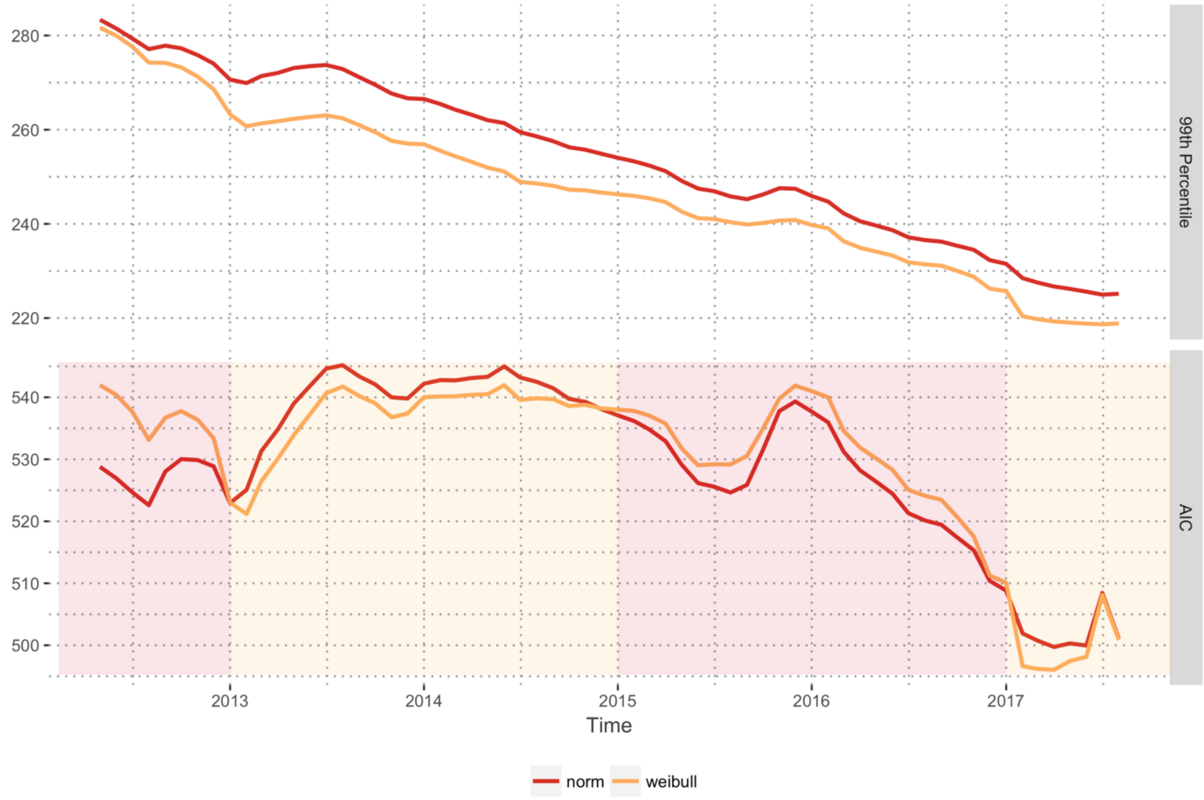

Figure 3: AIC calculated on a 5-year rolling basis for Rural Short SAIDI

There are four transition points over the historic time period in which the lowest AIC could trigger selection between the Normal and Weibull distributions.

To overcome the volatility introduced by small changes to the data, Western Power proposes averaging all distributions considered to be a good fit. Good fit is determined as having an AIC close to the lowest AIC. Four thresholds of 'closeness' were tested with the following results.

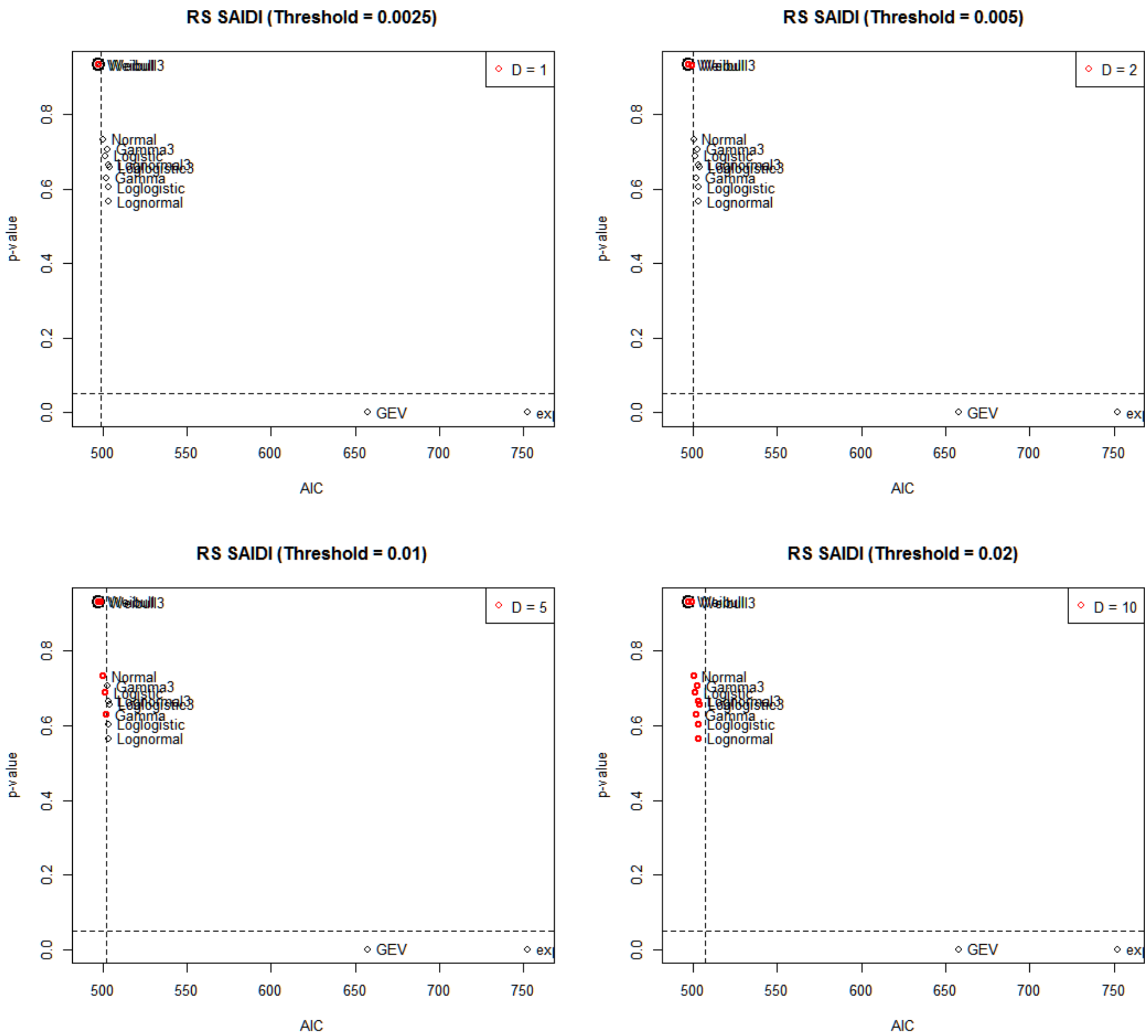

Figure 4 **A visualization of good fit distributions (red dots) for Rural Short SAIDI. D is the number of good fit distributions. A good fit is determined as a fitted distribution pass the goodness-of-fit test and having an AIC within 0.25% (upper left), 0.5% (upper right), 1% (lower left) and 2% (lower right) close to the lowest AIC (circled red dot).**

By way of illustration, Figure 4 demonstrates how close the statistics are across the candidate distributions. Each quadrant reflects a different threshold of closeness (i.e. 0.0025, 0.05, 0.01, 0.02) to the lowest AIC. The red dots indicate the statistical distributions included in averaging the service standard benchmarks and targets.

Count of AIC < Min(AIC)*(1+x) [7]	x=0.0025	x=0.005	x=0.01	x=0.02
CBD SAIFI	1	1	2	3
Urban SAIFI	1	1	1	1
Rural Short SAIFI	1	1	1	1
Rural Long SAIFI	2	3	4	6
CBD SAIDI	1	3	5	7
Urban SAIDI	1	1	6	9
Rural Short SAIDI	1	2	5	10
Rural Long SAIDI	1	2	4	8
Call Centre Performance	1	1	1	1
Average Outage Duration	4	6	9	10
Circuit Availability	1	1	1	1
LOSEF >0.1	1	1	1	1
LOSEF >1	2	2	3	3

Table 2: Measures of closeness for the AIC and number of statistical distributions included

The results indicate that there are some metrics (e.g. Average Outage Duration) that can be fit well with many distributions, while others (e.g. LOSEF >0.1) have one distribution alone that provide a good fit. Based on the intention to stabilise the outcome given small changes to the input data; the 0.01 threshold is proposed.

Note that this sensitivity analysis was conducted before the availability of the May and June 2017 data.

3.6 Summarised AA4 approach

The following proposed approach will be implemented for each service standard measure:

1. Establish the data series relating to historic service standard performance (5 years of 12 month rolling average data);
2. Determine the statistical distribution of best fit;
3. Discard any distributions with an Anderson Darling p-value less than 5%;
4. Sample the distribution to obtain the 50th and 99th percentiles (or 1st percentile where lower value reflects poorer performance);
5. Average the results for all distributions with an AIC within 1% of the distribution with the lowest AIC.

[7] Based on 5 years of proposed metric data, sampled May 2012 to April 2017, inclusive.

3.7 Major Event Days

Western Power is proposing an adjustment to the definition of Major Event Days (**MED**). As such, historic data based on both the current and proposed definition of MED (AA3 MED and Proposed MED) are shown for all distribution metrics.

4. Distribution Network Performance Measures

4.1 CBD SAIDI

CBD SAIDI performance has varied over the last 10 years; a result not well aligned with historic investment. Given the size of the network and that customer numbers in the CBD are low, there remains a probability of highly variable performance. While 10 years of performance is shown, the distributions are fitted to the last 5 years only.

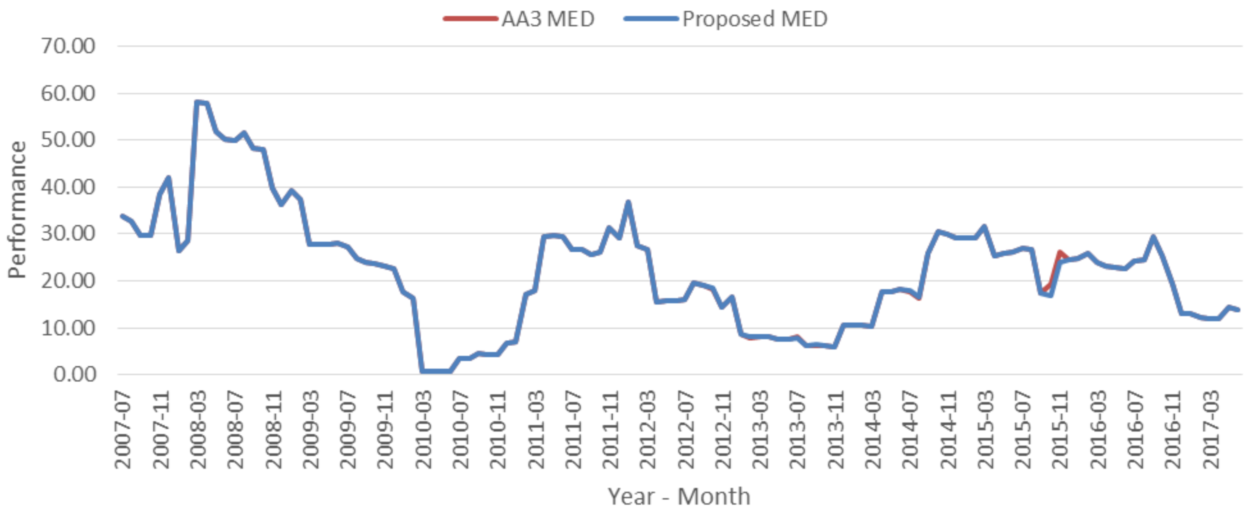

Figure 5: CBD SAIDI

All available distributions were fitted to both sets of data for the period July 2012 – June 2017. The performance statistics are shown below for the statistical distributions fitted to the data with the proposed MED definition.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					6.13	6.28	17.73	30.25	30.88
Average					2.41	4.39	17.82	34.04	37.18
Weibull	1	Yes	0.31	415.06	3.58	5.08	17.85	33.69	36.65
Weibull3	2	Yes	0.25	416.58	5.16	6.14	17.28	35.24	38.96
GEV	3	Yes	0.33	416.84	-1.64	1.76	18.84	30.60	31.76
Normal	4	Yes	0.31	418.23	0.40	3.20	18.17	33.15	35.95
Gamma	5	Yes	0.24	418.76	4.53	5.77	16.94	37.53	42.59
Gamma3	6	No	0.30	420.26	1.12	3.66	18.00	33.63	36.70
Lognormal3	7	No	0.30	420.32	1.26	3.75	17.96	33.77	36.91
Lognormal	8	No	0.15	423.04	5.27	6.29	16.34	42.39	50.66
Logistic	9	No	0.30	424.05	-3.29	1.06	18.18	35.31	39.66
Loglogistic3	10	No	0.28	426.24	-0.76	2.60	17.93	37.24	42.91
Loglogistic	11	No	0.18	427.09	4.50	5.89	17.01	49.09	64.28
exp		No	0.00	469.98	0.18	0.46	12.59	67.03	83.68

Table 3: Performance measures by candidate statistical distribution for CBD SAIDI

Table note: P indicates percentile

The process attempted to fit 12 statistical distributions to the proposed performance data, 5 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome by averaging the results from these five statistical distributions is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	20.30	39.90
AA4 proposed	17.82	37.18

Table 4: Proposed performance measures for CBD SAIDI

4.2 Urban SAIDI

Urban SAIDI performance has been improving for the last 10 years. This improvement aligns with historic investment in asset replacement and network augmentation. While historic investment should maintain current performance, the proposed network investment is not expected to result in continued improvement. While 10 years of performance is shown, the distributions are fitted to the last 5 years only, which has been relatively stable.

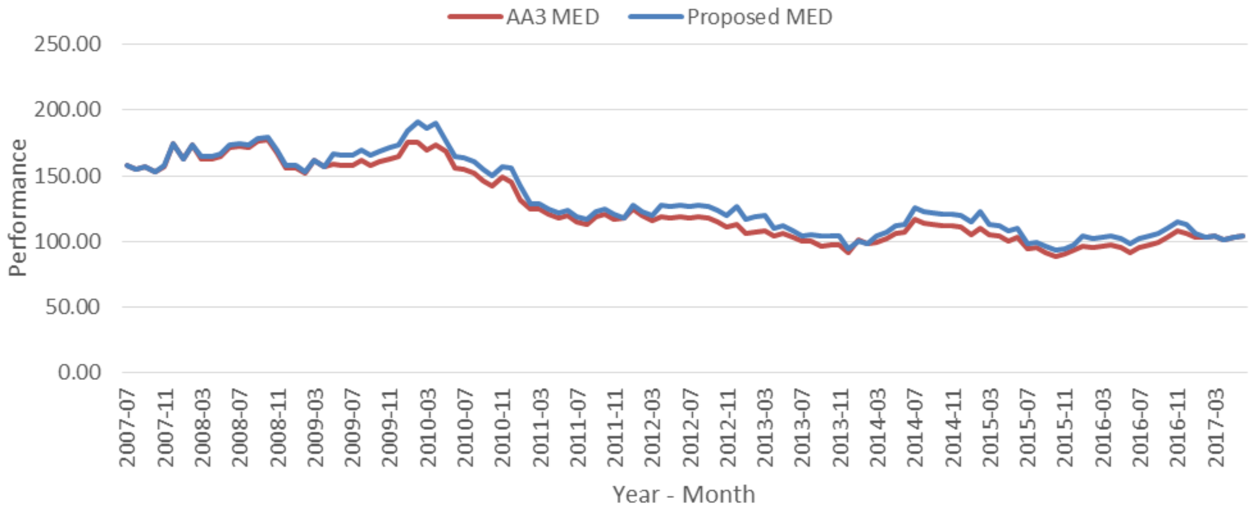

Figure 6: Urban SAIDI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the statistical distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					94.18	94.77	106.78	126.81	127.06
Average					91.03	93.30	108.66	130.00	134.70
Weibull3	1	Yes	0.64	438.52	93.85	94.96	108.25	130.62	135.33
Gamma3	2	Yes	0.64	440.31	93.44	94.88	107.97	132.33	138.32
Lognormal	3	Yes	0.36	441.04	89.54	92.37	109.04	128.73	132.78
Lognormal3	4	Yes	0.59	441.17	92.73	94.52	108.10	132.12	138.33
Gamma	5	Yes	0.33	441.50	88.97	91.96	109.16	128.38	132.20
Normal	6	Yes	0.27	442.72	87.65	91.08	109.43	127.79	131.22
Loglogistic3	7	No	0.57	444.80	92.31	94.33	107.88	138.45	151.07
Loglogistic	8	No	0.34	445.31	86.1	90.25	108.60	130.69	136.99
Logistic	9	No	0.27	446.81	83.27	88.45	108.84	129.22	134.40
Weibull	10	No	0.13	450.22	78.04	84.19	110.47	126.73	129.06
GEV		No	0.00	584.93	-1995.77	-1386.74	8.87	126.63	127.20
exp		No	0.00	685.44	1.10	2.77	75.86	403.70	503.97

Table 5: Performance measures by candidate statistical distribution for Urban SAIDI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 6 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	136.60	183.00
AA4 proposed	108.66	134.70

Table 6: Proposed performance measures for Urban SAIDI

4.3 Rural Short SAIDI

Rural Short SAIDI performance has been improving over the last 10 years. This improvement aligns with historic investment in asset replacement and network augmentation. While historic investment should maintain current performance, the proposed network investment is not expected to result in continued improvement. While 10 years of performance is shown, the distributions are fitted to the last 5 years only, which has been relatively stable.

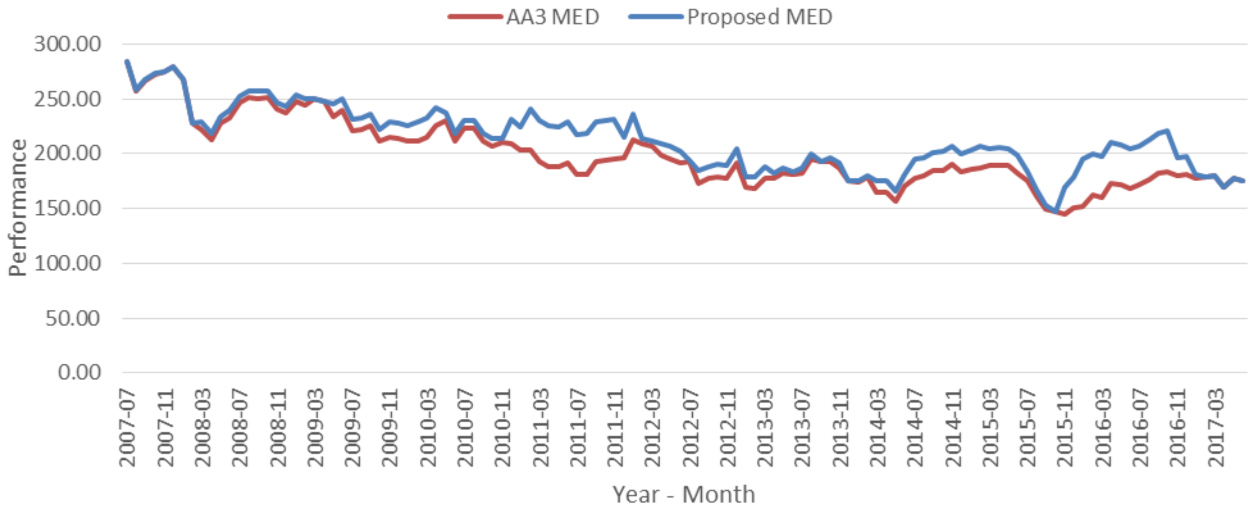

Figure 7: Rural Short SAIDI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017, inclusive. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					150.49	159.10	190.85	216.17	219.74
Average					151.25	158.22	190.44	220.23	226.30
Weibull	1	Yes	0.87	499.41	143.29	152.73	191.94	215.43	218.75
Normal	2	Yes	0.84	500.24	154.87	160.41	190.06	219.71	225.26
Gamma	3	Yes	0.76	501.71	156.25	161.22	189.63	221.19	227.46
Logistic	4	Yes	0.74	501.90	150.35	158.50	190.52	222.54	230.69
Lognormal	5	Yes	0.71	502.65	156.73	161.48	189.43	222.22	228.95
Gamma3	6	Yes	0.81	502.65	155.25	160.62	189.92	220.21	225.99
Lognormal3	7	Yes	0.81	502.77	155.42	160.73	189.88	220.30	226.13
Weibull3	8	Yes	0.78	503.13	135.44	147.86	192.40	214.95	217.94
Loglogistic	9	Yes	0.69	503.15	153.64	160.43	190.21	225.51	235.49
Loglogistic3	10	No	0.72	504.43	151.97	159.45	190.37	223.91	232.88
GEV		No	0.00	658.91	-2174.86	-1558.89	31.34	218.63	220.22
exp		No	0.00	751.67	1.91	4.81	131.73	701.06	875.20

Table7: Performance measures by candidate statistical distribution for Rural Short SAIDI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 9 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	207.80	227.80
AA4 proposed	190.44	226.30

Table 8: Proposed performance measures for Rural Short SAIDI

4.4 Rural Long SAIDI

Rural Long SAIDI performance has been worsening for the last 10 years. This performance decline occurred despite the historic investment in asset replacement and network augmentation. Rural Long performance is not expected to improve and proposed network investment is to maintain performance as per the most recent 5 years performance. While 10 years of performance is shown, the distributions are fitted to the last 5 years only.

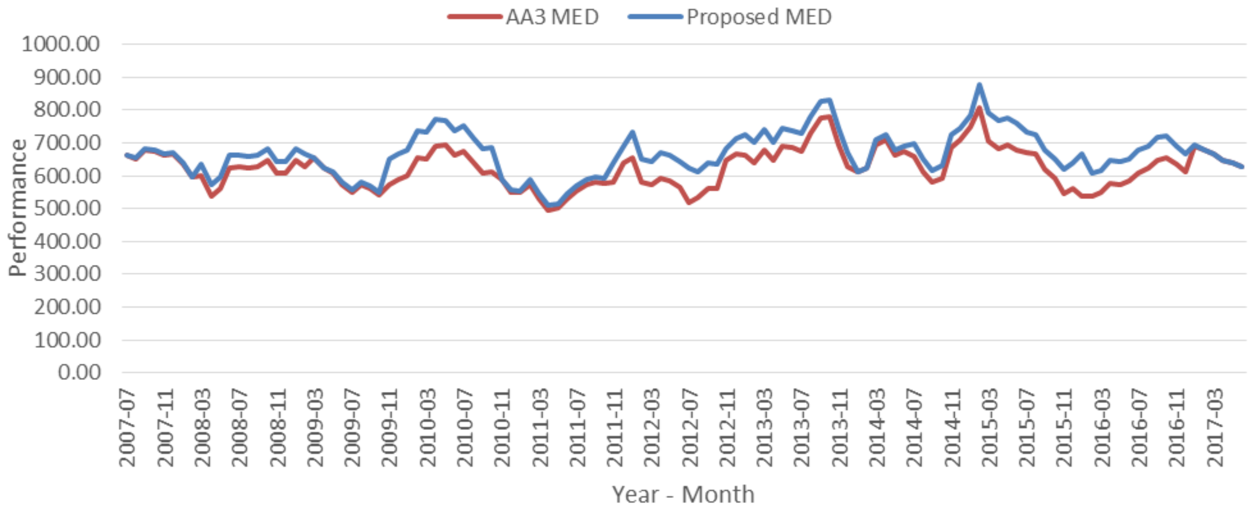

Figure 8: Rural Long SAIDI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					608.72	610.05	690.07	828.47	849.52
Average					604.72	610.26	681.27	855.73	902.93
Weibull3	1	Yes	0.88	655.9	608.43	612.17	681.57	848.28	888.48
Gamma3	2	Yes	0.76	658.12	607.43	611.74	679.33	864.99	916.02
Lognormal3	3	Yes	0.82	660.65	598.30	606.86	682.90	853.91	904.28
Lognormal	4	No	0.78	663.34	568.01	585.98	692.25	817.79	843.67
Gamma	5	No	0.72	664.24	564.64	583.73	693.26	815.72	840.08
Loglogistic3	6	No	0.71	664.39	593.41	604.05	683.83	894.53	989.51
Loglogistic	7	No	0.75	665.56	549.27	575.19	689.54	826.63	865.64
Normal	8	No	0.62	666.41	554.51	576.64	695.01	813.39	835.52
Logistic	9	No	0.68	667.38	532.54	564.63	690.84	817.06	849.15
Weibull	10	No	0.17	679.98	475.04	516.96	699.88	815.61	832.34
GEV		No	0.00	838.56	-12559.00	-8599.98	195.66	875.13	878.00
exp		No	0.00	907.23	6.98	17.59	481.60	2563.02	3199.65

Table 9: Performance measures by candidate statistical distribution for Rural Long SAIDI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 3 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	582.20	724.80
AA4 proposed	681.27	902.93

Table 10: Proposed performance measures for Rural Long SAIDI

4.5 CBD SAIFI

CBD SAIFI performance has been highly variable over the last 10 years (Figure 9), which is not well aligned with historic investment. Given the size of the network and that customer numbers in the CBD are low, there remains a probability of highly variable performance. Proposed network investment for CBD is not expected to improve performance. While 10 years of performance is shown, the distributions are fitted to the last 5 years only.

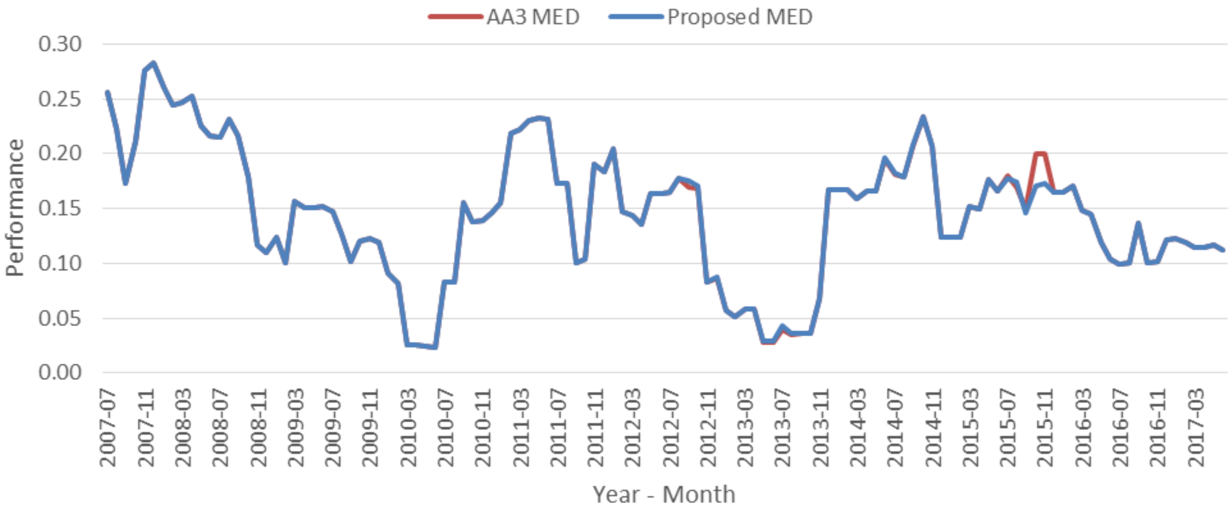

Figure 9: CBD SAIFI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					0.03	0.03	0.14	0.21	0.22
Average					0.00	0.02	0.14	0.22	0.23
Weibull3	1	Yes	0.40	-185.59	-0.01	0.02	0.14	0.21	0.23
GEV	2	Yes	0.38	-185.10	0.00	0.02	0.13	0.22	0.23
Normal	3	No	0.21	-183.60	0.01	0.03	0.13	0.23	0.25
Weibull	4	No	0.12	-183.29	0.03	0.04	0.13	0.23	0.25
Gamma3	5	No	0.19	-180.90	0.01	0.03	0.13	0.23	0.25
Logistic	6	No	0.23	-180.61	0.00	0.02	0.13	0.24	0.27
Loglogistic3	7	No	0.22	-178.19	0.00	0.03	0.13	0.24	0.27
Gamma		No	0.04	-171.97	0.03	0.04	0.12	0.27	0.31
Loglogistic		No	0.04	-164.02	0.03	0.05	0.13	0.35	0.46
Lognormal		No	0.01	-160.88	0.03	0.04	0.12	0.33	0.40
exp		No	0.00	-122.96	0.00	0.00	0.09	0.48	0.60
Lognormal3		No	0.00	-11.70	0.03	0.03	1.03	426.31	1322.16

Table 11: Performance measures by candidate statistical distribution for CBD SAIFI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 2 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	0.14	0.26
AA4 proposed	0.14	0.23

Table 12: Proposed performance measures for CBD SAIFI

4.6 Urban SAIFI

Urban SAIFI performance has been improving consistently for the last 5 years. This improvement aligns with historic investment in asset replacement and network augmentation. While historic investment should maintain current performance, the proposed network investment should not result in continued improvement. While 10 years of performance is shown, the distributions are fitted to the last 5 years only, which has been relatively stable.

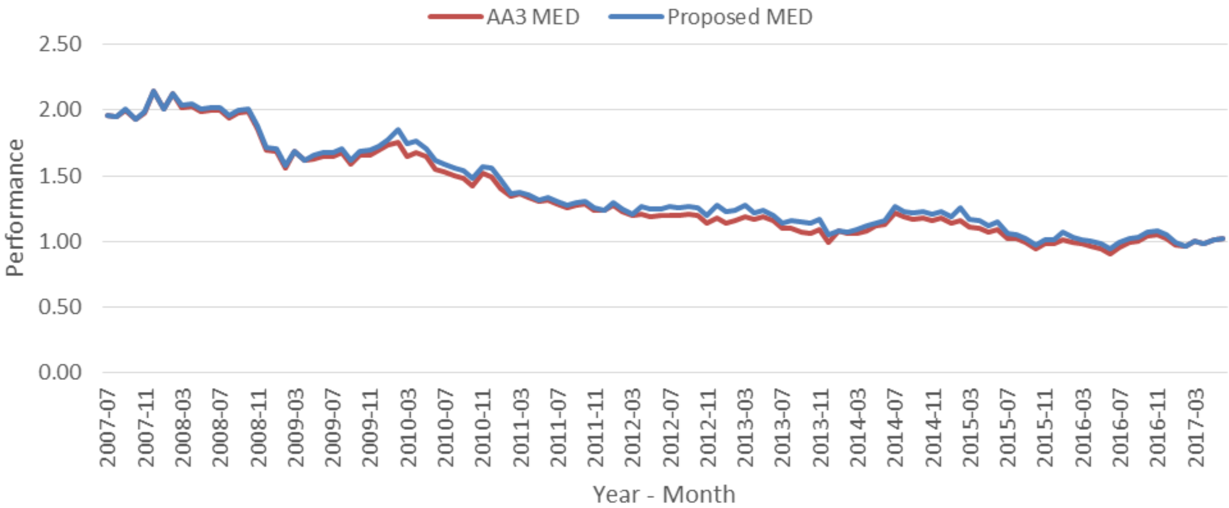

Figure 10: Urban SAIFI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					0.95	0.97	1.13	1.27	1.27
Average					0.89	0.92	1.12	1.3	1.33
GEV	1	Yes	0.12	-105.36	0.84	0.89	1.13	1.27	1.28
Weibull3	2	Yes	0.28	-104.88	0.94	0.96	1.11	1.33	1.38
Gamma	3	No	0.25	-103.29	0.90	0.93	1.12	1.32	1.36
Normal	4	No	0.24	-103.23	0.89	0.93	1.12	1.31	1.35
Lognormal	5	No	0.25	-103.18	0.91	0.94	1.11	1.33	1.37
Weibull	6	No	0.22	-101.79	0.81	0.88	1.13	1.29	1.31
Gamma3	7	No	0.26	-101.40	0.93	0.95	1.11	1.35	1.40
Loglogistic	8	No	0.24	-97.04	0.87	0.91	1.12	1.36	1.43
Logistic	9	No	0.24	-97.02	0.84	0.90	1.12	1.34	1.40
Loglogistic3	10	No	0.24	-95.05	0.86	0.91	1.12	1.36	1.42
Lognormal3		No	0.00	32.84	0.95	0.95	1.94	69.05	150.89
exp		No	0.00	135.51	0.01	0.03	0.78	4.13	5.15

Table 13: Performance measures by candidate statistical distribution for Urban SAIFI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 2 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	1.36	2.12
AA4 proposed	1.12	1.33

Table 14: Proposed performance measures for Urban SAIFI

4.7 Rural Short SAIFI

Rural Short SAIFI performance has been improving for the last 10 years (Figure 11). This improvement aligns with historic investment in asset replacement and network augmentation. While historic investment should maintain current performance, the proposed network investment should not result in continued improvement. While 10 years of performance is shown, the distributions are fitted to the last 5 years, which has been relatively stable.

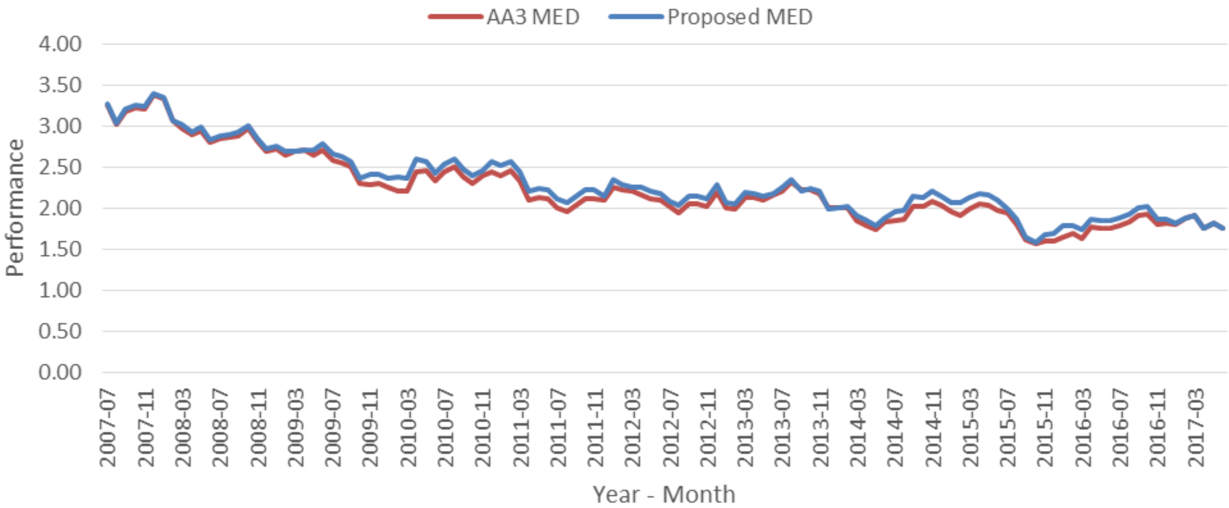

Figure 11: Rural Short SAIFI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					1.63	1.67	2.01	2.28	2.32
Average					1.52	1.60	2.01	2.32	2.38
Normal	1	Yes	0.66	-33.40	1.58	1.65	2.00	2.34	2.41
Weibull	2	Yes	0.72	-33.30	1.45	1.56	2.02	2.30	2.34
Weibull3	3	No	0.73	-33.06	1.59	1.65	2.00	2.32	2.37
Gamma	4	No	0.62	-32.64	1.60	1.66	1.99	2.36	2.43
Lognormal	5	No	0.59	-32.09	1.61	1.67	1.99	2.37	2.45
Lognormal3	6	No	0.64	-31.17	1.59	1.65	1.99	2.35	2.42
Gamma3	7	No	0.62	-30.75	1.60	1.66	1.99	2.36	2.43
Logistic	8	No	0.54	-29.22	1.51	1.61	2.00	2.39	2.49
Loglogistic	9	No	0.51	-28.37	1.56	1.64	1.99	2.43	2.55
Loglogistic3	10	No	0.54	-27.10	1.52	1.62	2.00	2.39	2.50
GEV		No	0.00	12.22	-0.16	0.47	2.14	2.35	2.36
exp		No	0.00	204.88	0.02	0.05	1.38	7.36	9.19

Table 15: Performance measures by candidate statistical distribution for Rural Short SAIFI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 2 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	2.27	2.61
AA4 proposed	2.01	2.38

Table 16: Proposed performance measures for Rural Short SAIFI

4.8 Rural Long SAIFI

Rural Long SAIFI performance has been varied for the last 5 years. This performance does not correlate closely with historic investment in asset replacement and network augmentation. While 10 years of performance history is shown, only the last 5 years has been used in distribution fitting.

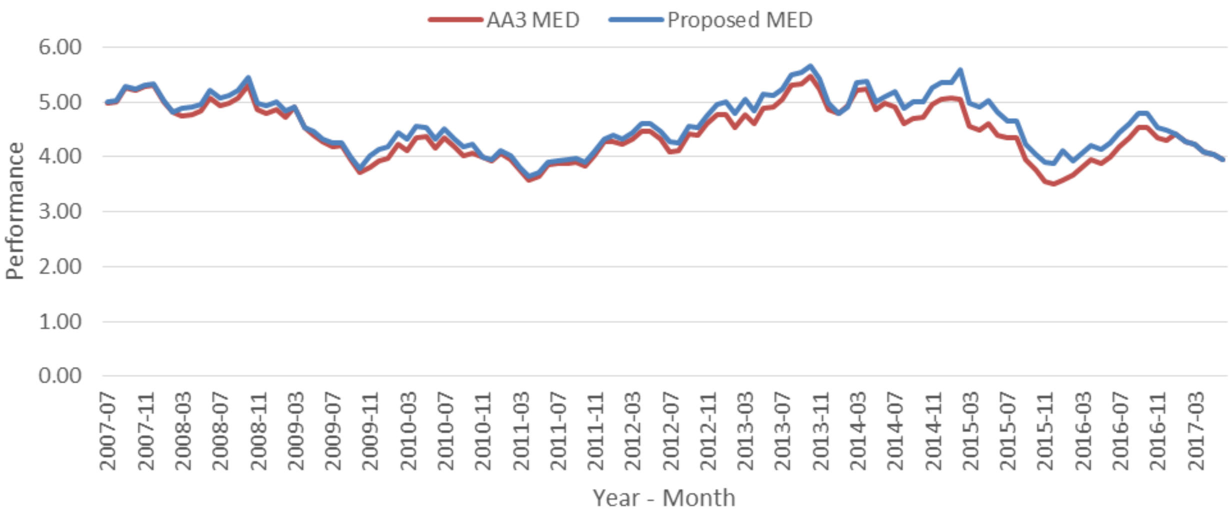

Figure 12: Rural Long SAIFI

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					3.89	3.91	4.80	5.56	5.62
Average					3.69	3.84	4.73	5.71	5.90
Weibull3	1	Yes	0.55	86.89	3.78	3.89	4.73	5.70	5.88
Normal	2	Yes	0.60	87.08	3.62	3.79	4.74	5.69	5.86
Gamma	3	Yes	0.53	87.50	3.68	3.83	4.72	5.74	5.95
Lognormal	4	No	0.48	87.91	3.71	3.85	4.72	5.77	6.00
Weibull	5	No	0.72	88.01	3.28	3.56	4.80	5.57	5.69
Gamma3	6	No	0.58	89.22	3.64	3.81	4.73	5.71	5.90
Lognormal3	7	No	0.57	89.25	3.64	3.81	4.73	5.71	5.90
Logistic	8	No	0.54	91.70	3.42	3.69	4.75	5.81	6.08
Loglogistic	9	No	0.45	92.47	3.56	3.77	4.73	5.94	6.29
Loglogistic3	10	No	0.52	93.83	3.45	3.71	4.75	5.84	6.12
GEV		No	0.00	169.88	-34.1	-22.43	3.60	5.64	5.65
exp		No	0.00	308.73	0.05	0.12	3.29	17.49	21.83

Table 17: Performance measures by candidate statistical distribution for Rural Long SAIFI

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 3 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	4.06	4.51
AA4 proposed	4.73	5.90

Table 18: Proposed performance measures for Rural Long SAIFI

4.9 Call Centre Performance

Call Centre Performance has varied over the last 10 years. The result is directly related to the efficiency of answering calls, and the size of the proposed work program. Although 10 years of data is shown, only 5 years of data is included in distribution fitting.

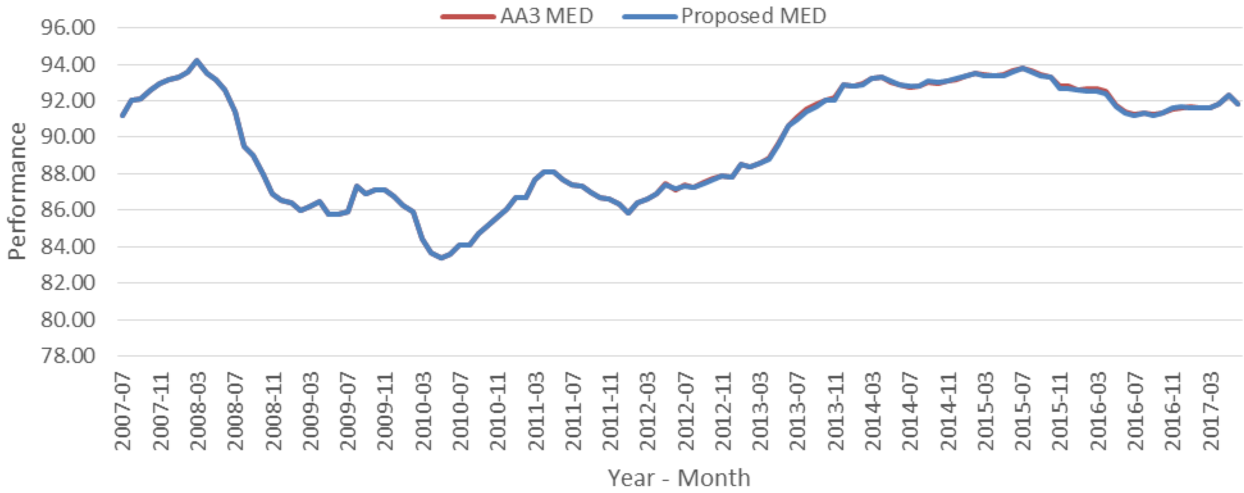

Figure 13: Call Centre Performance

All available distributions were fitted to both sets of data for the period July 2012 - June 2017. The performance statistics are shown below for the distributions fitted to the proposed MED history.

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					87.33	87.43	92.15	93.60	93.68
Average					85.26	86.76	92.17	93.72	93.77
GEV	1	Yes	0.59	216.11	85.26	86.76	92.17	93.72	93.77
Weibull	2	No	0.11	228.59	86.85	87.95	91.99	94.11	94.39
Weibull3	3	No	0.12	229.97	86.76	87.90	92.01	94.09	94.36
Logistic		No	0.04	247.14	87.32	88.26	91.95	95.65	96.59
Loglogistic		No	0.03	248.26	87.38	88.28	91.94	95.76	96.75
Normal		No	0.01	248.73	87.33	88.01	91.66	95.30	95.98
Gamma		No	0.01	249.65	87.36	88.02	91.65	95.37	96.08
Loglogistic3		No	0.03	249.99	87.36	88.28	91.95	95.73	96.71
Lognormal		No	0.01	250.13	87.36	88.02	91.64	95.40	96.12
Lognormal3		No	0.01	251.45	87.35	88.02	91.65	95.35	96.06
Gamma3		No	0.01	251.67	87.34	88.01	91.64	95.38	96.09
exp		No	0.00	664.17	0.92	2.32	63.53	338.11	422.10

Table 19: Performance measures by candidate statistical distribution for Call Centre Performance

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, only 1 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	87.60%	77.50%
AA4 proposed	92.17%	85.26%

Table 20: Proposed performance measures for Call Centre Performance

5. Transmission Network Performance Measures

5.1 Circuit Availability

Circuit Availability performance has improved over the last 10 years. As circuit availability includes planned work, the result is directly related to the size of the proposed work program; and a negative correlation can be seen with historic capacity expansion spend. Given the size of the work plan is similar to the past, performance should hold constant. Although 10 years of data is shown, only 5 years of data is included in distribution fitting.

Figure 14: Circuit Availability

All available distributions were fitted to the data for the period July 2012 - June 2017. The performance statistics are shown below:

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Emprical					98.03	98.03	98.52	98.90	98.90
Average					97.63	97.79	98.51	98.88	98.91
GEV	1	Yes	0.23	17.86	97.63	97.79	98.51	98.88	98.91
Weibull	2	No	0.21	21.73	97.50	97.72	98.51	98.92	98.97
Normal	3	No	0.12	22.59	97.81	97.91	98.46	99.02	99.12
Gamma	4	No	0.12	22.61	97.80	97.91	98.46	99.02	99.12
Lognormal	5	No	0.11	22.62	97.81	97.91	98.46	99.02	99.12
Weibull3	6	No	0.21	23.76	97.49	97.72	98.51	98.92	98.97
Lognormal3	7	No	0.11	24.62	97.81	97.91	98.46	99.02	99.12
Logistic	8	No	0.13	28.75	97.68	97.84	98.47	99.11	99.27
Loglogistic	9	No	0.13	28.78	97.68	97.84	98.47	99.11	99.27
Loglogistic3	10	No	0.13	30.78	97.68	97.84	98.47	99.11	99.27
Gamma3		No	0.02	26.49	98.03	98.04	98.35	99.53	99.88
exp		No	0.00	672.76	0.99	2.49	68.25	363.22	453.44

Table 21: Performance measures by candidate statistical distribution for Circuit Availability

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, only 1 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	98.10%	97.70%
AA4 proposed	98.51%	97.63%

Table 22: Proposed performance measures for Circuit Availability

5.2 Average Outage Duration

Average Outage Duration (**AOD**) performance has been declining for the last 10 years (Figure 15). Actual performance will be affected by work practices, availability of spares/parts, and the type of network faults that occur, and may therefore be highly variable. No material change to performance is expected. Although 10 years of data is shown, only 5 years of data is included in distribution fitting.

Figure 15: Average Outage Duration

All available distributions were fitted to the data for the period July 2012 - June 2017. The performance statistics are shown below:

Distribution	Rank	Included in Average	AD p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					643.44	650.49	870.57	1167.10	1225.17
Average					580.89	622.18	871.37	1236.31	1332.83
Weibull3	1	Yes	0.94	771.64	640.61	657.07	865.79	1234.01	1312.88
Lognormal	2	Yes	0.81	773.11	591.30	628.87	874.36	1215.67	1292.92
Gamma3	3	Yes	0.93	773.48	631.77	654.27	862.49	1258.13	1356.16
Gamma	4	Yes	0.73	773.81	576.04	618.04	878.66	1203.97	1272.38
Lognormal3	5	Yes	0.91	774.34	618.88	647.48	865.33	1251.80	1351.99
Normal	6	Yes	0.48	776.49	538.36	592.85	884.33	1175.81	1230.30
Loglogistic	7	Yes	0.77	776.65	551.92	605.35	870.62	1252.13	1373.34
Loglogistic3	8	Yes	0.88	777.61	607.18	641.64	863.79	1335.43	1523.55
Logistic	9	Yes	0.52	779.35	471.91	554.02	876.95	1199.88	1281.99
Weibull	10	No	0.27	781.69	452.40	525.39	897.83	1176.90	1219.94
exp		No	0.00	936.53	8.91	22.45	614.76	3271.69	4084.36
GEV		No	0.00	947.61	-28271.00	-20301.00	-750.81	1246.51	1260.00

Table 23: Performance measures by candidate statistical distribution for Average Outage Duration

Table note: P indicates percentile

The process attempted to fit 12 distributions to the proposed performance data, 9 of which had both an Anderson Darling p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	698.00	886.00
AA4 proposed	871.37	1332.83

Table 24: Proposed performance measures for Average Outage Duration

5.3 Loss of Supply Event Frequency (>0.1 System Minutes Interrupted to ≤1.0 System Minutes Interrupted)

Loss of Supply Event Frequency > 0.1 and ≤1.0 system minutes interrupted (**LOSEF>0.1**) performance has improved over the period although the change is not material. No performance improvement is expected. Although 10 years of data is shown, only 5 years of data is included in distribution fitting.

Figure 16: Loss of Supply Event Frequency (>0.1 System Minutes Interrupted to ≤1.0 System Minutes Interrupted)

All available discrete distributions were fitted to the data for the period July 2012 - June 2017. The performance statistics are shown below:

Distribution	Rank	Included in Average	CS p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					11	11.47	17.00	23.00	23.41
Average					9.00	10.00	17.00	25.00	27.00
Binomial	1	Yes	0.42	307.03	9.00	10.00	17.00	25.00	27.00
Poisson	2	No	0.17	312.46	9.00	10.00	17.00	26.00	28.00
Negative binomial	3	No	0.17	314.46	9.00	10.00	17.00	26.00	28.00
Geometric		No	0.00	468.38	0.00	0.00	12.00	66.00	82.00

Table 25: Performance measures by candidate statistical distribution for Loss of Supply Event Frequency (>0.1 System Minutes Interrupted to ≤1.0 System Minutes Interrupted)

Table note: P indicates percentile

The process attempted to fit 4 distributions to the proposed performance data, only 1 of which had both a Chi Square p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	24.00	33.00
AA4 proposed	17.00	27.00

Table 26: Proposed performance measures for Loss of Supply Event Frequency (>0.1 System Minutes Interrupted to ≤1.0 System Minutes Interrupted)

5.4 Loss of Supply Event Frequency (>1 System Minute)

Loss of Supply Event Frequency > 1 (**LOSEF>1**) is a measure of the events resulting in customer lost supply of a very large size. Given the events are very rare, performance has been volatile over the 10 year period. No performance improvement is expected. Although 10 years of data is shown, only 5 years of data is included in distribution fitting.

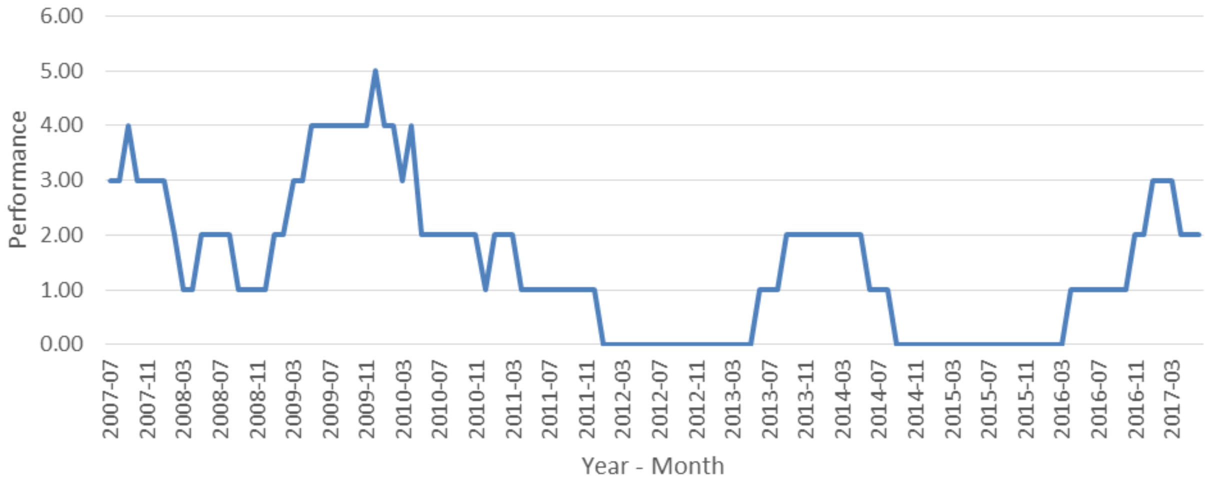

Figure 17: Loss of Supply Event Frequency (more than 1 minute)

All available discrete distributions were fitted to the data for the period July 2012 - June 2017. The performance statistics are shown below:

Distribution	Rank	Included in Average	CS p-value	AIC	p=0.01	p=0.025	p=0.5	p=0.975	p=0.99
Historical					0.00	0.00	0.50	3.00	3.00
Average					0.00	0.00	1.00	3.00	4.00
Poisson	1	Yes	0.07	150.39	0.00	0.00	1.00	3.00	4.00
Binomial	2	Yes	0.07	150.44	0.00	0.00	1.00	3.00	4.00
Negative binomial	3	No	0.09	152.07	0.00	0.00	1.00	3.00	4.00
Geometric		No	0.02	153.58	0.00	0.00	0.00	4.00	5.00

Table 27: Performance measures by candidate statistical distribution for Loss of Supply Event Frequency (more than 1.0 System Minutes Interrupted)

Table note: P indicates percentile

The process attempted to fit 4 distributions to the proposed performance data, 2 of which had both a Chi Square p-value above 0.05 and an AIC within 1% of the distribution with the lowest AIC. The outcome is shown below:

Scenario	Service Standard Target (SST)	Service Standard Benchmark (SSB)
AA3	2.00	4.00
AA4 proposed	1.00	4.00

Table 28: Proposed performance measures for Loss of Supply Event Frequency (more than 1.0 System Minutes Interrupted)

6. References

Castro, R 2013: Lectures 2 and 3 – Goodness-of-Fit (**GoF**) Tests; URL: https://pdfs.semanticscholar.org/1b85/f2b451b8bbcbc9e5cfbfe317b0fd2464b2e9.pdf

Deligenett-Muller, M. L. and C. Dutang, 2015: fitdistplus: An R Package for Fitting Distributions. Journal of Statistical Software. Vol 64, Issue 4, 1-34.

Western Power Report, 2015: Reliability KPIs-Setting the Service Standard Benchmark (**SSB**) and Service Standard Target (**SST**).