An Exploratory Comparison of Metrics for Line Error Measurement

Daniel HÖBLING, Maria-Lena PERNKOPF, Thomas JEKEL and Florian ALBRECHT

Abstract

This paper comprises a discussion of methods for line error measurement. It focuses on three known methods and suggests a forth one based on the evaluation of equally distributed sample points (EDSP) on line features. It is argued that this method of error measurement does not only provide the extent of error but also indicates systematic displacement of lines. While originally developed for a research project on GI-based learning, EDSP may contribute to a wider range of applications. One of them is the detection of errors occurring while digitizing processes as it raises awareness of common errors related to the delineation of features.

1 Metrics of line error measurement

The problem of line error measurement can be seen as an aspect of the more general issue of measuring accuracy and uncertainty within GIS. ZHANG & GOODCHILD (2002) and TVEITE & LANGAAS (1999) discussed a number of different methods for line error measurement. Three of them are compared in this contribution, namely (1) deterministic error band method, (2) gross misfit method and (3) buffer-overlay-statistics (BOS) method. All of them are quantifying the deviation from an estimated line to a “true” reference line:

First, errors in line placement can be described with epsilon ($\varepsilon$) error band models (PERKAL 1956, GOODCHILD & HUNTER 1997). Such approaches assume that drawn (or estimated) lines lie within a certain distance $\varepsilon$ of its corresponding “true” reference line. For the reference line an error band with a width of $2\varepsilon$, one $\varepsilon$ for each side, is calculated so that the digitized line is completely located inside this error band. Therefore the deterministic error band metric (Fig. 1) indicates the maximum deviation between the two compared lines.

Second, the gross misfit method of quantifying line error sums up the intersection area(s) between the estimated and the “true” line (Fig. 2).

Third, the buffer-overlay-statistics (BOS) method (Fig. 3) is creating buffers around the reference lines as well as around the estimated lines and compares them in order to assess spatial accuracy quantitatively (see TVEITE & LANGAAS 1999).
These three methods do measure the extent of error between line data sets. However, they do not indicate systematic error, i.e. systematic distortions in certain directions.

Hence a fourth metric, equally-distributed-sample-points (EDSP) method is presented which is based on comparing pairs of sample points that are evenly distributed on the “true” reference line and on the estimated line (Fig. 4). Each point on the “true” line is compared to its corresponding point on the estimated line. The distance and the direction of the vector from one point to the other can be analyzed. This method is supposed to allow for a deeper understanding of error as it does not only measure the extent of deviation of a line but also systematic distortions in lines and therefore, direction. It may contribute to the framing of hypotheses of the phenomena considered by Goodchild & Hunter (1997), who worked on techniques for accuracy assessment of linear features hampered by generalization.

2 Comparison of methods

The methods presented have been tested on data collected during a project on GI-based learning. Pre-/post tests were carried out to determine the success of a specific treatment (visionary urban planning on digital globes) towards topography learning via aggregated mental maps (Kitchin & Fotheringham 1997). Besides other dimensions, learners were
asked to prepare to-scale sketches before and after the treatment which then have been compared. This included two tasks: first, students had to delineate a specific path within the planning area known to all students. This estimated line was drawn on a blank piece of paper with only three points given for orientation (starting point, endpoint and additional waypoint). Second, students were asked to locate a set of landmarks. While the comparison of sets of landmarks in mental maps, and therefore point data, has proven a relatively easy task (see Lloyd & Heivly 1987), defining metrics for line errors between reference and estimated lines has been rather difficult. Within this context the reference line can be seen as the accurate path between a predefined starting and endpoint.

Data here was based on the sketches of 36 students who finalised their mental maps at t₁ and at t₂. A general overview of the data showed that the length of the drawn lines was slightly overestimated at both times of the assessment. At t₁ the mean of the estimated lines was 2,207.60 m and at t₂ 2,195.12 m, whereas the reference line has a length of 2,072.08 m.

The comparison of the results of the pre test (t₁) to those of the post test (t₂) indicates a significant change that can be traced with all four methods.

1) The deterministic error band has been calculated in order to identify the maximum deviation between the two lines. For calculation a python script for ArcGIS 9.2 was written, which iteratively created buffers around the reference line until the buffer completely included the estimated line. The method therefore constructs a minimum buffer around the reference line. The results show a reduction of the mean error band width from 302.39 m (t₁) to 232.56 m (t₂). The error was lowered by 23.09%.

Table 1: Results of deterministic error band method and change from t₁ to t₂ in meters

<table>
<thead>
<tr>
<th></th>
<th>Sum</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error band (t₁)</td>
<td>10,886</td>
<td>302.39</td>
<td>161.83</td>
<td>260.5</td>
</tr>
<tr>
<td>Error band (t₂)</td>
<td>8,372</td>
<td>232.56</td>
<td>122.39</td>
<td>195</td>
</tr>
<tr>
<td>Change in m</td>
<td>-2,514</td>
<td>-69.83</td>
<td>-39.44</td>
<td>-65.5</td>
</tr>
</tbody>
</table>

(2) Gross misfit. The second method applied produces results comparable to the first one. Therefore the “true” and the estimated lines were converted to polygons. This was possible, because both lines had the same defined starting and endpoints. The resulting areas of the polygons were calculated. The sum gross misfit here decreases from approximately 7.37 km² to 6.33 km², i.e. a decline of 14.15%.

Table 2: Results of gross misfit method and change from t₁ to t₂ in square meters

<table>
<thead>
<tr>
<th></th>
<th>Sum</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross misfit (t₁)</td>
<td>7,370,024.67</td>
<td>204,722.91</td>
<td>94,043.45</td>
<td>182,609.19</td>
</tr>
<tr>
<td>Gross misfit (t₂)</td>
<td>6,326,833.66</td>
<td>175,745.38</td>
<td>98,643.19</td>
<td>151,672.74</td>
</tr>
<tr>
<td>Change in m²</td>
<td>-1,043,191.01</td>
<td>-28,977.53</td>
<td>+4599.74</td>
<td>-30936.45</td>
</tr>
</tbody>
</table>
(3) **Buffer-overlay-statistics.** The BOS method has been used for computing two measures for the geometric accuracy (TVEITE & LANGAAS 1999). To begin with, buffers with a constant width of 75 m were created around the “true” and the estimated lines. Several different buffer distances have been tested to find an adequate buffer for all of the 36 estimated delineations, which show a strong variation in the absolute deviation from the “true” line. The distance of 75 m seemed to be most appropriate. Second, line-polygon overlays were done. The lengths of the lines lying within respectively outside the buffer were considered for calculation of the following measures (see Fig. 5): (1) **completeness,** which compares the length of the “true” line inside the buffer of the estimated line to the total length of the true line and (2) **miscoding,** which compares the length of the estimated line lying outside the buffer of the “true” line to the total length of the estimated line. The average completeness of the examined data sets rises from 56.86 % at \( t_1 \) to 62.31 % at \( t_2 \). Miscodings decline from 47.78 % to 42.83 %. The results are of course heavily dependent on buffer width chosen. An advantage of this method is the adaptability to specific research questions.

![Fig. 5: Completeness (left) and miscoding (right)](image)

(4) **Equally-distributed-sample-points (EDSP)** method. Here, the same number of points was distributed evenly on both the reference line and the estimated lines. Vectors were calculated between corresponding points. The mean lengths of the vectors that resulted from the EDSP method declined by 14.99 % from test \( t_1 \) to test \( t_2 \).

<table>
<thead>
<tr>
<th></th>
<th>Sum (all vectors)</th>
<th>Mean (per sketch)</th>
<th>Std. deviation (per sketch)</th>
<th>Median (per sketch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDSP (( t_1 ))</td>
<td>123,219.02</td>
<td>171.14</td>
<td>80.14</td>
<td>154.90</td>
</tr>
<tr>
<td>EDSP (( t_2 ))</td>
<td>104,750.71</td>
<td>145.49</td>
<td>74.01</td>
<td>129.93</td>
</tr>
<tr>
<td>Change in m</td>
<td>(-18,468.31)</td>
<td>(-25.65)</td>
<td>(-6.12)</td>
<td>(-24.97)</td>
</tr>
</tbody>
</table>
The added value of EDSP results from its possibility to visualize the direction of error. The distribution of the mean directions at \( t_1 \) and \( t_2 \) are visualized in figure 6 and 7.

An obvious tendency to the northwest appears. Vectors getting shorter from \( t_1 \) to \( t_2 \) and a slight shift of the orientation to the northwest can be seen. Detailed analysis of single points within the EDSP method is still to be carried out in order to get a better understanding of the nature of error and form hypotheses about possible treatments.

Fig. 6 and 7: Comparison of mean orientation and mean vector length at \( t_1 \) and \( t_2 \)

### 3 Discussion and further applications

All of the four applied methods show a decline in case of line errors from \( t_1 \) to \( t_2 \) and so indicate an enhancement of spatial accuracy with learners. From the point of topography learning, all methods suggest that the treatment had a reasonable learning effect towards spatial orientation.

From a more general perspective of line error measurement, all methods seem to give valid results and reveal the same trends. The three existing methods of error measurement give reasonable indicators for the extent of line error. They do not indicate systematic nature of line error. If we consider digitizing (either automatic or by hand) from an orthophoto or satellite image, we all remember our uncertainties when it came to delineating borders which may have shadows cast over them. For example, when digitizing forest cover/meadow boundaries, we did of course put down an estimated line. Are there common features between digitizers to the estimates we have been putting down? And are there common systematic distortions to specific complex line features and objects? EDSP line error measurement allows tracing both, the extent as well as the systematic of line error. However, the EDSP method delivers plausible results only if the reference and the estimated lines show at least a certain similarity. This can be considered true for most problems of digitizing. Due to the even distribution of the sample points problems occur if the accuracy of the estimated line strongly varies along its path as in the case of mental maps.
It is still suggested that this tool therefore provides a new quality of line error measurement leading to further applications. Results can be used in training of digitizers as they make aware of common and systematic errors. Results can also be applied to develop better rules for automated feature extraction for both polygons and lines.

Further research has to be done on:

(1) testing the EDSP metric in more detail, especially in respect of directional shifts. An algorithm may be devised that gives an indication of positional error for which EDSP is still a useful tool.

(2) the identification of relevant line sections (e.g., which line sections tend to have specific errors when digitizing (sun/shade; along clear-cut borders, close to landmarks) and

(3) a combination of EDSP with a probabilistic error band method to reduce the effects of little similarity of reference line and estimated line.

Acknowledgements

Thanks are due to Robert Marschallinger (GIScience, Salzburg) for general discussions on line error and Dirk Tiede (Z_GIS Salzburg) for the script calculating the deterministic error band.

References


