Pit-centric depression removal methods

John B. Lindsay§
Department of Geography, Environment, and Geomatics
The University of Guelph
50 Stone Rd. East, Guelph, ON, Canada, N1G 2W1
§ jlindsay@uoguelph.ca

Abstract—**Topographic depressions are problematic for digital elevation model (DEM) based flow-path modelling applications. Two new depression removal methods are presented in this paper, including a depression filling and a depression breaching algorithm. These new methods adopt an approach to depression removal that is in contrast to the simulated landscape flooding approaches that dominate current leading methods. Instead, these methods start by identifying pit cells to each depression, and apply flood-simulation (filling) and least-cost breaching only to a relatively small area of grid cells contained within (filling) or around (breach channels) depressions. Algorithm performance was tested using two large LiDAR DEMs of watersheds in Southwestern Ontario, Canada. Both of the new pit-centric methods were found to be between 5.0 and 10.7 times faster than the widely used algorithms based on simulated landscape flooding. In addition to computational efficiency, other benefits of the new methods are discussed.**

I. Introduction

Topographic depressions are bowl-like features of digital elevation models (DEMs) with undefined flow directions [1]. Depressions are particularly abundant in fine-resolution DEMs [1] and may represent actual landforms or can result from DEM error [2]. Road and rail embankments are a common source of artifact depressions in DEMs, particularly those derived from LiDAR, because of the inability of these data to represent sub-surface drainage infrastructure (e.g. culverts) and flow beneath bridges [3-5]. Depression are particularly problematic for geomorphometric applications involving modelled surface flow paths [6-7], including contributing area mapping, watershed mapping, stream mapping, and the calculation of wetness index, stream power, sediment transport index, and other common terrain attributes.

Numerous depression removal methods have been developed over the past three decades to automatically adjust elevations within depressions to ensure continuously defined flow paths (see [8] for a good review). Three broad approaches exist for depression removal, including filling, which raise elevations within depressions to their outlet heights, breaching, which carves channels from depression interior pit cells to outside points, and hybrid methods [9]. Ongoing development of these methods has been primarily driven by improved algorithm efficiency and reduced modification to the DEM.

While there have been numerous advancements in depression removal methods in recent years, the two most widely implemented removal methods, and therefore, the most commonly used in practice, include the techniques of Planchon and Darboux (P&D) [10] and the priority flood method, first proposed by Soille and Gratin [7] and popularized by Wang and Liu (W&L) [11] and Barnes et al. [8]. Many recent techniques are modifications of these two basic approaches [12-14]. Most of the advanced techniques for depression removal therefore operate by either shedding water from an inundated landscape toward its edges [10], or by progressively flooding the digital landscape from the data edges inward using the priority flood approach [7, 11].

Two novel depression removal techniques are presented in this paper, including a depression filling method and a least-cost depression breaching algorithm. These techniques adopt a pit-centric approach that does not require visiting each cell in the DEM and therefore has potential efficiency advantages over previous alternatives.

II. Methods

Depression removal techniques that are based on the P&D and W&L methods identify depressions as a bi-product of the simulated water shedding or flooding operations. Any DEM grid cells for which the elevation after the operation is higher than the input elevations are contained within depressions. Simulated water shedding and flooding are both relatively costly computational operations and their processing times scale with the total number of DEM grid cells. Because the number of cells contained within topographic depressions is always less (usually significantly so) than the total number of DEM cells, a method that explicitly identifies depressions, without the need for a broader simulation, could potentially be more efficient.

Jenson and Domingue [6] proposed a method (J&D) for depression filling that was based on identifying the pit cells (i.e. interior cells with undefined flow) contained within each depression and then mapping full depression extents.

Unfortunately, this early filling technique was computationally inefficient because the method for mapping depression extents relied on identifying each pit cell's watershed, the majority of which lies outside of the feature, and special handling of nested depressions, in the case of features containing multiple pit cells, was required. Rieger [15], Martz and Garbrecht [16, 17], and Lindsay and Dhun (L&D) [6] each proposed a pit-centric (i.e. a method based on identifying depression pit cells first) depression breaching method. The L&D technique was demonstrated to provide a low-impact solution that was particularly well suited to carving through embankments and predicting culvert/bridge locations. However, this method was also very computationally inefficient owing to the need to perform expensive cost-accumulation operations on each pit cell. Furthermore, the performance of the algorithm degrades significantly when larger search distances for breach paths are specified. These computational challenges with least-cost depression breaching were the motivation for Lindsay [18] proposing the 'efficient breaching' (EB) method, based on the W&L priority flood algorithm. While significantly improving computational performance, this method often provides much higher impact than EB and can yield unintuitive breach channels that follow the flood-order of the priority flood operation on which it is based.

The following two sections describe two new pit-centric depression removal methods. Both of the new depression removal methods, *FillDepressions* and *BreachDepressionsLeastCost*, have been implemented as tools within the open-source geospatial analysis software WhiteboxTools (WBT) [19]. Brief descriptions of each method are presented below and the source code of the tools is available for inspection online (https://github.com/jblindsay/whitebox-tools) for further detail.

A. Pit-Centric Filling Algorithm

The pit cells in a DEM can be mapped by identifying interior cells with no lower neighbors, an operation that is readily parallelized. Once pits are located, they are placed into a list and sorted from highest to lowest order based on cell elevations. The highest pit cell is popped from the list and a check is completed to determine whether the pit cell has already been raised in the output raster during a previous iteration. If the pit is unmodified, a region-growing operation is initiated to identify the pour point(s) and the depression interior of the feature to which the pit cell belongs. A priority queue, based on lowest cell heights, is initialized with the pit cell. All undiscovered neighbors are scanned and pushed into the queue if they are higher. Once a previously undiscovered neighboring cell of lower elevation (than the discovering cell's elevation) is located, the discovering cell is flagged as the depression's pour point. This priority-queue region-growing operation is only then terminated once there are no further cells of equal elevation to the outlet cell contained in the queue—this allows for the possibility of multiple pour-point cells. Each cell that was visited during the search for pour point cells are then raised to the pour point elevation. Importantly, this operation does not require visiting any cells contained outside of depressions (except for the single cell used to identify the pour-point cell) and there is no special handling required for nested depression. The region-growing operation raises the interior of topographic depressions to the level of their pour-point cells, leaving behind a flat surface. This flat surface may optionally have a slope applied to enforce flow across the extent of the depression. The flats-correcting algorithm applies another priority-queue based region-growing operation, initiated at outlets and with priority values set by the elevations contained within the original input DEM. This has the impact of forcing flow directions within depressions to follow similar patterns to the pre-flooding land-surface. This same procedure is iterated for each pit cell in the DEM.

The pit-centric depression filling technique identifies pit cells and their corresponding pour-point cells prior to filling. Therefore, it is possible to know a depression's depth in advance. If a depression's depth is greater than a user-specified maximum value, the feature can be left unfilled; this characteristic can be useful when working within closed geological basins. The P&D and W&L methods, in contrast, begin to fill features before their depths are known and can only mimic this property using post-processing.

B. Pit-Centric Least-Cost Breaching Algorithm

This method is inspired by the L&D least-cost breaching method and follows a similar method for identifying least-cost breach channels for each pit cell in the DEM. Here, the notion of cost is the height to which a cell's elevation would need to be lowered to ensure continuous flow from the pit cell along a breach channel. The iterative technique used by L&D for performing the cost-accumulation operation, needed to identify least-cost paths connecting pit cells to downstream cells exterior to depressions, has been replaced with a priority-queue based cost-accumulation operation that offers significant efficiency improvements. This new approach is in effect a region-growing operation that expands outward from each pit cell, from areas of lowest-cost to areas of increased cost, and accumulating the cost of breaching continuously descending channels to each newly discovered cell. The operation ceases once a cell with an elevation lower than the calculated breach channel height is identified. Importantly, this priority-based region-growing operation does not require a constant search window (unlike the L&D method) and will continue only until either a low enough breach channel end-point is located or no suitable target is identified within the constraints of a user-specified maximum breach distance or breach depth. The L&D method, by comparison, uses constant-sized search windows for each least-cost operation regardless of how far a breach end cell is from the pit, and therefore exhibits a relatively high computational cost. Because the vast majority of pits in a DEM can be resolved after relatively short-running region-growing search operations, this represents a significant speed up for this new breaching method. If any unresolved pits exist after the region-growing operation, because potential candidates for breach end cells exceed the specified maximum distance/depth, pit-centric depression filling may optionally be applied afterwards—thus, this method is a potentially hybrid solution.

C. Study Site and Data

Two test DEMs were used to evaluate the relative performance of the new pit-centric depression removal methods. Catfish Creek and Big Otter Creek are adjacent coastal watersheds draining to Lake Erie in Southwestern Ontario, Canada (Figure 1). The watersheds are dominated by agricultural land-uses, although there are small urbanized areas in each site. The physiography of the watersheds is composed of clay, sand, and till plains with an area of till moraines in the north [20]. Local relief is low, with elevations ranging from approximately 175 m at the outlets to 345 m in the headwater areas. The embankments of several major roads/railways transect both watersheds, which create apparent dams in the DEMs and extensive artificial topographic depressions.

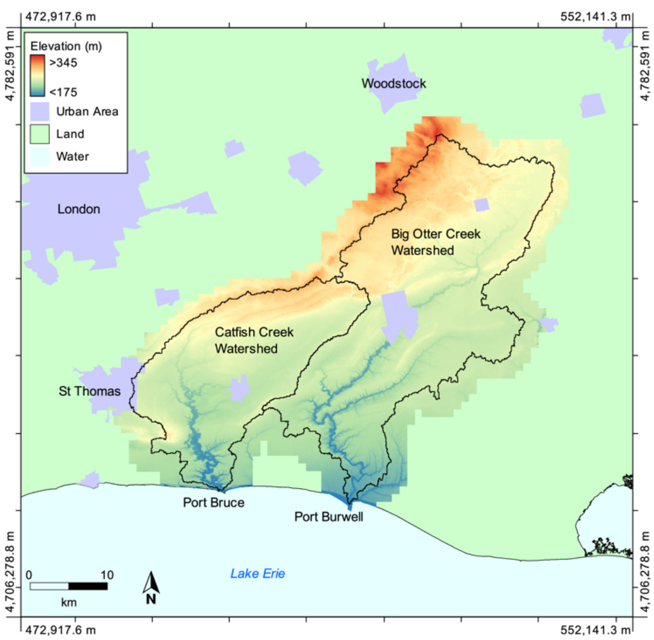

Figure 1. The test DEMs of the Catfish Creek and Big Otter Creek watersheds.

The test DEMs were interpolated using Delaunay triangulation from ground-classified returns of the source Lake Erie Watershed LiDAR point cloud data set. The source data were collected during leaf-off and snow-free conditions in the spring of 2018 by a private contractor commissioned by the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA) and the Ministry of Natural Resources and Forestry (MNRF). The average density of the data set is 8 points·m−2 and the overall vertical accuracy was estimated to be 0.05 m in unforested areas. The LiDAR data were stored in the NAD83 UTM zone 17N (EPSG:2958) coordinate system. Both DEMs were treated using the feature-preserving smoothing method [21], owing to the high degree of surface roughness common with fine-resolution LiDAR data. Table 1 summarizes the salient characteristics of the DEMs. Importantly, both DEMs are large data sets, with greater than one billion grid cells. Therefore, they provide suitable tests for algorithm efficiency and the practicality of the methods.

Table 1. Properties of the test DEMs for Catfish Creek and Big Otter Creek.

DEM Property	Catfish Creek	Big Otter Creek
Area (km^2)	685.0	1201.4
Resolution (m)	1.00	0.75
DEM file size[1] (GB)	4.55	15.29
Total cells in raster	1.138×10^{9}	3.822×10^{9}
Non-NoData cells	6.850×10^{8}	2.136×10^{9}
Pit cells	1.033×10^{6}	3.745×10^{6}
Depressions	5.054×10^{5}	1.633×10^{6}

[1] The input DEMs were stored as 32-bit floating point values, however, the output DEMs were 64-bit floats because fixing flat areas uses very small elevation increments. Thus, output DEM sizes were double the values reported here.

III. RESULTS AND CONCLUSIONS

The computational performance of the two pit-centric depression removal methods, were compared with other methods that are available in common geospatial analysis software (Table 2). Each of the tests were conducted using a computer system with a 3.0-GHz 8-core Intel processor and 64 GB of 1866-MHz memory. The P&D implementation in WhiteboxTools was used for testing, however, this is the same algorithm that is implemented in the widely used *Fill* tool in ArcGIS. The WhiteboxTools and SAGA GIS (*Fill sinks xxl*) implementations of the W&L depression filling method were both tested and were found to perform similarly. Lindsay's [15] EB method was the only depression breaching algorithm used for comparison. Being based on a priority-flood operation, this breaching method was found to have a broadly similar performance profile to the W&L algorithm. Attempts were made to test the L&D least-cost breaching method [5], however, this tool could not successfully process the large test DEMs.

The pit-centric filling and breaching tools were able to remove depressions in the Catfish Creek test site in less than one minute and the larger Big Otter Creek DEM in less than five minutes.

Both of the tools had similar efficiency, with the breaching solution providing the fastest solution for the Catfish Creek DEM, while pit-centric filling provided the fastest solution for the Big Otter Creek DEM. Notice that the pit-centric breaching tool was run using a maximum breach channel length (2048 cells) that was sufficiently long to ensure complete removal of depressions by breaching, with no subsequent filling required. Both of the new pit-centric methods were between 5.0 and 10.7 times faster than the P&D and the W&L algorithms (Table 1). The J&D method, one of the earliest depression removal techniques and implemented in GRASS GIS's *r.fill.dir* tool, was unsurprisingly the slowest tested method. The J&D Big Otter Creek test appeared suspended until the process was manually terminated after nearly 30 hours of processing. Unexpectedly, the tool did successfully produce a filled DEM after the process was terminated, and therefore, the reported time for this test is suspect.

Table 2. Depression removal algorithm performance. Processing times are in seconds and exclude file input/output. The two pit-centric techniques are compared to various software implementations of the Planchon & Darboux (P&D) filling, Wang & Liu (W&L) filling, Jenson & Domingue (J&D) filling, and Lindsay (2016) efficient breach (EB) methods.

Software / Algorithm	Catfish Creek	Big Otter Creek
WBT / Pit-centric filling	45.89	227.13
WBT / Pit-centric breaching	39.56	269.19
WBT / P&D	341.60	1340.97
WBT / W&L	300.50	1450.42
SAGA / W&L	422.08	1451.67
GRASS / J&D	6975.09	108,202.65
WBT / EB	324.62	1902.28

The pit-centric filling tool provided a very similar depression removal solution to all other filling methods (Figure 2B), although the way that this tool enforces slopes on flats using the pre-fill surface provided an improved flow accumulation pattern in flooded areas (see Figure 3A and Figure 3B). The pit-centric least-cost breaching tool was able to breach through embankments and dams at appropriate locations near stream crossings (Figure 2C), and therefore, was found to provide flow accumulation patterns that simulated flow through culverts (Figure 3C). Breach channels were found to be substantially shallower, shorter, and more appropriately located compared with the EB method. The differences were particularly evident where an embankment raises toward stream-road crossings to accommodate wide-diameter culverts. Under such conditions, EB can breach the road embankment long distances from the actual stream road crossing, while the pit-centric least-cost method typically follows a closer path to the mapped stream network (and presumed culvert location) in these cases.

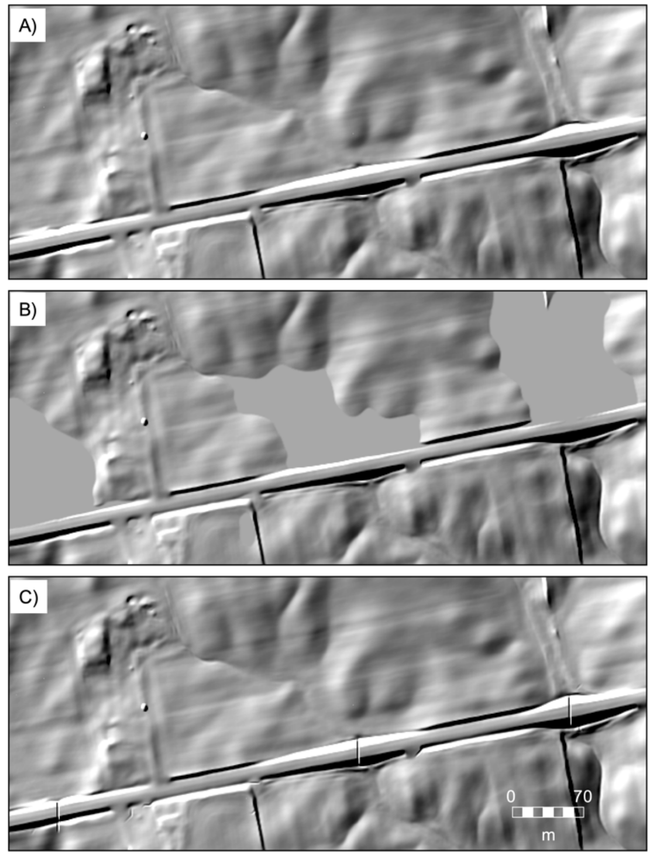

Figure 2. Hillshade images of a road embankment in the Catfish Creek Watershed, derived from A) the raw DEM, B) the pit-centric filled DEM, and C) the pit-centric least-cost breached DEM.

In conclusion, the two pit-centric depression removal methods presented in this paper were found to be significantly more efficient than methods that simulate landscape-wide flooding because they restrict the relatively costly simulation component to fewer grid cells in and around depression features. These methods can be used to resolve flow in multi-billion cell DEM data sets, offering the advantages of more natural flow enforcement on flats and least-cost breach channels, with the additional flexibility provided to users by allowing for maximum depression depth and breach channel length/depth parameters. Both pit-centric filling and breach methods offered very similar computational performance. Therefore, unless an application requires depression extents (i.e. sink mapping), the lower-impact breaching solution (*BreachDepressionsLeastCost*), which was found to resolve flow at stream-road crossings more naturally, should be preferred by practitioners. Lastly, both algorithms benefited from the parallelization of the pit discovery step, but were sequential for their more computationally expensive simulation steps. Future work may focus on further exploration of parallel version of these workflows.

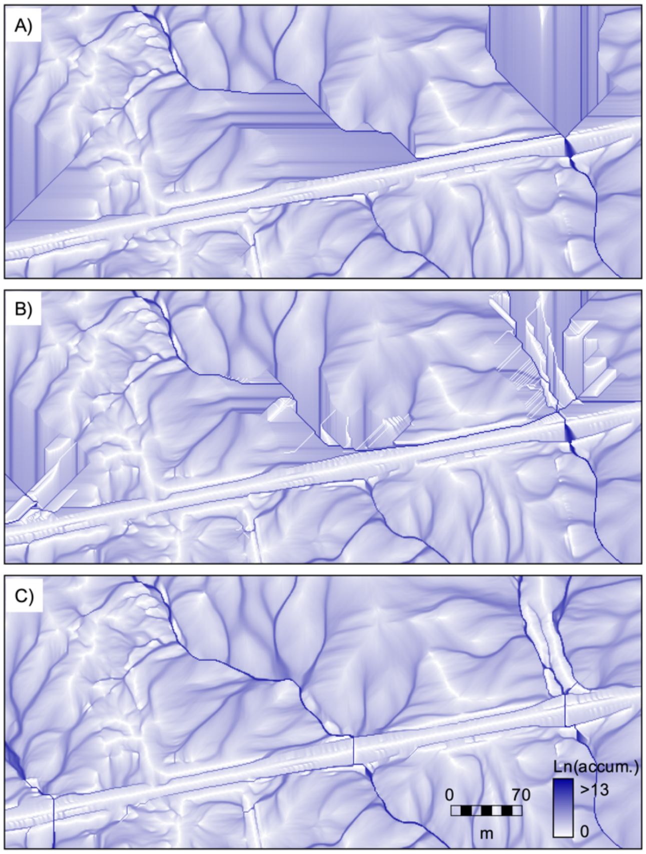

Figure 3. D∞ flow accumulation [22] rasters derived from DEMs treated using the A) Wang & Liu filling, B) pit-centric filling, and C) pit-centric least-cost breaching depression removal methods. The same road embankment found in Figure 2 is mapped here.

REFERENCES

[1] Lindsay, J.B. and Creed, I.F., 2005. "Sensitivity of digital landscapes to artifact depressions in remotely-sensed DEMs." Photogrammetric Engineering & Remote Sensing, 71(9), pp.1029-1036.

[2] Lindsay, J.B. and Creed, I.F., 2006. "Distinguishing actual and artefact depressions in digital elevation data." Computers & Geosciences, 32(8), pp.1192-1204.

[3] Duke, G.D., Kienzle, S.W., Johnson, D.L. and Byrne, J.M., 2003. "Improving overland flow routing by incorporating ancillary road data into digital elevation models." Journal of Spatial Hydrology, 3(2).

[4] Barber, C.P. and Shortridge, A., 2005. "Lidar elevation data for surface hydrologic modeling: Resolution and representation issues." Cartography and Geographic Information Science, 32(4), pp.401-410.

[5] Lindsay, J.B. and Dhun, K., 2015. "Modelling surface drainage patterns in altered landscapes using LiDAR." *International Journal of Geographical Information Science*, *29*(3), pp.397-411.

[6] Jenson, S.K. and Domingue, J.O., 1988. "Extracting topographic structure from digital elevation data for geographic information system analysis." Photogrammetric engineering and remote sensing, 54(11), pp.1593-1600.

[7] Soille, P. and Gratin, C., 1994. "An efficient algorithm for drainage network extraction on DEMs." Journal of Visual Communication and Image Representation, 5(2), pp.181-189.

[8] Barnes, R., Lehman, C. and Mulla, D., 2014. "Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models." Computers & Geosciences, 62, pp.117-127.

[9] Lindsay, J.B. and Creed, I.F., 2005. "Removal of artifact depressions from digital elevation models: towards a minimum impact approach." Hydrological Processes, 19(16), pp.3113-3126.

[10] Planchon, O. and Darboux, F., 2002. "A fast, simple and versatile algorithm to fill the depressions of digital elevation models." Catena, 46(2-3), pp.159-176.

[11] Wang, L. and Liu, H., 2006. "An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling." International Journal of Geographical Information Science, 20(2), pp.193-213.

[12] Wei, H., Zhou, G. and Dong, W., 2019. "A Variant of the Planchon and Darboux Algorithm for Filling Depressions in Raster Digital Elevation Models." ISPRS International Journal of Geo-Information, 8(4), p.164.

[13] Barnes, R., 2016. "Parallel Priority-Flood depression filling for trillion cell digital elevation models on desktops or clusters." Computers & Geosciences, 96, pp.56-68.

[14] Yong-He, L., Wan-Chang, Z. and Jing-Wen, X., 2009. "Another fast and simple DEM depression-filling algorithm based on priority queue structure." Atmospheric and Oceanic Science Letters, 2(4), pp.214-219.

[15] Rieger, W., 1998. "A phenomenon-based approach to upslope contributing area and depression in DEMs." Hydrological Processes, 12(6), pp.857-872.

[16] Martz, L.W. and Garbrecht, J., 1998. "The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models." Hydrological Processes, 12(6), pp.843-855.

[17] Martz, L.W. and Garbrecht, J., 1999. "An outlet breaching algorithm for the treatment of closed depression in a raster DEM." Computer & Geosciences, 25(7), pp.835-844.

[18] Lindsay, J.B., 2016. "Efficient hybrid breaching‐filling sink removal methods for flow path enforcement in digital elevation models." Hydrological Processes, 30(6), pp.846-857.

[19] Lindsay, J.B., 2019. "WhiteboxTools User Manual, v.1.1.0." Available online at https://jblindsay.github.io/wbt_book/intro.html (accessed Feb. 9, 2020).

[20] Chapman, L.J. and Putnam, D.F., 1973. "Physiography of Southern Ontario". Published for the Ontario Research Foundation by University of Toronto Press.

[21] Lindsay, J.B., Francioni, A. and Cockburn, J.M., 2019. "LiDAR DEM smoothing and the preservation of drainage features." Remote Sensing, 11(16), p.1926.

[22] Tarboton, D.G., 1997. "A new method for the determination of flow directions and upslope areas in grid digital elevation models." Water resources research, 33(2), pp.309-319.