NBSIR 78-1457

Statistical Model for Random Errors of Position Location Based on Lines of Position

Joan R. Rosenblatt

Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

March 1978

Final

Prepared for
U.S. Coast Guard
400 Seventh Street, S.W.
Washington, D.C. 20590

NBSIR 78-1457

STATISTICAL MODEL FOR RANDOM ERRORS OF POSITION LOCATION BASED ON LINES OF POSITION

Joan R. Rosenblatt

Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

March 1978

Final

Prepared for
U.S. Coast Guard
400 Seventh Street, S.W.
Washington, D.C. 20590

NBS REPORT

Statistical Model for Random Errors of Position Location Based on Lines of Position

Joan R. Rosenblatt

Statistical Engineering Laboratory

Introduction

It has been proposed that procedures for routine determination of the position of an aid to navigation should be revised so that the position is "over-determined" by observations. Redundant data are to be obtained so that the precision and accuracy of the position location may be assessed. Redundant observations can also provide on-site checks against gross errors, and can be used in evaluation of systematic errors, but only the imprecision due to random errors is considered here.

The purpose of this report is to provide, for the U.S. Coast Guard's Aids to Navigation Positioning Project, a statistical model and procedure that provides (a) an estimate of the position location and (b) a confidence region for the position location. Using the model, it is shown how the analysis should be modified for the following special circumstances:

(i) replications of some or all observations,

(ii) estimates of individual-observer standard deviations,

(iii) systematic-error checks obtainable from "closing the circle" or from other known relationships among observed angles.

Various positioning methods are used (radar, sextant angles, etc.), but all can be described in general terms as follows. With respect to a known "assumed position" (in practice very close to the unknown position), each observation determines a "line of position (LOP)." If there were no measurement errors, and if the assumed position were correct, each LOP would pass through the assumed position with a known direction (calculated from the assumed position and the known locations of fixed objects on which observations are made). In practice, an observed LOP is taken to be parallel to the calculated LOP, and the distance from the observed LOP to the assumed position is calculated from the measurements.

Remark on geometry: The calculated lines of position are often derived as lines tangent to small circles, or sometimes to hyperbolas or great circles. Errors of the assumed position and of measurement would in general imply that the observed LOP's should be associated with slightly different circles. Since the errors are in practice relatively small, it is customary to ignore the latter complication, which is negligible when the assumed position is sufficiently close.

Coordinate System and Notation

Let the origin be at the known assumed position and let the positive x-axis lie in a fixed direction (North) from the origin. In this coordinate system, each calculated LOP is determined by its angle with the x-axis, taken to be positive when measured in the clockwise direction. Since the problem is invariant when the direction of an LOP is reversed ($180°$ change), the convention is adopted that the angle between each LOP and the x-axis is between $0°$ and $180°$. The framework, then, is the given coordinate system and the angles $\Theta_1, \Theta_2, \ldots, \Theta_n$ of n calculated LOPs.

An observed LOP differs from the corresponding calculated LOP by being translated parallel to the calculated LOP by a distance d. The algebraic sign (positive or negative) of the distance d is taken to be positive if the observed LOP intersects the positive x-axis. Thus the observed LOP corresponding to the angle Θ_i is specified by a distance d_i(i=1, 2, ..., n). The equation of the ith observed LOP may be written

$$y=d_i \sec \Theta_i - x \tan \Theta_i$$

or

$$x \sin \Theta_i + y \cos \Theta_i = d_i \; .$$

Let (μ_x, μ_y) denote the coordinates of the unknown actual position P at which observations were made. A line through P parallel to the ith LOP has the equation

$$x\sin\Theta_i + y\cos\Theta_i = \mu_x \sin\Theta_i + \mu_y \cos\Theta_i = E_i ,$$

where $(-E_i)$ is seen to be the unknown error of the assumed position (origin) in the direction perpendicular to Θ_i. Finally, the unknown distance D_i from P to the observed LOP is $d_i - E_i$,

$$D_i = d_i - \mu_x \sin\Theta_i - \mu_y \cos\Theta_i .$$

Statistical Model

The basis for the statistical model is the set of observation equations

$$d_i = \mu_x \sin\Theta_i + \mu_y \cos\Theta_i + D_i , \; i=1, \ldots, n.$$

The angles Θ_i are known and the coordinates of P are to be estimated.

The errors $D_1, D_2, \ldots, D_n$ are assumed to be statistically independent and normally distributed with zero means. The variances of the D_i are unknown, but their ratios are known; that is, constants W_i are given such that the variance of D_i is

$$\tilde{\sigma}_i^2 = W_i^2 \sigma^2.$$

The constants W_i are defined and interpreted as follows. There is given a "gradient" G_i depending on the geometry of the measurement situation for the line of position, which relates measurement-reading errors (such as angle readings in minutes) to distance errors. If the observers are all equally precise, then $W_i = G_i$, σ^2 is the observer variance for angle determinations, and $\tilde{\sigma}_i^2$ is the variance of corresponding distance errors.

If the measurement errors for different observers are different, then there is an additional constant H_i such that the variance for the particular observer who measured the ith line is $H_i^2 \sigma^2$, and

$$W_i = G_i H_i.$$

If observer variances are taken to be known in advance, then the constants H_i can be defined so that $\sigma^2 = 1$. The constant H_i may also depend on the instrument used.

Estimation of Position and Variance

Under the above assumptions, the weighted least squares estimates of μ_x and μ_y are unbiased minimum variance and maximum likelihood estimates. These estimates are calculated by minimizing

$$\Sigma_i (d_i - \mu_x \sin\Theta_i - \mu_y \cos\Theta_i)^2 / W_i^2,$$

which means solving the equations

$$I\hat{\mu}_x+K\hat{\mu}_y=\Sigma_i d_i \sin\Theta_i/W_i^2,$$

$$K\hat{\mu}_x+J\hat{\mu}_y=\Sigma_i d_i \cos\Theta_i/W_i^2,$$

where

$$I=\Sigma_i \sin^2\Theta_i/W_i^2,$$

$$J=\Sigma_i \cos^2\Theta_i/W_i^2,$$

$$K=\Sigma_i \sin\Theta_i \cos\Theta_i/W_i^2.$$

The underlying variance σ^2 is estimated by

$$s^2=\Sigma_i (d_i-\hat{\mu}_x \sin\Theta_i-\hat{\mu}_y \cos\Theta_i)^2/W_i^2(n-2),$$

where $\hat{\mu}_x$, $\hat{\mu}_y$ are the estimated values. The divisor is n-2 because two quantities have been estimated, leaving n-2 degrees of freedom for evaluation of measurement errors.

Even if σ^2 is taken to be equal to one (observer variances known), it is advisable to calculate s^2. A control chart plotting the value of s for successive position determinations would provide a check on the assumed known values. The procedure for establishing control limits on the control chart is given, for example, in the ASTM manual on control chart analysis, Part 3, Section 9 and Table 27, where "three-sigma" limits are suggested. (The control limits for an ordinary standard deviation are equally applicable to the square root of the residual variance.) Also useful are "warning limits" at plus/minus "two-sigma." When a point falls outside the warning limits, this is a signal to investigate. For instance, if it appears that the points have been falling above the center line too often, it may be necessary to reinvestigate the estimation of the H-factors, i.e., the individual observers' variances and instrument variances.

Confidence Region

There are two cases, according as σ^2 is known or unknown.

The confidence ellipse for the point (μ_x, μ_y) at the 0.95 probability level is given by setting the quadratic form below equal to a constant:

$$I(\mu_x-\hat{\mu}_x)^2+J(\mu_y-\hat{\mu}_y)^2+2K(\mu_x-\hat{\mu}_x)(\mu_y-\hat{\mu}_y)=C.$$

When σ^2 is known, the constant is

$$C=\sigma^2\chi^2(2, .95),$$

and when σ^2 is unknown, the constant is

$$C=2s^2F(2, n-2, .95),$$

where $\chi^2(2, .95)$ denotes the 0.95 point (upper 5% point) of the chi-square distribution with 2 degrees of freedom and $F(2, n-2, .95)$ denotes the 0.95 point (upper 5% point) of the F-distribution with 2 and n-2 degrees of freedom.

When the number of lines is small, the number (n-2) of degrees of freedom for estimating σ^2 is small and the estimate is highly variable. The effect of this on the size of the uncertainty ellipse is seen in the following values of the probability distribution constants.

$\chi^2(2, .90) = 4.61$	$\chi^2(2, .95) = 5.99$	
$F(2, 1, .90) = 49.0$	$F(2, 1, .95)=199.5$	(n=3)
$F(2, 2, .90) = 9.00$	$F(2, 2, .95)=19.00$	(n=4)

The elliptical locus of points (μ_x, μ_y) bounding the confidence region has one of its axes at an angle $\emptyset$ with the x-axis (measured counter clockwise) defined by

$$\tan 2\emptyset = 2K/(I-J), \ |\emptyset| \leq 45°.$$

For $K \neq 0$, $I \neq J$, the half-length of the axis in the direction $\emptyset$ is

$$a = C^{1/2}[J + K \cot \emptyset]^{-1/2} .$$

The half-length of the axis in the perpendicular direction is

$$b = C^{1/2}[I - K \cot \emptyset]^{-1/2} .$$

If $I<J$, the first of these is the major axis; if $I>J$, the second is the major axis.

If $K \neq 0$ and $I=J$, then $\emptyset = 45°$ and

$$a = C^{1/2}(I+K)^{-1/2} , \quad b = C^{1/2}(I-K)^{-1/2} ;$$

the first of these belongs to the major axis if K is negative.

If $K=0$, then $\emptyset=0$ and

$$a = (C/I)^{1/2} , \quad b = (C/J)^{1/2} .$$

Dependent Lines of Position

Suppose three LOPs L_1, L_2, L_3 are related in the following way: L_1 is based on the bearing difference taken to two fixed points A and B. Next, the bearing difference between B and C determines L_2 and finally, the difference between A and C determines L_3. The three measurements are statistically independent; the mathematical relationship among the three angles has no effect on independence. If the observed bearing

differences are changed in some way to force the observed angles to add correctly, the resulting altered values of d_1, d_2, d_3 would indeed be correlated; but such an adjustment is not recommended since it is equivalent to changing the estimation procedure.

In general, there seems to be no special advantage to be gained from using the information about relations among the angles. Whenever there are three or more lines, $n \geq 3$, then μ_x and μ_y can be estimated from any subset of two or more. The data contain n-2 degrees of freedom for estimating s^2 and ordinarily it would be preferable to use all lines symmetrically rather than to look at the distance from one line to the point determined by other lines.

If there are <u>four</u> or more lines, however, the sum of squares of residuals having n-2 degrees of freedom (d.f.) can be written as the sum of two terms, one with n-3 d.f. associated with the residual variance for a position P* located by n-1 lines, and a second term with one d.f. for the distance from P* to the n<u>th</u> line. This comparison might be utilized in an attempt to isolate an "outlier" measurement if the s^2 calculated from all n lines appears to be too large. Specifically, let r_n denote the distance from the n<u>th</u> line to the position calculated from the other n-1 lines and let s_n^2 denote the estimate of σ^2 calculated with the n<u>th</u> line omitted. Then compute the ratio of r_n^2/W_n^2 to s_n^2. Repeat the calculation for each subset of n-1 lines. If one of the ratios is much larger than all the others, suspicion points to the corresponding omitted line.

<u>Closing the Circle</u>

As in the case of the sum of two adjacent angles, just considered, it is not recommended that the observed lines of position be adjusted by forcing the angles between bearings to add up to 360°. The check on the sum of angles is, of course, useful as a check for gross errors. But an adjustment would induce correlations among the observations.

Selection of lines to provide for checks on the angle measurements is desirable for investigation of systematic errors, even though it does not lead to any improvement of estimates of random error.

Replications of some or all of the LOP's

Suppose the data for each LOP are

$$d_{ij},\ \Theta_i, \quad j = 1,2,\ldots,n_i, \quad i = 1, \ldots, n,$$

where $n_i \geq 1$. The maximum likelihood estimates may be obtained from $N=\Sigma n_i$ lines in exactly the same way as before, but it is known from least squares theory that mathematically identical results (estimated values of μ_x and μ_y) will be obtained using the n equations

$$\bar{d}_i = \mu_x \sin \Theta_i + \mu_y \cos \Theta_i,$$

with appropriate weights. If the n_i measurements on the ith line are all made with equal variances (e.g., by the same observer), then

$$\bar{d}_i = (d_{i1} + \ldots + d_{in_i})/n_i,$$

and the solution is obtained by least squares with weights n_i/W_i^2. That is, one uses the n lines specified by $\bar{d}_i$ and Θ_i, taking the number of replications into account in the variances.

If the variances of the replicated d_{ij} are different, then the weights for the individual observations are the squared reciprocals of

$$W_{ij}=G_iH_{ij}$$

and $\bar{d}_i$ is the weighted mean

$$\bar{d}_i=(d_{i1}H_{i1}^{-2} +\ldots+d_{in_i}H_{in_i}^{-2})/(H_{i1}^{-2}+\ldots+H_{in_i}^{-2}).$$

The weight for the <u>ith</u> equation in the least squares calculation is

$$G_i^2/\Sigma_j H_{ij}^{-2} .$$

If the W_i are very different, replications can be used to equalize the weights.

Also, additional information about σ^2 can be obtained. In addition to the estimate of σ^2 obtained from the analysis using the "average" lines, there is an approximate estimate of σ^2 calculated from the sum of squares of deviations within groups of replications,

$$\Sigma_i \Sigma_j (d_{ij} - \bar{d}_i)^2/(N-n)G_i^2 H_{ij}^2 .$$

Large differences between these two estimates of σ^2 would indicate either systematic errors of measurement or flaws in the statistical model.

<u>Standard deviations of individual observers</u>

An estimate based on exactly 3 LOP's provides only one degree of freedom for estimating error, so that the effects of individual observers are confounded. The estimates of σ^2 obtained from a series of different position location operations performed by the same team of observers could be plotted on a control chart and a long-run estimate could be obtained. It would also be useful to plot such series of estimates of σ^2 against other variables that might be suspected to be associated with systematic errors (one might be visibility; there are surely many others).

The design of an experiment including replications, to isolate individual observer variabilities, depends on the details of field operations. For example, suppose each line of position is calculated from observations by one observer. Two designs could be considered:

(1) Three different LOP's, one by Observer 1, one by Observer 2, and one by both observers.

(2) Three different LOP's, two by Observer 1 and the third determined twice by Observer 2.

The first type of experiment provides information about the difference between observers. The second type provides information about an individual observer's variability. There would be many other possibilities.

Each such experiment provides one difference measurement; a sequence of such experiments would yield estimates based on any required sample size.

References

N.R. Draper and H. Smith, Applied Regression Analysis, John Wiley, 1966.

ASTM Manual on Presentation of Data and Control Chart Analysis, ASTM Special Technical Publication 15D, American Society for Testing and Materials, 1976.

NBS-114A (REV. 7-73)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBS IR 78-1457	2. Gov't Accession No.	3. Recipient's Accession No.
4. TITLE AND SUBTITLE Statistical Model for Random Errors of Position Location Based on Lines of Position			5. Publication Date March 1978
			6. Performing Organization Code
7. AUTHOR(S) Joan R. Rosenblatt			8. Performing Organ. Report No.
9. PERFORMING ORGANIZATION NAME AND ADDRESS NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234			10. Project/Task/Work Unit No. 2050597
			11. Contract/Grant No. MIPRZ70099-7-73509
12. Sponsoring Organization Name and Complete Address *(Street, City, State, ZIP)* U.S. Coast Guard 400 Seventh Street, S.W. Washington, D.C. 20590			13. Type of Report & Period Covered Final
			14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

The U.S. Coast Guard's Aids to Navigation Positioning Project has proposed that procedures for routine determination of positions be revised to include redundant observations for assessing precision and accuracy of the results. This report provides a statistical model and a method for calculating a confidence region for the position location.

17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*

Least squares; Navigation, aids to; Position location; Statistics.

18. AVAILABILITY [x] Unlimited	19. SECURITY CLASS (THIS REPORT)	21. NO. OF PAGES
[ ] For Official Distribution. Do Not Release to NTIS	UNCLASSIFIED	13
[ ] Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, SD Cat. No. C13	20. SECURITY CLASS (THIS PAGE)	22. Price
[x] Order From National Technical Information Service (NTIS) Springfield, Virginia 22151	UNCLASSIFIED	$4.00

USCOMM-DC 29042-P74