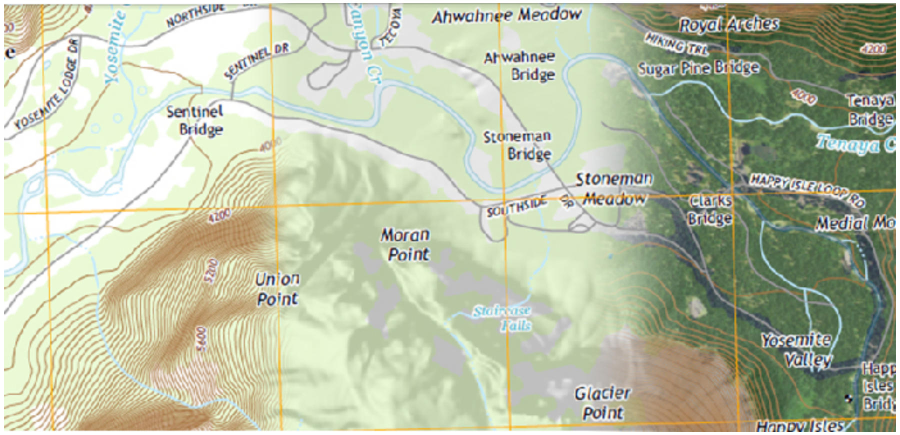
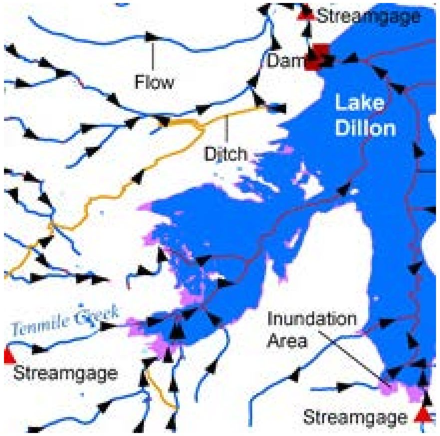
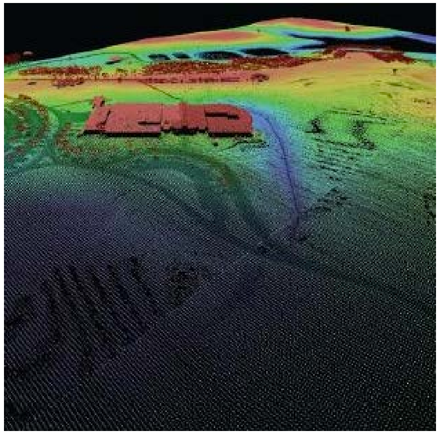
Generalizing Linear Stream Features to Preserve Sinuosity for Analysis and Display: A Pilot Study in Multi-Scale Data Science







Lawrence V. Stanislawski[1], Barry J. Kronenfeld[2], Barbara P. Buttenfield[3], Ethan Shavers[1], and Tyler Brockmeyer[1]

[1]U.S. Geological Survey, Center of Excellence for Geospatial Information Science, lstan@usgs.gov, eshavers@usgs.gov, tbrockmeyer@usgs.gov

[2]Eastern Illinois University, Department of Geology and Geography, bjkronenfeld@eiu.edu

[3]University of Colorado-Boulder, Department of Geography, babs@colorado.edu

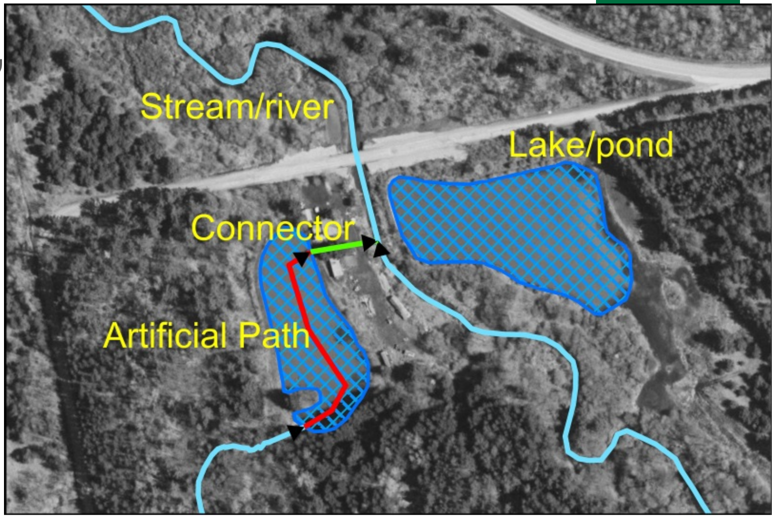
USGS science for a changing world

The National Map Your Source for Topographic Information

+ Outline

- Motivation
- Methods
- Results
- Summary

+ The National Hydrography Dataset (NHD)



- NHD is a database of vector features (points, lines, polygons) that represent surface water features in the United States.
- NHD Flowline includes the linear flow network with features of type stream/river, canal/ditch, pipeline, artificial path, and connector.
  - Artificial path represents a flow path through a polygonal water feature that is connected to other flowline features.
  - Connector represents a path where surface flow is known to exist but was not included in the source material.
  - Flowline features are oriented, where possible, in the direction of surface water flow.
- Within conterminous U.S., NHD compiled from 24K or larger source material

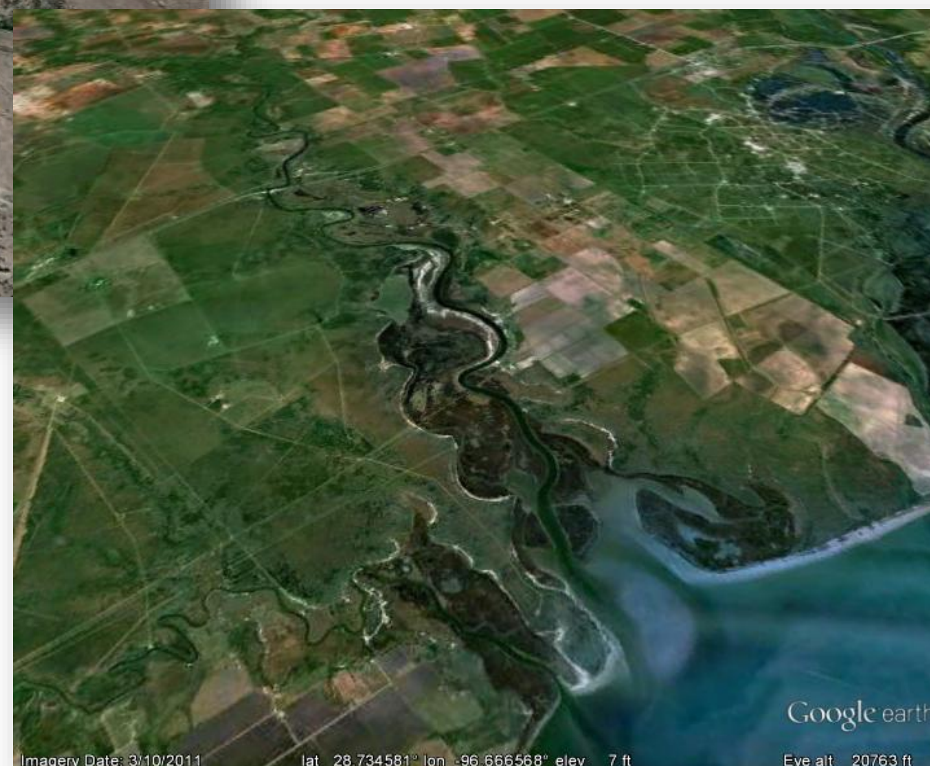
+ Problem Statement

- Cartographic generalization reduces the content and details of feature representations to reduce clutter and improve legibility and display speed.
- Regional and smaller scale analysis is feasible with generalized data, but
- Simplification and other generalization operations can impact data and analysis regarding stream geometric or geomorphic conditions, particularly at smaller scales.

Stream sinuosity varies in different landscapes

Rough terrain

Flat terrain



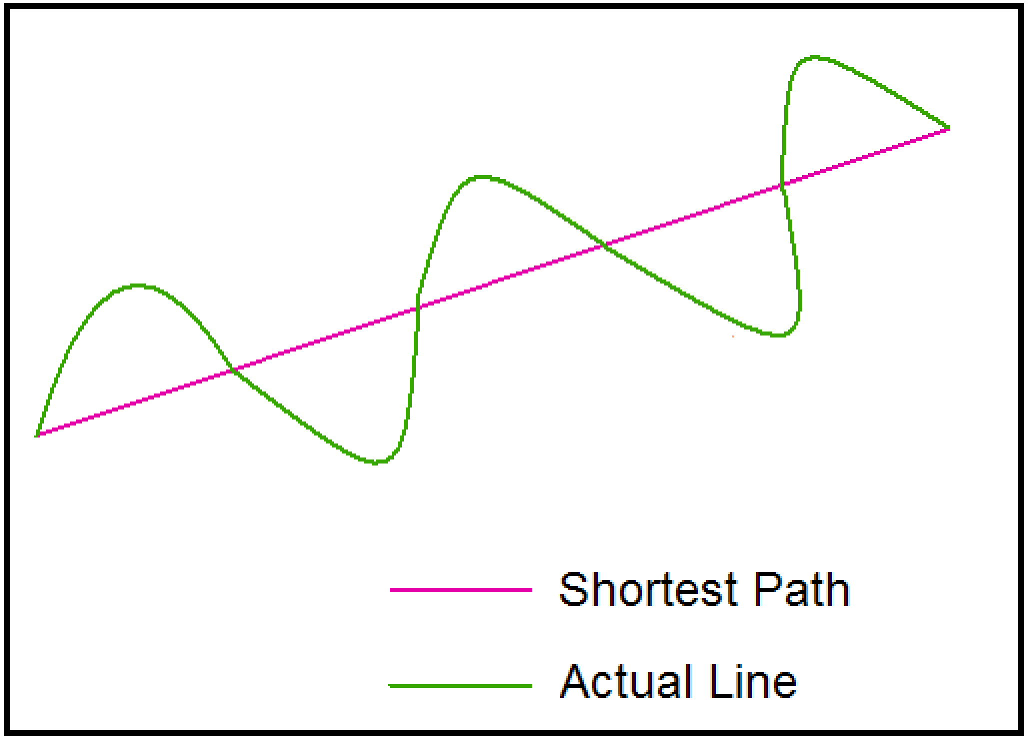
USGS The National Map

Sinuosity

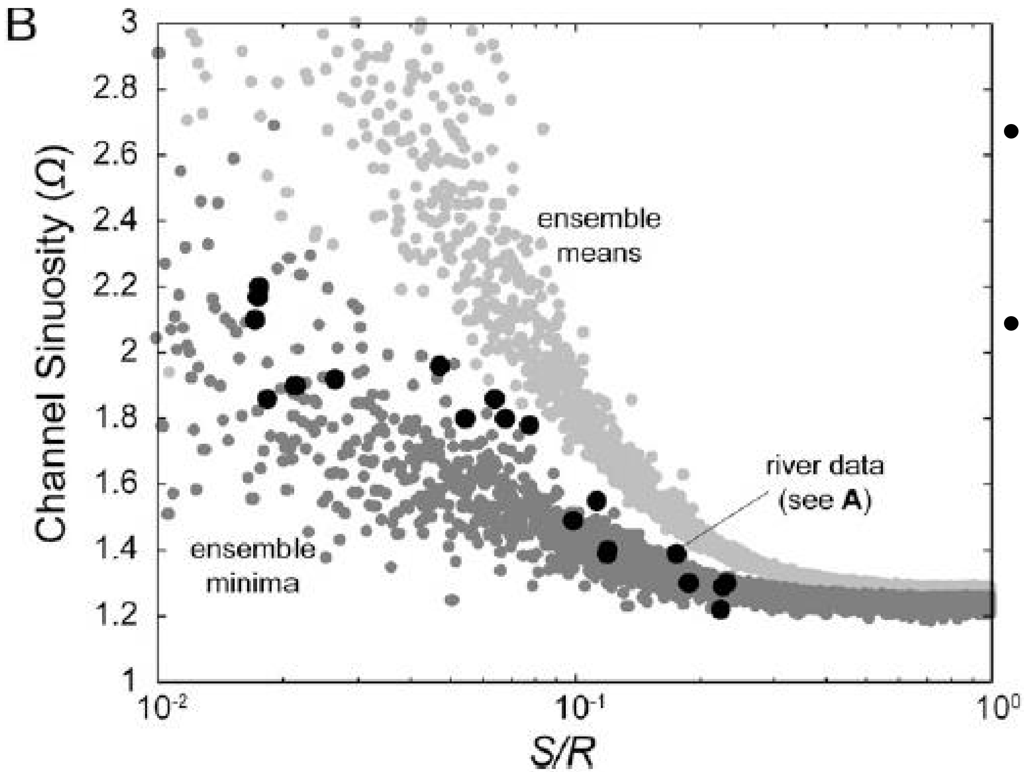
Sinuosity is defined as the length of a line (feature) divided by the distance between its endpoints.

Straight line sinuosity =
    1;
else
    sinuosity > 1

Residual sinuosity is sinuosity - 1



+ Stream sinuosity relation to geomorphic conditions



- **Within local conditions, channel stability, shape, and sinuosity are influenced by sediment load.**
- **Over the course of the river channel, critical thresholds in sediment load and slope alter a channel's pattern, which cause variations in channel sinuosity**

Schumm (1973)

As shown above, Lazarus and Constantine (2013) propose generic theory for channel sinuosity whereby sinuosity is related to landscape slope relative to flow resistance (topographic roughness and vegetation).

- Lazarus E D, and Constantine J A, 2013, Generic theory for channel sinuosity, Proceedings National Academy of Sciences 110(21): 8447-8452.
- Schumm S A, 1973, Geomorphic thresholds and complex response of drainage systems. Fluvial Geomorphology 6:69-85.

+ Stream classification system that includes sinuosity

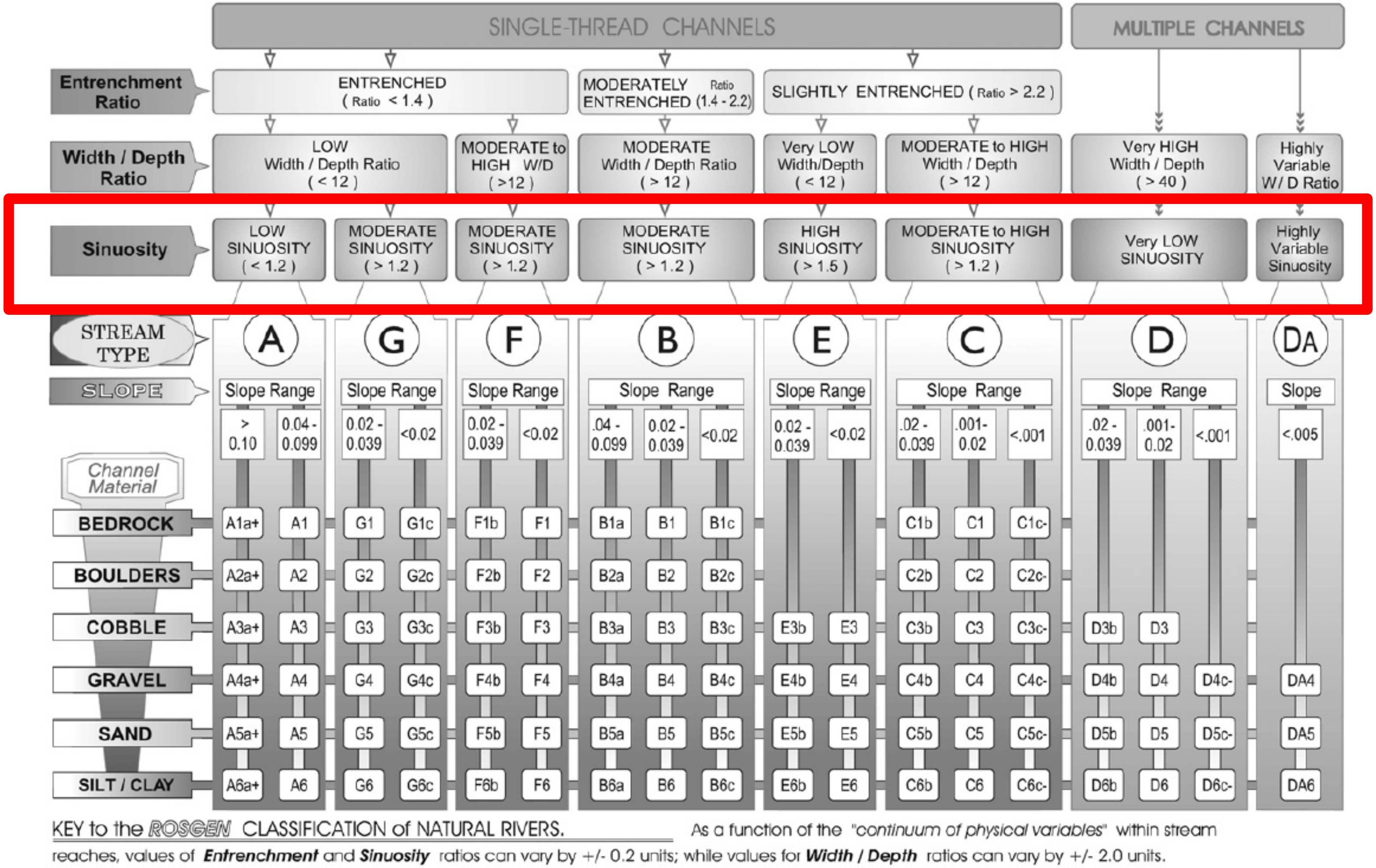


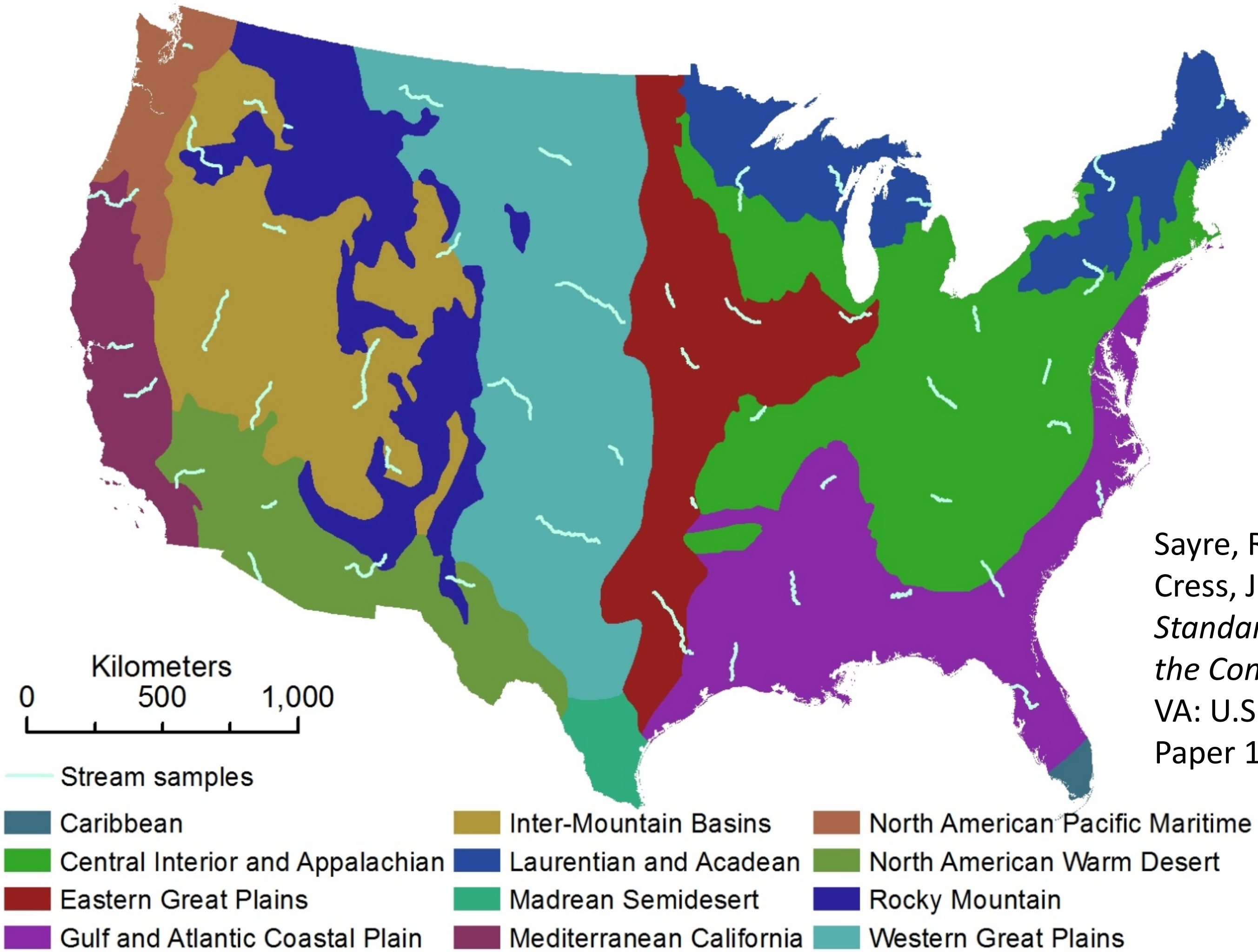
Figure 12 Rosgen (1994, 1996b) stream type classification. Reprinted with permission from Figure 5.3 in Rosgen, D.L., 1996b. Applied River Morphology. Wildland Hydrology, Pagosa Springs, CO.

Figure 12 from Buffington, J. M.; Montgomery, D. R. 2013. Geomorphic classification of rivers. In: Shroder, J.; Wohl, E., ed. Treatise on Geomorphology; Fluvial Geomorphology, Vol. 9. San Diego, CA: Academic Press. p. 730-767.

+ Objectives

- Evaluate methods to simplify linear stream features in a manner that more accurately retains the geometric variability (along with positional accuracy of the data)
  - Preserve sinuosity
  - Minimize horizontal (positional) and areal displacement

Methods: Test data 50 stream segments of NHD HR flowlines



- Over 15,000 flowline features
- Nearly 310,000 vertices
- Distributed in 10 of 12 ecological divisions

Sayre, R., Comer, P., Warner, H., and Cress, J. (2009). *A New Map of Standardized Terrestrial Ecosystems of the Conterminous United States.* Reston VA: U.S. Geological Survey Professional Paper 1768, 24pp.

USGS The National Map

+ Methods: Comparison of five line simplification algorithms

With Steiner points (points not on the original feature)

- Raposo Spatial Means (RSM)
- Kronenfeld’s Area-Preserving Segment Collapse (APSC)

Without Steiner points (retains subset of original vertices)

- Wang and Muller’s Bend-Simplify (BS)
- Ramer-Douglas-Peucker (RDP), and
- Visvalingam Effective-Area (VIS)

+ Methods: Simplify all features to a scale-dependent number of points

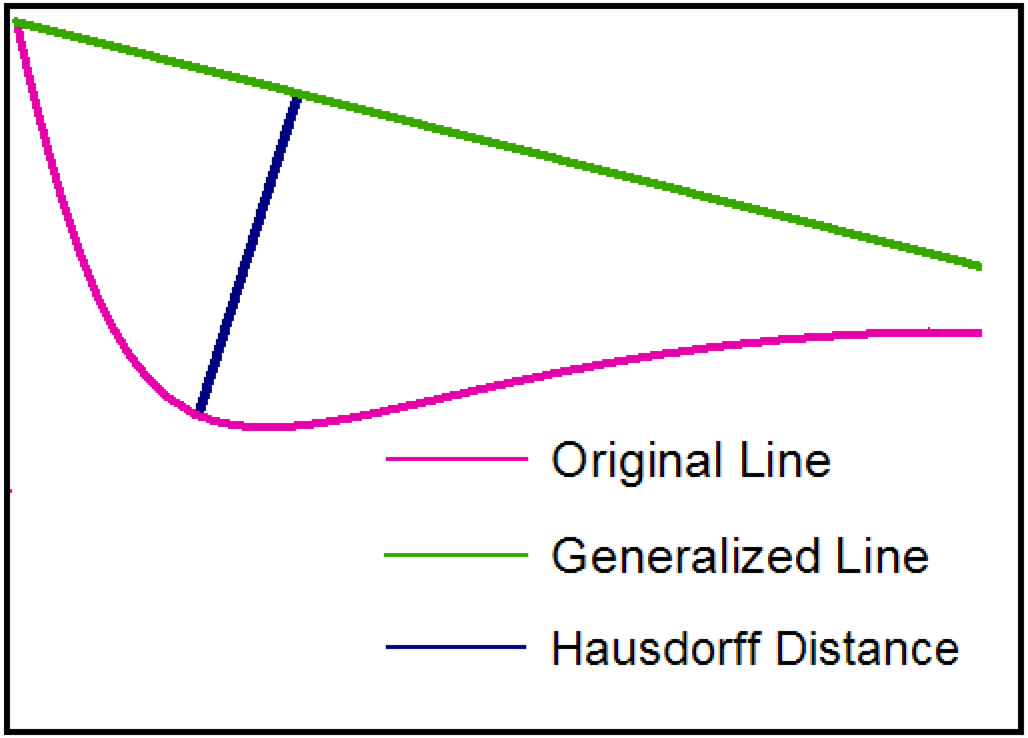
	legibility constraint		NMAS
Scale denominator	0.4	mm	0.02 inches in meters
5,000	2	meters	4.2
10,000	4	meters	8.5
24,000	9	meters	12.2
50,000	20	meters	25.4
100,000	40	meters	50.8
250,000	100	meters	127.0
500,000	200	meters	254.0
1,000,000	400	meters	508.0
2,000,000	800	meters	1016.0
5,000,000	2000	meters	2540.0
10,000,000	4000	meters	5080.0
20,000,000	8000	meters	10160.0

- Simplification was constrained by legibility, which was estimated as the diagonal of a pixel of a 19-inch monitor, or 0.4 mm.
- Simplify features to 12 levels of detail using Bend-Simplify algorithm with legibility constraint as tolerance.
- Number of retained points was determined for each BS simplified feature at each tolerance.
- Simplify features with the four other algorithms constraining each simplified feature to the number of points retained by BS algorithm.

Horizontal positional displacement: Hausdorff Distance

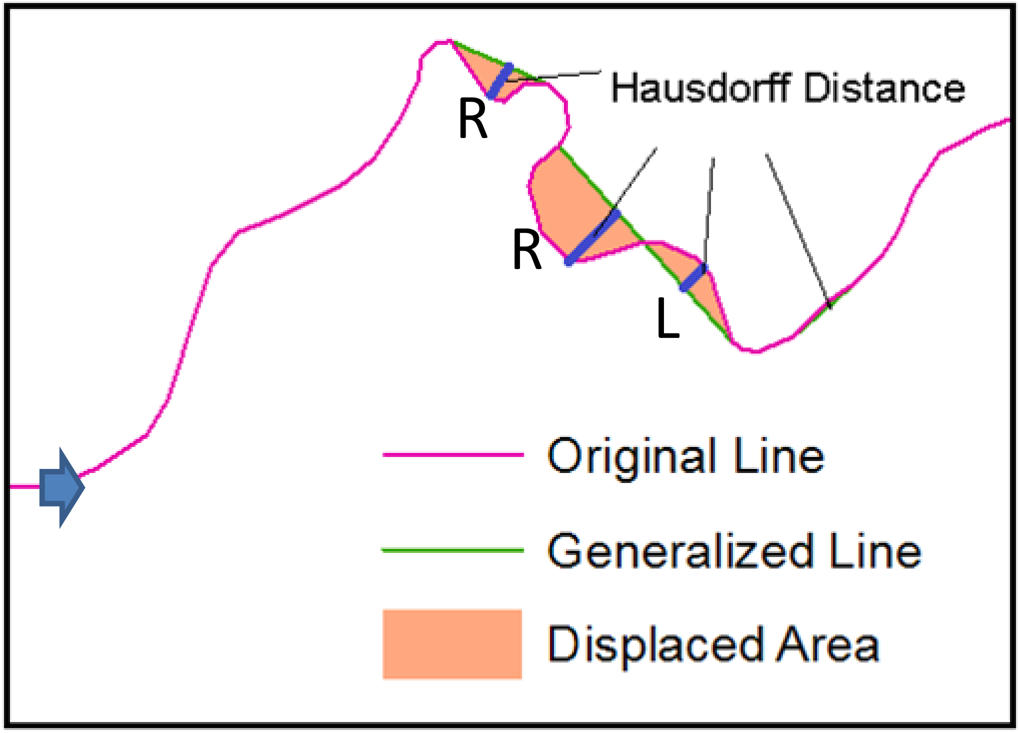
Defined as the maximum of the shortest distances between any point on the source feature and any point on the generalized feature.

This is the largest distance that a feature or segment has been moved in the simplification process.

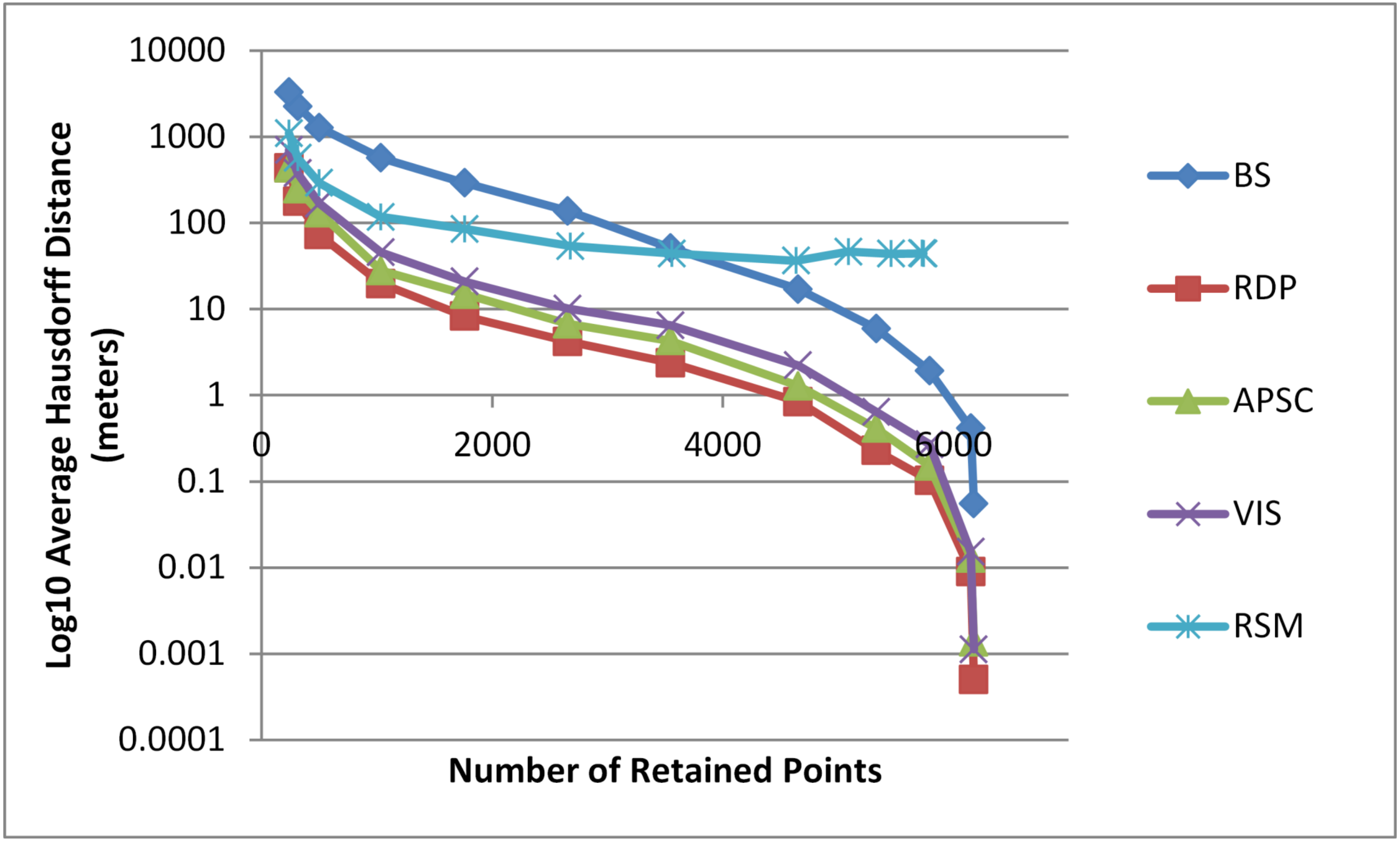


Positional and Areal Displacement

Vector displacement is the average of the Hausdorff distances of each displacement polygon for a feature.



Results: Modified Hausdorff Distance



Smaller scale

Larger scale

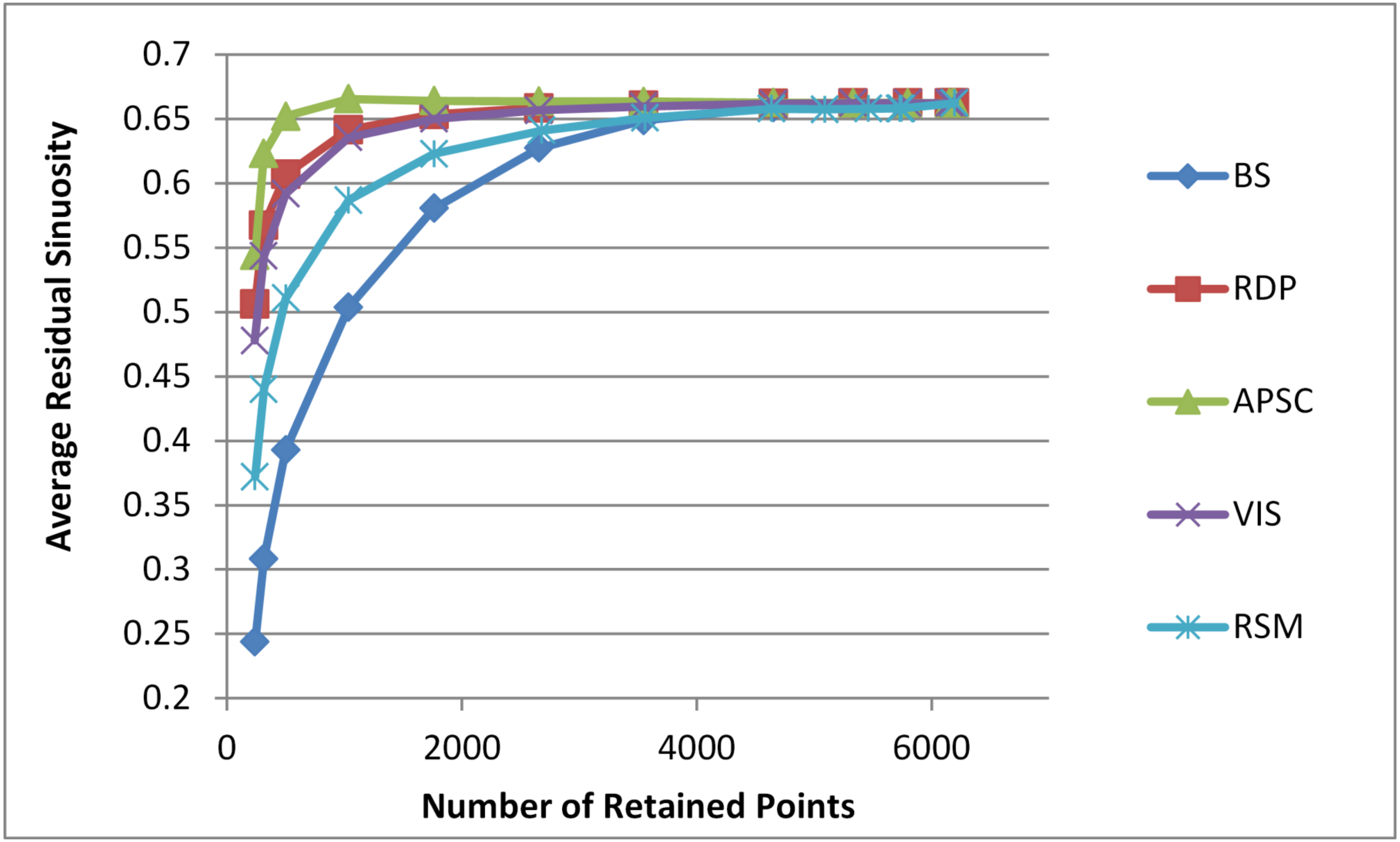
Results: Modified Hausdorff Distance

Paired different t-test with Holm's reduction for multiple hypothesis testing

Confidence Level: 95%										
Hausdorff										
Compare to:	BS				RDP			APSC		VIS
LOD	RDP	APSC	VIS	RSM	APSC	VIS	RSM	VIS	RSM	RSM
1:5000				TRUE			TRUE		TRUE	TRUE
1:10000	TRUE	TRUE	TRUE	TRUE			TRUE		TRUE	TRUE
1:24000	TRUE	TRUE	TRUE	TRUE			TRUE		TRUE	TRUE
1:50000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:100000	TRUE	TRUE	TRUE		TRUE	TRUE	TRUE		TRUE	TRUE
1:250000	TRUE	TRUE	TRUE		TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:500000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:1000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:2000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:5000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE		TRUE	TRUE
1:1000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:2000000	TRUE	TRUE	TRUE	TRUE		TRUE	TRUE	TRUE	TRUE	TRUE

Holm S. (1979) A simple sequentially rejective multiple test procedure. Scand. J. Stat 6(2): 65-70.

Results: Residual Sinuosity



Smaller scale Larger scale

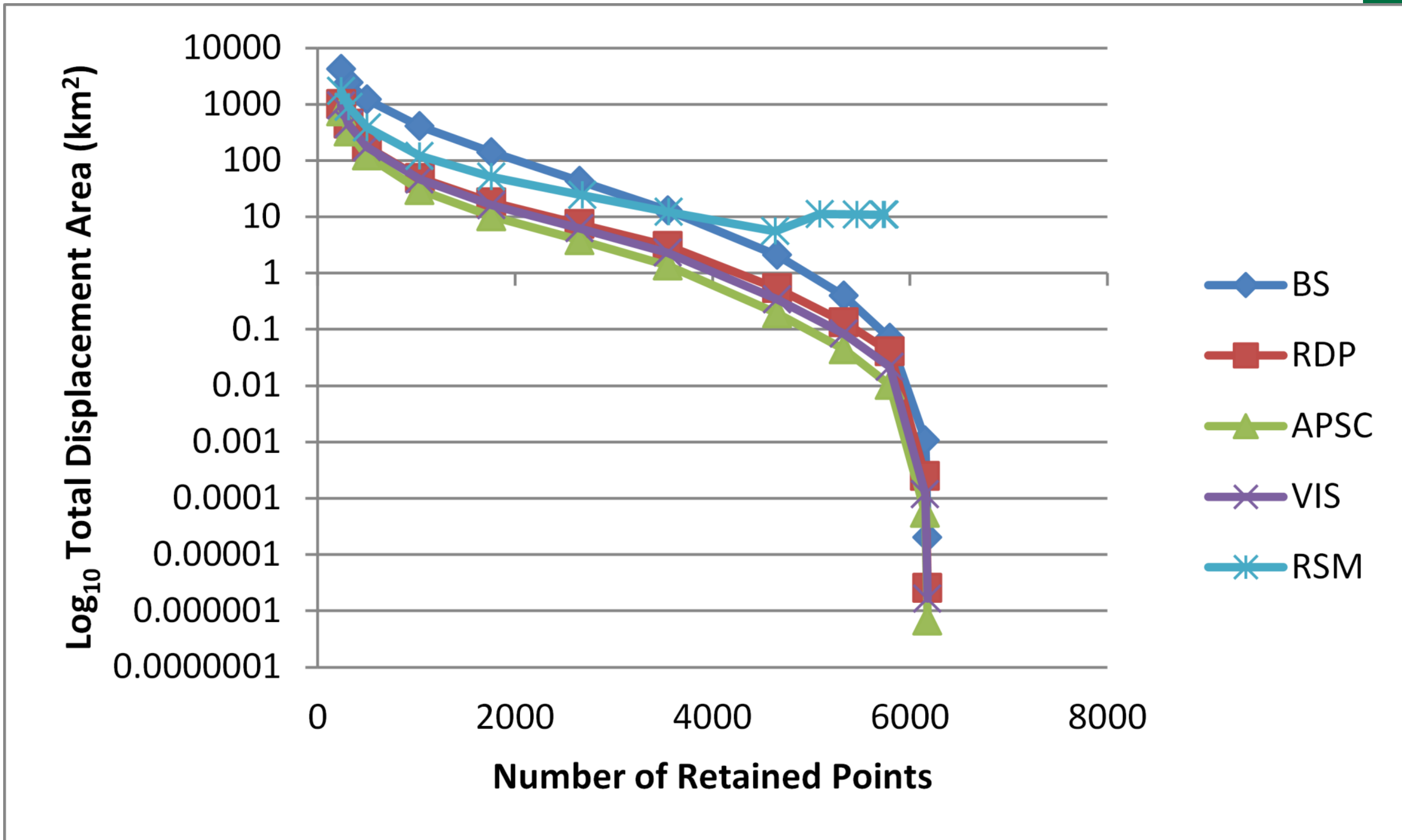
Results: Residual Sinuosity

Paired different t-test with Holm's reduction for multiple hypothesis testing

Confidence Level: 95%										
Residual Sinuosity										
Compare to:	BS				RDP			APSC		VIS
LOD	RDP	APSC	VIS	RSM	APSC	VIS	RSM	VIS	RSM	RSM
1:5000				TRUE			TRUE		TRUE	TRUE
1:10000				TRUE			TRUE		TRUE	TRUE
1:24000				TRUE			TRUE		TRUE	TRUE
1:50000	TRUE	TRUE		TRUE	TRUE		TRUE	TRUE	TRUE	TRUE
1:100000						TRUE	TRUE	TRUE	TRUE	TRUE
1:250000	TRUE	TRUE	TRUE				TRUE		TRUE	TRUE
1:500000	TRUE	TRUE	TRUE		TRUE		TRUE	TRUE	TRUE	TRUE
1:1000000	TRUE	TRUE	TRUE	TRUE	TRUE		TRUE	TRUE	TRUE	TRUE
1:2000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:5000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:1000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
1:2000000	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

Holm S. (1979) A simple sequentially rejective multiple test procedure. Scand. J. Stat 6(2): 65-70.

+ Results: Displacement Area

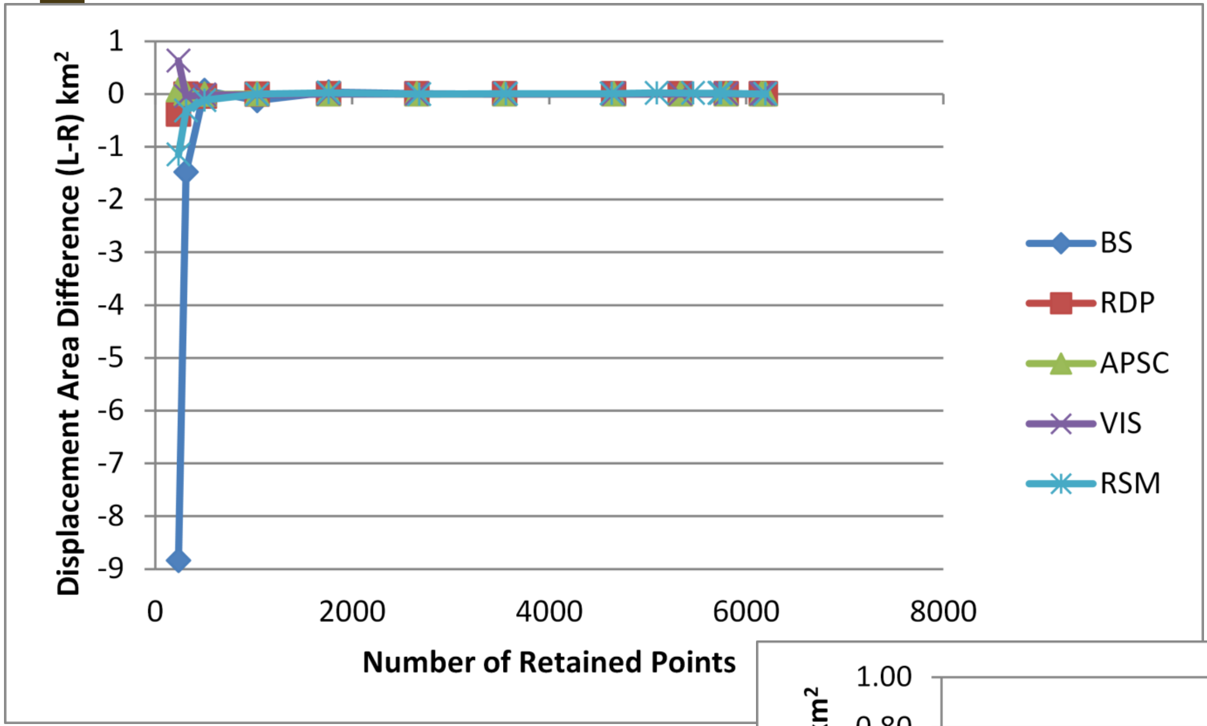


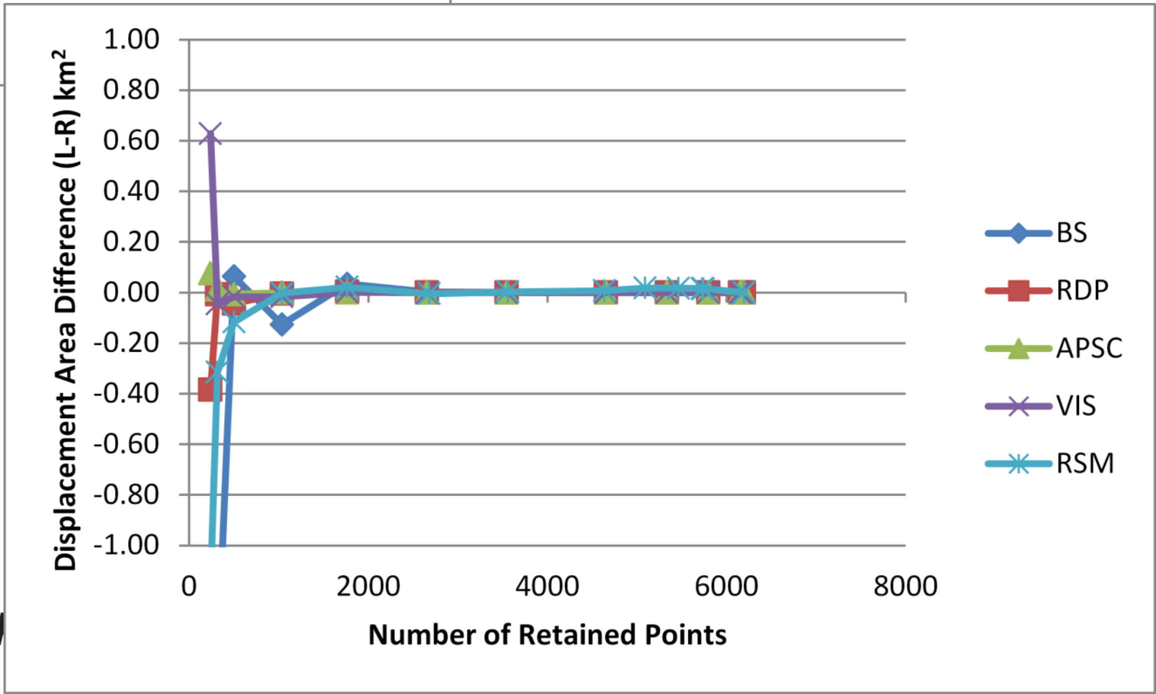
Smaller scale

Larger scale

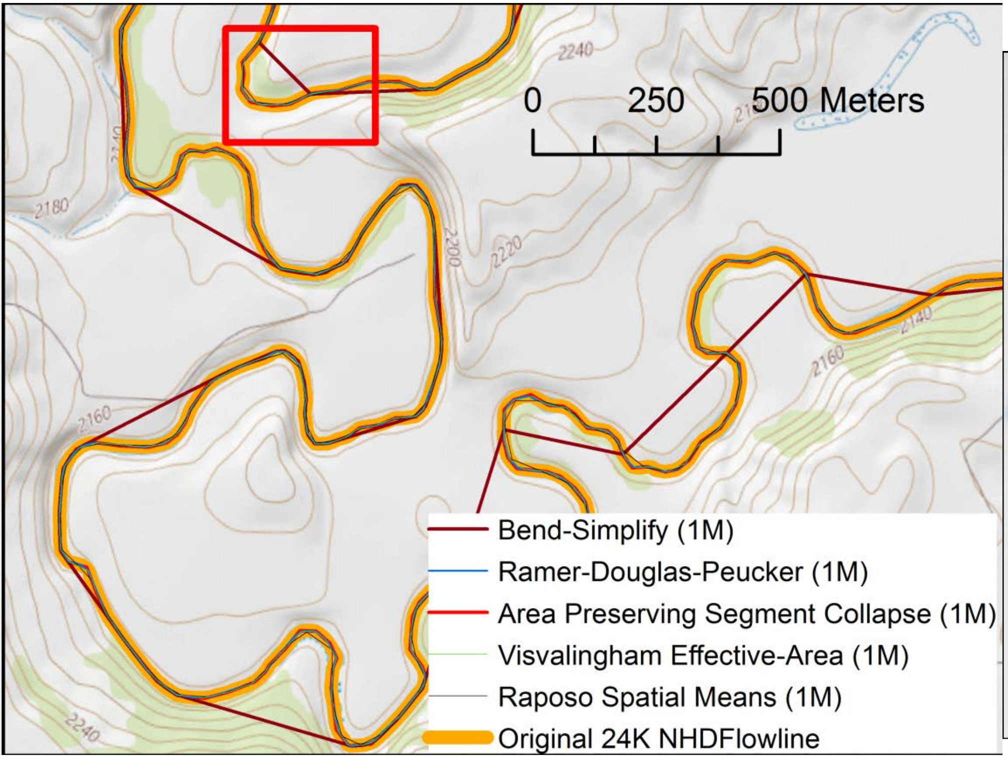
USGS The National Map

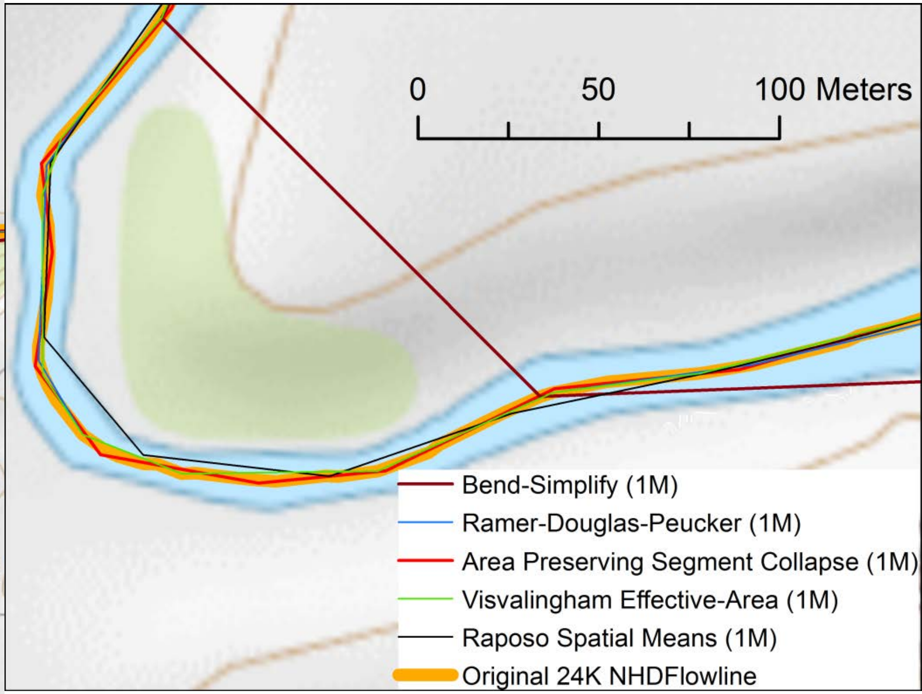
Preliminary Results: Displacement Area Difference (left – right)





+ Visual comparison





+ Summary

- Compared five different line simplification algorithms, two of which use Steiner points (RSM and APSC), and three which do not (BS, RDP, VIS).
- Simplified 50 sections of streams from the HR NHD flowlines distributed over most of the ecological conditions within the conterminous United States.
- Features were simplified to 12 levels of detail based on a 0.4 mm legibility at 12 scales ranging from 1:5,000 to 1:20,000,000.
  - Applying BS method with the legibility tolerance determined the number of vertices to retain for each feature, and all algorithms were simplified to this number of vertices by the other algorithms
- Results indicate the RDP is most efficient at minimizing horizontal positional displacement but APSC is the next best solution.
- APSC optimally simplifies linear stream features with minimal horizontal and areal displacement and the best retention of sinuosity

+ Further work

- Complete statistical evaluation/verification of results
- Improve APSC for better topological control (eliminate possibilities of line intersections)
- Comparison of Richardson plots (fractal dimension) of the stream sections to quantify relative amounts of high and low frequency detail.