

Comments on Chapter 7 and Appendix C of the Commerce Commission
Draft Decisions Paper

**Initial Reset of the Default Price-Quality Path for Electricity
Distribution Businesses (8 September 2009)**

by

Peter Thomson
Statistics Research Associates Ltd
Wellington, New Zealand

8 October 2009
Revised 12 October 2009

Summary of recommendations

This commentary is focussed exclusively on Chapter 7 and Appendix C (Quality Standards) of the New Zealand Commerce Commission draft decisions paper ComCom (2009b). It builds directly on, and should be read in conjunction with, Thomson (2009), a previous submission to the Commerce Commission. It is informed by the results of research undertaken by Statistics Research Associates on the quality of supply by New Zealand electricity distribution networks. In addition to other more general comments, it makes the following specific recommendations.

Recommendation 1 *With regard to the normalisation of daily SAIDI and SAIFI values, it is recommended that:*

- (a) zero SAIDI values be accounted for when determining T_{MED} ;*
- (b) the assumption that non-zero daily SAIDI or SAIFI values follow a log-normal distribution be replaced by the assumption that they follow a mixture of log-normal distributions;*
- (c) normal variation be assessed using truncated daily data (MEDs excluded) in keeping with IEEE standard 1366TM-2003;*
- (d) MEDs are carefully scrutinised, monitored and assessed separately from the assessment of normal variation;*
- (e) the same procedure used to normalise daily SAIDI data be used to normalise daily SAIFI data, leading to MEDs for SAIDI data and MEDs for SAIFI data that are not necessarily coincident.*

Recommendation 2 *With regard to the setting of SAIDI and SAIFI reliability targets, it is recommended that:*

- (a) *a common network-independent false-positive rate α be set in advance;*
- (b) *the reliability targets for each network be given by equation (2) where z_α is the solution to equation (1) and the correlation ρ is estimated from normalised data over the five year reference period. Equations (1) and (2) are given on page 12 of this submission.*

Further details are given in the following sections.

1 Introduction

On 8 September 2009, the New Zealand Commerce Commission (Commission) released a draft decisions paper ComCom (2009b) in which it sets out its draft decisions and current views in relation to the reset of the default price-quality path that expires on 31 March 2010. The Commission has invited submissions on the draft decisions paper which should be received by the Commission before 11am on Monday 12 October 2009.

Given the short time frame, these comments are brief and are focussed exclusively on Chapter 7 and Appendix C (Quality Standards) of ComCom (2009b). They build directly on Thomson (2009), a previous submission to the Commission on the discussion paper ComCom (2009a).

This commentary should be read in conjunction with Thomson (2009) and is informed by the results, now largely completed, of ongoing research on the quality of supply by New Zealand electricity distribution networks. This research, undertaken by Statistics Research Associates (SRA) and commissioned by the New Zealand Electricity Networks Association (ENA), is concerned with establishing a robust statistical framework for measuring performance and setting performance targets for quality of supply. Much of the focus of this research has been on the statistical properties of the daily SAIDI (System Average Interruption Duration Index) and SAIFI (System Average Interruption Frequency Index) reliability indices for New Zealand electricity distribution networks. The Commission was provided with a progress report on this research on 8 April 2009.

Recommendations were made in Thomson (2009) concerning normalisation, dead-bands and assessing compliance for network reliability. It would seem that the Commission have largely taken up many of these recommendations, but not all. The major points of difference concern the normalisation of daily SAIDI and SAIFI values, and the procedures laid down by the Commission for assessing compliance.

ComCom (2009b) further clarifies and extends the methodology proposed in ComCom (2009a). This has created additional issues and highlighted philosophical differences between the approach advocated by the Commission and the IEEE standard 1366TM-2003 (IEEE 2004) on which it is based. Such issues need to be directly addressed and resolved.

Section 2.4 of ComCom (2009b) lists implementation principles that it seeks to use to guide its decision making on the reset of the default price-quality path. These include the following key principles among others.

Consistency Principle (ComCom 2009b, paragraph 2.51). The Commission considers that consistency in decision making is important and that this consistency should apply over all electricity distribution networks and through time. This principle presumably embraces the *fairness principle* embodied in the IEEE standard 1366TM-2003 and Recommendation 2(b) of Thomson (2009).

Transparency Principle (ComCom 2009b, paragraph 2.55). The Commission seeks to use sound analytical foundations to ensure robust decisions are made. This principle presumably embraces the statistical methodology on which the decisions are based, and the statistical properties of the New Zealand reliability data assessed.

The comments and recommendations made in the following sections are intended to assist the Commission develop assessment and monitoring procedures that have a sound statistical basis, are relevant for New Zealand reliability data, and comply with the Consistency and Transparency principles, at least to a safe and reasonable level. SRA is willing to meet with the Commission to provide further information on this submission and the research it is based on, should that be of assistance.

2 Normalisation

The Commission acknowledges that rare events such as extreme weather can lead to daily SAIDI and SAIFI values that are sufficiently large and unrepresentative of a network's normal day-to-day operations that they can significantly distort the average of daily values over any assessment year. The Commission's draft decision on extreme variation (ComCom 2009b, paragraph 7.52) is that the reference and annual assessment data be normalised to take account of such extreme events, where normalisation is defined in paragraphs 7.34–7.51 and Appendix C of ComCom (2009b).

2.1 Zero SAIDI days

Thomson (2009) noted that many research papers and IEEE Working Group reports (Christie 2003a, 2003b, 2003c, Warren and Saint 2005, for example) have addressed the issue of zero SAIDI days, but so far without leading to a change in IEEE standard 1366TM-2003 (IEEE 2004) which excludes zero SAIDI days in the calculation of the threshold T_{MED} . In essence, the IEEE Beta Method assumes that the number of zero SAIDI days has little impact on the calculation of T_{MED} and, more importantly, the annual totals of normal daily SAIDI values where MEDs (Major Event Days) are excluded.

In ComCom (2009b), paragraph 7.37, the Commission quotes results from Warren and Saint (2005) relating to zero SAIDI days and selected US data sets, but not the background literature leading up to that magazine article. In terms of the statistical methodology and analyses adopted, this literature appears to be reliant, for the most part, on work undertaken by Christie (2003a, 2003b, 2003c). Christie (2003a), a peer refereed journal article, is the most authoritative and reliable reference which correctly accounts for zero SAIDI days. Christie (2003b, 2003c) are working papers available from the IEEE Working

Group’s website which have not, it would seem, been peer reviewed by any journal. Christie (2003b) advocates replacing zero SAIDI values by the minimum SAIDI value in the data, whereas Christie (2003c) advocates a maximum likelihood estimation procedure for censored log-normal data or, failing that, replacing zero SAIDI values by the minimum SAIDI value as before. Earlier drafts of the IEEE standard (IEEE 2004) also advocated replacing zero SAIDI values by the minimum SAIDI value.

From the discussion given in these papers and reports, it would appear that large US utilities rarely experience zero SAIDI days even over periods as long as five years (Christie 2003b) and, judging from the examples given, those that do experience zero SAIDI days have relatively few (up to 30%). Note also the comment in Warren and Saint (2005) that “Unfortunately, not many small utilities such as cooperatives were included in the initial analysis”. However it is just such utilities that are most likely to experience zero SAIDI days and which might have a bearing on the New Zealand experience. In short, the US literature on the impact of zero SAIDI days seems confused, deficient, unreliable and far from authoritative, perhaps because it is not a major issue for most US utilities. With the exception of Christie (2003a), this literature is of debatable value in terms of providing guidance in the New Zealand context where zero SAIDI days and small utilities are more the norm than the exception.

Table 1: Summary of the proportions \hat{p}_0 of zero SAIDI days for 17 New Zealand electricity distribution networks over the period 1 April 2004 to 31 March 2009.

Minimum	Lower Quartile	Median	Upper Quartile	Maximum
0.015	0.202	0.467	0.595	0.700

What is the New Zealand experience with regard to zero SAIDI days? Table 1 provides summary statistics for the proportions \hat{p}_0 of zero SAIDI days experienced by 17 New Zealand electricity distribution networks over the proposed five year reference period 1 April 2004 to 31 March 2009. They show a wide variation across networks. The typical \hat{p}_0 was approximately 47% zero SAIDI days, the lowest \hat{p}_0 was 1.5% and the highest 70%, and approximately half the networks had values of \hat{p}_0 that ranged between 20% and 60%.

Zero SAIDI days are evidently a major feature of New Zealand reliability data. It is shown below that accounting for zero SAIDI days significantly reduces the SAIDI and SAIFI thresholds proposed by the Commission for almost all the 17 New Zealand electricity distribution networks considered by SRA. Yet the reliability assessment procedures proposed by the Commission ignore this excellent record of perfect reliability and, instead, focus only on those days where outages occur. This would lead, for example, to networks with the same non-zero SAIDI day performance, but widely differing proportions of zero SAIDI days, being treated alike in contravention of the Commission’s Consistency Principle, the IEEE fairness principle, and customer expectations.

These considerations show that it is both necessary and desirable to include zero SAIDI days in any assessment of the reliability of New Zealand electricity distribution networks. If this is not done, then it is not hard to conceive of operational strategies being developed by individual networks which would ensure that they did have predominantly non-zero SAIDI days, like their US counterparts.

2.2 Distribution of non-zero daily SAIDI and SAIFI data

The Commission notes the SRA finding that Auckland and Wellington non-zero daily SAIDI values are inadequately approximated by the log-normal distribution. They comment (ComCom 2009b, paragraph 7.39) that this finding “appears to have been based on a small subset of EDB data”. However the findings for Auckland and Wellington are replicated across almost all of the 17 electricity distribution networks considered for both SAIDI and SAIFI data.

Table 2: Summary of descriptive statistics for the logarithms of non-zero daily SAIDI ($\log D_+(t)$ with $D_+(t)$ measured in minutes), and the logarithms of non-zero daily SAIFI ($\log F_+(t)$) for data from 17 New Zealand electricity distribution networks over the period 1 April 2004 to 31 March 2009. For each network, the descriptive statistics used are the mean, standard deviation, Bowley coefficient of skewness, and p-value of the Shapiro-Wilk test for normality.

	$\log D_+(t)$				$\log F_+(t)$			
	Mean	SD	Skew	p-value (%)	Mean	SD	Skew	p-value (%)
Minimum	-2.974	1.405	-0.200	3.6×10^{-24}	-7.577	1.221	-0.295	9.6×10^{-27}
Lower quartile	-2.010	1.595	-0.163	8.5×10^{-12}	-6.636	1.788	-0.104	9.8×10^{-15}
Median	-1.614	1.772	-0.076	1.3×10^{-8}	-5.865	1.916	-0.045	1.2×10^{-10}
Upper quartile	-1.053	1.911	-0.042	0.104	-5.549	2.007	0.013	2.2×10^{-6}
Maximum	-0.676	2.101	0.008	29.360	-5.176	2.354	0.161	9.874

Denote the non-zero daily SAIDI and SAIFI values by $D_+(t)$ and $F_+(t)$ respectively. Table 2 summarises descriptive statistics for the $\log D_+(t)$ and $\log F_+(t)$ data from the 17 networks over the five year period 1 April 2004 to 31 March 2009 (the proposed reference period). The statistics considered are the mean, standard deviation, Bowley coefficient of skewness, and p-value of the Shapiro-Wilk test for normality (see Royston, 1982). The Bowley coefficient of skewness is a robust measure of skewness given by

$$\frac{(q_{0.75} - q_{0.5}) - (q_{0.5} - q_{0.25})}{q_{0.75} - q_{0.25}} = \frac{q_{0.75} - 2q_{0.5} + q_{0.25}}{q_{0.75} - q_{0.25}}$$

where $q_{0.25}$, $q_{0.5}$ and $q_{0.75}$ denote the lower quartile, median and upper quartile of the variable considered. The Bowley coefficient of skewness varies from -1 to +1 with a zero value indicating symmetry, positive values indicating a longer right-hand tail, and negative values indicating a longer left-hand tail.

For both the $\log D_+(t)$ and $\log F_+(t)$ the coefficients of skewness are close to zero (typically around -0.05) and predominantly negative. This suggests that the logarithm has slightly over-transformed the data, although this should help ensure that the right-hand tail of the transformed data is light-tailed as desired. The means vary more than the standard deviations which are typically around 1.8. This suggests that the distribution of the $\log D_+(t)$ for each network overlaps with those from the other networks, with a similar observation holding for the $\log F_+(t)$. The p-values of the Shapiro-Wilk test for normality emphatically reject the hypothesis of normality for almost all of the $\log D_+(t)$ and $\log F_+(t)$ data sets. There is one exception in each case (corresponding to the maximum p-values in Table 3) with the second largest p-values being 0.717% ($\log D_+(t)$) and 0.008% ($\log F_+(t)$). The Shapiro-Wilk statistic has good power properties and is widely used to check for

normality. In practice p-values of less than 5% or 1% are used to indicate non-normality. Here the p-values are typically percentages of the order of 10^{-8} ($\log D_+(t)$) or 10^{-10} ($\log F_+(t)$) which present overwhelming evidence of non-normality.

Furthermore, for all the 17 networks considered by SRA, a mixture of normals fits the logarithms of non-zero daily SAIDI and SAIFI data very well, especially in the right-hand tail, with fits comparable to those shown in Thomson (2009) for the Auckland and Wellington SAIDI data. As noted in Thomson (2009), a mixture of two normal distributions provides a very flexible family of distributions that includes the normal distribution as a special case. It can account for bimodality as well as skewness and is widely used in practice to provide accurate, flexible approximations of data distributions. Indeed, for the 17 networks considered, around a third had bimodal distributions over some periods, typically for the logarithms of daily SAIFI, but also the logarithms of daily SAIDI in some cases. The mixture of normals typically fitted these cases very well, unlike the normal which cannot cope with either bimodality or skewness. Furthermore, the model is readily fitted to data, and its component normal distributions often admit useful interpretations.

If the logarithms of positive data follow a mixture of normal distributions, then the original data follow a mixture of log-normal distributions. Given the excellent results obtained for the 17 networks considered by SRA, there is now very little doubt that a mixture of log-normal distributions will provide a much more accurate approximation to the distribution of non-zero daily SAIDI and SAIFI reliability data than the log-normal (a special case of the log-normal mixture). Since the differences are highly significant and, like the failure to account for zero SAIDI days, lead to significant reductions in the SAIDI and SAIFI thresholds proposed by the Commission, adopting the log-normal will again lead to contravention of the Commission’s Consistency Principle and the IEEE fairness principle.

2.3 Threshold estimation

Does the choice of distribution and inclusion or exclusion of zero SAIDI days make a significant difference (practically as well as statistically) to the estimation of the threshold T_{MED} ? If not, then such issues are of little practical importance. However, if they do, then such differences must be dealt with since they are in direct contravention of the Commission’s Consistency Principle and the IEEE fairness principle.

As noted in Thomson (2009), the threshold T_{MED} (B_{SAIDI} and B_{SAIFI} in ComCom 2009b) is the 99.38% quantile of the daily SAIDI or SAIFI data. It can be estimated by the corresponding quantile of the best-fitting log-normal distribution (including or excluding zero SAIDI days) and the best-fitting log-normal mixture, among other possibilities. These model-based quantiles can be benchmarked against the empirical 99.38% quantile which is computed directly from the data as the value, actual or interpolated, which has 99.38% of the daily values below and 0.62% above. The empirical quantile is a non-parametric estimator of T_{MED} which, for large sample sizes, is unbiased with reasonable sampling properties. Here the sample size is 1826 observations (5 years of daily data) and so the empirical quantile would be expected to give good estimates of T_{MED} , although not as

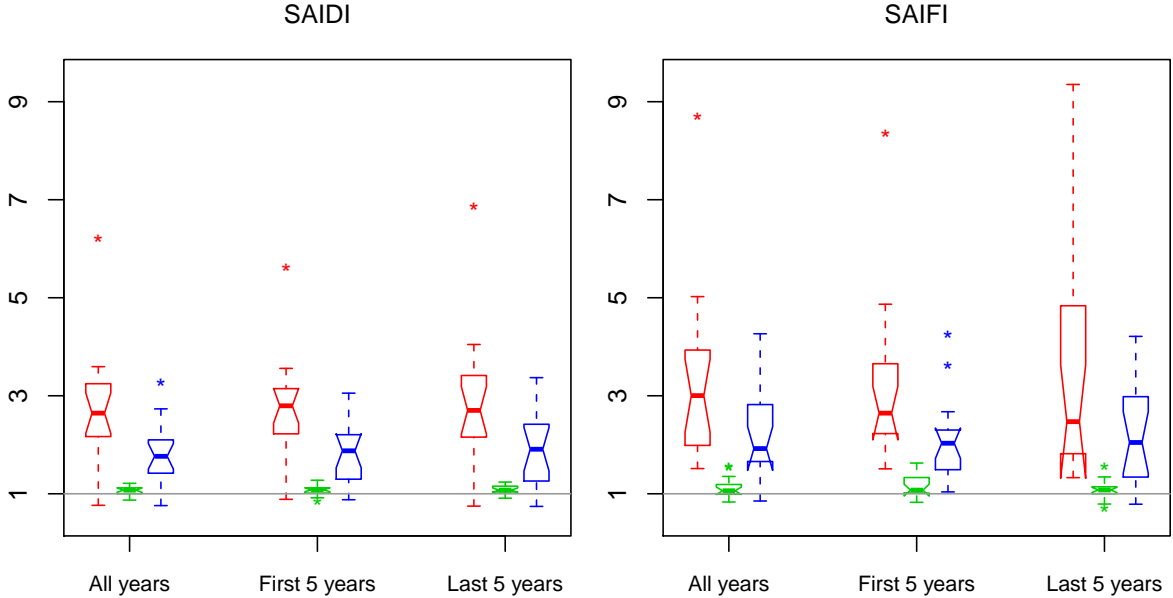


Figure 1: Boxplots of the ratios of estimated T_{MED} values to the empirical 99.38% quantile for daily SAIDI and SAIFI data from 17 New Zealand electricity distribution networks typically for the decade ending 31 March 2009, and the first and last five years of this period. The estimated T_{MED} values are determined from the best-fitting log-normal distribution with zero SAIDI days excluded (red), the best-fitting log normal distribution adjusted for zero SAIDI days (blue), and the best fitting log-normal mixture adjusted for zero SAIDI days (green).

precise as those calculated from the parametric distribution that best fits the data (in this case the log-normal). An advantage of the empirical quantile, particularly for benchmarking, is that it is not subject to model bias since it is determined directly from the empirical distribution function of the data.

The impact on the thresholds T_{MED} of ignoring zero SAIDI days and using the log-normal, rather than the mixture of log-normals, is shown in Figure 1 for the 17 electricity distribution networks considered by SRA. Here the estimates of T_{MED} determined from the best-fitting log-normal (with or without adjusting for zero SAIDI days) and the best-fitting log-normal mixture are shown as a ratios to the empirical 99.38% quantile over three periods. For almost all the networks (14 out of 17) the three periods considered are 1 April 1999 to 31 March 2009 (all years), and the first and last five years of this decade. There were two networks with data from 1 April 2003 to 31 March 2009 and one from 1 April 2001 to 31 March 2009.

Figure 1 shows the boxplots of the various T_{MED} ratios for the 17 networks over the respective periods. In a boxplot, the central box shows the lower quartile (lower line), median (middle line), upper quartile (upper line) and the ends of the vertical notches give the upper and lower limits of a confidence interval for the median (these can occasionally extend beyond the lower or upper quartiles). The whiskers (dashed lines extending above and below the boxes) give the effective range of the data with the isolated points (asterisks) indicating potential outliers. For large samples from a Gaussian distribution, roughly 1% of values would be classified as potential outliers.

Ideally one would like to see T_{MED} ratios close to 1 (boxplots with very little spread concentrated around a median of 1) indicating close agreement with the data determined empirical 99.38% quantile. This is achieved only by the log-normal mixture adjusted for zero SAIDI days (the very small positive bias above 1 is probably an artifact of the way that the empirical quantile has been calculated). The two estimates of T_{MED} based on the log-normal distribution show significantly more variability than the estimates based on the log-normal mixture adjusted for zero SAIDI days (typically 8 times as variable judging from the interquartile ranges). More importantly the two log-normal based estimates of T_{MED} are significantly higher than those based on the log-normal mixture. In each of the three periods the respective medians are significantly different (practically and statistically) with the log-normal estimates ignoring zero SAIDI days being around three times larger than they should be, and the log-normal estimates adjusted for zero SAIDI days being around twice as large as they should be.

Using the log-normal distribution to model non-zero daily SAIDI and SAIFI values with or without the inclusion of zero SAIDI days is no longer an option. It is both necessary and desirable to

- (a) account for zero SAIDI days;
- (b) use a log-normal mixture to model non-zero daily SAIDI and SAIFI values, or use a parametric family of distributions that is at least as flexible.

A less reliable alternative to (b) would be to use the empirical 99.38% empirical quantile of the daily SAIDI and SAIFI data (zero SAIDI days included). Although this estimate is less attractive from the point of view of precision (parametric estimators are usually considerably more precise, especially for extreme quantiles, provided the appropriate distribution is used), it is a much better option than either of the log-normal options considered in Figure 1. It is also the simplest to compute and gives results that are quite close to those based on the log-normal mixture (the best option). Christie (2003a) also explores this approach and notes its advantages and disadvantages by comparison to parameteric estimation.

It is noted that zero SAIDI days are simply accounted for when determining T_{MED} . If \hat{p}_0 is the proportion of zero SAIDI days estimated from the five year reference period, then T_{MED} is estimated as the

$$100\left(1 - \frac{1 - \Phi(2.5)}{1 - \hat{p}_0}\right) = 100\left(1 - \frac{1 - 0.9938}{1 - \hat{p}_0}\right)$$

percentile of the non-zero daily SAIDI or SAIFI data concerned, or its best fitting distribution, provided $\hat{p}_0 < \Phi(2.5) = 0.9938$ where $\Phi(z)$ is the standard normal distribution function. Furthermore, the parameters of a best-fitting mixture of log-normal distributions can be estimated straightforwardly using the method of maximum likelihood, a standard estimation procedure with well-documented statistical optimality properties. Indeed, for data that follow a log-normal distribution, the fitting procedure advocated by the IEEE standard is the same as maximum likelihood. For the log-normal mixture there are simple numerical recursions that can be used to calculate the maximum likelihood estimators (they could be programmed in a spreadsheet for example). It would take little to amend the steps laid out in Appendix C of ComCom (2009b) to incorporate zero SAIDI days and replace the log-normal by the log-normal mixture.

2.4 Censoring versus truncation

As noted in Thomson (2009), the so-called Beta Method described in IEEE (2004) is based on the following key principles.

Classification Each daily SAIDI value is classified as either a *Major Event Day* (MED), during which a network’s reliability is stressed beyond that normally expected, or a *normal day*. A daily SAIDI value falls in one or other category depending on whether it exceeds, or doesn’t exceed, a suitably chosen threshold T_{MED} .

Fairness The Major Event Day (MED) identification threshold T_{MED} is chosen to be an *agreed common percentile* of each network’s daily SAIDI distribution. Here T_{MED} is defined to be the 99.38% percentile of a network’s daily SAIDI distribution. Although the values of T_{MED} will be network dependent, the same proportion of MEDs (0.62%) will be identified for each network, on average, over any given period of time.

At the end of each annual reporting period, a network is expected to fully report on all MEDs to the regulator, but only provide a suitable summary of the normal daily performance (a total annual SAIDI value with MEDs excluded). Note that the Classification Principle is just that; it classifies data as MED or normal variation without modifying the data in any way.

While acknowledging the IEEE criteria, the Commission proposes replacing all MEDs by their boundary values (a process known as censoring in statistics) and justifies this procedure by the following statement (ComCom 2009b, paragraph 7.42).

The Commission considers that scrutiny of extreme events is likely to be a relatively complex and costly exercise, thus a mechanism for undertaking detailed analysis of such events is unlikely to be appropriate under the reset DPP. The Commission recognises that these events should be taken into account in some manner and considers that this can be achieved through boundary substitution.

This admission by the Commission now clarifies the purpose of its proposal to replace MEDs by the boundary values T_{MED} , something that does not appear to have been made explicit in its earlier documents. This is a major methodological and philosophical departure from the IEEE principles and process summarised in the previous paragraph.

Perhaps the best interpretation of the Commission’s approach within the IEEE framework is that the assessment of MEDs and normal variation (two separate assessments) is now replaced by one omnibus assessment of all the data (MEDs included). The latter is a censored mean which is, in this case, a penalised estimate of normal variation. If, for any assessment year, \bar{x}_{normal} denotes the mean of normal daily variation (excluding MEDs) then the censored mean $\bar{x}_{censored}$ is given by

$$\bar{x}_{censored} = (1 - p_{MED})\bar{x}_{normal} + p_{MED}T_{MED}$$

where p_{MED} is the proportion of MEDs in the assessment year considered. Note that this is an estimate, not of normal variation, but of all variation which has been censored at

the threshold T_{MED} . The replacement of MEDs by thresholds T_{MED} raises a number of important issues that would appear to require further discussion and clarification.

The Commission acknowledge (ComCom 2009b, paragraph 7.28) that extreme events such as severe weather can lead to daily SAIDI and SAIFI values that are unrepresentative of a network's underlying service quality, and it considers that quality standards should be structured so as to mitigate the potential for breaches caused by circumstances beyond the control of a network (ComCom 2009b, paragraph 7.27). However it is possible that the replacement of MEDs by T_{MED} may be at odds with these considerations and the Consistency and Transparency Principles advocated by the Commission. Comments and issues to consider include, but are not limited to, the following.

- (a) As noted above, the censored mean $\bar{x}_{censored}$ is a measure of all variation, not just normal variation. To describe such a measure as a measure of normal variation is incorrect since it incorporates a bias term due to the MEDs.
- (b) The use of a penalty such as T_{MED} can be construed as implying that a MED is not only attributable to an extreme event beyond the control of a network, but also in part due to the network itself, with T_{MED} assumed to be a proxy for that part of the MED attributable to the network. However the latter assumption seems dubious since the proportion of a MED's value attributable to the network is more likely to vary systematically according to the nature of the network (rural or urban, topography, structure etc) leading to inequity.
- (c) A similar interpretation can be framed in terms of insurance with T_{MED} playing the role of an excess and the network only liable for daily SAIDI values less than or equal to the threshold T_{MED} . However the premium (relative to T_{MED}) required to cover the excess SAIDI over the threshold T_{MED} is unlikely to be the same across networks. Rather, it will depend on known risk factors such as whether the network is rural or urban, its topography and structure etc.
- (d) As noted in Thomson (2009b), if MED values are replaced by large thresholds, many of the variability issues inherent in the original data will remain in the censored data. Not only will such variability be difficult for networks to manage operationally, it will also lead to additional variability in the reliability measure $\bar{x}_{censored}$ by comparison to the IEEE measure \bar{x}_{normal} . This increases the need for estimates of T_{MED} to be of the highest quality (unbiased, robust and relatively efficient) since variation in these estimates has a direct impact on the reliability of any quality assessment.
- (e) The most worrying implication of the Commerce Commission's approach is that they clearly do not wish to scrutinise and analyse the MEDs in order to resolve their causes, monitor their rate and severity etc. Given that MEDs provide the most extreme stress testing of network reliability, such monitoring and analysis would seem to be extremely important from an industry as well as regulatory perspective.

Unless the Commission can provide firm evidence to the contrary, it would seem that the strategy of replacing MEDs by the threshold T_{MED} is unsafe since it may lead to violations of its Transparency and Consistency Principles, including the IEEE fairness

principle. From a statistical viewpoint, the IEEE approach and the separate assessment of MEDs and normal variation provides the more secure strategy.

Paragraphs 7.49 and 7.50, ComCom (2009b), outline the steps involved in the data normalisation process. Apart from the issue of whether or not to replace MEDs by the threshold T_{MED} which has been discussed above, it would appear that the Commission propose to only normalise daily SAIFI data on the days corresponding to SAIDI MEDs. Although daily SAIDI and SAIFI are highly correlated, there will be instances when a SAIFI MED does not coincide with a SAIDI MED and, as a result, extreme SAIFI values will be retained in the normalised data. Since a breach will occur if either or both of the daily SAIDI or SAIFI values exceed their respective thresholds, this again adds unnecessary and unwarranted variability into the assessment process. There seems to be no obvious statistical or other reason why the daily SAIFI data cannot be normalised in exactly the same way as the SAIDI data, whatever normalisation procedure is adopted.

2.5 Recommendations

Zero SAIDI days and MEDs are highly informative measures of reliability. However the normalisation procedure proposed by the Commission ignores zero SAIDI days (perfect reliability) and fails to take proper account of reliability performance on MEDs (the most extreme stress testing of network reliability). Furthermore, it is based on incorrect distributional assumptions and a quality assessment measure, a censored mean, that is flawed and at odds with the principles embodied in the IEEE standard (IEEE 2004).

These shortcomings, discussed fully in the previous sections, contravene the Commission's Consistency and Transparency Principles. They are addressed in the following recommendation.

Recommendation 1 *With regard to the normalisation of daily SAIDI and SAIFI values, it is recommended that:*

- (a) *zero SAIDI values be accounted for when determining T_{MED} ;*
- (b) *the assumption that non-zero daily SAIDI or SAIFI values follow a log-normal distribution be replaced by the assumption that they follow a mixture of log-normal distributions;*
- (c) *normal variation be assessed using truncated daily data (MEDs excluded) in keeping with IEEE standard 1366TM-2003;*
- (d) *MEDs are carefully scrutinised, monitored and assessed separately from the assessment of normal variation;*
- (e) *the same procedure used to normalise daily SAIDI data be used to normalise daily SAIFI data, leading to MEDs for SAIDI data and MEDs for SAIFI data that are not necessarily coincident.*

3 SAIDI and SAIFI reliability targets

As noted in Thomson (2009), hypothesis testing provides an appropriate statistical framework within which reliability targets (dead-band limits) can be determined. Following the discussion in ComCom (2009a), Thomson (2009) developed appropriate reliability targets for normalised daily SAIDI or SAIFI data for an assessment year. These were of the form

$$\bar{x}_c = \mu + z_\alpha \sigma / \sqrt{n}$$

where n is the sample size (the number of SAIDI or SAIFI values in the assessment year), $\Phi(z_\alpha) = 1 - \alpha$ and $\Phi(z)$ is the standard normal distribution function. The constant z_α is 1.645 when $\alpha = 0.05$, and the mean μ and standard deviation σ of normalised SAIDI and SAIFI data are estimated using analytical formulae derived from the statistical model adopted, or directly using suitable empirical estimates from the five year reference period. These limits give individual false-positive rates of

$$\alpha = P(\bar{X} > \bar{x}_c | \text{network is compliant})$$

where \bar{X} denotes the sample mean of either the SAIDI or SAIFI normalised data for the assessment year. In practice, α is specified in advance and the same for all networks with z_α chosen to satisfy $\Phi(z_\alpha) = 1 - \alpha$.

The Commission now proposes (ComCom 2009b, paragraph 7.67) that these two compliance checks be applied jointly to SAIDI and SAIFI normalised data using the same value of z_α so that a network is deemed compliant only if both SAIDI and SAIFI reliability targets are not exceeded. However, since SAIDI and SAIFI values are dependent (the 17 networks examined by SRA yielded correlation coefficients that varied around 0.7), the false-positive rate or Type I error of this joint test is now given by

$$\begin{aligned} \alpha &= P(\bar{X} > \bar{x}_c \text{ or } \bar{Y} > \bar{y}_c | \text{network is compliant}) \\ &= 1 - P(\bar{X} \leq \bar{x}_c, \bar{Y} \leq \bar{y}_c | \text{network is compliant}) = 1 - \Phi(z_\alpha, z_\alpha | \rho) \end{aligned} \quad (1)$$

where $\Phi(x, y | \rho)$ is the distribution function of the bivariate normal distribution with zero means, unit standard deviations, and correlation ρ . Here \bar{X} and \bar{Y} denote the sample means of the normalised daily SAIDI and SAIFI data for the assessment year, and the reliability targets \bar{x}_c, \bar{y}_c are given by

$$\bar{x}_c = \mu_{SAIDI} + z_\alpha \sigma_{SAIDI} / \sqrt{n}, \quad \bar{y}_c = \mu_{SAIFI} + z_\alpha \sigma_{SAIFI} / \sqrt{n} \quad (2)$$

with $\mu_{SAIDI}, \mu_{SAIFI}, \sigma_{SAIDI}, \sigma_{SAIFI}$ and ρ denoting the respective means, standard deviations and correlation of normalised SAIDI and SAIFI data. The latter are estimated using analytical formulae derived from the statistical model adopted, or directly using suitable empirical estimates from the five year reference period. When $\rho = 1$ (perfect linear dependence) the individual and joint tests coincide, and when $\rho = 0$ (independence) the false-positive rate of the joint test becomes

$$1 - P(\bar{X} \leq \bar{x}_c | \text{network is compliant}) P(\bar{Y} \leq \bar{y}_c | \text{network is compliant})$$

which always exceeds the individual false-positive rate.

Unlike the individual tests, the false-positive rate of the joint test is now a function of the correlation ρ which varies across electricity distribution networks. Thus, to ensure that all networks have the same false-positive rate, it will be necessary to use different z_α values for each network according to the value of the correlation ρ of their normalised daily SAIDI and SAIFI values. This is readily done numerically by solving (1) for z_α and any given value of ρ , but is more complicated than a procedure based solely on SAIDI values, for example.

This leads to the following recommendation.

Recommendation 2 *With regard to the setting of SAIDI and SAIFI reliability targets, it is recommended that:*

- (a) *a common network-independent false-positive rate α be set in advance;*
- (b) *the reliability targets for each network be given by equation (2) where z_α is the solution to equation (1) and the correlation ρ is estimated from normalised data over the five year reference period.*

Further issues that the Commission may need to address are as follows.

Normal variability In ComCom (2009b), Section 7.5, the Commission states that reliability indices will demonstrate statistical variation and go on to describe this as “normal variability”. A better description, especially with regard to the variability of sample means, is the statistical term “sampling variability”. This would avoid a clash with the concept of normal variation defined earlier in the context of normalisation.

Definition of standard deviation For large samples of n independent observations, the sample mean of normalised daily SAIDI or SAIFI data will have a normal distribution with mean μ (either μ_{SAIDI} or μ_{SAIFI}) and standard deviation σ/\sqrt{n} (where σ is either σ_{SAIDI} or σ_{SAIFI}). Here σ is the standard deviation of the normalised daily SAIDI or SAIFI data, and σ/\sqrt{n} is the standard deviation of the sample mean. The latter standard deviation is often called the standard error of the sample mean. It is important to clearly distinguish between these two standard deviations since they are quite different conceptually and in size. However the two are confused in ComCom (2009b), Section 7.5 and Appendix C, especially in the draft decision given in ComCom (2009b), paragraph 7.67. A clearer terminology would correct existing errors and avoid misconceptions.

Sample size n and zero SAIDI days Currently the procedure described in ComCom (2009b), Appendix C, is correct if zero SAIDI days are included since Steps RT5 and RT6 assume that normalised SAIDI or SAIFI values are available for every day of the year ($n = 365$ or occasionally $n = 366$). If zero SAIDI days were excluded then n would need to be reduced considerably to reflect this. This inconsistency with the Commission’s proposed procedure, which ignores zero SAIDI days, is eliminated if Recommendation 1 is adopted by the Commission.

With regard to ComCom (2009b), paragraph 7.67 and Appendix C, it is noted that compliance is framed in terms of average annual totals of normalised daily SAIDI and SAIFI values, instead of average normalised daily SAIDI and SAIFI values. Such comparisons are equivalent, as noted in Thomson (2009), Section 4, since one is just a rescaled version of the other. However, retaining the same time scale throughout (days) has many advantages both conceptually and practically. In particular it minimises the likelihood of errors and the potential for confusion.

4 Assessing compliance

The Commission wishes to retain its “two out of three” assessment scheme noting that it is similar in philosophy to the simpler “two in a row” scheme proposed in Thomson (2009). Indeed, although the “two out of three” scheme always detects more breaches than the “two in a row” scheme, the statistical properties and detection rates of the two schemes are very similar as will be shown below. It is for this reason and to enhance transparency that the “two in a row” scheme was suggested as an alternative.

In ComCom (2009b), Appendix C, the Commission has undertaken analyses which endeavour to determine the differences between competing schemes. The schemes considered are:

Scheme 1 based on the compliance test alone (“single year”);

Scheme 2 the “two in a row” assessment scheme;

Scheme 3 the “two out of three” assessment scheme.

However the Commission’s comparisons of these three schemes are not valid since they failed to ensure that the overall false-positive rates of breaching the quality standard were the same for all three methods.

Let α_1 , α_2 and α_3 denote the false-positive rates for the compliance tests given by (1) under Schemes 1, 2 and 3 respectively. Then the false-positive rates for a breach in quality standards are given in Table 3. Consider one of the last three assessment years in the regulatory period. If the required false-positive rate for a breach in the quality standards was set at a constant α , independent of the scheme adopted, then to achieve this α_1 would need to be set to α , α_2 would need to be set to $\sqrt{\alpha}$ and α_3 would need to satisfy $\alpha = \alpha_3^2(2 - \alpha_3)$. Similar considerations apply to the first two years. Now all schemes have the same false-positive rates for a breach in quality standards and can be compared.

Table 3: False-positive rates for a breach in quality standards under Schemes 1, 2 and 3 where α_1 , α_2 and α_3 are the respective false-positive rates for the compliance tests given by (1).

Assessment year	Scheme 1	Scheme 2	Scheme 3
1	α_1		
2	α_1	α_2^2	α_3^2
3, 4 and 5	α_1	α_2^2	$\alpha_3^2(2 - \alpha_3)$

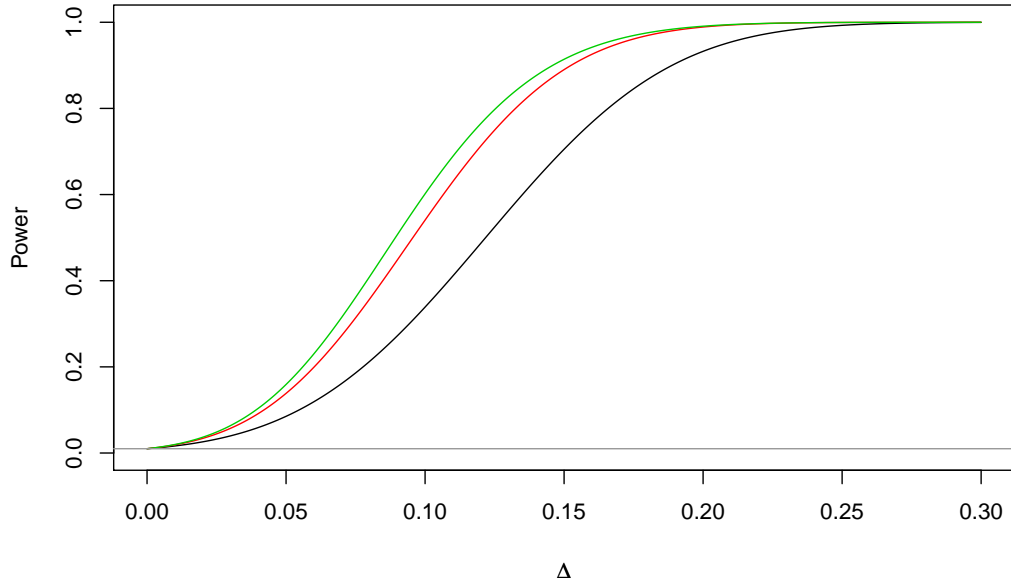


Figure 2: Power functions of the test for a breach in quality standards based on SAIDI compliance alone, and for one of the last three assessment years in the regulatory period. For Scheme 1 (black), Scheme 2 (red) and Scheme 3 (green) the false-positive rate is 1% (horizontal grey line) and the non-compliance alternative assumes that the mean of the network is $\mu_{SAIDI} + \Delta\sigma_{SAIDI}$ with all other parameters remaining the same.

How can schemes with the same false-positive rates be compared? Consider again the compliance test for each scheme i ($i = 1, 2, 3$) and the probability

$$\beta_i = P(\text{network assessed as compliant} | \text{network is non-compliant})$$

which is sometimes referred to as the false-negative rate. Here β_i is the Type II error and $1 - \beta_i$ the power of the compliance test. The probabilities β_i can now be used to determine the power of the test for a breach in quality standards in very much the same way as was done for the false-positive rates of a breach in Table 3. The more powerful the test for a breach in quality standards the better it is.

The power of a test under appropriate non-compliance alternatives (the power function) can often be evaluated analytically or, failing that, by simulation. An example of a power function is given in Figure 2 for the test of a breach in quality standards based on SAIDI compliance alone, and for one of the last three assessment years in the regulatory period. Here the false-positive rate for all three schemes has been set at a common value of $\alpha = 0.01$ (1%) and the non-compliance alternative assumes that the mean of the network is no longer μ_{SAIDI} , but $\mu_{SAIDI} + \Delta\sigma_{SAIDI}$ ($\Delta > 0$) with all other parameters remaining the same. There is little practical difference in the power functions for Schemes 2 and 3, with Scheme 3 slightly better at detecting departures from compliance than Scheme 2 as expected. However Schemes 2 and 3 are markedly better than Scheme 1 based on the compliance test alone.

Further comments that the Commission may wish to consider are as follows.

Number of compliant networks in breach of quality standards Consider any assessment scheme with false-positive rate α of a breach in quality standards. If the number of networks assessed in an assessment year is m and all are compliant, then the number of networks assessed as being in breach of the quality standards follows a binomial distribution with parameters α and m . In particular, this distribution has mean $m\alpha$ and standard deviation $\sqrt{m\alpha(1-\alpha)}$. This result and Table 3 can be used to determine results such as those considered in ComCom (2009b), Table 7.

Initialisation Schemes 2 and 3 are unable to make an assessment of a breach in the first year of the regulatory period. If it was important to extend either of Schemes 2 and 3 to the first year then the simple expedient is to augment these schemes by applying Scheme 1 in the first year.

Consistency over time The Commission's Consistency Principle is intended to apply over networks and over time. In this regard it seems important to ensure that, whatever assessment scheme is adopted, the false-positive rate for a breach in the quality standards should be the same for all networks and for all assessment years of the regulatory period.

References

- Christie, R.D. (2003a) Statistical classification of major event days in distribution system reliability. *IEEE Transactions on Power Delivery* **18**, 1336-1341.
- Christie, R.D. (2003b) Zero SAIDI days issue - response to WMECO. Technical report. <http://www.ee.washington.edu/faculty/christie/>
- Christie, R.D. (2003c) Maximum likelihood estimates for alpha and beta with zero SAIDI days. Technical report. <http://www.ee.washington.edu/faculty/christie/>
- ComComm (2009a) *Reset of Default Price-Quality Path for Electricity Distribution Businesses*. Regulation of Electricity Lines Businesses, Targeted Control Regime. Commerce Commission Discussion Paper.
- ComComm (2009b) *Initial Reset of the Default Price-Quality Path for Electricity Distribution Businesses*. Regulation of Electricity Lines Businesses, Targeted Control Regime. Commerce Commission Draft Decisions Paper.
- IEEE (2004) IEEE Guide for Electricity Power Distribution Reliability Indices. IEEE Std 1366TM-2003, IEEE Power Engineering Society.
- Royston, J.P. (1982) An extension of Shapiro and Wilk's W test for normality to large samples. *Journal of the Royal Statistical Society C* **31**, 115-124.
- Thomson, P.J. (2009) Comments on Chapter 8 of the Commerce Commission Discussion Paper *Reset of Default Price-Quality Path for Electricity Distribution Businesses (19 June 2009)*. A submission to the Commerce Commission commissioned by the Electricity Networks Association.

Warren, C.A. and Saint, R. (2005) IEEE reliability indices standards: major event day calculations and how they relate to small utilities. *IEEE Industry Applications Magazine* **11**, 16-22.