

Landscape-scale geomorphic change detection: Quantifying spatially variable uncertainty and circumventing legacy data issues



Keelin R. Schaffrath*, Patrick Belmont, Joseph M. Wheaton

Department of Watershed Sciences, 5210 Old Main Hill, Utah State University, Logan, UT 84322, USA

ARTICLE INFO

Article history:

Received 13 May 2015

Received in revised form 23 September 2015

Accepted 24 September 2015

Available online 28 September 2015

Keywords:

Lidar

Change detection

DEM/DTM

Fuzzy logic

ABSTRACT

Repeat surveys of high-resolution topographic data enable analysis of geomorphic change through digital elevation model (DEM) differencing. Such analyses are becoming increasingly common. However, techniques for developing robust estimates of spatially variable uncertainty in DEM differencing estimates have been slow to develop and are underutilized. Further, issues often arise when comparing recent to older data sets, because of differences in data quality. Airborne lidar data were collected in 2005 and 2012 in Blue Earth County, Minnesota (1980 km²) and the occurrence of an extreme flood in 2010 produced geomorphic change clearly observed in the field, providing an opportunity to estimate landscape-scale geomorphic change. Initial assessments of the lidar-derived digital elevation models (DEMs) indicated both a vertical bias attributed to different geoid models and localized offset strips in the DEM of difference from poor coregistration of the flightlines. We applied corrections for both issues and describe the methods we used to discern those issues and correct them. We then compare different threshold models to quantify uncertainty. Poor quantification of uncertainty can erroneously over- or underestimate real change. We show that application of a uniform threshold, often called a minimum level of detection, overestimates change in areas where change would not be expected, such as stable hillslopes, and underestimates change in areas where it is expected and has been observed, such as channel banks. We describe a spatially variable DEM error model that combines the influence of slope, point density, and vegetation in a fuzzy inference system. Vegetation is represented with a metric referred to as the cloud point density ratio that assesses the complete point cloud to describe the density of above ground features that may hinder bare-earth returns. We compare the significance of spatially variable versus spatially uniform DEM errors on change detection by thresholding the DEM of Difference at a 95% confidence interval (2σ). Results indicate significant geomorphic change in relatively predictable locations, such as erosion on the outside and deposition on the inside, of bends. Final totals indicated net erosion of $[2,625,100] \pm 2,389,000$ m³ in the county between 2005 and 2012. Of this, 39% was generated from bluffs, 1% from ravines, and the remainder came from banks and flood-plain areas.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Quantifying spatial patterns of erosion and deposition at a landscape scale is vitally important for improving models of water quality and sediment routing, predicting channel dynamics at the network scale, advancing landscape evolution models, and providing constraints for watershed sediment budgets (Smith et al., 2011; Mueller and Pitlick, 2013; Passalacqua et al., 2015). The past decade has witnessed rapid advances in our ability to collect high-resolution topographic (HRT) data and concomitant improvements in computing power, which have increased our ability to acquire, store, and analyze large data sets. As

such, detection of spatial and temporal patterns of landscape-scale erosion and deposition is increasingly feasible (Collins and Walling, 2004).

Differencing sequential sets of HRT can be used to detect and quantify geomorphic change (Brasington et al., 2003; Lane et al., 2003; Wheaton et al., 2010b; Day et al., 2013a) in order to better understand river form and process (Brasington et al., 2000) or to infer rates of sediment transport and channel morphology (Ashmore and Church, 1998; Wheaton et al., 2013). Previous studies involving DEMs of Difference (DoDs) have been conducted on reach- and small catchment-scales, utilizing DEMs derived from airborne lidar, total station, or real-time kinematic GPS (rtkGPS) surveys (Brasington et al., 2000; Bangen et al., 2014; Goodwell et al., 2014), aerial photography and photogrammetry (Lane et al., 1993, 2003; Rumsby et al., 2008; James et al., 2012), or terrestrial lidar (Milan et al., 2011; Abellán et al., 2014; Leyland et al., 2015). At all scales and for all methods, users must understand: (i) the technology

* Corresponding author.

E-mail address: keelin.r.schaffrath@gmail.com (K.R. Schaffrath).

and its limitations at the time of data collection; (ii) how post-processing steps (point cloud classification and generation of the gridded product) for each individual data set might affect the results; and (iii) georeferencing information for the original data, as systematic errors can be introduced at any one of these steps.

We present a landscape-scale analysis of repeat lidar to demonstrate methods to identify and rectify issues that may arise with older, 'legacy' data and illustrate the value of spatially variable DEM error modeling. Specifically, we present methods to address common data issues with legacy lidar data sets, and we quantify DEM error using uniform and spatially variable uncertainty models. We develop and test a 'cloud point density' metric to accurately quantify uncertainty owing to vegetation density. We discuss important considerations for developing robust estimates of uncertainty in similar large data sets and illustrate how different approaches may exclude useful information, or incorporate erroneous information, into the DoD. Finally, we investigate the significance of the above analyses in the context of interpreting geomorphic change in a highly dynamic landscape that has recently experienced a historically unprecedented flood bracketed by the two lidar data sets.

2. Study area

The Minnesota River basin is located in southern Minnesota and is a tributary to the Mississippi River (Fig. 1). The major tributaries to the Minnesota River, especially those in Blue Earth County, are experiencing rapid geomorphic change as a result of a 70-m drop in base level 13,400 years ago as Glacial Lake Agassiz catastrophically drained through the proto-Minnesota River, compounded by recent agricultural land use and climate change (Belmont et al., 2011; Gran et al., 2013; Schottler et al., 2014). Presently, tributaries such as those in the Greater Blue Earth River basin continue to respond to the drop in base level as a knickzone has propagated 35–80 km upstream through the channel network (Belmont, 2011; Gran et al., 2013). Fine-grained, semi-consolidated glacial till behaves like weak bedrock (Jennings, 2010; Runkel et al., 2011) and incision in the knickzone has left behind 3–70 m tall bluffs that line about 30% of the channel. Bluffs are subjected to multiple erosion processes including fluvial toe erosion, as well as slumping or block failure from groundwater sapping or freeze-thaw processes (Day et al., 2013a). Additionally, about 275 ravines have been identified in the region that connect the uplands to the incised mainstem tributary channels within the knickzone.

An extreme flood occurred in September 2010 in the basin. This flood was the second largest on record (1910 to present) in the Blue Earth River and the flood of record (1940 to present) in the Le Sueur

River (23% larger than the next largest flood, in 1965). In the Blue Earth River, this flood event has a return period of about 25 years using a Log Pearson Type 3 distribution. For the Le Sueur, the event is essentially undefined (>200 -year return period) if all years of record are analyzed using a Log Pearson Type 3 distribution. Acknowledging that the Le Sueur River exhibits highly nonstationary hydrology (Schottler et al., 2014), we conducted a Log Pearson Type 3 analysis using only post-1975 peak flows. That analysis indicates a return period exceeding 100 years. Field evidence indicated dramatic geomorphic change throughout the county. Such a large flood in a geomorphically active area provides a unique opportunity to analyze geomorphic change over this 1980 km² area.

3. Methods and methodological results

Airborne Light Detection and Ranging (lidar) data were collected for Blue Earth County, MN in 2005 and 2012. The 2005 data were collected April 13–15 and 24–25 using an Optech 3100 ALTM 70 kHz laser system and Realm Terrascan and Geocue Survey processing software mounted on a fixed-wing aircraft (Optimal Geomatics, 2005a). Aerial orthophotography and lidar data were simultaneously collected in 2005 at 1800 m above ground level. Vertical accuracy of the data was estimated by the vendor (Optimal Geomatics) by calculating the root mean square error (RMSE) of elevations for 350 checkpoints collected within four subareas in different land cover types (open terrain, tall weeds and crops, brush and low trees, forested areas, and urban areas). Checkpoint locations are shown in Fig. 2. Reported RMSE ranged from 0.17 m (brush) up to 0.25 m (tall weeds) and averaged 0.24 m, but the 'forested' cover error was not reported (Optimal Geomatics, 2005b). Lidar intensity data and lidar grammery were used to generate breaklines for hydrologic features regardless of size. Additional checkpoints were collected by other entities at the time of data collection (total 695 points). Mean point density of the 2005 bare-earth point cloud is 0.16 pts m⁻² (standard deviation 0.08 m). Data were distributed in North American Datum 1983 in Universal Transverse Mercator, zone 15N, coordinates with orthometric heights converted using the GEOID03 model.

In 2012, data were collected April 4–8 using a Leica ALS 70 lidar system and processing software (TerraTec TerraPOS, TPAS TC GNSS/INS Processor version 3.10, IPAS Pro version 2.01.02) mounted on a fixed-wing aircraft flying at 2260 m above ground level with a swath width of 40° and 30% sidelap (Blue Earth County, 2012). Vendors collected 145 checkpoints spatially distributed throughout the county, and reported vertical accuracy (RMSE) was 0.11 m for 26 points in the 'open' cover category. Hydrologic breaklines were hand-digitized for rivers

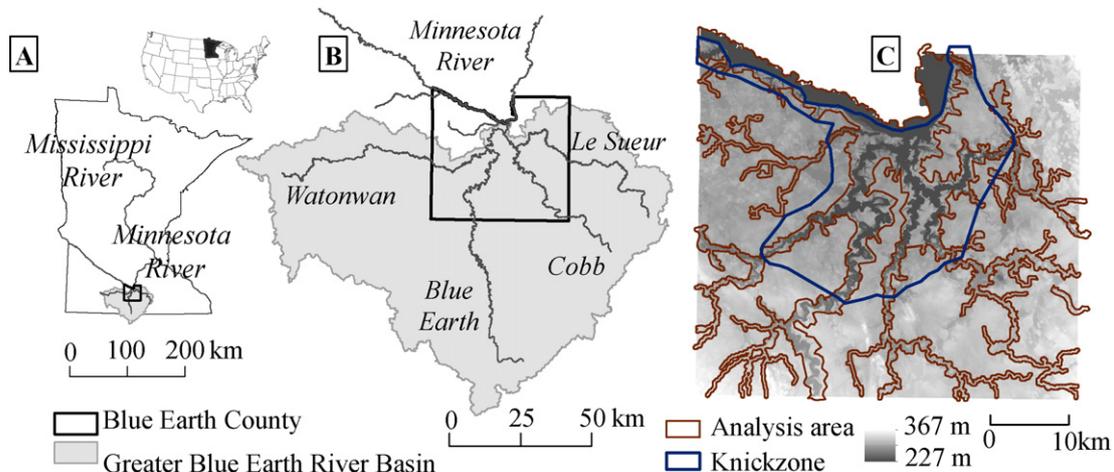


Fig. 1. Maps of the (A) state of Minnesota showing the Minnesota and Mississippi Rivers, (B) location of Blue Earth County in the Greater Blue Earth River Basin, and (C) 2012 digital elevation model of Blue Earth County with a blue line indicating the area inside the knickzone and a brown line indicating the area near the channels where the analysis was completed.

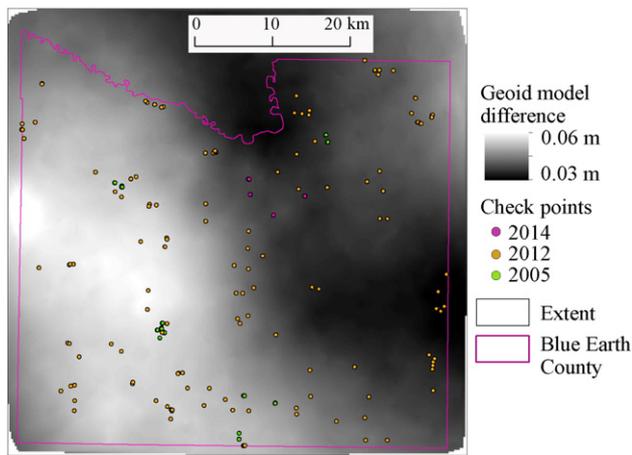


Fig. 2. Systematic vertical offset shown by the error surface generated from differences in the GEOID03 and GEOID09 models; and locations of checkpoints for each year of data collection and additional points collected for this study in 2014.

wider than 30 m using the lidar intensity data and the unenforced elevation model (R. Merry, Aerometric, April 26, 2013; U.S. Geological Survey National Geospatial Program, 2010). Mean point density of the 2012 bare-earth point cloud is 1.4 pts m^{-2} (standard deviation 0.34 m). Data were distributed in North American Datum 1983 (Blue Earth County High Accuracy Reference Network) in Universal Transverse Mercator, zone 15N, coordinates with orthometric heights converted using the GEOID09 model.

3.1. Digital elevation model offset

A raw, uncorrected DoD was generated by subtracting the two DEMs, adjusted accordingly to ensure concurrency, orthogonality, and dimensional divisibility. We started with the vendor-generated DEM raster and used the geomorphic change detection (GCD) add-in for ArcGIS (found at gcd.joewheaton.org; version 6; Wheaton et al. 2010a) to adjust the DEMs to ensure concurrency, orthogonality, and dimensional divisibility of the data sets. The raw, uncorrected, DoD suggested pervasive deposition throughout the county, including flat areas where no measurable change is expected to occur over the seven years between surveys, suggesting a systematic bias. Initial attempts to identify the source and spatial distribution of the bias were to evaluate the difference between the two DEMs at the rtkGPS checkpoints collected for each flight and digitized points in areas where we did not expect change. However, metadata for the two data sets indicated that different GEOID models were applied to the data for each year to convert ellipsoid heights to orthometric heights.

Differences between lidar and rtkGPS checkpoints collected in 2012 tend to be positive relative to the 2005 data, and a systematic positive bias pervades over all elevation differences. An independent measure to identify a systematic vertical offset between the 2005 and 2012 DEMs was attained by comparing elevations of points where we were confident no vertical change (erosion or deposition) had taken place. The median difference between elevations in the 2012 and 2005 DEMs is 9.2 cm for points digitized on roads and 3.7 cm for randomly generated points. Results were stratified by geographic area or 2012 flight date in order to determine whether geography or flight date explained spatial variability in the error, but no systematic trends were observed.

Lidar vendor metadata indicated that GEOID03 was applied in 2005 and GEOID09 was applied in 2012. To correct for the offset between the two models, we generated a surface of the difference in geoid models and applied the correction to the 2005 DEM (Fig. 2). Specifically, we generated 1700 random points and fed them into the National Geodetic Survey's interactive calculators (see: <http://www.ngs.noaa.gov/GEOID/>) for each geoid model. The two geoid models were differenced, and the increase to ellipsoid heights associated with the GEOID09 model is 3–7 cm higher than the GEOID03 model, though the difference is spatially heterogeneous (Fig. 2). Adding the geoid correction surface to the original 2005 DEM provided an adjusted 2005 DEM.

Vendor-reported vertical RMSE is 0.24 m for 2005 data and 0.11 m for 2012 data. However, generic RMSE values are not effective indicators of uncertainty for geomorphic change detection at the landscape scale because they fail to indicate systematic bias, dispersion above or below zero, or the spatial patterns of uncertainty (Fisher and Tate, 2006), all of which are critical to identifying real geomorphic change.

In 2005 (695 checkpoints), mean error was 0.004 m with a standard deviation of 0.11 m. In 2012 (145 checkpoints), mean error was 0.08 m and standard deviation was 0.16 m, indicating a positive bias in the 2012 data. Kessler et al. (2012) detected a bias when they compared lidar-derived DEMs from the same 2005 data as in this study and data collected in 2009 (not available for this study). In quantifying bank erosion, they reported the bias but did not adjust their data because it was less than the vertical accuracy guaranteed by vendors (Kessler et al., 2012). In the case of a large study area such as ours, even small systematic offsets can yield large errors in estimated volumes of sediment eroded or deposited. Moreover, despite what lidar vendors may report for vertical accuracies, there are no guarantees that this is not exceeded locally.

Each vendor also distributed 1-m DEMs with the lidar data. For 2005 data, bare-earth points and the hydrolines were used to create a terrain data type from which a DEM was derived using nearest neighbors interpolation. With point density of 0.16 pts m^{-2} , 'NoData' holes are likely quite prevalent throughout the county, but vendor explanations of DEM generation did not address this issue. Sparse ground returns introduce uncertainty because elevations are interpolated over large areas with no data. Metadata for 2012 indicate that the DEM was generated using TerraModeler. The 2012 DEM was transformed from the NAD83 (HARN) datum to NAD83 to match the 2005 DEM. The UTM 15N coordinate system was retained for all points and grid products. We adjusted the 2005 DEM using the geoid correction surface and used that adjusted 2005 DEM for all differencing. All differencing was then completed on DEMs adjusted by the GCD tool to ensure orthogonality and dimensional divisibility (extents were rounded to the nearest 10 m).

3.2. Horizontal uncertainty between digital elevation models

Horizontal offset between DEMs is difficult to quantify, and most studies either ignore that source of error or report the manufacturer-specified minimum error (Hodgson and Bresnahan, 2004; Kinzel et al., 2007). In 2005, horizontal error was reported as 1/2000th of flight altitude (1836 m) which was 0.91 m. In 2012, the vendor reported a horizontal accuracy of 0.44 m based on manufacturer specifications. While both of these numbers are $<1 \text{ m}$, the cell resolution of the DEM, even small horizontal shifts can introduce additional vertical error especially in areas of steep slope, such as bluffs or stream banks (Hodgson and Bresnahan, 2004; Streutker et al., 2011; James et al., 2012). To determine if there was a systematic horizontal bias, we compared x and y coordinates of 63 digitized points on well-defined structures such as buildings, berms, or dams on the two DEMs. Differences in northing and easting were used to calculate the angle (direction) and the magnitude of the vector between the two points. Measured offset varied from 0.4 to 5 m. However, direction of error was randomly distributed.

While we did not detect any systematic horizontal error (Fig. 3, plot generated using heR.Misc package developed by Klepeis, 2004), the DoD indicated localized anomalies. Most notably, east–west strips of elevation differences unlikely to be real erosion or deposition were visible (Fig. 4A). Additionally, many ravines showed erosion on one aspect and deposition on the other. Such spatial patterns indicative of systematic errors have been documented in other studies (Bretar et al., 2004; Habib et al., 2008; Akca, 2010; Streutker et al., 2011; DeLong et al.,

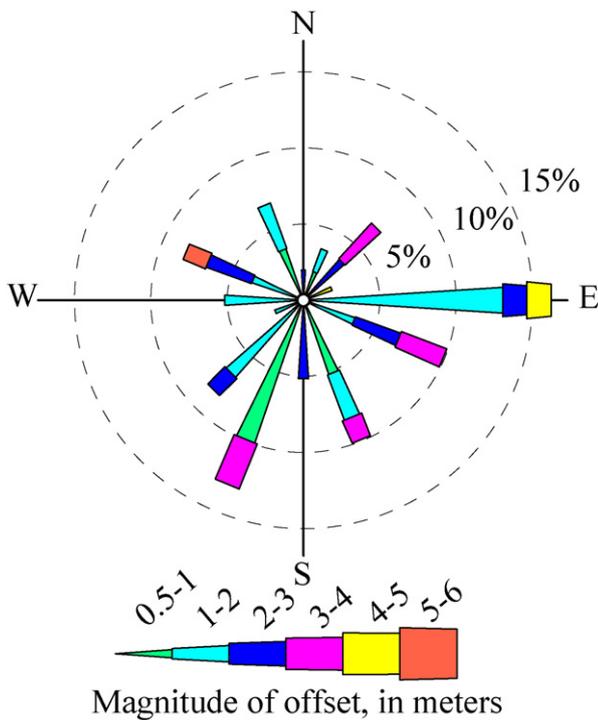


Fig. 3. Horizontal offset between 2005 and 2012 DEMs as a rose diagram. Direction of offset is shown based on the position around the compass and magnitude is shown by color. For example, the majority (~17% of points evaluated) of error is in the easterly direction. The plot shows about 13% of the easterly error is 1–2 m, 2% is 2–3 m, and 2% is 4–5 m in magnitude.

2012; Glennie et al., 2014; Goodwell et al., 2014) and attributed to poor coregistration of flightlines.

Lidar data are collected in strips along the flightlines and, with each strip, vendors generate areas of overlap, typically 10–30%. When calibration parameters are not precisely determined, discrepancies between the flightlines can occur, known as coregistration errors, which are common in older ‘legacy’ lidar data sets. Errors may not be detectable in the standalone surface. However, when differenced with recent and more precisely calibrated data, the offset strips might be obvious, as was the case for the Blue Earth County data (Fig. 4A). The most accurate method of correction would be to use the original GPS/IMU and laser measurement data to recalibrate the raw data (Glennie et al., 2014). However, such data are rarely available for legacy data sets. Another technique uses local slope and local elevation difference between adjacent points in the overlap area to realign flightlines (Streutker et al., 2011). To apply this approach, users need overlap points (points with the same x and y coordinate) and to be able to separate overlap points by flightline. Separate flightlines might be identified using an associated flightline number or using a time stamp associated with the timing of collection of each point. With sufficient information to separate flightlines, the Iterative Closest Point algorithm (Besl and McKay, 1992) would be another approach. However, 2005 data included only x , y , and z coordinates for bare-earth and first return points, making these techniques impossible. Goodwell et al. (2014) applied a fast Fourier transform to the differential DEM in cross sections perpendicular to the observed stripping problem. They detected two dominant wavelengths associated with the strips and filtered them out of the data set. A fast Fourier transform of north to south transects of our Blue Earth County DoD indicated many dominant wavelengths that would require filtering and the filtered profile and power spectra still showed evidence of the systematic stripping error (Fig. 4D–E). Thus, this approach was not effective in our case.

To correct the DoD for the flightline stripping error, we iteratively generated multiple correction surfaces based on random points

generated throughout the county. Points were randomly generated outside of a 200-m buffer of the vendor-generated hydrolines, in areas where we were confident little to no real change had occurred. Strips trend east–west, while channels in Blue Earth County trend south–north so we expected interpolation error across the buffer to be minimal. A total of 12 surfaces were iteratively generated using a range of points (33,000, 150,000, 200,000, or 500,000) and different interpolation methods. Differences between surfaces using 200,000 and 500,000 points were minimal, yet computational time was reasonably short with 500,000 points, so we used that point density to generate our correction surface. For each point, we extracted the value of the DoD and interpolated a surface using inverse distance weighting (IDW), universal and ordinary kriging, and regularized splining (all tools are in the ArcGIS toolbox). In the process, artifacts were identified, particularly associated with lakes and reservoirs where large changes in water-surface elevations appear in the DoD. Those areas were eliminated from the correction surfaces. Of the 12 surfaces, three were chosen for comparison based on the quantity of detectable artifacts present in the surface and qualitative matching with the visible strips in the DoD. Each of the three correction surfaces were subtracted from the DoD. Elevation profiles (Fig. 4E) were extracted to evaluate the data retained in each surface relative to the raw surface. Also, power spectra (Fig. 4D) of the fast Fourier transform of the north–south transect (seen in Fig. 4A) were compared to evaluate how much of the stripping had been removed. The north–south transect was intentionally drawn perpendicular to the strips and with minimal intersections with water bodies and therefore represents a surface where real geomorphic change is minimized. The final surface (DODc, with c indicating ‘corrected’) removed evidence of stripping, while retaining the inherent variability of the original data (Fig. 4A), as can be seen at the scale of the entire data set in Fig. 4C and at finer scales, according to the power spectrum (Fig. 4D, green line) and elevation transects (Fig. 4E, green line). The final correction surface was generated using 500,000 points interpolated with the spherical model of ordinary kriging, a 5-m fixed search radius, and a minimum of 12 points. This method of correcting a DoD is effective for applying a spatially variable correction to the stripping problem when data needed for other approaches are not available, as might be the case for legacy data.

3.3. Digital elevation model error modeling

Sources of uncertainty in airborne lidar are spatially heterogeneous and include positional errors realized in the original point cloud as well as vertical errors and interpolation errors ultimately expressed in the raster DEM. Proxies for these sources of error include survey point quality and density, land-surface composition (vegetation type and density and topographic complexity), sampling strategy, and decisions in point classification and DEM interpolation (Hodgson et al., 2003, 2005; Hodgson and Bresnahan, 2004; Heritage et al., 2009; Wheaton et al., 2010a; Milan et al., 2011; Schwendel et al., 2012). Various methods exist for estimating DEM errors, but spatially uniform estimates based on an RMSE summary with checkpoint data are by far the most common. At the other end of the spectrum, spatially variable error modeling (including full error budgeting and developing statistical models of error) often require data density and additional information typically not supported nor available from legacy lidar data sets. Wheaton et al. (2010b) developed a simpler, more pragmatic technique for estimating spatially variable DEM errors that allows proxies for factors contributing to DEM error to be provided as spatially variable inputs to a fuzzy inference system (FIS). The FIS allows flexibility in the specification of inputs that can reasonably be modeled spatially, and their combined influence on DEM quality is described heuristically with simple rules. The FIS differs from a rule system in that input and output categories are represented as fuzzy numbers, allowing explicit overlap between categories where ambiguity exists. In practice, this translates to multiple rules often applying for a combination of inputs

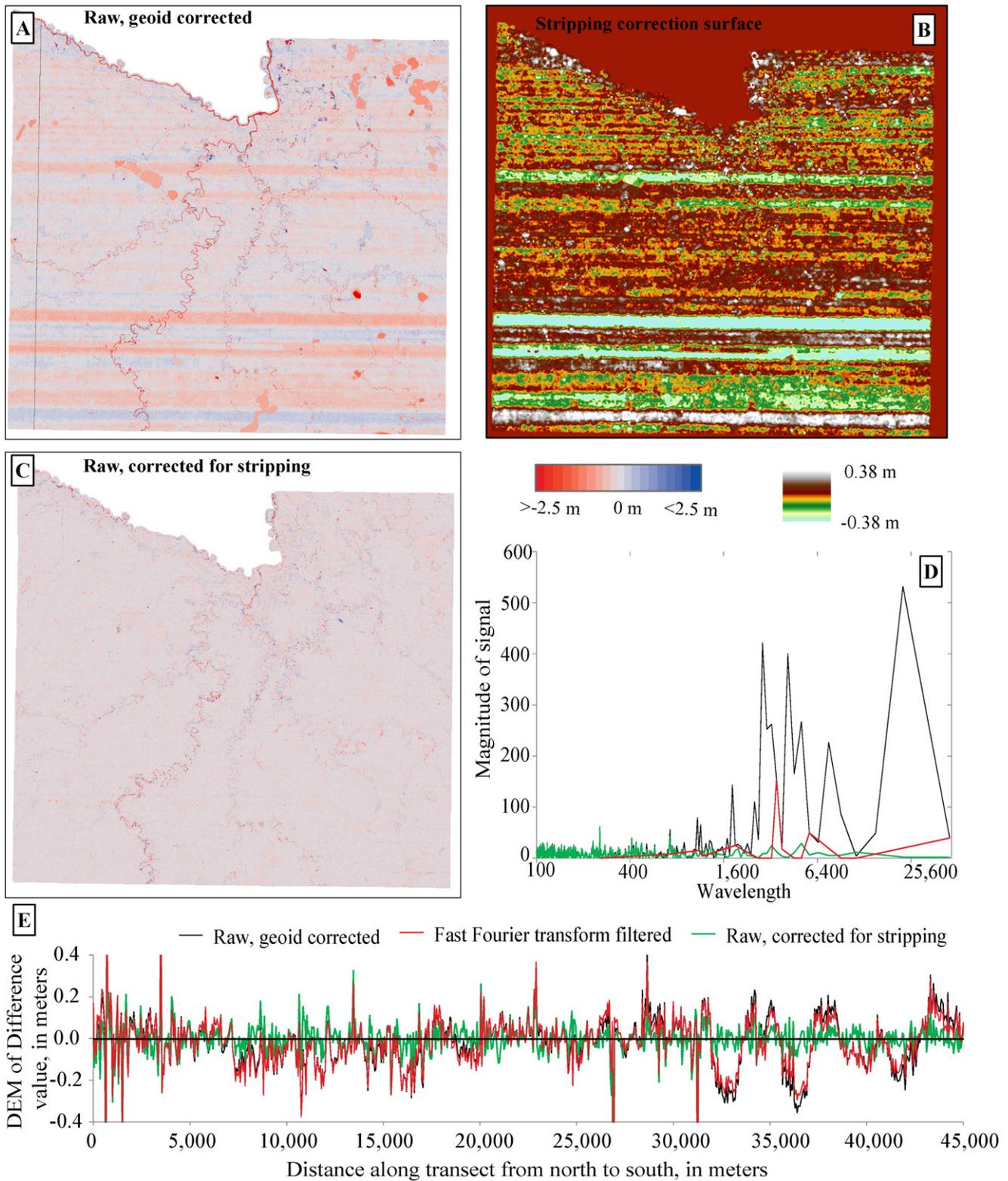


Fig. 4. Maps and plots showing the initial characteristics associated with stripping and the correction applied. (A) Geoid-corrected DEM of Difference showing distinct strips in the east–west direction. (B) Error surface generated to correct stripping. (C) DEM of Difference corrected for stripping using the surface shown in B. Note the east–west strips are no longer evident and the surface is a mix of red (erosion) and blue (deposition). Two plots (D–E) correspond with data collected at points on a 50-meter interval on the black line shown in panel A. (D) Power spectra plot where the black line refers to the data from the geoid-corrected DEM of Difference, the red line is the output after applying a filter based on a fast Fourier transform, and the green line is the output after the application of the stripping correction. (E) Elevation profile shows the DEM of difference values for the geoid-corrected (black), fast Fourier transform filtered (red), and final stripping correction (green).

and giving rise to more nuanced behavior in the output (the total consequent). The outputs themselves are calibrated to an independently determined range of error estimates specific to the survey methods.

Wheaton et al. (2010b) demonstrated the development and application of an FIS for an rtkGPS survey, but suggested FIS models should be developed for different survey techniques and applications. Here, we

developed FIS DEM error models for the lidar surveys described above. For each of the two years of lidar data, an FIS was developed using inputs of slope, point density, a vegetation metric, and inundation. Each FIS output was calibrated using independent estimates of DEM error from

checkpoint data and simple univariate correlations associated with the inputs of slope, point density, vegetation height or density, and whether the pixel was inundated (Fig. 5). Checkpoints were initially collected by the lidar vendors at the time of survey, with 695 points collected in 2005

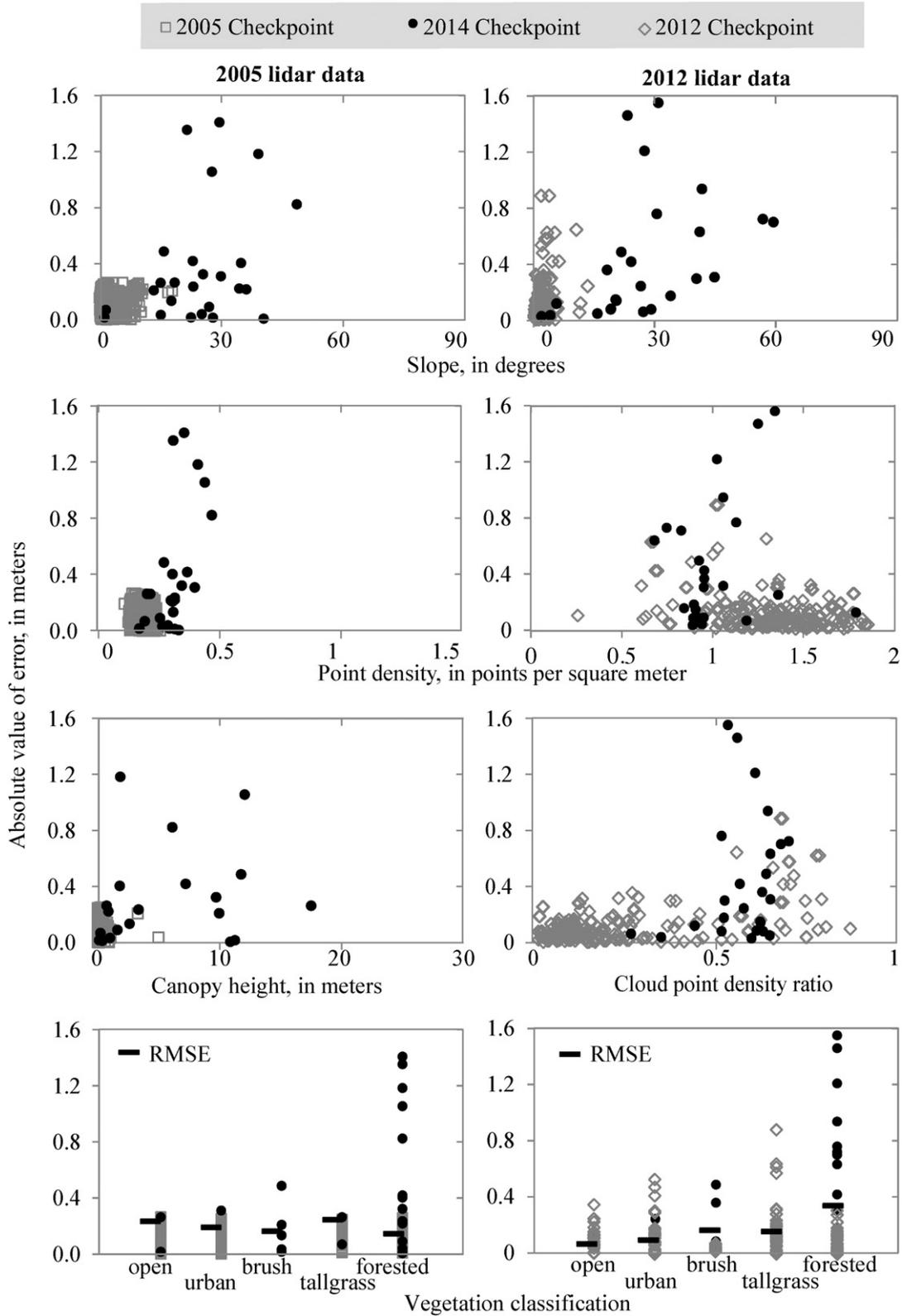


Fig. 5. Plots showing absolute value of elevation error (lidar elevation minus GPS elevation) relative to values of the input parameters to the 2005 and 2012 Fuzzy Inference Systems for 2005 checkpoints (column on the left, data shown in gray boxes); 2012 checkpoints (column on the right, data in gray diamonds); and additional points collected in 2014 (shown in all plots, black circles).

and 145 points in 2012. Because the 2005 checkpoints were clustered in only a few topographically simple areas (Fig. 2), ranges of error for each parameter are limited (Fig. 5, left column) and do not include the full range of error associated with more complex topography present within the survey extents. Checkpoints associated with the 2012 flight were better distributed spatially, but still did not cover the full range of conditions controlling error (Fig. 5, right column).

To better constrain uncertainty in more topographically complex environments, in June 2014 we collected additional checkpoints in areas of high slope using rtkGPS and densely vegetated areas using a total station (Fig. 2). We targeted high-slope and high-vegetation density areas near ravines and low-slope, high-vegetation density areas near roads where the initial DoD showed little to no elevation change. Reoccupation of some of the original 2005 and 2012 checkpoints confirmed our 2014 data collection effort was within 11 cm, on average, of the 2012 original data collection and within 14 cm of the 2005 data collection. Elevation differences between all checkpoints and the 2005 and 2012 DEMs were related to the values of slope, point density, or either of the two vegetation parameters (vegetation height in 2005 or vegetation density in 2012), and those relations provided initial guidelines for calibrating the FIS rules and outputs (Table 1, Fig. 5). Data collected in 2014 greatly extended our characterization of uncertainty associated with slope (both years) and point density and canopy height (for 2005 data, see black dots, Fig. 5).

Where possible, we wanted to represent the impact of vegetation on DEM error in our FIS. Vegetation can be an important source of uncertainty in bare-earth lidar DEMs because the probability a laser pulse will penetrate down to the ground decreases as vegetation density increases (Su and Bork, 2006; Reutebuch et al., 2014). Last returns are typically classified as bare-earth points, but the likelihood that the return is

Table 1

Rules for the Fuzzy Inference Systems used to infer elevation uncertainty in 2005 and 2012 data from slope, point density, vegetation, and whether the area was inundated or dry. Each column reads as part of an if/then statement: If slope is ____, and point density is ____, and vegetation is ____, then elevation uncertainty is ____. Units associated with the header are the units of the inputs and outputs.

Slope, degrees	Point density, points per square meter	Vegetation metric ^a	Inundated or dry	Elevation uncertainty, meters
Low	Low	Low	Dry	Average
Low	Low	Medium	Dry	Average
Low	Low	High	Dry	Average
Low	Medium	Low	Dry	Low
Low	Medium	Medium	Dry	Low
Low	Medium	High	Dry	Low
Low	High	Low	Dry	Low
Low	High	Medium	Dry	Low
Low	High	High	Dry	Low
Medium	Low	Low	Dry	High
Medium	Low	Medium	Dry	High
Medium	Low	High	Dry	High
Medium	Medium	Low	Dry	Average
Medium	Medium	Medium	Dry	Average
Medium	Medium	High	Dry	High
Medium	High	Low	Dry	Low
Medium	High	Medium	Dry	Average
Medium	High	High	Dry	Average
High	Low	Low	Dry	High
High	Low	Medium	Dry	High
High	Low	High	Dry	High
High	Medium	Low	Dry	Average
High	Medium	Medium	Dry	High
High	Medium	High	Dry	High
High	High	Low	Dry	Average
High	High	Medium	Dry	High
High	High	High	Dry	High
-	-	-	Inundated	High
-	-	Extreme	-	Extreme

^a In 2005, the vegetation metric was vegetation height in meters; and in 2012, it was the cloud point density ratio.

actually from the bare earth decreases as vegetation density increases. Vegetation height has typically been used as a surrogate for uncertainty in elevation owing to vegetation (Hodgson and Bresnahan, 2004; Hopkinson et al., 2005; Bangen et al., 2014). However, tree height and vegetation density are not always well correlated. For example, consider a deciduous forest with a dense understory relative to a sparse, coniferous forest with no understory (see Hodgson and Bresnahan, 2004). In these two examples, vegetation height may not differ very much, but vegetation density may differ substantially. We postulate laser penetration should be more sensitive to vegetation density than height. Data in 2005 were limited to bare-earth and first returns, making vegetation height the only possible surrogate for uncertainty owing to vegetation for the 2005 DEM. However, the LAS 1.4 format used for the 2012 data provided access to all returns and associated metadata including scan angle, return number, total returns, and a classification based on an industry standard. To consider the influence of a model for uncertainty owing to vegetation for the 2012 data, we created a metric called the cloud point density ratio, which uses the majority of returns with their classifications. Where vegetation is dense, we expect the point cloud above the ground to also be dense.

The cloud-point density ratio is the ratio of aboveground point density to total point density. Industry standard point-cloud classifications included bare-earth, vegetation, water, buildings, overlap, error or noise, and unclassified. Total point cloud density is the point density of all points except those classified by the vendor as noise or overlap. Aboveground points are all points from the total point cloud not classified as bare-earth. This metric depends on the vendor classifications. Vendors consider this information proprietary so reporting does not include details on the procedures used. For the purposes of this study, we chose to compare vendor-generated products, and our metric depends on the accuracy of those vendor classifications. In part, this was because we did not have the full point cloud in 2005. When full point clouds are available to users for each data set, we suggest that users reprocess point clouds and classify points using identical procedures and parameter choices.

Cloud point densities closer to 1 indicate a high density of aboveground points relative to the total point cloud and therefore a higher uncertainty owing to vegetation (Fig. 5). Checkpoint data shown in Fig. 5 have canopy height from 2005 positioned next to the cloud point density ratio of 2012. Both data sets indicate at least a slight positive trend, though the trend in 2012 is more clearly evident, suggesting that the cloud point density is a more accurate representation of uncertainty owing to vegetation.

Once the inputs for our FIS error models were prepared, we needed to build our rule tables and calibrate membership functions for the different FIS models (Table 1). The GCD tool is provided with built-in FIS models for various data collection techniques including airborne lidar. One built-in FIS for airborne lidar includes slope, point density, vegetation height, and inundation as input parameters. The rule table for the 2005 and 2012 FIS is shown in Table 1. We specified membership functions for the inputs and outputs in 2012 and 2015 as shown in Fig. 6. Three separate FISs were tested on a small area where an existing bluff is known to have eroded during the study period (Fig. 7, panels 3A and 3B; Table 2). Each FIS varied in the range of slope or elevation error values in each membership function. Slope and elevation error values were modified based on checkpoint data. Net change results for bluffs were compared to each other and also to the reported erosion rate measured by Terrestrial Laser Scanning (TLS) for one bluff from another study in the same area, 132 m³ y⁻¹ (TrB1 in Day et al., 2013b), which is not a perfect comparison because the time periods do not coincide exactly, but is a useful comparison nevertheless. Net change on the bluff using the chosen FIS was 223 m³ y⁻¹ which is higher than the TLS estimate (FIS 3 in Table 2). This discrepancy was expected because the TLS estimates pre-date the extreme 2010 flood. The TLS estimates were over a period from August 2007 to May 2010 when the maximum peak flow was only 9180 cfs in the Blue Earth River and 5610 cfs in the

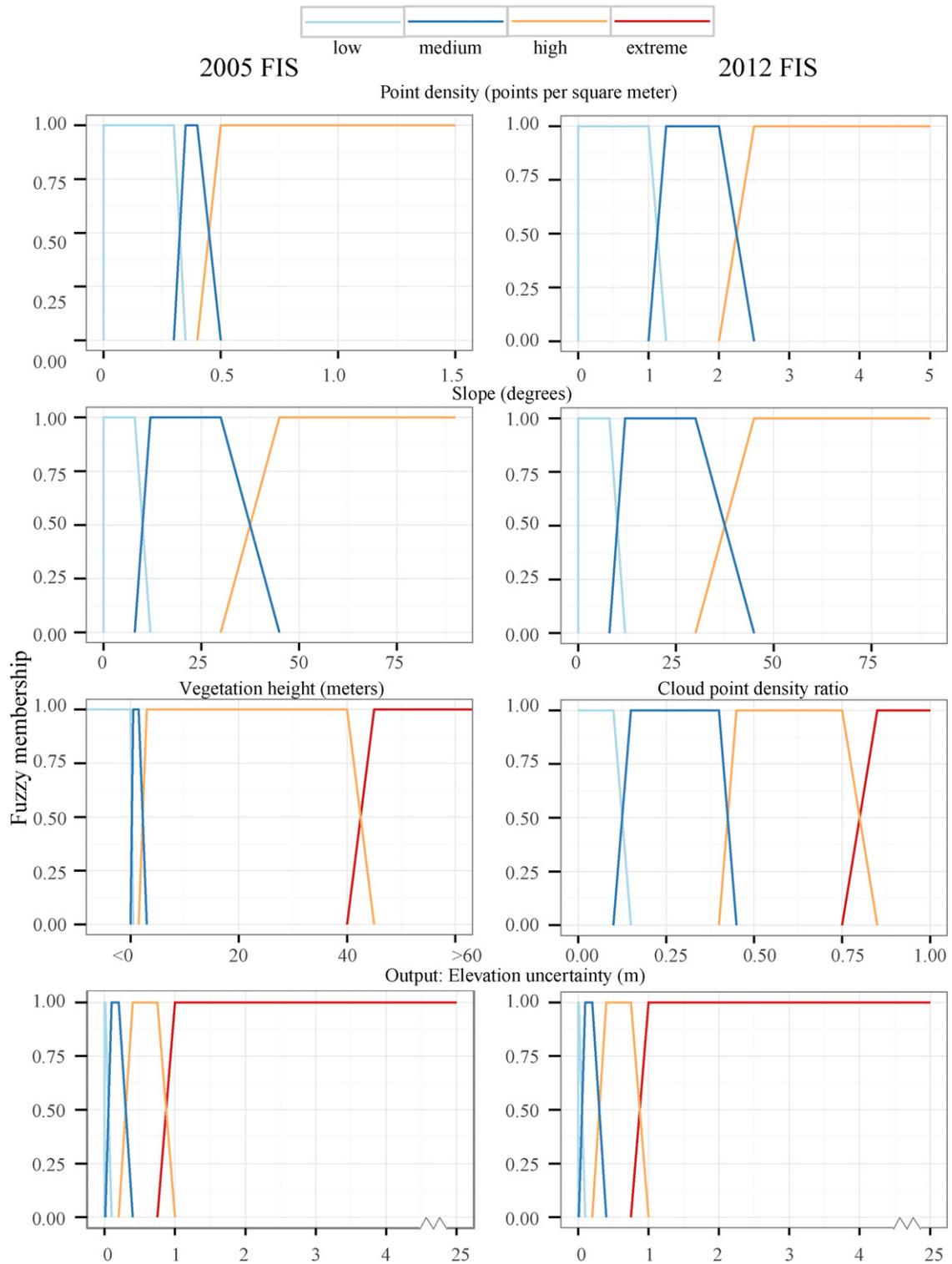


Fig. 6. Ranges associated with each membership function for the parameters used to develop the final Fuzzy inference system (FIS) used in the analysis.

Le Sueur River. From 2005–2012, peak flows exceeded or equaled these numbers 6 of the 7 years in both rivers. We built FIS models using the Matlab Fuzzy Logic Toolbox FIS file format (Jang and Gulley, 2009), which is also compatible with the Geomorphic Change Detection software (<http://gcd.joewheaton.org>) and posted our FIS files to the public BitBucket FIS-DEM-Error Repository (<https://bitbucket.org/pipbailey/fis-dem-error-repository>).

3.4. DoD error propagation and thresholding

Most DoD change detection analyses are based on principles of simple error propagation, and some form of thresholding to define a minimum level of detection ($_{min}LoD$) above which calculated changes can be assumed to be real and differentiated from noise (Passalacqua et al., 2015). The three most common techniques are to define a spatially

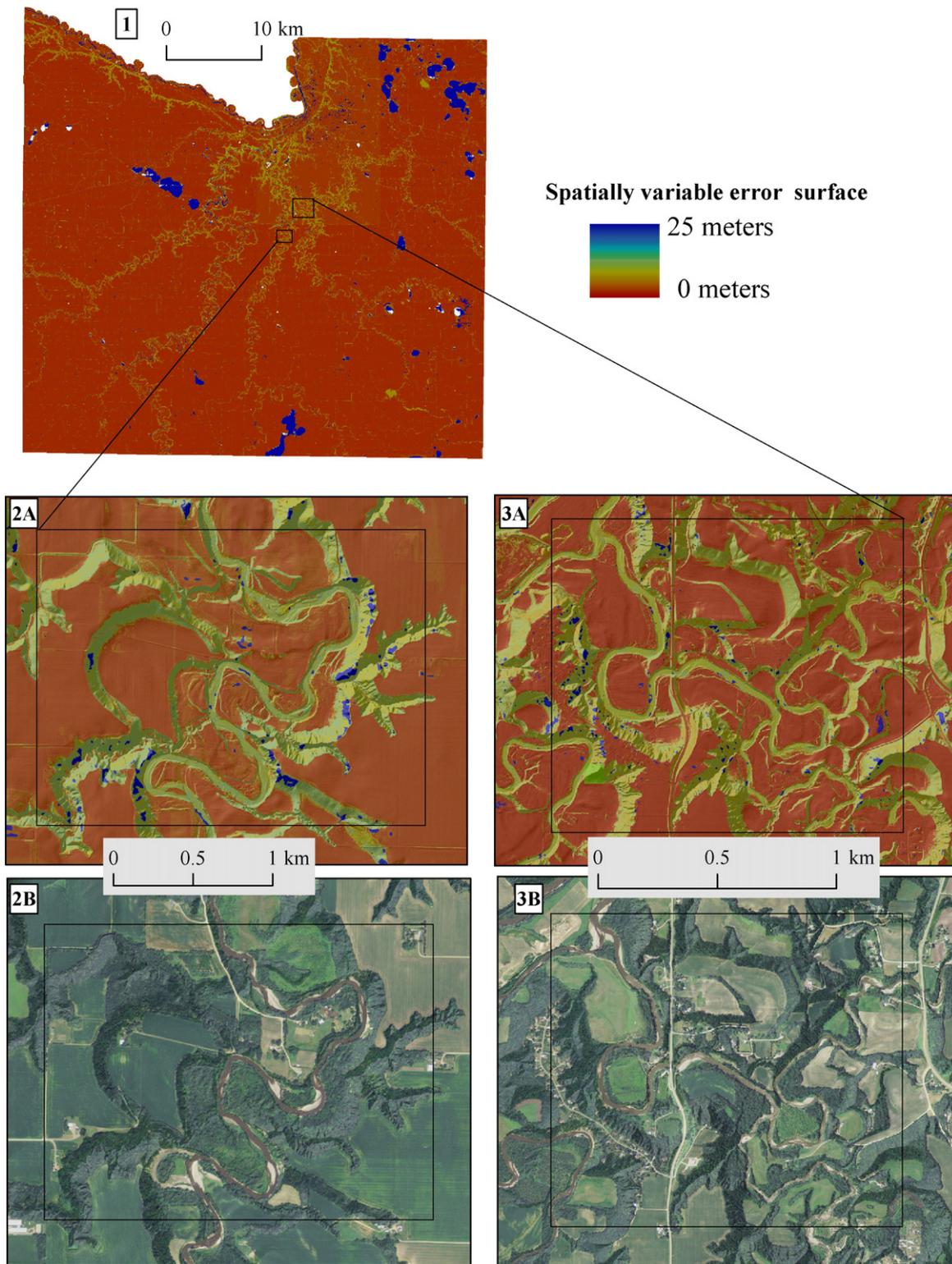


Fig. 7. Final surface of spatially variable error model for Blue Earth County (panel 1). Eq. (1) in the text describes how this surface was calculated. For additional detail, two additional panels are shown at finer scale and underlain by the 2012 hillshade model in areas also used to test the thresholding methods. Panels labeled 2 show the (A) spatially variable error model and (B) orthophoto for the area used to test different values of minimum level of detection. Panels labeled 3 are of the area used to test different fuzzy inference systems.

uniform minLoD as the threshold, use a propagated error value estimated from independent estimates of DEM error as the minLoD (notionally equivalent to 1σ), or threshold probabilistically at any chosen confidence interval (Brasington et al., 2003; Lane et al., 2003; Wheaton et al., 2010b; Bangen et al., 2014).

The process is automated within the GCD software, but from a geoprocessing perspective; a propagated error surface is derived by

taking the sum in quadrature (square root of the sum of squares) of an error surface for each individual DEM, where 'surface' refers to a raster product with a unique value for each pixel. Here, we were interested in comparing the impact of the spatially variable error (SVE) using an FIS and spatially uniform error models on change detection in a consistent fashion. Additionally, we chose to threshold probabilistically at a 95% confidence interval. In essence, probabilistic thresholding just

Table 2

Maximum values (minimum is 0 for each) and the range for each membership function for parameters in the three Fuzzy Inference Systems (FIS) developed; italics are used to indicate the changes between each FIS; the commit number is listed for each of the three FIS for both years (FIS are stored in the FIS DEM Repository at: <https://bitbucket.org/pipbailey/fis-dem-error-repository>).

Parameter	Maximum	Range of values for membership function				Repository record ^a	
		Low	Medium	High	Extreme	2005 FIS	2012 FIS
FIS 1							
2005 Point density, points per square meter	1.5	0.0–0.35	0.30–0.50	0.40–1.5	–		
2012 Point density, points per square meter	5	0.0–1.25	1.0–2.5	2.0–5.0	–		
Slope, degrees	90	0.0–15	10–25	20–90	–		
Vegetation height (2005) ^b , meters	600	0.0–0.50	0.0–3.0	1.5–45	>40	2047f9f	722110d
Point cloud density ratio (2012)	1	0.0–0.15	0.10–0.45	0.40–0.85	0.75–1.0		
Elevation error, meters	50	0.0–0.25	0.01–0.75	0.5–2	1.5–50		
FIS 2							
2005 Point density, points per square meter	1.5	0.0–0.35	0.30–0.50	0.40–1.5	–		
2012 Point density, points per square meter	5	0.0–1.25	1.0–2.5	2.0–5.0	–		
Slope, degrees	90	<i>0.0–12</i>	<i>8.0–45</i>	<i>30–90</i>	–		
Vegetation height (2005) ^b , meters	600	0.0–0.50	0.0–3.0	1.5–45	>40	71fb827	0379e2c
Point cloud density ratio (2012)	1	0.0–0.15	0.10–0.45	0.40–0.85	0.75–1.0		
Elevation error, meters	25	<i>0.0–0.10</i>	<i>0.01–0.40</i>	<i>0.20–1.0</i>	<i>0.75–25</i>		
FIS 3							
2005 Point density, points per square meter	1.5	0.0–0.35	0.30–0.50	0.40–1.5	–		
2012 Point density, points per square meter	5	0.0–1.25	1.0–2.5	2.0–5.0	–		
Slope, degrees	90	0.0–12	8.0–45	30–90	–		
Vegetation height (2005) ^b , meters	600	0.0–0.50	0.0–3.0	1.5–45	>40	953e684	44f4bb4
Point cloud density ratio (2012)	1	0.0–0.15	0.10–0.45	0.40–0.85	0.75–1.0		
Elevation error, meters	25	<i>0.0–0.20</i>	<i>0.05–0.50</i>	<i>0.30–1.0</i>	<i>0.75–25</i>		

^a Files are named ALS_Zerror_PD_Sldeg_VegHeight_WD.fis (2005) and ALS_Zerror_PD_Sldeg_VegDensity_WD.fis (2012).

^b Vegetation height was the difference between the first return and bare-earth surface. The large range of values is a reflection of the poor classification and processing of the first return surface. The data outside of the normal range of vegetation height (0–50 m) made up a small portion of the overall data set.

defines a new $\min\text{LoD}$ by multiplying the propagated error values (on a cell-by-cell basis) by a t -score. For example, a confidence interval of 95% is achieved by multiplying the entire propagated error surface by a t -value of 1.96 (roughly 2σ) to provide a threshold surface that allows us to detect differences at a 95% confidence interval:

$$\delta_{\text{DoD}} = 1.96 * \sqrt{(\delta z_{\text{DEM}12})^2 + (\delta z_{\text{DEM}05})^2} \quad (1)$$

where δ_{DoD} can act as the $\min\text{LoD}$ threshold and is the propagated error value, which corresponds with a $\geq 95\%$ probability (i.e., 2σ , assuming a two-tailed distribution) of being real change, and the δz_{DEM} refers to the individual DEM error estimates on a cell basis.

For the spatially variable FIS-based error estimates, the propagated error term was calculated by taking the individual δz_{DEM} on a cell-by-cell basis, and δ_{DoD} was then calculated using Eq. (1). The final surface is shown in Fig. 7: panel 1 shows the entire county and panels 2A and 3A show smaller areas in more detail. For spatially uniform estimates, we used the RMSE values from each year as the δz_{DEM} . The sum in quadrature of the two RMSE values is ~ 20 cm (roughly 1σ), which is equivalent to 39 cm when multiplied by 1.96. We also evaluated results using spatially uniform $\min\text{LoDs}$ of 29 and 49 cm. The three values are roughly equivalent to an 85%, 95%, and 98.5% confidence interval around the original 20-cm estimate. We compare the results of these three $\min\text{LoDs}$ on a smaller area within the county to show the effects of changing the value (Fig. 7, panels A and B). We then compare results from the 39-cm $\min\text{LoD}$ (20 cm with a 95% confidence interval, 2σ) threshold to spatially variable error (SVE) threshold (also 2σ).

Visual inspection is a useful first step to evaluate the efficacy of the various techniques to account for uncertainty. Deposition on the insides of meander bends and erosion on the outside of every bend is consistent with expectations, given the magnitude of the flood, and consistent with field observations. There is also a large cutoff in the southern part of this area, indicated by deposition in the old channel and severe erosion cut across what was previously a floodplain. Fig. 8 includes six panels showing (from top to bottom) the DoDc, SVE-DoDc (black color used to indicate where we detected change), and a bar graph

(left column Fig. 8A–C) and results from three $\min\text{LoDs}$ (29, 39, and 49 cm) on the right column (Fig. 8D–F). Bright green colors on Fig. 8D–F indicate areas where the $\min\text{LoD}$ threshold detected change while the SVE threshold did not. Orange indicates areas where the SVE threshold detected change but the $\min\text{LoD}$ threshold did not. Black areas are where both analyses detected change. With a 29-cm threshold, the $\min\text{LoD}$ approach greatly overestimates erosion on steeper, vegetated topography away from the channel. With a 49-cm threshold, the $\min\text{LoD}$ approach is missing some real change that occurs near the channel, while still overestimating change away from the channel.

Generally, the bar graph (Fig. 8C) and panels on the right side (Fig. 8D–F) indicate that as the uniform threshold value for error increases, less change is mapped. The plots also show that for all three uniform thresholds, more change is detected relative to the spatially variable threshold. Note that the primary difference in the areas where change was detected is within ravines in each of the $\min\text{LoD}$ panels. Fig. 8D–F show each of the three $\min\text{LoD}$ values detected change in the ravines not detected by the SVE model (Fig. 8B), likely because of the high vegetation density and higher slopes in the ravines. Where the SVE model detects change in ravines it is in the channel area of those ravines, which is consistent with field observations, rather than on the slopes, as seen in the $\min\text{LoD}$ models (Fig. 8D–F). Application of the 39- and 49-cm $\min\text{LoD}$ results in a failure to detect change along the banks (circled in Fig. 8E and clearly visible in Fig. 8F) detected by the SVE model. The uniform application of any $\min\text{LoD}$ over the entire area of analysis is likely to mask real change (from the banks in this example) while also detecting change that is not real (from the ravines in this example).

3.5. Final DoD processing and reporting

We created an analysis area of interest to focus on the portion of the data set tied to potential geomorphic changes to the fluvial system. To do this, we buffered the combined hydrolines by 200 m on either side and then manually edited that file to also include all areas that appeared in the 2009 orthophotos or in either DEM to be connected to the main

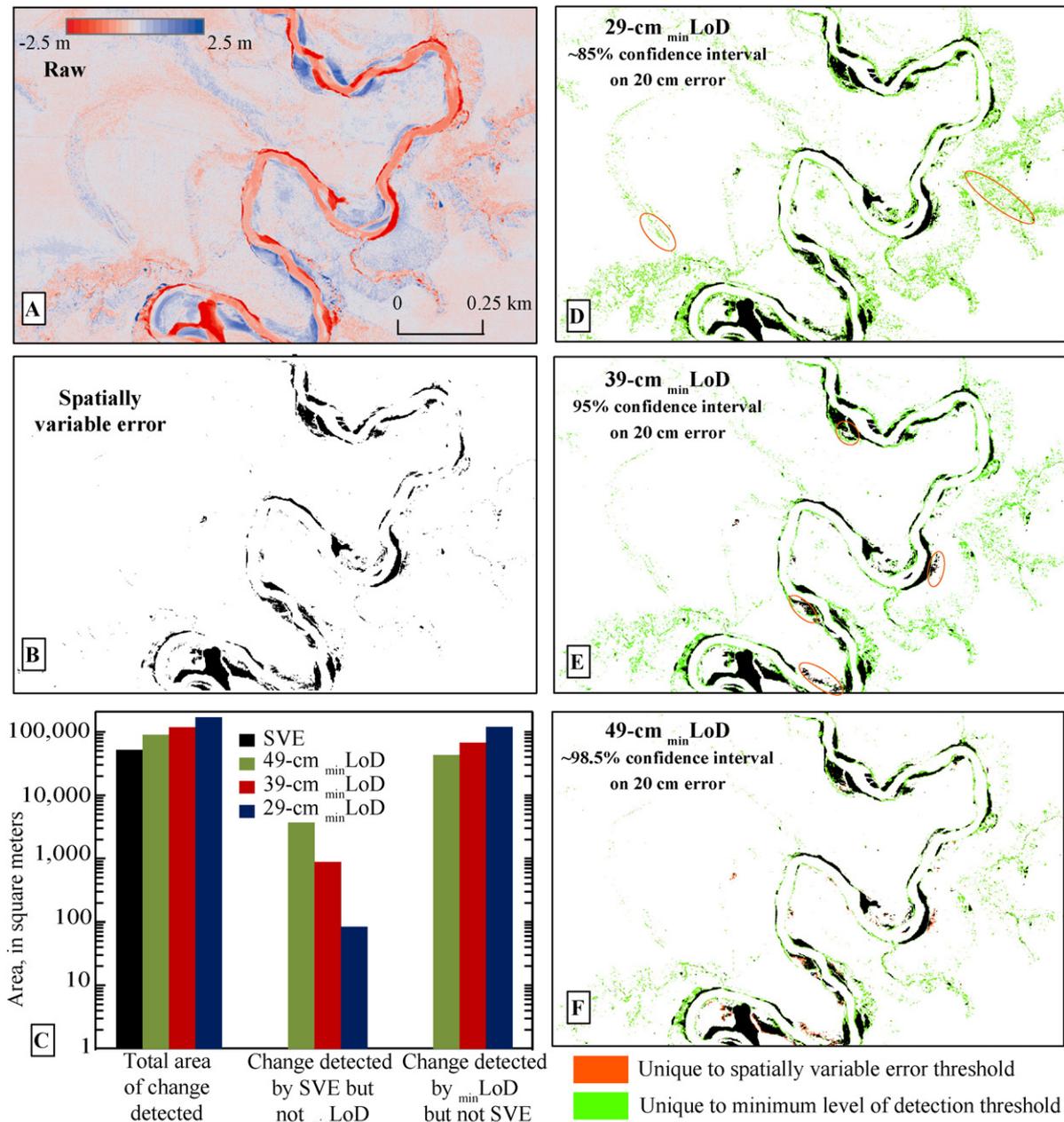


Fig. 8. Comparison of three error values for a minimum level of detection ($_{\min}\text{LoD}$) (29, 39, and 49 cm, panels D, E, and F, respectively) to the spatially variable error (SVE) model (panel B). Bright green colors in panels D–F show areas where the $_{\min}\text{LoD}$ detected change but the SVE did not. Orange colors show areas where the SVE threshold detected change but the $_{\min}\text{LoD}$ did not. For the two smaller values for error (panels D and E), orange circles were drawn to more clearly show the results as the area is small. Raw, stripping-corrected DEM of Difference results are in panel A and are provided for comparison. Panel C is a bar graph showing the results in common relative to unique results to each analysis. y -axis is on a log scale.

channel. Inundated areas in either year were then removed. From that area, we identified and eliminated anthropogenic features (e.g., gravel pits, landscaping features, and road maintenance and construction) that resulted in large amounts of change that do not contribute to the sediment budget related to the efflux at the watershed mouth. The full extent of the lidar coverage is 2590 km², but all change results reported are from within this 610 km² final analysis area (maroon outline in Fig. 1C).

Bluffs, ravines, and channel migration and widening have previously been identified as major contributing sediment sources in the Le Sueur River (Belmont et al., 2011). Bluffs were delineated by Day et al. (2013a) and ravines were mapped by Gran et al. (2011). There are 275 ravines delineated in the county, 76% of which are located within the knickzone. To quantify channel change within the limits of airborne lidar

(incapable of penetrating water), we first created a 100-m buffer on either side of the combined hydrolines file (2005 and 2012 as one file). From that file, areas delineated as bluffs or ravines were removed. We refer to the remaining area as 'banks and floodplain'.

Estimates of erosion and deposition from areas that were dry in both years are the most accurate, and unless otherwise specified, we report values from areas that were dry in both years. Estimates of change in areas inundated in only one year have inherent error related to the false-water surface elevation from the wet year, which is not real geomorphic change and is therefore eliminated from our sediment budget numbers. Analysis of stream flow and stage during each flight indicates that the water surface was ~1 m higher during the 2005 flight. Thus, comparisons of areas that were inundated in both years are erroneous and were eliminated from the analysis. Differences in areas that were

inundated in one year will be biased, but still provide coarse estimates of change. Areas that were dry in 2012 and inundated in 2005 are influenced by the high flows in 2005 and therefore overestimate erosion and underestimate deposition. Areas that were inundated in 2012 and dry in 2005 are almost exclusively places where channel erosion occurred between the flights. Yet, estimates in these locations slightly underestimate erosion because the false-water surface in 2012 obscures the channel bottom. The two hydroline files were combined to generate one shapefile with areas of inundation in both years, dry in both years, inundated in 2012 and dry in 2005, or dry in 2012 and inundated in 2005. We use this combined hydrolines shapefile to report coarse estimates of change in areas that were inundated in either 2005 or 2012 and as an input parameter in the FIS. Because inundation was an input parameter in the FIS, results from the spatially variable threshold only include a small portion of change from inundated areas (see Fig. 8) so erosion and deposition estimates for areas that were inundated in one year were generated using the results from the spatially uniform threshold of 39-cm (95% confidence interval on 20-cm).

4. Geomorphic results and interpretation

As we increase our capability to collect repeat HRT data on larger spatial scales, so increases the challenge of quantifying and reporting error in a robust and meaningful way. Our results demonstrate that generic error values reported by vendors may underestimate actual uncertainty and do not capture the spatially variable nature of uncertainty that is needed for meaningful geomorphic change detection. This problem can be lessened if approaches that are more rigorous are required of vendors, as discussed extensively by Passalacqua et al. (2015).

We first compare the results of an unthresholded DoD (essentially, the raw, geoid-corrected data) to a thresholded DoD using a minLoD of 39-cm, which we refer to as Min-DoD. We also corrected for a stripping problem detected in the DoD and applied the 39-cm minLoD (Min-DoDc) and an SVE model (SVE-DoDc). All results are reported with a 95% confidence interval in Table 1. For the remainder of this document, negative numbers, indicating erosion, are presented in brackets.

The DoD indicates over [33,164,800] m^3 of erosion and 28,104,200 m^3 of deposition from 2005 to 2012 (Table 1). This simple calculation would suggest net erosion of [5,060,600 m^3]. Simply applying the spatially uniform 39-cm minLoD (Min-DoD), erosion decreased by 75% to [8,092,600] \pm 3,300,000 m^3 and deposition decreased by 85% to 4,459,200 \pm 2,714,900 m^3 , indicating a net [3,633,400] \pm 4,273,300 m^3 of erosion (error includes zero, indicating results are not significant). Note a direct subtraction of the two DEMs may approximate the results from a minLoD analysis, but the overall erosion and deposition results are grossly overestimated. Also, as mentioned in Section 3.2 (Horizontal uncertainty between DEMs), evidence of stripping is clearly visible (Fig. 4). Applying the stripping correction reduced estimates of erosion and deposition in the DoDc and Min-DoDc. Erosion estimates from the DoDc were reduced by about 25% to [24,281,300] m^3 while deposition estimates were reduced by close to 15% to 23,535,800 m^3 . The slight difference in reduction brought estimates of erosion and deposition closer together and reduced the net change estimate substantially to just [745,600] m^3 of erosion. Relative to the Min-DoD, applying the spatially uniform 39-cm minLoD (Min-DoDc), estimates of erosion, deposition and net change all decreased by around 10% to [7,418,300] \pm 2,823,600 m^3 , 4,099,000 \pm 2,462,400 m^3 , and [3,319,300] \pm 3,746,500 m^3 , respectively (Table 1). The similar decrease indicates the stripping correction more or less equally affected erosion and deposition when we thresholded the results. However, with a spatially uniform threshold, the error includes zero, and we are unable to say there was significant change with 95% confidence.

Consideration of error is important in any study of geomorphic change, but especially for an analysis on such a large scale. As the area analyzed and heterogeneity of the topography increase, the potential for incorporating erroneous estimates of erosion or deposition (or

missing real change) also increases. By spatially varying error, we address error separately for each square meter pixel over the entire area analyzed. The result (Table 1, Fig. 9) indicates erosion along the banks and deposition along the insides of bends. Final estimates for the entire area analyzed were [4,153,300] \pm 2,176,200 m^3 of erosion and 1,528,300 \pm 985,700 m^3 of deposition. Net change in the dry areas analyzed was [2,625,100] \pm 2,389,000 m^3 of erosion for the 7-year study period. In this case, the error on the net change does not include zero, and we can say with 95% confidence that significant change has occurred.

Visually, results of the SVE-DoDc indicate erosion and deposition in areas we might expect (Fig. 9). Fig. 9A shows meander migration and cutoffs, with erosion on the outside of bends and deposition on the inside bends, as expected and confirmed on air photos. Fig. 9B shows multiple examples in one area where we consistently observe erosion on the outside of bends and deposition on the inside of bends. Fig. 9C shows ravine erosion, where the upstream end of the ravine to the west is eroding headward, which was confirmed in air photos and field observation. The ravine to the north also shows headward erosion and some deposition in the channel. The apparent, but dubious, erosion and deposition patterns seen on the slopes of ravines in the Min-DoDs have been eliminated in the SVE-DoDc (Figs. 8 and 9), primarily from high uncertainty associated with high slope and dense vegetation.

Spatially varying error would have addressed the majority of the stripping problem (data not shown), but actually applying the correction for stripping error before applying the SVE model provides even better results that exclude the obvious overestimates of erosion and deposition from the strips. Including erosion and deposition known to be erroneous artificially inflates estimates of geomorphic change, which is especially problematic over large areas. Also by spatially propagating error, results are in line with our expectation; namely, erosion and deposition occurred within a small buffer of the main channels and little to no erosion occurred in the upland areas (Fig. 9). Belmont et al. (2011) showed that only 2.9 mm of vertical erosion would occur throughout the upland areas of the Le Sueur watershed as a result of 100,000 Mg y^{-1} of soil loss (well above the modern upland erosion rate) for 150 years. So no widespread detectable change is expected in our analysis.

4.1. Spatial patterns of change

In the previous section, we compared DoDs and provided estimates of erosion, deposition, and net change for the analysis extent. In this section, we provide estimates of erosion, deposition, and net change for areas potentially targeted by land managers trying to address the sources of eroded material. Near-channel sediment sources such as banks and bluffs have been shown to be the main source of sediment in the Le Sueur River basin (Belmont et al., 2011). Using the SVE-DoDc, we calculated the amount of erosion and deposition from bluffs, ravines, and floodplains and banks inside and outside of the knickzone (Table 3).

We present the results of erosion in bluffs, ravines, and the bank and floodplain area for the full county and segregated by location inside or outside the knickzone. In areas delineated as bluffs in the entire county, there was net erosion of [1,032,400] m^3 . Net erosion from ravines was [31,900 m^3]. The greatest erosion and deposition occurred in the area delineated as banks and floodplains (net change of [2,585,100 m^3]). Kessler et al. (2012) differenced lidar from 2005 and 2009 and estimated [318,750 $\text{m}^3 \text{y}^{-1}$] of erosion from bluffs (referred to as 'tall banks' in their analysis). Our analysis, which includes another three years and the impacts of the major 2010 flood, indicates [147,500] $\text{m}^3 \text{y}^{-1}$ of erosion from bluffs. We speculate that Kessler et al. (2012) may have overestimated bluff erosion because they did not address systematic biases caused by differences in geoid models or poor coregistration of the 2005 flight lines (i.e., the stripping error). While Kessler et al. (2012) did not report the geoid models used, they did detect a 6-cm

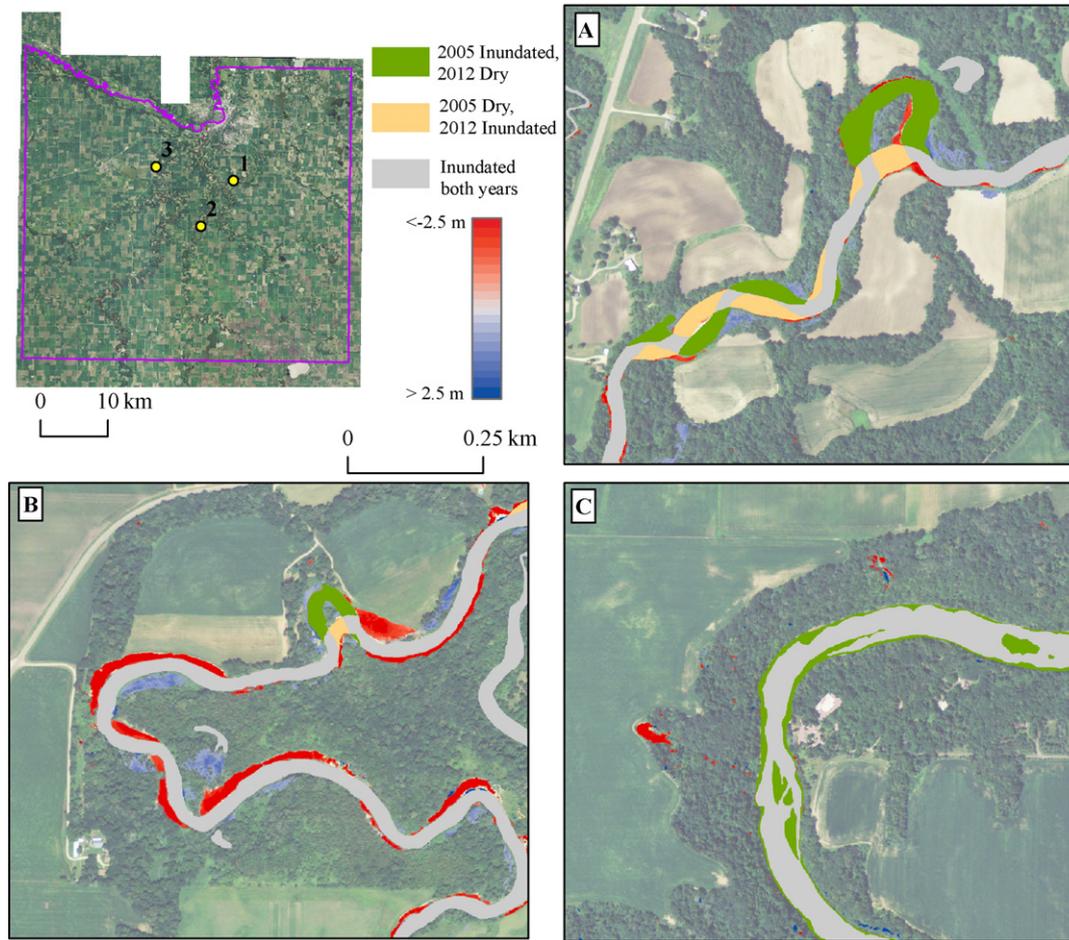


Fig. 9. DEM of Difference results after applying the spatially variable error model (SVE-DoDc) underlain by the aerial orthophoto from 2013 (National Agriculture Imagery Program). Panels are associated with the locations in the map on the top right and described further in text. (A) shows an area where meander migration and a cutoff occurred; (B) shows an example of deep erosion on the outside of bends and shallow deposition on the inside of bends; and (C) shows an area of headcut erosion in the tops of ravines.

bias, as observed (and rectified) in our study. The discrepancies highlight the importance of independently assessing the errors between different vendors and survey techniques that may be contributing systematic errors to a change detection analysis.

The area referred to as the actively incising knickzone is 260 km² (see area outlined in blue in Fig. 1), which is 43% of the area analyzed, and contributed 55% of eroded material. Net change within the knickzone was [1,419,300] m³, including [2,171,600] m³ of erosion and 752,300 m³ of deposition (Table 4). Net erosion was observed specifically in bluffs, banks/floodplains, and ravines. Bluffs contribute 58% ([827,200] m³) to the net eroded amount, while ravines contribute

Table 3

Erosion, deposition, and net geomorphic change results for the entire Blue Earth County using three methods of error analysis: none, minimum level of detection, and spatially variable error; results from minimum level of detection and spatially variable error models are probabilistic (95% confidence interval).

	Erosion, in cubic meters	Deposition, in cubic meters	Net change, in cubic meters
Raw (DoD)	[33,164,800]	28,104,200	[5,060,600]
Minimum level of detection (Min-DoD)	[8,092,600] ± 3,300,000	4,459,200 ± 2,714,900	[3,633,400] ± 4,273,300
Corrected for stripping			
Raw (DoDc)	[24,281,300]	23,535,800	[745,600]
Minimum level of detection (Min-DoDc)	[7,418,300] ± 2,823,600	4,099,000 ± 2,462,400	[3,319,300] ± 3,746,500
Spatially variable error (SVE-DoDc)	[4,153,300] ± 2,176,200	1,528,300 ± 985,700	[2,625,100] ± 2,389,000

only 3% ([46,100] m³). Net erosion in the bank and floodplain area inside the knick was [1,441,500] m³.

Outside the knickzone, our analysis indicates [1,981,700] m³ of erosion and just 776,000 m³ of deposition, yielding net erosion of [1,205,700] m³. Most of this change (81%) was in the bank and floodplain area ([1,511,500] m³ erosion, 367,900 m³ deposition, net erosion: [1,143,600] m³). In comparison, bluffs only contributed [215,000 m³] with just 9800 m³ of deposition, meaning net change of [205,200] m³. Outside the knickzone, change in ravines was positive, 14,200 m³ of net deposition.

Deposition might be difficult to accurately quantify for the entire study area using any error model, because in many areas, deposition occurs in vertically small quantities. Inside the knickzone, the channel is wider, bar deposits are relatively larger and actively changing, and riparian vegetation tends to be farther outside of the channel area on top of bluffs. Vegetation inside the floodplain is generally less dense. Outside the knickzone, channels are still relatively narrow, bar deposits are much smaller, and floodplains are also covered in dense vegetation. Spatially uniform error models are insensitive to differences in vegetation density and are therefore more prone to miss real change in less dense vegetation and incorporate erroneous change in dense vegetation. The SVE model is less likely to result in such errors because uncertainty increases with vegetation density.

5. Summary

This study describes essential practices for supporting lidar-based change detection, in which surveying and post-processing

Table 4

Erosion, deposition, and net change results segregated by area in the county; upland areas are not included in this list so the sum of bluffs, ravines, and banks will not be equal to the totals.

	Erosion area, square meters	Deposition area, square meters	Erosion, cubic meters	Deposition, cubic meters	Net change, cubic meters
Dry both years	2,570,200	1,793,200	[4,153,300]	1,528,300	[2,625,100]
Dry in 2005, inundated in 2012	734,100	11,000	[2,055,200]	7100	[2,048,200]
Inundated in 2005, dry in 2012	3,357,400	674,100	[4,105,200]	577,300	[3,527,900]
Bluffs	354,800	29,100	[1,110,500]	78,100	[1,032,400]
Ravines	70,700	43,600	[127,200]	95,300	[31,900]
Banks and floodplains	1,638,200	900,900	[3,347,900]	762,800	[2,585,100]
Inside knickzone					
Bluffs	282,000	25,100	[895,500]	68,300	[827,200]
Ravines	56,400	28,900	[100,100]	54,000	[46,100]
Banks and floodplains	723,300	445,500	[1,836,400]	394,900	[1,441,500]
Total, inside knickzone	1,010,800	814,100	[2,171,600]	752,300	[1,419,300]
Outside knickzone					
Bluffs	72,700	4000	[215,000]	9800	[205,200]
Ravines	14,300	14,700	[27,100]	41,300	14,200
Banks and floodplains	915,000	455,400	[1,511,500]	367,900	[1,143,600]
Total, outside knickzone	1,559,400	979,100	[1,981,700]	776,000	[1,205,700]

discrepancies may introduce significant, systematic bias. The methods and conceptual approach also have implications for other change detection studies using different topographic survey acquisition techniques. Identification, quantification, isolation, and correction of systematic errors is essential. We specifically addressed vertical bias by correcting the geoid model. The geoid model is not always reported in the metadata, but in our case it caused a systematic (and spatially variable) bias of 3–7 cm. We also detected erroneous east–west strips of erosion or deposition caused by poor coregistration of 2005 flight lines. This problem was resolved by generating an ad hoc stripping-correction surface. Our comparisons show the consequences of ignoring potential issues that may arise with older data collection methods. For the 2012 data, we developed a metric, cloud point density, to quantify uncertainty owing to vegetation that uses the entire point cloud as it was collected and classified. As a result, this metric accounts for multiple layers of vegetation and other potential barriers to successfully obtaining an accurate estimate of the bare-earth elevation in the square meter represented by each pixel. Where there is dense vegetation, the cloud point density metric is close to 1 indicating that the above-ground point density is much higher than the bare-earth point density. We suggest this is a better metric than canopy height to characterize uncertainty associated with vegetation because cloud point density captures various levels of canopy and understory found in a vegetated area while canopy height only represents the top layer of vegetation (canopy) found in an area.

Results from this analysis illustrate the value of considering the spatial variability of error. We quantified uncertainty in a spatially variable manner, accounting for heterogeneity in point density, slope, and vegetation height or density for each pixel analyzed. In our comparison of spatially variable versus spatially uniform DEM error models on a subsection of the county, the uniform models produce inaccurate estimates of erosion and deposition. Results from the spatially variable error model indicated significant net erosion, while significant net change (with 95% confidence) was not detected using uniform error models. The spatial distribution of erosion and deposition patterns resulting from the spatially variable error model were consistent with field observations and can be physically explained.

Using the geoid- and stripping-corrected DoD, we compared three methods of thresholding DoD results for the full 1980 km² of the county: no threshold, spatially uniform error (minLoD), and SVE. Estimates of erosion and deposition from the analysis with no threshold were an order of magnitude larger than estimates from the other two thresholds, while net change was grossly underestimated. We compared three values for a minLoD error model uniformly applied for the entire area of analysis. We showed that smaller values for error detected erroneous change, such as erosion on the slopes of the majority of ravines in the

study area. As the uniform error value increased, change on the slopes of ravines was still erroneously detected; but real change along the main channel was inappropriately eliminated. When we thresholded the results using an SVE model, detected change agreed with field observations. Additionally, countywide estimates of net change from the analysis with SVE were within 10% of estimates from the statewide network of stream gages, and results from bluffs and ravines closely match estimates from a separate analysis for the Le Sueur watershed. We assert that the accuracy of results of change detection increases when data are treated rigorously in terms of correcting legacy data issues (geoid model and stripping problem, in our example) and by applying a probabilistic, spatially variable error model.

Acknowledgments

This work was funded by National Science Foundation (NSF ENG1209445) and the Utah State University Agricultural Experiment Station (paper #8839). The authors would like to thank Blue Earth County, the Minnesota State University Water Resources Center, and the State of Minnesota for collecting and sharing lidar and water quality data. Chris Garrard, in the Remote Sensing and GIS lab at Utah State University, provided invaluable guidance and support in the data processing stages. Philip Bailey, North Arrow Research, and James Hensleigh, Utah State University, were also enthusiastically available to help with the GCD tool. Nancy Glenn, Boise State University, and Chris Crosby, OpenTopography.org and UNAVCO, helped us identify the stripping problem and geoid offset, respectively, and continued to provide guidance in addressing those issues in our data. We also thank Steve DeLong and two anonymous reviewers for their constructive comments, which greatly improved the clarity of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.geomorph.2015.09.020>.

References

- Abellán, A., Oppikofer, T., Jaboyedoff, M., Rosser, N.J., Lim, M., Lato, M.J., 2014. Terrestrial laser scanning of rock slope instabilities. *Earth Surface Processes and Landforms* 39, 80–97.
- Akca, D., 2010. Co-registration of surfaces by 3D least squares matching. *Photogrammetric Engineering & Remote Sensing* 76, 307–318.
- Ashmore, P.E., Church, M.A., 1998. Sediment transport and river morphology: a paradigm for study. In: Klingeman, P., Beschta, R.L., Komar, P., Bradley, J. (Eds.), *Gravel Bed Rivers in the Environment*. Water Resources Publications, Oregon, pp. 115–139.

- Bangen, S.G., Wheaton, J.M., Bouwes, N., Bouwes, B., Jordan, C., 2014. A methodological inter-comparison of topographic survey techniques for characterizing wadeable streams and rivers. *Geomorphology* 206, 343–361.
- Belmont, P., 2011. Floodplain width adjustments in response to rapid base level fall and knickpoint migration. *Geomorphology* 128, 92–102.
- Belmont, P., et al., 2011. Large shift in source of fine sediment in the Upper Mississippi River. *Environ. Sci. Technol.* 45, 8804–8810.
- Besl, P.J., McKay, H.D., 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, 239–256.
- Blue Earth County, 2012. Lidar elevation. Blue Earth County, Minnesota 2012.
- Brasington, J., Rumsby, B.T., McVey, R.A., 2000. Monitoring and modelling morphological change in a braided gravel-bed river using high resolution GPS-based survey. *Earth Surface Processes and Landforms* 25, 973–990.
- Brasington, J., Langham, J., Rumsby, B., 2003. Methodological sensitivity of morphometric estimates of coarse fluvial sediment transport. *Geomorphology* 53, 299–316.
- Bretar, F., Pierrrot-Desilligny, M., Roux, M., 2004. Solving the strip adjustment problem of 3D airborne lidar data. *IEEE International Geoscience and Remote Sensing Symposium* 7, 4734–4737.
- Collins, A.L., Walling, D.E., 2004. Documenting catchment suspended sediment sources: problems, approaches and prospects. *Progress in Physical Geography* 28, 159–196.
- Day, S.S., Gran, K.B., Belmont, P., Wawrzyniec, T.F., 2013a. Measuring bluff erosion part 1: terrestrial laser scanning methods for change detection. *Earth Surface Processes and Landforms* 38, 1055–1067.
- Day, S.S., Gran, K.B., Belmont, P., Wawrzyniec, T.F., 2013b. Measuring bluff erosion part 2: pairing aerial photographs and terrestrial laser scanning to create a watershed scale sediment budget. *Earth Surface Processes and Landforms* 38, 1068–1082.
- DeLong, S.B., Prentice, C.S., Hilley, G.E., Ebert, Y., 2012. Multitemporal ALSM change detection, sediment delivery, and process mapping at an active earthflow. *Earth Surface Processes and Landforms* 37, 262–272.
- Fisher, P.F., Tate, N.J., 2006. Causes and consequences of error in digital elevation models. *Progress in Physical Geography* 30, 467–489.
- Glennie, C.L., Hinojosa-Corona, A., Nissen, E., Kusari, A., Oskin, M.E., Arrowsmith, J.R., Borsari, A., 2014. Optimization of legacy lidar data sets for measuring near-field earthquake displacements. *Geophysical Research Letters* 41, 3494–3501.
- Goodwell, A.E., et al., 2014. Assessment of floodplain vulnerability during extreme Mississippi River flood 2011. *Environmental science & technology* 48, 2619–2625.
- Gran, K.B., Belmont, P., Day, S.S., Jennings, C.E., Lauer, J.W., Viparelli, E., Wilcock, P.R., Parker, G., 2011. An Integrated Sediment Budget for the Le Sueur River Basin.
- Gran, K.B., Finnegan, N., Johnson, A.L., Belmont, P., Wittkop, C., Rittenour, T., 2013. Landscape evolution, valley excavation, and terrace development following abrupt post-glacial base-level fall. *Geological Society of America Bulletin* 125, 1851–1864.
- Habib, A.F., Kersting, A.P., Ruifang, Z., Al-Durgham, M., Kim, C., Lee, D.C., 2008. Lidar strip adjustment using conjugate linear features in overlapping strips. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 37, 385–390.
- Heritage, G.L., Milan, D.J., Large, A.R.G., Fuller, I.C., 2009. Influence of survey strategy and interpolation model on DEM quality. *Geomorphology* 112, 334–344.
- Hodgson, M.E., Jensen, J.R., Schmidt, L., Schill, S., Davis, B., 2003. An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. *Remote Sensing of Environment* 84, 295–308.
- Hodgson, M.E., Bresnahan, P., 2004. Accuracy of airborne lidar-derived elevation: empirical assessment and error budget. *Photogrammetric Engineering & Remote Sensing* 70, 331–339.
- Hodgson, M.E., Jensen, J.R., Raber, G., Tullis, J., Davis, B.A., Thompson, G., Schuckman, K., 2005. An evaluation of lidar-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric Engineering & Remote Sensing* 71, 817–823.
- Hopkinson, C., Chasmer, L.E., Sass, G., Creed, I.F., Sitar, M., Kalbfleisch, W., Treitz, P., 2005. Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment. *Canadian Journal of Remote Sensing* 31, 191–206.
- James, L.A., Hodgson, M.E., Ghoshal, S., Latiolais, M.M., 2012. Geomorphic change detection using historic maps and DEM differencing: the temporal dimension of geospatial analysis. *Geomorphology* 137, 181–198.
- Jang, J.S.R., Gulley, N., 2009. Fuzzy Logic Toolbox 2: User Guide. Matlab, Matlab, Natick, MA.
- Jennings, C.E., 2010. Digital Reconnaissance Surficial Geology and Geomorphology of the Le Sueur River Watershed.
- Kessler, A.C., Gupta, S.C., Dolliver, H.A.S., Thoma, D.P., 2012. Lidar quantification of bank erosion in Blue Earth County, Minnesota. *Journal of Environment Quality* 41, 197–207.
- Kinzel, P.J., Wright, C.W., Nelson, J.M., Burman, A.R., 2007. Evaluation of an experimental lidar for surveying a shallow, braided, sand-bedded river. *Journal of Hydraulic Engineering* 133, 838–842.
- Klepeis, N.E., 2004. Exposure Science Website: Miscellaneous Plotting and Statistics.
- Lane, S.N., Richards, K.S., Chandler, J.H., 1993. Developments in photogrammetry; the geomorphological potential. *Progress in Physical Geography* 17, 306–328.
- Lane, S.N., Westaway, R.M., Hicks, D.M., 2003. Estimation of erosion and deposition volumes in a large, gravel-bed, braided river using synoptic remote sensing. *Earth Surface Processes and Landforms* 28, 249–271.
- Leyland, J., Darby, S.E., Teruggi, L., Rinaldi, M., Ostuni, D., 2015. A Self-limiting Bank Erosion Mechanism? Inferring Temporal Variations in Bank Form and Skin Drag From High Resolution Topographic Data. *Earth Surface Processes and Landforms*.
- Milan, D.J., Heritage, G.L., Large, A.R.G., Fuller, I.C., 2011. Filtering spatial error from DEMs: Implications for morphological change estimation. *Geomorphology* 125, 160–171.
- Mueller, E.R., Pitlick, J., 2013. Sediment supply and channel morphology in mountain river systems: 1. Relative importance of lithology, topography and climate. *Journal of Geophysical Research: Earth Surface* 118, 2325–2342.
- Optimal Geomatics, 2005a. Data Processing Procedures Report. Huntsville, AL.
- Optimal Geomatics, 2005b. Map Accuracy Report: Countywide Imagery and DEM. Blue Earth County, Huntsville AL.
- Passalacqua, P., et al., 2015. Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: a review. *Earth-Science Reviews* 148, 174–193.
- Reutebuch, S.E., McGaughey, R.J., Andersen, H.-E., Carson, W.W., 2014. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing* 29, 527–535.
- Rumsby, B.T., Brasington, J., Langham, J.A., McLelland, S.J., Middleton, R., Rollinson, G., 2008. Monitoring and modeling particle and reach-scale morphological change in gravel-bed rivers: applications and challenges. *Geomorphology* 93, 40–54.
- Runkel, A.C., Meyer, G.N., Lusardi, B.A., 2011. C-26 Geologic Atlas of Blue Earth County, Minnesota [Part A].
- Schottler, S.P., Ulrich, J., Belmont, P., Moore, R., Lauer, J.W., Engstrom, D.R., Almendinger, J.E., 2014. Twentieth century agricultural drainage creates more erosive rivers. *Hydrological Processes* 28, 1951–1961.
- Schwendel, A.C., Fuller, I.C., Death, R.G., 2012. Assessing DEM interpolation methods for effective representation of upland stream morphology for rapid appraisal of bed stability. *River Research and Applications* 28, 567–584.
- Smith, S.M.C., Belmont, P., Wilcock, P.R., 2011. Closing the gap between watershed modeling, sediment budgeting, and stream restoration. *Geophysical Monograph Series* 194, 293–317.
- Streutker, D.R., Glenn, N.F., Shrestha, R., 2011. A slope-based method for matching elevation surfaces. *Photogrammetric Engineering & Remote Sensing* 77, 743–750.
- Su, J., Bork, E., 2006. Influence of vegetation, slope, and lidar sampling angle on DEM accuracy. *Photogrammetric Engineering & Remote Sensing* 72, 1265–1274.
- U.S. Geological Survey National Geospatial Program, 2010. Lidar guidelines and base specification, version 13-ILMF 2010.
- Wheaton, J.M., Brasington, J., Darby, S.E., Sear, D.A., 2010a. Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. *Earth Surface Processes and Landforms* 35, 136–156.
- Wheaton, J.M., Brasington, J., Darby, S.E., Merz, J., Pasternack, G.B., Sear, D.A., Vericat, D., 2010b. Linking geomorphic changes to salmonid habitat at a scale relevant to fish. *River Research and Applications* 26, 469–486.
- Wheaton, J.M., Brasington, J., Darby, S.E., Kasprak, A., Sear, D.A., Vericat, D., 2013. Morphodynamic signatures of braiding mechanisms as expressed through change in sediment storage in a gravel-bed river. *Journal of Geophysical Research: Earth Surface* 118, 759–779.