Automatic geomorphic feature extraction from lidar in flat and engineered landscapes

Paola Passalacqua,1 Patrick Belmont,2 and Efi Foufoula-Georgiou3

Received 21 May 2011; revised 13 February 2012; accepted 15 February 2012; published 29 March 2012.

[1] High-resolution topographic data derived from light detection and ranging (lidar) technology enables detailed geomorphic observations to be made on spatially extensive areas in a way that was previously not possible. Availability of this data provides new opportunities to study the spatial organization of landscapes and channel network features, increase the accuracy of environmental transport models, and inform decisions for targeting conservation practices. However, with the opportunity of increased resolution topographic data come formidable challenges in terms of automatic geomorphic feature extraction, analysis, and interpretation. Low-relief landscapes are particularly challenging because topographic gradients are low, and in many places both the landscape and the channel network have been heavily modified by humans. This is especially true for agricultural landscapes, which dominate the midwestern United States. The goal of this work is to address several issues related to feature extraction in flat lands by using GeoNet, a recently developed method based on nonlinear multiscale filtering and geodesic optimization for automatic extraction of geomorphic features (channel heads and channel networks) from high-resolution topographic data. Here we test the ability of GeoNet to extract channel networks in flat and human-impacted landscapes using 3 m lidar data for the Le Sueur River Basin, a 2880 km2 subbasin of the Minnesota River Basin. We propose a curvature analysis to differentiate between channels and manmade structures that are not part of the river network, such as roads and bridges. We document that Laplacian curvature more effectively distinguishes channels in flat, human-impacted landscapes compared with geometric curvature. In addition, we develop a method for performing automated channel morphometric analysis including extraction of cross sections, detection of bank locations, and identification of geomorphic bankfull water surface elevation. Using the slope plotted along each channel-floodplain cross section, we demonstrate the ability to identify and measure the height of river banks and bluffs. Finally, we present an example that demonstrates how extracting such features automatically is important for modeling channel evolution, water and sediment transport, and channel-floodplain sediment exchange.


1. Introduction

[2] High-resolution topographic data derived from light detection and ranging (lidar) technology enables detailed geomorphic observations to be made on spatially extensive landforms in a way that was previously not possible. This provides new opportunities to study the spatial organization of landscapes and channel network features, increase the accuracy of environmental transport models (e.g., water, sediment, and nutrients) and inform decisions for targeting conservation practices by using data in conjunction with modeling. However, with the opportunity of increased resolution topographic data come formidable challenges in terms of automatic geomorphic feature extraction, analysis, and interpretation. Low-relief landscapes are particularly challenging because the topographic gradients (and signal-to-noise ratio) are low and in many places both the landscape and the channel network are heavily modified by humans. This is especially true for agricultural landscapes, which dominate the midwestern United States.

[3] Agricultural landscapes are very active geomorphologically and contribute a significant amount of fine sediment to rivers [Hooke, 2000; Wilkinson and McElroy, 2007; Montgomery, 2007]. Accelerated erosion is problematic from the standpoint of losing productive and valuable soil as well as degrading water quality and aquatic habitat. However, several attempts to mitigate fine-sediment loading in rivers that drain agricultural areas have so far demonstrated
only marginal success [Langland et al., 2005; Engstrom et al., 2009], partly because of a general inability to specifically identify the locations of sediment sources, particularly near-channel sources of sediment. Development of an accurate sediment budget and routing model requires explicit specification and measurement of near-channel terrain features that function as sources and sinks. Typically, this is done by hand digitization of maps and manual specification of the coordinates of such features. On a large watershed scale (e.g., Chesapeake watershed 166,000 km²) this can be tedious and time consuming and introduces a level of subjectivity and human error (S. S. Day et al., Rates and mechanisms of bluff erosion in a rapidly incising river system, submitted to Earth Surface Processes and Landforms, 2012). Automated techniques, such as those discussed in this paper, will inevitably play an important role in exploiting the wealth of information contained in lidar data, which will likely enhance our understanding of sediment sources and dynamics and therefore enable more effective conservation efforts.

The goal of this paper is to demonstrate a general approach to extract geomorphic features in flat and heavily engineered landscapes. Several authors have proposed methods for classification, feature detection and analysis of associated uncertainties in agricultural landscapes [e.g., Bailly et al., 2008; Lagacherie et al., 2010; Levavasseur et al., 2010; Bailly et al., 2011]. The approach described here is parallel and distinct. We have developed a suite of new techniques, specific for flat landscapes, to be used within GeoNet, a recently proposed channel network extraction method [Passalacqua et al., 2010a, 2010b] (available as free software at https://sites.google.com/site/geonethome/). These techniques allow the identification of both natural and artificial channels (ditches) that comprise the drainage network of a landscape as well as the differentiation between the drainage network and engineered structures, such as roads and bridges. In addition, our method enables extraction of channel cross sections, automatic detection of bank locations, identification of geomorphic bankfull water surface elevation, and measurement of channel width and of bank and bluff heights. We test our techniques in the Le Sueur River watershed, a 2880 km² agricultural basin in southern Minnesota (see Figure 1). While the specific thresholds used here will most certainly vary from landscape to landscape, the statistical approach proposed is readily portable to other locations.

The paper is organized as follows. In section 2 we discuss a number of challenges in geomorphic feature extraction from lidar data. Section 3 describes the study area. In section 4 the use of different metrics of curvature for identification of channels and artificial landscape features is analyzed. Section 5 presents the channel network extraction method with consideration of specific challenges presented in flat landscapes. In section 6 we present the technique for automatic extraction of channel cross sections, channel bank locations, channel width, geomorphic bankfull water surface elevation, and height of adjacent banks and bluffs. Section 7 presents an example to show how feature extraction can be used to assess morphodynamic evolution. Finally, we conclude and present ideas for further research in section 8.

2. The Need for Automated Techniques in Analysis of Lidar Data

High resolution topographic data contains an immense amount of information, making it possible to capture in detail channel networks and other geomorphic features of interest. As these data sets continue to be collected over larger and larger areas, automated and semiautomated extraction
techniques are essential for the sake of objectivity, efficiency, and reproducibility of the analyses. The development and application of these techniques is not trivial though, as numerous challenges are encountered in the analysis of lidar data sets, some common to all topography, others specific to flat landscapes. Some of these challenges are discussed below.

[7] An accurate channel centerline is critical for automated analysis because it provides a static reference point from which channel cross section measurements can be made. In addition, it establishes a robust river distance coordinate system, which is essential for modeling and conservation planning. Centerlines derived from typical flow accumulation algorithms on 30 m digital elevation models (DEMs) greatly oversimplify the channel path, thereby underestimating sinuosity and actual channel length (see Figure 2a, red line). In contrast, channel centerlines derived using typical flow accumulation algorithms (D8 or Dinf) on high-resolution lidar DEMs tend to overexaggerate channel sinuosity and length because the centerline meanders around within the actual channel (see Figure 2a, blue line) as a result of “bumpiness” in the interpolated false water surface of the channel (Figure 2b).

![Figure 2a](image1.png) ![Figure 2b](image2.png)

**Figure 2.** A visual example of some of the challenges encountered in geomorphic feature extraction from digital elevation models (DEMs). (a) Channel centerlines derived from typical flow accumulation algorithms on 30 m resolution data greatly oversimplify the channel path (red line). In contrast, channel centerlines derived using typical flow accumulation algorithms on high-resolution topographic data tend to overexaggerate channel sinuosity and length (blue line). (b) Banks can be identified by visual inspection, but interpolation techniques used to generate lidar DEMs from point cloud data often result in “bumpy” water surfaces. The area in Figure 2a is part of study area C in Figure 1, while Figure 2b is located southeast of study area A in Figure 1.

[8] Complications introduced by bridges for traditional extraction techniques applied to lidar are not trivial, especially in flat landscapes. Low bridges in steep, mountainous landscapes may not be problematic because the bridges can be “overtopped” by simple pit-filling techniques, without much loss of fidelity to channel location. However, pit filling does not work in flat landscapes because the bridges are often significantly higher than the surrounding landscape. The typical solution to this problem is manual removal of bridge crossings, which is time consuming and impractical for large watersheds. For example, 14 bridge crossings exist in the small area represented by box A in Figure 1. In addition, positive curvature along the sides of the road can erroneously be mistaken as an incoming tributary.

[9] Another challenging feature of interest for channel network extraction in agricultural landscapes is artificial drainage ditches, either intermittent or perennial. Ditches are very distinct from a sediment transport standpoint. They serve as sediment sinks for much of the year, trapping sediment contributed from fields, but they are readily flushed during large events, so they can also serve as substantial sources of field sediment over relatively short periods of time. The distinction between artificial and natural channels is thus an important step toward identifying the major sediment sources of the basin as artificial ditches represent very distinct management challenges/opportunities.

[10] Banks delineate the geomorphic boundaries of the channel and represent a critical, discrete interface for exchange of sediment between channel and floodplain. Accurate identification of the banks allows for measurement of channel width and changes in channel width along the river, as well as measurement of bank elevations, which can be used to compute the net sediment contribution from stream banks as a result of meander migration [Lauer and Parker, 2008]. Although banks can readily be identified by visual inspection of a given channel cross section (see Figure 2b), we know of no automated technique for bank identification at the river network scale. In addition, lidar data typically contains “lumpy” false water surfaces as a result of poor reflectivity of water and imperfect algorithms for interpolating surfaces between lidar returns on the channel edge. A consistent bankfull water surface longitudinal profile would provide more accurate measurements of channel gradient than do the noisy, interpolated false water surfaces typical of lidar data sets. Here, we are applying the geomorphic definition of bankfull in which the lower of the two channel banks is used to determine the elevation at which water begins to inundate the floodplain.

[11] Bluffs are tall (3–60 m) features adjacent to the river that are commonly found in incised landscapes, such as our study area. Accurate identification of bluffs and measurement of their length and height are essential for modeling potential sediment contributions from this source.

[12] While most of the tasks listed above can be completed manually, automated procedures are likely more efficient and objective. The development of such techniques is the goal of this work.

3. Description of the Study Area

[13] The Le Sueur watershed presents a rich setting for illustration of our techniques as it contains a wide range of
hydrogeomorphological environments as the result of the Quaternary evolution of the southern Minnesota landscape, described in detail by Fisher [2003], Gran et al. [2009], and Belmont et al. [2011a]. In brief, a relatively flat landscape, underlain by about 60 m of interbedded fine-grained till and glaciofluvial sediments, was uncovered when the Wisconsinan ice sheet retreated from southern Minnesota about 16,000 years before present. Meltwater from the retreating ice sheet was stored in Glacial Lake Agassiz, which covered more than 400,000 km² of North Dakota, Minnesota, Manitoba, Ontario, and Saskatchewan. Around 13,400 years ago, Glacial Lake Agassiz catastrophically drained through the proto-Minnesota River Valley (MRV), incising the MRV by 70 m. This incision caused a drop in base level for Minnesota River tributaries, spawning knickpoints at the confluence of each tributary with the main stem MRV.

[14] The knickpoint that initiated at the mouth of the Le Sueur River has propagated nearly 40 km up each of the three branches of the river (the Le Sueur, Cobb, and Maple rivers; locations of knickpoints are indicated by orange arcs in Figure 1) over the past 13,400 years, causing vertical incision at a rate of 4–5 m kyr⁻¹ [Gran et al., 2011]. We refer to these three rapidly incising channels below the knickpoints as the knick zone. Within this knick zone, tall bluffs (up to 60 m) have formed that currently represent the largest source of fine and coarse sediment to the river. Aside from the three incising river valleys, the landscape is composed of vast, flat upland terrain, nearly all of which has been heavily engineered with roads, ditches, and artificial subsurface drainage conduits to support agricultural row crop production.

[15] The modern channel network can be divided into four distinct regions: relatively steep (average gradient 0.002 m m⁻¹) trunk channels within the knick zone, low-gradient (0.0004 m m⁻¹) natural channels above the knick zone, steep first- and second-order tributaries (ravines) connecting the incised river valley to the uplands, and, for the past 150 years, agricultural ditches. The morphology of agricultural ditches varies depending on location, age, and maintenance history, but typical ditches are shown in Figure 3. Each of these components presents different challenges in extraction of the sediment-relevant terrain features, including bluffs, ravines, floodplains, channel banks, and natural and artificial levees.

[16] The Le Sueur River is a major source of sediment to the Minnesota River, both of which are listed as impaired for turbidity by the U.S. Environmental Protection Agency under section 303(d) of the Clean Water Act. For the purpose of identifying sediment sources and supporting decisions for conservation actions, a sediment budget [Belmont et al., 2011b] and a morphodynamic sediment routing model [Viparelli et al., 2011] have been developed for the Le Sueur River Basin. Despite the fact that the landscape is primarily agricultural, and therefore assumed to have high soil erosion rates, the sediment budget concluded that near-channel erosion (banks, bluffs, and ravines), rather than soil erosion, comprises the dominant source of suspended sediment. As much as 70% of suspended sediment is derived from the small fraction of the landscape (<1%) surrounding channels, which emphasizes the need to model this most dynamic portion of the landscape well.

4. Curvature Analysis for Feature Detection in Flat Landscapes

[17] Previous work [Passalacqua et al., 2010a, 2010b] has shown that the isohyet contours curvature or geometric curvature, defined as

\[
\kappa = \nabla \cdot \left( \frac{\nabla h}{|\nabla h|} \right)
\]  

(1)

is preferable to the commonly used Laplacian in steep and natural landscapes. In (1) \( h \) is the elevation, \( \nabla h \) is the gradient, and \( |\nabla h| \) the magnitude of the gradient. Since the Laplacian is defined as

\[
\gamma = \nabla^2 h.
\]  

(2)

[18] The two curvature definitions differ in the normalization performed on the gradients before the application of the divergence operator in (1). Normalized gradients become of the same order of magnitude throughout the area in analysis, thus enhancing all the features present. However, it is noted that the presence of locally flat regions in a landscape, which would typically be regions of local maxima or local minima, would render \( \nabla h \) very close to zero, thus making the geometric curvature in (1) undefined. In practice, these values
Figure 4. Illustration of the differences between the Laplacian and the geometric curvature on a synthetic example. (a) The sinc function, (b) Laplacian curvature, and (c) geometric curvature. The profiles plotted on the right are obtained for the cross section $x = 100$. The geometric curvature, because of the gradient normalization, enhances all the features present, while the Laplacian assigns a higher value of positive curvature to more convergent features.
would be “squeezed” to a narrow range that would make detection of features difficult. This is demonstrated in an example below and is intuitively expected to play an important role in landscapes which contain naturally flat (e.g., low-slope floodplain) and artificially flat (e.g., road, levee top) features.

[19] To illustrate the main differences between these two curvature operators, a synthetic example is shown in Figure 4. Figure 4a shows the function \( h(R) = \frac{\sin(R)}{R} \) (with \( R = \sqrt{x^2 + y^2} \)) (first column) and a cross-sectional profile through the middle of the field (second column). This function is used here as a synthetic example since it resembles a convergent feature in the landscape. The Laplacian curvature and its profile are shown in Figure 4b, while the geometric curvature and its profile are shown in Figure 4c. Note that each of the ridges and valleys are very discrete and nearly of the same value in the geometric curvature (in Figure 4c), while the Laplacian (in Figure 4b) illustrates that the actual curvature values of ridges and valleys differ significantly, with the curvature of the center being threefold higher than the curvature of the outer valley. If all the features present on the landscape are of interest for the analysis performed, then the geometric curvature represents a valuable definition of curvature, as it is able to enhance all the features present on the landscape. However, it cannot be assumed that geometric curvature will be the most effective metric for extracting features in all types of landscapes. As discussed above, the presence of both natural (e.g., channels) and artificial (e.g., roads) features in flat landscapes could make the normalization operation of the gradients not optimal, as enhancing all the features present in the landscape would result in enhancing roads, which may then become indistinguishable from the channel network.

[20] To further investigate this issue, we focused on part of the 25 km² tributary of the Le Sueur River, shown in Figure 1 (labeled A) and in Figure 5. This tributary was chosen because it exhibits the same topographic variability of the whole Le Sueur River basin (broad flat uplands, low-gradient channels and ditches, as well as a steep, incised river valley) and thus can be considered as a microcosm of the whole system. The first column of Figure 6 shows 70 m by 70 m hill-shaded DEMs cut in correspondence of a channel (Figure 6a), a road (Figure 6b), and an artificial ditch (Figure 6c), while the second column shows the corresponding cross sections. The third and the fourth columns contrast the behavior of the geometric curvature (third column) as defined in (1) versus the Laplacian (fourth column) as defined in (2). The geometric curvature assumes approximately the same range of values for the channel, the ditch, and the road because of the normalization operated on the gradient, which makes all three features equally detectable (the road itself has negative curvature, but there are small hillslopes on the sides of the road that exhibit positive curvature). The Laplacian, instead, as shown in the most right column of Figure 6, is of about the same order of magnitude for the channel and ditch, while largely decreases for the road. This characteristic of the Laplacian makes it an optimal measure of curvature for low-gradient landscapes where both natural and artificial features are present.

[21] The quantile-quantile plot of a variable of interest, in this case curvature, compares its statistical distribution to a normal distribution. If the data plot along the straight
line, then the variable has a normal distribution, otherwise
the point at which the curve deviates from the straight line
indicates a transition in the statistical behavior of the sys-

Following the work of Lashermes et al. [2007], this
deviation can be interpreted as the transition from hillslope
to valley and the corresponding curvature value can be
used as a threshold to extract the set of likely channelized
pixels (pixels with curvature greater than the threshold).
The quantile-quantile plot for study area A in Figure 7a
shows that the transition from normal behavior happens at
a value of the standard normal deviate $z_{i}$ The curve
remains at around zero for a fairly long portion, showing
the presence of a large number of pixels with curvature
about zero (to see a comparison with other quantile-quan-
tile curves, see Figure 5 of Passalacqua et al. [2010a] and
Figure 6 of Passalacqua et al. [2010b]). To further investi-
gate this behavior, we focused on 1.5 by 1.5 km of study
area A, shown in Figure 7b. We indicate as zone 1 the
range of curvature values corresponding to $z < 1$ and with zone 2 the range for $z > 1$. The pixels in
zone 1 and zone 2 are shown in white in Figures 7c and 7d,
respectively. The set in Figure 7d represents the set of
likely channelized pixels, as defined by Lashermes et al.
[2007], but we can see that in this case it erroneously
includes convergent features on either side of the road
(small ditches and/or breaks in slope) as part of the channel
network. This observation calls for a different interpreta-
tion of the quantile-quantile plot of curvature in a landscape
that contains natural and artificial features.

We propose that such a new interpretation can be
based on the curvature analysis previously described and
shown in Figure 6. In fact, what we observed previously
was that while channels and ditches both have curvature
values exceeding a specific quantile, convergent features
on either side of roads have indeed much smaller curva-
ture. On the basis of the curvature analysis, we identified
0.05 m m as the maximum value of curvature in corre-
spondence of the roadside (as can be seen from the example
in Figure 6b, Laplacian curvature map’s color bar), thus
suggesting further analysis of the quantile-quantile plot in
zones 3 and 4 (see Figure 7a). Figure 7e shows in white
the pixels with curvature above the value corresponding to
deviation from normal behavior but less than 0.05 m m (zone 3), while Figure 7f shows the pixels with curvature
Figure 7. (a) Quantile-quantile plot of Laplacian curvature. The middle part of the plot corresponding to standard normal deviate $z$ between $-1$ and $1$ is flat, highlighting the presence of a large number of pixels with curvature almost zero. (b) To analyze this behavior further, we have extracted the pixels that belong to the two areas, 1 and 2, in a box $1.5 \times 1.5$ km. Pixels belonging to (c) zone 1 ($-1 < z < 1$) and (d) zone 2 ($z > 1$) are plotted in white. Notice that the middle flat part of the quantile-quantile plot is due to the very flat topography that characterizes this basin (Figure 7c). Zone 2 is composed of the convergent pixels in channels and roads, thus using the deviation from normal behavior at $z = 1$ as the curvature threshold would result in a skeleton that includes both channels and roads. To be able to distinguish among the two, we have used the information obtained from the analysis shown in Figure 6. As can be seen, the curvature corresponding to the road reaches a maximum of approximately 0.05 m m$^{-2}$. (e) Zone 3 shows the pixels above deviation from normal behavior $z > 1$ but with curvature $\gamma < 0.05$ m m$^{-2}$, while (f) zone 4 is composed of the pixels with $\gamma > 0.05$ m m$^{-2}$. We can notice that in zone 4 the pixels belonging to the road are nearly all eliminated.
greater than 0.05 m m$^{-2}$ (zone 4). The curvature value 0.05 m m$^{-2}$ is specific to our landscape in analysis, being dependent on the scale and the type of features. However, the approach we have taken to detect the appropriate threshold from a curvature analysis, can be performed in any landscape to identify features that exhibit distinct curvature ranges. By defining 0.05 m m$^{-2}$ as the approximate value of curvature threshold to be used for identifying the likely channelized pixels in this landscape, we obtain the set shown in Figure 7f, which includes only a small number of roadside pixels.

5. Automatic Channel Network Extraction in Flat and Human-Impacted Landscapes

[23] The first step of GeoNet is a nonlinear filtering operation [Perona and Malik, 1990] to focus the analysis on the scale of interest, by smoothing variability at smaller scales (e.g., bumpiness of the ground) and enhancing features at the scale of interest or larger. The definition of the scale of interest depends on the application and on the landscape analyzed. In this work, we are interested in the channel network and channel morphology extraction, thus, the scale of interest can be defined as the channel bank, while smaller scales, for example, the bumpiness detected along the channel, should be smoothed out. For a complete discussion on the nonlinear filtering, see Passalacqua et al. [2010a, 2010b]. After the operation of nonlinear filtering, the set of likely channelized pixels is identified by applying the curvature threshold of 0.05 m m$^{-2}$, as suggested by the previous analysis. This set of pixels can be further narrowed by applying a small contributing area threshold (smaller than the value of area at channel initiation), able to exclude small convergent areas which do not belong to channels, without affecting channel head locations. In our case, we applied an area threshold of 5000 m$^2$ (see Passalacqua et al. [2010b] for discussion and selection of this area threshold), obtaining the skeleton of likely channelized pixels shown in Figure 8. This skeleton is used within GeoNet to detect channel heads as the upstream end points of the skeleton branches [see Passalacqua et al., 2010a]. Channels are then traced between the outlet and the channel heads as geodesic curves, or curves that minimize a certain cost $\psi$, which can be defined as a combination of flow accumulation and geometric curvature [Passalacqua et al., 2010b]:

$$\psi = \frac{1}{\alpha A + \delta \kappa}$$  \hspace{1cm} (3)

where $\alpha$ and $\delta$ are constants used to balance dimensionality and normalize the difference in order of magnitude between area and curvature. On the basis of the analysis described in [Passalacqua et al., 2010b], $\alpha$ can be fixed to 1 and for $b$ the order of magnitude of the mean accumulation area per contour length can be used. In our application, we used a cost function of the form

$$\psi = \frac{1}{\alpha A + \kappa + a S}$$  \hspace{1cm} (4)

where $a$ is the mean area per unit contour length (taken as the grid size) and $S$ is the skeleton based on curvature and geometric curvature $\kappa$. Similarly, $A$ is the mean area per contour length. In Figures 9c and 9d we show two conventional channel network extractions based on area threshold of 5000 m$^2$ (Figure 9c) and 10$^6$ m$^2$ (Figure 9d), compared to the reference network. The network in Figure 9c is shown to demonstrate how the same value of area threshold is used differently by GeoNet and a conventional area threshold method. The area threshold in Figure 9d is more appropriate for study area A and results in a more realistic network density, while, however, substantially shortening and straightening the extracted channels. The definition of an area threshold in this region is particularly challenging as the range of drainage areas over which first-order streams are initiated varies significantly (<1 to 217 km$^2$), in part because of the geologically young evolutionary stage of landscape development and human modifications to the hydrologic system [Smith et al., 2011]. A conventional method based on area threshold only would not be appropriate to define channel initiation in this landscape.

[24] To analyze the behavior of the GeoNet channel network extraction method in the presence of bridge crossings, we focus on the bridge shown in Figure 10a indicated by the letters $A$ and $B$, which crosses the channel perpendicularly. This bridge crossing is part of study area A (see Figures 1 and 5). Figure 10b shows that the skeleton of area, linearly filtered to increase its continuity (the skeleton is very disrupted as seen in Figure 8). The filtered skeleton was added to help the geodesics in following the meandering channels.

Figure 8. Skeleton of likely channelized pixels of study area A in Figure 1. The pixels with curvature and contributing area above threshold ($\gamma > 0.05$ m m$^{-2}$ and $A > 5000$ m$^2$) are plotted in black (skeleton is thicker to increase visibility).
likely channelized pixels correctly appears disrupted at the location of the bridge as at that location the threshold criteria imposed on curvature and contributing area are not satisfied. In Figure 10c we show the channel extraction obtained by applying a conventional extraction method based on an area threshold of 5000 m². Here part of the roadside is detected as tributary and the main channel erroneously splits in two at the location of the bridge. The GeoNet extraction, instead, as shown in Figure 10d, correctly depicts the channel path before and after the road crossing.

To analyze the behavior of GeoNet in tracing artificial drainage ditches, we focus on part of the Maple River basin, tributary of the Le Sueur River and located southwest...
Figure 10.  (a) Analysis of a road crossing in a 800 m × 800 m focus area. Hill-shaded DEM. The letters A and B indicate the horizontal extent of the road in the focus area. (b) The skeleton of likely channelized pixels correctly detects a channel disruption due to road crossing, while the extraction based on area threshold fails to do so. (c) The channel splits in two in correspondence of the road. (d) GeoNet correctly traces the channel path through the road crossing.

Figure 11. Location map of part of the Maple River basin, tributary of the Le Sueur River (box B in Figure 1).
of the tributary analyzed above (see box B in Figure 1 and see Figure 11). Figure 12a shows the manual delineation of ditch and natural channel from the USGS blue lines (obtained from http://deli.dnr.state.mn.us/). In particular, the left, more linear tributary is a ditch, while the more sinuous line on the right is a natural channel (the Maple River main stem). One challenge encountered in the application of a conventional channel extraction method is the fact that extra sinuosity is added, as the channel bounces around within the channel banks, an artifact of the uneven water surface and line of steepest descent technique utilized in the conventional method. An example of this is shown in Figure 11b (area threshold 30,000 m²), where the extracted path deviates from the actual one. The GeoNet extraction, shown in Figure 11c, although detects part of a road in the lower left side of the box, correctly traces the linear path corresponding to the artificial ditch. An automatic distinction of natural channels and artificial ditches is challenging because they do not always differ in sinuosity. In some cases, in fact, the artificial ditches have been designed on top of natural channels. While development of an automatic technique to distinguish these features is the subject of current research, the ability of GeoNet to trace ditches correctly, as shown in Figure 12, represents a promising starting point in that direction.

6. Automatic Extraction of Channel Morphology: Cross Section, Banks, Water Surface Elevation, and Bank and Bluff Height

[27] We develop below a method for the automatic extraction of channel cross section, detection of bank location, identification of geomorphic bankfull water surface elevation, and measurement of channel width and of bank and bluff height. Note that, as mentioned in section 2, we are applying here the geomorphic definition of bankfull in which the lower of the two channel banks is used to determine the elevation at which water begins to inundate the floodplain.

[28] For demonstration purposes, we focus this part of the analysis on the area shown in Figure 13 (see also box C in Figure 1), near the confluences of the Maple and Cobb Rivers with the Le Sueur main stem. This reach is characterized by tall bluffs, whose locations have been detected in the map by thresholding the local relief (minimum of 3 m relief in an 81 m² area; see mapped features in Figure 13) and offer a good test case for the analysis herein. Note that Figure 13 includes only bluffs that are directly connected to the channel.

[29] Let us assume we want to extract the channel cross section at a generic location \((x_1, y_1)\) along the main channel (extracted using GeoNet and shown in Figure 13). First we can compute the vector between location \((x_1, y_1)\) and the immediate upstream (or downstream) location along the main channel, which we can indicate as \((x_2, y_2)\). From vector algebra, the vector connecting these two points is given by

\[
\vec{r} = (x_2 - x_1) \hat{i} + (y_2 - y_1) \hat{j}
\]

After obtaining the vector \(\vec{r}\), we can compute the orthogonal vector \(\vec{q}\) as

\[
\vec{q} = (-r_2) \hat{i} + (r_1) \hat{j}
\]

In our analysis, the above steps are performed using not one, but five upstream and downstream locations (about 30 m along the main channel) to make the computation of the orthogonal transect less sensitive to local changes in channel sinuosity. To trace an orthogonal transect, in fact,
Figure 13. Location map of the Le Sueur River at the Maple and Cobb confluences (box C in Figure 1). The channel morphology extraction is applied along the Le Sueur main stem, which is characterized by a large number of bluffs (yellow patches in the map indicate pixels with elevation greater than 3 m along the channel in analysis).

Figure 14. Example of geomorphic feature extraction along the main stem. (a) Location of the transect along the main stem. (b) Original cross section. (c) Slope along the 300 m transect. The water surface is traced at the elevation of the lower channel bank. (d) The letters A and B indicate the detected bank locations.
one needs to go upstream and downstream some distance. While 5–10 m would be too short (because of variability in the centerline at that scale), a distance greater than 30 m could be problematic in meandering sections. This length scale will probably change with the channel size and the 30 m here employed corresponds to the order of magnitude of one channel width. Then, the unit vector is given by

$$\hat{e} = \frac{\vec{q}}{|\vec{q}|}$$

(7)

where $|\vec{q}|$ is the magnitude of the vector $\vec{q}$. The coordinates $(x, y)$ of a transect of about 300 m (150 m on each side of the centerline) can be obtained by multiplying a vector $[-50, 50]$ by the components of the unit vector $\hat{e}$. This range of 300 m was defined on the basis of the fact that it corresponds to the extent of the channel, floodplain and contributing bluffs in this area. Figure 14a shows an example of such a transect orthogonal to the main stem of the Le Sueur tributary extracted with the procedure described above. To extract the channel cross section, we simply have to plot the topographic data corresponding to the transect orthogonal to the channel. Figure 14b shows the cross section correspondent to the transect shown in Figure 14a. Once the cross section has been extracted, there are numerous automated, semiautomated, and manual analyses that may be performed, depending on the questions of interest. We discuss only a few, broadly applicable examples here.

Our first objective is to develop a method for automated detection of channel banks. These features are typically associated with a dramatic change in slope in the cross-section view. In particular, we expect the slope to be large along the “walls” of the channel and smaller on the adjacent floodplain. The magnitude of local slope, plotted in Figure 14c for the pixels along the transect, shows peaks on each side of the channel centerline, indicating where the slope is large. This information can be used to identify the channel bank locations. First, we identify the peaks of the slope and we take into consideration only the ones within approximately 50 m on each side of the centerline, since 100 m is the order of magnitude of the maximum channel width in this region. Then, among the two maximum peaks on each side, we pick the smaller one as the one corresponding to the lower geomorphic bank, thus the one defining the geomorphic water surface elevation at bankfull stage, just below the elevation at which floodplain inundation occurs. We expect that the assumption made here (i.e.,

Figure 15. Example of geomorphic feature extraction along the main stem. (a) Location of the transect along the main stem. (b) Original cross section. (c) Slope along the 300 m transect. The water surface is traced at the elevation of the lower channel bank. (d) The letters A and B indicate the detected bank locations, while C indicates the point that could be considered the channel bank location instead of A. Air photos show the presence of mature trees between A and C, thus suggesting that the detected bank locations are correct.
that the smaller of the two localized peaks on the slope-distance plot in Figure 14c corresponds to the depositional geomorphic bank, while the higher of the two localized peaks corresponds to the cutbank) will hold for most meandering, single-threaded rivers following the geometric argument laid out by Lauer and Parker [2008].

Once the water surface elevation has been identified, we determine the location of the other bank as the point at which the identified water surface elevation projects on the other side with respect to the channel centerline. Figure 14d shows the channel cross section after tracing the geomorphic bankfull water surface elevation. Note that this operation constitutes a smoothing of the water surface, thereby improving the quality of the water surface profile compared to that originally interpolated from the lidar point cloud. The letters A and B in Figure 14d indicate the detected bank locations from which the channel width can be measured as the distance between the two banks along the transect.

Figure 15 shows another example of cross-section extraction, automatic detection of bank locations, and identification of geomorphic bankfull water surface elevation. The letters A and B indicate the detected bank locations, while C indicates the location that, by looking at the cross section, could be considered the bank location instead of A. From visual inspection of recent air photos, it is apparent that mature trees are located between A and C, thus concluding that the active channel is the one between A and B, while the section between A and C receives flow only occasionally. This example shows that, despite the availability of lidar data, air photos, and field experience, the definition of channel boundaries can be very challenging.

The operations described above can be automatically repeated at any user-specified frequency along the channel network, obtaining values of geomorphic bankfull water surface elevation and channel width along the channel. Figure 16 shows the comparison of the geomorphic bankfull water surface elevation identified from raw data (Figure 16a) with the one obtained using GeoNet on the nonlinearly filtered data (Figure 16b). As can be seen, the water surface in Figure 16b appears smoother and less noisy. The reach average channel width based on field-measured cross sections (six measured within the reach here analyzed, but for a different purpose, which biased the sampling toward narrower channels) is 37 m (range between 29 and 50 m), and the reach average channel width measured from five sets of air photos (2003, 2004, 2005, 2006, 2008) is 47 m. The latter is a more robust measure of average because it was calculated by digitizing a polygon along 10 km of river and dividing by the length. Data for field surveyed and air photo measured widths are given by Belmont et al. [2011b, supplementary information]. The average width as extracted by our method is 55.7 m. There clearly will be cross sections that are misdelinated by our automated technique because of the complex nature of channel topography, which can be further complicated depending on the water depth at the time of the lidar flight and the algorithm used to convert the raw point cloud data to a raster DEM. In our case, we are using a relatively coarse resolution (3 m horizontal) of the lidar DEM and the flat topography may be contributing additional complications in the detection. We visually checked the cross sections extracted with our technique and approved as correct 70% of the total cross section analyzed. While better performance is expected at higher resolutions and in steeper areas, errors in detection are infrequent and appear to be smoothed out even with a very small filtering window.

Channel and floodplain morphology are important indicators of landscape history and the dynamic relationship between flow, sediment supply, and base level change. Therefore, identification and automated measurement of bank and bluff heights on either side of the channel are useful for understanding basic processes governing channel evolution, sediment transport, and channel-floodplain exchange as well as answering applied questions such as identification and quantification of near-channel sediment sources for construction of a sediment budget. The measurement of the difference in elevation between opposing channel banks (which we refer to as $\Delta H$) provides important information about channel processes and sediment dynamics. For example, Lauer and Parker [2008] show that an estimate of net sediment inputs from banks as a result of meander migration can be computed as a function of $\Delta H$ and the meander migration rate. Correct and automated measurement of bank and bluff height is essential for

![Figure 16. Water surface elevation computed at cross sections 3 m apart. The gray line indicates the fitted least squares line. (a) Original data. (b) Nonlinearly filtered data.](image-url)
achieving this goal at the network scale. Below we further discuss the ability of the proposed technique to automatically identify banks and bluffs, measure the difference in elevation between opposing banks/bluffs, and we explore how this information can be used for modeling sediment supply to the channel.

Figure 17 shows an example of extraction of channel cross section and channel morphology similar to the two examples shown in Figures 14 and 15. The letters A and B indicate the detected bank locations and the height of the bluff is correctly given by the difference in elevation between these two points. As previously discussed, complex channel, bank, and bluff morphology can make this automatic detection challenging. In the cross section shown in Figure 14, for example, location B is identified as the first slope peak on the left of the channel centerline. However, the bluff on the left side of this cross section extends out to 150 m from the channel centerline; thus, in this case our procedure has failed in identifying the bluff height. This could be solved by modifying the algorithm to search along the cross section until it passes a peak and comes down to some low end-member threshold (indicative of a floodplain, terrace, or upland area). In Figure 15 the floodplain, indicated as B, is located between the channel and the bluff. Even if the bluff extends over 150 m to the right of the channel centerline, the difference in elevation between the banks is correctly computed as the one corresponding to A and B, as the sediment coming from the bluff is buffered by the floodplain and does not interact directly with the channel.

7. Example Application: Using Geomorphic Features for Sediment Flux Computation

Geomorphic features, automatically extracted from lidar data, can be used to model channel evolution, water and sediment transport, and channel-floodplain sediment exchange. The bluff and bank heights, for example, can be computed along the main stem at each cross section, resulting in a plot of bank and bluff height versus distance as shown in Figure 18a. Combining this data with meander migration rates (Figure 18b) from historic air photos (1938–2005) measured every 10 m along the channel, the net sediment flux to the channel can be computed as:

$$ S = \sum_{i}^{N} \Delta \eta_{i} c_{i} \Delta l \rho $$

where $S$ is the net cumulative sediment input for each 10 m increment along the channel, measured in Mg yr$^{-1}$, $\Delta \eta$ is the difference in elevation between the channel banks in m, $c$ is the meander migration rate in m yr$^{-1}$, $\Delta l$ is the length increment along the channel over which the calculation is performed (10 m in this case), and $\rho$ is the bulk density (1.3 Mg m$^{-3}$) [Belmont et al., 2011b]. Figure 18c shows
the cumulative sediment inputs with distance downstream for one reach of the Le Sueur River within the knick zone. It should be noted that this prediction is not a complete mass balance sediment budget for the channel reach because it does not account for overbank deposition. However, it does provide a useful, decadal-scale estimate of net sediment inputs from bank and bluff erosion.

Measurement of $\Delta \eta$ is also useful for extracting information about river valley development over landscape evolution timescales. Bluff height can be viewed as a surrogate for the amount of incision since the river last occupied a particular position. Because the process of vertical incision is constantly generating taller and taller bluffs, and the process of lateral migration is constantly widening the river valley, eroding bluffs down to the current elevation of the river and therefore continually decreasing the probability that the river will encounter a bluff, the distribution of bluff heights along the river reflects the competition between vertical incision and lateral migration of the river over valley-forming timescales. For a river that is incising vertically much more rapidly than it is migrating laterally, $\Delta \eta$ should be high (70 m in this case as dictated by base level fall) and the proportion of the river that is eroding into a tall bluff should be high. For a river that is migrating laterally at a rate much faster than it is incising, $\Delta \eta$ should be relatively small and is primarily a function of deposition rather than incision, following the rationale of Lauer and Parker [2008]. Thus, it is possible to semiquantitatively assess the competition between lateral migration and vertical incision in the Le Sueur River from a cumulative probability plot of $\Delta \eta$, as shown in Figure 19. The observation that only 10% of the bluffs exceed 10 m in height suggests a predominance of lateral migration versus vertical incision for the area in analysis.

8. Concluding Remarks

Automatic extraction of detailed and localized geomorphic features of interest, such as channel networks, channel banks, and bluffs, is very important for accurate estimation of sediment sources and flux transport in watersheds. Flat and human-impacted landscapes present additional challenges in such an automatic extraction process, as bridges across channels, artificial ditches and other engineered structures interfere with the process of feature extraction based on geomorphologic attributes such as slopes, curvatures, etc. The increasing availability of high-resolution lidar topography over large areas, e.g., whole states and large watersheds, makes it imperative that we develop robust and reliable automatic methods for extraction of features of interest to be used in hydrologic, geomorphic, and water quality studies.

On the basis of previous development of a novel nonlinear diffusion filtering method combined with a local-to-global (geodesic) optimization technique for channel network extraction [Passalacqua et al., 2010b] packaged under the name GeoNet, we presented in this paper

Figure 18. Bluff height $\Delta \eta$ plotted as a function of (a) distance along the main stem, (b) migration rate, and (c) cumulative sediment inputs from bank and bluff erosion. This analysis refers to the area in Figure 13 and box C in Figure 1.

Figure 19. Probability of exceedance of the bluff height $\Delta \eta$. The shape of the curve indicates a slight predominance of lateral migration over vertical incision.
important methodological extensions that address challenges presented in flat and human-impacted landscapes. In particular, we presented a detailed analysis of a flat watershed in the upper Midwest (the 2880 km² Le Sueur River basin). We proposed and tested techniques that allow the extraction of channels interrupted by bridges and other crossings, extraction of channel morphology properties such as channel width, bank elevations, bankfull water surface elevations, and extraction of discontinuous features in the landscape such as steep bluffs, that are important localized sources and sinks of sediment.

The innovation of the extended method lies on subtle use of geomorphologic attributes compared to a more straightforward and less demanding use in steep terrains [Passalacqua et al., 2010a, 2010b]. For example, we demonstrated that while the use of geometric curvature presented distinct advantages in steep terrains because of its normalization properties, in flat terrains with artificial features this is a disadvantage, and one should use the unnormalized Laplacian curvature to depict regime transitions. Second, while possible channelized pixels (skeleton of river network) in steep terrains were extracted via a simple threshold in the probability distribution of curvatures, in flat terrains we demonstrated that a more refined curvature interrogation is needed to differentiate between natural and manmade features. Specifically, we identified two regions of curvature threshold which depict natural versus human engineered features. While the thresholds themselves will differ for other locations, the approach here proposed is portable. Third, zooming down to specific cross sections of channels, we demonstrated how nonlinear filtering offers advantages in depicting channel banks and as a result allows the estimation of smooth water surface elevation along channels, important for hydrologic and hydraulic modeling computations, channel width, and river bank and bluff height.

The methods presented in this contribution will be incorporated in the previously developed channel network extraction package GeoNet, which is publicly available for testing and use by the community.

Acknowledgments. We gratefully acknowledge support from the National Center for Earth-Surface Dynamics (NCEQD), a Science and Technology Center funded by NSF under agreement number EAR-0120914, NSF (supplement to a CDI grant EAR-0835789), and Google Inc. We thank the Editor, the Associate Editor, two anonymous referees, and Stefano Orlandini for insightful reviews of our paper, which improved the presentation of our results.

References


