Non-iterative Method for Modelling Systematic Data Errors in Power System Risk Assessment

C.J. Dent *Member, IEEE* and J.W. Bialek, *Senior Member, IEEE*

***Abstract*—This paper provides a new framework for modelling uncertainty in the input data for power system risk calculations, and the error bars that this places on the results. Differently from previous work, systematic error in unit availability probabilities is considered as well as random error, and a closed-form expression is supplied for the error bars on the results. This closed-form expression reveals the relative contribution of different sources of error much more transparently than iterative methods. The new approach is demonstrated using the thermal units connected to the Great Britain transmission system. The availability probabilities used are generic type availabilities, published rounded to the nearest 5% by the system operator. Very wide error bars on the results of risk calculations result from the use of these probabilities; however, this is only revealed by modelling of the systematic error caused by the rounding. The approach is also used to investigate quantitatively the widely acknowledged view that comparing relative risks is a more robust use of simulated risk indices than stating absolute risk levels.**

***Index Terms*—Power system reliability, Measurement errors**

I. Introduction

INTEREST in power system reliability calculations has increased over recent years, due to the increasing installed capacity of variable output renewables. Applications have included estimating the capacity credit of wind generation [1] and calculation of operating reserve requirements [2]. In any such risk calculation, it is necessary to provide a realistic model of the conventional thermal and hydro plant availability, as well as the variable renewables of direct interest. This paper provides a new framework for modelling uncertainty in the availability probabilities assumed for conventional plant, and the error bars that this places on the results of risk calculations.

The low-availability tail of the probability distribution for available capacity decays rapidly. The relationship between system risk and unit availabilities is therefore highly nonlinear, and it is possible for even small errors in input data to have a substantial effect on the results of risk calculations. As an example, the installed generation capacity in Great Britain is around 75 GW; if all the unit availability probabilities change by 1% in the same direction, then the mean available capacity will change by almost 1 GW. This is substantial when compared with the standard deviation of the available capacity distribution, which (as will be seen later) is around 2 GW. Understanding and robust quantification of this effect are therefore important.

This work is funded through the EPSRC Supergen V, UK Energy Infrastructure (AMPerES) grant in collaboration with UK electricity network operators working under Ofgem's Innovation Funding Incentive scheme – full details on http://www.supergen-amperes.org/.

The authors are with the School of Engineering and Computing Sciences, Durham University, South Road, Durham DH1 3LE, UK (Email: chris.dent@durham.ac.uk, janusz.bialek@durham.ac.uk).

Power system reliability indices considering conventional generation only are usually calculated via the Capacity Outage Probability Table (COPT) method [3]. This involves a iterative calculation, adding the generating units to the model one at a time. The COPT iterative method was extended in [4] to situations where the unit forced outage rates (FORs) are not known precisely; a method for translating error bars on FORs to error bars on reliability indices is provided. Subsequent developments have included faster approximate methods based on Taylor expansions, for small data uncertainties, of the formulae for reliability indices, e.g. [5]–[7], and pp 63-67 of [3]. More recent work has extended this methodology to network reliability studies [8] and multi-area systems [9], and to the use of fuzzy set theory to model data errors [10].

All of this previous work has assumed that there is no systematic error in the FORs, i.e. the errors are independent, with no tendency for them to lie in the same direction. The errors may then be modelled as statistically independent random variables. This assumption was questioned by P.F Albrecht and W.J. MacFarland in a discussion of [4] (see pp 1334-1335):

> Recall that these random variables quantify the state of knowledge of the investigator – therefore, all forced outage rates have one thing in common – they were made by the same individual or group. Hence, it seems more likely that deviations in forced outage rate will all tend to be in the same direction.

It will be shown later that another situation where systematic error can also occur when using rounded generic data for each type of generating unit.

This paper introduces consideration of systematic error in FORs into power system reliability index calculation. Differently from earlier related work, which has used iterative calculations, here a direct method for translating uncertainty in FORs into error bars on Loss of Load Probability (LOLP) is provided. The new method is applicable as long as a continuous approximation to the distribution for available generating capacity is reasonable; while the individual unit capacities are modelled as discrete, this continuous approximation will be valid as long as the number of units is large and their capacities are sufficiently diverse. The stronger assumption of a Normal distribution is not required. The benefits of such a direct method, in terms of revealing transparently the contributions of random and systematic components of data errors, are great. Iterative calculations are focused on obtaining a result, but may not offer substantial insight into why the choice of input parameters caused a particular numerical value to be obtained.

Following a review of previous work in Section II, Section III introduces the idea of systematic error via an example based on the Great Britain power system, for which data is supplied rounded to the nearest 5%. Section IV introduces the new method, which is applied to model errors in the Great Britain example in Section V. The widely acknowledged view (see for instance page 5 of [11]) that comparing relative risks is a more robust use of simulated risk indices than stating absolute risk levels is investigated quantitatively in Section VI. Finally, conclusions are presented in Section VII; in the Great Britain example used, it is demonstrated that using rounded data introduces large systematic error in addition to that arising from uncertainty in the best point estimates.

II. Previous Work

A. FORs Known Precisely: Capacity Outage Probability Table

The method for calculating the probability distribution for the generating capacity on outage is well established, and is described in detail in [3]. It works by adding units to the system iteratively; if $p_u(x)$ is the probability that total capacity of at least x is on outage from the first u units, and units are modelled using a 2-state model (i.e. all or nothing available) then

$$p_u(x) = (1 - r_u)p_{u-1}(x) + r_u p_{u-1}(x - c_u), \qquad (1)$$

where r_u and c_u are respectively the FOR and capacity of unit u. This is generally known as the Capacity Outage Probability Table (COPT) method.

B. Uncertain FORs

1) Exact Method for Independent Errors: [4] models the errors in the FOR of unit u as a random variable R_u, with mean μ_{R_u} and variance $\sigma^2_{R_u}$. The probabilities in the COPT are then themselves random variables[1]. An iterative method of calculating their means $E[P_u(x)]$ and covariances $\mathrm{Cov}[P_u(x), P_u(y)]$, again adding one unit at a time, is derived. Error bars may be placed on the probabilities $p_u(x)$ by evaluating the variances $V[P_{n_U}(x)] = \mathrm{Cov}[P_{n_U}(x), P_{n_U}(x)]$, where n_U is the number of units.

2) Use of Assumption of Independence: The derivation of this iterative method for calculating $\mathrm{Cov}[P_u(x), P_u(y)]$ involves expressions such as

$$E[P_{u-1}(x)P_{u-1}(y)(1 - R_u)^2]. \qquad (2)$$

The assumption of independence allows this to be expanded as

$$E[P_{u-1}(x)P_{u-1}(y)]E[(1 - R_u)^2], \qquad (3)$$

which may be expressed in terms of $\mathrm{Cov}[P_{u-1}(x), P_{u-1}(y)]$. There is no obvious way of extending this to situations involving systematic error, as the $\{R_u\}$ are then not independent.

[1]Throughout this paper, the convention that random variables are denoted by capital letters, and fixed parameters by lower case, will be used.

TABLE I
Conventional unit types on the GB transmission system, the sums of unit Operational Realisable Capacities for each type, and the assumed availabilities from National Grid's Winter Outlook 2008/09 [12].

Type	No. Units	Capacity (GW)	Availability
Nuclear	22	10.5	80%
Hydro	42	1.1	60%
Coal	62	28.5	85%
Oil	7	3.7	95%
Pump storage	16	2.9	95%
OCGT	46	1.5	95%
CCGT	51	25.8	90%
	246	**74.0**	

3) Approximate Method: A more computationally efficient approximate approach, which requires an iterative calculation for the probabilities only (i.e. not the covariances), is presented in [7]. This uses a Taylor expansion for small FOR errors to express the covariances in terms of the entries in the COPT (1) and the variance of the $\{R_i\}$. To first order:

$$\mathrm{Cov}[P_{n_U}(x), P_{n_U}(y)] = \sum_{u=1}^{n_U} \left(\frac{\partial p_{n_U}(x)}{\partial r_u}\right)\left(\frac{\partial p_{n_U}(y)}{\partial r_u}\right)\sigma^2_{R_u}. \qquad (4)$$

The partial derivative with respect to r_u may be expressed in terms of entries from a COPT with unit u removed. This approach is appropriate for modelling systematic errors; (4) would then become

$$\mathrm{Cov}[P_{n_U}(x), P_{n_U}(y)] = \sum_{u,v=1}^{n_U} \left(\frac{\partial p_{n_U}(x)}{\partial r_u}\right)\left(\frac{\partial p_{n_U}(y)}{\partial r_v}\right)\mathrm{Cov}[R_u, R_v]. \qquad (5)$$

The approach in this paper is related to this approximate COPT-based method. The key advance is that by making the further approximation of assuming a particular form for $p_n(x)$, the new method provides a transparent closed-form expression for the error bar on calculated LOLP values.

III. Motivation: Rounded Availability Data for the Great Britain System

A. Units and Availability

1) Model for GB Conventional Plant: The conventional generation connected to the Great Britain transmission system is summarised in Table I. The unit capacities used are the Operational Realisable Capabilities (ORCs), as determined by National Grid, the System Operator [13].

2) Unit Availabilities: The availabilities in the table are described as 'assumed availabilities' in [12]. They are rounded to the nearest 5%, and their original use is as the typical proportion of the capacity of each type of unit which is available at winter peak. Using the two-state model for each unit as in Sections II and IV (i.e. either unit ORC or no capacity available), the mean and standard deviation of the distribution for total available capacity are 64.16 and 2.09 GW respectively. If these availability figures are to be used as unit availability probabilities in a risk calculation, it is important to

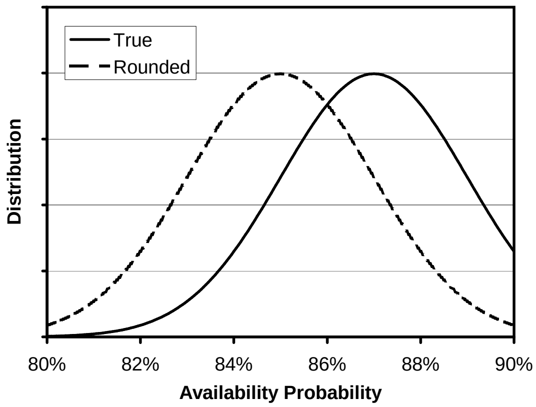

Fig. 1. Illustration of how using rounded typical type availabilities can lead to systematic error. The true availability probabilities are scattered around a central value of 87%. If a rounded typical probability of 85% is used, this introduces a common error across all units of the type, termed a systematic error.

have a means of estimating the consequences of such rounding errors for the result of the calculation.

B. Systematic Error in Rounded Data

The ‘assumed availabilities’ are average availabilities across the units of each type. This paper uses the reasonable picture that actual availability probabilities for each type of unit are distributed about some typical value (illustrated in Fig. 1 for a typical availability of 87%, along with the ‘modelled’ spread if the typical availability is rounded to 85%). Using the rounded assumed availabilities as unit availability probabilities for all units causes two types of error:

- *Rounding errors in the typical availabilities.* If the ‘assumed availability’ for a type of unit is given as 85%, then the best point estimate might be anywhere between 82.5% and 87.5%. Moreover, exactly the same error would apply to all units of the type if a probability of 85% is used for each such unit. Hence, this is a source of systematic error.
- *Scatter about the typical availabilities.* The actual availability probabilities for a type of unit will be scattered about the typical availability. As the deviations from the typical type value for different units are independent, this is a source of random error.

The modelling of errors for this example will be discussed in detail in Section V, after the methodology for treating systematic errors is introduced in the next section.

IV. Theory

A. Model

1) Normalised Capacity Variable: Here, a distribution for the available capacity (denoted X), rather than the capacity on outage, will be derived. If the unit availability probabilities ($a_u = 1 - r_u$ for unit u) are known exactly, then the mean and variance of X are:

$$\mu_X = \sum_u c_u a_u \tag{6}$$

$$\sigma_X^2 = \sum_u c_u^2 a_u (1 - a_u). \tag{7}$$

It is convenient in the derivation to work with the a normalised[2] capacity variable Z, which has mean 0 and standard deviation (SD) 1:

$$Z = \frac{X - \mu_X}{\sigma_X}. \tag{8}$$

This is familiar for large systems, where the Central Limit Theorem may be invoked to justify assuming a Normal distribution for X in a window about its mean; however, in this context it is useful even if a Normal approximation is not made. For the rest of this paper, these values for the mean available capacity μ_X, and its SD σ_X, calculated using central estimates for the availabilities will be referred to as μ_0 and σ_0 respectively. Given a fixed capacity level x, z_x^0 will be defined as $(x - \mu_0)/\sigma_0$.

2) Error Model: Uncertainty in the unit availability data may be quantified by modelling the availability probabilities as random variables $\{A_u\}$. The random and systematic errors which result from using a generic availability probability across each unit type are modelled by setting

$$A_u = a_{t(u)} + \Delta_{t(u)} + \Delta_u. \tag{9}$$

The unit types are indexed by t. The availability probability used for all units of type t is denoted a_t, and the difference between this and the true typical probability for type t is represented by Δ_t (the two components of error are illustrated in Fig. 1.) The $\{\Delta_t\}$ for different types are assumed to be independent. Finally, the random scatter of the true individual unit probabilities about the typical type value is represented by Δ_u. The $\{\Delta_u\}$ for different units are assumed to be independent. The expectation values of Δ_t and Δ_u are both zero, and their variances will be denoted σ_t^2 and σ_u^2. Earlier work essentially used this error model with $\sigma_t^2 = 0$ for all types.

B. Mean and Variance

1) Linearisation: When the unit availability probabilities are not known precisely, the mean and standard deviation of the total available capacity become random variables, which will be written as $\mu_0 + \Delta\mu$ and $\sigma_0 + \Delta\sigma$ respectively. Making a Taylor expansion for small uncertainties, the random variable representing uncertainty in the normalised capacity equivalent to x GW is:

$$Z_x = z_x^0 \left(1 - \frac{\Delta\mu}{x - \mu_0} - \frac{\Delta\sigma}{\sigma_0} \right). \tag{10}$$

[2]There is an unfortunate clash of notation between the term *normalised* and the *Normal distribution*. If a random variable is *normalised* to mean 0 and standard deviation 1, this carries no implication that it is assumed to be Normally distributed, unless explicitly stated. A normalised capacity variable has been used for a similar purpose in capacity credit calculations [14].

Expanding $\Delta\mu$ in terms of the unit errors,

$$\Delta\mu = \sum_u \left[(\Delta_{t(u)} + \Delta_u) \left. \frac{\partial \mu_0}{\partial a_u} \right|_{a_u = a_{t(u)}} \right], \tag{11}$$

and making a similar expansion for $\Delta\sigma$, the following expression is obtained (the partial derivative $\partial\mu/\partial a_u$ is simply c_u, see (6)):

$$Z_x = z_x^0 - \frac{1}{\sigma_0} \sum_u c_u k_u (\Delta_{t(u)} + \Delta_u), \tag{12}$$

where

$$k_u = 1 + \frac{z_x^0 c_u}{2\sigma_0}(1 - 2a_u). \tag{13}$$

2) Mean and Variance of Z_x: Finally, the mean and variance of Z_x are:

$$\mu_{Z_x} = z_x^0 \tag{14}$$

$$\begin{aligned} \sigma_{Z_x}^2 &= \sum_u \left[\sigma_u^2 \left(\frac{c_u}{\sigma_0} \right)^2 k_u^2 \right] \\ &\quad + \sum_t \left[\sigma_t^2 \left(\sum_{u \in U_t} \left(\frac{c_u}{\sigma_0} \right) k_u \right)^2 \right]. \end{aligned} \tag{15}$$

For units of a given type, this clearly indicates how correlated systematic errors can be much more significant than independent random errors; the contribution of units of type t in the first (independent error) sum in (15) is proportional to the number of such units, whereas in the second (systematic error) term it is proportional to the number of units squared. The closed form expression (15) is the key contribution of this paper, as it gives greater insight into how errors combine than previous iterative approaches. The next paragraphs show how it may be used to derive error bars on the result of risk calculations; if a Normal approximation for available generating capacity is assumed, this may be done using direct (i.e. non-iterative) expressions throughout.

C. Error Bars on Loss of Load Probability

1) Error Bars on Z: (15) allows the calculation of the standard deviation of Z, given the standard deviations representing the random and systematic parts of the data errors (in principle, it is possible to deduce the form of the distribution for Z given the distributions for the $\{\Delta_t\}$ and $\{\Delta_u\}$, but it is almost certainly unrealistic to model the input errors in sufficient detail to make this worthwhile.) In this paper, an error bar covering the likely range of values of a calculation result will be taken as 2 standard deviations above and below the central estimate. For the value of z corresponding to a generating capacity x (see Paragraph IV-A1 for definitions), this means an error bar of $z_0 \pm 2\sigma_Z$.

2) Error Bars on LOLP: Normal Approximation: For a demand level d, the Loss of Load Probability (LOLP) is simply the cumulative distribution function $F_Z(z) = p(Z \leq z)$, evaluated at the value of z corresponding to an available capacity d. Given a form for this distribution, it is straightforward to convert error bars on z into error bars on the LOLP. As stated earlier, for large systems and sufficiently high LOLPs (as discussed in Paragraph V-A1, for the GB model this means LOLPs greater than about 3%), the Central Limit Theorem may be invoked to justify assuming a Normal distribution for the available generating capacity. The cumulative distribution for the standard Normal distribution (with mean 0 and standard deviation 1) is

$$F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\bar{z}^2/2} \mathrm{d}\bar{z}. \tag{16}$$

The error bar on the LOLP is then the interval $[F_Z(z_0 - 2\sigma_Z), F_Z(z_0 + 2\sigma_Z)]$, with a central estimate of $F_Z(z_0)$.

3) Error Bars on LOLP: Other Distributions: As long as the system is sufficiently large, and the unit capacities sufficiently diverse, the distribution for available capacity may be regarded as continuous, irrespective of whether a Normal approximation is valid (for smaller systems, or those with less diverse generation connected, methods such as those described in Section II may be more appropriate.) Provided that the shape of the distribution for available capacity does not change substantially over the range of possible unit availability probabilities, this same approach could then be applied (albeit with less formal justification). Paragraph V-A2 will show that this condition is indeed satisfied for the Great Britain example used here. A normalised (i.e. rescaled to mean 0 and standard deviation 1) version of the distribution function obtained through the COPT method would be used in (16), instead of the standard Normal distribution (one and only one COPT calculation is required to obtain the form of the distribution.) The form of the distribution for Z does not affect any earlier parts of the derivation.

4) 'Error Bars' or 'Confidence Interval'?: In the original paper on risk assessment with uncertain availability probabilities [4], the term 'confidence intervals' was used for the range of reasonable LOLP vales found based on the SD of the distribution for the LOLP. In their discussion on that paper, Albrecht and MacFarland commented that these might more properly be called Bayesian confidence intervals. Some other work, e.g. [3], presents only the SD of the LOLP distribution. We prefer the term 'error bar', as used above, due to the difficulty in defining precisely the variances of the distributions representing availability probability errors. We believe that this term captures the fact that the interval derived gives an order-of-magnitude estimate for the uncertainty in the calculated LOLP, rather than a more precise statement of the possible error.

V. Example – Rounded Availability Data in GB

A. Test System

The Great Britain test system used in this paper is described in Section III. The mean and standard deviation of the distribution for available capacity, calculated using the unit availability probabilities in Table I, are 64.16 and 2.09 GW respectively.

1) COPT versus Normal Approximation: The COPT-derived cumulative distribution function for available capacity is compared with the Normal approximation in Fig. 2 (upper panel). It may be seen that the Normal approximation is reasonable down to capacities of about two standard deviations

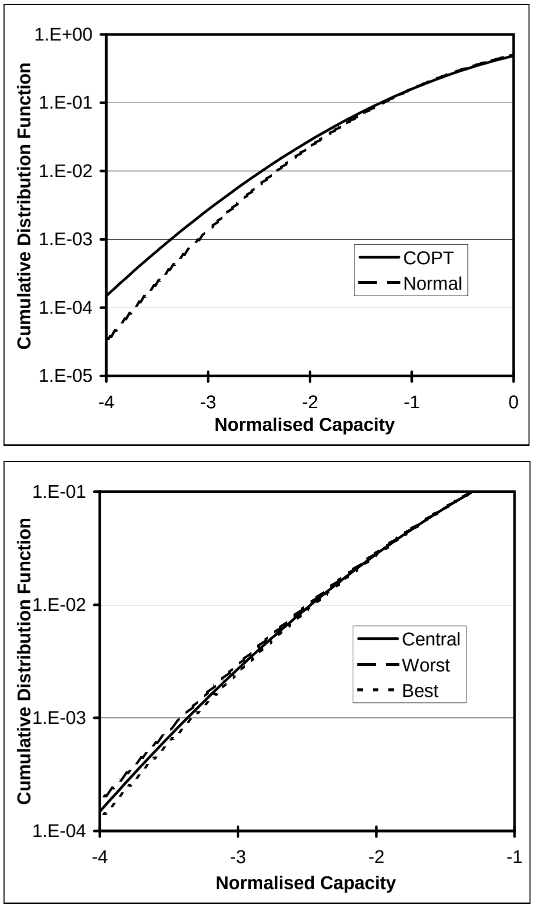

Fig. 2. Upper panel: comparison between COPT-derived cumulative distribution function and Normal approximation. Lower panel: comparison between the shapes of the COPT-derived central, best and worst case distributions, as described in Paragraph V-A2. In each case the x-axis the the normalised capacity $(X - \mu_X)/(\sigma_X)$.

below the mean (corresponding to LOLP of around 3%). This is consistent with the expectation that a Normal approximation based on the Central Limit Theorem will be best near the mean, and that the tails of the exact distribution will usually be fatter than those of the Normal distribution.

2) Central, Best and Worst Case Probabilities: In order to examine whether the method described in this paper can be used when the Normal approximation does not suffice (see Paragraph IV-C3), it is necessary to investigate how the form of the COPT distribution depends on the unit availability probabilities. The probabilities used are given rounded to the nearest 5%. As discussed in Section III-B, it might therefore reasonably be assumed that (e.g.) a probability given as 85% can be taken to lie between 82.5% and 87.5%. Cumulative distribution functions for available capacity using

- the central case (the given probabilities)
- the best case (all probabilities increased by 2.5%)
- the worst case (all probabilities decreased by 2.5%)

are compared in the lower panel of Fig. 2. By plotting the distribution functions against normalised capacity, it may be seen that the shapes of the distributions are indeed very similar. It is reasonable to deduce that this will be the case for any combination of availability probabilities which may be rounded to the given values. For this test problem, the direct method described in Section IV may therefore be used to derive error bars on LOLP for a demand d (i.e. not just on the corresponding value of z), even when a COPT-based LOLP calculation is performed.

3) Limitations: All the models described in this paper assume that if a unit is mechanically available, then its full capacity is available to the system. This might not quite reflect reality due to transmission constraints, limits on energy output at a station due to emissions constraints, derated operating states, or other operational factors. In Great Britain, some of these effects are encapsulated in station Operational Realisable Capacities (see Paragraph 49 of [15]); it is assumed that a station's total output cannot exceed its station ORC. The generating capacities given by type in the Winter Outlook [12] are based on these station ORCs, hence the difference from those listed in Table I. The COPT method may be extended to model these station output limits if the availability probabilities are assumed to be known exactly [16], but the (both previous iterative and new non-iterative) methods for dealing with availability probability uncertainty discussed here do not extend in such a straightforward manner. However, results for the order-of-magnitude of uncertainty in LOLP due to uncertainty in availability probabilities, as opposed to the precise width of error bars, are expected to be robust even given these caveats.

B. *Quantification of Data Uncertainty*

1) Modelling Rounded Data: If a data point is given rounded to the nearest 0.05 as a, then it is reasonable to assume that the true best point estimate may lie anywhere between $a - 0.025$ and $a + 0.025$. This uncertainty may naturally be modelled as a Uniform distribution on $[a-0.025, a+0.025]$. As the probability density function for this distribution is constant within that range, and zero elsewhere, it is easy to verify that it has mean a and standard deviation (SD) $0.05/\sqrt{12} = 0.0144$.

2) Random and Systematic Components: As discussed in Section III, the 'assumed availabilities' in National Grid's Winter Outlook are intended for use as average availabilities across each type of unit; there is no implication that all units of a given type should be modelled as having the same availability probability. Therefore:

- a_t, the central estimate of the typical availability probability for units of type t, is taken to be the Winter Outlook assumed availability for that type.
- Δ_t, the distribution representing error in the typical availability probability for type t, is modelled as a Uniform distribution on $[-0.025, 0.025]$ as described in the previous paragraph. This gives $\sigma_t = 0.0144$ for all t.
- Random variations of individual unit availabilities about these typical values are modelled by σ_u, the SD of the distribution for Δ_u.

C. *Results*

1) Variation of LOLP With Demand: The central estimate for, and upper and lower ends of the error bar on, LOLP are

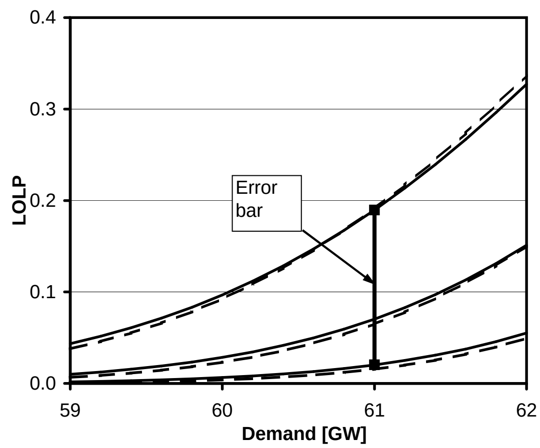

Fig. 3. Central estimate for, and error bar on, LOLP, plotted against fixed demand. The solid lines show the limits error bar calculated using the COPT-derived distribution, and the dashed lines the limits of the error bar using the Normal approximation.

plotted against fixed demand in Fig. 3. The distributions for the systematic errors are chosen to model the rounded availability data, as described in the previous paragraph. The random error in all units' availabilities are modelled as having standard deviations of 0.01. Results using both a COPT-derived distribution for available capacity, and a Normal approximation, are shown; particularly at low LOLP levels, the Normal approximation to the central case may differ substantially from the COPT distribution, but these differences are much smaller than the error bars on either. The graph illustrates just how much uncertainty the use of rounded generic data induces in the final LOLP results; at a fixed demand of 59 GW, the LOLP varies by a factor of more than 40 from 0.17% to 4.4%, and at 62 GW it varies from 5.5% to 33%. For this example, while at low LOLP the COPT calculation is required for the central estimate, the fully closed form approach using the Normal approximation gives the key information on the width of the error bar. The coarse rounding adds substantial additional error above that due to uncertainty in the best point estimates.

2) Importance of Systematic and Random Errors: The variation of the error bars on the LOLP with the error distribution SDs is demonstrated in Fig. 4. The central case used has σ_t (which drives the systematic error) based on the rounding model discussed above, and random error $\sigma_u = 0.01$; this is marked with black squares in both panels. The figure clearly demonstrates how the uncertainty in LOLP depends much more strongly on the size of the systematic error than it does on the random component of the error; this confirms the prediction following (15). Indeed, a random component of the same order of magnitude as the systematic 'type' components hardly affects the LOLP error bar at all. For the central case, where the possible systematic error is quite large, the precise values for the $\{\sigma_u\}$ chosen to quantify the random error components are therefore not very important. It may also be observed in the upper panel that inaccuracies due to a Normal approximation for the capacity distribution are more significant when the possible error is small.

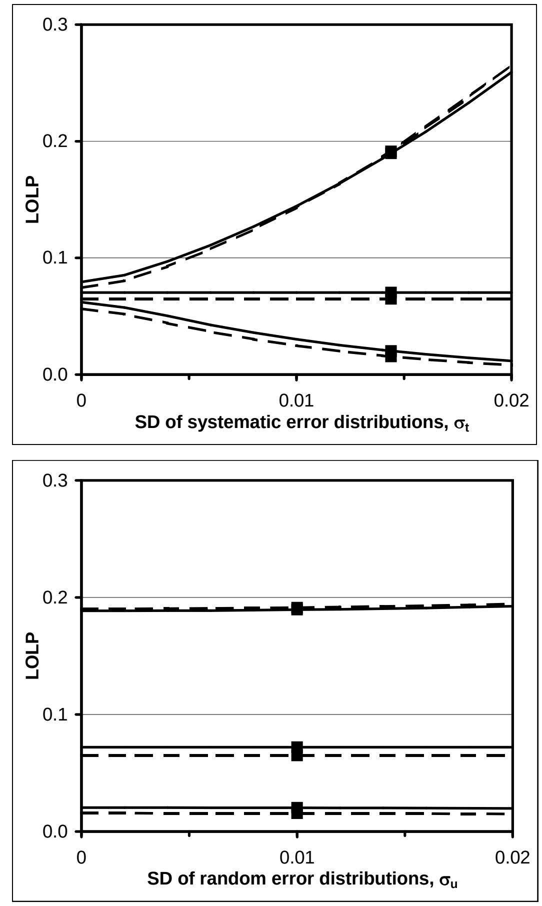

Fig. 4. Central estimate for, and upper and lower bounds on, LOLP, plotted against the SD of error distributions. A fixed demand of 61 GW is assumed. All types are assumed to have the same σ_t, and all units the same σ_u. Upper panel: variation of LOLP with σ_t for fixed σ_u of 0.01. Lower panel: variation of LOLP with σ_u for fixed σ_t of $0.05/\sqrt{12}$. The black squares indicate the common point between the two graphs, where $\sigma_u = 0.01$ and $\sigma_t = 0.05/\sqrt{12}$. The COPT calculation is represented by solid lines, and the Normal approximation by dashed lines.

3) Implications for Practical Risk Calculations: It is clear from the results presented that any risk calculation based on the model used in this paper, and the rounded data considered, will yield highly approximate results. However, as discussed in Paragraph V-A3, it may not be possible for all power stations to supply capacity equal to the sum of their unit capacities. As this reduces the effective capacity connected to the system by about 2 GW (around one standard deviation of the distribution for available capacity), this will have a very substantial effect on the results of risk calculations. The method in this paper will however still give a good idea of the degree of uncertainty in the results resulting from uncertainty in availability probabilities alone.

4) Error Quantification: For the example presented here, the error in the final results is dominated by rounding errors in the input data. As discussed earlier, rounding errors may naturally be modelled as Uniform random variables. In more general situations, for instance where a best point estimate is used for the typical type availabilities but where systematic error might still be possible, there might not be such a natural choice of distribution family. Partly as a consequence (as discussed in Paragraph IV-C1) the choice of error bars (effectively the $\{\sigma_t\}$) on the input data is necessarily rather approximate. This and the previous paragraph emphasise still further our suggestion in Paragraph IV-C4 that our and related methods for calculating errors in risk indices should be regarded as order-of-magnitude estimates for the errors, rather than calculations of precise error distributions.

5) Validity of Method: As demonstrated by Fig. 2 and Paragraph V-A2, the method presented here does not rely on a Normal approximation for its validity. It does however rely on the possibility of approximating the discrete distribution for available capacity with a continuous random variable. This approximation is clearly valid for the test system used, as would be expected from the system size and diversity of units, and is demonstrated by the smooth nature of the curves in Fig. 2. One benefit of using the Normal approximation however is that the various formulae are then completely closed form, and require no reference to an iterative COPT calculation at all; in this case, where the possible error is rather large, the Normal approximation provides all the key information (Figs. 3 and 4). For smaller systems, or ones with less diversity of unit capacities, it might therefore be necessary if quantitative results are required to use an iterative method more closely based on the COPT approach. The direct calculation should still however yield valuable insight into the order of magnitude of errors, and the relative importance of systematic and random errors.

VI. Error in Relative Risks

A. Model

1) Absolute and relative risks: It is widely acknowledged in the literature that comparing the relative values of simulated risk under different scenarios is more robust than quoting absolute values of risk (see for example page 5 of [11].) The same approach as above may be used to quantify the relative uncertainties in calculations of relative and absolute risk arising from uncertain availability probabilities; in the former, data uncertainties cancel out to some extent.

It is not possible to transform these relative z-values into relative risks in such an instructive way as in the previous sections, and it is therefore necessary to express results in terms of uncertainties in relative z-values as opposed to system risk probabilities. While this section provides important theoretical support for the robustness of comparing relative risk levels, it therefore not so useful in quantifying error bars in practical risk calculations.

2) Definitions and Derivation: The difference in normalised capacities equivalent to (x, y) GW will be denoted $Z_{xy} = Z_x - Z_y$; as in Section IV, the uncertainty in availability probabilities is modelled using random variables. It follows that

$$Z_{xy} = z^0_{xy}\left(1 - \frac{1}{2}\sum_u \left[\left(\frac{c_u}{\sigma_0}\right)^2 (1-2a_u)(\Delta_{t(u)} + \Delta_u)\right]\right) \tag{17}$$

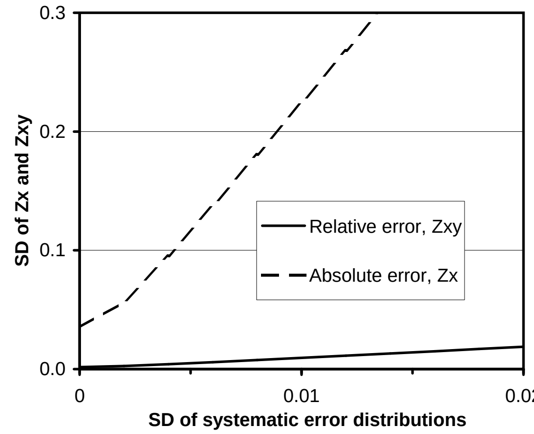

Fig. 5. SD of Z_{xy} against systematic error, for $x = 61$ GW and $y = 60$ GW. For comparison, the variance of Z_x is also plotted.

It is clear that $\mu_{Z_{xy}} = z^0_{xy}$. The uncertainty may again be estimated from the standard deviation:

$$\sigma^2_{Z_{xy}} = \frac{(z^0_{xy})^2}{4}\left[\sum_u \sigma_u^2 \left(\frac{c_u}{\sigma_0}\right)^4 (1-2a_u)^2 + \sum_t \sigma_t^2 \left(\sum_{u\in U_t}\left(\frac{c_u}{\sigma_0}\right)^2 (1-2a_u)\right)^2\right] \tag{18}$$

It may be seen immediately that, when taking the difference of two normalised capacities, much of the systematic error has cancelled. The second (systematic error) term is vastly the more important in (15) and (18). In the relative risk result (18), k_u (which for availability probabilities of at least 0.5 and $z^0_u < 0$ is always greater than 1) has been replaced by

$$\left(\frac{z^0_{xy}}{2}\right)\left(\frac{c_u}{\sigma_0}\right)(1-2a_u),$$

in which for most relevant values of z_{xy} all three terms have magnitude less than 1; because these terms are summed and squared, the relative error (18) will be much smaller than the absolute error (15).

B. Numerical Results

The SD of Z_{xy} is plotted in Fig. 5 against the SD σ_t which quantifies the systematic error; the two demand levels for which the risks are compared are $x = 61$ GW and $y = 60$ GW. For comparison, the SD of Z_x is also plotted. It may be seen that the uncertainty in the relative risk calculation is more than an order of magnitude less than that found in the absolute risk calculation. This verifies the algebraic inference in the previous paragraph, and also provides quantitative support for the commonly stated view that comparing relative risk levels is much more robust than quoting absolute risk levels.

VII. Conclusions

This paper has presented a new direct approach to estimating uncertainty in power systems reliability calculations, based on the uncertainty in the unit availability probabilities.

Differently from previous approaches, it provides a closed-form (non-iterative) expression for the error bar on the central estimate for loss-of-load probability. A further new contribution is to consider the effect of systematic error, where there may be a tendency for errors in the availability probabilities of multiple units to be in the same direction, as well as random errors.

Systematic error is considered by dividing availability probability errors into two components: the systematic component, which is modelled by a single random variable for each unit type, and the random component, which is modelled by a separate independent random variable for each unit. The analytical expression for the error bar on the LOLP demonstrates clearly and quantitatively how, for the same degree of uncertainty, systematic errors are far more significant than random ones. The transparency of such calculations is a major advantage over earlier iterative approaches. Moreover, the method is applicable for any system where the distribution for total available capacity may be approximated as continuous; the stronger assumption of a Normal distribution is not required. However, where a Normal distribution is assumed, the method is completely non-iterative (otherwise, a single capacity outage probability table calculation is required to obtain the form of the available capacity distribution). The new approach is also used to verify the commonly stated view that comparing relative risk levels is much more robust than quoting absolute risk levels.

The method has been demonstrated using as a test system the conventional plant connected to the Great Britain transmission network. Data on generic availabilities for each type of unit is published by the system operator rounded to the nearest five percent. It is demonstrated that this places an error bar on the result of typical LOLP calculations which can extend over an order of magnitude or more about the central value; the coarse rounding adds additional error above that due to uncertainty in the best point estimates.

Acknowledgment

The authors acknowledge valuable discussions with K. Bell, D. Eager, C.M. Gibson, G.P. Harrison, A. Keane, their partners in the Supergen-AMPerES consortium, the Probability and Stochastic Models group at Heriot-Watt University, and the Energy Requirements team at the National Grid Company. They express particular thanks to National Grid for sharing data on generating units' operational capabilities.

References

[1] M. Milligan, "Modelling utility-scale wind power plants. Part 2: Capacity credit," *Wind Energy*, vol. 3, 2001.

[2] "A New Approach to Quantify Reserve Demand in Systems With Significant Installed Wind Capacity, author=R. Doherty and M. O'Malley, journal=IEEE Trans. Power App. Syst., volume=20, number=2, month=May, year=2005, pages=587-595,."

[3] R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems, 2nd Edition*. Plenum, 1994.

[4] A. D. Patton and A. Stasinos, "Variance and Approximate Confidence Limits on LOLP for a Single-Area System," *IEEE Trans. Power App. Syst.*, vol. PAS-94, no. 4, pp. 1326–1336, July/Aug 1974.

[5] A. D. Patton and N. H. Tram, "Sensitivity of genaration reliability indices to generator parameter variations," *IEEE Trans. Power App. Syst.*, vol. PAS-97, no. 4, pp. 1337–1343, July/Aug 1978.

[6] G. J. Anders, "A Novel Approach to Frequency and Duration Analysis and Uncertainty Considerations for Radial and Two Interconnected Systems," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 7, pp. 2165–2172, July 1983.

[7] G. Hamoud and R. Billinton, "An Approximate and Practical Approach to Including Uncertainty Concepts in Generating Capacity Reliability Evaluation," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 3, pp. 1259–1265, March 1981.

[8] A. M. Oliveira, A. C. G. Melo, and L. M. V. G. Pinto, "Consideration of equipment failure parameter uncertainties in bus composite reliability indices," *Power Engineering Society 1999 Winter Meeting, IEEE*, vol. 1, pp. 448–453 vol.1, Jan-4 Feb 1999.

[9] Suhartono, Y. Zika, and S. Sasaki, "The Evaluation of Confidence Limit on LOLP for Multi Area System," *Electric Utility Deregulation and Restructuring and Power Technologies, 2000. Proceedings. DRPT 2000. International Conference on*, pp. 592–595, 2000.

[10] J. O. Kim and C. Singh, "Including Uncertainty in LOLE Calculation Using Fuzzy Set Theory," *Power Systems, IEEE Transactions on*, vol. 17, no. 1, pp. 19–25, Feb 2002.

[11] W. Li, *Risk Assessment Of Power Systems: Models, Methods, and Applications*. Wiley-IEEE, 2005.

[12] "Winter Outlook Report 2008/9," National Grid, Tech. Rep., 2 October 2008, available from http://www.nationalgrid.com/uk/Gas/TYS/outlook/.

[13] National Grid Company, personal communication.

[14] K. Dragoon and V. Dvortsov, "z-Method for Power System Resource Adequacy Applications," *IEEE Trans. Power Syst.*, vol. 21, no. 2, May 2006.

[15] "Summer Outlook Report 2009," National Grid, Tech. Rep., April 2008, available from http://www.nationalgrid.com/uk/Gas/TYS/outlook/.

[16] P. E. Olmos Aguirre, C. J. Dent, and J. W. Bialek, "Realistic Calculation of Wind Generation Capacity Credits," Cigré/IEEE PES Symposium on Integration of Wide-Scale Renewable Resources into the Power Delivery System, Calgary, July 2009.

Chris Dent (M'08) is a Research Fellow in the School of Engineering and Computing Sciences, Durham University, U.K. He obtained his BA in Mathematics from Cambridge University in 1997, PhD in Theoretical Physics from Loughborough University in 2001, and MSc in Operational Research from the University of Edinburgh in 2006. From 2007-2009 he was with the University of Edinburgh. His research interests lie power system optimisation, risk modelling, economics, and renewables integration.

Dr. Dent is a Chartered Physicist, an Associate of the OR Society, and a member of the Institution of Engineering and Technology, the Operational Research Society and Cigré.

Janusz W. Bialek (SM '08) is Professor of Electrical Power and Control at Durham University, UK. He holds M.Eng (1977) and Ph.D. (1981) degrees from Warsaw University of Technology, Poland. From 1981-1989 he was with Warsaw University of Technology, from 1989-2002 with Durham University, UK, and from 2003-2009 he held the Chair of Electrical Engineering at the University of Edinburgh. He has co-authored 2 books and over 100 technical papers. His research interests include power system economics, power system dynamics and sustainable energy systems.