# Environmental data for future generations: storage formats for multi-parameter, spatiotemporal data

DEWEY W. DUNNINGTON[1] AND IAN S. SPOONER[2]

*1. Centre for Water Resources Studies, Department of Civil and Resources Engineering, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada <dewey.dunnington@dal.ca>*

*2. Department of Earth and Environmental Science, Acadia University, Wolfville, Nova Scotia B4P 2R6, Canada*

Long-term records of environmental change are made up of complex, multi-parameter data, often collected from many sites; the complexity of these datasets can make the storage, visualization, and manipulation of such data inherently challenging. Metadata documenting the how and why of data collection are commonly omitted or stored separately from the data. Dedicated programs have attempted to ameliorate these problems, but the storage format used can be inflexible and/or proprietary, limiting the future reuse of data. In essence, environmental data are comprised of measurements, each having qualifiers (e.g., location identifier, depth below surface, and measured parameter), a value, and tags (e.g., amount of error, number of replicates, written notes pertaining to the value). When data are stored in a table with one row per measurement, the maximum amount of measurement data is retained; when data are stored in a table with one row per time interval per location (one column per parameter), some information is lost but the data are more amenable to visualization in spreadsheet software. The conversion between these structures is easily accomplished using both interactive (e.g., spreadsheet software) and programmatic (e.g., R and Python) mechanisms. As more advanced statistical treatment of data becomes common in long-term environmental studies, storing data in a way that does not result in data loss is advantageous to enhance the replicability of visualizations and statistical analyses. As datasets are increasingly combined with others and reused in future analyses, formats that store metadata with the data itself are particularly important for data collectors to consider.

Atlantic Geology, 2018, Volume 54
Atlantic Geoscience Society Abstracts – 44th Colloquium & Annual General Meeting 2018
doi: 10.4138/atlgeol.2018.003