An Automated Approach to Compare Generalized Hydrographic Network Features to a Benchmark Dataset

Lawrence Stanislawski

ABSTRACT

Two primary levels of detail of the United States National Hydrography Dataset (NHD) are available for all areas of the country: the high-resolution and 1:100,000-scale layers. Additional resolutions are required to support the automated cartographic mapping and web services desired by the United States Geological Survey (USGS). Automated methods are being developed and tested to prune the high-resolution NHD to a consistent level of detail, where necessary, and to prune it to other levels of detail that are appropriate for mapping at smaller scales. An automated process is presented that summarizes omission and commissions errors of pruned network lines when compared to a benchmark dataset. The index is known as the coefficient of linear correspondence (CLC). The CLC is estimated for each 0.5 degree by 0.5 degree cell covering a 48 subbasin area in the central United States, where the high-resolution NHD network features were pruned to the 1:100,000-scale level of detail and compared to a benchmark, which is the 1:100,000-scale NHD network features. The CLC ranges from 0.40 to 0.92 in the 110 cells covering the study area, with an average of 0.75. The automated process of estimating the CLC and associated values is a reasonably fast method to compare pruned, or otherwise generalized, linear features to a set of benchmark lines. This repeatable, consistent process provides spatially distributed indices that isolate areas with relatively poor matching to the benchmark lines.