Accuracy assessment of georectified aerial photographs: Implications for measuring lateral channel movement in a GIS

Michael L. Hughes*, Patricia F. McDowell, W. Andrew Marcus

Department of Geography, University of Oregon, Eugene, OR 97403-1251, USA

Received 3 September 2004; received in revised form 8 July 2005; accepted 11 July 2005

Abstract

Aerial photographs are commonly used to measure planform river channel change. We investigated the sources and implications of georectification error in the measurement of lateral channel movement by testing how the number (6–30) and type (human versus natural landscape features) of ground-control points (GCPs) and the order of the transformation polynomial (first-, second-, and third-order) affected the spatial accuracy of a typical georectified aerial photograph. Error was assessed using the root-mean-square error (RMSE) of the GCPs as well as error in 31 independent test points. The RMSE and the mean and median values of test-point errors were relatively insensitive to the number of GCPs above eight, but the upper range of test-point errors showed marked improvement (i.e., the number of extreme errors was reduced) as more GCPs were used for georectification. Using more GCPs thus improved overall georectification accuracy, but this improvement was not indicated by the RMSE, suggesting that independent test-points located in key areas of interest should be used in addition to RMSE to evaluate georectification error.

The order of the transformation polynomial also influenced test-point accuracy; the second-order polynomial function yielded the best result for the terrain of the study area. GCP type exerted a less consistent influence on test-point accuracy, suggesting that although hard-edged points (e.g., roof corners) are favored as GCPs, some soft-edged points (e.g., trees) may be used without adding significant error. Based upon these results, we believe that aerial photos of a floodplain landscape similar to that of our study can be consistently georectified to an accuracy of approximately \( \pm 5 \) m, with \(~10\%\) chance of greater error. The implications of georectification error for measuring lateral channel movement are demonstrated with a multiple buffer analysis, which documents the inverse relationship between the size of the buffers applied to two channel centerlines and the magnitude of change detected between them. This study demonstrates the importance of using an independent test-point analysis in addition to the RMSE to evaluate and treat locational error in channel change studies.

© 2005 Published by Elsevier B.V.

Keywords: Channel change; Channel migration; Georectification; Aerial photographs; Geospatial error; GIS

* Corresponding author. Fax: +1 541 346 2067.
E-mail address: mhughes1@uoregon.edu (M.L. Hughes).
1. Introduction

Aerial photographs are rich sources of information on historical river conditions (Trimble, 1991; Lawler, 1993) and have been widely used to track the historical planform evolution of river systems (e.g., Lewin and Weir, 1977; Petts, 1989; Gurnell, 1997; Surian, 1999; Graf, 2000; Winterbottom and Gilvear, 2000; O'Connor et al., 2003; plus many others). Historical planform channel analysis typically involves the co-registration of aerial photos and maps from different years so channel positions can be analyzed in overlay. Since the 1980s, the development of desktop GIS software and improvements in remote sensing and digital scanning technology have enabled users to more efficiently scan and co-register aerial photos; however, spatial error in digital imagery (including scanned aerial photos) is inevitable and can impart inaccuracies in measurements of lateral channel movement.

While there is widespread recognition in the GIScience community of the sources, types, and implications of locational error in geospatial data sets (Chrisman, 1982, 1992; Goodchild and Gopal, 1989; Unwin, 1995; Leung and Yan, 1998), fluvial geomorphologists have generally ignored the magnitude of geospatial error in relation to geomorphic change or have used only Root Mean Square Error (RMSE) as a measure of this error (e.g., Urban and Rhoads, 2004). Only recently have fluvial geomorphologists begun to embrace geospatial error as an independent research topic (e.g., Mount and Louis, 2005). Consequently, despite the development of approaches for measuring positional accuracy of linear features (e.g., Goodchild and Hunter, 1997; Leung and Yan, 1998) and recognition of the inherent problems of positional error on maps of rivers (Hooke and Redmond, 1989; Locke and Wyckoff, 1993) and lakes (Butler, 1989), there is no widely supported conceptual framework for evaluating and treating positional error on digital imagery in the measurement of lateral channel movement.

In this article, we seek to identify the magnitude and controls of geospatial error in georectified aerial photos and to address the implications of this error for measuring lateral channel movement. Accordingly, we raise the following questions:

(i) How is the locational accuracy of georectified aerial photos affected by the number and type of ground control points (GCPs) and the order of polynomial transformation used in georectification?
(ii) Is root-mean-square error (RMSE) a good proxy of overall georectification error?
(iii) What are the implications of georectification error for quantifying lateral channel movement and how can such error be minimized?

We address these questions using repeated georectification of an aerial photo showing the Umatilla River in northeastern Oregon. The quality and scale of this imagery is typical of those used throughout North America and many other parts of the world to reconstruct river histories. This article is the first phase of a broader study to evaluate channel and floodplain change resulting from large floods in selected rivers of the U.S. Pacific Northwest.

2. Background

GIScience and remote sensing play an increasingly significant role in geomorphological studies. Some recent examples of topics that have benefited from advances in the generation and handling of digital geospatial data include (but are not limited to) mapping and modeling of: fluvial erosion (Finlayson and Montgomery, 2003), complex terrain (Wilson and Gallant, 2000), mass wasting (Roering et al., 2005), mountain topography (Schroder and Bishop, 2004), historical channel change (Leys and Werrity, 1999; Collins et al., 2003), and river habitats (Marcus et al., 2003) and depths (Fonstad and Marcus, 2005). While many studies have developed methods for using digital data (e.g., aerial photos, satellite images, historical maps, and digital elevation models) to address traditional research topics, relatively few studies have rigorously addressed the effects of geospatial data quality on the results of geomorphic analyses (although see Holmes et al., 2000; Mount et al., 2003; Mount and Louis, 2005). Therefore, geomorphologists currently using digital geospatial need to better understand how the quality of geospatial data may affect analyses of digital data sets and to understand what factors control such data quality. Development of error-sensitive change detection methods depends on this knowledge. As GIScience continues
to better establish a theoretical basis in geography, opportunities are emerging for geomorphologists to undertake GIScience studies aimed at better understanding the applicability and limitations of digital geospatial data in their research.

2.1. General notes and terminology

Before aerial photos can be overlaid to map channel change in a GIS, they must be scanned and co-registered. Aerial photo co-registration refers to the conversion of digitally scanned photos to a common projection and coordinate system. Co-registration is usually achieved by georegistering individual photos to the same base layer. Digital orthophotographs (DOQs) and topographic maps (DRGs, digital raster graphics) are typically used as base layers.

Several techniques are available for co-registration of digital aerial photographs in a GIS, including aerotriangulation, orthorectification, and polynomial transformation. Each of these techniques has advantages and disadvantages that make it appropriate for specific applications. Aerotriangulation and orthorectification are typically used only when polynomial georectification fails to yield acceptable results. During aerotriangulation, GCPs are forced to have identical coordinates on the target (unregistered) layer and (georeferenced) base layer, thereby causing the image to be warped along triangulated edges rather than at point locations. This process requires a large number of GCPs for high accuracy and can therefore be difficult to apply in river change analysis because the number and distribution of GCPs are often limited. Moreover, error on triangulated photos varies in a nonsystematic fashion, complicating error analysis and application of buffers for reducing error and uncertainty during change detection. By contrast, orthorectification can provide high degrees of geospatial accuracy, but is less commonly employed by geomorphologists because it requires sophisticated software and is generally more labor- and data-intensive.

In this article, we evaluate polynomial georectification, which is readily applied to large sets of aerial photos (e.g., photos from flight lines along a river), can be performed with most commercially available GIS software packages, and is widely used for co-registration of aerial photos. When coupled with pixel resampling to correct for image warping during transformation, the process is called polynomial georectification. After scanning the original paper photo to create a digital file, polynomial georectification is performed in three steps: (i) matching of ground-control points (GCPs) on the scanned photo image and base layer, (ii) transformation of the GCP coordinates on the scanned image from a generic raster set to a geographical projection and coordinate system, and (iii) pixel resampling.

2.2. Aerial photo scanning

During the scanning procedure, the user defines the type (color versus gray scale) and resolution (dots per inch or d.p.i.) of the scan. Color and gray scale photos are customarily scanned into color and gray scale digital images, respectively. Because some data are “lost” in this digital conversion, users tend to maximize the resolution of the scan to improve image quality, however, users should consider the resolution of the base layer to which the digital photo will be registered before selecting a scan resolution. Scanning to a pixel resolution of 0.1 m, for example, makes little sense if the base-layer resolution is 2.0 m. Data loss during photo georectification, which includes pixel resampling (discussed below), may be minimized if the resolution of the scanned photo and georeferenced base layer are similar.

2.3. GCP selection for channel change analysis

The number, distribution, and type of GCPs can affect the accuracy of polynomial georectification, and researchers investigating river channel change have offered different guidelines for GCP selection. In examining historical planform change using scanned maps, Leys and Werrity (1999) noted that GCPs should be widely distributed across the image to provide a “stable warp,” while Richards (1986) and Campbell (2002) advised that the majority of control points should be located around the edge of the image with several uniformly spaced points in its central portion. While these suggestions may be appropriate for satellite images that have relatively little error due to topographic variations, or for scanned maps with constant scale variations across their projections, they are not necessarily well suited for historical aerial

M.L. Hughes et al. / Geomorphology 74 (2006) 1–16
photos, which usually have GCPs and areas of analytical interest that are unevenly distributed across the image over space and time, particularly in rural or forested settings. Moreover, better accuracy may be obtained by concentrating GCPs near the features of interest rather than across the entire aerial photo. This is particularly true with river channels, which tend to flow through floodplains of low relief and may be surrounded by valley walls of relatively high relief. Selecting GCPs that are far removed from the river channel may unnecessarily skew the transformation toward topographically complex areas not representative of the river channel and floodplain.

In addition to GCP distribution, GCP type can affect georectification accuracy. For the purposes of this study, we define two types of GCPs: hard and soft points. Hard points are features that have a sharp edge or corner, so their locations can be pinpointed. Hard points may include features such as building corners, road intersections, fences, and sidewalks. Soft points are features with irregular or fuzzy edges, such as rock outcrops and the centers of individual trees and shrub clusters. Because it is more difficult to pinpoint a soft point and because soft points may change over time (e.g., as when a tree grows larger), the choice of soft rather than hard points can affect overall georectification accuracy. However, in order to have enough GCPs for polynomial georectification, particularly in riverine environments, it is sometimes necessary to intermix hard and soft GCPs; therefore, soft points often cannot be categorically excluded.

Another challenging aspect of locating GCPs on historical aerial photos is that the correspondence between features on photos collected years or decades apart is sometimes poor. Buildings, roads, fences, trees, and other similar features can be moved, obliterated, or altered over time. Even in developed areas, GCPs may be difficult to locate and users are often faced with using a sub-optimal number, type, or spatial distribution of GCPs.

### 2.4. Polynomial georectification, transformation order, and RMSE

Polynomial transformation is applied to unregistered raster images (including scanned aerial photos) using linear and nonlinear functions. Polynomial transformations are named by their order, or the numerical value of the highest exponent used in the polynomial function. Therefore, first-order, second-order, and third-order transformations are linear, quadratic, and cubic transformations, respectively. When curvilinear (i.e., quadratic or higher) functions are used, the term “rubbersheeting” is sometimes applied, although this term may also be applied to aerotriangulation. Transformations using curvilinear functions are popular for aerial photos of the scale and terrain of this study because they can correct for some of the effects of both radial error (related to curvature of the earth) and geometric error (related to topography and camera lens distortion) and can therefore lend map-like qualities to a georectified photo without orthorectification. Remote sensing textbooks and photogrammetry manuals tend to emphasize the use of first-order and second-order transformations (e.g., Campbell, 2002; Leica Geosystems, 2003), because third- and higher order transformations tend to excessively warp digital images.

During polynomial transformation, a least-squares function is fit between GCP coordinates on the scanned image and base layer. This function is then used to assign coordinates to the entire photo. After transformation, GCPs on the photo and base layer will have slightly different coordinates, depending on the degree to which the overall transformation affects the proximal area of each GCP. The difference in location between the GCPs on the transformed layer and base layer is often represented by the total root-mean-square error (RMSE), a metric based in the Pythagorean Theorem and calculated for a coordinate pair by the equation (Slama et al., 1980)

\[
\text{RMSE} = \left[ (x_s - x_r)^2 + (y_s - y_r)^2 \right]^{1/2}
\]

where \(x_s\) and \(y_s\) are geospatial coordinates of the point on the source image; and \(x_r\) and \(y_r\) are coordinates of the same point on the transformed aerial photo. The RMSE for the whole image is the sum of the RMSE for each coordinate divided by the square root of the number of coordinate pairs.

### 2.5. Pixel resampling

Spatial transformations typically generate a different number of pixels in the transformed image than in the original image. Moreover second-order or higher
transformations can create pixels of variable size across the transformed image. A resampling step is necessary to equalize pixel size throughout the image and to assign values from the original image to the transformed image. There are a number of resampling approaches; nearest neighbor, bilinear, and cubic convolution (Campbell, 2002) resampling schemes are most common and are included in almost all GIS programs. We found that cubic convolution produced output photos best suited for interpretation of fluvial features because it smooths jagged edges along linear boundaries (e.g., river banks). Nearest neighbor resampling can create jagged feature boundaries, but does not alter the original pixel values, a critical element if spectral analysis of the image is planned. Bilinear resampling provides intermediate results in comparison to the other two techniques. If the reference and transformed images are approximately the same resolution, variations in resampling methods should not alter spatial location by more than approximately +0.5 pixels; however, because resampling methods affect image interpretation, we recommend experimentation with different resampling methods to select a method that works best for specific photo sets and research applications.

3. Study area

The Umatilla River is a gravel-bed river originating in the Blue Mountains of northeastern Oregon and flowing into the Columbia River at Umatilla, OR (Fig. 1). Its channel pattern ranges from meandering to anabranching, making it laterally mobile, particularly in reaches that are naturally unconfined or that have not been channelized. Because of ongoing efforts to improve water quality and restore native fisheries, the Umatilla River has been the focus of several completed and ongoing geohydrologic investigations, including a thermal TMDL study (ODEQ, 2001) and a hydrogeomorphic classification of riverine wetlands (Adamus, 2002). These studies have identified a need to better understand the river’s historical fluvial processes, how these processes have influenced contemporary fluvial landforms, and how river process-form relationships affect aquatic and wetland habitats.

Fig. 1. Location map of the Umatilla River Watershed.
important to native species. Channel modifications, including levees and revetments, are believed to degrade physical habitats and water quality by physically constraining the river channel and hampering lateral channel movements that may otherwise benefit habitat quality. Therefore, a detailed understanding of lateral channel movement serves a variety of river science and management needs.

4. Study design and methods

We hypothesized that georectification accuracy would improve when larger numbers of GCPs are used, when hard rather than soft GCPs are selected, and when a second-order polynomial is applied for spatial transformation. To test these hypotheses, we repeatedly georectified a 1964, 1:20,000 black-and-white aerial photo of the Umatilla River at Pendleton, OR (ASCS, 1964), varying the hypothesized controls to evaluate their relative effects. The quality and scale of this photo was typical of historical aerial photos used for analysis of channel change. The photo was scanned at a resolution of 600 dots per inch (DPI) and saved as a JPEG file (Fig. 2). Although TIFF format is best for complete data preservation, the JPEG file format generated much smaller file sizes and did not compromise the ability to precisely locate GCPs at normal compression ratios (Zhilin et al., 2002). The 600 DPI scan resolution was chosen because it produced pixels of about 1 m, the same resolution as the base DOQ.

During each experiment, the image was georectified to the USGS 7.5-minute Digital Orthophoto Quad (DOQ) of Pendleton, OR using the georeferencing toolbar in ESRI’s ArcGIS 8.2 ArcMap software. For each experiment, we conducted trials whereby one of the three variables (number of GCPs, type of GCP, or polynomial order) was changed and the other two were held constant (Table 1). All images were rectified using cubic convolution resampling. After each trial, we used ArcMap’s field

Fig. 2. A portion of the aerial photo used for analysis. Photo was shot in 1964 by the Agricultural Stabilization and Conservation Service (ASCS) at a scale of 1:20,000. Location of the photo portion relative to entire photo shown by outline at upper left. The Umatilla River flows from right to left in this and subsequent images.
calculation utility to measure the distance between 31 corresponding test-points (Fig. 3H) on the georectified photo and DOQ. The distance between the corresponding test-points on the photos and DOQ represented locational error; a zero distance between points would indicate perfect co-registration (although we never experienced this result in practice). Only hard points were used for the 31 test-points. GCPs and test-points were located on or immediately adjacent to the river’s floodplain, according to availability, and within approximately 0.75 km of the river channel.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Factor addressed</th>
<th>Treatment</th>
<th>Control</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number of GCPs</td>
<td>Georectified same image with 6, 8, 10, 12, 14, 20, and 30 GCPs; measured positional error of 31 independent test points on image and DOQ</td>
<td>Used second-order transformation function; used hard GCPs only</td>
<td>Fig. 4</td>
</tr>
<tr>
<td>2</td>
<td>GCP type</td>
<td>Georectified same image with 10, 20, 30 soft and hard GCPs; measured positional error of 31 independent test points on image and DOQ</td>
<td>Used second-order transformation function on same number of GCPs</td>
<td>Fig. 5</td>
</tr>
<tr>
<td>3</td>
<td>Polynomial order</td>
<td>Georectified same image with 14 GCPs; using first-, second-, and third-order polynomial transformation functions; measured positional error between 31 independent test points</td>
<td>Used identical GCP for each transformation function on same number of GCPs</td>
<td>Fig. 6</td>
</tr>
</tbody>
</table>

4.1. Experiment 1: number of GCPs

Experiment 1 evaluated the degree to which the number of GCPs affected the overall georectification accuracy (Table 1). Trials with 6, 8, 10, 12, 14, 20, and 30 GCPs were conducted (Fig. 3A–G). The number and locations of GCPs used for the experiments approximately corresponds to the number and locations of GCPs that are typically available for this type of application. During these trials, only hard GCPs were used and the images were transformed using a second-order polynomial function, which

![Fig. 3. Spatial distribution of GCPs (A–G) and 31 independent test points (H) with respect to the Umatilla River channel (line) and floodplain (hatched). Boxes show extent of georectified aerial photo.](image-url)
yielded the best results during pilot trials. We plotted five indicators to evaluate the magnitude of and controls on georectification error: the RMSE of GCPs and the mean, median, 90th percentile cumulative error value and maximum distances between test-points on the georectified image and DOQ. The degree of correspondence between the reported RMSE and the summary statistics for the 31 test-points provided the basis for evaluating georectification accuracy.

4.2. Experiment 2: GCP type

Experiment 2 tested how using hard- versus soft-edged GCPs affected georectification accuracy. Hard-edged GCPs were defined as landscape features with permanent, easily identified corners or edges and mainly included building corners, but also included fence corners and street and sidewalk intersections. Soft-edged GCPs were defined as features with “soft” or fuzzy edges; in this study we used only isolated tree canopies for soft-edged GCPs. Trials were conducted to compare test-point error resulting from transformations based on 10, 20, and 30 hard or soft point GCPs (Table 1). A second-order polynomial transformation was used for all the experimental trials. Differences in median and range of test-point values from trial to trial provided the basis for evaluating the effects of test-point type on georectification accuracy.

4.3. Experiment 3: polynomial order

Experiment 3 tested how polynomial order affected georectification accuracy. Aerial photos were georectified with 14 identical GCPs using a first-, second-, and third-order polynomial transformation function. We chose 14 GCPs based on the results of Experiment 1, which showed that RMSE did not substantially improve when more than eight GCPs were used. Therefore, we believed that 14 GCPs would be more than sufficient to limit the number of GCPs as a factor affecting comparisons of photos georectified with different polynomial functions. Differences in the median and range of test-point values from trial to trial provided the basis for evaluating the effects polynomial order on georectification accuracy.

5. Results

5.1. Number of GCPs

Fig. 4 displays the results of Experiment 1. RMSE initially increased from <1.0 to ~4.0 m as the number of GCPs increased from six to eight, while the independent test-point mean, median, and 90th percentile cumulative error decreased. With eight or more GCPs, the RMSE and the mean and median test-point errors were relatively insensitive to number of GCPs, remaining at ~4.0 +0.75 m; however, the 90th percen-
tile value of test-point errors continued to improve as GCP number increased to 30. When 30 GCPs were used the RMSE converged with the mean, median, and 90th percentile error values of test-points to $4.0 \pm 1.0$ m.

5.2. GCP type

Comparison of test-point distributions shows that GCP type has little effect on the median value of test-point error; however, soft-point transformations dis-

![Boxplot showing GCP type versus distribution of test-point error, stratified by 10, 20, and 30 GCPs. GCP types include hard (hd) and soft (sf) points. Central tendency is the median, and boxes represent inner quartile ranges (25th–75th percentile) of test points. Vertical lines indicate 1.5 times the interquartile range or the median plus or minus the extreme value, depending on which value is less. Asterisks indicate individual extreme values.](image1)

![Boxplot showing polynomial order versus error distribution for the 31 test points. Interpretation of boxplot bars and lines is as in Fig. 5.](image2)
played a greater range of error with higher outliers (i.e., larger errors) than the hard-point transformations (Fig. 5). For soft-point transformations, the median and upper range of test-point values increased from 10 to 20 GCPs, but then decreased from 20 to 30 GCPs. In contrast, the median and upper range of test-point values from hard point transformations consistently decreased as more GCPs were added.

5.3. Polynomial order

Fig. 6 shows the effect of polynomial order on test-point error. The second-order transformation yielded the best results with the lowest and smallest inner quartile range, although the median error was similar to that of the first-order transformation. The third-order transformation displayed much higher error values than either the first- or second-order transformations.

6. Discussion

6.1. Experimental results

Results of this study support the hypotheses that georectification accuracy improves when larger numbers of GCPs are concentrated within an area of interest (although this effect is not reflected by the RMSE values), when hard rather than soft GCPs are selected, and when a second-order transformation is used. While these hypotheses may be intuitive, results of this study reflect the relative sensitivity of georectification accuracy to its user-defined controls.

With respect to the number of GCPs, RMSE remained approximately the same when 8 or more GCPs were used (Fig. 4) and displayed little variability when 12 or more GCPs were used. The lack of significant improvement in the RMSE with additional GCPs is not surprising in the riverine landscape of the Pendleton area (Fig. 2). RMSE will improve with more GCPs only if the additional GCPs improve the fit of the polynomial function. In our low lying, relatively flat river landscape, adding more than 8 GCPs provided little additional information necessary to correct for average image displacement and topography across the photo. In fact, adding more GCPs can increase the RMSE, because the polynomial must be fit through a larger scatter of points, potentially creating larger residuals (e.g., note the ~1 m increase in RMSE moving from 10 to 12 GCPs in Fig. 4). This increase in RMSE may arise from displacement error from the addition of more topographic variation or from the use of additional GCPs that are imprecisely located.

As with the RMSE, the mean and median errors associated with the 31 test-points continued to improve as more GCPs were added. This result is consistent with the statement of Unwin (1995, p. 552) that RMSE does not capture spatial variations in error. This phenomenon is reflected in the 90th percentile values of test-points, which continued to improve as more GCPs were used and local topography was better represented in the transformation. Also, the 31 test-points were concentrated in one side of the photo (Fig. 4H) because of the clustering of hard points in that area; as more GCPs in this area were used, the error improved (note the locations of the GCPs in Fig. 4A–G relative to the test-point locations in Fig. 4H). Thus, the RMSE provided a reasonable estimate of the central tendencies of the error for the 31 independent test-points when 12 or more GCPs were used (Fig. 4), but was a poor indicator of the upper range of test-point error, which is driven by topographic variability in relation to GCP locations.

Like the number of GCPs, the order of the transformation polynomial exerted a clear influence on test-point error. The second-order transformation yielded the best results, probably because it was best able to capture spatial variations resulting from GCPs located both on and adjacent to the floodplain. A first-order transformation might work as well in areas where all GCPs could be located on the floodplain; but limiting GCPs to the immediate river area may not be an option with historical imagery and users are often faced with placing GCPs on terraces and hillslopes.

The third-order transformation generated poor results because of the excessive warping near the outer boundary of GCP locations, a classic problem with higher order transformations. Third and higher order transformations require GCPs far removed from the key features of interest in order to avoid boundary effects. Use of outlying points for GCPs would contradict our finding above that river studies should con-
strain GCPs to the area of the interest near the river. In general, it is hard to imagine a scenario where third or higher order transformations would be appropriate for studies of areas with similar topography.

In comparison to the number of GCPs and transformation order, GCP type exerted a less consistent influence on georectification accuracy. The median values of test-points derived from hard- and soft-point transformations were generally similar. However, the quartile ranges and outlying values were greater for the soft-point transformations when 20 or 30 GCPs were used. In contrast, with 10 GCPs both the median and inner quartile range were lower for the soft-point transformation, probably because the distribution of soft points was more favorable with respect to the 31 independent test-points. Results suggest that hard points should ideally serve as the basis for polynomial georectification, but that some soft points may be used without significantly changing the average transformation error or overall georectification accuracy.

These results have significant implications for understanding the positional accuracy of rivers and other landscape features on georectified aerial photos. First, GCPs on historical aerial photos are typically limited in number, so transformations are often generated from a limited number of GCPs that may or may not be representative of key areas of interest. The “average” positional accuracy in such cases may therefore be acceptable, but local errors, perhaps critical to the measurements, may be missed. Second, users tend to remove “rogue” points to improve RMSE. Our results suggest that, contrary to intuition, this practice may actually diminish georectification accuracy in key areas where the additional GCP(s) may otherwise improve accuracy. Third, tracking the relation between RMSE and number of GCPs may be misleading because using more GCPs can result in better transformations, even when the RMSE appears to have stabilized. In general, increasing the spatial density of GCPs within an area of interest (when possible) can reduce the overall range of error for that area and potentially for the entire image.

6.2. Implications for measuring lateral channel movement in GIS

Most approaches for measuring lateral channel movement with aerial imagery fall into one of two categories. Leopold (1973) introduced the concept (since used by many authors: e.g., Gurnell et al., 1994; O’Connor et al., 2003) of measuring the change in distance of the intersection of the channel centerline (or margin) with a series of floodplain or cross-valley transects. This method generates a set of change-distance measurements, the number of which depends on stream length and transect spacing. A second approach treats the floodplain and channel as rasters or polygons that can be mapped on aerial imagery to determine migration rates over time (e.g., Graf, 1984; Urban and Rhoads, 2004). In this approach, channel locations from sequential images are overlaid to calculate changes in channel area (m²) per unit channel length (m), and therefore a distance of migration (m²/m=m) for each river-length unit. Both approaches rely on image overlay, making them sensitive to geospatial error on component layers.

Alongside these two approaches of channel change detection, researchers have adopted several approaches to treat geospatial error in the measurement of channel change. Two approaches are common: (i) treating error as negligible with respect to the magnitude of geomorphic change, and (ii) applying buffers within which any apparent “change” is attributed to error and therefore disregarded. Until recently, many authors have adopted the first of these approaches without evaluating the effects of error on change measurements; however, the growing emphasis on remote sensing and GIS techniques in fluvial geomorphology has begun to shed light on issues of scale and error in geomorphic analyses (e.g., Gilvear and Byant, 2003; Marcus et al., 2004), prompting some researchers to recognize the value of error-sensitive change detection methodologies. For example, Urban and Rhoads (2004) presented an approach for buffering channel centerlines during measurement of lateral channel movement by applying a value of twice the RMSE error of the georectified photo; however, because our results indicate that RMSE can be a poor metric of georectification accuracy, we suggest that when possible buffer size be based on an analysis of independent test-points distributed across an area of interest.

To illustrate this concept, we calculated cumulative error probabilities (using a cumulative frequency function) for georectification errors of the 31 test
points in Experiment 1 (Fig. 7; see description of data in Section 5.1). These data can be used to specify channel centerline buffers according to the “risk” of error deemed acceptable by the user. In this case, we believe that aerial photos similar to the test photo can be georectified to an accuracy of approximately ± 5 m of the base layer coordinates with approximately 30 GCPs and an approximate 10% chance of encountering greater error within the area of interest; however, the relation between the optimal number and location of GCPs will vary among photos of different scale and regions of different topography, so the results from our analysis should not be used to prescribe a minimum number of GCPs in other studies. Rather, Fig. 7 should be viewed as one approach to defining error probabilities and change detection thresholds. In general, the magnitude of errors we documented in this study is consistent with that of other channel change studies that employed aerial photos (e.g., Lewin and Hughes, 1976; Gurnell et al., 1994; Winterbottom and Gilvear, 2000; Urban and Rhoads, 2004;) and digitally georeferenced satellite imagery (Zhou and Li, 2000), suggesting the existence of error thresholds across remote sensing platforms.

Buffer size can strongly affect change detection capability. Fig. 8 demonstrates the effects of buffer size on the measurement of lateral channel movement on a 2-km test reach of the Umatilla River. Pre- and post-flood aerial photos dated 1964 and 1971 were georectified with 10, 20, 30 GCPs. Wetted channel centerlines were then digitized from each of these photos. Buffers corresponding to the 90th percentile value of test-point error for 10, 20, and 30 GCPs (5-, 7.8-, and 10.8-m buffers, respectively; see Fig. 7) were applied to each side of the corresponding centerlines and a series of lateral movement polygons were generated by extracting from the GIS areas between the two centerline buffers. These polygons, representing areas of lateral channel movement, were then cut into smaller polygons along 50-m cross-valley transects. Finally, the area of these transect polygons was plotted versus distance downstream.

Fig. 8 demonstrates the inverse relationship between buffer size and the magnitude of measurable lateral movement. Where lateral channel movement is greatest (e.g., transects 11–14), percent differences in measured lateral movement across buffer sizes are small. In comparison, percent differences in measured channel change across buffer sizes are large where channel movement is more subtle (e.g., transects 16–20). In areas of limited channel movement, estimated rates of channel change may be more sensitive to buffer size than to actual channel movement.

While these results suggest that buffers based on RSME values can lead to channel-change measurements that are significant in error, the use of RMSE for buffer delineation has another other problematic tendency: RSME-based buffers tend to be used to
Fig. 8. (Above) Cross-valley transects and lateral channel movement polygons for three channel centerline buffers (5-, 7.8-, and 10.8-m) along a 2-km test reach of the Umatilla River. Transects were generated approximately every 50 m along the valley floor centerline. (Below) Bar graph of lateral channel movement versus distance downstream for the three buffer scenarios. Superimposed are line graphs showing percent difference in channel movement between the 5- and 7.8-m buffers and 5- and 10.8-m buffers. Figures are spatially aligned.
determine whether change has taken place despite the possibilities of true channel change within the RMSE buffer and no channel change outside it. Alternatively, we suggest that change detection be viewed in the context of the probability that measured change is real and that error probability be based on analyses of independent test points (Fig. 7). Termed the “empirical probability approach”, this approach avoids the assumption that all channel movements within the buffer size are not real and that all movements outside the buffer are real. Alternatively, researchers using the empirical probability approach can specify the probability of measuring actual change at their discretion and proceed with channel measurements knowing the likelihood that georectification error is affecting their measurements. This approach may be particularly useful in areas where channels are relatively confined (e.g., transects 16–21) and measured changes are often less than the RSME. Also, this approach consistent with the probability-based approaches for reporting change advocated by Graf (1984, 2001) and implemented in GIS by Graf (2000) and Winterbottom and Gilvear (2000).

Despite its shortfalls as an error indicator, RMSE is still quite useful in reconstructing channel change with aerial photos. In particular, because RSME is readily calculated for each individual photo as the image is georectified, providing a basis for varying the buffer size from image to image if necessary. In the case of the Umatilla River, we believe the error probability functions we developed for the Pendleton photo (Fig. 7) can be applied across many stream segments in that basin because the RMSE on other photos is similar, the topography from photo to photo is reasonably constant, and georectification methods have followed a consistent protocol; however, in basins (or portions of basins) with variable topography or inconsistent photo resolution and quality, development of probability functions for multiple photos would likely be necessary. In these cases, RMSE is a useful tool to screen photos that may require more detailed error analyses. We recognize the time costs associated with developing multiple probability functions and corresponding buffers must be weighed against the benefits of their application. In many fluvial hazard and river restoration studies, we believe that this cost–benefit would be justified by the improvements in information on channel movement rates and processes allowed by the empirical probability approach.

7. Conclusions

Results of this study show that the RMSE and the central tendency of locational error for 31 test-points were relatively insensitive to GCP number when eight or more GCPs were used. The 90th percentile cumulative error values of test-points, however, consistently decreased (i.e., improved) as more GCPs were used (Fig. 4), indicating that the upper range of georectification error can be significantly reduced by using more GCPs. We attribute the reduction in test-point error to a higher spatial density of GCPs within the area of interest and a better fit to local topography. Using more GCPs improves georectification accuracy only when additional points are positioned to better incorporate the topography of the area of interest.

A second-order polynomial transformation generated the best fit (Fig. 6), providing sufficient flexibility to correct for the range of topographic variation typical of the terrace-floodplain environment of this study. A first-order polynomial transformation generated a similar median error, but had higher outliers from poor transformation in areas of higher elevation near the river. First-order transformations may be appropriate for channel change studies if GCPs could be limited to the floodplain, but this may be impractical with historic photos of rural or forested settings. A third-order polynomial transformation generated poor results because of image warping at the outer GCP locations. The need to avoid edge effects by including GCPs far from the river suggests that third or higher order polynomial transformations are probably inappropriate for most river change studies.

The use of hard or soft GCP points did not dramatically affect median rectification errors, although the hard points generated fewer high-error values (Fig. 5). The similarity of results across GCP types indicates they can be intermixed without introducing spurious amounts of error.

Results clearly demonstrate that while RMSE may be an acceptable proxy of average error, it is generally a poor indicator of overall georectification accuracy across a photo. Therefore, using RMSE for error esti-
mates and determination of buffer size may lead to over- or under-estimating the amount of true change, depending on the correspondence of the RMSE and the upper range of true error on the photo in an area of interest. We recommend that lateral movement measurements be based on empirical probability functions (e.g., Fig. 7), which are generated from a set of test-point errors independent of the GCPs. According to this study of a 1:20,000 image transformed with 30 GCPs and a second-order polynomial, a buffer distance of 5 m on each side of the channel centerline would remove ~90% of georectification error that may otherwise affect measurements of lateral channel movement. A 5-m value is equivalent to 1.25 times the RMSE for the 30 GCPs. Buffers of similar magnitude are likely to be necessary for error-sensitive photo-based studies of lateral channel movement. Researchers using aerial photos to measure channel change are encouraged to conduct similar error analyses in order to assess the magnitude of georectification error relative to the magnitude of channel migration. Accordingly, error probability should be explicitly stated so that photo-based studies of channel change may be better understood in the context geospatial error.

Acknowledgements

Research on this project was supported by the National Science Foundation, Geography and Regional Science award BCS 0215291 to P.F. McDowell and W.A. Marcus.

References


