SCIENCE

Statistical analysis of flashover data using a generalised likelihood method

A.J. Davies, BSc, PhD, CEng, FIEE
A.R. Rowlands, BEng(Tech), BA, DipEE, CEng
R. Turri, Dott Ing
Prof. R.T. Waters, BSc, PhD, CEng, FIEE

Indexing terms: Arcing, Breakdown and gas discharge

Abstract: In order to ensure consistency in the treatment of flashover data, it is desirable that a standard method of analysis of test results be established. Previous workers have suggested that the maximum likelihood method would be appropriate for estimating the parameters of the breakdown probability distribution, such as V_{50} and σ in the normal distribution. The present paper shows how this method may be applied in practice and can be extended to determine the confidence region associated with sets of parameters, and the confidence interval on a single parameter. It is also shown how other parameters, such as the voltage level corresponding to a specified probability of breakdown, may be determined. An example is given of the application of the method to a typical data set for flashover in air together with a listing of a suitable computer program in Fortran 77. It is recommended that this method be considered for adoption as a standard procedure.

1 Introduction

A variety of different test strategies may be employed in investigating the breakdown characteristics of a discharge gap subjected to impulse voltages.

In the Class I multilevel test procedure [1], for example, a number of impulses are applied for a fixed value of the crest voltage and the number of breakdown and withstand events recorded. The time between the voltage applications must be sufficient for each application to be independent (i.e. the same initial conditions apply) and the test is repeated for various crest voltages in the region where the breakdown probability undergoes a transition from a low to a high value.

In analysing the results of such tests it is normally desired to evaluate the crest voltage V_{50} at which there is a 50% probability of flashover and also some measure of the range of voltages over which the probability of breakdown $P_B(V)$ increases from very low to very high values (Fig. 1*a*).

Most workers in the past have taken the form of the distribution of Fig. 1*a* to be that of the cumulative

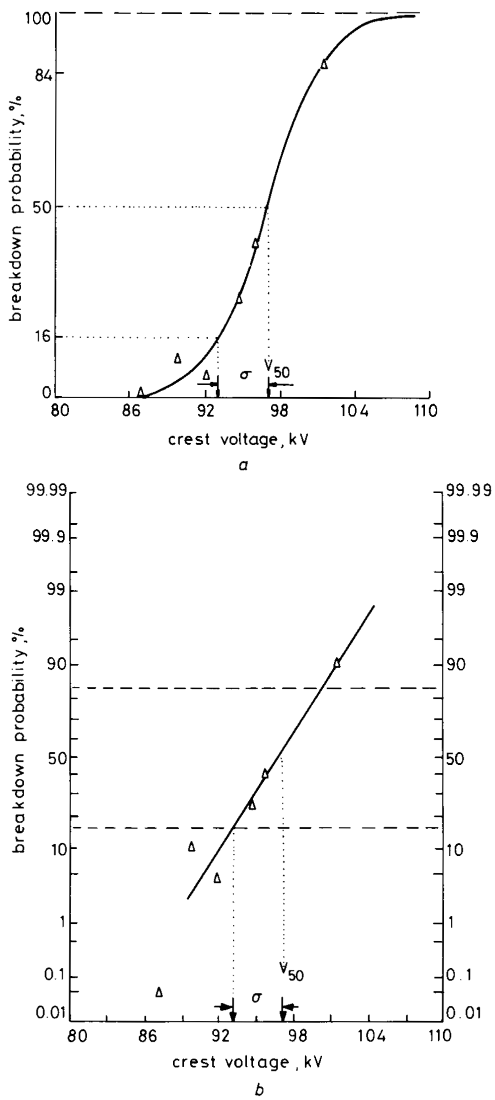

Fig. 1 *Typical breakdown probability distribution on linear axes (a) and cumulative normal probability paper (b)*
Δ = observed frequencies of breakdown (data of Table 2)

Paper 5723A (S3), first received 21st November 1986 and in revised form 21st July 1987

Dr. Davies and Dr. Turri are with the Department of Physics, University College of Swansea, Singleton Park, Swansea SA2 8PP, United Kingdom

Prof. Waters and Mr. Rowlands are with the Department of Physics, Electronics & Electrical Engineering, University of Wales Institute of Science and Technology, PO Box 25, Cardiff CF1 3XE, United Kingdom

normal distribution:

$$P_B(V) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{V} \exp\left[-(V - V_{50})^2/2\sigma^2\right] dV \qquad (1)$$

since most breakdown data agrees well with this distribution for $P_B(V)$ greater than about 2% [2, 3]. Some authors [4–6] have suggested that alternative distributions such as the Weibull, Gumbell or extreme-value distributions are more appropriate at very low probability levels. All these distributions, however, are very close to the cumulative normal distribution for probability levels greater than 2%.

Whatever distribution is assumed to be appropriate for a given case, some method must be adopted for estimating the parameters defining the distribution.

For multilevel tests the parameters V_{50} and σ in the distribution (eqn. 1) have most commonly been estimated by plotting the observed frequencies of breakdown on normal probability paper, drawing the 'best' straight line by eye (Fig. 1*b*) and determining the values of V_{50} and σ from

$$V_{50} = V[P_B(V) = 0.5]$$
$$\sigma = V[P_B(V) = 0.5] - V[P_B(V) = 0.16] \qquad (2)$$

It should be pointed out at this stage that σ must not be interpreted in any sense as the standard deviation of V_{50}. It is simply the parameter in eqn. 1 which gives a measure of the rate at which $P_B(V)$ goes from low to high values with increasing crest voltages.

A number of alternative test strategies have also been adopted which have the aim of establishing an accurate estimate of one particular parameter with fewer overall impulses than that necessary in the full multilevel test procedures. Typical standard procedures are the Class III 'up-and-down' method [1] for determining V_{50} and the 'extended up-and-down' test [7] for evaluating crest voltages corresponding to low breakdown probability levels. The basic philosophy of such strategies is illustrated by the following example.

Suppose that n impulses are applied at a given crest voltage. If no breakdowns occur then the voltage is increased by a predetermined amount and n impulses again applied. If one or more breakdowns occur then the voltage is reduced by the same amount. Repeating this procedure, the average crest voltage V_P will approach a level at which there is a 50% probability that no breakdowns occur in the n impulses, i.e.

$$(1 - P)^n = 0.5 \qquad (3)$$

where P is now the probability that breakdown occurs in any single impulse at the voltage V_P. The number of impulses is therefore chosen to give an average crest voltage corresponding to the desired probability level. For example, $n = 7$ and 20 give $P = 9.4\%$ and 3.4%, respectively, i.e. $V_P = V_{9.4}$ and $V_{3.4}$ [8]. Another modification to the above procedure is to reduce the crest voltage immediately after the first flashover in a series.

It should be noted that, even when the test procedure is aimed at measuring a single V_P, the investigator very often wishes to extrapolate the results in order to obtain an estimate of the crest voltage corresponding to a very low breakdown probability level. Some method for evaluating the best probability distribution from such data is therefore highly desirable.

Whatever testing strategy is adopted, providing all the impulses are independent of each other, the test data will give the frequencies of breakdown at a number of crest voltages. The number of impulses at a given level, however, may vary widely.

Brown [3] and Carrara and Yakov [6] have suggested that the maximum likelihood method is the most appropriate technique for estimating the parameters of a breakdown probability distribution (e.g. V_{50} and σ in eqn. 1). Brown estimated the confidence intervals associated with V_{50} and σ using both the concept of tolerance limits and a semiempirical graphical method, and also made a detailed comparison of his procedures with other accepted techniques. The confidence intervals for a single parameter, estimated by the different techniques, were in good agreement and further details may be obtained in [3]. Carrara and Yakov [6] pointed out that it should be possible to extend the maximum likelihood method to estimate the confidence intervals, but their suggested procedure was somewhat empirical. In the present work a standard procedure, based on the likelihood function, is established for estimating the parameters of the distribution together with their corresponding confidence regions and intervals.

In the following it will be assumed that the cumulative normal distribution is appropriate but the techniques described are equally applicable to any probability distribution.

2 Estimation of V_{50} and σ

In standard statistical texts [9–12] it is stated that the most powerful method for estimating the parameters of an assumed distribution based on a given set of data is the method of maximum likelihood. Brown [3] was the first worker to apply this technique to flashover data and it has also more recently been used by Hylten-Cavallius *et al.* [4], Carrara and Yakov [6] and Eriksson *et al.* [14].

For multilevel tests consisting of n_k impulses at each of N voltage levels, the likelihood function is given by

$$L(y/H(V_{50}, \sigma)) = C \prod_{k=1}^{N} P_k^{b_k}(1 - P_k)^{n_k - b_k} \qquad (4)$$

where b_k is the number of observed breakdowns at the kth level, P_k is the probability of breakdown as determined from eqn. 1 and C is an arbitrary constant.

We note that each series of impulses is a binomial trial of whether or not the gap breaks down and that $L(y/H(V_{50}, \sigma))$ is proportional to the probability that the observed data set y occurs on the hypothesis, $H(V_{50}, \sigma)$, that V_{50} and σ take on certain values. In the maximum likelihood method the likelihood function L is maximised and thus yields the 'maximum likelihood' estimates, V_{50}^* and σ^*, of the true parameters. Note that in this procedure the data set remains constant while the hypothesis $H(V_{50}, \sigma)$ is chosen so as to maximise L. Alternatively it is usually more convenient to maximise the natural logarithm of L, in which case the multiplicative constant C becomes an arbitrary additive constant.

Estimates of parameters using this procedure have large sample properties which are desirable in the sense that they are approximately unbiased and normally distributed and no other estimate has a smaller limiting value of [(sample size) × (variance)] [10].

It is a relatively straightforward procedure to write a computer program to find the maximum likelihood estimates V_{50}^* and σ^*, and thus enable the straight line representing the corresponding best fit cumulative normal distribution to be drawn, on probabilistic paper, for a given data set. A listing of a simple algorithm in Fortran 77 is given in Appendix 7.3.

A measure of how well the estimated distribution fits the observed breakdown frequencies can be obtained from a χ-square test which, for example, in the multilevel procedure takes the form [3]

$$\chi^2 = \sum_{k=1}^{N} \frac{(b_k - n_k P_k)^2}{n_k P_k(1 - P_k)} \quad (5)$$

where the P_ks are obtained from eqn. 1 with $V_{50} = V_{50}^*$ and $\sigma = \sigma^*$. The number of degrees of freedom is $\gamma = N - 2$ and the parameter $h = \chi^2/\gamma$ is expected to be of order of, or less than, unity. Values of h appreciably greater than 1 either indicate that the assumed distribution is inappropriate, or that a particular data set may not be reliable and that care must be taken in interpreting the results.

In any testing it would seem highly desirable to carry out a χ-square check on the results as they are taken in order to verify that consistent data have been obtained.

3 Confidence region and confidence intervals

In his studies Brown [3] compared several different methods, all based on the concept of variances, for estimating the confidence interval associated with a single distribution parameter. Carrara and Yakov [6] pointed out that the likelihood approach could be extended to determine the confidence region associated with the simultaneous evaluation of the distribution parameters. They concluded, however, that further work was necessary to put their semiempirical procedure on to a sound statistical basis.

In the following it is shown how the likelihood function may be used to estimate both the confidence region and confidence intervals associated with the maximum likelihood estimates evaluated as in Section 2.

3.1 The confidence region for V_{50}^* and σ^*

The P% confidence region for the maximum likelihood estimators V_{50}^* and σ^* is defined by the boundary in the V_{50}/σ plane for which there is a P% probability that the true values of V_{50} and σ lie within this boundary. It can be demonstrated [10–12] that the inequality

$$\frac{L(y/H(V_{50}, \sigma))}{L(y/H(V_{50}^*, \sigma^*))} > K \quad (6)$$

defines the $100(1 - \alpha)$ per cent confidence region for the hypothesis $H(V_{50}^*, \sigma^*)$, provided that the constant K is chosen to correspond to a confidence level of the likelihood ratio (i.e. the left-hand side of eqn. 6) equal to $(1 - \alpha)$.

The determination of K is not straightforward and in [6], for example, it was suggested on semiempirical grounds that a value of K equal to 0.2 was appropriate for a 90% confidence region. There is, however, a well established theorem [12, 13] that in the limit of large samples

$$2[\ln L(y/H^*) - \ln L(y/H)] \quad (7)$$

may be approximated by a limiting χ-square distribution $\chi^2(r)$ with a degree of freedom, r, equal to the number of parameters in the hypothesis H. If we denote by $C(\alpha)$ the value of $\chi^2(r)$ corresponding to a probability of $(1 - \alpha)$, the $100(1 - \alpha)$ per cent confidence region will be given by

$$2 \ln \frac{L(y/H(V_{50}^*, \sigma^*))}{L(y/H(V_{50}, \sigma))} < C(\alpha) \quad (8)$$

that is

$$\frac{L(y/H(V_{50}, \sigma))}{L(y/H(V_{50}^*, \sigma^*))} > e^{-C(\alpha)/2} \quad (9)$$

so that K may be identified with the corresponding value of $e^{-C(\alpha)/2}$ and is easily calculated from standard χ-square tables.

Table 1 shows the values of K for the 90%, 95% and 99% confidence regions at various degrees of freedom.

Table 1: Values of *K* in relation (6) corresponding to different confidence levels (1 − α)

	Degrees of freedom r	90% (α = 0.1)	95% (α = 0.05)	99% (α = 0.01)
K	1	0.258	0.1465	0.03625
	2	0.100	0.0500	0.01000
	3	0.044	0.0200	0.00343

We thus see that the locus of points in the V_{50}/σ plane for which the likelihood ratio is equal to 0.1 defines a 90% confidence region for the simultaneous estimates V_{50}^* and σ^*. This value of the likelihood ratio was also employed by Eriksson *et al.* [14]. For large samples this confidence region will have an elliptic shape and Fig. 2 shows the results obtained for the typical set of flashover data, Table 2, taken from experimental results reported in

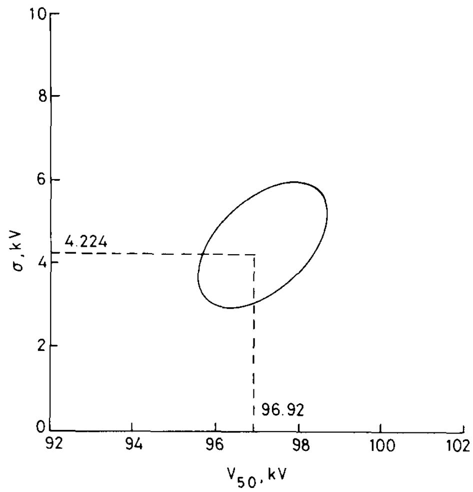

Fig. 2 *Maximum likelihood estimation of V_{50} (=96.92 kV), σ (=4.224 kV) and the associated 90% confidence region for the typical data set of Table 2*

Table 2: Typical set of flashover data [15]

Crest voltage, kV	Number of breakdowns (out of 20 impulses)
87	0
90	2
92	1
94	5
96	8
102	18

a companion paper [15], in which the generalised likelihood method is extensively applied.

3.2 Confidence intervals for the individual parameters V_{50}^* and σ^*

In many circumstances it is more convenient to evaluate confidence intervals for the individual parameters V_{50}^* and σ^* since they are more easily presented in tabular form. Brown has, for example, made a detailed examination of the various methods of determining the confidence intervals from the variances of the maximum likelihood estimators of V_{50} and σ [3].

The simplest (and most optimistic) method assumes that the error in the estimation of a parameter is normally distributed and that the sample is sufficiently large for tolerance limits to be used. Thus, for example, the $100(1-\alpha)\%$ confidence interval for V_{50}^* is defined by

$$V_{50}^* \pm t\sqrt{v_{V_{50}^*}} \tag{10}$$

where t is the $100(1-\alpha)\%$ tolerance limit and $v_{V_{50}^*}$ the variance of V_{50}^* (see Appendixes 7.1 and 7.2).

It is also possible, however, to estimate the confidence interval for a single parameter from the likelihood function [10, 12]. In this case the generalised likelihood ratio [10] must be used since the hypothesis concerning the distribution and the parameters is composite rather than simple. A hypothesis is termed simple when it defines the distribution of the random variable exactly, otherwise it is termed composite. Thus, for example, the hypothesis $H(V_{50}, \sigma)$ of a normal distribution with the parameters equal to certain specific values is simple, whereas the hypothesis $H(V_{50})$ of a normal distribution with V_{50} specified but σ unspecified is composite.

In order to establish a confidence interval for V_{50}^*, the maximum value of $L(y/H(V_{50}))$, obtained by varying σ, may be compared with the maximum of the likelihood function $L(y/H(V_{50}^*, \sigma^*))$ and the generalised likelihood ratio test becomes

$$\frac{\max L(y/H(V_{50}))}{L(y/H(V_{50}^*, \sigma^*))} > K, \tag{11}$$

where the limiting χ-square distribution for the 2 log-likelihood ratio now has only one degree of freedom because only one parameter (V_{50}) is being tested. $L(y/H(V_{50}))$ is now proportional to the probability that the given data set y will occur on the hypothesis, $H(V_{50})$, that V_{50} takes a particular value with σ unspecified and arbitrary.

For one degree of freedom we see from Table 1 that a 90% confidence interval would be obtained with $K = 0.258$, and a 95% confidence interval with $K = 0.1465$.

The same procedure may be followed to establish an appropriate confidence interval for σ.

Table 3: Maximum likelihood estimates and condidence intervals for the typical data set of Table 2 as computed from the variances (*a*) and the generalised likelihood ratio (*b*)

	Maximum likelihood estimaters, kV	Confidence intervals, kV 90% *a*	90% *b*	95% *a*	95% *b*
V_{50}	96.92	95.7–98.1	95.8–98.2	95.5–98.4	95.5–98.6
σ	4.224	3.0–5.4	3.1–5.5	2.8–5.6	2.9–5.8

In Table 3 the 90% and 95% confidence intervals of V_{50}^* and σ^* computed both from the variances and the generalised likelihood test are compared and very good agreement is obtained between the two methods.

It should be noted that the simultaneous confidence intervals of the two parameters define a rectangle in the V_{50}/σ plane which could also be regarded as a confidence region [10], but that simultaneous $100(1-\alpha)\%$ intervals lead to a $100(1-\alpha)^2\%$ confidence region (e.g. simultaneous 90% confidence intervals lead to a 81% confidence region). The rectangular regions corresponding to simultaneous confidence intervals of 90% and 95% for the typical data set are compared in Fig. 3, with the 90% confidence region determined from the likelihood ratio.

Carrara and Yakov [6] estimated the 90% confidence interval associated with a single parameter by drawing the extreme tangents to the corresponding elliptic confidence region. We see from Fig. 3 that this can lead to

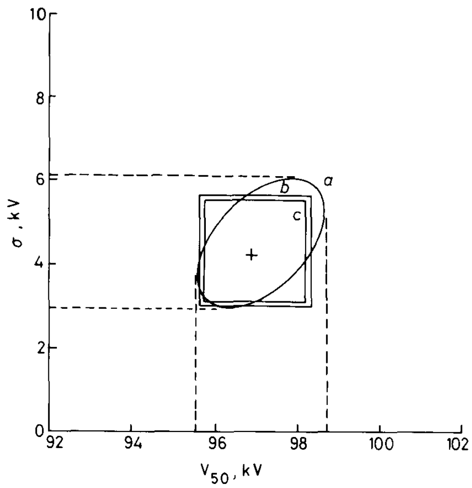

Fig. 3 *Confidence regions and intervals*
a 90% confidence region from likelihood ratio
b region in σ/V_{50} plane defined by simultaneous 95% confidence intervals
c region in σ/V_{50} plane defined by simultaneous 90% confidence intervals
- - - - intervals defined by extreme tangents to the 90% confidence region
(Data as **Table 2**)

very pessimistic estimates, since the intervals defined by the tangents to the 90% confidence region are even greater than the 95% confidence intervals computed either with the variances or the generalised likelihood ratio.

3.3 Estimation of other parameters of the distribution

The methods described in Sections 2, 3.1 and 3.2 may be applied to any two-parameter distribution and are easily expanded in principle to n-parameter distributions.

In addition to estimating V_{50} and σ the method of maximum likelihood may also be employed for the estimation of other parameters of interest, such as voltage levels corresponding to low or high breakdown probabilities, and the generalised likelihood ratio used to determine the corresponding confidence intervals.

Suppose, again for a cumulative normal distribution, that we wish to determine the maximum likelihood estimate of the voltage level having a 5% probability of breakdown V_5^*. This can be achieved in two ways. First, one may maximise the likelihood function, where L is now expressed as a function of V_5 and any other parameter, e.g. σ. In this case L can be written

$$L(y/H(V_5, \sigma)) = C \prod_{k=1}^{N} P_k^{b_k}(1 - P_k)^{n_k - b_k} \quad (12)$$

where P_k is given by

$$P_k = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{V_k} \exp\{[V - (V_5 + 1.65\sigma)]^2/2\sigma^2\}\, dV \quad (13)$$

Alternatively, since the cumulative normal distribution has been assumed to represent the data, V_5^* may be computed from

$$V_5^* = V_{50}^* - 1.65\sigma^* \quad (14)$$

provided that V_{50}^* and σ^* have already been determined.

It can easily be demonstrated that both these procedures lead to the same value of V_5^* and that the maximum value of the likelihood function is also the same.

The confidence intervals associated with V_5 can be estimated from the generalised likelihood ratio which now becomes

$$\frac{\max L(y/H(V_5))}{L(y/H(V_{50}^*, \sigma^*))} > K \quad (15)$$

where the limiting χ-square distribution has one degree of freedom. Table 4 shows the maximum likelihood estimates of several voltage levels at high and low breakdown probabilities together with the 90% confidence intervals determined both from eqn. 15 and from using the concept of tolerance limits [3] for the assumed data set of Table 2.

Table 4: Confidence intervals associated with various voltage levels as found from (*a*) variances and (*b*) generalised likelihood ratio. Data as in Table 2

	Maximum likelihood estimates kV	90% confidence intervals, kV	
		a	*b*
V_1	87.1	84.8–89.2	84.8–89.2
V_5	90.0	88.2–91.5	88.3–91.5
V_{10}	91.2	90.2–92.8	90.1–92.8
V_{90}	102.4	100.5–104.7	100.4–104.8
V_{95}	103.9	102.2–106.6	101.6–106.7
V_{99}	106.8	104.7–109.9	104.0–110.0

Very good agreement is obtained between the two procedures, but it should be noted that the use of tolerance limits implies the assumption of a normally distributed error in the estimates (which, however, is usually a good assumption) and large samples, while use of the generalised likelihood ratio requires only large samples.

In principle the above method could be used to estimate the voltage levels, confidence regions and intervals associated with any breakdown probability. It should be remembered, however, that the form of the probability distribution function must be specified and that extrapolation to very low or very high breakdown probability levels can only be carried out reliably provided that the assumed distribution remains valid at these levels.

Following the procedure adopted by Brown [3], it is also possible to estimate from the given data set the confidence interval associated with the breakdown probability at any voltage level. Assuming that a normal breakdown probability distribution applies and that the only experimental errors are in the estimates of the breakdown frequency (these errors having a binomial distribution), the method of variances may be applied to compute the 90% confidence interval for the breakdown probability (eqn. 24 in Appendix 7.1). Thus at the voltage level 87.1 kV, corresponding to the maximum likelihood estimate of V_1 for the data set of Table 2, $P_B(V)$ is equal to 1% and the 90% confidence interval on $P_B(V)$ is 0%–3%. If a narrower confidence interval on $P_B(V)$ is required then the number of shots in the data set must be greatly increased, especially at low probability levels.

4 Conclusions

The establishment of a standard procedure for the statistical analysis of flashover data is highly desirable since different procedures can lead to quite different results for the same data.

The likelihood approach is a very powerful method for analysing such data, and not only estimates the parameters of the breakdown probability distribution but also their associated confidence regions and intervals. In order to apply the analysis, however, the form of the breakdown probability distribution must be specified and the approach does not provide a criterion for comparing different possible distributions [9].

As previously suggested [6], the likelihood analysis should be considered for adoption as a standard procedure and be incorporated in the appropriate IEC Standard.

Provided that the impulses are independent, the method can treat both Class I and Class III testing procedures, and is uniformly applicable to any assumed breakdown probability distribution.

In all testing it is highly desirable that a χ-square test is carried out in order to check the consistency of the test data.

Finally, it should be remembered that, although the likelihood analysis provides the best estimates of the parameters of a distribution, it is the responsibility of the investigator to ensure that the optimum testing strategy is adopted to minimise the confidence intervals for the desired parameters. The likelihood method provides a very convenient means of comparing different strategies.

5 Acknowledgments

This work has been carried out with the aid of a Research Grant GR/B/4892.5 from the UK Science and Engineering Research Council. We are grateful to members of the EEC Group on Energy Exchange Processes in Electrical Discharges for much helpful discussion.

6 References

1 'High voltage testing techniques. Part 2: Test procedures'. IEC Publication 60-2: First edition 1973
2 SUZUKI, T., KISHIJIMA, J., OHUCH, Y., and ANJO, K.: 'Parallel multigap flashover probability', *IEEE Trans.*, 1969, **PAS-88**, pp. 1814–1821

3 BROWN, G.B.: 'Method of maximum likelihood applied to the analysis of flashover data', *IEEE Trans.*, 1969, **PAS-88**, pp. 1823–1830
4 HYLTEN-CAVALLIUS, N., and CHAGES, F.A.: 'Possible precision of statistical insulation test methods', *IEEE Trans.*, 1983, **PAS-102**, pp. 2372–2376
5 TRINH, N.G., and VINCENT, C.: 'Statistical significance of test methods for low probability breakdown and withstand voltages', *IEEE Trans.*, 1980, **PAS-99**, pp. 711–719
6 CARRARA, G., and YAKOV, Y.: 'Statistical evaluation of dielectric test methods', CESI Technical Issue, L'Energia Eletrica, 1983, 1
7 CARRARA, G., and DELLERA, L.: 'Accuracy of an extended up-and-down method in statistical testing of insulation', *Electra*, 1972, **23**
8 JONES, B., and WATERS, R.T.: 'Air insulation at large spacings', *IEE Proc.*, 1978, **125**, pp. 1152–1176
9 EDWARDS, A.W.F.: 'Likelihood' (Cambridge University Press, 1972)
10 JOHNSON, N.L., and LEONE, F.C.: 'Statistics and experimental design' (John Wiley & Son, Inc., New York, 1964)
11 COOPER, B.E.: 'Statistics for experimentalists' (Pergamon Press, Oxford, 1975)
12 FRASER, D.A.S.: 'Probability and statistics: theory and applications' (Duxbury Press, Massachussetts, 1976)
13 BREIMAN, L.: 'Statistics with a view towards applications' (Houghton Mifflin Co., Boston, 1973)
14 ERIKSSON, A.J., Le ROUX, B.C., GELDENHUYS, H.J., and MEAL, D.V.: 'Study of airgap breakdown characteristics under ambient conditions of reduced air density', *IEE Proc. A*, 1986, **133**, (8), pp. 485–492
15 DAVIES, A.J., DUTTON, J., TURRI, R., and WATERS, R.T.: 'Effects of humidity and gas density on switching impulse breakdown of short airgaps', *IEE Proc. A*, 1988, **135**, (1), pp. 59–68

7 Appendix

7.1 Variances of distribution parameters

Brown [3] has shown that the variances of the maximum likelihood estimates of the parameters (assuming a cumulative normal distribution) may be calculated as follows:

variance of V_{50}

$$v_{V_{50}} = \frac{\sigma^2}{D} \sum_{k=1}^{N} r_k Y_k^2 \tag{16}$$

variance of σ

$$v_\sigma = \frac{\sigma^2}{D} \sum_{k=1}^{N} r_k \tag{17}$$

where

$$Y_k = \frac{V_k - V_{50}}{\sigma} \tag{18}$$

$$Z_k = \frac{e^{-Y_k^2/2}}{\sqrt{2\pi}} \tag{19}$$

$$P_k = \int_{-\infty}^{Y_k} Z_k \, dY_k \tag{20}$$

$$r_k = \frac{n_k Z_k^2}{P_k(1 - P_k)} \tag{21}$$

$$D = \left(\sum_{k=1}^{N} r_k Y_k^2\right)\left(\sum_{k=1}^{N} r_k\right) - \left(\sum_{k=1}^{N} r_k Y_k\right)^2 \tag{22}$$

The variances depend on the true values of V_{50} and σ but are estimated by using the values V_{50}^* and σ^*.

Similar expressions to eqns. 16 and 17 will give the variances of other variables of interest. For example the variance of the estimator of the voltage V_p corresponding to a P% breakdown probability is given by

$$v_{V_p} = \frac{\sigma^2}{D} \sum_{k=1}^{N} r_k (Y_k - Y_p)^2 \tag{23}$$

and

$$v_{P_B} = \frac{Z_B^2}{D} \sum_{k=1}^{N} r_k (Y_k - Y_B)^2 \tag{24}$$

is the variance of the estimate of the probability of breakdown P_B at a voltage level V_B, where

$$Y_B = \frac{V_B - V_{50}}{\sigma} \tag{25}$$

7.2 Definition of tolerance limits

Suppose that a certain variable x is normally distrubuted with a known mean value, μ, and a standard deviation σ. Then the probability that a measurement x will lie within t standard deviations from the mean is

$$P = \frac{100}{\sqrt{2\pi}\,\sigma} \int_{-t}^{t} \exp\left[-(x - \mu)^2/2\sigma^2\right] dx \text{ per cent} \tag{26}$$

where the limits $\pm t$ are called the P% tolerance limits.

7.3 Fortran 77 program listing

```
      PROGRAM LHOOD
C
C
C
C  A PROGRAM TO PERFORM STATISTICAL ANALYSIS OF
C  FLASHOVER DATA USING A MAXIMUM LIKELIHOOD
C  METHOD
C
C
C
C  THE PROGRAM CALLS FIVE SUBROUTINES AS FOLLOWS:
C
C  1. INPUT
C       ALLOWS THE USER TO SUPPLY THE NUMBER OF
C       VOLTAGE LEVELS AND FOR EACH LEVEL THE
C       TEST VOLTAGE, THE NUMBER OF BREAKDOWNS
C       AND NUMBER OF WITHSTANDS
C
C  2. ESTIM
C       OBTAINS ESTIMATES FOR THE PARAMETERS WHICH
C       BEST FIT THE CHOSEN DISTRIBUTION
C
C  3. CHIS
C       CALCULATES H=CHI**2/(NO. OF DEG. OF FREEDOM)
C       AS A CHECK ON THE RELIABILITY OF THE RESULTS
C
C  4. CONREG
C       PLOTS THE CONFIDENCE REGION FOR A SELECTED
C       CONFIDENCE LEVEL
C
C  5. CONINT
C       CALCULATES CONFIDENCE INTERVALS USING THE
C       GENERALISED LIKELIHOOD RATIO
C
C
C
C  VARIABLES IN COMMON FOR ALL ROUTINES ARE:
C
C  M          NUMBER OF VOLTAGE LEVELS
C  V50        ESTIMATE OF V50
C  SIG        ESTIMATE OF SIGMA
C  U(1000)    TEST VOLTAGES
C  ID(1000)   NUMBER OF BREAKDOWNS
C  IW(1000)   NUMBER OF WITHSTANDS
C
C
C
C  THE PROGRAM USES A NUMBER OF NAG ROUTINES
C  AND NAG GRAPHICAL ROUTINES.
C
C  DOUBLE PRECISION DECLARATIONS MAY NOT BE
C  REQUIRED FOR SOME IMPLEMENTATIONS
C
C
C
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      EXTERNAL HEIGHT
      COMMON M,V50,SIG,U(1000),ID(1000),IW(1000)
      CALL INPUT
      CALL ESTIM
      CALL CHIS
      CALL CONREG
      CALL CONINT
      STOP
      END
```