

This is a brief explanatory document (readme) for appendix 1
Characterization of Major-Ion Chemistry and Nutrients in Headwater Streams Along the Appalachian
National Scenic Trail and Within Adjacent Watersheds, Maine to Georgia
USGS Scientific Investigations Report 2011--5151

Appendix 1. Ancillary and water-quality variable definitions and catchment data

The data in this appendix are a compilation of selected water-quality data collected from first- and second-order streams along or near the Appalachian National Scenic Trail. Water-quality data include pH, specific conductance, acid neutralizing capacity, and concentrations of bicarbonate, calcium, chloride, fluoride, magnesium, potassium, silica, sodium, sulfate, ammonia, nitrate, total nitrogen, and total phosphorus.

The data compiled for this project were collected by Federal, State, and university water-quality research programs. Data sources include the U.S. Geological Survey; the U.S. Environmental Protection Agency; the National Park Service; the U.S. Forest Service; the National Science Foundation, including the Long-Term Ecological Research Program; the Water Resources Research Center at the University of Massachusetts, Amherst; and the Department of Environmental Sciences at the University of Virginia, Charlottesville. Water-quality data from all of these sources were obtained in a variety of electronic formats.

The data are provided in a Microsoft Excel workbook (ANST_dataset) that includes two worksheets: variable definitions and catchment data. The variable-definition worksheet describes the variables in the catchment-data worksheet. Where appropriate, references were included in the variable definitions. The catchment-data worksheet lists the water-quality and ancillary variables used in the analysis for the project. Ancillary variables include those for the geospatial information system data that were assigned to and used in conjunction with the water-quality data.

The water-quality data in the catchment-data worksheet are computed values. To mitigate a potential bias resulting from disparities in the spatial locations of sampling sites and in the frequencies of sampling, the data were aggregated in four steps: 1) data were aggregated by sampling site and date to determine a daily median value for each constituent; 2) daily median values were aggregated by sampling site and year to determine an annual median value for each constituent (for example, if 300 daily median values of pH were recorded at a sampling site in one year, the median of these 300 pH values was determined as the annual median value for that year); 3) annual median data were aggregated by site, thus reducing the data to one median annual value for each constituent at that site for the entire sampling period in years; and 4) data were aggregated by National Hydrography Dataset *Plus* (NHD*Plus*) catchments to a single median value for each constituent in each catchment. Sampling sites in the same catchment are on the same stream segment (fig. 10). The number of sampling sites within a catchment ranged from 1 to 55; however, 542 of the 831 catchments contained only one sampling site. The resulting values (hereafter referred to as catchment values or concentrations) for each constituent were used to assess water quality along the Appalachian Trail and are available in addition to ancillary data in the catchment-data worksheet of appendix 1.