David I. Bower*

A method of estimating mean errors in areas on a map from the errors in point separations.

Keywords: Robert Saxton; accuracy of areas; *MapAnalyst*; planimetric accuracy

Summary
The computer program MapAnalyst allows a fit of a test map to a standard map by a bi-dimensional linear regression on a set of corresponding points in the two maps and provides as one output the scale of the test map, assuming that the scale of the standard map is known. Another output is a distortion grid. It is shown by consideration of a specific example, Robert Saxton's map of Manningham, dated 1613, that the standard deviation of the errors in area of plots of land shown on the test map cannot be reliably determined from the distortion grid. It can, however, be determined from the statistics of the errors in separations of the points on the test map, which can be calculated from the rescaled map. The result is only reliable provided that corresponding points can be found on the two maps such that their mean separations are comparable with the linear dimensions of the plots whose areas are considered.

Introduction

Robert Saxton's map of Manningham, dated 1613, measures approximately 77½ cm E-W by 52 cm N-S and represents an area to the NW of Bradford, England, measuring approximately 4.0×2.7 km. It shows 285 individual plots of land. In a previous publication (Bower, 2009) the scale and accuracy of the map were studied using *MapAnalyst* (Jenny and Weber 2005-9, Jenny *et al*. 2007) to compare it with the earliest six-inch (1:10 560) Ordnance Survey map of the area (1852, parts of sheets 201 and 206). It was necessary to use such an early, but accurately surveyed, map because of the rapid development of the area, which is now almost completely built-up, in the 19th century. One of the purposes of the study was to compare the areas of land-holdings that could be determined from the map with those given by Saxton in the written survey which accompanied the map. A second purpose was to assess the absolute accuracy of area determinations by Saxton. Even as late as 1658 it was pointed out that surveyors frequently made serious errors in determining areas, which could be as much as 25% (Atwell 1658), and clearly this would have been of significance at the time for the value of the land. The reasons why such studies are of interest today is discussed in the paper referred to above, which also gives some further information about Robert Saxton and his work.

As explained in that publication, it was possible to identify 56 corresponding plots of land on Saxton's map and in his written survey. Having determined the scale of the map with *MapAnalyst* it was possible to determine the areas of these plots, as shown on the map, in statute acres. From the scale of the map and Saxton's scale bar it was possible to determine the length of the linear unit, the perch that was used by him and thence to calculate the factor required to convert the areas stated in the written survey to statute acres. The areas so converted agreed well with those determined from the map, with a standard deviation of 0.44 acres on a mean area of 4.45 acres, or about 10% (1 statute acre = 0.4047 hectare). This did not, however, provide any information about the absolute accuracy of either the areas determined from the map or those determined from the survey: the latter must have been determined either from the map or from measurements used in constructing it. In order to study the accuracy of the areas determined from the map, two types of method were used. In the first, areas of plots determined from the Saxton map were compared with the corresponding areas deter-

* Formerly Reader in Polymer Spectroscopy in the University of Leeds, England. [d.i.bower@e-plus.co.uk]

mined from the six-inch map, and in the second an attempt was made to calculate the expected standard deviation of the areas on the Saxton map from these latter areas by using the distortion grid available from the *MapAnalyst* program. The results showed that the standard deviation deduced from the distortion grid was much smaller than that found from the comparison with the six-inch map, and this led to the development of an alternative method for deducing the standard deviation in areas. This method, which is described in the present paper, uses the errors in point separations.

Theory of the methods for estimating errors in area

I: Estimates from map distortion using MapAnalyst

This method gives an indication of the area distortion over the map as well as providing a value for the standard deviation of the error in area measurement. The program produces a distortion grid calculated by the multiquadratic interpolation method (see e.g. Beineke 2007). The method here is to output the map and distortion grid for a set of squares of area a, or side $l = \sqrt{a}$, rescale to the same scale as the original and measure the areas of all (distorted) squares lying wholly or very nearly wholly within the boundary of the mapped area.. The standard deviation of the areas can then be found and compared with that found by direct comparison with the standard (OS) map**.**

II: Estimates from errors in point separations

The second method for estimating the standard deviation of areas is based on considering squares of side l and how their area changes because of the displacements of their vertices due to errors in the map. It is shown in the appendix that the standard deviation in area of such a square is given, to a good approximation, by $\sqrt{2}\sigma_d l$, where σ_d is the standard deviation of the vertices from their 'correct' positions. The important point now is that the centres of the squares under consideration are not fixed. The squares will in general be displaced bodily, and rotated, as well as being distorted, and these bodily displacements do not contribute any change in area. We must therefore consider the displacements of the vertices from the *displaced* position of the centre of each square, i.e. we must take account of the possible correlation of the displacements of points separated by distances comparable with the semi-diagonal of the square $d = l/\sqrt{2}$. The appropriate value for σ_d will then be the standard deviation s of separations d on the map, provided that $s << d$, so that the modified and original semi-diagonals are nearly parallel.

One way of estimating s, and thus σ_d, is to assume that $s = \sigma_d = \sqrt{2} f \sigma$, where σ is the standard deviation of the point positions determined by *MapAnalyst* from a relatively small number P of corresponding points on the two maps and $f = d/D$, where D is the typical separation for close neighbours among the P points. The factor √2 relates the standard deviation s in the separation of two points to the standard deviation of the point positions, provided that s is very much less than the separation between the points, so that the lines of separation are approximately parallel in the two maps. The assumption here is that the standard deviation of separations falls linearly with the separation. Then $\sigma_d = (l/D)\sigma$. A better way of determining σ_d is to use a much larger number of points in *MapAnalyst*, so that D becomes less than $l/\sqrt{2}$. It is then possible to calculate a very large number of separations and to determine σ_d directly.

Results

Before discussing the results of the application of the two methods described above for estimating errors in area, the results of the earlier direct comparison between areas determined from the Saxton map and those determined from the OS map are described so that the new results can be assessed.

Results I: Comparison with the six-inch map

Only 41 plots of land could be found which seemed to be, at least approximately, corresponding plots on the two maps. There was a lack of such plots in the southern and eastern regions of the maps, but five composite corresponding areas were identified in these regions. Figure 1 shows a comparison of the areas from the two maps. The line shown is the simple linear regression line $y = (0.968 \pm 0.013)x + (0.04 \pm 0.16)$.

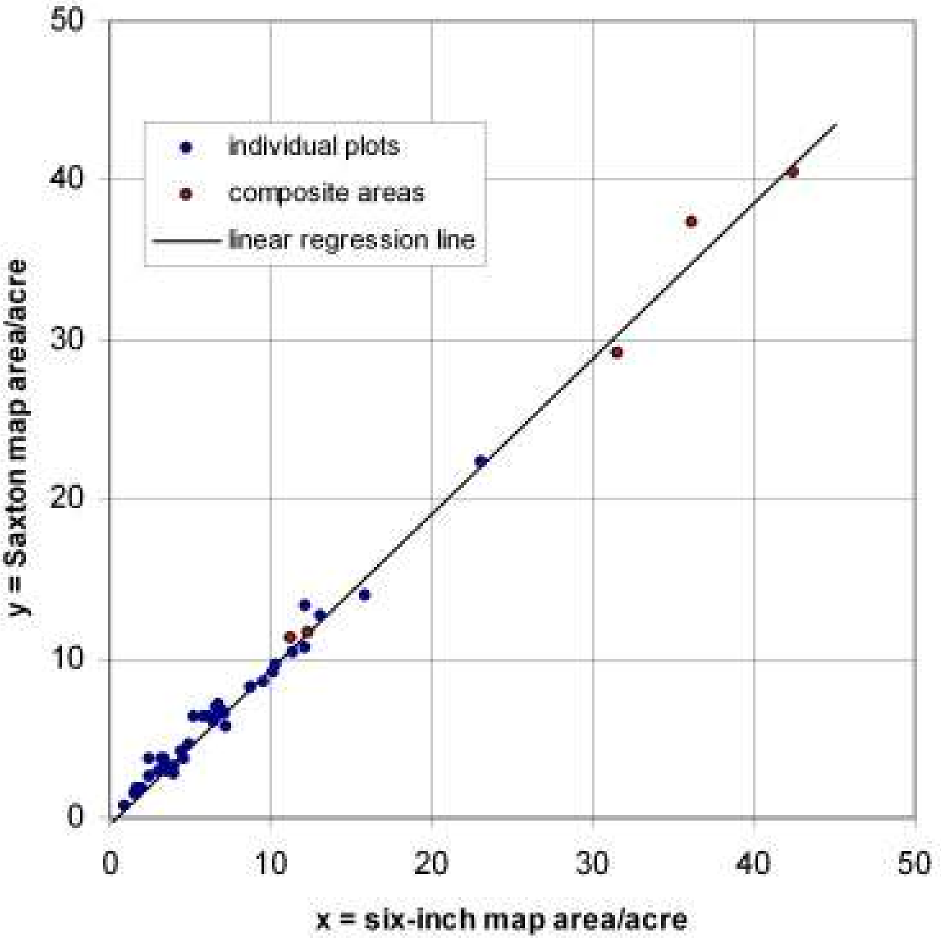

Figure 1. Comparison of areas of corresponding plots on Saxton and six-inch maps.

The gradient is expected to be slightly less than unity because the areas of the lanes occupied about 8% of the total area of Manningham in Saxton's day, but in 1852 the roads were narrower, on average, so that plot boundaries must on average have been moved outwards slightly for plots adjacent to tracks, like most of those considered here. The average size of the plots was 8.6 acres and the standard deviation from the fit was 0.74 acres, or 8.6%. The smallest 30 of the 41 individual plots whose areas were compared with those from the six-inch map had average area 4.27 acres, comparable to the average of 4.45 acres for the 56 plots used in the comparison of Saxton's map and written survey. The standard deviation of the fit for these 30 plots is 0.53 acres, or 12%.

Results II: Estimate of errors in area from distortion grids

In the original comparison with the six-inch map, using *MapAnalyst*, 27 corresponding points were used. It was shown that the Helmert four-parameter fit was adequate and that the scale of the Saxton map was 1: 5140. The standard deviation σ of the 27 points on the map from their correct positions was 21 m, corresponding to a distance of approximately 4 mm on the Saxton map or about 2 mm on the six-inch map. Assuming a Gaussian (normal) distribution of errors in point positions the probability that a point lies in any infinitesimal unit area at distance r from its true position is proportional to $\exp[-r^2/(2\sigma^2)]$. The most probable value of r is σ and the probability that r is less than 2.45σ for any point is 0.95.

Figure 2 shows the results of applying the distortion-grid method with squares of area 7 acres using the original 27-point fit. The results are consistent with the gross variation of scale across the map found earlier (Bower 2009), and the standard deviation of the actual errors in area is 0.25_5 acres, which is much smaller than that found in the direct comparison with the six-inch map. In order to understand this discrepancy the positions of all the angular points of the 46 comparison plots on the Saxton map were compared with their corresponding points on the six-inch map, using *MapAnalyst*. Including the original 27 points, this gave a total of 164 points of comparison. The four-parameter fit in *MapAnalyst* gave a scale 1:5130, less than 0.2% different from that with the 27 point fit, and a slightly reduced standard deviation σ of 19 m for the point position errors. Using the distortion grid with 7 acre squares as before to estimate the errors in area gave a larger standard deviation than before, viz. 0.47 acres.

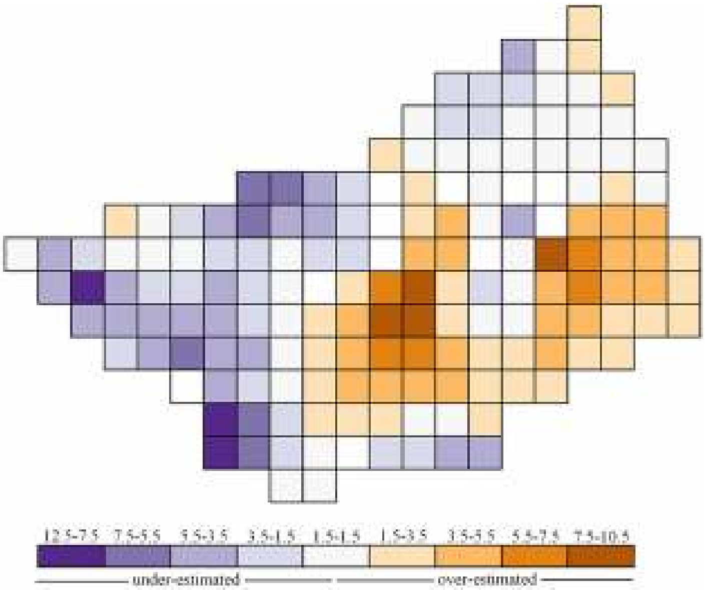

Figure 2. Percentage errors in area for 7-acre squares as deduced from the first *MapAnalyst* fit

Even for the larger 7 acre squares this is less than the 0.53 acres found for plots of average area 4.27 acres in Results I. The predicted absolute standard deviation must fall with the area of the squares used, so that the distortion-grid method of estimating uncertainties in area is clearly inadequate.

*Results III: Estimate of errors in area from errors in point separations***.**

Using the expression $\sigma_d = (l/D)\sigma$ from the theory section II to estimate σ_d and the expression $\sqrt{2}\sigma_d l$ from the appendix for the estimated standard deviation σ_a of the areas leads to $\sigma_a = \sqrt{2}\, l^2\sigma / D$. Estimating D as the mean value of a Gaussian fitted to the distribution of the 67 lengths of the joins in the Delaunay triangulation of the original 27 fitted points in the Saxton map, 600 m, with $\sigma = 21\,\text{m}$ leads to $\sigma_d = 5.88\,\text{m}$ and $\sigma_a = 0.34_5$ acres for squares of area 7 acres. This value is larger than but in rough agreement with the value 0.25_5 acres found by the distortion-grid method with the original 27 points, which is not surprising, because both methods are based on smooth interpolations. It is, however, much smaller than the 0.53 acres found in the direct comparison of plot areas.

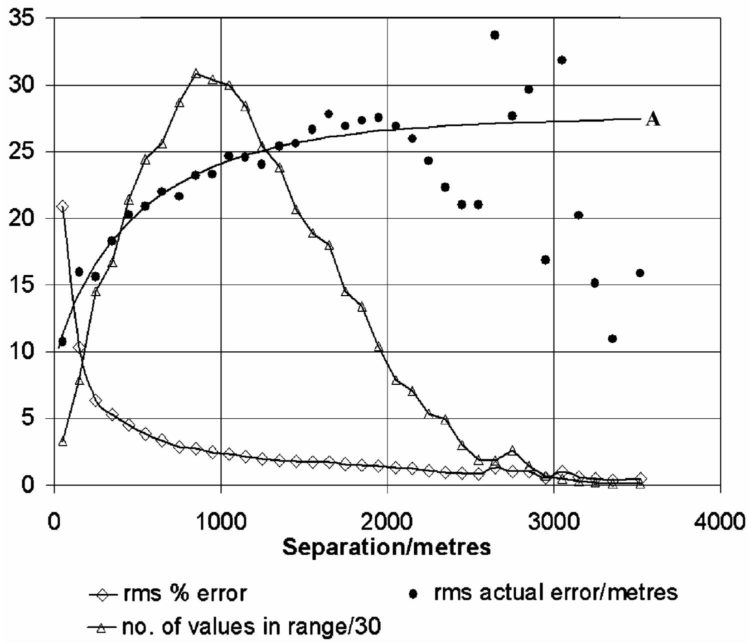

Figure 3. Errors in separations of points on the Saxton map. The rms values are those for separations within a range of 100 m centred at the separation shown. The curve A through the rms actual errors is the best fit of the empirical function $y = a - b.\exp(-c.x^g)$, where y is the rms value and x is the separation, for $50\ \text{m} \leq x \leq 2250\ \text{m}$

In order to get a better estimate of σ_d, a spread-sheet was used to calculate all 13366 separations between pairs of points for each map so that the relationship between separation and error in separation of points on the Saxton map could be examined. A plot of the separations on the Saxton map against those on the six-inch map gave a straight line with gradient 0.9998 passing only 0.52 metres from the origin, as expected, confirming the correct mean scaling of the map to much better than 1%. The statistics of the errors in the distances shown on the Saxton map (assuming all 164 points really are corresponding points) are shown in Figure 3.

As expected, the root-mean-square (rms) percentage error in separation falls monotonically towards zero with increasing separation, except for a very small anomaly near 3 km separation. By 450 m separation it is below 5%, and by 1 km it is below 2.5%. The behaviour of the actual errors is more interesting. It should first be noted that when the separation is large compared with σ the rms error or standard deviation is expected to approach $\sqrt{2}\sigma$ because the lines of separation in the two maps are then almost parallel. The rms error increases steadily until the separation reaches about 2 km, after which it begins to vary in a fairly random manner. It is to be noted, however, that the rms error is far from being linearly related to the separation, as assumed in the first method of estimating σ_d.

The part of the curve between 50 m and 2250 m separation was found to be well fitted by the empirical function $y = a - b\exp\left(-cx^g\right)$, where y is the rms error and x is the separation, as shown in Figure 3. This expression gives the rms value 27.6 m at large separations (cf $\sqrt{2}\sigma = 19\sqrt{2} = 26.9$ m) and 13.6 m for a separation of

119 m (semi-diagonal for 7 acre square). Because 13.6 m<<119 m, this rms value is σ_d for d = 119 m and is very much greater than that used in the first paragraph of this section, viz. 5.88 m. Using this higher value leads to a standard deviation on the area of squares of average area 7 acres of 0.80 acres. For plots of average area 4.27 acres, however, d = 93 m, σ_d = 12.8 m, and the standard deviation of area is 0.59 acres, very much closer to the value 0.53 acres found by direct comparison with the six-inch map (Results I).

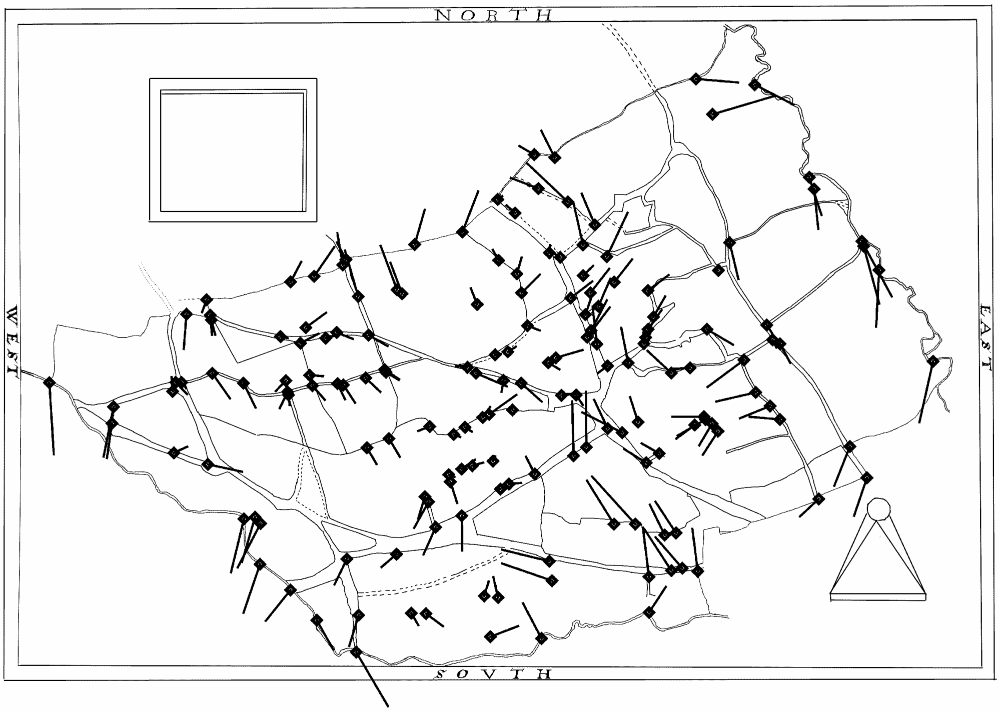

Figure 4. The displacements of points on the Saxton map scaled relatively ×5. The head of the symbol represents the point on the Saxton map and the tip of the tail the point on the 6-inch map. The correlation of the directions of the displacements for nearby points is clear.

The erratic behaviour of the rms error beyond about 2 km can be accounted for by the falling off of the number of values of the separation within each 100 m range, leading to poor statistics. At 2 km this number is already less than 300, compared with the peak value of over 900 near 1 km separation, and by 3 km it is below 20. The low values for the last few points can be readily understood by reference to Figure 4, which shows that the displacements of points near the extreme east and west of the Saxton map tend to be approximately parallel and of similar magnitude.

Conclusions and discussion

It can be concluded that the standard deviation of the errors in area on a map can be deduced fairly reliably from the standard deviation of relative placement errors of points using the expression given in the appendix, provided that the correct value of σ_d is used. This value must, however, be determined by a method that allows correctly for the degree of correlation of displacements of points separated by distances comparable with the linear dimensions of the areas involved, such as the one described. It cannot generally be determined from the displacements of a set of much more widely spaced points by either of the interpolation methods considered, nor can it be determined from the displacement grid produced by *MapAnalyst* even when points separated by distances comparable with the linear dimensions of the areas involved are used.

If the method suggested here could be automated in a program such as *MapAnalyst* it might be more readily applicable than the direct comparison of individual areas on the two maps, but before such implementation was attempted it would be necessary to show that the statistics of the errors in separations of points always follow exponential dependencies of the form found in the present example, and further work on this topic is therefore desirable.

The method could not be applied to the determination of errors in area for a map where the only reliable comparison points that can be found are separated by much greater distances than the linear dimensions of the areas of interest. For such a map there is no method available for determining the mean errors in area, because it is clear that large numbers of corresponding plots of land cannot be found on the two maps for direct comparison of areas. It may, however be possible to find a few plots for direct comparison.

Acknowledgement

I wish to thank Dr Bernhard Jenny and an anonymous referee for useful comments on earlier versions of this paper and the referee in particular for suggesting the present derivation of the result in the Appendix, which is simpler than my original derivation.

References

Atwell, G. (1658). *The Faithful Surveyor*, London.

Beineke, D. (2007). Zur Bestimmung lokaler Abbildungsverzerrungen in Altkarten mit Hilfe der multiquadratischen Interpolationsmethode. *Allgemeine Vermessungsnachrichten* 1**:** 19-27.

Bower, D. I. (2009). The accuracy of Robert Saxton's survey and map of Manningham, dated 1613. *The Cartographic Journal*, 46(2), 115-25.

Jenny, B. and Weber, A. (2005-9). Institute of Cartography, ETH Zurich, http://mapanalyst.cartography.ch/

Jenny, B., Weber, A. and Hurni, L. (2007). Visualizing the Planimetric Accuracy of Historical Maps with MapAnalyst. *Cartographica*, 42: 89-94.

Appendix

For any square with vertices $(x_1, y_1)\ldots(x_4, y_4)$ the area a is given by

$$2a = x_1(y_4 - y_2) + x_2(y_1 - y_3) + x_3(y_2 - y_4) + x_4(y_3 - y_1),$$

which follows from the general expression given by Burgh (1724) for any closed, right-lined polygon.

If the centre of the square is fixed and the x and y coordinates of each vertex are subject to random errors δx_i and δy_i, it follows that the change in area δa is given by

$$\delta a = l\,(-\delta x_1 + \delta y_1 - \delta x_2 + \delta y_2 + \delta x_3 + \delta y_3 + \delta x_4 - \delta y_4)/2\,,$$

where l is the length of the sides of the square and it is assumed that the vertices are at (0,0), (0,l), (l,l) and (l,0), respectively, for $i = 1\ldots4$. If σ_d is the variance of all δx_i and δy_i and $\sigma_d << l$ it follows from the law of propagation of variances that the standard deviation of δa is given by $\sigma_{\delta a} = \sqrt{2}\,l\sigma_d$.

Burgh, Thomas (1724). *A Method to Determine the Area of Right-lined Figures Universally*, London.