Pruning of Hydrographic Networks: A Comparison of Two Approaches

Lawrence V. Stanislawski[1], Sandro Savino[2]

[1] Center of Excellence for Geospatial Information Science (CEGIS),
United States Geological Survey (USGS), Rolla, Missouri USA
Email: lstan@usgs.gov

[2] Department of Information Engineering, University of Padova, Italy
Email: sandro.savino@dei.unipd.it

Abstract: Hydrography is one of the most important themes in a map. As such, many different approaches to generalize it have been developed over the years. This paper describes two automated approaches to prune hydrographic drainage networks and metrically compares the pruned results. The first approach, referred to as stratified pruning, partitions the target network into line density classes (or strata) and separately prunes stream reaches from each stratum based on upstream drainage area, until the stratum target density is achieved. The second approach, referred to as length and density pruning, uses a stroke-based method to build river courses based on enriched geometric values (e. g., length, stream order, and branching) and available attributes for feature type and name. River courses are subsequently pruned based on length and density thresholds, with inflowing courses simultaneously removed. The two pruning approaches are applied to the flowline networks of two datasets from the high-resolution United States National Hydrography Dataset to produce 1:100,000-scale data. To evaluate and compare the results of the two algorithms, the pruned data are compared to the 1:100,000-scale National Hydrography Dataset using the coefficient of line correspondence. Results indicate the stratified pruning approach better maintained feature density variations caused by natural terrain conditions than the length and density pruning approach. Both methods appear equally capable of removing density variations caused by inconsistent data compilation.

Introduction

Generalization of cartographic data involves several operations categorized as cartographic and database generalization operations. Generally speaking, tasks concerned with abstraction (including selection) of cartographic features in a database are referred to as model generalization operations, whereas tasks that affect the visible display of features on a map are called cartographic generalization operations (Mackaness, 2007). Pruning is an operation that involves eliminating less important sections of a single feature, for instance, removal of lesser tributary parts of a polygon representing a lake or stream, or removal of parts of a compound feature, such as a surface water drainage network (Mustière et al., 2000; Stanislawski, 2009).

This paper compares two different approaches for pruning hydrographic networks: stratified pruning (SP) based on upstream drainage area (Stanislawski, 2009), and length and density pruning (LaDP) based on the "best continuation" principle (Savino et al., 2011a). The approaches are compared to ascertain the strengths and weaknesses of each approach, and identify possible future enhancements. The first approach was developed at the U.S. Geological Survey (USGS) Center of Excellence for Geospatial Information Science to automatically produce less detailed data from the high-resolution (HR) National Hydrography Dataset (NHD) to support multi-scale mapping and analysis. The HR NHD generally is compiled for use at 1:24,000-scale (24K), but is undergoing piece-meal revisions that densify the data in some areas for use at more detailed, local scales (Stanislawski et al., 2009). The second approach was developed at the Department of Information Engineering of the University of Padova, Italy, through a research project on generalization of 1:5000-scale (5K) data to 1:25000-scale (25K) data.

To compare the two network pruning strategies, parameters for each approach were trained to produce 1:100,000-scale (100K) data from 24K data, and tested on two sets of HR NHD data: a single subbasin dataset in Maine, and a four-subbasin dataset in Iowa. The objectives of pruning are to reduce data content in a cartographically appropriate manner that maintains feature-density variations that depict natural terrain variability while removing, where needed, variations caused by inconsistent data compilation. The 24K Maine dataset exhibits obvious feature density variations caused by inconsistent data compilation, whereas glacial processes form natural density variations in the Iowa data. Consequently, parameters were selected for the two approaches to best mimic the existing benchmark dataset, the 100K NHD, which does not exhibit data compilation variations in the test areas. Pruning results from the two strategies are compared to the associated 100K NHD through the coefficient of line correspondence [(CLC), Stanislawski, 2009] and bootstrapped confidence intervals for the CLC (Stanislawski et al., 2010). Detailed descriptions of CLC and associated confidence interval computations are given in Stanislawski (2009) and Stanislawski et al. (2010), respectively.

Related work

Due to its relevance, many algorithms have been developed for generalization of the hydrography network. Generalization of river networks involves different operators: typification may be necessary to generalize some parts having special patterns (Zhang, 2007; Savino et al., 2011b), vertex weeding should be applied to simplify the line geometries (Douglas and Peucker, 1973; McMaster, 1987; Wang and Muller, 1998), and selection is necessary to reduce the number of elements in the network. Most of the known approaches resort to pruning, or removing less prominent features of the hydrographic network. The main difference between the approaches is in the methods applied to estimate feature prominence.

One network pruning approach is to create a fine hierarchy and filter the data based on this classification (Richardson, 1994). Stream order, or number, is a common technique to assign a hierarchy to components of a river network (Horton, 1945; Strahler, 1952; Shreve, 1966). In a network each river is represented by one or more edges; each represents a section of the river. During selection it is important that a river is processed as a whole, and that single sections are not separately pruned, which may disconnect the graph. The information to reconstruct a river from the sections may be stored in the original data model (e. g., as the river name, or a unique identification code, such as a reach code), or else it is necessary to calculate it. One method that has been applied to road networks is the "best continuation" principle (Thomson and Brooks, 2000) derived from the Gestalt theory (Thomson and Richardson, 1999; Wertheimer, 1923).

Algorithms for river selection usually first enrich the source data model with feature prominence information that guides the selection process. Subsequently, the network is pruned based on one or more thresholds and metrics related to prominence, such as length (Thomson and Brooks, 2002; Brewer et al., 2009), watershed (Ai et al., 2006), order (Touya, 2007), density (Stanislawski et al., 2009; Savino et al., 2011a), or upstream drainage area (Stanislawski, 2009).

Methods

Table 1 summarizes the main aspect of the SP and LaDP hydrographic network pruning algorithms that are compared in this paper. Following sections describe the details of each method, the CLC metric used to compare pruned data to benchmark data, and the results of pruning tests performed on two datasets.

Table 1: Overview of the two hydrographic network pruning algorithms compared in this paper. [SP, stratified pruning; LaDP, length and density pruning]

Algorithm	Preprocessing	Parameter training	Pruning principle	Applied to
SP	Compute upstream drainage area Partition into density classes (areas with specific density ranges)	Estimation of local (within partition) target densities automatically through Radical Law, or manually from existing data	Remove less prominent rivers until target density reached Remove short dangling rivers	Partitions, Reach codes
LaDP	River segment enrichment Building of river courses	Manual Scale based	Remove short dangling river courses Remove river courses too “close” to others.	River courses (“best continuation” of river segments)

Stratified pruning (SP)

In stratified pruning, features in the hydrographic drainage network (NHD flowline network) are enriched with upstream drainage area (UDA) estimates that are assigned using an augmented directed graph approach and Thiessen polygon-derived catchment areas (Stanislawski, 2009). UDA values estimate feature prominence. The network is then subdivided into strata, or partitions, based on line density. A raster-based partitioning algorithm is applied, which uses predetermined density class breaks to define the partitions (Stanislawski and Buttenfield, 2011). After enrichment, a target density is determined for each partition using a best available estimate, which may be derived through a variation of the Radical Law (Töpfer and Pillewizer, 1966) based on stream length, or determined from existing data. In this paper, target densities for each partition were estimated from the existing 100K NHD network features within each partition. Subsequently, less prominent segments composing full reaches (confluence-to-confluence segments with similar feature types) are iteratively pruned from each partition based on a minimum UDA tolerance. The UDA tolerance is increased through an iterative process until the target density is achieved. Upon pruning each partition, some reaches may be added back into the pruned dataset to maintain proper connectivity along partition boundaries. After initial pruning, dangling tributaries that are retained but too short to be included in the dataset, as determined from NHD data standards (USEPA and USDOI, 1999), are removed.

Length and density pruning (LaDP)

In 2006 the Department of Information Engineering at University of Padova, Italy, started a research project, funded by the Regione Veneto, to develop algorithms for the automated generalization to form the 25K Italian national geodatabase from the regional 5K data. The LaDP process, developed to generalize hydrography, borrows ideas from the literature, and also adopts some original solutions for the selection of the rivers.

The process is comprised of three main steps: the reconstruction of the river courses, the selection of short rivers, and the pruning in dense regions of the network. The first step is a preprocess that groups the edges of the hydrography graph into river courses in a manner similar to the construction of "strokes" on a road network (Thomson and Brooks, 2000). Selection algorithms operate on the full river courses and will not delete single edges of the graph, which avoids the possibility of disconnecting a part of the network.

A river course can be defined as a group of consecutive edges that runs from a source to either a sink or a confluence into another river course, following the water flow direction; each edge belongs to one river course only, and a river course may comprise one or more edges. Similar to strokes, river courses are built by trying to find, for each edge, the consecutive edge that mostly "resembles" it; both semantic data and numeric quantities are evaluated at each fork to find the best continuation of the river course. Semantic data include river class and river name, which exist as attributes for each network edge in the source data model. Numeric quantities are Strahler order (S), the length to furthermost source (L), the number of branches uphill (B), and the number of edges uphill (N). These quantities are not present in the source data model but are calculated for each edge with a top-down enrichment process starting from each source and following the flow direction. For bifurcations, all braids inherit B, L, and N of the uphill edge, and S is not increased if two braids from the same river join downstream.

Once values have been computed for each network edge, river courses are built with a bottom-up approach. Starting from the sink having the highest Strahler order, the algorithm goes "uphill" (i.e., opposite to flow direction), deciding at each fork the branch that is the best continuation of the river course. The semantic and numeric data are weighted to bias the choice toward the branch having the same river name, the same Strahler order, the same river class, similar width (if present), the longest path to the furthermost source uphill, and the greatest number of branches uphill. The process stops when the river course arrives at its most uphill point, which may be another river course (in the case of a bifurcation) or a source; the course inherits the maximum value of S, L, B, and N of the associated edges. Next, a new river course is built from another sink (or confluence), and this process continues until every edge of the graph is part of one river course.

River courses are pruned based on the enriched length and density values. Each river course shorter than the minimum length threshold is pruned, along with every river course stemming from it.

Pruning by density requires a measure of how close each river course is to nearby river courses. This is done by building a buffer around each river course and calculating the area of overlap with buffers around nearby river courses. The ratio of the overlap area over the total buffer area is used as a measure of local density around each river course. If this density is larger than a density threshold, the river course is a candidate for pruning. The algorithm relies on the enriched values to decide whether or not a candidate is actually deleted. Thresholds for S, L, B, and N constrain pruning to only those river courses that are not a prominent part of the network. Thresholds are set in a manner that avoids deletion of rivers that are long, have many branches, or a high Strahler order.

The density pruning algorithm evaluates whether or not each candidate is prominent enough to be kept; in the latter case it prunes a candidate river course and all courses stemming from it, updating the percentage of overlap area of the nearby river courses. This process continues until all remaining river courses have densities below the density threshold or are too relevant to be removed, meaning the network cannot be pruned any further.

The original algorithm was developed to generalize data to 25K; nevertheless, by modifying the threshold values (i.e., buffer radius, density threshold, and length threshold), the algorithm can be adapted to generalize data to other scales. For this paper, parameters have been trained to generalize data to the 100K based on NHD data standards (USEPA and USDOI, 1999).

Comparison to benchmark: coefficient of line correspondence (CLC)

The CLC is a measure of the quality of conflation between two sets of line data. It was developed as an adaptation of Taylor's (1977) coefficient of area correspondence, but it applies to linear features. It estimates how well the two line datasets, representing the same set of features, match (Stanislawski, 2009; Buttenfield el al. 2010). An accepted standard dataset representing features at or near the target scale typically is used as a benchmark dataset for CLC computations, which can help validate and refine generalization procedures. CLC values range from 0 for complete mismatch to 1 for a perfect match. For this paper, the 100K NHD flowlines are used as the benchmark for comparing pruning results. A non-parametric bootstrapping approach is used to generate a confidence interval for each CLC value (Stanislawski et al., 2010).

Test on Maine subbasin with inconsistent collection

The first test site is the Lower Penobscot subbasin in Maine, where HR NHD data density variations are apparent because of inconsistent compilation (figure 1 c). The subbasin includes 13,000 HR NHD flowlines and covers nearly 6,130 square kilometers (km^2).

Flow-directed HR flowlines were pruned to 100K using SP with target densities in each partition modeled to fit the density of 100K NHD data. Four line-density partitions were determined for the HR flowlines using a raster partitioning process. Average line-densities for the four HR partitions were 0.62, 1.41, 2.42, and 4.24 kilometers per square kilometer (km/km^2) for the low- to high-density partitions. Target densities modeled to achieve the 100K densities were 0.70, 0.87, 0.78, and 0.71 km/km^2, respectively for the low- to high-density partitions. These densities account for simplification differences between the two scales and subsequent removal of retained dangling tributaries that are too short for the 100K.

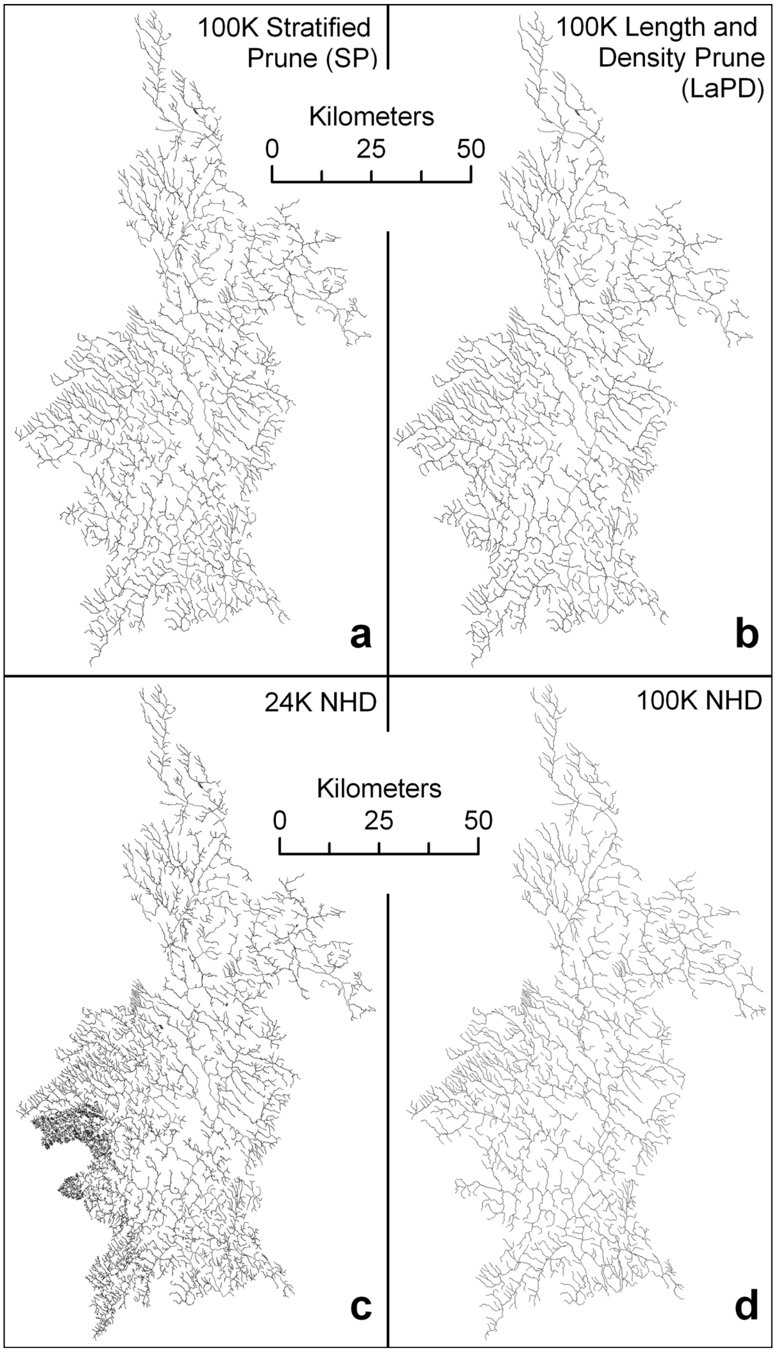

Figure 1. High-resolution National Hydrography Dataset (NHD) flowline features that remain after pruning the 1:24,000-scale (24K) Lower Penobscot subbasin in Maine to 1:100,000-scale (100K). Results of stratified pruning (SP) are shown in panel a, and length and density pruning (LaDP) results are shown panel b. The original, unpruned 24K NHD is shown in panel c, and the 100K NHD is shown in panel d for comparison.

New parameters to generalize to 100K were chosen manually for the LaDP algorithm by comparing a small part of the HR dataset with the corresponding 100K NHD data, and with reference to the 100K NHD data standards (USEPA and USDOI, 1999). The minimum length chosen was 1,600 meters (m) with a buffer radius of 1,200 m for density pruning and an allowed maximum density of 50 percent. Candidates for selection were not pruned if their length (L) was longer than 3,200 m, or had order (S) larger than 3, or number of uphill branches (B) bigger than 16, or more than 10 uphill arcs (N).

Pruning results of the two approaches are displayed in figure 1, along with the source HR NHD and benchmark 100K NHD flowlines. Subbasin CLC values are presented in table 2, and bootstrapped 90 percent confidence intervals for the subbasin CLC values are presented in table 3. Results from the two approaches remove local density variations caused by inconsistent data compilation, and seem to match the 100K data with a similar level of quality. The distribution of weighted CLC values for the 204 6.31-by-6.31 km cells covering the subbasin is shown in figure 2. Cell values are weighted to eliminate subbasin edge effects (Stanislawski et al., 2010). The SP approach provides relatively better correspondence in the areas with high density, whereas relatively greater correspondence appears evenly distributed over the subbasin with the LaDP approach.

Table 2. Sums of cell values for weighted CLC cell values and weighted omission and commission errors for the 204 6.31-by-6.31 km cells covering the Lower Penobscot HR NHD subbasin. Sums are shown for the two tested pruning methods.
[CLC, coefficient of line correspondence; HR, high-resolution; NHD, National Hydrography Dataset]

Pruning method	Weighted CLC	Weighted omissions	Weighted commissions
Stratified	0.779	0.080	0.141
Length and Density	0.775	0.130	0.095

Table 3. Bootstrapped 90 percent confidence interval for the weighted CLC for the two pruning methods applied to the Lower Penobscot HR NHD subbasin.
[CLC, coefficient of line correspondence; HR, high-resolution; NHD, National Hydrography Dataset]

	Bootstrapped 90 Percent Confidence Interval		
Pruning method	Lower bound	Mean	Upper bound
Stratified	0.763	0.779	0.796
Length and density	0.757	0.775	0.793

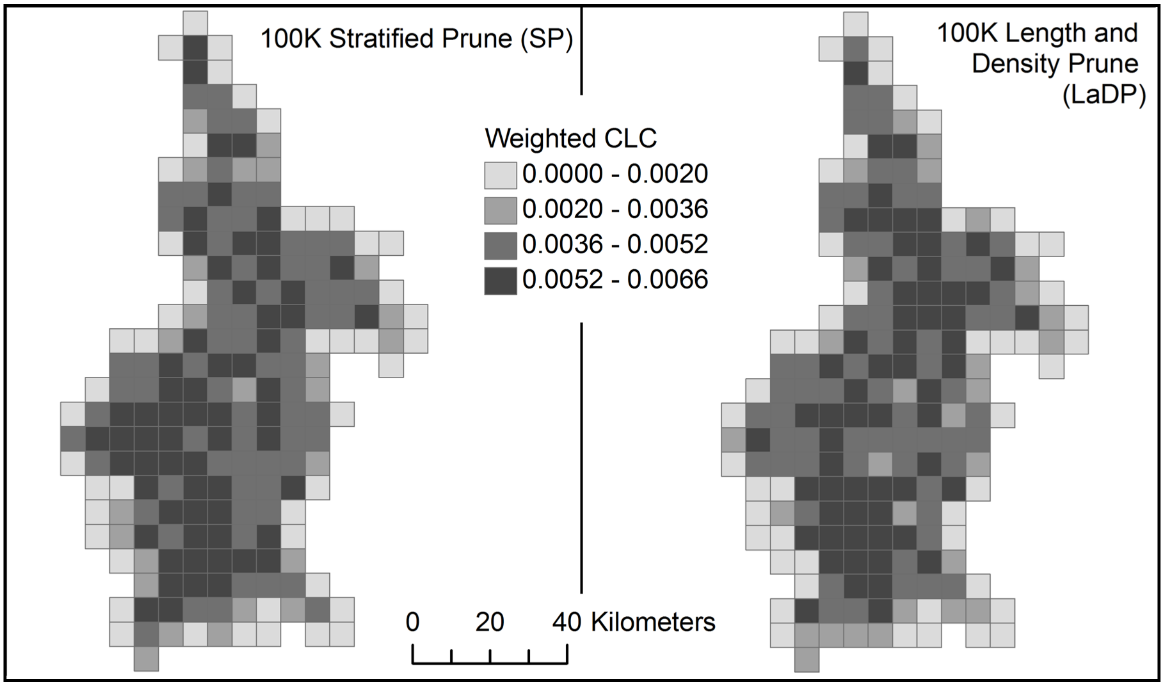

Figure 2. Distributions of weighted coefficient of line correspondence (CLC) values for the 204, 6.31 by 6.31 km cells covering the National Hydrography Dataset (NHD) Lower Penobscot subbasin in Maine. Weighted CLC values measure the amount of matching line features between the high-resolution NHD flowlines pruned to the 1:100,000-scale (100K) and the benchmark 100K NHD flowlines. Greater values (darker shades) represent a better match between features in the two data sets. Results from stratified pruning are shown on the left, and results from length and density pruning are on the right.

Test on four-subbasin glaciated area in Iowa

HR NHD flowlines from a four-subbasin area in central Iowa were used for the second test dataset. This study area in the midwest United States covers about 20,172 km^2 and includes the drainage area for the Raccoon River and the Middle Des Moines River upstream from the confluence with the Raccoon River. The four subbasins, containing more than 22,000 flow-directed HR flowlines, straddle two physiographic regions where a glacial lake borders a till plain. The hydrography shows a clear distinction between glaciated areas and the area known as the dissected till plains.

As with the Maine test case, the flow-directed HR flowlines were pruned to 100K through SP, fitting target densities for each partition to densities of the 100K NHD data. In this case, only two line-density partitions were determined for the HR flowlines using the raster partitioning process. Average line-densities for the two HR partitions were 0.40 and 1.21 km/km^2 for the low- and high-density partitions. Target densities modeled to achieve the 100K densities were 0.39, and 0.94 km/km^2, respectively for the low- and high-density partitions. These target densities account for subsequent removal of retained dangling tributaries that are too short for the 100K. In this case, target densities were not adjusted to account for different granularities in the source and benchmark datasets because doing so tends to generate a larger proportion of commission than omission errors, as seen in the Maine dataset (table 2).

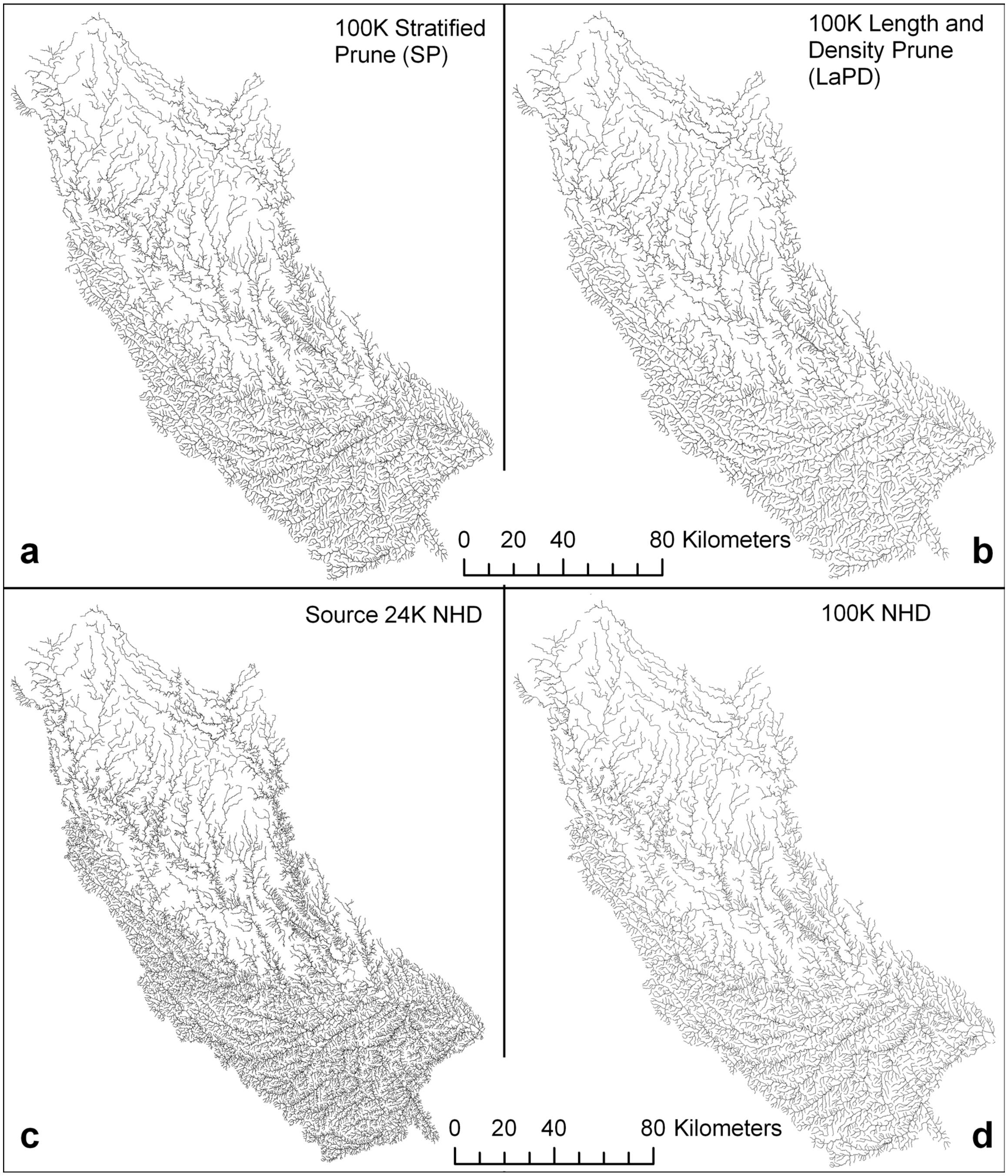

Figure 3. High-resolution National Hydrography Dataset (NHD) flowline features that remain after pruning four Iowa subbasins to the 1:100,000 (100K) level of detail. Stratified pruning (SP) results are shown in panel a. Length and density pruning (LaDP) results are shown in panel b. For comparison, the unpruned, source 1:24,000-scale (24K) NHD flowlines are shown in panel c, and the 100K NHD is shown in panel d.

The LaDP approach was developed with scale-dependent parameters in mind and presently (2011) considers only a single set of parameters for each scale. Consequently, the LaDP was applied to the Iowa data using the same parameters that were applied to the Maine subbasin, which helps assess the effects on the different datasets.

Pruning results for the two approaches are displayed in figure 3. Subbasin CLC values are presented in table 4, and bootstrapped 90 percent confidence intervals for the subbasin CLC values are presented in table 5. The weighted average CLC from the SP approach was 0.91, which is significantly greater, at the 90 percent confidence level, than the 0.84 average produced by the LaDP approach. This indicates the pruned data from SP approach matches the 100K NHD better than the results from LaDP approach. The distribution of weighted CLC values determined for the 200 44.92-by-44.92 km cells covering the subbasin is shown in figure 4. High CLC values are solidly distributed over the subbasin with the SP approach, whereas the LaDP approach shows slightly lower values distributed over the high-density area. This result indicates the LaDP approach may diminish some natural density variations that should be retained in hydrographic networks more so than the SP approach.

Additional tailoring of the parameters could improve LaDP results for the Iowa dataset, but may worsen results for the Maine dataset. Extensive processing was required to enrich the Iowa data (only 2 hours for the Maine dataset but several hours for Iowa) through the LaDP method, which is necessary to test each modification of parameters. On the other hand, enrichment is performed once for the SP approach, and required about 2 hours for Iowa and 1 hour for Maine. Various alternatives for density partitions and target density values can be evaluated relatively quickly (about 10 to 15 minutes for each alternative) with the SP approach.

Table 4. Sums of cell values for weighted CLC and weighted omission and commission errors for the 200 44.92-by-44.92 km cells covering four HR NHD subbasins in Iowa. Sums are shown for the two tested pruning methods.
[CLC, coefficient of line correspondence; HR, high-resolution; NHD, National Hydrography Dataset]

Pruning method	Weighted CLC	Weighted omissions	Weighted commissions
Stratified	0.914	0.048	0.038
Length and Density	0.841	0.140	0.019

Table 5. Bootstrapped 90 percent confidence interval for the weighted CLC for the two pruning methods applied to four HR NHD subbasins in Iowa.
[CLC, coefficient of line correspondence; HR, high-resolution; NHD, National Hydrography Dataset]

	Bootstrapped 90 Percent Confidence Interval		
Pruning method	Lower bound	Mean	Upper bound
Stratified	0.907	0.914	0.921
Length and Density	0.830	0.842	0.853

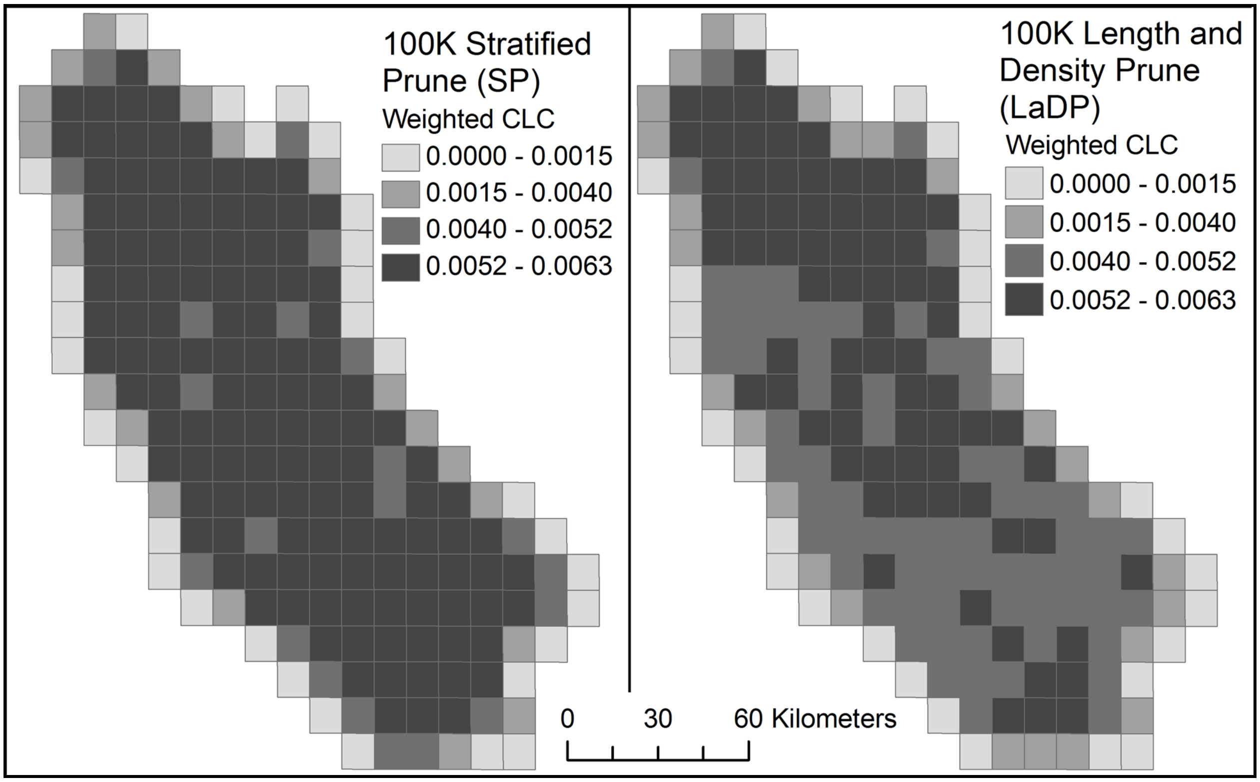

Figure 4. Distributions of weighted coefficient of line correspondence (CLC) values for the 200, 44.92 by 44.92 km cells covering four National Hydrography Dataset (NHD) subbasins in Iowa. Weighted CLC values measure the amount of matching line features between the high-resolution NHD flowlines pruned to the 1:100,000-scale (100K) and the benchmark 100K NHD flowlines. Greater values (darker shades) represent a better match between features in the two data sets. Results from stratified pruning are shown on the left, and results from length and density pruning are on the right.

Summary and future work

This paper compared the SP and LaDP approaches for pruning hydrographic networks from the 24K to the 100K level of detail. The approaches were tested on two United States NHD datasets: one with data inconsistencies that should be removed, and one with natural density variations that should be retained. The 100K NHD was used as a benchmark for comparison, and the CLC was used to metrically evaluate how well the pruned results from each approach matched the benchmark. Results indicate the two approaches worked equally well at removing data inconsistencies, but the SP approach was slightly better at retaining density variations that depict natural terrain differences. In addition, for these two datasets, enrichment processing is faster through the SP approach than the LaDP approach. The SP approach includes ancillary automated tools for evaluating various partitioning and pruning alternatives relatively quickly; however, arriving at appropriate target densities for the SP approach remains a topic for further research.

The LaDP approach is tailored to maintain the full extent of river courses, whereas the SP approach may eliminate less prominent sections of river courses that are deemed prominent by the LaDP approach. Further analysis comparing how well these two approaches prune networks to less detailed target scales, such as 1:500,000 or smaller, may indicate the LaDP approach provides the more cartographically appealing result. The SP approach may need to be enhanced to better consider inclusion of full river courses when pruning to small scales, similar to the LaDP approach.

Some insight gained for the LaDP approach is to consider the possible use of collinearity (Touya, 2007) to create better-looking river courses (avoiding some bad angles), which may affect pruning and possibly CLC results. The "one size fits all" idea is wrong; the same parameters are not able to generalize different areas. Although it is good that parameters are scale dependent (set once for every target scale), the results clearly indicate that the algorithm should be able to adapt its behavior to the structure (especially the density) of the original data. Furthermore, importance is relative; the thresholds that determine whether a candidate for river course deletion is relevant or can be pruned, could be "dynamic." Depending on how "seriously" the river course is in conflict (in overlap) with nearby rivers, the thresholds could be "relaxed" to delete river courses that otherwise would be deemed too important to be pruned.

References

Ai, T., Liu, Y. and Chen, J. (2006) The hierarchical watershed partitioning and data simplification of river network, in Progress in Spatial Data Handling Part 11, p.617-632, doi:10.1007/3-540-35589-8_39.

Brewer, C.A., Buttenfield, B.P. and Usery, E.L. (2009) Evaluating generalizations of hydrography in differing terrains for *The National Map* of the United States, in Proceedings of 24th International Cartographic Conference, 15-21 November 2009, Santiago, Chile.

Buttenfield, B.P., Stanislawski, L.V. and Brewer, C.A. (2010) Multiscale representations of water: Tailoring generalization sequences to specific physiographic regimes. GIScience 2010, 14-17 September 2010, Zurich, Switzerland.

Douglas, D. and Peucker, T. (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, in Cartographica: The International Journal for Geographic Information and Geovisualization, 10, p.112-122.

Horton, R.E. (1945) Erosional development of streams and their drainage basins; Hydrophysical Approach to Quantitative Morphology, in Bulletin of the Geological Society of America, 56, p.275-370.

Mackaness, W.A., (2007) Understanding geographic concepts, in: Mackaness, W.A., Ruas, A., and Sarjakoski, L.T., (eds.), Generalisation of Geographic Information: Cartographic Modelling and Applications. Elsevier for International Cartographic Association, p.1-10.

McMaster, R.B. (1987) Automated line generalization, in Cartographica, 24 (2), p.74-111.

Mustière, S., Saitta, L. and Zucker, J. D., (2000) Abstraction in cartographic generalization, in: Raś, Z. W. and Ohsuga, S. (eds.): ISMIS '00 Proceedings of the 12th International

Symposium on Foundations of Intelligent Systems, Springer-Verlag London, UK 2000, p.638-644.

Richardson, D. (1994) Generalisation of spatial and thematic data using inheritance and classification and aggregation hierarchies, in Advances in GIS Research 2, Taylor and Francis, p.957-972.

Savino, S., Rumor, M., Canton, F., Langiù, G., and Raineri, M. (2011a) Model and cartographic generalization of the hydrography network in the CARGEN project. 25th Conference of the International Cartographic Association, 2011, Paris, France.

Savino, S., Rumor, M. and Zanon, M. (2011b) Pattern recognition and typification of ditches. 25th Conference of the International Cartographic Association, 2011, Paris, France.

Shreve, R.L. (1966) Statistical law of stream numbers. Journal of Geology, 74, p.17-37.

Stanislawski, L.V. (2009) Feature pruning by upstream drainage area to support automated generalization of the United States National Hydrography Dataset, in Computers, Environment and Urban Systems, 33, p.325-333.

Stanislawski, L.V. and Buttenfield, B.P. (2011) A raster alternative for partitioning line densities to support automated cartographic generalization. 25th International Cartography Conference, 3-8 July 2011, Paris, France.

Stanislawski, L.V., Buttenfield, B.P., Finn, M.P. and Roth, K. (2009) Stratified database pruning to support local density variations in automated generalization of the United States National Hydrography Dataset, in Proceedings of 24th International Cartographic Conference, 15-21 November 2009, Santiago, Chile.

Stanislawski, L.V., Buttenfield, B.P. and Samaranayake, V.A. (2010) Generalization of Hydrographic Features and Automated Metric Assessment Through Bootstrapping in 13th Workshop of the ICA commission on Generalisation and Multiple Representation, 12-13 September 2010, Zurich, Switzerland.

Strahler, A.N. (1952) Hypsometric (area-altitude) analysis of erosional topography, in Geological Society of America Bulletin, 63 (11), p.1117-1142.

Taylor, P.J. (1977) Quantitative methods in geography: an introduction to spatial analysis, Chapter 5: Areal association. Houghton Mifflin, Boston, p.396.

Thomson, R.C. and Brooks, R. (2000) Efficient generalization and abstraction of network data using perceptual grouping, in Proceedings of the 5th International Conference on GeoComputation: Greenwich, August 23-25, online: http://www.geocomputation.org/2000/GC029/Gc029.htm. accessed 3 June 2011.

Thomson, R.C. and Brooks, R. (2002) Exploiting perceptual grouping for map analysis, understanding and generalization: The case of road and river networks, in Graphics Recognition, Algorithms and Applications, p.148–157.

Thomson, R.C. and Brooks, R. (2007) Generalization of geographical networks, in Generalisation of Geographic Information: Cartographic Modelling and Applications, William A. Mackaness, Anne Ruas, L. Tiina Sarjakoski (eds.), Elsevier Science, p.255-267.

Thomson, R.C. and Richardson, D.E. (1999) The 'good continuity' principle of perceptual organisation applied to the generalisation of road networks, in Proceeding of the 19th International Cartographic Conference, Ottawa, Canada, p.1215-1225.

Töpfer, F. and Pillewizer, W. (1966) The principles of selection: a means of cartographic generalization, in The Cartographic Journal 3(1), p.10-16.

Touya, G. (2007) River network selection based on structure and pattern recognition, in Proceedings of the 23rd International Cartographic Conference, 4-9 August 2007, Moscow, Russia.

USEPA and USDOI (1999) Standards for National Hydrography Dataset, United States Environmental Protection Agency and United States Department of the Interior, United States Geological Survey, National Mapping Program Technical Instructions, July 1999. http://rockyweb.cr.usgs.gov/nmpstds/acrodocs/draft/dlg-f/nhd/NHD0799.PDF (accessed 24 Feb. 2011).

Wang, Z., and Muller, J.C. (1998) Line generalization based on analysis of shape characteristics, Cartography and GIS, 25(1), p.3-15.

Wertheimer, M. (1923) Laws of organization in perceptual forms. First published as Untersuchungen zur Lehre von der Gestalt II, in Psycologische Forschung, 4, p.301-350. Translation published in Ellis W. (1938) A source book of Gestalt psychology, p.71-88.

Zhang, Q. (2007) Qingnian drainage typification based on dendritic decomposition, in The Cartographic Journal 44(4), p.321-328.