Identifying Wetland Depressions in Bare-Ground LIDAR for Hydrologic Modeling

Janet H. Gritzner

Abstract

Processing Digital Elevation Models (DEMs) for hydrologic modeling generally requires filling of depressions (sinks), features that inhibit surface flow. Majority of sinks are spurious artifacts of DEM creation. Yet some sinks represent real features (i.e., lakes, depressional wetlands, and karst landforms) and should not be filled. To maintain sinks, a threshold grid (wetland mask) that identifies real depressions is first generated and, from it, a locate (or seed) mask. Creating a realistic mask grid requires DEMs that properly resolve topographic depressions, a methodology for sink selection, and data for validation.

This paper evaluates bare-ground LIDAR (Light Detection and Ranging) data for its ability to resolve surface features in Devils Lake Basin, North Dakota, a region characterized by thousands of shallow, water-filled depressions with sizes ranging from a fraction of a hectare to several square kilometers. It uses National Wetland Inventory (NWI) and Landsat data to aid and validate wetland identification and raster operations to calculate sink depth and area.

Background

Hydrologic modeling with Geographic Information Systems (GIS) technology starts with good elevation data. Choice of data source is application dependent. United States Geological Survey (USGS) provides Level 1 DTEDs (Digital Terrain Elevation Data), 30-meter profile Level 2 DEMs, seamless 10 and 30-meter National Elevation Dataset (NED), and now is incorporating bare-ground LIDAR data into NED.

The surfaces of low-lying, extremely flat areas, and/or containing numerous ponds and wetlands may not be adequately represented using DEMs created using hypsography data extracted from 1:24,000 scale USGS topographic maps (Blackwell and Wells, 1999). Bare-ground LIDAR shows potential for resolving subtle landscape features and in this project is used the constructing a hydrologically-correct DEM. Creating a hydrologically correct DEM generally follows the flow model depicted in Figure 1.
Creating a Hydrologically Correct DEM

Figure 1: Procedures for creating a hydrologically correct DEM

Conditioning phase of hydrologic modeling in ArcGIS requires the development of three general utility datasets. In the order that they are produced, they are a DEM with sinks filled, a grid indicating the flow direction for each cell, and a flow accumulation dataset in which each cell receives a value equal to the total number of cells that drain into it. Depressions or sinks in a DEM hinder flow routing and need to be handled prior to developing flow direction or flow accumulation grids. Many sinks are spurious, a result of DEM creation errors and should be eliminated (Jensen and Domingue, 1988). Some sinks need to be maintained.

Real sinks include lakes (i.e., the Great Salt Lake), karst, and depressional wetlands in glacial, playa, and coastal terrain (Winter and LaBaugh, 2003). To fill or not to fill becomes the question. A procedure is in place to selectively fill sinks. A prerequisite is the development of a threshold grid or sink mask. Production of Elevation Derivatives for National Applications (ETNA) data employs a partially automated process, whereby all sinks with cell groupings less than 1000 are filled; only the larger depressions are maintained (Franken et al., 2000). Depth and volume of sinks are not considered, nor are ancillary data employed in constructing threshold masks. Alternate approaches would perhaps yield better results by resolving critically a greater number of depressional features.

The purpose of this work is to focus on identifying wetland features in a LIDAR-derived DEM which in turn will be used to construct a realistic sink mask. This mask will subsequently be used with a set of ArcInfo GRID commands to selectively fill a LIDAR
DEM. The resultant DEM is employed in generating flow direction, flow accumulation, and drainages line grids.

Study Area

The Prairie Pothole Region (PPR), a 715,000 sq kilometer area, is full of real sinks, i.e., depressional wetlands. The PPR extends from north-central Iowa to central Alberta, and contains parts of Minnesota, South Dakota, North Dakota, Montana, Manitoba, and Saskatchewan.

The landscape of the PPR was formed some 12,000 years ago during the Pleistocene Epoch. When the last glaciers retreated, they left behind a multitude of small saucer-like depressions which we have given the name prairie potholes or sloughs.

The PPR contains an estimated 25 million depressional wetlands, which fill with snowmelt and rain in spring. The highest densities of wetlands are found in areas of terminal and recessional moraines, where there may be as many as 40 wetlands per sq. kilometer (Sethre et al., 2005).

Prairie potholes vary in size and a permanency. Cowardin et al. (1979)’s classification system used in the coding of the National Wetland Inventory (NWI) dataset places them into four types of water regime or permanency: semi-permanently flooded, seasonally flooded, temporarily flooded, and intermittently exposed (Cowardin et al., 1979). Hydrologically, pothole wetlands can function as groundwater recharge sites, flow-through systems, or groundwater discharge sites. What function a pothole wetland performs is dependent on climate variation, position in the landscape, groundwater flow characteristic, soils, and geologic substrate. The regional hydrology of the PPR is poorly understood. Euliss et al. (2002) speak of a wetland continuum defined by the groundwater system. Wetlands then are the surface-water expressions of larger ground-water watershed, in which wetlands serve and discharge functions with respect to ground water (Euliss et al., 2002). Isolated wetlands can spill over their surface divides into adjacent wetlands during periods of abundant precipitation and high water levels (Winter and LaBaugh, 2003). The size of a prairie potholes ranges from a fraction of an acre to several square miles. Potholes larger than about 40 acres are arbitrarily called lakes by the U.S. Fish and Wild Wildlife Service. Most depressional wetlands are small. The median wetland in Eastern South Dakota was 0.16 ha. (Johnson and Higgins, 1997), which as Sethre et al (2005) pointed out is only a little larger than the 0.09 ha pixel size of Landsat TM data and is not resolved at all in 30 meter NED.

The study area is located in the Devils Lake Basin of North Dakota, some 90 miles west of Grand Forks, North Dakota. Choice for the location of the study was dictated by the availability of high resolution bare-ground LIDAR for the PPR. A LIDAR-derived DEM for 2,524 sq kilometers of the Devils Lake Basin is freely distributed by the St. Paul District of the U.S. Army Corps of Engineers (USACE). It provided by the Federal Emergency Management Agency (FEMA) to USACE to support hydrologic studies of the problematic Devils Lake Basin.
The Devils Lake Basin covers a total of about 9,870 square kilometers (3,814 square miles) containing all or portions of nine counties, a small portion of the Turtle Mountain Chippewa Reservation, and a large portion of the Spirit Lake Sioux Reservation. Devils Lake consists of several connected basins that are care enclosed within a larger basin currently with no surface outlet (Aber et al. 1997). Since glaciations, Devils Lake has been fluctuating from overflowing to being dry. As a result of increased snowfall and higher than normal precipitation rates since 1993, the water level of Devils Lake has risen nearly 8 meters to reach its highest lake elevation of 1,447.2 ft above sea level in July of 1999 (Perkins 2000). It has nearly tripled in size and since 2001 has overflowed into Stump Lake, a smaller terminal lake (Aber 2003). The outflow from Devils Lakes has led to the 6 m rise of Stump Lake and extensive flooding throughout the basin (Perkins 2000). This has led to the increase in number, size, and permanence prairie wetlands. For this study, the changes in the Basin present additional problems for identifying wetland depressions. NWI for the PPR in North Dakota was largely compiled in the early to mid 1980s, a time when the elevation of Devils Lake was lower and lakes and pothole wetlands less extensive. NWI data though still important for developing a processing mask and for validation had to be supplemented with more current wetland information.

Over 83 percent of the wetlands in PPR are smaller than 0.8 ha (9 TM pixels) and are generally not resolved by Level II USGS DEMSs. Even larger wetlands may not be resolved or appear only as flat areas in the USGS DEMs. Remaining depressions are often filled during standard conditioning procedures that create a depressionless DEM. Given these statistics, the accuracy of any hydrologic modeling in the region using normal methods is doubtful.

High resolution bare-ground LIDAR DEMs could potentially identify wetland features and is used in this study as raw elevation data. Selective filling of the DEM would make the DEM more realistic and therefore, derivative products would have greater validity.

Materials

Four basic datasets are employed in this study. A DEM for the Devils Lake Basin derived from LIDAR, National Inventory Data (NWI) for area, a Landsat 7 ETM+ image, and a 2003 color digital orthophoto mosaic for Ramsey County, ND.

Bare-Ground LIDAR of Devils Lake

A Digital Elevation Model (DEM) derived from LIDAR, provided the raw elevation data for the conditioning phase of application. The Devils Lake Elevation Mosaic was collected in 2000, revised in 2002 by FEMA contractor TerraPoint Inc. FEMA provided the Corps of Engineers – St. Paul District in 2004 with revised floating point LIDAR grids as quarter-quad areas after reprocessing the original bare earth data. The GIS Center, St. Paul, merged the quads together to produce a seamless data set in ESRI grid format, covering over 25.5 percent of the Devils Lake Basin. The dataset is distributed as 5-meter data with a stated 30 cm vertical accuracy. Projection is in UTM, Zone 14,
NAD 83. Missing data lines in the northern part mar the otherwise high quality dataset. These data can be downloaded from GIS Center Data Depository, St. Paul District. URL: http://www.mvp.usace.army.mil/gis/.

The LIDAR data has a much higher resolution than the existing, publicly available National Elevation Dataset (NED), formulated by the U.S. Geological Survey (USGS). A disadvantage of NED is that the resolution, the vertical accuracy, is anywhere from 1 to 3 meters and prone to errors especially in flat areas.

DEMs produced using LIDAR have the capability of capturing small variations in relative surface relief with a vertical accuracy of 0.1-0.2 meter, and could potentially resolve pothole features, particularly during dry cycles.

Hydrologic modeling with LIDAR has its own problems. LIDAR data must be carefully filtered to eliminate reflectances from vegetation and other features unrelated to surface topography. In many instances, "bare-earth" LIDAR data retains far too much detail for current hydrologic modeling software to manage over areas of any substantial size. Moreover LIDAR does not penetrate water. The high level of vertical accuracy of LIDAR technology means that LIDAR products contain enormous numbers of sinks, most being errors in the data. (Blackwell and Wells, 1999).

Blackwell and Wells (1999) presented noted that resampling 1-meter, bare-ground filtered LIDAR to 5 meter and 10 meter cells made the data in their Bellaire, Texas study area more useable. No resampling of the 5 meter Devils Lake LIDAR elevation dataset was done in the first testing of the data in this study.

National Wetlands Inventory Data

National Wetlands Inventory (NWI) data were used to help to identify depressional wetlands in the DEM. These data were downloaded as USGS quad-based shapefiles from the U.S. Fish and Wildlife National Wetland Inventory publishing website. Some 30 shapefiles were required to complete the coverage represented by the Devils Lake Elevation mosaic. These data were merged into a single NWI feature dataset. The primary attribute in the dataset are polygons codes that relate to Cowardin’s Classification of Wetlands and Deepwater Habitat, henceforth referred to as the Cowardin Classification. A major problem with NWI is the age of the data. Most NWI data for the PPR in North Dakota and South Dakota was acquired from 1977 to the late 1980s. While there were moderately wet years in the1980s, it was really due to the high precipitation in 1993 and the years that followed that Devils Lake rose, overflowed and flooded the basin. Wetland shapes and even classifications changed during drastically during this period. In this study, we depend on NWI data for counts of wetlands and have added Landsat TM to help establish size and shape of the larger wetlands. URL: http://www.fws.gov/nwi/
Landsat 7 ETM+ Image

Landsat 7 ETM+ is used in combination with NWI data to develop a wetland identification dataset. A partial 7-band Landsat TM scene (Path/Row 31/27) for May 20, 2003 was downloaded from the Upper Midwest Aerospace Consortium (UMAC) website. Band 5 was subsetted in preparation for density slicing. The 2003 year of the TM scene was relatively close to collection date of the LIDAR data and in May wetlands would be easily detected in Landsat as they would be filled with snowmelt and spring rains. URL: [http://www.umac.org](http://www.umac.org)

National Agriculture Imagery Program (NAIP) Data for Ramsey County, ND (1:12,000) for 2003

NAIP 2003, 1:12,000 scale data for Ramsey County are used for wetland identification and for visualization. These data are described as natural color digital orthophoto mosaics, corrected to UTM, Zone 14 NAD 83. They were provided to the North Dakota GIS hub by the National Resources Conservation Service (NRCS) and the North Dakota Geological. The address for North Dakota Geographic Information Systems website is: URL: [http://www.nd.gov/gis/mapsdata/download/](http://www.nd.gov/gis/mapsdata/download/)

Methods

The general methodology for the study is to: first to delineate a pilot basin, next identify sink features within the LIDAR DEM for the basin, locate wetland features using ancillary datasets, code real sinks (wetland depressions) in the LIDAR DEM, and produce a wetland mask. This mask is used in the selective filling of sinks.

Delineating a pilot basin

The Devils Lake Elevation Mosaic covers a reasonably good size portion of the substantial Devils Lake Basin. For the purpose of detailed assessment, a smaller dataset was required. Standard conditioning processes were run on the Devils Lake elevation mosaic. The processes include filling all sinks, developing a flow direction grid from the filled DEM, deriving a flow accumulation grid, and finally creating watersheds. A small watershed, 78.15 sq kilometers, located between several extensive lakes that avoided areas of the data with missing scan lines was chosen as a study site. Size, shape, and distribution of wetlands within this basin appear to be more representative of the PPR then in other parts of the Devils Lake Basin.

Identifying Sink Features in the DEM

Initially the original DEM, filled DEM, flow direction grid, and flow accumulation grids were clipped to the boundary of the pilot basin using ArcINFO extract by a mask tool. A difference grid, whereby the original LIDAR DEM grid, is subtracted from the filled,
depressionless version, was created for the pilot study area. This procedure is a common method of identifying sinks and described in Jenson and Dominque, Danielson, Gusman, Voigt, and Forman).

Next a sink mask, where all cell values were coded to 1 using a con statement, was derived from the difference grid. The ArcINFO RegionGroup algorithm was then applied to the sink mask to create unique sink regions. Some 30,324 sink groups were identified in the basin. Cell counts within sink groups varied from a single cell (25 sq meters) to as many as 278,185 cells (6.95 sq kilometers). Figure 2 displays the regiongroup grid and the syntax of GRID commands used in this phase of the study.

**Sinks in LIDAR DEM**

**Pilot Basin**

\[
\text{Difference Grid} = \text{Filled DEM} - \text{Original DEM}
\]

\[
\text{Sink Mask} = \text{Con} (\text{Difference} > 0, 1)
\]

\[
\text{Sink Regions} = \text{Region Group} (\text{Sink Mask by 8})
\]

\[
30,324 \text{ Sink Groups Identified}
\]

Figure 2: Sinks in the LIDAR- derived DEM for the pilot basin

**Attributing the Sink Mask**

The sink mask was fully attributed with data on maximum depth, average depth, volume, and surface area for each sink. Information on codes, counts, surface area of NWI wetlands were added as well. This gave us a dataset that could be manipulated, used, and assessed in a variety of ways and for many purposes including classification (Gusman et al., 2001).

**Locating Wetland Features**

Few sink features in the region group grid were indeed wetlands. But question is of course which cell groupings represent real wetlands. Most single cells were likely to be spurious sinks, errors in the data; yet given the small size of wetlands in PPR, at least to
begin with even single cells need further assessment. The large groupings of cells were thought to be flat areas or areas bounded by elevated road grades, which only under extreme conditions would water filled. Other data were needed to help identify wetland features.

Merged and clipped NWI data for the basin provided a basic count and classification of wetlands. NWI records 1,613 separate wetland polygons and 1,503 wetland basins in the pilot study area. Greater credibility was given to count of NWI wetlands, rather than size or shape, given changes taking place in the Devils Lake Basin since 1993. Summary of NWI wetlands for the study area are included on Table 1.

Table 1: NWI Wetland Polygons for Pilot Study Area: Attributes, Water Regime, Counts, and Hectares

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Water Regime</th>
<th>Counts</th>
<th>Hectares</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEMA, PEMAd, and</td>
<td>A - Temporary</td>
<td>1,080</td>
<td>522.08</td>
</tr>
<tr>
<td>PEMC, PEMCd, PEMCx, and PFO/EMCd</td>
<td>C - Seasonal</td>
<td>455</td>
<td>498.17</td>
</tr>
<tr>
<td>PABF, PABFx, PEM/ABF, and PEM/ABFd</td>
<td>F – Semi-Permanent</td>
<td>40</td>
<td>270.9</td>
</tr>
<tr>
<td>L2ABG and L2ABGd</td>
<td>G – Intermittently Exposed</td>
<td>4</td>
<td>295.85</td>
</tr>
<tr>
<td>R4SBF and R4SBFx</td>
<td>F – Semi-Permanent</td>
<td>2</td>
<td>27.8</td>
</tr>
</tbody>
</table>

Classification Source: Cowardin Classification, 1979

NWI wetlands in the basin totaled 1,793.24 hectares with an average size of 105.58 hectares. These wetlands constitute approximately 23 percent of the basin.

A Landsat TM image (partial scene) for May 20, 2003 with 30 m pixels was used to supplement the NWI data for the pilot basin. This study employed techniques similar to those reported by Sethre et al. (2005), who in part based their work on earlier studies by Best and Moore, 1981; and Frazier and Page, 2000. They noted wetland features could be extracted from band 5 with simple density-slice techniques with a high degree of accuracy.

For the first test, a threshold value of 59 was selected for the density slice, which was done in ERDAS Imagine. Extracted pixels ranging from 0 – 59 in value were coded to 1 using a Con statement and resampled to 5 meter pixels. Cells were then grouped into regions using the ArcGIS RegionGroup tool. Some 526 wetland groups were identified using this technique. Groupings ranged from 5 cells (.01 h) to 32,573 cells (814 h) with an average grouping of 1,062 (2.66 h) cells. Band 5 density-slice identified 1,397.81 hectares of possible wetland features or 18 percent of the basin. Figure 6 displays the NWI vector data and Landsat Band 5 density slice grids for pilot basin.
Creating a Wetland Mask

Initial step in creating the wetland mask involved combining NWI delineated wetlands with the band 5 density slice results. The NWI data were converted to a raster data with Features to Raster Tool. To combine NWI gridded data and the density-sliced dataset, it was necessary to use Isnull to change NoData values to a value. Isnull returned 1 if the input value was NoData and 0 for cells that were not NoData. When the isnull grids were added together, i.e., NWI isnull plus band 5 density slice isnull, the resulting codes were 0 where both codes were 0 or data, 1 where NoData (1) and data (0) were combined, and 2 for NoData cells in both inputs.

This combined dataset was reclassed so the 0 and 1s were both coded to 100,000 and 2 coded to NoData. This wetland mask was then added to region grouped sink data. The resulting values ranged from 130,300 to 100,002. This resulted in 6016 different cell groupings. Next a dataset was constructed where 100,000 was subtracted from the cell values of the sink plus wetland data.

The attribute from this dataset was exported to a dbf file. A field was added to this stand-alone table called wetlink. An arbitrary value of 10 was given to the field for all records.

Because of the difficulty of permanently joining tables to grids in ArcGIS, the dBASE table (dbf) was converted to an INFO table. In ArcInfo workstation, this INFO table was joined to the original sink mask.

After the table was joined, sink pixels were selected by the wetlink attribute, i.e. wetlink = 10, using the Extract by Attribute spatial analyst tool. The resulting dataset was the new sink mask, that will be used the selective filling of sinks in LIDAR dem. The sink mask, given the name wetland_mask, had 6,016 sink groups that correspond to wetland features from NWI and a band 5 density slice. While this grid dataset needed further refining to more closely correspond to NWI data and wetlands appearing on the NAIP 2003 DOQ, it was a starting point to selective filling of the LIDAR DEM.

Selective Filling of the DEM

Selective filling of the LIDAR DEM involved a number of steps, which translated into a sequence of ArcInfo GRID commands. Selective filling, normally run through an AML, was developed by scientists in Topographic Science Program at the USGS EROS data center, Sioux Falls, SD. Danielson (2000) detailed the process as used in developing a hydrologically correct one-kilometer DEM of Africa. We adapted these procedures to the selective filling of the LIDAR DEM for pilot study area. The principal difference in methodology was how the sink mask was derived. For the Africa research, sinks were selected according to size with a final threshold value set at 200 sq. kilometers. A threshold value of 1000 30-meter cells or 90 hectares is standard for EDNA products. In the PPR, all but the larger wetland depressions would be filled.
Procedures for selective filling were as follows with ArcInfo GRID commands capitalized. First, the minimum elevation were identified for each masked wetland. This lowest elevation location was taken as the point to which flow would be routed. ZONALMIN identified the lowest elevation. A CON statement created a locate grid where minimum elevation cells within the masked sinks were given a value of -9999 and all other cells were NoData. Assigning a value of -9999 allowed these cells to be selected out as a group later in the processing (Danielson, 2000).

Since the methodology treated each sink as a unique feature, the next processing step created a grid with distinct values inside and outside of each sink zone. ISNULL was in used combination with CON to produce a grid (DEM_type) based on the wetland_mask, wherever there was a sink zone, these areas were given original DEM values; otherwise the original filled elevation data were used (Danielson, 2000).

The next step of the processing involved merging the locates grid -9999 cells with DEM_Type, and selecting for all cell values not equal to -9999. In the resulting grid, the formally -9999 cells became NoData. This step was required to deceive the FILL algorithm into thinking this point was the edge of the data set or if continent, the ocean. In the second to last step, the new grid was then filled using the ArcInfo grid FILL function. As a result of the process whenever a NoData cell was encountered the elevations were raised around the point creating a concave feature in the DEM. When flow directions were computed, the cells in the concave feature sloped downhill in the direction of the NoData point, representing the minimum elevation in sink region.

To complete the process, the original DEM data values for minimum elevation in the sinks were reinserted into the now filled grid. Retaining NoData values in the DEM would cause problems with algorithms such as Flowdirection and Flowaccumulation. Figure 8 depicts the selectively filled DEM.

The steps above are summarized with the following ArcInfo GRID commands (Danielson, 2000).

Grid: Minimum_Elev = ZONALMIN (Wetland_Mask, Original_DEM)

Grid: Locates = CON (Minimum_Elev eq Original_DEM, -9999)

Grid: DEM_type = CON (ISNULL (Wetland_Mask), Filled_DEM, Original_DEM)

Grid: NoData_mask = CON (ISNULL (Wetland_Mask), DEM_type)

Grid: NoData_points = SELECT (MERGE (Locates, DEM_type) ‘value ne -9999’)

Grid: FILL NoData_points NoData_fill sink
Grid: \text{Merge\_DEM} = \text{MERGE (NoData\_fill, Original\_DEM)}

**Deriving Drainage Lines**

The final grid produced by the selective filling process was used to compute flow direction and flow accumulation. Drainage lines were derived from the flow accumulation grid. Various threshold values were tried, with best results being at flow accumulation value of 300 (.75 h).

**Results and Summary**

Drainage lines were overlain over subsets of the NAIP DOQ, depicting the different NWI classes of wetlands. Paired examples showed drainage derived from a DEM, processed without a wetland mask and drainage coming from a DEM, processed with a wetland mask. Figures 3 depicts drainage an intermittently exposed wetland. The filled DEM gave us information on trends of stream flow, while the selective filled DEM provided detailed wetland hydrography.

**Drainage Line Derivatives**

<table>
<thead>
<tr>
<th>Drainage No Mask</th>
<th>Drainage With Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Figure 3: Drainage lines for a R2SBFx wetland

In summary, some 30,324 sinks were initially identified in the LIDAR DEM. Some 6,016 sinks were possibly wetland features.
Disparity in wetland counts between NWI data and the wetland mask was problematic. Count difference most often occurred because a wetland might be made up of multiple cells groupings, although some sink groups contained more than one NWI wetland, some with 2 or 3. However, the accuracy of NWI was likewise in doubt due to age of the data and how it was produced. Landsat band 5 density slice result proved somewhat disappointing, in that it failed to resolve the smaller wetland features. This might in part be due the extraction method. Sethre et al. (2005) had good success with creating binary images consisting water and upland classes by sampling known water pixels and then classing pixels with digital numbers in that range as “water.” Remaining pixels were coded to “upland.”

The first use of the wetland mask to produce drainage lines showed promising results. The choice of threshold value in the flow accumulation grid to produce drainage lines might need to be adjusted and there was evidence that flow direction was significantly influenced by with diking effect of roads. Future work will focus on refining the wetland mask and setting up scenarios whereby depth and volume of sinks are additional considerations in the selective filling process.

References


Contact:

Janet H. Gritzner,
Department of Geography
South Dakota State University
Brookings, SD 57007

Janet.Gritzner@sdstate.edu