Useful Statistics for Land Surveyors

By Urho A. Uotila
Department of Geodetic Science
the Ohio State University

Foreword

In earlier days when only transits and tapes were used for traversing, surveyors had guidelines for detecting blunders in their surveys, utilizing closures in a traverse. Modern equipment has changed accuracies obtainable to surveyors; therefore, new statistical quantities should be used to detect possible gross errors in measurements. Some of these statistics are discussed and simple formulas are given for their computations. It is also shown that a closure is not to be used as a measure of accuracy, but only for detection of possible blunders in traversing.

Statistics for Land Surveyors

Most textbooks on surveying have a section on the theory of measurements and errors—for example, Bouchard and Moffitt, 1965; Brown and Eldridge, 1969; Brinker, 1969—therefore, only some specific points will be discussed in this paper.

The first question which might be asked is why anyone uses statistics and, more specifically, why land surveyors should know anything about statistics. Before dealing with either of these questions, the word "statistics" should be defined. According to the Kendall and Buckland, 1971 dictionary, the word statistics is the "Numerical data relating to an aggregate of individuals; the science of evaluating, analyzing and interpreting such data." A surveyor uses statistics when he wants to find out how good his measurements and results are or will be. Statistics can be used also in the process of making decisions about actions to be taken, and as some surveyors do already, statistics can be used to convey results more precisely to a third party.

Every surveyor is familiar with the concept that by measuring a quantity several times, some feeling emerges as to the repeatability or precision and, it is hoped, the accuracy of the measurements and mean value. The following examples demonstrate how to translate this feeling into numbers through the use of statistics. For example, the results of measuring a line 50 times are given in Table 1. The average of 50 measurements, x, is 1462.100 feet. The sample standard deviation of one measurement is computed with the following formula:

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} (\bar{x} - x_i)^2}{n-1}}$$

where $\bar{x}$ = average value, x_i = *ith* observation and n = number of observations. The difference of $\bar{x} - x_i$ is called a residual. This means that the sample standard deviation of one observation is the positive value of the square root of the quotient of the square sum of all residuals divided by a number which is one less than the number of observations. The sample standard deviation for a measurement given in Table 1 is: $\hat{\sigma} = 0.033$ foot. The corresponding sample standard deviation of the mean value is obtained by using the following formula:

$$\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}.$$

The standard deviation of the mean value of the 50 observations given in Table 1, is

$$\hat{\sigma}_{\bar{x}} = 0.005 \text{ foot.}$$

What do these numbers mean? It has been common practice that observed values having residuals larger than three times the standard deviation are rejected as blunders. When this idea is related to the normal distribution of residuals and when the *true* standard deviation is known, it can be shown that 99.73% of the residuals will be less than three times the standard error. What is the difference between true standard deviation and a sample standard deviation? It can be said, at this point, that the sample standard deviation is a progressively better estimate of the true standard deviation when it is computed by using more and more observations. When the n (number of observations) approaches infinity, the sample value approaches the true value.

This statement is also correct for the *mean* value. This is demonstrated in Table 2 where the sample means and the sample standard deviations for various combinations of the values given in Table 1 have been computed. The second column indicates which observations have been included in the computations. For example, the first row gives the results where only the first and second observations were used. On the second row, observations one to three have been used and so on. Under column $\bar{x}$, sample means are given for each case and corresponding sample standard deviation of one observation is given under heading $\hat{\sigma}$, and the sample standard error of the mean is given under $\hat{\sigma}_{\bar{x}}$. All values under column $\bar{x}$ are estimates of the same length of the line and under column $\hat{\sigma}$ all values are estimates of the same theoretical value—standard deviation of one observation. It can be seen that values in the same column differ considerably.

Here is a set of numbers representing the length of the line; what can be said about them? With experience it would be concluded that when the number of observations increases, then the mean is a *better* estimate for the quantity itself. In reality, this does not necessarily mean that estimates derived from more observations are closer to true value than those estimates derived from fewer observations. However, if a confidence interval is set up for derived values, the interval is found to be also a function of the number of observations. A confidence interval of a quantity is usually very broad when the quantity has been derived from a small number of observations as compared to the interval related to a large number of observations.

Table 1.

i	Observed Values x_i in feet	Residuals $\bar{x} - x_i$ in feet
1	1462.121	−0.021
2	1462.165	−0.065
3	1462.071	0.029
4	1462.095	0.005
5	1462.097	0.003
6	1462.088	0.012
7	1462.163	−0.063
8	1462.045	0.055
9	1462.108	−0.008
10	1462.091	0.009
11	1462.052	0.048
12	1462.056	0.044
13	1462.145	−0.045
14	1462.085	0.015
15	1462.088	0.012
16	1462.057	0.043
17	1462.081	0.019
18	1462.094	0.006
19	1462.144	−0.044
20	1462.085	0.015
21	1462.115	−0.015
22	1462.113	−0.013
23	1462.071	0.029
24	1462.078	0.022
25	1462.126	−0.026
26	1462.164	−0.064
27	1462.088	0.012
28	1462.101	−0.001
29	1462.084	0.016
30	1462.164	−0.064
31	1462.128	−0.028
32	1462.121	−0.021
33	1462.087	0.013
34	1462.121	−0.021
35	1462.096	0.004
36	1462.106	−0.006
37	1462.130	−0.030
38	1462.107	−0.007
39	1462.051	0.049
40	1462.118	−0.018
41	1462.096	0.004
42	1462.122	−0.022
43	1462.034	0.066
44	1462.117	−0.017
45	1462.021	0.079
46	1462.080	0.020
47	1462.074	0.026
48	1462.140	−0.040
49	1462.097	0.003
50	1462.100	−0.000

Average $\bar{x} = 1462.100$ feet. $\hat{\sigma}_{\bar{x}} = 0.005$ foot.

Table 2.

Set#	Obs. incl.	$\bar{x}$	$\hat{\sigma}$	$\hat{\sigma}_{\bar{x}}$	$P(x_i < \mu < x_j) = 0.95$		$P(\sigma_i < \sigma < \sigma_j) = 0.95$	
1	1–2	1462.143	0.0312	0.0221	1461.863 < μ <	1462.423	0.014 < σ <	0.987
2	1–3	1462.119	0.0471	0.0272	1462.002	1462.236	0.024	0.296
3	1–4	1462.113	0.0403	0.0202	1462.049	1462.177	0.028	0.150
4	1–5	1462.110	0.0356	0.0159	1462.066	1462.154	0.021	0.102
5	1–6	1462.106	0.0331	0.0135	1462.071	1462.141	0.021	0.081
6	1–7	1462.114	0.0369	0.0139	1462.080	1462.148	0.023	0.081
7	1–8	1462.106	0.0421	0.0149	1462.071	1462.141	0.028	0.086
8	1–10	1462.104	0.0375	0.0119	1462.077	1462.131	0.026	0.068
9	1–15	1462.098	0.0372	0.0096	1462.077	1462.119	0.027	0.059
10	1–20	1462.097	0.0353	0.0079	1462.081	1462.114	0.027	0.052
11	1–30	1462.101	0.0346	0.0063	1462.088	1462.114	0.028	0.046
12	1–40	1462.103	0.0321	0.0051	1462.093	1462.113	0.026	0.041
13	1–50	1462.100	0.0333	0.0047	1462.091	1462.110	0.028	0.041

When the normal distribution is assumed, the confidence interval of the mean is defined as

$$P\left(\bar{x} - \frac{\hat{\sigma}\, t_{n-1,\,\alpha}}{n^{\frac{1}{2}}} < \mu < \bar{x} + \frac{\hat{\sigma}\, t_{n-1,\,\alpha}}{n^{\frac{1}{2}}}\right) = 1 - \alpha$$

where

$\bar{x}$ = sample mean
$\hat{\sigma}$ = sample standard error
n = number of observations
$t_{n-1,\,\alpha}$ = Student's distribution for $n-1$ degree of freedom and α significance level
μ = true or theoretical value.

Usually, α is taken as 5% and t is obtained from tables published in textbooks or handbooks on statistics (for example, Hamilton, 1964). The values corresponding to $\alpha = 0.05$ are given in Table 3 where *DF* means degree of freedom, in this case, $n-1$. Using the above formula and $\alpha = 0.05$ the confidence interval for the first mean value in Table 2 is computed as follows:

$$P\left(1462.143 - \frac{0.0312 \cdot 12.706}{\sqrt{2}} < \mu < 1462.143 + \frac{0.0312 \cdot 12.706}{\sqrt{2}}\right) = P(1461.863 < \mu < 1462.423) = 0.95$$

Strictly speaking, this interval means the probability that either 1461.863 will exceed μ or 1462.423 will be less than μ is 5%. Engineers usually use a less accurate—but more understandable—statement: The probability that the true value, μ, is between 1461.863 and 1462.423 is 95%. Confidence intervals for the mean values are given in column 6 of Table 2. These samples indicate how fast the interval decreases when the number of observations increases, for example, from 2 to 5. If a very good estimate of the true standard error is available, such as $\sigma = 0.033$, then $\hat{\sigma}$ can be replaced with this value and it can be assumed that $DF = \infty$ when using Table 3. In that case, $t_{n-1,\,\alpha}$ would be 1.96 for $\alpha = 0.05$, and the corresponding confidence interval for the first mean:

$$P\left(1462.143 - \frac{0.033 \cdot 1.96}{\sqrt{2}} < \mu < 1462.143 + \frac{0.033 \cdot 1.96}{\sqrt{2}}\right) = P(1462.097 < \mu < 1462.189) = 0.95$$

For the fourth set, with observations from 1 to 5, follow the same procedure to get:

$$P\left(1462.110 - \frac{0.033 \cdot 1.96}{\sqrt{5}} < \mu < 1462.110 + \frac{0.033 \cdot 1.96}{\sqrt{5}}\right) = P(1462.081 < \mu < 1462.139) = 0.95$$

These examples show that a much narrower confidence interval results when the standard deviation is derived from a larger sample or from earlier experience, rather than the sample standard deviation derived from a small number of observations. A confidence interval for the variance can be computed, using the following formula:

$$P\left(\frac{DF\check{\sigma}^2}{\chi^2_{DF,\ \alpha/2}} < \sigma^2 < \frac{DF\hat{\sigma}^2}{\chi^2_{DF(1-\alpha/2)}}\right) = 1 - \alpha$$

The values for χ^2 are given for $\alpha/2 = 0.025$ and $(1 - \alpha/2) = 0.975$ in Table 3 for establishing an 0.95 probability confidence interval. Corresponding intervals are given for the standard deviations in Table 2.

There is another point which deserves attention. One residual, for example, as given in Table 1, does not at any time indicate how accurately the measurement of the quantity corresponding to the residual has been made. A residual is related to the sample mean which changes its value when a number of observations is changed. This means that the value of residuals also changes even though the numerical value of the original measurement does not change. For example, by computing the sample mean and residuals of the first seven measurements, the data in Table 4 result. Comparing the corresponding values in Tables 1 and 4, it can be seen that the values of the measurements have not changed but that the residuals have. This demonstrates beyond doubt that one residual is not to be used as an indication of accuracy or precision. However, if the standard deviation is known, it can be ascertained whether one residual is larger than it should be with a certain probability. A confidence interval gives an idea of the limits with a certain probability, but according to Mandel (1964), it should not be used for rejection when the sample mean and sample standard errors are used. A better way of handling the rejection is through tolerance limits; however, that subject is outside the scope of this paper.

It may be of interest to persons involved in practical work to discuss limits which have been established for closures in different classes of control work. The fulfillment of these requirements is important in fieldwork, but at the same time, any blunders in observations need to be detected. When the same instrument is used repeatedly, the surveyor develops a feeling for the size of closure which will be obtained under certain circumstances. For example, if he ran a closed traverse he would know about what

Table 3.

DF	Student's *t* Distribution	χ^2 – Distribution $\alpha = 0.05$	
	$\alpha = 0.05$	$(1 - \alpha/2) = 0.975$	$\alpha/2 = 0.025$
1	12.706	0.00+	5.02
2	4.303	0.05	7.38
3	3.183	0.22	9.35
4	2.776	0.48	11.14
5	2.571	0.83	12.83
6	2.447	1.24	14.45
7	2.365	1.69	16.01
8	2.306	2.18	17.53
9	2.262	2.70	19.02
10	2.228	3.25	20.48
11	2.201	3.82	21.92
12	2.179	4.40	23.34
13	2.160	5.01	24.74
14	2.145	5.63	26.12
15	2.132	6.27	27.49
20	2.086	9.59	34.17
25	2.060	13.12	40.65
30	2.042	16.79	46.98
40	2.021	24.43	59.34
60	2.000	40.48	83.30
120	1.980		
∞	1.960		

Table 4.

x_i	$\bar{x} - x_i$
1462.121	−0.007
1462.165	−0.051
1462.071	+0.043
1462.095	+0.019
1462.098	+0.016
1462.088	+0.026
1462.163	−0.049
$\bar{x} = 1462.114$	$\hat{\sigma} = 0.0369$

size angle closure to expect, providing he had used the same instrument many times. He would also be able to judge when the size of the closure would be too large. Usually the number of angles measured and the length of the traverse legs would be taken into consideration by the surveyor. With a change of instruments, new numbers must be established. Now the question is: How are these numbers established for each case, using information based on the experience of someone else? If a closed traverse is run, for example, the sum of the interior angles, α:*s*, should be

$$\sum_{i=1}^{n} \alpha_i - (n-2) \times 180° = 0;$$

when all observed interior angles are summed up and $(n-2)\,180°$ substracted, the result is not zero, but a closure:

$$\sum_{i=1}^{n} \alpha_{bi} - (n-2) \times 180° = w;$$

The expected value for the angle closure is zero as it is for the residuals given in Table 1. As stated earlier, a residual is not a measure of accuracy; likewise, an angle closure is not a measure of accuracy. As confidence limits are established for a measured length, limits for the angle closure can also be established. In order to do this, a variance equal to the square of standard error for the closure must be obtained. Variance for the closure can be computed from the following formula:

$$\sigma_w^2 = \sigma_{a_1}^2 + \sigma_{a_2}^2 + \ldots + \sigma_{a_n}^2$$

provided that angle measurements are independent of each other and σ_{a1} is the standard error of *ith* angle. σ_{ai}:*s* are related to: What instrument was used, how long the sides of the angle were, what observation methods were used, and who was observing, etc. If there are good estimates for the variance of each angle, a good estimate for the variance of the closure will result from the above formula. A 95% confidence interval for the closure is

$$P(-1.96 \cdot \sigma_w < w < +1.96 \cdot \sigma_w) = 0.95,$$

and 99.73% confidence interval, which corresponds with the so-called 3σ interval is

$$P(-3 \cdot \sigma_w < w < +3 \cdot \sigma_w) = 0.9973.$$

As an example, assume that each angle has been measured with the same accuracy:

$$\sigma_{a_1} = \sigma_{a_2} = \sigma_{a_3} = \ldots = \sigma_{a_1} = \ldots \sigma_{a_n} = \sigma_A,$$

then

$$\sigma_w^2 = n\sigma_A^2$$

where n = number of angles. As a numerical example, further assume that a closed traverse has 16 interior angles and each angle has been measured with standard error $\sigma_A = 2''$. The corresponding confidence intervals are

$$P(-1.96 \cdot 2'' \cdot \sqrt{16} < w < +1.96 \cdot 2'' \cdot \sqrt{16}) = P(-15''.7 < w < +15''.7) = 0.95,$$

and

$$P(-3 \cdot 2'' \sqrt{16} < w < +3 \cdot 2'' \sqrt{16}) = P(-24'' < w + 24'') = 0.9973.$$

If the closure actually computed is outside the confidence interval, say 24″ in this case, it is evident that there is a blunder in the observations of angles, provided, of course, that the estimates for σ_A:*s* have been good. On the other hand, it cannot be said that a small closure indicates very accurate angle measurements because only one estimate was available. For example, assuming a normal distribution of errors and a case where the sample traverse was measured 100 times, the expected outcome for the size of the closures is given in Table 5.

From Table 5, it can be concluded that 47 times out of 100 the closure will be expected between −5″ and +5″.

Taking the same traverse and computing expected closures when σ_A is 4″ instead of 2″ results in the data given in Table 6.

Under these circumstances, a closure between −5″ and +5″ can be expected in 24.5 times out of 100. It is obvious that if the

Table 5.

Size of the Closure	*Number of Times*
< − 25″	0.1
− 25″ to − 15″	3.0
− 15″ to − 5″	23.6
− 5″ to + 5″	46.8
+ 5″ to + 15″	23.6
+ 15″ to + 25″	3.0
> + 25″	0.1

Table 6.

Size of the Closure	*Number of Times*
< − 25″	5.9
− 25″ to − 15″	11.5
− 15″ to − 5″	20.3
− 5″ to + 5″	24.5
+ 5″ to + 15″	20.3
+ 15″ to + 25″	11.5
> + 25″	5.9

closure were between −5″ and +5″, it would not be known which one of the above cases it would be or if it were either. Therefore, it can be concluded that one closure does not give any measure of work accuracy, but a closure can be used to determine whether any blunders are present.

By running a closed traverse, a so-called *error of closure* can be computed as follows:

$$\textit{error in closure} = \sqrt{(\Sigma\, Lat)^2 + (\Sigma\, Dep)^2} = \sqrt{(\Delta Lat)^2 + (\Delta Dep)^2}.$$

Often the ratio between the error in closure and the total length of the traverse is computed. Sometimes, erroneously, this ratio is believed to represent a measure of accuracy. From earlier discussions it can be seen easily that one error in closure is not a measure of accuracy, but can be used for detecting possible blunders.

The expected value for $\Delta\, Lat$ and $\Delta\, Dep$ is zero, but for $(\Delta\, Lat)^2 + (\Delta\, Dep)^2$ it is something else. Theoretically, it is easy to propagate the errors through a traverse and derive the quantities which can be used to compute expected closures at the end point of the traverse. The formulas to follow can be used for an open traverse where the starting point is numbered 1 and considered errorless, and the last point n. For simplification, let North coordinate = N, East coordinate = E, Latitude = L, and Departure = D. The formulas are:

$$\sigma_{N_n}^2 = \sum_{i=1}^{n-1} \left(\frac{L_{i,\,i+1}}{s_{i,\,i+1}}\right)^2 \sigma_{s_{i,\,i+1}}^2 + \sum_{i=1}^{n} \frac{(E_n - E_i)^2}{\rho^2} \cdot \sigma_{a_i}^2$$

$$\sigma_{E_n}^2 = \sum_{i=1}^{n-1} \left(\frac{D_{i,\,i+1}}{s_{i,\,i+1}}\right)^2 \sigma_{s_{i,\,i+1}}^2 + \sum_{i=1}^{n} \frac{(N_n - N_i)^2}{\rho^2} \sigma_{a_i}^2$$

$$\sigma_{N_nE_n} = \sum_{i=1}^{n-1} \frac{L_{i,\,i+1} D_{i,\,i+1}}{s_{i,\,i+1}} \sigma_{s_{i,\,i+1}}^2 - \sum_{i=1}^{n} \frac{(N_n - N_i)\,(E_n - E_i)}{\rho^2} \sigma_{a_i}^2$$

where

$\sigma_{N_n}^2$ = variance in N at point n

$\sigma_{E_n}^2$ = variance E at point n

$\sigma_{N_nE_n}$ = covariance between N and E at point n

$L_{i,\,i+1}$ = Latitude from point i to point $i+1$ = $N_{i+1} - N_i$

$D_{i,\,i+1}$ = Departure from point i to point $i+1$ = $E_{i+1} - E_i$

$s_{i,\,i+1}$ = distance between point i and point $i+1$.

N_i = North coordinate at point i

N_n = North coordinate at point n

E_i = East coordinate at point i

E_n = East coordinate at point n

$\sigma_{s_{i,\,i+1}}^2$ = variance for distance from point i to point $i+1$

$\sigma_{a_i}^2$ = variance of angle at point i

ρ = conversion factor depending on units of σ_a; if σ_a in seconds, $\rho'' = 206264''.81$; likewise, if σ_a in minutes, $\rho' = 3437'.75$.

In a closed traverse, the above computation formulas can be used by changing the subindex n to 1. After computation of

$$\sigma_N^2,\ \sigma_E^2 \text{ and } \sigma_{N,\,E}\ .$$

the elements of an error ellipse at the point can be computed by the following formulas:

$$w = \sqrt{(\sigma_N^2 - \sigma_E^2)^2 + 4\sigma_{N,\,E}^2}$$

$$\tan 2t_o = \frac{\sin 2t_o}{\cos 2t_o} = \frac{2\sigma_{N,\,E}}{\sigma_N^2 - \sigma_E^2}$$

$$A^2 = \tfrac{1}{2}(\sigma_N^2 + \sigma_E^2 + |w|)$$

$$B^2 = \tfrac{1}{2}(\sigma_N^2 + \sigma_E^2 - |w|)$$

where A and B are semi-major and semi-minor axes of the error ellipse, respectively, and t_0 is azimuth of the semi-major axis clockwise from the North. The proper quadrangle for $2t_0$ is selected by taking the quadrangle where the sine has the same sign as $\sigma_{N,E}$ and the cosine has the same sign as

$$\sigma_N^2 - \sigma_E^2.$$

In the above formulas w is an auxiliary quantity. The propagation of errors and computations of error ellipses can be done when

$$\sigma_{s_{i,\,i+1}} : s, \sigma_{\alpha_i} : s,$$

and approximate N and E for each point are known. For this purpose, $s_{i,i+1}$ can be computed from approximate values of $N:s$ and $E:s$. A simple closed traverse can demonstrate the computation. Approximate coordinates for the points are given in Table 7 and a diagram of the traverse is given in Figure 1.

The first task is an estimation of

$$\sigma_{s_{i,\,i+1}} : s, \text{ and } \sigma_{\alpha_i} : s.$$

If it is assumed that the standard error in the centering is σ_c ft. and in observing one direction, σ_t, the following formulas result:

$$\sigma^2_{s_{i,\,i+1}} = (0.01 + \frac{1}{100000} \cdot s_{i,\,i+1})^2 + \sigma_c^2$$

$$\sigma^2_{\alpha_i} = 2 \cdot \sigma_t^2 + \rho^2 \cdot \sigma_c^2 \left(\frac{1}{(s_{i,\,i-1})^2} + \frac{1}{(s_{i,\,i+1})^2}\right)$$

$$\rho = 3437'.75 = 206264''.81$$

is a conversion factor.

These formulas are approximate formulas which take into consideration the length of the sides of angles in evaluating $\sigma_\alpha : s$. The formulas will differ depending on the in-

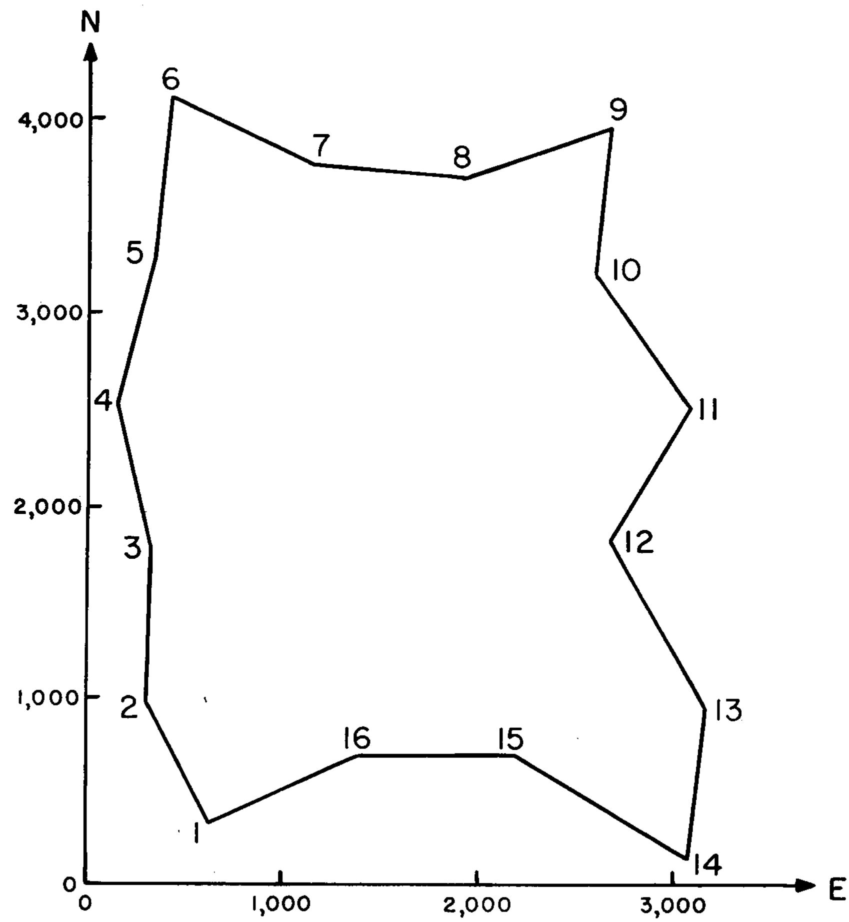

Figure 1.

Table 7.

Point	N in feet	E in feet
1	290	690
2	900	200
3	1700	300
4	2500	180
5	3250	275
6	4060	320
7	3750	1100
8	3700	1900
9	3960	2690
10	3220	2680
11	2500	3050
12	1780	2690
13	980	2910
14	190	2820
15	700	2150
16	710	1380

struments and techniques used. Formulas more or less rigorous than those above might be used. The quality of the end results depends on how well σ:s have been determined.

Using $\sigma_c = 0.005$ foot and $\sigma_t = 2''$ for the closed traverse, starting from point one, the results are as follows:

$$\sigma_w^2 = 181.07\ ('')^2$$
$$\sigma_N^2 = 0.012013\ \text{ft.}^2$$
$$\sigma_E^2 = 0.024431\ \text{ft.}^2$$
$$\sigma_{NE} = -0.006776\ \text{ft.}^2$$

These results were obtained very easily through a small program prepared for a Hewlett-Packard 9810A calculator. Using the same calculator, results for the error ellipse were obtained as follows:

$$w = 0.018381$$
$$t_o = 113^\circ.75$$
$$A = 0.1656\ \text{ft.}$$
$$B = 0.0950\ \text{ft.}$$

The corresponding standard ellipse is given in Figure 2. Assuming that

$$\sigma_a^2{:}s \text{ and } \sigma_s^2{:}s$$

were very good estimates of true variances, then it can be expected that a point corresponding to the coordinates (which were obtained by computing through the traverse without balancing the angles) will fall inside the standard ellipse with a .373 probability or 37.3% of the repeated cases.

Using factors, c, given in Table 8 as a multiplier of A and B, error ellipses can be constructed for various probabilities. Some of them are given in Figure 2. A 50% probability error ellipse is easiest to understand, where there is a 50–50 chance that the point will fall inside the error ellipse. For example, if the point does not fall inside a 99% error ellipse, it is then quite certain that a blunder exists somewhere along the line.

It must be emphasized that the factors given in Table 8 apply only when σ_s:s and σ_a:s are known. If a scalar (the so-called variance of unit weight) derived from an adjustment is used, a different set of multipliers related to an F-distribution (Richardus, 1966) must be utilized.

Table 8.

Values of c for P% probability ellipse	
P	c
25	0.7585
50	1.177
75	1.665
90	2.146
95	2.448
99	3.035

When the angles are not balanced before adjustment, the size of the closure is a function of a starting point, i.e., different closures will develop, depending from which point the computations start. Similarly, different values result for

$$\sigma_{N_n}^2,\ \sigma_{E_n}^2, \text{ and } \sigma_{N_n E_n},$$

depending on which point is selected as a starting point. For example, if the start is from point 11 in the traverse given in Figure 1 and the same input data is used as in the previous computations, the results are:

$$\sigma_w^2 = 181.074\ \ ('')^2$$
$$\sigma_N^2 = 0.017805\ \text{ft}^2$$
$$\sigma_E^2 = 0.010461\ \text{ft}^2$$
$$\sigma_{N,E} = -0.002029\ \text{ft}^2$$

To have the smallest error ellipse and also the smallest expected closure without balancing angles, computation is made by using as a starting point one which has coordinates closest to the mean value of all the

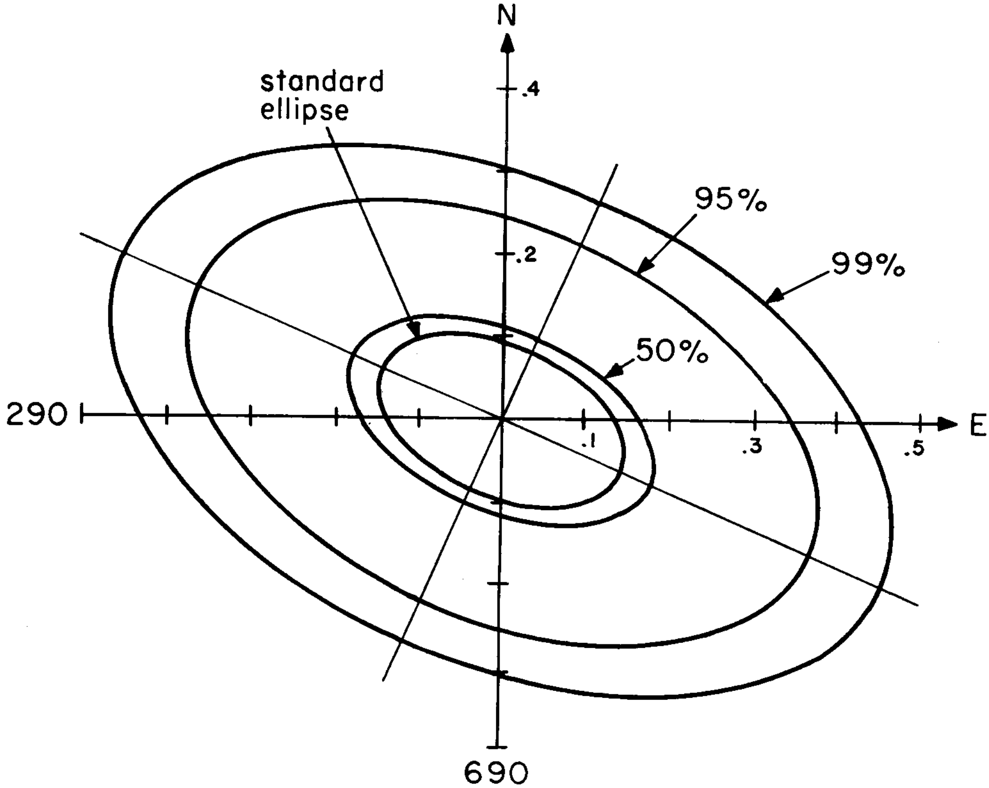

Figure 2.

coordinates in the traverse. In that case, errors in the angles have the least effect as compared to their effects in other cases. The variance of an angle closure and the angle closure itself are, in fact, invariant.

If it is found that the point falls inside the error ellipse, it does not guarantee that blunders do not exist, because blunders may have balanced each other; however, those cases are very rare. If blunders are found to exist, then how would the surveyor proceed to locate them? Pinpointing blunders is unlikely where several exist; but, if there is only one, the following steps are suggested for locating it:

1) It has been learned above that confidence limits for the angle closure can be used for checking whether or not blunders exist in angle measurements.

2) If there is no indication of blunders in the angles and the point still falls outside the 99% or 95% error ellipse, the surveyor may then suspect a blunder in the distance measurements.

How can a blunder be found? Many surveyors already know the trick of checking the direction of a closure and whether any sides are in this direction. The following method might also be useful, especially with a high-speed computer. Compute coordinates commencing with point 1 to the last point, $n-1$—i.e., one before closing. For example, in Figure 1, start at point 1, compute clockwise until point 16 is reached; take computed coordinates of 1 and 16 and compute:

$$\alpha_{1,16,15} = \tan^{-1}\frac{E_1 - E_{16}}{N_1 - N_{16}} - \tan^{-1}\frac{E_{16} - E_{15}}{N_{16} - N_{15}}$$

$$\alpha_{2,1,16} = \tan^{-1}\frac{E_2 - E_1}{N_2 - N_1} - \tan^{-1}\frac{E_1 - E_{16}}{N_1 - N_{16}}$$

$$s_{16,1} = \sqrt{(N_1 - N_{16})^2 + (E_1 - E_{16})^2}$$

a) If angles $\alpha_{1,16,15}$ and $\alpha_{2,1,16}$ are approximately within the confidence limits for the observed angles, there are no blunders in the angles. If the computed side is not equal within a reasonable tolerance (use error ellipse as a scale), the blunder is in that side.

b) If one angle does not check but the other angle and side checks reasonably, the blunder is in the angle which did not check.

c) If distance checks but angles do not, either both angles may be in error, or the error may be somewhere else.

d) If both angles differ from the measured angles and the length of the side does not check, a blunder is usually somewhere else.

Start over at the next point until the blunder is found.

As stated, the formulas which were given above for computation of

$$\sigma_W^2, \sigma_N^2, \sigma_E^2 \text{ and } \sigma_{N,E}$$

can also be used for an open traverse, where the n-point is the last point for determination of the expected errors in the coordinates and azimuth or bearing. An error ellipse can also be drawn at that point. What does an error ellipse—for example, a 99% error ellipse—mean at that point? Usually, even though it is not quite correct, it could be stated with 99% probability that the correct coordinates for the point are inside of that error ellipse.

It is interesting to examine the distribution of the size of relative closures obtained when the expected closures are known. It is somewhat difficult to give these numbers in usual cases, but a demonstration for the case in which

$$\sigma_N^2 = \sigma_E^2$$

and

$$\sigma_{N,E} = 0$$

may be useful. In that case, the error ellipse is a circle.

If the

$$\text{ratio} = \frac{\sqrt{\sigma_N^2 + \sigma_E^2}}{\Sigma s} = \frac{1}{24300}$$

the cumulative distribution of ratios given in column 2 of Table 9 would be obtained. By comparison, the cumulative distribution for another traverse having

$$\text{ratio} = \frac{\sqrt{\sigma_N^2 + \sigma_E^2}}{\Sigma s} = \frac{1}{12150}$$

is given in column 3 of the same table.

Table 9.

Cumulative Distribution when $\sigma_N^2 = \sigma_E^2$ and $\sigma_{NE} = 0$.

Ratio $\frac{CLOSURE}{\Sigma_s}$	Expected % when $\frac{\sigma_N^2 + \sigma_E^2}{\Sigma_s}$	
Equal or Smaller	1/24300	1/12150
1/100000	5.8	1.5
1/90000	7.0	1.8
1/80000	8.8	2.3
1/70000	11.4	3.0
1/60000	15.2	4.0
1/50000	21.1	5.7
1/40000	30.9	8.8
1/30000	48.2	15.2
1/20000	77.2	30.9
1/15000	92.8	48.2
1/10000	99.7	77.2
1/5000	100.0	99.7

This demonstrates again that from a single closure one cannot judge the accuracy of a survey. For example, if

$$\frac{\sqrt{\sigma_N^2 + \sigma_E^2}}{\Sigma s}$$

is one part in 24300, a ratio

$$\frac{\sqrt{\Delta L^2 + \Delta D^2}}{\Sigma s}$$

one part in 50,000 or smaller results 21.1% of the time, but only 5.7% of the time when

$$\frac{\sqrt{\sigma_N^2 + \sigma_E^2}}{\Sigma s}$$

is one part in 12,150. If

$$\frac{\sqrt{\Delta L^2 + \Delta D^2}}{\Sigma s}$$

were one part in 50,000 or less, the surveyor would not know which one of the situations he had, or perhaps it would be neither. This ratio is adequate roughly for checking for blunders in a survey, but that is all. It can also be used with some confidence to set up the limits for acceptable work. To summarize the results from Table 9, one could say that in the first case (1/24,300), 25% of the time a relative closure is equal to 1/45,400 or smaller; 50% of the time, 1/29,200 or smaller; and 75% of the time 1/20,700 or smaller. The corresponding numbers for the second case are 1/22,700, 1/14,600 and 1/10,300. For cases where

$$\sigma_N \neq \sigma_E \text{ and } \sigma_{NE} \neq 0,$$

similar analyses can be made, but would necessitate the computing of semi-major axis, *A*, and semi-minor axis, *B*, of error ellipses and the use of somewhat more complicated procedures.

Now let us discuss another problem: a closed traverse such as the one in Figure 3. When points 2 to 7 are connected with one measurement, what size of closure can be expected? There is not a simple formula for checking this, but both closed traverses can be checked by using the technique employed in the previous example of a closed traverse. In fact, three possible closed traverses can be considered here, one original and two formed by the part of the original traverse and the new line. Had the least squares adjustment been used, checking would have been easier.

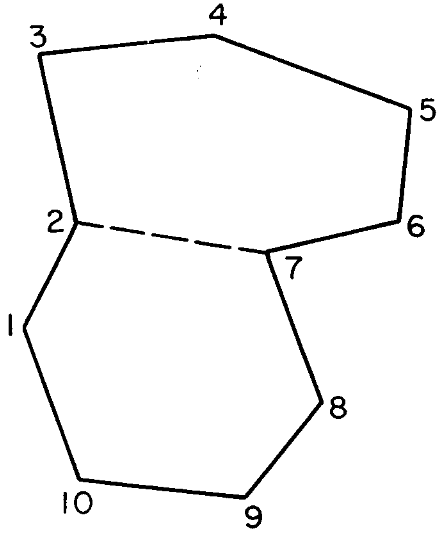

Figure 3.

Traverses in land surveying need not be adjusted by least squares, but when several checklines, angles, distances or combinations must be contended with, it might be the easiest solution. The least-squares method is not complicated (Hirvonen, 1971) if high-speed computer services are available. Furthermore, with very little extra computer time, automatic checks can be provided and the acceptability or inacceptability of all data can be ascertained by the computer. In addition, these services can supply standard errors for all adjusted coordinates, as well as compute accuracy estimates for any distances and bearings computed from the adjusted coordinates. When surveyors are called upon to be expert witnesses in courts, they must be able to attest as to the integrity of their results. Statistics and least-squares adjustments give data that can be used to convey information to a third party in a professional manner.

in a triangle are linearly dependent on each other. Using the above formula and computing the correlation coefficient it can also be proved. The correlation coefficient, which is

$$\rho_{AB} = \frac{\sigma_{AB}}{\sigma_A \sigma_B}$$

has a value equal to –1.00. This is one of the extreme values of correlation coefficients and reflects a linear correlation. Similar formulas can be derived for the cases where there is one common side involved in the computations of the angles.

The point should be emphasized that by deriving new fictitious observations from a set of real observations, no additional meaningful observations are obtained. Through computations, correlations, which must be taken into account in further operations are introduced. The new computed quantities do not control the azimuth any better than the set of original observations.

Furthermore, when some approximate formulas are used in a simple case like this, one is inviting more trouble than one can handle. For example, if Ingram's angles as a set of "adjusted" angles are accepted, then which side would be taken to give a scale for the computations of coordinates? Each one's choice might be different. Why invite trouble when a simple method exists which will give the same results—the least-squares method. It is this writer's feeling that a least-squares solution for this problem in hand is less time consuming than the suggested method and much less confusing. The least-squares solution to the problem gives almost the same adjusted angles in this case, but the results are more impersonal and the adjusted values for the distances are also obtained. It is urged that everyone spend their time studying the least-squares method which gives a consistent set of results and all types of statistical quantities rather than new methods which will be more complicated and which are subject to personal interpretations.

Are there any other comments on this method?—Ed.

Re: 'Useful Statistics for Land Surveyors' by Urho A. Uotila

Published in Surveying and Mapping, Vol. XXXIII, No. 1, March 1973, pp. 67–77

Author Uotila has brought to the editorial staff's attention that several changes should be made in his paper as well as the inclusion of the list of reference works used to prepare it:

—*Page 67,* first formula, col. 2, should read:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x - x_i)^2}{n-1}}$$

—*Page 74,* last two lines, col. 1, should read:

. . . the standard ellipse with a .39347 probability or 39.3% of the repeated cases.

—*Page 76,* col. 2, Table 9, rt. head should read:

$$\text{Expected \% when } \frac{\sqrt{\sigma_N^2 + \sigma_E^2}}{\Sigma_S}$$

References

Bouchard, H. and Moffitt, H. Surveying. International Textbook Company. 1965.

Brinker, R. C. Elementary Surveying. International Textbook Company. 1969.

Brown, C. M., and Eldridge, W. H. Evidence and Procedures for Boundary Location. John Wiley & Sons, Inc. 1967.

Hamilton, W. C. Statistics in Physical Science. The Ronald Press Company. 1964.

Hirvonen, R. A. Adjustment by Least Squares in Geodesy and Photogrammetry. Frederick Ungar Publishing Company. 1971.

Kendall, M. G., and Buckland, W. R. A Dictionary of Statistical Terms. Hafner Publishing Company, Inc. 1971.

Mandel, J. The Statistical Analysis of Experimental Data. Interscience Publishers. 1964.

Richardus, P. Project Surveying. American Elsevier Publishing Company, Inc. 1966.