Insights on Great Lakes Water Quality From Geospatial Data Analytics

With modern sensor technology, large amounts of sensor data on lake water chemistry are being generated to help researchers understand the spatial and temporal patterns of lake processes. Each sensor generates different data streams and effectively utilizing the resulting large and diverse datasets to improve understanding of lake processes and optimize sampling strategies is essential to protect and improve lake resources. For example, in the Great Lakes, the case study examined in this work, the US Environmental Protection Agency (USEPA) conducts several monitoring programs with various sensors, including the TRIAXUS undulating vehicle, the Sea-Bird CTD (Conductivity, Temperature, Depth) depth profiler, and a dissolved oxygen (DO) logger network. In this work, we developed three data analysis frameworks to support limnologists in more effectively collecting and analyzing these types of datasets, providing a lake system perspective. Applying the analytics to the Great Lakes datasets yielded new insights on river-lake mixing zones, lake stratification, and trends in lake warming and hypoxia. The frameworks have been made available to the research community as open-source code, including three prototype interactive Web applications.