Geomorphic change detection using historic maps and DEM differencing: The temporal dimension of geospatial analysis

L. Allan James *, Michael E. Hodgson, Subhajit Ghoshal, Mary Megison Latiolais

Geography Department, University South Carolina, Columbia, SC, 29208, USA

A R T I C L E   I N F O

Article history:
Received 19 June 2010
Received in revised form 29 September 2010
Accepted 6 October 2010
Available online 17 June 2011

Keywords:
Cartometry
Geomorphometry
Change detection
Historic maps
DEMs
Error analysis

A B S T R A C T

The ability to develop spatially distributed models of topographic change is presenting new capabilities in geomorphic research. High resolution maps of elevation change indicate locations, processes, and rates of geomorphic change, and provide a means of calibrating temporal simulation models. Methods of geomorphic change detection (GCD), based on grided models, may be applied to a wide range of time periods by utilizing cartometric, remote sensing, or ground-based topographic survey data to measure volumetric change. Advantages and limitations of historical DEM reconstruction methods are reviewed with a focus on coupling them with subsequent DEMs to construct DEMs of difference (DoD), which can be created by subtracting one elevation model from another, to map erosion, deposition, and volumetric change. The period of DoD analysis can be extended to several decades if accurate historical DEMs can be generated by extracting topographic data from historical data and selecting areas where geomorphic change has been substantial. The challenge is to recognize and minimize uncertainties in data that are particularly elusive with early topographic data. This paper reviews potential sources of error in digitized topographic maps and DEMs. Although the paper is primarily a review of methods, three brief examples are presented at the end to demonstrate GCD using DoDs constructed from data extending over periods ranging from 70 to 90 years.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The time domain is an important dimension of geomorphic mapping and geospatial modeling. The application of temporal analysis in GIScience has been anticipated for almost five decades and is receiving growing attention (Langran, 1992; Raper, 2000; Wilde and Cressie, 2000; Peuquet, 2003; O’Sullivan, 2005). Cartography – and by extension, much of geospatial science – is potentially four-dimensional with the planimetric dimensions, X-Y, forming the traditional basis, and the third dimension consisting of elevation or other attributes describing a statistical surface. Time may be regarded conceptually as the fourth dimension (Langran, 1992). Just as traditional cartography maps space in bounded areas, so the time dimension may have abrupt or transitional temporal boundaries. For historical reconstructions, the sequent snapshots produced by available maps or imagery define a space-time cube, but the temporal resolution tends to be coarse, so rates of change must be interpolated (Fig. 1A). Discrete temporal periods are often defined by the availability of reliable data rather than the occurrence of events. The time of each map or image brackets the period in which change occurred but does not specify the time or the agents of change (Langran, 1992). Where change occurred in space can be identified specifically, but how and precisely when and why the change happened must be inferred from other information. In geomorphic change detection (GCD), inferences about processes and times of events may often be made from knowledge of the record of natural events such as storms, floods, or earthquakes, and these inferences can improve estimates of rates of change (Fig. 1B).

Historic changes in geomorphic systems can be quantified with geospatial processing of empirical data from historical maps, airborne or satellite imagery, or field surveys. Where accurate historical topographic data are available, time-discrete elevation surfaces can be developed and registered to topographic data from one or more other times for quantitative comparisons. The development methods for digital elevation models (DEMs), described in this paper, generate static data layers, but differencing two or more sequential DEMs is a rudimentary form of spatially distributed dynamic geomorphological analysis. Even in the static mode, time-discrete DEMs can be used to identify locations of geomorphic stability or change, past trends, processes and rates of change, as well as to construct sediment budgets. They may also be used to calibrate dynamic models of change for greater time integration in GIScience. For example, accurate historic topographic reconstructions can be used to establish initial boundary conditions for continuous simulation models at higher temporal resolutions (Rumsby et al., 2008).

2. Volumetric geomorphic change detection (GCD) by DEM differencing

Change detection in remote sensing of environmental systems includes a wide range of techniques, including changes in spectra (surface brightness values), planimetry (2-dimensional position), or elevations...
where \( \Delta E_{ij} \) is the i, j grid value of the change in elevation model, \( Z_{1ij} \) is the i, j value of the early DEM, and \( Z_{2ij} \) is the i, j value of the later DEM. The resulting DoD represents reductions in elevation as negative values and increases in elevation as positive values. Determining the cause of this change (e.g. erosion, deposition, subsidence, anthropogenic modification, precision, accuracy, or uncertainty) is more challenging.

Of particular interest to this study is the extraction of topographic data from historical contour maps to allow construction of historical DEMs for DoD analysis. Most DoD studies have been concerned with temporal scales less than decadal based on field surveys or remote sensing data (Heritage et al., 2009). Historic reconstructions of greater duration require historical imagery or the use of cartographic data. Cartographic data are especially important for reconstructions of surfaces prior to the availability of stereoscopic aerial photographs or in heavily vegetated areas where conventional remote sensing methods cannot penetrate the canopy. If a high-resolution historic topographic map is available from a ground or canopy penetrating photogrammetric survey, this map may be used to develop an early DEM for the area. Historical reconstructions based on analysis of aerial photography can extend the time dimension back several decades under favorable conditions. The challenge of using historic remotely sensed imagery or topographic maps is first dependent on the source materials and processing methods for their construction (Hodgson and Alexander, 1990). For example, contour lines on many early (i.e. pre 1940s) topographic maps were “artistically” drawn with little intervening field observations between field measurements. Modern methods of topographic map construction (e.g. remote sensing based) use a comparatively dense set of observations (e.g. every few meters planimetrically) for contour construction.

The methods of volumetric change detection described in this paper are a subset of a broader set of comparison methods for spatial data. Four traditions in map or imagery comparison can be identified (Table 1). Comparisons may be made using cell-based or feature-based statistics or with spatial patterns. This paper is primarily concerned with the first two methods — accuracy of historic spatial data and methods of change detection. It begins with the importance of extending historical geomorphic research and GCD back in time, followed by a brief sampling of past studies and examples of historical reconstructions and DoD analysis. Limitations and uncertainties associated with historical reconstructions using maps, airborne imagery, DEMs, and DoDs are described, emphasizing topographic maps that can be used to extend volumetric GCD back in time. Finally, application of DoD analysis to an extended temporal scale is demonstrated with three case studies of fluvial and hill-slope systems. DEMs are developed from early 20th century large-scale topographic maps and differenced with modern DEMs from aerial photographic stereo pairs or Light Detection and Ranging (LiDAR) data to construct DoDs. These studies demonstrate the utility and limitations of the method for volumetric analyses of decadal to centennial change.

### Table 1

Methods of map comparison.

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Map accuracy assessments</td>
<td>comparisons with a reference map.</td>
</tr>
<tr>
<td>2. Change detection</td>
<td>differences between maps or images collected at different times.</td>
</tr>
<tr>
<td>3. Model comparisons</td>
<td>comparisons of model output to observed landscapes or to other model outputs.</td>
</tr>
<tr>
<td>4. Landscape comparisons over time</td>
<td>similar to change detection, but focus is on global (i.e., area wide) spatial metrics calculated from map data, and may be used to compare different geographical areas. Primarily used in landscape ecology but a growing trend in geomorphometry.</td>
</tr>
</tbody>
</table>

(Hodgson et al., 2000).
under consideration increases, the average rates of change decrease and the relevant time span that needs to be considered increases (Fig. 2):

“As the size and age of a landform increases, fewer of its properties can be explained by present conditions and more must be inferred about the past” (Schumm, 1991; p.52).

Given the strong scale dependency of process, accurate characterizations of global geomorphic changes and calibrations of landscape evolution models require a greater emphasis on historical studies. ‘Historical’ in the context of this discussion of scale should be extended to ‘stratigraphic’ or ‘geologic’ definitions of history, although this paper is primarily limited to cartographic records no more than 150 years in age. A perspective that links event-based processes to historical evolution is needed for a better understanding of geomorphic responses that are globally relevant. This article addresses GCD using geospatial technology with historical maps and imagery to measure morphological change and identify changes in processes over decadal and centennial time scales. Examples of these methods are reviewed and case studies are provided for a variety of environments and geomorphic systems.

4. Previous studies of volumetric Geomorphic Change Detection (GCD)

Using historical spatial data to measure changes has a long history. Cartometry has been practiced on relatively old maps to push GCD back in time (Hooke and Perry, 1976). The mainstay of GCD methods over the late twentieth century has been planimetric analysis based on aerial photogrammetric and field survey methods to generate DEMs at the meso scale (Lane, 2000; Lane et al., 2003; Hughes et al., 2006; Heritage et al., 2009). These methods have facilitated volumetric GCD not only for computing sediment budgets, but also to estimate sediment transport rates. For example, a morphometric method has been adopted by several river scientists as a means of computing bed-material transport rates in gravel-bed rivers. Measurement of bedload transport in gravel-bed rivers is notoriously difficult (Gomez, 1991), so methods have been developed based on sediment budgets from volumetric GCD. Topographic surveys can be combined with digital terrain-modeling to quantify and monitor river-channel changes (Lane et al., 1994). Increasingly, GCD is being conducted with measurements from remote sensing as methods improve for measuring volumetric changes in shallow submerged bars (Gauman et al., 2003; Fonstad and Marcus, 2005; Marcus and Fonstad, 2008). The morphologic method or inverse method uses volumes of sediment computed from change detection to develop morphological sediment budgets to infer the rates of sediment transport in gravel-bed rivers (Ferguson and Ashworth, 1992; Lane et al., 1995; Ashmore and Church, 1998; Brasington et al., 2003; Martin and Ham, 2005).

5. Assessing data quality

Accurate GCD requires the reconstruction of one or more historic geomorphic surfaces from which elevation changes can be computed. The quality and confidence in the historical topographic data available are usually the limiting factors in the accuracy and confidence in the resulting GCD. Quantitative assessments of uncertainty range from the precision of an instrument that was used to a full-blown uncertainty or error-budget analysis (Gottsegen et al., 1999; Hodgson and Bresnahan, 2004; Wheaton et al., 2009). Uncertainties in source materials, processing methods, classification, resolution, completeness, image registration, interpolation, and other forms of error and error propagation should be recognized before comparisons of sequential images are interpreted. A critical evaluation of data sources, exercised early in the process, may lead to the rejection of potential cartographic or DEM data sources for morphometric analysis. In some cases, qualitative evaluations of features depicted
on maps may alter interpretations of apparent changes on maps. For example, rectification of an 1859 map that was not sufficiently accurate for quantitative analysis of change was useful for qualitative referencing with a more precise 1881 map (Fig. 3). The north half of the 1859 source map was georeferenced to section corners of federal land survey resulting in a rectification accuracy of 44 m root mean square error (RMSE). The actual uncertainties are presumably larger because of other errors not accounted for, such as positions of survey corners on the map. Superimposition of the 1859 channel locations onto the more precise 1881 map shows that the 1881 dashed lines correspond to the 1859 channels, which allows dating of the 1881 features.

The feasibility of producing an accurate GCD increases as the magnitude of expected change increases; i.e., GCD quality depends on the strength of the signal being measured relative to the data quality. This relationship may be expressed as a signal-to-noise ratio in which the signal is actual geomorphic change and noise is introduced by error variability (adapted from Griffith et al., 1999):

$$S / N = V_{GC} / V_{EP}$$

where $S/N$ is the signal-to-noise ratio, $V_{GC}$ is variability caused by geomorphic change, and $V_{EP}$ is variability caused by errors. A high S/N value is desirable. At low values approaching one, geomorphic change is rarely known. Observed changes include errors and actual change, so $V_{GC}$ must be estimated after errors are known. Nevertheless, the concept is useful as it expresses the feasibility of a successful GCD as an inverse relationship between the degree of change and data quality. An analyst may be able to reduce data uncertainty, but geomorphic variability is inherent to the system. Thus, systems with large geomorphic change are more conducive to GCD. Errors are difficult to determine, so a conservative estimate of S/N based on data uncertainty in the denominator may be the appropriate measure. An important research area for DEM analyses going forward will be to develop standard methods of computing S/N values and, perhaps, establishing one or more threshold values for including analyses based on this criterion (cf. Wheaton et al., 2009).

Difficulties in distinguishing change from error can be seen in lateral shifts in a pair of lines measured from sequential maps or images. The lines may represent the edge or crest of a landform, such as a stream bank or dune ridge. Even with no geomorphic change, some degree of line offset is to be expected as inherent error introduced by a variety of cartographic or photogrammetric sources. Line offsets are commonly used as an error metric (‘sliver’ analysis) in assessments of cartographic error (Chrisman, 1989). The likely magnitude of such errors should be evaluated to assess the confidence in the GCD. Some, but not all error may be removed or minimized by image registration. Least square polynomial transformation, commonly used in map/image rectification, distributes the errors across the study area and it is important to report these error statistics. Spatial error statistics are often reported as a RMSE to estimate the potential error component in a GCD analysis. Alternatively, the Circular Map Accuracy Standard (CMAS) may be derived from the RMSE and reported for two-dimensional uncertainties (FGDC, 1998). These are not the only sources of uncertainty, however, and other error contributions will increase the resulting error (and decrease confidence). Therefore, values of spatial accuracy should be considered a minimum value of uncertainty in the data.

Change detection should be performed using historic data of sufficient quality to decrease the likelihood that much of the observed differences are caused by errors. In this context, data quality should be defined on the basis of uncertainty; a broader concept that includes errors, accuracy, and precision. Errors are used to describe differences between measured or recorded values and the actual value (i.e. the reference data), whereas uncertainties represent a broader assessment of discrepancies, imperfect knowledge, or vagueness of the data (Gottsegen et al., 1999; Mowrer, 1999). In most spatial databases used in geomorphic work, errors are not fully specified and uncertainty may be considerable. Uncertainties in surveys, maps, or imagery will be propagated onto DEMs and DoDs, so an evaluation of the quality of the end product depends upon knowledge about the spatial data used for computations and measurement and interpolation methods. For example, Butler (1989) found evidence of substantial “apparent” elevation change of + 20 m (65 ft) and extent in a mountain lake that was “field surveyed” for a 1904 topographic map and propagated to a 1938 topographic map. A variety of ways can be used to categorize geospatial error depending on the ultimate purpose (Veregin, 1989). The goal of some studies may be to determine amounts of error from a source, define confidence bounds for a dataset, or predict cumulative errors for new studies or scenarios. A general categorization often used is to separate error into two broad categories — errors in source materials and errors produced from the processing approaches to create products (i.e. DEMs, contour maps, slope maps, etc.). Walsh et al. (1987 and 1989) refer to the former as inherent error and the latter as operational error. In DEM analysis for GCD, inherent error is in topographic source data such as topographic maps or LiDAR point clouds. Operational errors may be introduced by filtering bare-earth points from LiDAR point clouds, interpolating from contour or point data, mis-registration during coordinate transformations, or computing topographic derivatives such as slope, aspect, or roughness. Recent work categorizes geospatial error into more refined subcategories of source errors and processing errors. For example, error budget modeling (Hodgson and Bresnahan, 2004) uses separate error categories for the LiDAR system, horizontal error, slope-related error, interpolation error, and reference data error. Separating error sources into refined categories for which errors are known (or may be modeled) allows for (1) computing individual contributions of unknown error sources and (2) predicting errors (and confidence) in similar studies with different error amounts.

Quality assessments require consideration of accuracy and precision. Accuracy is the degree to which measurements conform to reality. Accuracy metrics characterize bias or systematic error and may be estimated by comparisons with a reference map or data layer that best indicates the true value. Precision is the degree to which measurements conform to one another (Fig. 4). The precision of map measurements is limited by such factors as instrumental resolutions and spatial resolution of the source data. Outliers (blunders) are a third type of spatial data error. Seven elements of spatial data accuracy have been identified (Table 2). Common sources of uncertainty encountered with historical topographic data arise from positional inaccuracies, incomplete coverage, or temporal discrepancies in the source data. Lineage describes the source materials,
Table 2
Elements of Spatial Data Quality.
Adapted from Morrison (1995); cf. NIST (1994); FGDC (1998).

1. Lineage – data sources, time period, processing, transformations, etc.
2. Positional Accuracy – horizontal and vertical accuracies of features.
3. Attribute Accuracy – characteristics of facts about locations (features or thematic elements) including name, classification of objects, etc.
4. Completeness – selection criteria, generalization, definitions used, or omissions.
5. Logical Consistency – fidelity of relationships encoded in data structure; unique identifiers, consistent topology, etc.
7. Temporal Information – date of observation, updates, and valid period for data.

dates or times, processing methods, and decisions for producing the datasets. Lineage often helps subsequent scientists understand ‘apparent’ shortcomings of digital spatial data. Lineage, attribute accuracy, and other aspects of spatial data quality are covered elsewhere (Harley, 1988; G uptill and Morrison, 1995).

Measurements made from maps are the domain of cartometry, the science of extracting quantitative information from maps (Maling, 1989). Historical topographic data may be derived from cartographic, remote sensing, or field sources. Thus, the digitization of topographic data may involve cartometry if derived from maps, but also falls within the domains of photogrammetry and geomorphometry, a branch of the sciences of geomorphology and geomatics concerned with the quantitative measurement of topographic information (Pike, 1995; Pike et al., 2009). Since the 1940s, remotely sensed data have become the fundamental method for broad area mapping. Photogrammetry is the science of making reliable measurements from aerial photographs or remotely sensed imagery. The distortions inherent to airborne and satellite sensors are different than planimetric (orthometric) maps and, thus, require different approaches. Measuring three-dimensional attributes from such remotely sensed sources is also different than maps.

6. Uncertainties in cartometrics

Quantitative measurements of change require accurate spatial data for two time periods. In many cases, data for one or both of the two periods may have been collected relatively recently by modern remote sensing or mapping methods. The further back in time that information is sought, however, the more likely it is that historical reconstructions will rely on cartographic data. Cartometry, the “measurement and calculation of numerical values for maps” (ICA, 1973; cf. Maling, 1989), relies on the planimetric and topographic accuracy of maps that varies greatly. This discussion is focused on large-scale maps constructed within the past 200 years using rigorous contemporary cartographic standards. The precision and accuracy of cartographic data tends to decrease with age, although modern maps should not be assumed to be of sufficient quality for GCD. Whereas old maps often contain large uncertainties, the large historical information content may warrant analysis, especially where geomorphic change has been substantial.

6.1. Precision

Precision in GIScience may be characterized in the spatial domains (X–Y and Z) and temporal domain. DEMs are commonly characterized by the cell size, whereas irregular tessellations are characterized by the average sampling density (e.g. LiDAR post spacing). Unfortunately, the resolution of the source materials is often ignored, although this should be reported and included with interpretations. For example, using a GIS, DEMs may be created at any spatial resolution desired, even at spatial resolutions much finer than the original data (e.g. creating a 1 m × 1 m DEM from spatial observations at 100-m intervals). Thus, high data precision should not be assumed from cell size without knowledge of the source materials and interpolation methods used. The precision of the vertical dimension is less well understood. Geomorphologists often assume elevations are measured with high precision whereas, in practice, vertical precision from remote sensing sources, such as aerial photography or imagery, is typically in decimeters at best. Until the late 1990s most of the USGS DEMs were created and stored in whole feet or meters in the vertical dimension. Misuse of these DEMs often results in spatial artifacts, such as long plateaus and sharp discontinuities observed on surfaces in areas of low slopes (Carter, 1988). Such artifacts merely result from rapid changes in whole-number elevations across short distances (e.g. adjoining 30-m DEM cells).

6.2. Positional accuracy

Horizontal positional accuracy is of obvious importance to quantitative GCD, although historical maps with poor positional accuracy may still be valuable. Historical information from early maps for which accurate cartometry is not feasible may constrain the timing of specific geomorphic events. For example, approximate positions of alpine glaciers during the Medieval cold period can be documented from maps that are relatively imprecise. The primary limitation to quantitative cartometry, however, is positional accuracy — including the accuracy of original data collection, map production, quality of the media (e.g., stability and condition of paper maps), and errors introduced by data extraction and processing (Maling, 1989). Scale is a limiting factor in positional precision and governs the relative size of symbolization such as line widths and contour intervals. Even high quality early maps may have substantial bias with regard to positioning. Maps greater than 300 years old often contain relatively large inaccuracies such as inconsistent scale or orientation (Tobler, 1966; Livieratos, 2006).

Several methods have been used to measure positional accuracy on historical maps including coordinate methods that calculate correlations with modern latitudes and longitudes, and digitally converting map positions to a modern coordinate system and plotting on a reference map to derive error vectors (Hu, 2001). Cumulative positional errors on digital maps (RMSEp) can be estimated from three general sources:

\[
\text{RMSE}_H = \sqrt{\text{RMSE}_h^2 + \text{RMSE}_e^2 + \text{RMSE}_m^2 + \text{RMSE}_d^2}
\]

where RMSEh is the error introduced by the ground survey, RMSEe is the map error, RMSEd is the digitization error, and M is the map scale factor (Cheung and Shi, 2004). The error budget model, expressed in Eq. (3), assumes a linear increase in map error and digitization error with changes in the map scale factor. Some studies approximate ground survey error, map error, or digitization error simply from the precision of the instruments used in surveying, cartographic construction, or digital conversion. These instrumental parameters underestimate map uncertainty, however, because they do not account for the high variability in operator care, judgment, or ability. For example, the precision of digitizer tablets exceeded the ability of operators to generate accurate lines (Jenks, 1981; Keefer et al., 1991). Traylor (1979), Bolstad et al. (1990), and Henderson (1984) examined the error introduced by digitizing from tablets, a primary source of early digital data transformed from hardcopy maps. Most hardcopy-to-digital map conversions in the last couple decades utilize scanned maps and ‘heads-up digitizing,’ allowing for virtual zooming and typically result in less spatial error than observed from tablet digitizing. The resulting error in early topographic maps is a function of the plotting of control points, compilation of the map, redrafting, photographic reproduction, and printing. Maling (1989) suggested errors from survey and the various map creation errors to range from .42 mm to .73 mm (RMSE). In summary, it is important to understand the lineage of historic maps or digital databases to estimate the spatial errors.

Positional uncertainties may be constrained by cartographic standards for some 20th century maps that were constructed according to
standards. Error analyses often employ map accuracy standards as the value of map error for maps conforming to standards. In the U.S.A., national cartographic accuracy standards were established in the early 1940s (Marsden, 1960) and similar standards have been adopted internationally. U.S. National Map Accuracy Standards (NMAS) use an observed error threshold of 90% for all maps that bear the statement “This Map Meets National Map Accuracy Standards.” The error tests that 90% of ‘tested’ points on such maps (e.g. U.S. Geological Survey topographic quadrangles) are within a specified distance of the true position (Table 3). Standards for modern geospatial data in the U.S.A. are given by the Federal Geographic Data Committee (FGDC 1998). Horizontal standards are stated at the 95% confidence level defined by the radius of a circle assuming a normal distribution. Similarly, vertical standards are defined at the 95% confidence level. Of particular interest to geomorphologists are the reference data used in the statistical tests. The ‘tested’ points for such maps meeting NMAS or used in the FGDC NSSDA statistics are “well-defined”, such as street intersections, benchmarks, etc. rather than more ill-defined features such as ridgelines, riverine banks, etc. The reliance on well-defined points in statistical tests suggests that the ill-defined features typically of interest to geomorphologists are likely to be less-accurately positioned on maps.

6.3. Rectification and co-registration

Volumetric change detection requires that two spatial data sources are co-registered horizontally and vertically with minimal error. The objectives of co-registration should recognize differences between geodetic coordinate systems related to astronomically oriented graticules, planimetric references that measure distances or directions between locations of identifiable objects, and elevation references (Laxton, 1976; Blakemore and Harley, 1980; Lloyd and Gilmartin, 1987). Rectification of historical maps and images to a base map or image with precise planimetry unleashes the potential for multivariate GIS analysis utilizing ancillary spatial data, so referencing to a universal positioning system has advantages.

Prior to the availability and utility of GIS tools, accurate rectification of historical maps and aerial photographs with different scales and projections or internal distortion often required the use of expensive and cumbersome optical analogue equipment. Digital corrections of historic maps began in the 1970s (e.g. Ravenhill and Gillig, 1974; Stone and Gemmell, 1977). Modern digital rectification methods, available in standard GIS software packages, allow much of the systematic positional errors to be removed by automated coordinate transfer methods. Historical maps can often be co-registered with other spatial data to overcome scale and projection differences if a set of geographic control points (GCPs) can be identified across both maps. Most geomorphologists are not trained in coordinate transfer methods, however, and may uncritically use the default values in common software packages. For example, common procedures require only four GCPs and employ an affine transformation with ordinary least squares as the default method. Such transformations may be adequate to correct for different map scales and orientations, but they are inadequate when the images are based on different projections or contain distortions (Chrisman, 1999). Furthermore, planar affine transformations are generally inappropriate for large scale aerial photography. A camera model and appropriate transformation available in softcopy photogrammetric solutions should be used for registering such remotely sensed data. The polynomial transformations available in a common GIS cannot reproduce the radial distortions, relief displacement, and camera orientation present in an aerial photograph with a central perspective.

Rectification of historical maps and aerial photographs are conventionally done using ground control points (GCPs) that can be identified on the map and on a reference image, such as building corners, road intersections, or large boulders. A well-distributed set of GCPs is needed to produce the most accurate rectification. Whereas a camera model is most appropriate for geometric rectification of aerial photographs derived from a film camera, this approach is foreign to many geomorphologists. For small geographic areas that cover a relatively small portion of an aerial photograph, simple polynomial transformations may produce reasonable results. Hughes et al. (2006) examined the sensitivity of aerial photograph rectifications to GCP selection in channel-change studies and provide guidance on the number and type of GCPs that are needed. Substantial difficulties may be encountered in rectifying large-scale historical maps for geomorphic applications where reliable point features suitable for GCPs are scarce. Cultural features such as road crossings generally stand out on imagery, whereas natural landscapes often lack distinctive point features, especially in heavily vegetated areas. Attempts to use linear features do not generally constrain the rectification in both X and Y, and points on natural boundaries may migrate through time. For example, attempts to use environmental features such as woodland boundary patterns are associated with difficulties because of shifting positions through time (Lindsay, 1980), so the stability of landscape features is an important consideration (Lloyd and Gilmartin, 1987). This issue is problematic for geomorphic studies where cultural features, such as buildings and roads, may be absent and particularly troublesome for GCD studies where many landscape features have changed.

6.4. Completeness and temporal accuracy of planimetric data

Incomplete coverage on maps can lead to errors in historic geomorphic reconstructions and distort temporal accuracy such as when geomorphic features appear to change on maps. Apparent geomorphic changes may be misinterpreted from errors of cartographic commission or omission inherent to the source data. Errors of commission represent the inclusion of an object that was not present at the time of mapping. These errors may represent cartographic blunders such as inclusion without field verification of a feature shown on an earlier map but no longer present. Errors of omission occur when a feature was not mapped but was present during map data collection. This type of incompleteness may lead to misinterpretations in the temporal dimension, such as when a feature appears or disappears, or in measurement errors, such as computations of erosion or deposition. A notable example in physical geography is the apparent appearance, ‘disappearance’, and reappearance of alpine lakes on subsequent topographic maps (Butler and Schipke, 1992). Similarly, lack of change may be erroneously inferred from a map that adopted information from a previous map without field verification. Thus, a later date on a map does not ensure geomorphic stability unless an independent field mapping survey was conducted. U.S. Geological Survey topographic maps have used a magenta color to indicate that features on the present map edition were mapped from remotely sensed imagery and not field surveys. Errors of cartographic omission are common where the area was incompletely mapped by ground surveys or penetration to the

<table>
<thead>
<tr>
<th>Scale</th>
<th>On map</th>
<th>On ground</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(inch)</td>
<td>(mm)</td>
</tr>
<tr>
<td>&gt;20,000</td>
<td>1/50</td>
<td>0.0200</td>
</tr>
<tr>
<td>&lt;20,000</td>
<td>1/30</td>
<td>0.0333</td>
</tr>
<tr>
<td>Examples:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250,000</td>
<td>1/50</td>
<td>0.0200</td>
</tr>
<tr>
<td>100,000</td>
<td>1/50</td>
<td>0.0200</td>
</tr>
<tr>
<td>62,500</td>
<td>1/50</td>
<td>0.0200</td>
</tr>
<tr>
<td>24,000</td>
<td>1/50</td>
<td>0.0200</td>
</tr>
<tr>
<td>12,000</td>
<td>1/30</td>
<td>0.0333</td>
</tr>
<tr>
<td>5,000</td>
<td>1/30</td>
<td>0.0333</td>
</tr>
</tbody>
</table>

a No more than 10% of well-defined points can be beyond these error limits. (These values may be interpreted as the 90% confidence limit with the understanding the NMAS do not assume a specific probability distribution.)
ground by remote sensing methods was incomplete. For example, studies of changes in river planform may erroneously record a shift from single channel to multithread channel based on an early map that shows only a single channel when additional channels simply were not included on the early map. Such errors of cartographic omission were common when maps relied entirely on ground surveys (Fig. 5). Time and resources often prevented comprehensive mapping of secondary geomorphic features such as headwater streams and gullies (James et al., 2007).

7. Uncertainties in geomorphometry

DEM have been used for many geomorphic and hydrologic metrics, such as mapping drainage networks (Mark, 1984), steepest slope lines (Chou, 1992), and shallow landslides (Duan and Grant, 2000). Townsend and Walsh (1998) combined DEMs with a variety of remote sensing and GIS data including synthetic aperture radar images to map areas of flood inundation and develop potential flood inundation models for the Roanoke floodplain. Simple rectangular grids (standard DEMs) are not ideal for studies of surfaces with abrupt steep surfaces separated by large relatively flat areas. These surface conditions require high spatial resolutions in the areas of slope changes that result in massive data redundancy in the flat areas. For example, steep stream banks surrounded by relatively flat floodplains call for either highly dense data sets or the use of other data models, such as a triangulated irregular network (TIN) supported with breaklines and drainlines (Lane, 2000).

Vertical geodetic positioning differs from vertical positioning relative to a local relative datum with an arbitrary elevation. For some cartometric purposes and for qualitative assessments of change, a relative datum may be adequate. For quantitative change detection and the construction of DoDs, however, geodetic control between two data sets may specify parameters for the vertical co-registration that can prevent systematic bias in elevation changes. Early maps often lack accurate geodetic control (or lack documentation), so empirical methods, such as using vertical GCPs in stable locations, may be needed for vertical registration.

Errors in topographic data are important to GCD because they are incorporated in the change-detection analysis and could be misinterpreted as geomorphic change. All DEMs have errors from sampling, measurement, and interpolation, and these will be propagated to products derived from them such as channel networks (Walker and Willgoose, 1999; Fisher and Tate, 2006). Unfortunately, error propagation is often poorly understood because error sources and the spatial variation of errors in the source materials, processing, and resulting maps are seldom documented. Some standard DEM products may not be well suited for DoD analysis, especially where geomorphic change is subtle (cf. Eq. (2)). Several design specifications of DEMs should be considered before they are employed in quantitative GCD, including the scale and contour interval of the source map or imagery, sampling interval, precision of the base map relative to terrain complexity, and interpolation method used (Walsh, 1989).

Errors in spatial data are conventionally reported as Circular Map Accuracy Standard (CMAS), RMSE, or more recently as “Accuracy.” The latter two statistics are often reported separately for the horizontal and vertical dimensions of the map. RMSE for the entire map, which may be the only error reported with the data, may mask systematic bias that can be measured by the mean error or error standard deviation (Fisher and Tate, 2006). Knowledge about where errors occur is also of great importance, yet the RMSE is devoid of spatial information (Wood and Fisher, 1993). Elevation errors refer to specific measurement problems related to differences with reference elevations, while uncertainty includes additional doubts in measurement accuracies that are difficult to assess, such as differences introduced by interpolation or rescaling (Fisher and Tate, 2006). Vertical errors are affected by accuracies of the reference data and by the number and spatial distribution of control points used for geometric rectifications or interpolations (Li, 1991). Estimates of vertical accuracies in the reference data may be constrained by vertical map accuracy standards for maps conforming to standards. In the U.S.A., NMAS require at least 90% of points to have elevations that differ by no more than 1/2 the contour interval (USGS, 1999).

Vertical errors in DEMs can be attributable to poor vertical registration, improper geodetic control, errors inherent to topographic data used, or errors introduced by horizontal error. These errors are compounded in the generation of vertical change data (e.g., DoDs). Vertical errors introduced by horizontal error are particularly important in geomorphic studies because they are common, can be large in magnitude, and are associated with steep slopes, so they often occur at critical locations such as terrace or fault scarps, stream banks, or dune

Fig. 5. Cartographic errors of omission should not be misinterpreted as geomorphic change. (A) Von Schmidt (1859) map shows a single-thread lower Yuba River channel with no southern channel. (B) Map showing anastomosed system with small southern channel two years later (Wescoatt, 1861). (C) Later map based on detailed field survey shows anastomosed channel system as former channel positions (Mendell, 1881). These images were manually edited to clarify linework and text from greatly enlarged originals and to remove artifacts introduced by the interpolation process that degrade the graphical quality of lines and lettering.
faces. The inter-dependency between vertical and horizontal error is well known and has interesting implications regarding the effects of DEM grid-cell size that influence variability in DEMs. Vertical errors caused by horizontal offsets in a map can be estimated as a function of local slopes and the horizontal errors (Maling, 1989, p. 154; Hodgson et al., 2005):

\[ \varepsilon_{Zij} = \varepsilon_{H} \tan \alpha \]  

(4)

where \( \varepsilon_{Zij} \) is the elevation error caused by horizontal offsets on a sloping surface, \( \varepsilon_{H} \) is the horizontal error, and \( \alpha \) is the local slope. Because most geomorphic slopes are less than 45°, most vertical errors caused by horizontal displacement, \( \alpha \), will be less than the horizontal displacement, \( \varepsilon_{H} \) (Maling, 1989). For slopes >45°, however, errors recorded by this metric increase rapidly with slope, and go to infinity in the limit as slope approaches 90°. Thus, it may be prudent to constrain the maximum error calculated by this method to no more than the relief of local features. A map of estimated vertical errors, attributable to a constant worst-case horizontal offset, can be computed for DEMs by computing slopes between cells and utilizing the horizontal RMSEh to estimate the horizontal offset error of individual grid cells:

\[ \text{RMSE}_{Zij} = \text{RMSE}_{H} \tan S_{ij} \]  

(5)

where \( \text{RMSE}_{Zij} \) is the vertical error in a cell of the error grid caused by the horizontal offset and slope, \( \text{RMSE}_{H} \) is the horizontal error for the map, and \( S_{ij} \) is the slope of a grid cell in the slope grid. The vertical error resulting from a horizontal error is directionally dependent — a horizontal error parallel to a slope will have no resulting vertical error. The fundamental Eq. (5) can be extended by Monte Carlo simulation modeling to incorporate random horizontal error directions with a known (mean and standard deviation) slope distribution (Hodgson and Bresnahan, 2004):

\[ \text{RMSE}_{Z} = (a_1 \tan (\text{mean}\_\text{slope}) + \text{RMSE}_{H}) + (a_2 \cdot \text{standard}\_\text{deviation}\_\text{slope} \cdot \text{RMSE}_{H}) \]  

(6)

where \( a_1 \) and \( a_2 \) are constants for the population of slopes in a study area.

### 7.1. Interpolating DEMs from topographic point, contour, and TIN data

DEM may be generated from topographic information based on field surveys, topographic maps, stereo pairs of aerial photographs, or other remotely sensed data (Hutchinson and Gallant, 2000; Lane, 2000; Jensen, 2007; Nelson et al., 2009). An important source of uncertainty in gridded DEMs is the method of interpolation used to generate a lattice of elevation points from other topographic data structures. Topographic data may be stored and displayed as contour lines, profiles, DEMs, TINs, or clusters of points. The structure, completeness, and planimetric and topographic accuracy of the data should be considered in the selection of an appropriate interpolation method and grid-cell size. Interpolation of contours can be done using automated procedures available on most commercial GIS software packages by a variety of interpolation methods (Carrara et al., 1997). Many studies have documented contour-to-grid interpolation algorithms including the addition of ancillary data (e.g. drainlines or hypsography) and the resulting quality of DEMs. An improper method can introduce artifacts in the DEM. Guth (1999) evaluated USGS DEMs that were based on a contour-to-grid algorithm and found systematic high (and corresponding low) frequencies of elevations (“contour line ghosts”) with values peaking at the contour line elevations. For example, micro-topographic “stepping” can be introduced in low-relief areas by interpolating between contour lines along north–south and east–west trend lines rather than following flow lines perpendicular to contours (Pelletier, 2008). Such early directional problems with USGS DEM production led to the adoption of the U.S. Forest Service’s Linetrace algorithm that utilized 8-directional interpo-
8. Uncertainties associated with DEM differencing

Uncertainties that arise from data acquisition, recording, and post-processing may be larger than actual geomorphic changes, and construction of DoDs compounds these uncertainties. The DoD method may not be appropriate where real geomorphic changes are small and uncertainties are large. This question of appropriate application is especially relevant to historical reconstructions using data such as historical maps that may have large errors. Cartographic-based applications of the DoD method are best reserved for cases where relatively accurate cartographic data are available and geomorphic change has been considerable. An evaluation of the uncertainty of a DoD should consider three steps (Wheaton et al., 2009):

1) quantification of uncertainty in each individual DEM,
2) propagation of these uncertainties into the DoD, and
3) an assessment of the importance of the propagated uncertainty.

The first step was covered under cartometric and geomorphometric uncertainties and can also be shown by an evaluation of the accuracy and precision of the data and procedures used to generate DEMs. Some researchers recommend that values of change be filtered by setting to zero values that are small relative to errors. Wheaton et al. (2009) suggested the use of a critical threshold, t, for screening the change model based on the ratio of observed elevation change to error:

\[ t_{ij} = \frac{|Z_{ij} - Z_{2ij}|}{\varepsilon_{2ij}} \]

where \( |Z_{ij} - Z_{2ij}| \) is the absolute value of the DoD and \( \varepsilon_{2ij} \) is the composite error associated with the DoD propagated by the two DEMs:

\[ \varepsilon_{2ij} = \sqrt{\varepsilon_{1ij}^2 + \varepsilon_{2ij}^2} \]

where \( \varepsilon_{1ij} \) and \( \varepsilon_{2ij} \) are the vertical errors in cells of the early and late DEMs, respectively. They recommend using a critical threshold (e.g., \( \alpha = 95\% \)) for values to filter out changes in the DoD that are not significantly different than the errors. This ratio has intuitive appeal because it is similar to a signal-to-noise ratio that measures the magnitude observed changes in the DoD (real change plus uncertainties) relative to uncertainties. If the cumulative error \( \varepsilon_{DoD} \) is normally distributed and based on a probability distribution, then Eq. (7) is similar to the equation for computing a z-score, and the t-value would reflect the confidence that the change was an actual change. Where topographic change is large and uncertainty is small \( \varepsilon_{DoD} \), values of \( |Z_{ij} - Z_{2ij}| \) and t values will be high, the likelihood of variability being generated primarily by error will be small, and the cell will not be filtered from the DoD model. The difficulty in applying this method is computing the actual errors for the individual DEM grid cells.

The process of filtering out all cells with small change should be undertaken with caution, because filtering may eliminate extensive real changes that are cumulatively important. In some cases, small vertical changes of the same sign may be widespread and comprise a large proportion of the sediment budget. For example, floodplain sedimentation by overbank processes may be thin but constitute a large volume of sediment. If errors can be assumed to be normally distributed, the distribution of small changes can be tested for symmetry around the mean to ensure against bias in the filtration process.

9. Examples of change detection in fluvial processes

Three case studies are presented to demonstrate the use of historical maps in conjunction with LiDAR or photogrammetric data to produce DoDs. The first example is a gully system in the upper Piedmont of South Carolina. The next two examples are large rivers and floodplains in California where historical sedimentation and channel change were extreme. Detailed historical topographic maps and modern high-resolution remote sensing data are available for all three sites. High-resolution time-sequential DEM pairs were generated for each area by digitizing early topographic data from topographic maps and differencing them with modern high-resolution topographic data (Eq. (1)) to produce high-resolution DoDs. Each site presented different limitations and challenges to the DoD method, so these studies provide a diverse set of examples of the method applied to fluvial landforms.

10. Cox gully

Gully erosion poses a serious threat to environmental systems by damaging productive lands, increasing flood magnitudes, and delivering non-point-source pollution. They generate dense, more efficient drainage networks, raise flood stages as a result of lowland channel filling, and degrade water quality by generating non-point-source (NPS) pollution. Gullies in the Piedmont region of the southeastern U.S.A. were highly active in the early 20th century because of forest clearance, lack of conservation measures, and intense rainfall. The Upper Piedmont region is characterized by moderate local relief, into which gully incision generates a striking vertical relief along sidewalks and headwalls. Gullies in this area were studied carefully during the early stages of soil conservation research in the United States (Ireland et al., 1939). In the 1930s, agriculture in the southern Piedmont region declined rapidly, soil conservation measures were introduced, reforestation ensued, and many gullies stabilized. Little study has been done of these systems since World War II because of the need for ground-based surveys, and it is often assumed that gullies are no longer active. The first example of DoD analysis is focused on the Cox Gully system, which was surveyed in 1938 (Ireland et al., 1939).

10.1. Map and LiDAR processing

A preliminary DoD for the Cox gully from 1938 to 2004 was developed using the historical map (Ireland et al., 1939) and airborne LiDAR data collected for Spartanburg County in August, 2004 for the South Carolina Flood Map Modernization Project. Topographic data with sufficient resolution to map gullies in this region were rare prior to the advent of airborne LiDAR because photogrammetric analysis is restricted by thick vegetation. Fortuitously, detailed topographic surveys, mapping, cross-sections, longitudinal profiles, and stratigraphic analysis of several active gully systems in Spartanburg County, South Carolina were conducted in the late 1930s (Ireland et al., 1939). Modern LiDAR mapping techniques have proven to be effective in mapping gully systems under forest canopy in the region, although LiDARbare-Earth data from standard processing can produce inaccurate cross-section morphologies of gullies (James et al., 2007). Several gullies studied by Ireland et al. (1939) were revisited in the field from 2000 to 2004. The Cox gully was singled out for a detailed study (Kolomechuk, 2001) that was followed by field surveys to measure topographic changes since the Ireland et al. (1939) surveys.

The 1939 map was scanned and preprocessed to remove speckling and shading (Fig. 6). Use of conventional ground control points (GCPs) for map registration was severely limited by the lack of distinctive points on the historic map. This is a common problem with historical map analysis that may limit the accuracy of reconstructions. Instead, registration was based on a combination of three types of registration points on the map and the digital orthophoto quadrangle (DOQQ): (1) two conventional GCPs distinctly identifiable on both images, (2) 15 synthetic GCPs were generated on an orthogonal grid of points, and (3) 35 approximately located GCPs at critical locations on the map to ensure that the 1939 gully rims did not extend beyond the 2004 rims. The synthetic grid was generated by using the graphical scale bar on the 1939 map to space points along the centerline of a road and on perpendiculars from the road. The same grid was constructed on the reference image using identical spacings and bearings. The advantage of this method is that it allows the map to be approximately registered to the DOQQ in spite of the lack of uniquely identifiable GCPs. Two key disadvantages of this method are that the synthetic GCPs are not based
on true positions so they may force map deformation and the accuracy of the registration cannot be quantified. The RMSE reported by the registration (0.27 m) represents only the success of fitting the map to the GCPs. Because most of the GCPs were simulated or approximate, this metric grossly underestimates errors and should not be used as an indicator of the planimetric accuracy of the resulting map. The general orientation of the gully system on the resulting map was validated using GPS points collected along the rim of the modern gully.

LiDAR data for Spartanburg County were collected by airborne scanning in three flights April 5, 6, 7, 2004 by Woolpert, LLP, for use compatible with developing 0.6-m (2-ft) contours. Counting the bare Earth points in a 14 ha rectangle around the gully yielded a mean point density of 0.10 pts/m² which converts dimensionally to a mean point spacing of 3.2 m/pt. Unfortunately, the bare Earth data in the study area – as delivered – have a non-uniform spatial distribution that is strongly biased towards clearings. The gullies are heavily wooded, so point densities within the gullies are sparse. The point cloud is available and future analysis will ultimately re-process a new bare Earth data set. For this preliminary study, the existing points were interpolated to a 70×70-cm DEM using inverse distance weighting (IDW) based on 12 points in a variable search radius. This interpolation worked well for the inner gullies, although the resulting DEM has a stair-stepping artifact that propagated into the DoD analysis. The 2004 DEM from the IDW

Fig. 6. Cox Gully, Spartanburg, South Carolina. Rectified excerpt from topographic map surveyed in 1938 by Ireland et al. (1939).

Fig. 7. Cox Gully cross section changes based on field surveys in 1938 and 2001 (see Fig. 8 for locations). Section D12 is representative of main-stem gully sections that filled and widened. Gully E did not exist in 1938 and represents post-1939 gully erosion >4 m deep and 9 m wide.

Fig. 8. Cox Gully DoD map with gully rims derived from contours for both periods. (A) Change model shows erosion by widening of sidewalls and deposition in gully bottoms. Two small gullies on south side were filled after 1938 as was the upper half of the large eastern gully, but the lower half of the eastern gully was not detected by LiDAR owing to dense canopy and backfilling. (B) Close-up of north branch with 1938 contours over DoD change model showing headward extension and more than 2 m of deposition on gully floors including deep fill in former plunge pool (center of image). Northeast gully branch with cross-section E did not exist in 1938 and was surveyed with a total station in 2001 (Fig. 7).
interpolation was used to generate a contour map that was more realistic than a contour map generated directly from a TIN of LiDAR points; i.e., overall contour patterns were similar but the IDW-derived contours were smoother and lacked the angularity of TIN-derived contours. The resulting contour map and DEM of gullies conforms approximately to contemporary topographic cross sections and other field observations (Fig. 7). A DEM of difference (DoD) was constructed by differencing the 1939 and 2004 DEMs (Fig. 8A).

In spite of large uncertainties arising from poor 1939 map registration and low point densities of LiDAR data within gullies, several important geomorphic changes and processes can be identified from this preliminary change model in addition to guidance towards future methodological improvements. LiDAR data detected but underestimated the depth of a new northeastern branch of the Cox Gully system. This underestimation of the depths of V-shaped gullies and rounding of gully rims typical of "off-the-shelf" LiDAR bare-Earth point clouds derived from landscapes under forest canopy in this region because of the low density of bare-Earth points within gullies (James et al., 2007). Hopefully, the mapping of forested gully morphologies can be improved by careful reprocessing of the point-cloud data testing multiple methods with special focus in the vicinity of gullies. Detection of the new gully branch indicates that substantial erosion occurred after the 1938 survey and that these features can be detected using LiDAR data. In addition, the GCD using historical maps reveals several local-scale geomorphic processes. For example, main gullies filled and widened in conformance with common gully evolutionary processes (Ireland et al., 1939; James et al., 2007). Moreover, a pocket of fill >3 m deep in the head of the main 1939 shoreline indicates the presence of the former plunge pool (Fig. 8B). Comparisons of 1939 and 2004 gully cross sections show substantial filling in the lower gully in that period (Fig. 7). Filling had begun in the lower gully prior to 1939 as was shown by core stratigraphy and photographs at that time (Ireland et al., 1939). The Cox Gully DoD, therefore, provides guidance for a sediment coring program of the gully floor.

11. Feather River at Shanghai Bend

This example examines geomorphic changes in a 1.8-km reach of the lower Feather River from 1909 to 1999. Prior to 1909, the river had been affected by sedimentation following hydraulic mining and a considerable amount of channel and floodplain engineering, including levees and channel dredging (James et al., 2009). At the time of the 1909 survey, channels in the area were avulsing from positions where they had been for at least 50 years to new channelized positions from where they subsequently migrated to the present configuration forming Shanghai Bend. Thus, this is a case of exceptional geomorphic change rather than a randomly selected site.

A modern (1999) DEM for the floodplain around Shanghai Bend was extracted from an extensive data set generated by airborne LiDAR as part of a study designed to compare LiDAR and photogrammetry (Stonestreet and Lee, 2000; Towill, 2006). LiDAR data were acquired by EarthData Aviation using a fixed-wing aircraft flying between 2440 and 2740 m above mean terrain, on March 27 and 28, 1999, using an AeroScan system with a scan rate of 50 Hz. The aerial survey was performed with a Phalanx inertial measurement unit (IMU) and a dual frequency GPS receiver, in conjunction with a mobile kinematic GPS survey. The LiDAR mission was designed to acquire data with a horizontal accuracy of 1 m RMSE and a vertical accuracy of 15 to 20 cm RMSE. The resulting points have mean postings varying from 1.5 to 6.1 m and, based on comparing points in overlapping areas between flight lines, a vertical accuracy of 19.3 cm RMSE (Towill, 2006). A third party (Douglas Allen written comm.) merged the data with channel bathymetry data collected in 1999 by a hydrographic boat survey using a real-time and post-processed kinematic GPS survey, a fathometer, and sonar transducer to meet the requirements of a Class 2 Hydrographic Survey (Ayres, 2003). The merged 1999 data set was interpolated to a 3 × 3 m DEM using kriging with a linear model.

A set of large-scale map sheets, produced from floodplain topographic surveys in 1909 (CDC, 1912), was used to generate a 1909 DEM for DoD analysis (Megison, 2008). The 1909 maps did not include channel bathymetry, but they did include frequent cross sections showing channel depths at a set of closely-spaced points across the channel. These point depths were associated with low-water elevations printed at several locations on the map. They were converted to point channel-bottom elevations and used to manually interpolate bathymetric contour lines that were combined with the terrestrial contours (Fig. 9). Contours were given a slight longitudinal orientation to simulate longitudinal or lingoid bars common to sand bedded streams (as opposed to random as automated procedures may generate or transverse as may occur in some other environments).

Points extracted from the 1909 contours (including bathymetry), field survey points, and low-water edge contours were used in a spatial interpolation to a 3 × 3 m grid co-registered with the 1999 LiDAR DEM. Four methods of interpolation were tested on the Shanghai Bend dataset: kriging, IDW, TIN to raster, and Topogrid, a tool in the ArcGIS Spatial Analyst toolbox (ESRI, Corp.). For kriging and IDW, point locations and elevation values were extracted from the contour line data at a 3-m spacing and produced cross-validated vertical RMSEs of 0.355 m and 0.420 m, respectively. In spite of the low RMSE values, both methods created a stepping effect on the floodplain where a gradual change in elevation is more realistic. TopoGrid uses a spline function to generate a hydrologically correct DEM from contour-line data by accentuating topographic variation in areas lacking elevation data. The 1909 Shanghai Bend floodplain map has broad areas of sparse data.
TopoGrid modeled elevations in these areas as unrealistic peaks or troughs, so this method was dropped from further analysis. The TIN method allows input of points and contour lines. Contour lines were input as soft breaklines, water-surface boundaries as hard breaklines, and survey points as mass points. The TIN-derived DEM was selected for further study because of the lack of stepping and realistic gradual slopes on such features as point bars.

The resulting 1909 DEM shows the channel in the initial process of an avulsion from an early eastern position into a dredged channel along the west levee (Fig. 10). Interestingly, a large natural levee had already formed along the east bank of the new channel, which presumably reflects the high loads of mining sediment carried by the river at this time (James et al., 2009). Eliza Bend, the old, high-amplitude meander to the northeast, was cut off by this avulsion, locally steepening the channel, and providing excess hydraulic energy that ultimately led to lateral channel migration and the formation Shanghai Bend at this site.

The 1909 and 1999 DEMs were vertically registered by an empirical test that compared elevations of stable 1906 and 1999 surfaces outside of levees along cross sections generated from both data sets and validated by comparisons with contemporaneous (1909) cross-section surveys and profiles extracted from the 1999 LiDAR data. The vertical datum on the 1909 map was assumed to be the United States Engineering Datum whereas the 1999 LiDAR-based DEM was referenced to the NGVD29. Tests confirmed a $-0.76 \text{ m}$ vertical offset between the DEMs and this adjustment was made to bring the 1909 elevations to NGVD29. The two DEMs were then differenced to produce a DoD that reveals net changes integrated over a period of 90 years (Fig. 11). These GCD models illustrate spatial patterns of sediment reworking in a very active fluvial system. Light shaded areas on the models represent net sediment deposition. Zones of maximum sedimentation occurred as channel fill near the top of the DoD where the thalweg of the former channel filled up to 7.9 m. Downstream, the former thalweg filled with up to 6 or 7 m of sediment in narrow strips. A 300-m stretch of the former canal along the west levee is now filled with 6–6.5 m of sediment. These sediment deposits are important because of the high concentrations of total mercury associated with the hydraulic mining sediment in the system (James et al., 2009). The natural levee, showing in 1909 (Fig. 10), is not present in 1999, which suggests that Shanghai Bend formed by lateral migration to the east rather than by avulsion. This interpretation is supported by the presence of a large gravel point bar on the inner bend. The large rectangular eroded areas in the center of the DoD are lagoons for a sewage treatment plant. The dark shaded areas of Fig. 11 indicate erosion depths up to 14.4 m primarily within the modern channel. The deepest erosion occurred in the southern part of Shanghai Bend along the outer cut bank at the southern end of the map. Other areas of deep erosion include the point bar – especially in a chute – and along the lower banks of the old channel where, the natural levee is now gone, possibly because of the migration of the secondary channel. Draping a 1952 aerial photograph over the DoD brings out the geomorphic context of the changes (Fig. 11C). This differs from draping over a standard DEM in that zones of deposition are shown in positive relief even where the present topography may be low. For example, the deeply alluviated paleochannels are shown as ridges. The light tonality indicates relatively recent sedimentation at the time of the photograph (e.g., on-going channel filling and point-bar construction in 1952).

11.1. Vertical uncertainties in the Shanghai Bend DoD

An analysis of vertical uncertainty in the 1909 and 1999 Shanghai Bend DEMs was conducted based on an error budget approach (Megison, 2008). Some uncertainty values were assumed to be uniform across the surface whereas others were associated with spatially distributed uncertainty grids (Table 4). Horizontal accuracies were computed and applied to the vertical assessment, as demonstrated here with the 1909 data. Horizontal accuracies were computed as a function of two factors. Cartographic uncertainty ($\text{RMSE}_{\text{Cart}}$) accounts for inaccuracies, such as line work, and was set to a uniform value based on the fair drawing error of less than 0.5 mm (Maling, 1989) or approximately 4.8 m on the ground at the 1909 map scale of 1:9600. Horizontal map registration uncertainties ($\text{RMSE}_{\text{Reg}}$) are potentially high on large paper sheets such as the 1909 maps. To estimate $\text{RMSE}_{\text{Reg}}$,
66 measurements along a railroad on the 1909 and 1999 reference data were used to compute one-dimensional RMSE for the 1909 map:

$$\text{RMSE}_{1D} = \sqrt{\frac{\sum (d_i - d_{i+1})^2}{n}}$$  \hspace{1cm} (9)$$

where $d_i - d_{i+1}$ is the horizontal difference between the 1909 and 1999 measurements. Assuming the error along this one dimension was the same in the orthogonal direction, this RMSE$_{1D}$ was converted to a two-dimensional RMSE for horizontal map registration uncertainty (RMSE$_{\text{Reg}}$):

$$\text{RMSE}_{\text{Reg}} = \sqrt{2 \times \text{RMSE}_{1D}^2}$$  \hspace{1cm} (10)$$

where RMSE$_{\text{Reg}}$, the potential error in X and Y. This uncertainty, a uniform value of 6.26 m on the ground (Table 4), is based on measurements from the 1909 to 1999 digital data, so it presumably includes all digitization, map compilation, reproduction, and registration uncertainties. Assuming the error sources are independent, cumulative horizontal uncertainties were estimated to be 7.89 m from cartographic and registration inaccuracies:

$$\text{RMSE}_H = \sqrt{\text{RMSE}_{\text{Cart}}^2 + \text{RMSE}_{\text{Reg}}^2}$$  \hspace{1cm} (11)$$

where RMSE$_H$ is the horizontal uncertainty that is assumed uniform across the map.

Vertical uncertainties were computed from three factors: inaccuracies introduced by horizontal offset, contour lines, and interpolations. Horizontal offsets were assumed constant and equal to RMSE$_H$ that was combined with a 1909 slope grid, $S_p$, to compute a map of vertical uncertainties caused by horizontal offset (RMSE$_{ZH_{ij}}$) (cf. Eq. (5)). Uncertainties introduced by vertical errors (i.e. as exhibited in the contour lines) were estimated based on U.S. National Cartographic Accuracy Standards; i.e., 90% of point elevations differs no more than one half of the contour interval. Although these maps were generated prior to establishment of the California Debris Commission, an affiliate of the War Department (now U.S. Army Corps of Engineers), presumably using equally high accurate and precise cartographic methods. The interval for the CDC (1912) topographic maps is 0.610 m (2 ft), so one half of the interval is 0.305 m. We assumed that the 1912 map met the one-half contour interval threshold and that less than 10% of the points on contour lines have vertical errors exceeding 0.305 m. To convert the one-dimensional 90% confidence level to a 68% confidence level, 0.305 m was divided by 1.645, resulting in a uniformly distributed contour uncertainty of 0.185 m across the 1909 DEM. Interpolation uncertainties were estimated empirically. The interpolation error was measured at a companion study site downstream and applied to the Shanghai Bend site. Vertices of the TIN along contours, combined with other known point features, were treated as ‘known’ z-values and compared with
corresponding 1909 DEM values using cross-validation. The resulting interpolation error was 0.765 m (RMSE).

Following the error budget model of Hodgson and Bresnahan (2004) and assuming independent error sources, cumulative vertical RMSE values for the 1909 data were computed and mapped from three sources of error:

\[
\text{RMSE}_{Zij} = \sqrt{\text{RMSE}_{Zcont}^2 + \text{RMSE}_{Zint}^2 + \text{RMSE}_{Zerr}^2}
\]  

(12)

where \(\text{RMSE}_{Zij}\) is the vertical uncertainty of a cell in the uncertainty grid, \(\text{RMSE}_{Zcont}\) is the contour interval uncertainty with a uniform value (0.185 m), \(\text{RMSE}_{Zint}\) is the interpolation uncertainty with a uniform value (0.765 m), and \(\text{RMSE}_{Zerr}\) is the grid of vertical errors caused by horizontal offsets and that vary spatially with slope. Values in parentheses are for the 1909 data and show that the uniform sources of error combine to a total of less than one meter (0.95 m). The resulting 1909 vertical uncertainty grid has values ranging from 0.8 to 9.3 m (Fig. 12). Most values are less than one meter and indicate that uncertainties on relatively flat floodplain surfaces tend to be low for a map of the quality assumed here. The large values in the uncertainty grid are dominated by potential errors associated with horizontal offsets along steep channel banks and terrace scars. Errors at these locations could inflate estimates of total erosion and deposition, especially in narrow channels with high banks. In this case, large-scale lateral migration, avulsion, and abandonment of large channels were the dominant processes of erosion and deposition, so uncertainties associated with local boundaries is of minor consequence relative to actual geomorphic changes.

The dominant vertical uncertainty in this analysis was associated with horizontal displacements near steep slopes. Wheaton et al. (2009) suggest that horizontal components of uncertainty may be negligible in gravel-bed rivers and that vertical errors are random and independent (Eq. (7)). The results of this study of a floodplain with fine-grained cohesive banks indicate that horizontal uncertainty can be an important element of vertical error and should not be neglected, in studies where sites of potential erosion are associated with steep slopes. Moreover, vertical error may have strong spatial autocorrelations and occur in association with features that persist between sample periods so that assumptions of error independence within or between DEMs should be questioned. In addition to vertical error correlations with landform features, correlations for certain land covers in a DEM are typically high (e.g. greater errors in forested versus non-forested land cover). This has been empirically documented many times for remote sensing derived data (e.g., Hodgson et al., 2003; 2005) and is now so well accepted that error estimates by class of land cover are often required for contract work. Fortunately, the horizontal component of vertical uncertainty can be estimated and mapped fairly easily. Unfortunately, the maximum of this uncertainty occurs along steep banks and terrace scars that are often the focus of GCD studies of rivers.

12. Lower Yuba River channels

The Yuba River is the largest tributary of the Feather River, which it joins about 3.6 river km above Shanghai Bend. The Yuba River has a drainage area (3470 km²) about one third of the drainage area of the Feather River at the confluence, but it received more hydraulic mining sediment than any of the Sierra Nevada Rivers. Wide levee setbacks were designed in the late 19th century to encourage capture and storage of the mining sediment and prevent it from progressing downstream to the navigable Feather and Sacramento Rivers. This policy resulted in deep aggradation of an active floodplain –4 km wide, channel avulsions, and morphogenesis from single- or dual-threaded channels to braided and back to single-thread low-water channels by the time of the 1906 floodplain survey (James et al., 2009; Ghoshal et al., 2010). When mapped in 1906, the channel in the study reach was approximately in its present location. Subsequently, channel incision left a broad historical terrace several meters above the active channel. Thus, this study site has experienced a great deal of change over the 93 years of the study period, but in a manner that is entirely different than the Feather River at Shanghai Bend.

A series of large-scale map sheets of the lower Yuba River (CDC, 1906), based on a topographic survey in 1906, are used as the historical basis of a third case study. The maps are similar to those used in the Shanghai Bend example but are from an earlier field survey that included low-flow channel bathymetric contours that were digitized directly. Geometric rectification of the CDC maps produced average rectification RMSEs of 5.5 m. These errors are substantially greater than rectification errors produced by preliminary geometric rectification of aerial photography (RMSEs 0.2 to 0.4 m) because of deformation of paper maps and the need to mosaic map sheets from multiple scans (Ghoshal et al., 2010). Whereas the resulting CDC maps are not as accurate as rectified aerial photographs, the information content is rich and the geomorphic change was substantial, especially in the vertical dimension. Thus, DoD models provide an important foundation for interpreting floodplain and channel geomorphic changes over a 93-year period. The contour data were interpolated to a 3×3-m DEM (Fig. 13) using a TIN interpolation with Erdas Imagine 9.2 software (©ERDAS, Inc.). The contour lines were used as soft break lines and river banks were used as hard break lines. This method introduced stair stepping but was considered to be superior to the IDW method used on an alternate GIS software package that generated more pronounced steps.

Previous research on this channel has shown that the channel returned to a position similar to where it was prior to avulsions during the mining era (James et al., 2009). The 1906 DEM reveals that the main channel had built a natural levee along much of its length. Braided channel scars to the northwest and anastomosed channel scars to the southeast belie the transformation of the floodplain back to a single-thread channel shortly before the 1906 survey.

Modern topographic data were obtained from the same sources as the data used for Shanghai Bend (Stonestreet and Lee, 2000; Towill, 2006; Ayers, 2003). The Yuba terrestrial data, however, were generated by photogrammetric rather than LiDAR methods as part of a comparison.
study to test the feasibility of LiDAR methods for floodplain mapping. The 1999 terrestrial and bathymetric data were merged and processed to the same specifications as the 1999 data used at Shanghai Bend. The two 3×3-m DEMs generated from the 1906 and 1999 datasets were differenced to show volumetric changes from 1906 to 1999 (Fig. 14).

The DoD documents substantial net main channel incision after 1906, typically from 5 to 9 m but up to 12.9 m along the thalweg (Fig. 14). Surprisingly, the high-water channels on the southeastern floodplain also experienced net erosion on the order of 2 m during the period. (This does not include occasional pits in the beds of these side channels that show ~4 m lowering where sand has been quarried.) Although sedimentation of these abandoned channels was expected, net incision may reflect decreased sediment loadings after closure of Englebright Dam upstream along with a decrease in main channel sediment deliveries to the terrace as channel incision decoupled over-bank flow processes. It may also represent an artifact of averaging changes over a long period; i.e., early incision has not yet been negated by subsequent in-filling.

A volumetric analysis was performed in this area by Ghoshal et al. (2010) by computing erosion and deposition as a simple product of DoD grid-cell depths and areas. Volumetric change was more than 12×10⁶ m³ of total erosion, mostly from within channels (including low-water channels and bars within the bankfull channel). Total deposition was more than 5×10⁶ m³, mostly on terraces, but also on bars within the channel above low water. Much of the terrace deposition was concentrated along natural levees near the main channel and, to a lesser extent, along secondary high-water channels where it is largely removed from active reworking. The net sediment change for the area, therefore, was about 7×10⁶ m³ of erosion over the 93-year period, reflecting the systematic removal of mining sediment stored in the area.

Arguments have been made in a ground-breaking DoD study for filtering out apparent changes in DoDs that do not exceed a threshold signal-to-noise ratio (Wheaton et al., 2009). In some cases, however, filtering out all small changes could obscure an important component of the sediment budget; for example, where thin sedimentation covers broad floodplains. The spatial pattern of sediment deposition along the lower Yuba River suggests that removing data solely on the basis of the magnitude of vertical change is not always appropriate.

Extensive sedimentation across the floodplain has been an ongoing process in the lower Yuba and Feather Rivers throughout the mining and post-mining periods as evidenced by buried soils in stream banks and other exposures. Even relatively thin deposits on the extensive terrace and bar surfaces sum to large volumes of net change, and the sediment budget would be incomplete without this component. By filtering out all small changes below a threshold, all thin deposits are systematically removed, which may result in a bias towards lower sedimentation volumes.

13. Conclusions

The principles of GCD may be applied to geospatial analyses over a wide range of time periods. For studies extending back more than 100 years, qualitative methods may provide important information about geomorphic change, or lack thereof, but care should be taken to recognize potential errors of omission and commission as well as accuracy limitations of the data. With high quality maps, cartometric and geomorphometric methods may be applied to extract quantitative planimetric and topographic data. More recent studies may utilize archives of satellite and airborne remotely sensed imagery or tap a growing archive of high resolution data from airborne laser and radar surveys, or terrestrial LiDAR. Combining multiple sources of historical spatial data to increase temporal resolutions will open the way for developing well-documented time-domain spatial analyses from which inferences of geomorphic rates and processes can be made.

Methods of GCD are likely to be of growing utility, because studies of global change require a deeper historical understanding of geomorphic transformations. DoD will be especially important in geomorphology where topographic change governs or reflects many critical processes. The use of DoD methods in geomorphometry is likely to increase as data availability improves and historical reconstruction methods become more simple and accurate. As DoD methods are developed and standardized, uncertainties can be reduced allowing greater confidence in the results. The limiting factor in most cases will likely be the availability of high quality historical data from which to build a DEM.

The science is in the early stages of debating what can be accurately done and what should not be attempted with this type of analysis. For
the present state of the science, uncertainties in historical cartographic data tend to be large. Under these circumstances, long-term DoD projects are best reserved for systems that have experienced large amounts of geomorphic change. If uncertainties can be reduced, however, these methods may be expanded to systems that have undergone more subtle changes.

When and how to neglect changes in DoDs that are smaller than some measure of uncertainty should be considered. The tradeoffs in this question are the ability to omit spurious measures that result from errors rather than change, versus adding a bias by systematically removing small changes in elevation that are cumulatively large when added over extensive areas. The spatial distribution of errors, in addition to the total magnitude, may be important to establishing confidence in results and interpreting processes. Geomorphic changes are not likely to be uniformly distributed. This study documents a case where errors in a historical DEM are dominantly generated by horizontal offsets and associated with steep scarps along stream banks and terraces. By modeling the spatial distribution and magnitude of cumulative errors and the combination of errors from multi-date change, it may be possible to spatially distribute the confidence in ‘apparent’ changes in an effort to determine and explain the real geomorphic changes.

The three case studies briefly documented here utilized detailed topographic surveys more than 70 years old to construct DEMs and compute difference maps. These studies had substantial limitations, each in different ways. Yet, each also provided considerable insight into geomorphic processes and quantitative, spatially distributed estimates of erosion and deposition — again, each in different ways. Although flawed by serious limitations in the historic map registration and LiDAR ground point densities, DoD analysis of the Cox Gully in South Carolina, demonstrates that new branches of gullies can be mapped and specific processes of erosion and deposition can be identified from the DoD. At Shanghai Bend on the Feather River, California, the incipient stages of a major channel avulsion was captured by a detailed 1909 topographic survey and collection of LiDAR and SONAR data in 1999 allow a detailed DoD to quantitatively document those changes in a spatially distributed modeling environment. This analysis shows that the sediment budget was dominated by main-channel erosion, but that substantial volumes of sediment were stored in abandoned channels. Patterns of erosion and sedimentation along the lower Yuba River document a 93-year period following a major episode of floodplain morphogenesis. Channel planform adjustments had already been largely made by 1906, the baseline for the period, but deep main channel incision and natural levee formation during the recovery is quantitatively documented. The latter two studies demonstrate the utility and limitations of volumetric analyses conducted by DoD over decades to centuries.

The methods and examples given in this paper focused on conventional geomorphic change or volumetric analysis, but this could easily be expanded. For example, computation of primary and secondary terrain derivatives from DEMs and other mainstay methods of geomorphometry could be performed on historical DEMs or DoDs.

Acknowledgments

This research benefitted greatly from the help by a number of individuals. The ideas and limitations of this paper are ours and we take full responsibility for them, but several people provided considerable intellectual input for which we are deeply thankful. Michael Singer and Rolf Aalto have been constant sources of stimulation in their discussions of floodplain morphogenesis and sedimentation processes in the region. Catherine DeMauro provided extensive logistical support during field work and data acquisition in the region. Historical maps were obtained from the University of California, Davis Map Library, the California Department of Water Resources, the Center for Sacramento History, and the California State Archives. The U.S. Army Corps of Engineers provided 1999 LiDAR and sonar data via the late Douglas Allen who reconstructed the LiDAR data set. The National Science Foundation provided funding

example, a better understanding is needed of the sources and spatial patterns of uncertainties and how to deal with them in the development and interpretation of DoDs. Comparing the magnitude of geomorphic change to uncertainties (e.g., by using a signal-to-noise ratio) is an important concept to consider at all phases of a DoD project. The DoD method is best suited where change is great relative to uncertainty.
for a project in this area (BCS 0520933) that helped us to gather the data from which much of the case study work was conducted.

References


California Debris Commission (CDC), 1906. Map of the Yuba River, California from the narrows to its mouth in the Feather River. Made under direction of Major Wm. W. Harr, U.S. Army Corps of Engineers, by G. C. McDaniel, Jr., August to Nov., 1906. 1:9600; 0.6-m contours.


