NASA
Technical
Paper
2252

January 1984

Statistical Analysis of Direct-Strike Lightning Data (1980 to 1982)

Larry D. Lee,
George B. Finelli,
Mitchel E. Thomas,
and Felix L. Pitts

NASA

NASA
Technical
Paper
2252

1984

Statistical Analysis of Direct-Strike Lightning Data (1980 to 1982)

Larry D. Lee,
George B. Finelli,
Mitchel E. Thomas,
and Felix L. Pitts

Langley Research Center
Hampton, Virginia

NASA
National Aeronautics
and Space Administration

**Scientific and Technical
Information Office**

1984

INTRODUCTION

The objective of the NASA Langley Research Center direct-strike lightning research is to aid in characterization of the direct-strike lightning environment affecting aircraft. The research is motivated by the need for refined characterization of the lightning-aircraft environment in support of advanced aircraft designs that contain composite structures and digital electronic systems.

A specially instrumented F-106B aircraft is used in the research to penetrate thunderstorms and elicit direct strikes. Measurements of the lightning-aircraft electromagnetic interaction process are made with sensors derived from those designs developed for nuclear electromagnetic pulse measurements. Nine sensors with bandwidths greater than 100 MHz are located about the aircraft and measure the rates of change of magnetic and electric flux density and current in the nose boom. Two transient recorders with bandwidths of 50 MHz each (100 MHz sample rates) record the output of two selected sensors, and the trigger configuration allows time correlation of simultaneously recorded waveforms. Two other transient recorders, triggered independently and operating at a sample rate of 60 MHz, record nose-boom and vertical tail-fin currents. An airborne field mill system gives a continuous indication of charge polarity on the aircraft as well as vertical and lateral electric fields from dc to 10 Hz. (A list of symbols used in this paper appears after the references.) Photographic time histories of lightning attachments to one wing and to the vertical tail have been acquired with a 16-mm movie camera with a wide-angle lens. A time-code generator enables time correlation of 1 msec between the photographic data and other recorded aircraft and ground-based measurements. Ground-based measurements include fast and slow electric field changes and radio frequency spherics generated by lightning in storms penetrated by the aircraft.

The instrumentation system, transient recorders, sensors, and a summary overview of the tests are presented in references 1 to 5. Interpretation, analysis, and generalization of the direct-strike data involves a methodology for deriving the lightning source from the response measurements using electromagnetic coupling codes as described in reference 6 along with time-domain reflectometry techniques using laboratory aircraft scale models for resonant frequency identification and comparison with flight data as described in references 7 and 8. The complete set of waveforms recorded in 1980, 1981, and 1982 are reported in references 9, 10, and 11.

Data from 10 strikes to the aircraft were recorded during each of the first 2 years of tests in 1980 and 1981. In 1982, data were recorded from 156 strikes. The significant increase in the number of strikes in 1982 was brought about by improved operational procedures. An aircraft fuel system modification provided an additional 1/2-hour mission time for a total mission time of 1 1/2 hours, the penetrations were made at generally higher altitudes of 25 000 to 35 000 ft, and guidance to specific altitudes of lightning activity was obtained based on UHF radar returns from lightning channel ionization.

The measurements made with the direct-strike data system are nose-boom current I and its rate of change $\dot{I}$, rate of change of electric and magnetic flux density $\dot{D}$ and $\dot{B}$, and static electric field E. The nominal sensor locations are shown in figure 1. The I and $\dot{I}$ sensors are located at the base of the nose boom at position 1; $\dot{D}$ sensors are located under the fuselage at position 2, under the wings at

positions 9 and 11, and on the port side of the vertical fin at position 12 corresponding to the respective sign conventions for D_F, $D_{W,L}$, $D_{W,R}$, and D_T; $\dot{B}$ sensors are located at opposite sides of the aft fuselage at positions 6 and 7 and under the wings at positions 8 and 10 corresponding to the respective sign conventions for B_L (starboard side), B_T (port side), $B_{W,L}$, and $B_{W,R}$; and E sensors are located at position 3 corresponding to sign conventions for E_Z and at positions 4 and 5 corresponding to E_Y. The sign conventions established for the measurements shown in figure 1 indicate positive quantities in the direction of the arrows; the relation between the sign conventions of the variables D, B, and I and their derivatives are also shown in figure 1 in the measurement polarity time history. All the data reported and analyzed in this paper are from the $\dot{B}$ sensor at position 6, the $\dot{D}$ sensor at position 2, and I and $\dot{I}$ sensors at position 1.

The statistical analyses and results presented in this report provide insight into the expected ranges of the electromagnetic environment when fighter-class aircraft are struck in thunderstorms at altitudes of 20 000 to 35 000 ft. It should be noted that the statistics characterize the first transient of the lightning process which is greater than the system threshold and which occurs within the first 1300 μsec of the time the threshold is exceeded. Data acquired subsequently but too late for inclusion in the 1982 statistical analysis data base indicate these statistics are actually representative of the entire lightning process. It is known from limited bandwidth continuous recordings that the total lightning process consists of a moderate number of discrete electromagnetic pulses distributed over a second or so (ref. 7). The continuous records indicate the transient recorder threshold is sufficiently high so that the time of its exceedance is distributed over the duration of the lightning process (and not limited to particular time intervals in the process), and wide-band peak detectors which monitor the peak amplitudes of the entire process yielded results commensurate with the 1982 data statistical predictions.

The primary objectives of the statistical analysis are: to identify a location-scale family of distributions consistent with the data, and to use the resulting model to compute estimates of the quantiles. Graphical display of the data suggests a lognormal distribution for some subsets. Statistical significance tests support the hypothesis that the $\dot{D}$ and $\dot{I}$ measurements are lognormally distributed. For these data, and for the $\dot{B}$ measurements, quantiles and confidence limits are established.

The cumulative direct-strike lightning data base includes all the waveforms recorded during 3 years of flight tests (1980 to 1982) with transient recorders. The recorded data are categorized as resulting from direct lightning strikes or from nearby lightning events. If the lightning attachment to the aircraft was observed by the flight crew or by the onboard cameras, or if the current sensors indicated current flow, the lightning event was identified as a "strike." Other lightning occurrences, which were not confirmed as strikes by the above criteria, were termed "nearby" events. The notation to identify data in these two categories for the statistical analyses is as follows: $\dot{B}_N$ and $\dot{D}_N$ identify $\dot{B}$ and $\dot{D}$ data for nearby events, and $\dot{B}_S$ and $\dot{D}_S$ identify $\dot{B}$ and $\dot{D}$ data for strikes; I and $\dot{I}$ data result only from strikes.

STATISTICAL RESULTS

The span (range) of the recorded waveforms was arbitrarily chosen as the random variable in the statistical analyses. The span is the difference between the maximum positive and the maximum negative amplitude of the measured waveforms. Table 1 shows the number of records acquired for each of the six distinct data sets (referred to as samples) and shows the range of span values for each sample.

For each of the samples, frequency-of-occurrence tables were generated to suggest the underlying distribution of the random variable. (See tables 2 to 7.) For each of these tables, the range of the particular data set is broken into 10 equally sized intervals. The lower limit of each interval is shown in the first column. The second column shows the frequency of data values which are greater than or equal to that lower limit but less than the lower limit of the next interval. Samples B_N and D_N were excluded from further analysis since the variable span is dependent on the distance to the lightning source, which was not known. The I sample was also excluded from the following analysis on the basis of insufficient sample size.

Probability Plots

The probability plotting method described in the appendix and known as the Q-Q plot (ref. 12) is used to judge the adequacy of the lognormal and the type II extreme value distribution of largest values. Since this plotting method is useful only for judging the adequacy of location and scale parameter models, the original sample values $x_1, x_2, \ldots, x_n$ are replaced by their natural logarithms, so that $y_i = \ln(x_i)$ for $i = 1, 2, \ldots, n$. This transformation also converts the lognormal and type II extreme value distributions to the normal and type I extreme value distributions, which are location and scale parameter models.

Members of the set $y_1, y_2, \ldots, y_n$ are ordered from smallest to largest, and the ith ordered value is paired with a measure of location of the ith ordered value in a sample from the theoretical distribution. In the plots, departures from linearity suggest that the lognormal and type II extreme value distributions are not adequate models. Tables 8 to 10 are also helpful in judging the adequacy of these distributions when compared with the theoretical distributions.

Figures 2 to 4 show plots of the ordered transformed samples versus the theoretical quantiles of the standard normal distribution. For the D_S and I samples (figs. 3 and 4), the plots are nearly linear, suggesting a good fit to the normal distribution. Thus, the lognormal distribution is an adequate model for the true parent distribution giving rise to the untransformed sample $x_1, x_2, \ldots, x_n$. In the case of the B_S sample (fig. 2), the plot exhibits a concave shape and suggests that negative skewness is a property possessed by the true parent distribution of $y_1, y_2, \ldots, y_n$. The significance test for normality, which is presented later, supports the conclusion that the concave shape is a significant departure from the linear shape expected under normality.

Figures 5 to 7 show plots of the ordered transformed samples versus the theoretical quantiles of the type I extreme value distribution of largest values. Generally, none of these plots indicate that the fit to the extreme value model is better than the fit to the normal distribution model. However, the heavier upper tail (positive skewness) of the extreme value model and a possible link of this model with the physical feature that span values are themselves largest values motivated the a priori selection of the extreme value model.

A significance test for adequacy of the lognormal model can be based on the probability plot correlation coefficient (ref. 13). The small observed significance level (less than 0.005) shown in table 11 for the $\dot{B}_S$ sample suggests a significant departure from the lognormal distribution model. As anticipated from the probability plots, the $\dot{D}_S$ and I samples do not yield significant values of the correlation statistic, indicating the data support the hypothesis that the underlying distributions are lognormal.

Quantile Estimates

Asymptotically unbiased quantile estimators as developed in reference 14 and described in the appendix are applied to estimate the upper quantiles y_ξ of the true parent distribution for $y_1, y_2, \ldots, y_n$. Then the transformation $x_\xi = \exp(y_\xi)$ gives an estimate of the quantile x_ξ of the true parent distribution for $x_1, x_2, \ldots, x_n$. Tables 12, 13, and 14 show the estimator $\hat{x}_\xi$ of the quantile x_ξ, its estimated variance, and the 95-percent lower and upper confidence limits L and U for the $\dot{B}_S$, $\dot{D}_S$, and I samples.

Sample Size

Of particular importance in direct-strike lightning experiments is the problem of estimating the "worst case" values of the distribution of the measurements. The extreme quantiles of most interest require large sample sizes to obtain "good" precision for the estimates. Precision is commonly measured by the length of its associated confidence interval. However, because the theoretical confidence limits have an expected length depending on the unknown quantile, precision as measured by length would also depend on the unknown quantile. To avoid this problem, we examine the effect of sample size on the ratio $R = U/L$. A general representation of the sample size needed to ensure that $(1 - \alpha)100$-percent confidence limits have a ratio not exceeding a specified R_o is given in the appendix.

Table 15 shows how the sample size varies with the required precision as measured by R. The computations were made for 95-percent confidence limits for x_ξ assuming that estimation is according to the normal distribution. For $\dot{D}_S$ samples and for estimating $x_{.95}$, a sample of size $n = 46$ yields an $R_o = 2.0$, whereas for estimating $x_{.99}$ it takes a sample of size $n = 144$ to yield the same precision. Similar conclusions are obtained for other samples by examining table 15.

CONCLUDING REMARKS

Statistical methods have been applied in the analysis and interpretation of electromagnetic data acquired during 176 direct strikes to the F-106B aircraft. Probability plots and significance tests suggest that the data for the rate of change of electric flux density $\dot{D}$ and for the current I are lognormally distributed. The data for the rate of change of magnetic flux density $\dot{B}$, on the other hand, appear more uniformly distributed. An explanation is that truncation occurred in the measurement process - some of the data points in the upper interval are at the full-scale limits of the measurement system, for example.

Techniques were presented for estimating the quantiles of the distributions of $\dot{B}$, $\dot{D}$, and I measurements. The estimate of the theoretical 90th quantile of the distribution yielding the $\dot{B}$ sample for direct lightning strikes is 1497.09 T/sec,

and 95-percent confidence limits for this quantile range from 1075.83 T/sec to 2083.31 T/sec. Similarly, for $\dot{D}$ the estimate is 16.10 A/m^2, with limits of 13.23 A/m^2 and 19.58 A/m^2, and for $\dot{I}$ the estimate is 9156.66 A, with limits of 7032.07 A and 11 923.15 A. The largest observed values for direct strikes were 1727 T/sec, 30.60 A/m^2, and 13 900 A for $\dot{B}$, $\dot{D}$, and $\dot{I}$. An insufficient number of samples for rate of change of current $\dot{I}$ measurements for statistical analysis have been recorded to date. The maximum zero-to-peak $\dot{I}$ recorded, however, exceeded the full-scale peak of 48 GA/sec; the corresponding maximum $\dot{I}$ derived graphically from the leading edge of the simultaneously recorded current waveform was 130 GA/sec.

Langley Research Center
National Aeronautics and Space Administration
Hampton, VA 23665
November 23, 1983

APPENDIX

STATISTICAL DATA ANALYSIS

Probability Plotting Method

The estimation method used assumes that the parent distribution for $y_1, y_2, \ldots, y_n$ belongs to some location-scale family. Such a family has the representation

$$G(y) = G_o((y-\eta)/\delta)$$

with density

$$g(y) = g_o((y-\eta)/\delta)\ \delta^{-1}$$

where $G_o(z)$ is a known distribution, and η and δ are unknown location and scale parameters. We refer to the generator distribution $G_o(z)$ as a standardized distribution because it is the distribution of the standardized variable $Z = (Y - \eta)/\delta$ and serves to specify the location-scale family.

Let Z be a standardized variable having distribution $G_o(z)$. Let $G(y)$ be the true distribution of $Y = \ln(X)$. If $G_o(z)$ and $G(y)$ differ only by location and scale parameters, then a plot of the theoretical z-quantiles versus the theoretical y-quantiles should yield a linear plot. Departures from linearity indicate that $G(y)$ does not belong to the family defined by $G_o(z)$.

The theoretical versus empirical plot is constructed from the pairs $(M_i, y_{(i)})$ for $i = 1, 2, \ldots, n$, where $y_{(1)} \leq y_{(2)} \leq \ldots \leq y_{(n)}$ is the ordered sample and $M_1, M_2, \ldots, M_n$ are the corresponding theoretical quantiles of $G_o(z)$. The plotting positions suggested in reference 13 are calculated with $M_i = G_o^{-1}(m_i)$, where

$$m_i = 1 - m_n \qquad (i = 1)$$

$$m_i = (i - 0.3175)/(n + 0.3650) \qquad (i = 2, 3, \ldots, n - 1)$$

$$m_i = (0.5)^{1/n} \qquad (i = n)$$

The plots shown in figures 2 to 4 are constructed by plotting the ordered transformed sample values against the theoretical quantiles of the standard normal distribution

$$G_o(z) = (2\pi)^{-1/2} \int_{-\infty}^{z} \exp(-w^2/2)\ dw$$

APPENDIX

Those plots shown in figures 5 to 7 are constructed with the theoretical quantiles of the type I extreme value distribution of largest values, $G_o(z) = \exp(-e^{-z})$ for $-\infty < z < \infty$. Whenever a linear plot suggests a good fit to the family of distributions defined by $G_o(z)$, the distribution $F(x) = G_o([\ln(x) - \eta]/\delta)$ for $x > 0$ is judged to be consistent with the original sample $x_1, x_2, \ldots, x_n$. For example, a standard normal distribution $G_o(z)$ leads to a lognormal distribution $F(x)$, whereas the type I extreme value distribution $G_o(z)$ gives a type II extreme value distribution $F(x)$.

The probability plot correlation coefficient (ref. 13) used earlier for testing the adequacy for the lognormal model is

$$r = \sum_{1}^{n} M_i y_{(i)} / C_n q$$

where

$$C_n = \left(\sum_{1}^{n} M_i^2 \right)^{1/2}$$

and

$$q = \left[\sum_{1}^{n} (y_i - \bar{y})^2 \right]^{1/2}$$

With this test the normal distribution (for $y_1, y_2, \ldots, y_n$) is rejected whenever r takes a value in the lower tail of its sampling distribution (r is always positive). Percentage points needed to define the upper limits of the rejection regions are available in reference 13 for various significance levels. A similar test is available (ref. 15) for testing the adequacy of the extreme value model.

Quantile Estimators

Let $y_i = \ln(x_i)$ for $i = 1, 2, \ldots, n$ denote the transformed sample values and let y_ξ denote the ξth quantile of the true distribution $G(y)$. Asymptotically unbiased estimators originally proposed in reference 14 are described below.

The estimator $\hat{y}_\xi$ of y_ξ is computed from the ordered sample $y_{(1)} \leq y_{(2)} \leq \ldots \leq y_{(n)}$ as the linear function

$$\hat{y}_\xi = c y_{(M)} + (1 - c) y_{(L)}$$

where $y_{(L)} < y_{(M)}$ are chosen from the ordered sample with indices

$$L = [nq] + 1$$

$$M = [np] + 1$$

and [np] and [nq] are the integral parts of the real values np and nq. For a specified ξ, the quantities $0 < q < \xi < p < 1$ and $0 < c < 1$ are chosen so that the variance of $\hat{y}_\xi$ is minimized. These quantities are reproduced from reference 14 in table A1 for the case of a parent normal distribution.

The variance τ^2 of the estimator $\hat{y}_\xi$ from reference 14 is $\tau^2 = (\delta^2/n)\, V(p,q)$, where

$$V(p,q) = c^2 p \frac{1 - p}{g_o^2(z_p)} + (1 - c)^2 q \frac{1 - q}{g_o^2(z_q)} + 2c(1 - c)q \frac{1 - p}{g_o(z_p)\, g_o(z_q)}$$

and $g_o(z)$ is the density function of the standardized variable $Z = (Y - \eta)/\delta$; z_p and z_q are theoretical quantiles of the distribution $G_o(z)$.

In the case of a normal distribution, $g_o(z)$ is the standard normal density and z_p and z_q are quantiles of the standard normal distribution. For this case, the estimator $\hat{\tau}^2$ of τ^2 is computed by replacing δ^2 by its estimator, the sample variance S_y^2, which is computed from $y_1, y_2, \ldots, y_n$. Computed values of $V(p,q)$ are shown in table A1.

The transformation $\hat{x}_\xi = \exp(\hat{y}_\xi)$ converts the estimator $\hat{y}_\xi$ of y_ξ to an estimator of the quantile x_ξ of the distribution $F(x)$. The approximate variance of $\hat{x}_\xi$ shown in tables 3 to 5 is computed with $\operatorname{var}(\hat{x}_\xi) = \tau^2 \exp(2\hat{y}_\xi)$, with the approximate form obtained from a Taylor series expansion of $\exp(\hat{y}_\xi)$. (See method discussed on p. 302 of ref. 16.) Confidence limits for x_ξ follow from the result that $\hat{\tau}^{-1} \ln(\hat{x}_\xi/x_\xi) = \hat{\tau}^{-1}(\hat{y}_\xi - y_\xi)$ has a limiting normal distribution (as $n \to \infty$) with mean equal to zero and variance equal to one (ref. 14). The upper and lower limits U and L for a $(1 - \alpha)100$-percent confidence interval are

$$L = \hat{x}_\xi \exp(-z_{(1-\alpha)/2}\hat{\tau})$$

$$U = \hat{x}_\xi \exp(z_{(1-\alpha)/2}\hat{\tau})$$

where $z_{(1-\alpha)/2}$ is the upper $(1 - \alpha)/2$ percentage point of the standard normal distribution.

APPENDIX

Sample Size

Let $R = U/L$. From the preceding representation of L and U, $\ln^2 R$ is a random variable having expected value $4z^2_{(1-\alpha)/2}\tau^2 = 4z^2_{(1-\alpha)/2}(\delta^2/n)\ V(p,q)$. The sample size needed to ensure that this expected value does not exceed a prescribed value $\ln^2 R_o$ is the smallest n satisfying $4z^2_{(1-\alpha)/2}(\delta^2/n)\ V(p,q) \leq \ln^2 R_o$. The solution is

$$n = 4z^2_{(1-\alpha)/2}\delta^2\ V(p,q)\ \ln^{-2} R_o$$

The sample size depends on the index ξ but not on the parameter x_ξ. It also depends on the level of confidence. An estimate of the scale parameter δ is needed to compute n.

The sample size for a required R increases with increasing ξ. This is not surprising because the samples yield relatively little information concerning distributional shape at the tails. For less extreme quantiles ($\xi < 0.95$), the confidence limits are narrower and the estimators have good asymptotic relative efficiencies (ref. 14).

TABLE A1.- OPTIMAL SELECTION OF INDICES FOR PARENT NORMAL DISTRIBUTION

ξ	c (a)	p (a)	q (a)	z_p	z_q	V(p,q)
0.99	0.302	0.997	0.984	2.75	2.15	11.44
.95	.313	.985	.920	2.17	1.41	3.64
.90	.324	.970	.840	1.88	.99	2.36
.80	.371	.930	.680	1.48	.47	1.63

[a]From reference 14.

REFERENCES

1. Thomas, M. E.: Direct Strike Lightning Measurement System. AIAA-81-2513, Nov. 1981.

2. Thomas, Robert M., Jr.: Expanded Interleaved Solid-State Memory for a Wide Bandwidth Transient Waveform Recorder. Lightning Technology, NASA CP-2128, FAA-RD-80-30, 1980, pp. 119-129.

3. Von Bokern, G. J.; Piszker, L. D.; and Brick, R. O.: In-Flight Lightning Data Measurement System for Fleet Application. Federal Aviation Administration - Georgia Institute of Technology Workshop on Grounding and Lightning Protection, FAA-RD-78-83, May 1978, pp. 345-363.

4. Trost, Thomas F.; and Zaepfel, Klaus P.: Broadband Electromagnetic Sensors for Aircraft Lightning Research. Lightning Technology, NASA CP-2128, FAA-RD-80-30, 1980, pp. 131-152.

5. Pitts, Felix L.: Electromagnetic Measurement of Lightning Strikes to Aircraft. J. Aircr., vol. 19, no. 3, Mar. 1982, pp. 246-250. (Available as AIAA-81-0083R.)

6. Rudolph, Terence; and Perala, Rodney A.: Interpretation Methodology and Analysis of In-Flight Lightning Data. NASA CR-3590, 1982.

7. Trost, Thomas F.; and Pitts, Felix L.: Analysis of Electromagnetic Fields on an F-106B Aircraft During Lightning Strikes. Proceedings of the International Aerospace Conference on Lightning and Static Electricity - Volume 1, Public Relations Section, Culham Lab., Mar. 1982, pp. B3-1 - B3-11.

8. Turner, C. D.; and Trost, T. F.: Laboratory Modeling and Analysis of Aircraft-Lightning Interactions. NASA Grant NAG-1-28, Texas Tech Univ., Aug. 1982. (Available as NASA CR-169455.)

9. Pitts, Felix L.; and Thomas, Mitchel E.: 1980 Direct Strike Lightning Data. NASA TM-81946, 1981.

10. Pitts, Felix L.; and Thomas, Mitchel E.: 1981 Direct Strike Lightning Data. NASA TM-83273, 1982.

11. Thomas, Mitchel E.; and Pitts, Felix L.: 1982 Direct Strike Lightning Data. NASA TM-84626, 1983.

12. Wilk, M. B.; and Gnanadesikan, R.: Probability Plotting Methods for the Analysis of Data. Biometrika, vol. 55, no. 1, 1968, pp. 1-17.

13. Filliben, James J.: The Probability Plot Correlation Coefficient Test for Normality. Technometrics, vol. 17, no. 1, Feb. 1975, pp. 111-117.

14. Kubat, Peter; and Epstein, Benjamin: Estimation of Quantiles of Location-Scale Distributions Based on Two or Three Order Statistics. Technometrics, vol. 22, no. 4, Nov. 1980, pp. 575-581.

15. Gerlach, B.: A Consistent Correlation-Type Goodness-of-Fit Test; With Application to the Two-Parameter Weibull Distribution. Math. Operationsforsch. Statist., Ser. Statist., vol. 10, no. 3, 1979, pp. 427-452.

16. Cox, D. R.; and Hinkley, D. V.: Theoretical Statistics. Chapman & Hall, Ltd., 1974.

SYMBOLS

B	magnetic flux density, T
$\dot{B}$	rate of change of magnetic flux density, T/sec
$\dot{B}_N$	rate of change of magnetic flux density for nearby lightning strike, T/sec
$\dot{B}_S$	rate of change of magnetic flux density for lightning attachment to aircraft, T/sec
D	electric flux density, c/m^2
$\dot{D}$	rate of change of electric flux density, A/m^2
$\dot{D}_N$	rate of change of electric flux density for nearby lightning strike, A/m^2
$\dot{D}_S$	rate of change of electric flux density for lightning attachment to aircraft, A/m^2
$D_F, D_{W,L}, D_{W,R}, D_T, B_L, B_T, B_{W,L}, B_{W,R}, E_X, E_Y, E_Z$	sign conventions for electromagnetic variables of figure 1
dc	direct current
E	electric field, V/m
$F(x)$	underlying distribution of X
$G(y)$	distribution of Y
$G_o(z)$	distribution of the standardized variable $Z = (Y - \eta)/\delta$
$g(y)$	density function of $G(y)$
$g_o(z)$	density function of $G_o(z)$
I	current, A
$\dot{I}$	rate of change of current, A/sec
L	lower confidence limit
M_i	theoretical quantiles of $G_o(y)$ indexed by integers
n	sample size
p,q	quantile indexes used when referring to the distribution $G_o(z)$
R	$= U/L$
R_o	prescribed value of R
r	probability plot correlation coefficient

S_y^2	variance of the sample $y_1, y_2, \ldots, y_n$
U	upper confidence limit
V(p,q)	a factor in the representation of τ^2
X	random variable representing a span measurement
x_ξ	theoretical ξth quantile of F(x)
$\hat{x}_\xi$	estimator of x_ξ
Y	natural logarithm of X
y	value of the random variable Y
$\bar{y}$	mean of the sample $y_1, y_2, \ldots, y_n$
$y_{(i)}$	ith ordered value in the sample $y_1, y_2, \ldots, y_n$
y_ξ	theoretical ξth quantile of G(y)
$\hat{y}_\xi$	estimator of y_ξ
z_p, z_q	theoretical pth and qth quantile of $G_o(z)$
$1 - \alpha$	level of confidence
δ	scale parameter of the distribution G(y)
η	location parameter of the distribution G(y)
ξ	quantile index
τ^2	theoretical variance of $\hat{y}_\xi$
$\hat{\tau}^2$	estimator of τ^2

TABLE 1.- SUMMARY OF DATA ANALYZED

Sample	n	Minimum	Maximum	$\bar{y}$	s_y^2
$\dot{B}_N$, T/sec	22	81.52	1 727		
$\dot{B}_S$, T/sec	46	54.82	1 727	6.570	0.553
$\dot{D}_N$, A/m^2	19	0.7656	19.21		
$\dot{D}_S$, A/m^2	93	1.524	30.60	2.056	0.393
$\dot{I}$, A	27	2003	13 900	8.394	0.207
$\dot{I}$, GA/sec.....	8	2.976	81.85		

TABLE 2.- $\dot{B}_N$ FREQUENCY-OF-OCCURRENCE DISTRIBUTION

Lower limit of grouping, T/sec	Frequency
81.0	7
245.6	6
410.2	2
574.8	2
739.4	3
904.0	0
1068.6	1
1233.2	0
1397.8	0
1562.4	1

TABLE 3.- $\dot{B}_S$ FREQUENCY-OF-OCCURRENCE DISTRIBUTION

Lower limit of grouping, T/sec	Frequency
54.0	4
221.3	4
388.6	6
555.9	7
723.2	4
890.5	6
1057.8	4
1225.1	2
1392.4	4
1559.7	5

TABLE 4.- $\dot{D}_N$ FREQUENCY-OF-OCCURRENCE DISTRIBUTION

Lower limit of grouping, A/m^2	Frequency
0	3
2.0	0
4.0	2
6.0	5
8.0	3
10.0	4
12.0	0
14.0	1
16.0	0
18.0	1

TABLE 5.- $\dot{D}_S$ FREQUENCY-OF-OCCURRENCE DISTRIBUTION

Lower limit of grouping, A/m^2	Frequency
1.0	15
4.0	23
7.0	15
10.0	21
13.0	10
16.0	1
19.0	5
22.0	1
25.0	1
28.0	1

TABLE 6.- I FREQUENCY-OF-OCCURRENCE DISTRIBUTION

Lower limit of grouping, A	Frequency
2 003.0	6
3 192.7	9
4 382.4	4
5 572.1	2
6 761.8	4
7 951.5	1
9 141.2	0
10 330.9	0
11 520.6	0
12 710.3	1

TABLE 7.- $\dot{I}$ FREQUENCY-OF-OCCURRENCE DISTRIBUTION

Lower limit of grouping, GA/sec	Frequency
2.976	2
10.860	1
18.750	1
26.640	0
34.520	0
42.410	0
50.300	1
58.180	1
66.070	1
73.960	1

TABLE 8.- FREQUENCY-OF-OCCURRENCE DISTRIBUTION FOR LOGARITHMICALLY TRANSFORMED $\dot{B}_S$ SAMPLE

Lower limit of grouping, T/sec	Frequency
4.0	1
4.4	1
4.8	0
5.2	2
5.6	4
6.0	6
6.4	11
6.8	12
7.2	9
7.6	0

TABLE 9.- FREQUENCY-OF-OCCURRENCE DISTRIBUTION FOR LOGARITHMICALLY TRANSFORMED $\dot{D}_S$ SAMPLE

Lower limit of grouping, A/m^2	Frequency
0	0
.4	3
.8	6
1.2	12
1.6	22
2.0	17
2.4	24
2.8	7
3.2	2
3.6	0

TABLE 10.- FREQUENCY-OF-OCCURRENCE DISTRIBUTION FOR LOGARITHMICALLY TRANSFORMED I SAMPLE

Lower limit of grouping, A	Frequency
7.0	0
7.3	0
7.6	5
7.9	5
8.2	8
8.5	3
8.8	5
9.1	0
9.4	1
9.7	0

TABLE 11.- NORMAL PROBABILITY PLOT
CORRELATION COEFFICIENT r

Sample	n	r	Significance level
$\dot{B}_S$	46	0.945	<0.005
$\dot{D}_S$	93	.933	>.250
I	27	.980	>.250

TABLE 12.- QUANTILE ESTIMATES FOR $\dot{B}_S$ SAMPLE

ξ	$\hat{x}_\xi$	var($\hat{x}_\xi$)	L	U
0.99	1726.97	410 369.59	834.71	3573.02
.95	1726.97	130 343.63	1146.40	2601.58
.90	1497.09	63 702.34	1075.83	2083.31
.80	1297.77	33 063.77	986.12	1707.90

TABLE 13.- QUANTILE ESTIMATES FOR $\dot{D}_S$ SAMPLE

ξ	$\hat{x}_\xi$	var($\hat{x}_\xi$)	L	U
0.99	27.33	36.16	17.76	42.07
.95	21.07	6.83	16.53	26.87
.90	16.10	2.59	13.23	19.58
.80	14.06	1.37	11.95	16.55

TABLE 14.- QUANTILE ESTIMATES FOR I SAMPLE

ξ	$\hat{x}_\xi$	var($\hat{x}_\xi$)	L	U
0.99	13 900.00	16 969 695.24	7775.83	24 847.52
.95	9 346.66	2 436 825.02	6736.88	12 966.02
.90	9 156.66	1 521 150.71	7032.07	11 923.15
.80	5 960.06	445 145.03	4785.88	7 422.31

TABLE 15.- SAMPLE SIZE VARIATION WITH PRECISION

ξ	R_o	Sample size, n, for -		
		$\dot{B}_S$	$\dot{D}_S$	I
0.99	1.5	592	421	222
	2.0	202	144	76
	2.5	116	82	43
	3.0	81	57	30
0.95	1.5	188	134	70
	2.0	64	46	24
	2.5	37	26	14
	3.0	26	18	10
0.90	1.5	122	87	46
	2.0	42	30	16
	2.5	24	17	9
	3.0	17	12	6

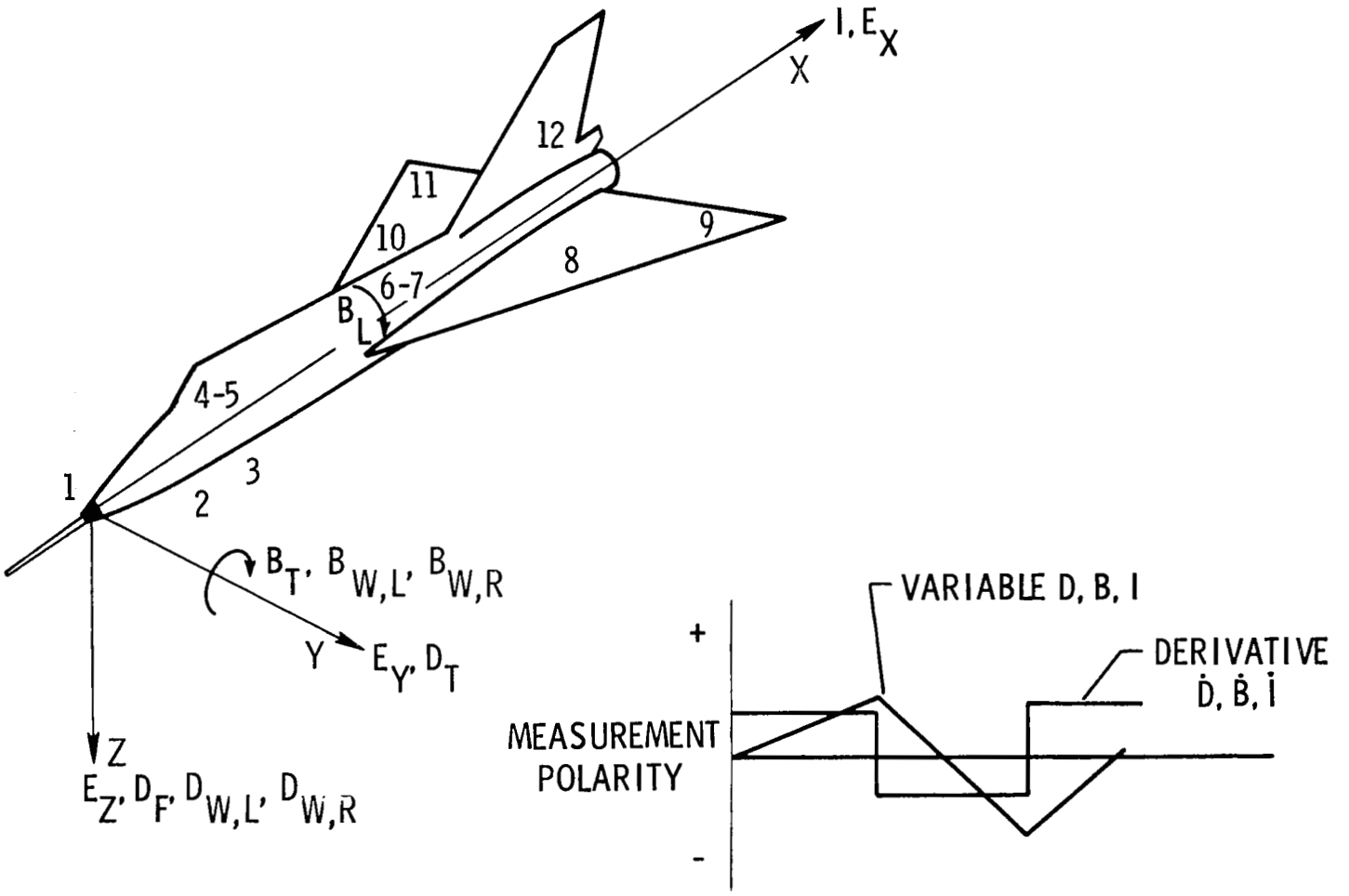

Figure 1.- Electromagnetic sign convention and sensor location.

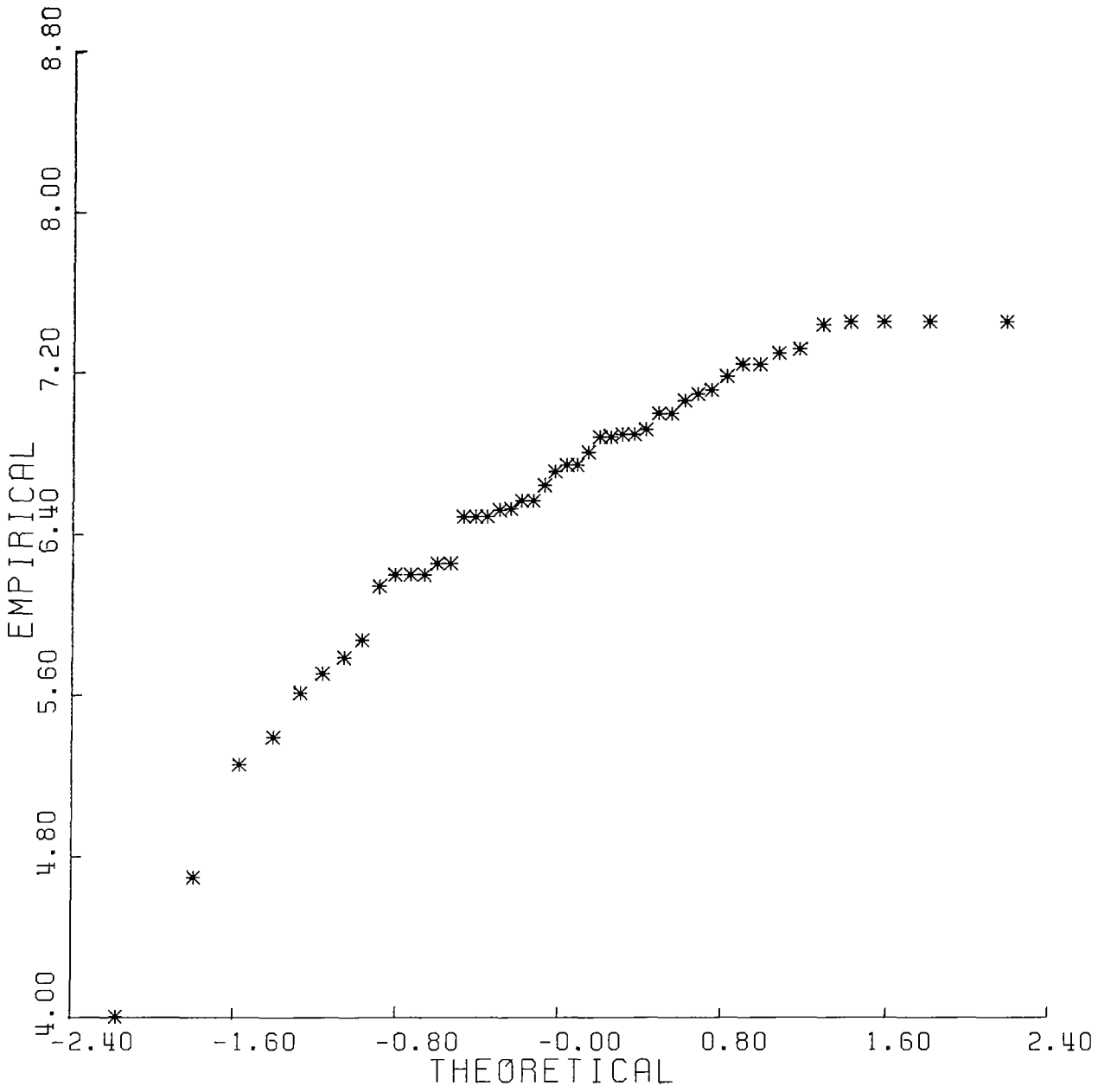

Figure 2.- Empirical transformed $\dot{B}_S$ sample versus theoretical quantiles of standard normal distribution.

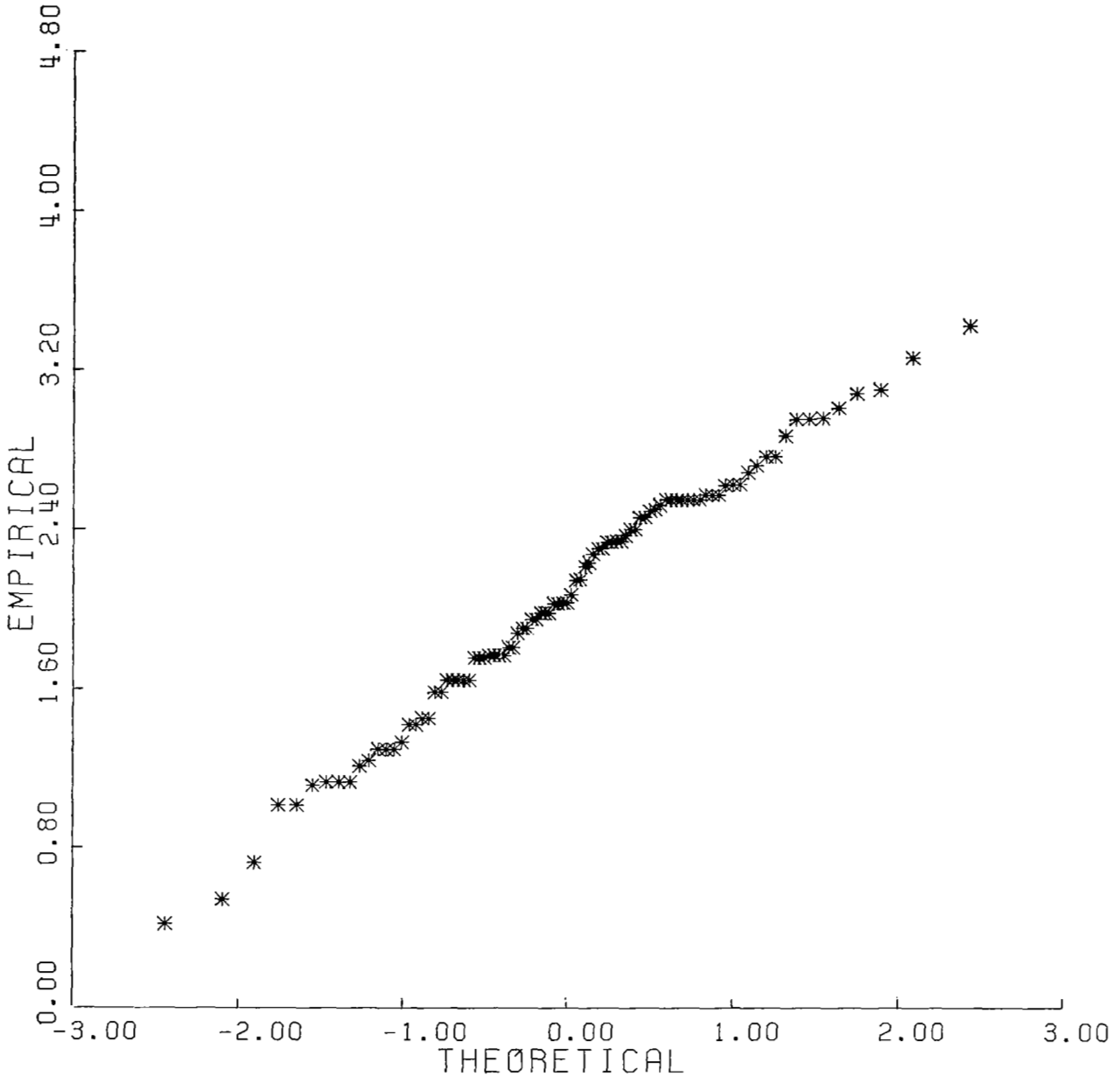

Figure 3.- Empirical transformed $\dot{D}_S$ sample versus theoretical quantile of standard normal distribution.

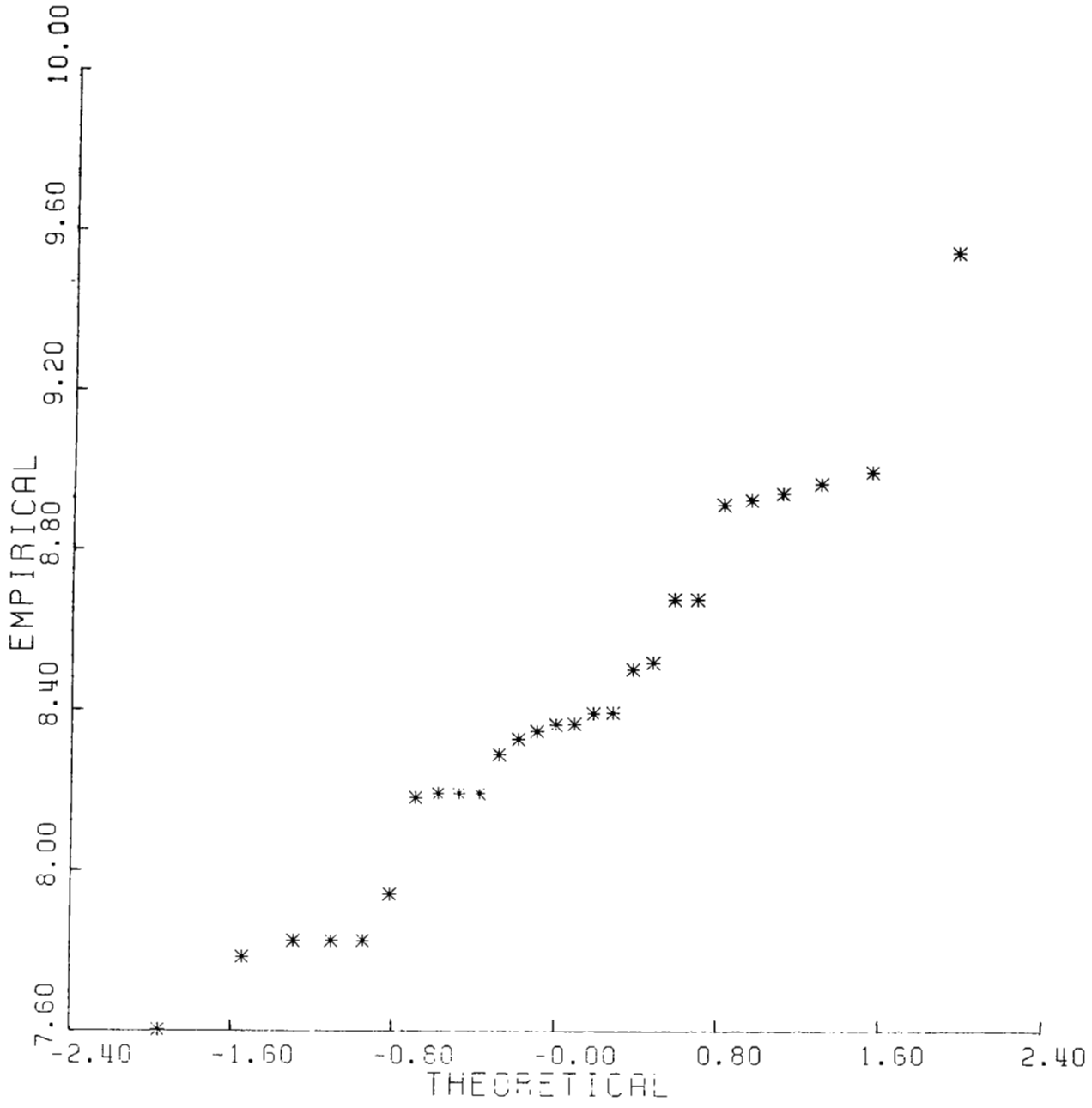

Figure 4.- Empirical transformed I sample versus theoretical quantile of standard normal distribution.

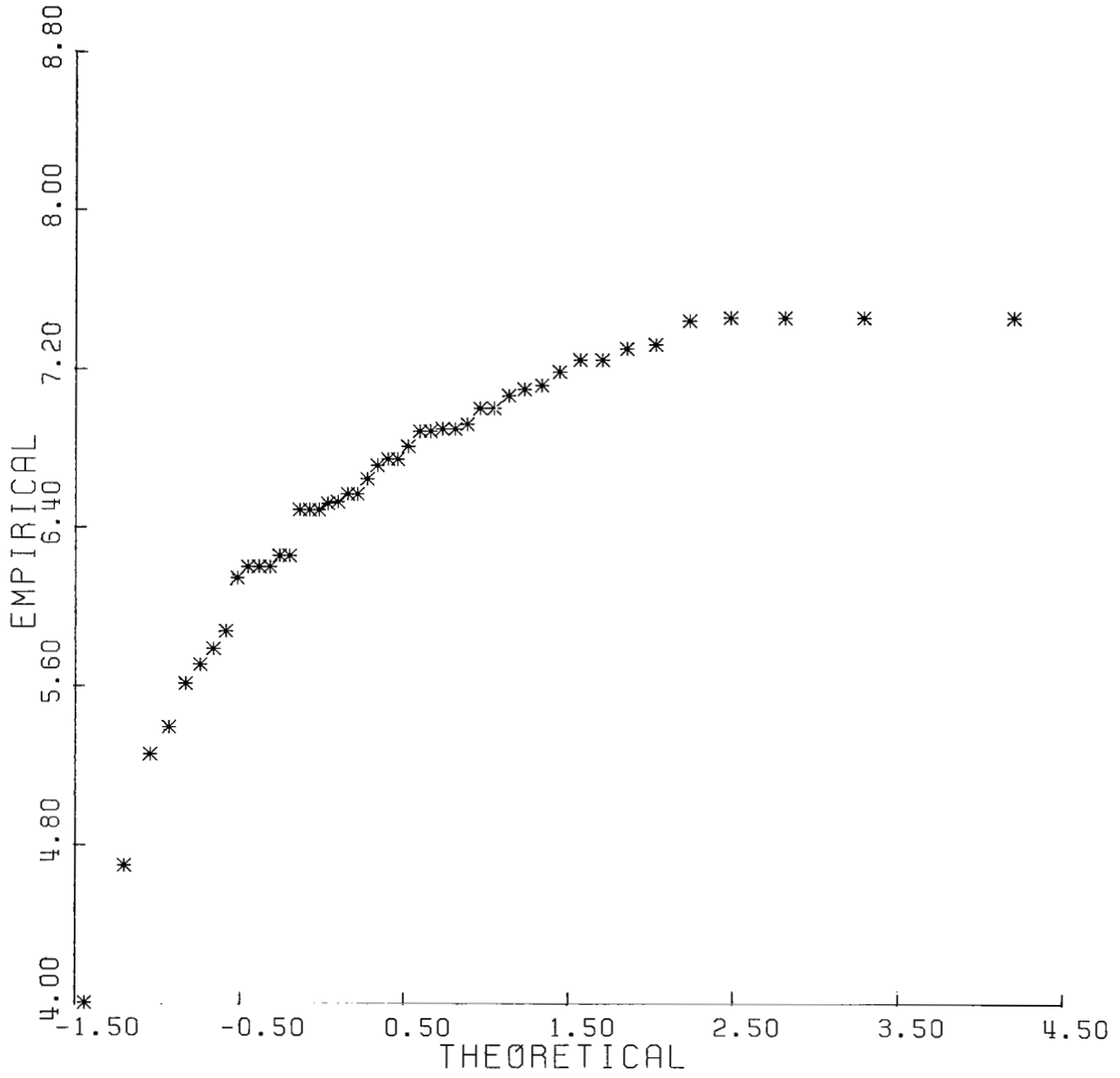

Figure 5.- Empirical transformed $\dot{B}_S$ sample versus theoretical quantile of extreme value distribution.

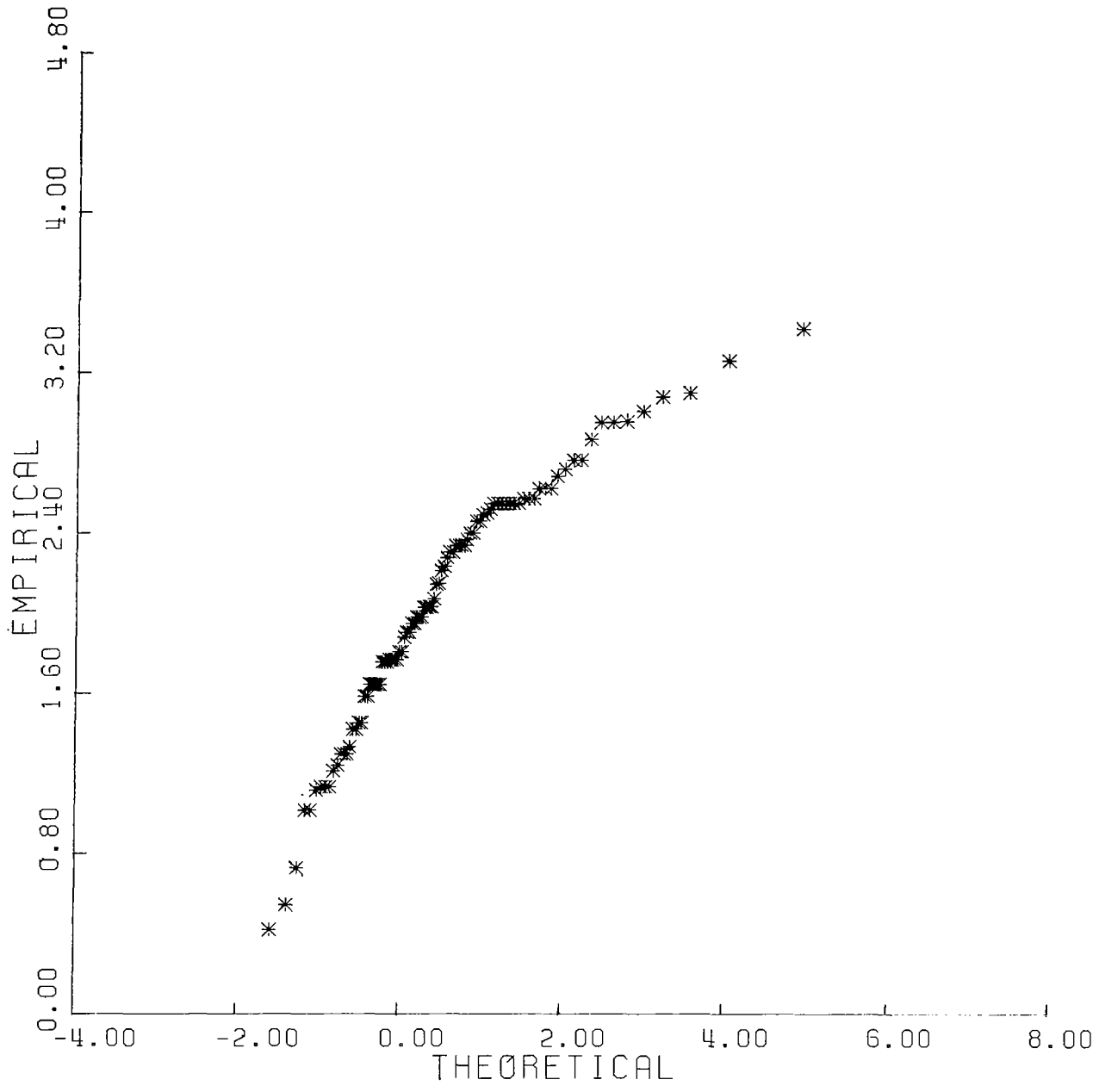

Figure 6.- Empirical transformed $\dot{D}_S$ sample versus theoretical quantile of extreme value distribution.

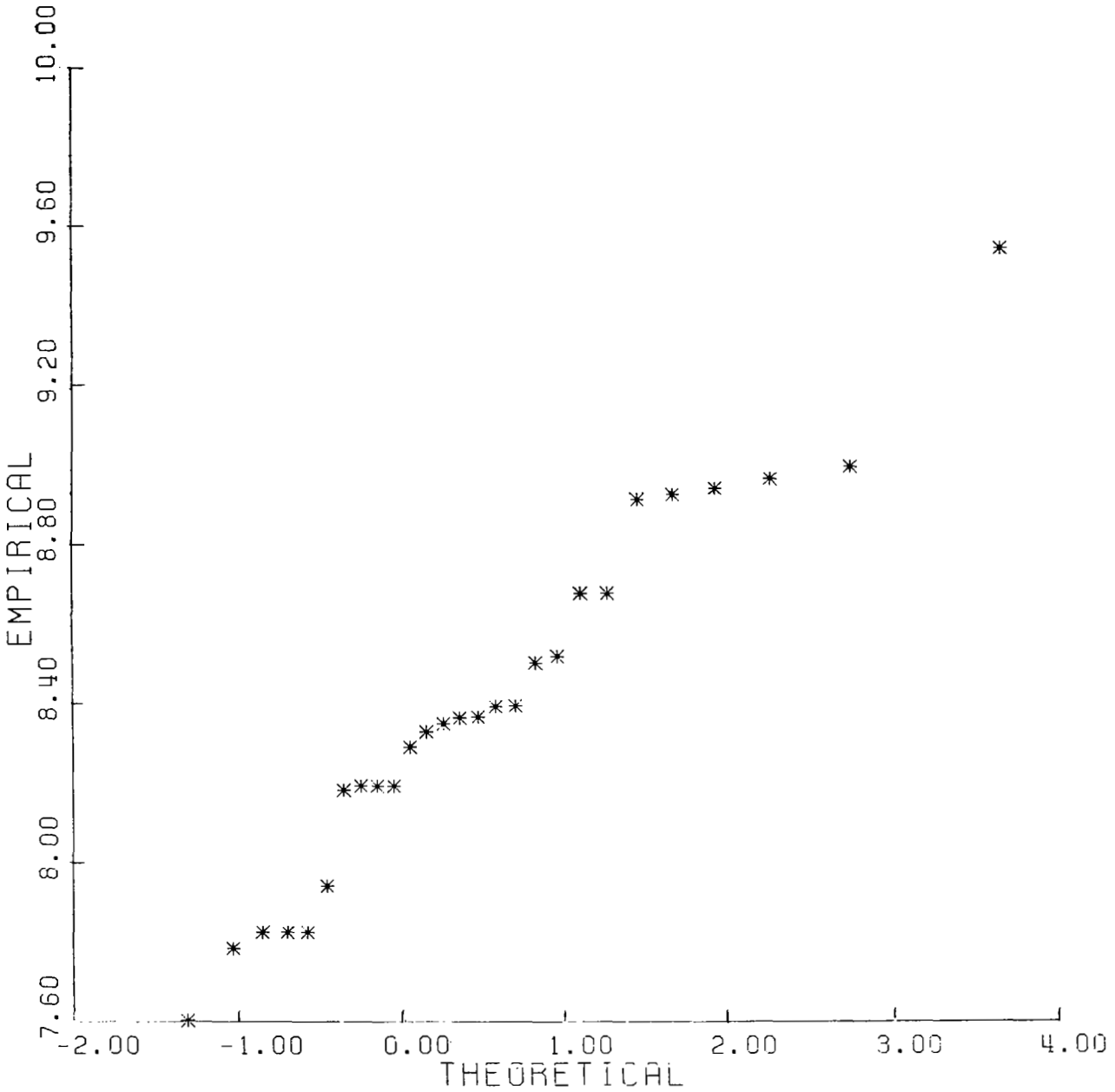

Figure 7.- Empirical transformed I sample versus theoretical quantile of extreme value distribution.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TP-2252		

4. Title and Subtitle	5. Report Date
STATISTICAL ANALYSIS OF DIRECT-STRIKE LIGHTNING DATA (1980 TO 1982)	January 1984
	6. Performing Organization Code 505-34-13-34
7. Author(s) Larry D. Lee, George B. Finelli, Mitchel E. Thomas, and Felix L. Pitts	8. Performing Organization Report No. L-15686
	10. Work Unit No.
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665	11. Contract or Grant No.
	13. Type of Report and Period Covered Technical Paper
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Electromagnetic measurements are being made during direct lightning strikes by NASA Langley Research Center using a specially instrumented F-106B aircraft. The research is to aid refinement, characterization, and understanding of the lightning-aircraft interaction process and the lightning hazards to aircraft. Statistical methods are applied to characterize some aspects of the lightning data obtained from 176 strikes to the aircraft. Specific attention is given to the problem of estimating the upper extreme quantiles of the distributions of peak-to-peak values for currents and rates of change in the magnetic and electric flux densities. A formal treatment via a general location-scale family of models allows the estimation method to be adapted to the realized shapes of the distributions. The shapes are examined by probability plotting methods.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
F-106B aircraft Span of recorded waveforms Cumulative direct-strike lightning data Quantile estimates Lognormal distribution Type II extreme value distribution of largest values	FEDD Distribution Subject Category 47

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	28	

Available: NASA's Industrial Applications Centers

NASA-Langley, 1984