Demystifying Area Uncertainty: More or Less

Abstract

This paper discusses techniques that can be used to analyze uncertainty in computed areas for the purpose of ascertaining the proper number of significant figures that should be shown in an area expression. It develops a statistical procedure based on the familiar coordinate method for area computations, and then proposes two less rigorous procedures that can be used to yield the same decision. It then demonstrates the use of these equations on three figures that vary in size from a urban/suburban lot to a typical midwest farm.

INTRODUCTION

The uncertainty of a computed area is seldom published in record documents, but is often implied by the number of significant digits given in the area expression. For instance, the value for a parcel area may be expressed as 6.285 acres, or 272,600 ft^2. Both expressions imply that there is uncertainty in the last significant digit. As stated in Wolf (1994), "In recording measurements, an indication of the accuracy attained is the number of digits (significant figures) recorded. By definition, the number of significant figures in any measured value includes the positive (certain) digits, plus one (*only one*) digit that is estimated or rounded off, and therefore questionable." A problem in the profession is that the last significant digit is often determined as a matter of experience, and insight by the professional, and thus varies from person to person.

It is common to hear about surveyor X who always uses too many digits when expressing the values for areas. It is also common, to be asked, "Just how many digits should be expressed in an area of size XXX?" Each of these statements by professionals begs an answer, but most surveyors want a simple rule-of-thumb, and not a full statistical analysis of the problem. This paper will investigate error propagation in computed areas, and propose two less rigorous solutions that can supply the answer to the number of significant figures to be used in an area expression.

SIMPLE ERROR ANALYSIS IN AREA COMPUTATIONS

To provide some insight in how measurement errors propagate into area uncertainties, consider a simple rectangular tract of land shown in Figure 1. For the sake of discussion, assume that this parcel has dimensions of 151.00 ft by 222.00 ft, and was measured by an EDM with manufacturer stated accuracies of ±(5 mm + 5 ppm). From Wolf (1997), it is seen that the uncertainty in the distance observation is dependent on the repeatability of the instrument, and the ability to center the instrument and reflector over the monuments. The uncertainty in an observed distance is provided by the well-known formula

$$\sigma_D = \sqrt{\sigma_i^2 + \sigma_r^2 + a^2 + \left(D \times \frac{b}{1{,}000{,}000}\right)^2} \tag{1}$$

where s_D is the uncertainty in the measured distance D, s_i is the instrument miscentering error, s_r is the reflector miscentering error, and a and b are the instrument's specified accuracy parameters.

In the example, assuming that the EDM is centered over the monument to within ±0.01 ft and that the reflector is centered over its monument to within ±0.02 ft, the estimated error in both the 151 ft and 222 ft distances is approximately ±0.03 ft. The area of the rectangular tract of land can be obtained by multiplying its width times length, and is therefore 33,522 ft^2. The estimated error in this computed area is obtained from error propagation, and is given by Wolf (1997) as

$$\sigma_{area} = \sqrt{\left(l \times \sigma_w\right)^2 + \left(w \times \sigma_l\right)^2} \tag{2}$$

where l is the length of the rectangle having an uncertainty of s_l, and w is the width having an uncertainty of s_w. Using Eq. (2), the uncertainty in the area of the rectangle is

$$\sigma_{area} = \sqrt{(151 \times 0.03)^2 + (222 \times 0.03)^2} = \pm 8 \text{ ft}^2$$

Using this value, it can be seen that the area of the rectangle should only be expressed to the nearest 10 ft^2, and should be written as 33,520 ft^2.

Figure 1 provides evidence in how the uncertainty in the two distances results in an uncertainty in the area. In this figure, the cross-hatched areas depict the uncertainty in the area. Thus it can be seen that the uncertainty in the length results in an uncertainty in the area that is 151 ft long by 0.03 ft wide, or approximately 4.5 ft^2. Similarly the uncertainty in the width of 0.03 ft, results in an area uncertainty that is 222 ft by 0.03 ft wide, or approximately 6.7 ft^2.

The computation of the uncertainty in the area is rather simple in this case. However, few tracts of land are simple rectangular plots. Moreover, most areas today are computed from coordinates, and not from direct measurements such as the distances used in this example. Still, the area is a computed quantity, and thus its uncertainty can also be computed. Since the coordinates are computed quantities, error propagation theory requires knowledge of the covariance matrix for the unknowns (Wolf, 1997). With the presence of least squares software, this is not an insurmountable problem, but it is more than most surveyors would like to use, and probably more than they truly need.

This paper will analyze the error propagation that results when computing area by coordinates, and then will seek implementation of simpler approximations that yield conservative estimates for the area uncertainty. The final goal of this paper is to answer the question, “How many digits should be used in an area of size XXX?”

ERROR ANALYSIS OF THE AREA BY COORDINATES TECHNIQUE

The formula for the computation of area by coordinates for the tract of land in Figure 2 is given in Wolf (1994) as

$$\text{area} = [X_A(Y_E - Y_B) + X_B(Y_A - Y_C) + X_C(Y_B - Y_D) + X_D(Y_C - Y_E) + X_E(Y_D - Y_A)] \tag{2}$$

where X_A, Y_A, X_B, Y_B, ... are the rectangular coordinates of the parcel corners. From the measurement values, it can be shown that the area for this particular parcel is approximately 272,588 ft^2. From Wolf (1997), it is shown that the computed coordinates are correlated, and that their errors propagate through Eq. (2) by the well-known equation $A\Sigma_{xx}A^T$, or in this case:

$$\Sigma_{area} = \begin{bmatrix} \frac{\partial area}{\partial X_A} & \frac{\partial area}{\partial Y_A} & \frac{\partial area}{\partial X_B} & \frac{\partial area}{\partial Y_B} & \cdots & \frac{\partial area}{\partial X_E} & \frac{\partial area}{\partial X_E} \end{bmatrix} \Sigma_{xx} \begin{bmatrix} \frac{\partial area}{\partial X_A} \\ \frac{\partial area}{\partial Y_A} \\ \frac{\partial area}{\partial X_B} \\ \frac{\partial area}{\partial Y_B} \\ \vdots \\ \frac{\partial area}{\partial X_E} \\ \frac{\partial area}{\partial Y_E} \end{bmatrix} \tag{3}$$

where $\frac{\partial area}{\partial X}$ and $\frac{\partial area}{\partial Y}$ are the partial derivatives of Eq. (2) with respect to each of the coordinates, and Σ_{xx} is the covariance matrix for the unknown *XY* parameters that is derived from a least squares adjustment.

The partial derivatives for Eq. (3) are shown in Table 1. Note that these partial derivatives are functions of the coordinates of the traverse. Using the data for the traverse shown in Figure 2, a least squares adjustment was performed, and the covariance matrix for the unknown parameters, Σ_{xx} was computed as

$$\Sigma_{xx} = \begin{bmatrix} 3.5 & -2.6 & 3.5 & -1.7 & 1.5 & -0.3 & 0.1 & -0.6 \\ -2.6 & 2.0 & -2.6 & 1.3 & -1.1 & 0.3 & -0.1 & 0.4 \\ 3.5 & -2.6 & 6.4 & -1.4 & 1.1 & 1.4 & -0.5 & -0.8 \\ -1.7 & 1.3 & -1.4 & 6.0 & -1.8 & 4.4 & -0.6 & 1.7 \\ 1.5 & -1.1 & 1.1 & -1.8 & 7.5 & -3.7 & 3.5 & 1.3 \\ -0.3 & 0.3 & 1.4 & 4.4 & -3.7 & 11.0 & -2.3 & 2.5 \\ 0.1 & -0.1 & -0.5 & -0.6 & 3.5 & -2.3 & 4.5 & 1.7 \\ -0.6 & 0.4 & -0.8 & 1.7 & 1.3 & 2.5 & 1.7 & 4.6 \end{bmatrix} \times 10^{-4} \tag{4}$$

It can be seen in Eq. (4) that the coordinates of control station *A* were held fixed during the adjustment. Thus, their covariance matrix elements can be set to zero, or the partial derivatives with respect to station *A* can be dropped from Eq. (3). The latter technique was used. Following this, it can be shown that Eq. (3) yields

$$\Sigma_{area} = \begin{bmatrix} 295.89 & 261.705 & -245.54 & 99.425 & \dots & 51.12 & -358.145 \end{bmatrix} \Sigma_{xx} \begin{bmatrix} 295.89 \\ 261.705 \\ -245.54 \\ 99.425 \\ \vdots \\ 51.12 \\ -358.145 \end{bmatrix} = \begin{bmatrix} 252.3 \end{bmatrix}$$

Thus, the uncertainty in the computed area for the tract of land in Figure 2 is ±16 ft^2 at a 68% level of certainty. In general, a higher probability is sought in practice–for example 95%–, and the computed value must be increased by a critical value from the *students t* table. Since this is a simple closed traverse, there are only three redundant measurements, and the desired *t* value is 3.183. Thus, the estimated error in the area is 3.183×16 ft^2, or approximately ±50 ft^2. From this analysis, it can be seen that if 50 ft^2 is added and subtracted from the computed area of 272,590 ft^2, the digit in the hundred's place varies, and thus following the guidelines for significant figures as given in Wolf (1994), the proper manner to express this area is 272,600 ft^2.

Although, this method is theoretically correct, few practicing surveyors will have the patience to derive a value that most would have chosen as the proper expression in the first place. Thus, another easier method is sought, that will result in the same rounded figure without the extensive work that is required by a complete statistical analysis of the problem.

SIMPLIFICATION OF ERROR PROPAGATION

Since the coordinates used in Eq. (2) are correlated, the covariance matrix in Eq. (4) has off-diagonal quantities. However, if these off-diagonals are ignored–effectively setting them equal to zero, Eq. (3) can be simplified to

$$\sigma_{area} = \sqrt{\left(\frac{\partial area}{\partial X_A}\sigma_{X_A}\right)^2 + \left(\frac{\partial area}{\partial Y_A}\sigma_{Y_A}\right)^2 + \left(\frac{\partial area}{\partial X_B}\sigma_{X_B}\right)^2 + \left(\frac{\partial area}{\partial Y_B}\sigma_{Y_B}\right)^2 + \cdots + \left(\frac{\partial area}{\partial X_E}\sigma_{X_E}\right)^2 + \left(\frac{\partial area}{\partial Y_E}\sigma_{Y_E}\right)^2} \tag{5}$$

Finally, if a single uncertainty is assumed for the computed coordinates–possibly the root mean square error of the coordinate standard deviations, the estimated error in the area can be simplified to

$$\sigma_{area} = \sigma_C \sqrt{\left(\frac{\partial area}{\partial X_A}\right)^2 + \left(\frac{\partial area}{\partial Y_A}\right)^2 + \left(\frac{\partial area}{\partial X_B}\right)^2 + \left(\frac{\partial area}{\partial Y_B}\right)^2 + \cdots + \left(\frac{\partial area}{\partial X_E}\right)^2 + \left(\frac{\partial area}{\partial Y_E}\right)^2} \tag{6}$$

where s_C is the estimated error in the coordinates.

In Table 2, the necessary coordinate differences for Eq. (6) are taken. In this table, the *X* and *Y* columns list the coordinates of the stations as shown in Figure 2. Columns (3) and (4) list one-half of the difference in the bounding coordinates for each station. That is, the value in(3) for station *A* is ½(517.44 − 125.72) which is ½(X_B − X_E). The value in column (4) for station *B* is ½(0.00 − 591.78) which is ½(Y_C − Y_A). The values in columns (5) and (6) are simply the squares of the corresponding values in columns (3) and (4), respectively. These computations are similar to traverse computations, and can be done without requiring the knowledge about the partial derivatives.

Using the column headings as shown in Table (2), Eq. (6) may be rewritten as

$$\sigma_{area} = \sigma_C \sqrt{\sum (DX)^2 + \sum (DY)^2} \tag{7}$$

From the least squares adjustment of the data in Figure 2, the root mean square error of the coordinates was 0.023 ft. Using this value for s_C in Eq. (6), the estimated error in the area is

$$\sigma_{area} = \pm 0.023\sqrt{284{,}543 + 434{,}039} = \pm 20\ \text{ft}^2$$

This computed value closely matches that computed in the previous section. Again, this value is at a 68% probability level, and should be multiplied by the appropriate *t* value from the *students t* distribution–3.183–to obtain a 95% probability level. Even though, this value is slightly larger than the value of ±16 ft^2 that was previously derived, it will result in a similar decision about the number of significant figures that should be used to express the area for the parcel. That is, the area for the parcel should only be expressed to the nearest 100 ft^2. In fact, this value is a conservative estimate of the value computed from Eq. (3), and thus does not yield an incorrect decision by the surveyor.

EQUIVALENT-AREA SQUARE

Even though Eq. (7) is rather straight-forward in computation, it still requires knowledge about the uncertainty in the station coordinates. Thus, a simpler model is desired. For this it might be appropriate to consider a square having equivalent-area, and perform the analysis using Eq. (2). To further simplify the computations, in the case of a square parcel, Eq. (2) can be rewritten as

$$\sigma_{area} = D\sigma_D\sqrt{2} \tag{8}$$

where *D* is the length of the side of the square, and s_D is the uncertainty in this distance.

To apply Eq. (8) to the parcel in Figure 2, the equivalent-area square must be determined. Since the area of the parcel is approximately 272,590 ft^2, a square of equal-area would have sides with lengths of approximately 522 ft. Equation (1) can be used to determine the uncertainty in the length of the side. In this example, the 522 ft distance has an uncertainty of approximately ±0.03 ft. Thus from Eq. (8), the uncertainty in the area is approximately

$$\sigma_{area} = 522 \times 0.03\sqrt{2} = \pm 22 \text{ ft}^2$$

Again, this value represents a probability of 68%, and should be increased in probability. As can be seen, the equal-area square yields an uncertainty that is close to those previously computed in this paper, and should result in the same decision to round the final value for the area of the parcel to the nearest 100 ft^2.

Since this value is greater than the value computed using Eq. (3), Eq. (8) provides an extremely simple procedure to answer the question of "How many significant figures should be expressed in the area?"

COMPARISON OF PROCEDURES FOR OTHER EXAMPLES

In Figure 3, there are three tracts of land having different areas. This section will compare the application of Eqs. (3), (7) and (8) to these parcels. Table 3 shows the results of the appropriate computations based on each of the equations. Note that while the magnitudes for each method slightly vary, the resulting decision on how many significant figures to retain in the final area remains the same.

Note in Table 3, that the simpler equations begin to vary from the statistically-derived value of Eq. (3) as the parcel size increases. Even though these discrepancies seem large, they result in the same decision about the rounding of the final area expression.

CONCLUSIONS

The uncertainty in a computed area is statistically well defined, and was demonstrated through the use of Eq. (3). However, this procedure requires knowledge about the inverse matrix from a least squares adjustment. Additionally, the paper investigated two less rigorous techniques that result in the same decision concerning the number of significant figures to use in an area expression.

The first procedure uses tabular lists similar to familiar traverse computations, but requires knowledge about the accuracy of the coordinates. This procedure varied most from the statistically derived value as the parcel size increased, but this variation did not affect the decision on the number of significant figures to use in an expression of the area.

The second procedure involved simple error analysis on an equal-area square. This procedure only requires that the accuracy of the distance measurements be known, or computed by Eq. (1). It yielded values that were sufficient to make a decision concerning the number of significant figures to use in an area expression.

While the latter two procedures do begin to differ from the statistically derived values as the parcels becomes large, this variation does not affect the decision on the number of significant figures to use in the area expression, but offer much simpler techniques to perform the analysis.

REFERENCES

Wolf, P.R. and Brinker, R.C. 1994. *Elementary Surveying*. New York:HarperCollins College Publishers, Inc.

Wolf, P.R. and Ghilani, C.D. 1997. *Adjustment Computations: Statistics and Least Squares in Surveying and GIS*. New York:John Wiley & Sons, Inc.

TABLE 1 PARTIAL DERIVATIVES FOR EQUATION 3.

$\frac{\partial area}{\partial X_A} = \frac{1}{2}\left(Y_E - Y_B\right) = 322.385$	$\frac{\partial area}{\partial Y_A} = \frac{1}{2}\left(X_B - X_E\right) = 195.86$
$\frac{\partial area}{\partial X_B} = \frac{1}{2}\left(Y_A - Y_C\right) = 295.89$	$\frac{\partial area}{\partial Y_B} = \frac{1}{2}\left(X_C - X_A\right) = 261.705$
$\frac{\partial area}{\partial X_C} = \frac{1}{2}\left(Y_B - Y_D\right) = -245.54$	$\frac{\partial area}{\partial Y_C} = \frac{1}{2}\left(X_D - X_B\right) = 99.425$
$\frac{\partial area}{\partial X_D} = \frac{1}{2}\left(Y_C - Y_E\right) = -423.855$	$\frac{\partial area}{\partial Y_D} = \frac{1}{2}\left(X_C - X_E\right) = -96.44$
$\frac{\partial area}{\partial X_E} = \frac{1}{2}\left(Y_D - Y_A\right) = 51.12$	$\frac{\partial area}{\partial Y_E} = \frac{1}{2}\left(X_A - X_D\right) = -358.145$

TABLE 2 COMPUTATION OF PARTIAL DERIVATIVE VALUES FOR EQ. (6).

STATION	*X* (1)	*Y* (2)	½D*X* (3)	½D*Y* (4)	*DX*^{2} (5)	*DY*^{2} (6)
A	0.00	591.78	195.860	322.385	38,361	103,932
B	517.44	202.94	261.705	295.890	68,490	87,551
C	523.41	0.00	99.425	−245.540	9,885	60,290
D	716.29	694.02	−198.845	−423.855	39,539	179,653
E	125.72	847.71	−358.145	51.120	128,268	2,613
				Totals =	284,543	434,039

TABLE 3 Comparison of methods on different size parcels.

Parcel	95% Uncertainty in Area (ft^2) Eq. (3)	Eq. (7)	Eq. (8)	Decision on Final area ft^2
(a)	±16	±20	±22	33,520
(b)	±118	±118	±127	1,038,700
(c)	±2300	±1050	±1500	7,994,000

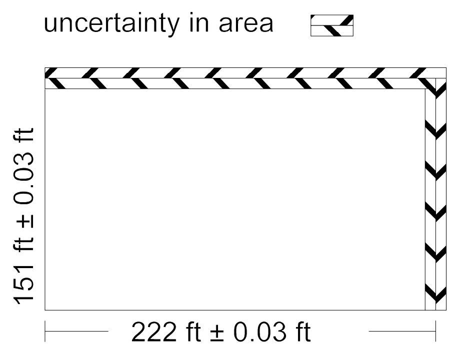

Figure 1 Uncertainty in a computed area for a rectangular tract of land.

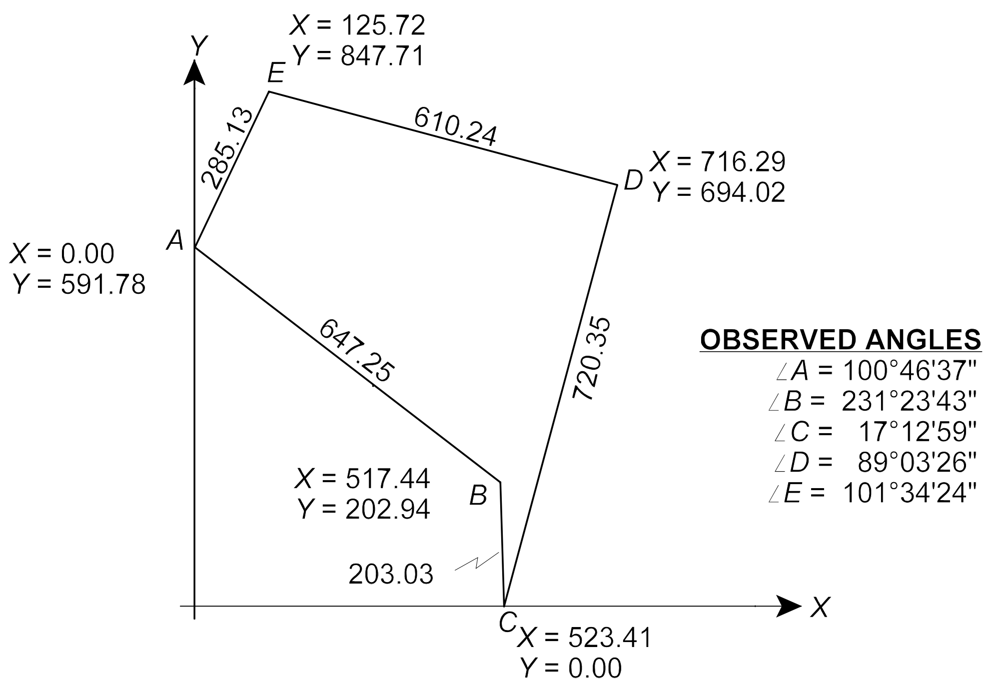

Figure 2 Traverse for computation of area by coordinates.

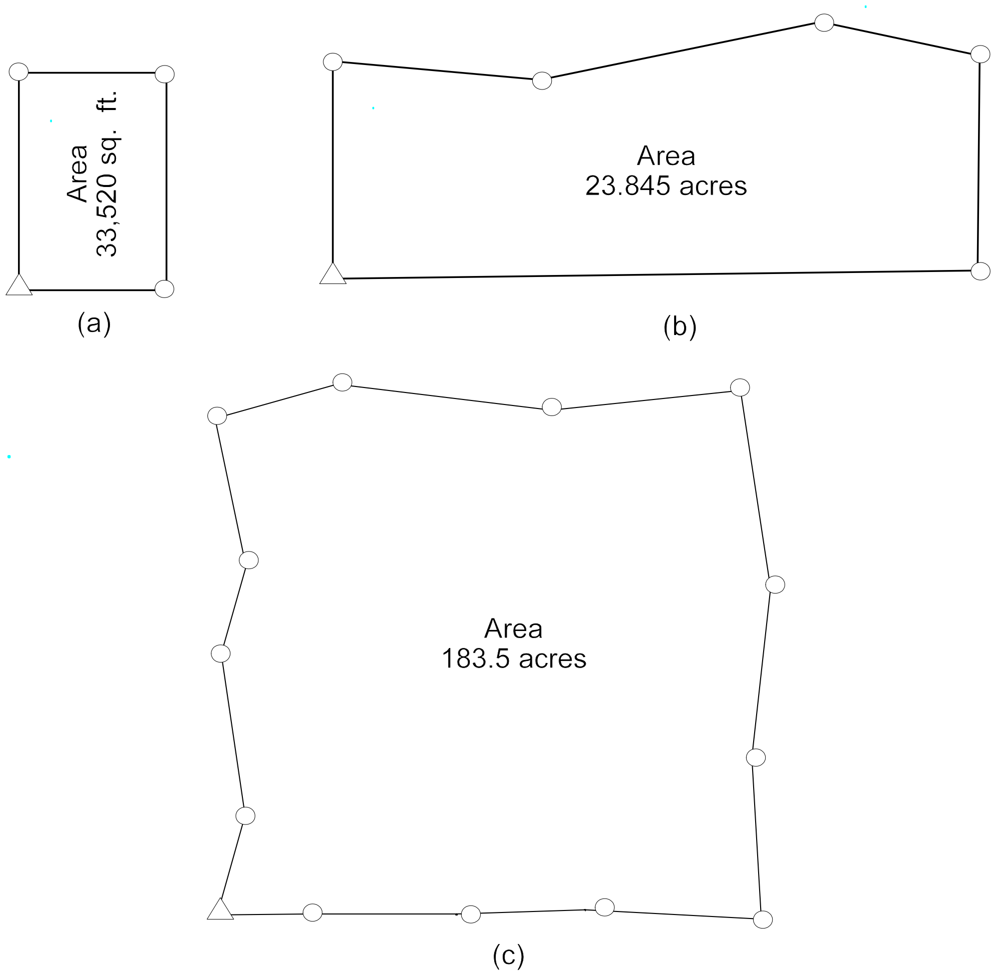

Figure 3 Varying tracts of land: (a) 33,520 ft^2 lot, (b) parcel of area 23.845 acres (1,038,700 ft^2), (c) parcel of area 183.5 acres (7,994,000 ft^2.)