Technical Communication

Estimation and propagation of error in measurements of river channel movement from aerial imagery

Nicholas Mount1* and John Louis2
1 School of Geography, Birkbeck College, Malet Street, London, WC1E 7HX, UK
2 Farrer Centre, Charles Sturt University, PO Box 588, Wagga Wagga, NSW 2650, Australia

*Correspondence to: N. Mount, School of Geography, Birkbeck College, Malet Street, London, WC1E 7HX, UK. E-mail: n.mount@bbk.ac.uk

Received 4 December 2003; Revised 30 September 2004; Accepted 30 September 2004

Abstract

Rates of river channel migration or widening are only valid if it can be demonstrated that the amount of change in the measured parameters exceeds the measurement errors. However, it is still uncommon to see estimates of errors presented alongside rates of river channel change measured from aerial photography. Here, detailed methods are presented for the estimation of errors in quantifying channel migration rates that can be easily applied to GIS-based analysis of planform change. In addition, the impact of relaxing assumptions regarding error independence between image dates and systematic registration errors is presented and discussed. Copyright © 2005 John Wiley & Sons, Ltd.

Keywords: aerial photography; channel change; error estimation; error propagation

Introduction

Analysis of vertical aerial photography has proven to be an essential tool in studies of medium- and short-term river planform change. Despite the continuing advances in, and focus on, short-term process-based studies of river channel form and function, planform change measurements continue to form the basis of many contemporary medium-term river channel studies (e.g. Gilvear et al., 2000; Winterbottom, 2000; Warburton et al., 2002; Liebault and Piegay, 2001, 2002). Advances in geographical information system (GIS) software capable of undertaking map and image registering, correction and planform change analysis (e.g. Leys and Werritty, 1999) have significantly improved the ease with which planform change quantification can be achieved. However, GIS seldom calculate, or report, spatial errors inherent in the quantification process and published rates of planform change are not normally accompanied by error estimates. This is despite Downward et al.’s (1994) assertion that rates of river channel change are only valid if it can be demonstrated that the amount of change in the measured parameters exceeds the errors in the measurement process.

In response to this lack of error quantification, Mount et al. (2003) present methods for estimating and propagating error in bankfull width measurements made from aerial photographs in a GIS, and identifying the range of temporal and spatial resolutions in which aerial photograph analysis of channel parameter change is valid. However, their error estimation methods are limited to bankfull width and ignore the other commonly used index for assessing planform change, lateral shift (Gurnell et al., 1994; Leys and Werritty, 1999). This paper presents a methodology to estimate error in lateral shift. It also investigates the impact of applying a range of simplifying assumptions in the error estimates and highlights the need for careful justification of assumptions.

Channel midpoint determination and errors

Analyses of lateral channel shift normally quantify the change in location of the channel midpoint along a predetermined section line (Figure 1) located perpendicular to the channel (e.g. Gurnell et al., 1994; Leys and Werritty, 1999).
The channel midpoint \((x_m, y_m)\) at any single section line in any single image can therefore be calculated from knowledge of the coordinates at which the section line and right and left bankfull lines intersect. This is expressed as:

\[
\begin{align*}
    x_m &= x_l + \frac{(x_r - x_l)}{2} = \frac{(x_l + x_r)}{2} \\
    y_m &= y_l + \frac{(y_r - y_l)}{2} = \frac{(y_l + y_r)}{2}
\end{align*}
\]

The measurement of the position of each bankfull intersect is subject to two components of error (Mount et al., 2003):

1. systematic, image registration component;
2. random feature identification component.

The United States National Standard for Spatial Data Accuracy states that the 95 per cent confidence interval of the root mean square error (RMSE) associated with individual image correction procedures should be used as a simple parameter to represent the image registration component (FGDC, 1998). However, this is a conservative parameter, and it may be considered reasonable to use the standard error of the RMSE. The feature identification component results from imprecision in the identification of bankfull. Identification of bankfull from aerial photography is extremely problematic. There are numerous available definitions of bankfull (Williams, 1978) including those that (1) associate the bankfull line with attributes of the visible sedimentary surfaces (e.g. the average elevation of the highest surfaces of the channel bars; Wolman and Leopold, 1957); (2) associate bankfull with the location of boundary features (e.g. the height of the lower limit of perennial vegetation; Schumm, 1960); (3) define bankfull according to attributes of a measured cross-section (e.g. the elevation at which the width-to-depth ratio becomes a minimum; Wolman, 1955). Where bankfull is to be defined according to the width-to-depth ratio of the channel or the attributes of sedimentary surfaces, bankfull measurement from two-dimensional aerial photography is likely to be inappropriate. However, where it is considered acceptable to define bankfull according to the location of boundary features, locating bankfull from aerial photography is likely to be appropriate so long as overhanging vegetation does not obscure banks. Similarly, where it is acceptable to define bankfull according to the location of the water–sediment interface, locating bankfull from aerial photo will be possible. However, defining bankfull in this way is likely to be appropriate only in situations where the channel has defined banks and will be problematic in channels containing numerous point and lateral barforms. The 95 per cent confidence interval, or the standard error obtained from repeated measurements of the bankfull location, may be used as a simple parameter to represent this component.

On the basis of the above sources of error, the \(x\) position of the left bank \((x_l)\) can be written:

\[
\bar{x}_l = x_l + e_s + e_r
\]

where \(x_l\) is the measured location of bankfull in the \(x\) direction, \(e_s\) is the systematic image registration component of error in the location of bankfull in the \(x\) direction, and \(e_r\) is the random feature identification component of error in the location of bankfull in the \(x\) direction.
The estimated total error in the $x$ position of the left bankfull measurement $e_x$ can hence be summarized:

$$e_x = e_r + e_s$$

(3)

Similar, expressions apply for the other positional components of bankfull.

### Propagation of error in a single image

If the systematic component of the error can be approximated by a uniform offset across the image then the propagation of this error will cancel in the calculation of the channel midpoint. This assumption is likely to be valid in most cases where the size of the river channel relative to the overall size of the image from which it is being measured is small and relief is low. The channel midpoint will thus retain the same systematic error $e_s$ as the left and right bank measurements. In this case, the problem can be simplified to one of propagation of the random error component and the subsequent simple addition of the systematic component.

If the feature identification error component is expressed in terms of pixels $p$ and is multiplied by the pixel resolution of the image $R$ to give a distance $pR$, then the random components of the midpoint location error can be written as:

$$e = pR$$

(4)

The random error components in the left and right bank positions can then be propagated in quadrature to find the random error components in the stream midpoint $x$-coordinate $x_m$, as:

$$e_{x_m} = \sqrt{\left(\frac{\partial x_m}{\partial x_l} \times e_x\right)^2 + \left(\frac{\partial x_m}{\partial x_r} \times e_x\right)^2}$$

$$= \sqrt{\left(\frac{1}{2} \times e_x\right)^2 + \left(\frac{1}{2} \times e_x\right)^2}$$

$$= \frac{1}{2} \sqrt{e_x^2 + e_x^2}$$

(5)

Similarly

$$e_{y_m} = \frac{1}{2} \sqrt{e_y^2 + e_y^2}$$

(6)

The total error in the midpoint coordinates $(x_m, y_m)$, including both the random and systematic error component, can now be written as:

$$e_{x_m} = \frac{1}{2} \sqrt{(e_x^2 + e_x^2) + e_s}$$

$$e_{y_m} = \frac{1}{2} \sqrt{(e_y^2 + e_y^2) + e_s}$$

(7)

where $e_s$ is the maximum systematic error over the $x$ and $y$ directions.

Further simplification can be achieved if it is assumed that the random error component is isotropic and invariant at each bank location. Such an assumption may be appropriate in infrequent cases where the bankfull location is defined with equal precision at both banks. However, it will not be appropriate in the majority of situations where the ease of identifying bankfull locations varies due, for example, to vegetation or one poorly defined bankfull location. Where the assumption is applied, the random components of the midpoint location error can be written as:

$$e_{x_m} = e_{x_r} = e_{x_l} = e_{y_r} = e_{y_l} = e_r = pR$$

(8)
Equation 5 for the random error component in the stream midpoint $x$-coordinate can now be simplified as:

$$e_{x_m} = \frac{1}{2} \sqrt{e_{x_1}^2 + e_{x_2}^2} = \frac{pR}{\sqrt{2}}$$

(9)

Similarly

$$e_{y_m} = \frac{pR}{\sqrt{2}}$$

(10)

The total error in the midpoint coordinates $(x_m, y_m)$, including both the random and systematic error component, can now be written as:

$$e_{x_m} = e_{y_m} = \frac{pR}{\sqrt{2}} + e_s$$

(11)

where $e_s$ is the maximum systematic error over the $x$ and $y$ directions.

### Propagation of error between image dates

Estimation of error in the channel midpoint for a single image date provides no information about uncertainty in measurements used to quantify process. Of much greater interest is the shift of the channel midpoint over time, here termed lateral movement $lm$. Estimation of error in lateral movement $e_{lm}$ requires propagation of the errors in the midpoint coordinates from two or more images.

Varying image quality and resolution, and the availability of ground control in registration procedures, mean that the total midpoint identification errors $e_{x_m}^{(i)}$ and $e_{y_m}^{(i)}$ can be expected to vary between the dates $i = 1, 2$. Hence it would not generally be safe to assume cancellation of systematic errors or that the random and systematic errors are invariant, between image dates.

The lateral movement of the channel midpoint between two images can be expressed as:

$$lm = \sqrt{(x_m^{(2)} - x_m^{(1)})^2 + (y_m^{(2)} - y_m^{(1)})^2}$$

(12)

where $(x_m^{(i)}, y_m^{(i)})$ are the stream midpoints on image dates $i = 1, 2$. The corresponding errors in these multi-date midpoint coordinates can be written as:

$$e_{x_m}^{(i)} = \frac{1}{2} \sqrt{(e_{x_1}^{(i)})^2 + (e_{x_2}^{(i)})^2 + e_s^{(i)}}$$

$$e_{y_m}^{(i)} = \frac{1}{2} \sqrt{(e_{y_1}^{(i)})^2 + (e_{y_2}^{(i)})^2 + e_s^{(i)}}$$

(13)

where $e = pR$ and where $R$ may vary between dates and $p$ may vary with both date and between banks.

In the general case, where there is a significant component of systematic error $e_s$, the between-date calculation of the error in lateral movement needs to be computed through simple addition of errors using the triangle inequality. In this case a maximum bound on the error in the inter-date lateral movement $lm$ can be written as:

$$e_{lm} \leq \frac{\partial lm}{\partial x_m^{(1)}} e_{x_m}^{(1)} + \frac{\partial lm}{\partial y_m^{(1)}} e_{y_m}^{(1)} + \frac{\partial lm}{\partial x_m^{(2)}} e_{x_m}^{(2)} + \frac{\partial lm}{\partial y_m^{(2)}} e_{y_m}^{(2)}$$

$$= \frac{x_m^{(2)} - x_m^{(1)}}{lm} \left\{ \frac{1}{2} \sqrt{(e_{x_1}^{(i)})^2 + (e_{x_2}^{(i)})^2 + e_s^{(i)}} \right\} + \frac{y_m^{(2)} - y_m^{(1)}}{lm} \left\{ \frac{1}{2} \sqrt{(e_{y_1}^{(i)})^2 + (e_{y_2}^{(i)})^2 + e_s^{(i)}} \right\}$$

$$+ \frac{x_m^{(2)} - x_m^{(1)}}{lm} \left\{ \frac{1}{2} \sqrt{(e_{x_1}^{(i)})^2 + (e_{x_2}^{(i)})^2 + e_s^{(i)}} \right\} + \frac{y_m^{(2)} - y_m^{(1)}}{lm} \left\{ \frac{1}{2} \sqrt{(e_{y_1}^{(i)})^2 + (e_{y_2}^{(i)})^2 + e_s^{(i)}} \right\}$$

(14)
Estimating error in channel movement

since

$$\frac{\partial m}{\partial x_m^{(1)}} = \frac{1}{2} \{ (x_m^{(2)} - x_m^{(1)})^2 + (y_m^{(2)} - y_m^{(1)})^2 \} \frac{1}{2} 2(x_m^{(2)} - x_m^{(1)}) = - \frac{(x_m^{(2)} - x_m^{(1)})}{lm}$$

and with similar expressions for \( \frac{\partial m}{\partial y_m^{(1)}} \), \( \frac{\partial m}{\partial x_m^{(2)}} \) and \( \frac{\partial m}{\partial y_m^{(2)}} \).

Equation 14 may be used directly to estimate an error bound on inter-date lateral movement \( lm \) and makes no assumptions of isotropy or error invariance between banks or dates. However, it is a complex and conservative method for estimating error, and alternative error bounds can be reasonably calculated following the consideration of a number of assumptions.

**Assumptions**

Assumption of directional independence in random error

Unless there is reason to suspect bias in the direction of the error associated with the left and right bank identification (e.g. effects of shadow), Equation 14 can be further simplified by incorporating the assumption that the random components of bank errors are isotropic:

$$e_{x_i}^{(i)} = e_{y_i}^{(i)} = p_{e_i}^{(i)} R_i^{(i)} \quad \quad i = 1, 2$$

then the error bound on the inter-date lateral movement \( lm \) can be written as:

$$e_{lm} \leq \left\{ \frac{x_m^{(2)} - x_m^{(1)}}{lm} + \frac{y_m^{(2)} - y_m^{(1)}}{lm} \right\} \left\{ \frac{R_i^{(1)}}{2} \sqrt{(e_i^{(1)})^2 + (e_i^{(1)})^2 + e_i^{(1)}} \right\}$$

$$+ \left\{ \frac{x_m^{(2)} - x_m^{(1)}}{lm} + \frac{y_m^{(2)} - y_m^{(1)}}{lm} \right\} \left\{ \frac{R_i^{(2)}}{2} \sqrt{(e_i^{(2)})^2 + (e_i^{(2)})^2 + e_i^{(2)}} \right\}$$

Using the triangle inequality, we note that if

$$(x_m^{(2)} - x_m^{(1)})^2 + (y_m^{(2)} - y_m^{(1)})^2 = lm^2$$

then

$$\left| x_m^{(2)} - x_m^{(1)} \right| + \left| y_m^{(2)} - y_m^{(1)} \right| \leq \sqrt{2} \cdot lm$$

The inequality 16 simply reflects the fact that the sum of the absolute values of the difference in the coordinates of the midpoints in the two image dates will be a maximum for a given \( lm \) when the coordinate pairs \((x_m^{(1)}, y_m^{(1)})\) and \((x_m^{(2)}, y_m^{(2)})\) form a 45° triangle with respect to the x and y axis directions.

Hence the maximum error bound on the inter-date estimate for \( lm \) (Equation 12) can be simplified to:

$$e_{lm} \leq \sqrt{2} \left\{ \frac{R_i^{(1)}}{2} \sqrt{(p_{e_i}^{(1)})^2 + (p_{e_i}^{(1)})^2 + e_i^{(1)}} \right\}$$

$$+ \sqrt{2} \left\{ \frac{R_i^{(2)}}{2} \sqrt{(p_{e_i}^{(2)})^2 + (p_{e_i}^{(2)})^2 + e_i^{(2)}} \right\}$$

Application of this assumption has the advantage that Equation 14 is simplified considerably. However, a slightly larger error bound results as a maximal value is applied to all of the random bank error components.

Assumption of homogeneous random error

If it is further assumed, as is the case for Equation 11, that within each date the random component of the error is the same for the left and right banks, \( p_{e_i}^{(i)} = p_{e_i}^{(i)} \), then the bound for \( e_{lm} \) can be further simplified as:
This assumption of homogeneous errors between banks is unlikely to be applicable in many cases as it will be extremely uncommon for the two bankfull locations within two or more individual images to be located to exactly the same degree of precision.

**Assumption of error independence between dates**

The error bound for $l_{m}$ given by Equation 17 is based on the conservative assumption that midpoint location errors may not be independent between image dates and errors from each date are combined using simple addition. The systematic component of the error may not be wholly independent if ground control points for image correction are limited and used to correct all images. However, where different ground control is used for different images it is more likely that the systematic component of the error will be independent. Similarly, where the main cause of the random component of the error occurs in all photographs (e.g. a tree obscuring bankfull) the conservative assumption is likely to be appropriate. However, where independence in the errors between image dates can be demonstrated the inter-date location errors can be added in quadrature, rather than in simple addition. This leads to a slightly smaller estimate for the lateral movement than that given in Equation 17, as:

$$
e_{lm} \leq \sqrt{\frac{R_{1}^{2}}{2} \left[ (p_{1}^{(1)})^2 + (p_{1}^{(2)})^2 + e_{i}^{(1)} \right]^2 + \frac{R_{2}^{2}}{2} \left[ (p_{2}^{(1)})^2 + (p_{2}^{(2)})^2 + e_{i}^{(2)} \right]^2}$$

**Case of insignificant systematic errors**

In cases where the systematic registration errors are negligible compared to the random feature identification errors, Equation 19 can be further reduced to give a smaller bound on the inter-date lateral movement error as:

$$
e_{lm} \leq \sqrt{\frac{R_{1}^{2}}{2} \left[ (p_{1}^{(1)})^2 + (p_{1}^{(2)})^2 \right]^2 + \frac{R_{2}^{2}}{2} \left[ (p_{2}^{(1)})^2 + (p_{2}^{(2)})^2 \right]^2}$$

Such a scenario is most likely to exist where large-scale imagery is available with excellent ground control for image correction, but identification of bankfull is made difficult through poorly defined banks or vegetation.

In choosing between the error estimates for $l_{m}$ given by Equations 14, 17, 18, 19 and 20 there is the need to carefully consider the nature of the assumptions and approximations underlying these estimates. In particular, independence of the factors causing error from image to image needs to be established.

**Worked example**

Use of the above equations, together with the impact of applying assumptions, is illustrated in the following worked example. The example assumes that the systematic error can safely be approximated by a uniform offset across each of the images. However, it is not assumed that the systematic error component is invariant between image dates. Initially, no assumptions of isotropy or error invariance between banks or dates are made and the conservative error bound is established using Equation 14. However, the assumptions are subsequently relaxed via application of Equations 17 and 19, and the impact of assuming isotropy in the random error component and error independence between image dates is highlighted.

Figure 2 shows a common fluvial scenario of channel migration around a meander. Bankfull lines, visible on two image dates, are shown together with the section line along which bankfull locations have been measured. Universal Transverse Mercator coordinates of the bankfull locations as measured from the imagery are given for each bankfull location and presented in Table I.

Midpoint $x$ and $y$ coordinates for each of the two image dates can be determined from Equation 1:
The lateral movement of the channel along the section line can then be determined by application of Equation 12:

\[
lm = \sqrt{(687990.75 - 687984.75)^2 + (573391.6 - 573417.2)^2}
\]

\[
= 26.3
\]

The channel centreline is therefore calculated as having moved 26.3 m. In order to calculate the error associated with this value, the magnitude of systematic error (i.e. the residual error following image correction procedures) and random error (i.e. the error in correctly identifying each bankfull location) must be known. The errors associated with this example are given in Table II.
Table II. Systematic and random errors associated with the identification of the bankfull locations given in Figure 2 and Table I

<table>
<thead>
<tr>
<th></th>
<th>$e_x$ (m)</th>
<th>$e_y$ (m)</th>
<th>$e_{x'}$ (m)</th>
<th>$e_{y'}$ (m)</th>
<th>$e_{x''}$ (m)</th>
<th>$e_{y''}$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date 1</td>
<td>3·45</td>
<td>0·85</td>
<td>0·77</td>
<td>0·85</td>
<td>0·77</td>
<td></td>
</tr>
<tr>
<td>Date 2</td>
<td>2·78</td>
<td>0·96</td>
<td>1·25</td>
<td>0·96</td>
<td>1·25</td>
<td></td>
</tr>
</tbody>
</table>

Error in the individual midpoint coordinates, determined above from Equation 1, can be calculated by application of Equation 13:

$$
e_{x}^{(1)} = \frac{1}{2} \sqrt{(085)^2 + (077)^2} + 345 = 402$$

$$
e_{x}^{(1)} = \frac{1}{2} \sqrt{(085)^2 + (077)^2} + 345 = 402$$

$$
e_{x}^{(2)} = \frac{1}{2} \sqrt{(096)^2 + (125)^2} + 278 = 357$$

$$
e_{x}^{(2)} = \frac{1}{2} \sqrt{(096)^2 + (125)^2} + 278 = 357$$

Where it is not safe to assume isotropy or error invariance between banks or dates, application of Equation 14 provides a conservative estimate of the error bound on inter-date lateral movement $l_m$:

$$e_{lm} = \left[ \frac{687990 - 687984}{263} \right] 402 + \left[ \frac{5733916 - 5734172}{263} \right] 402$$

$$= 912$$

This produces an error bound estimate of 9·12 m, more than 34 per cent of the total lateral movement. The size of the error bound estimate is surprising given the relatively small systematic error ($< 4$ m in all cases) and random error components ($< 1·25$ m in all cases) used in this example. It serves to demonstrate that measured lateral movement is unlikely to exceed error in cases where image registration is poor, where bankfull identification is problematic or where both registrations are poor and bankfull identification is problematic.

The estimation of the magnitude of the error bound can be simplified if the assumption of non-isotropy in the random error components is relaxed. Such a relaxation is likely to be reasonable in most analyses. Equation 17 can then be used to estimate the error bound:

$$e_{lm} \leq \sqrt{2} \left\{ \frac{1}{2} \sqrt{(085)^2 + (077)^2} + 345 \right\}$$

$$+ \sqrt{2} \left\{ \frac{1}{2} \sqrt{(096)^2 + (125)^2} + 278 \right\}$$

$$= 10·73$$

Note that as this is an error bound based on simplified isotropic assumptions it is slightly larger that the error bound of 9·12 m computed with known error estimated in both the $x$ and $y$ directions, as would be expected.

A reduction in the magnitude of the error estimation can be made if it can be demonstrated that errors in locating the channel midpoint are independent between dates (Equation 19). In such a scenario it is safe to add the error associated with each image date in quadrature rather than using simple addition as above:
\[ e_{lm} \leq \sqrt{\frac{1}{2} \sqrt{(0.85)^2 + (0.77)^2 + 3.45}^2 + \frac{1}{2} \sqrt{(0.96)^2 + (1.25)^2 + 2.78}^2} = 5.38 \]

Adding errors in quadrature reduces the error associated with the measured lateral movement to 5.38 m (20 per cent).

**Summary**

Estimation and propagation of error in measurements of river channel movement from aerial imagery requires knowledge of both the systematic error components (e.g. the root mean square error of correction procedures) and the random error components associated with the precision of feature identification. In addition, it is necessary to consider a number of possible assumptions relating to the homogeneity/heterogeneity of error components, directional and temporal independence of errors and the magnitude of systematic error components compared to random error components. The magnitude with which the error bound can be reduced following the application of certain assumptions presents a clear temptation for the analyst to estimate error using the method most favourable to their desired outcomes (in the worked example, Equation 19). The temptation for the analyst to apply assumptions inappropriately is likely to be particularly acute where the magnitude of the error estimated using the most conservative methods is close to the magnitude of the measured change. Consequently, it is vital that, where assumptions are applied and lower magnitude error bounds are presented, clear evidence to support the decision is also given. Where such evidence cannot be provided, analysts should choose more conservative estimates of error.

**References**


