Open-File Report 2008-1163
Data file README

Data Files

Geospatial datasets are available to download as Arc Info shapefiles (zipped using Winzip). Data files of attributes are available to download in Excel (version 2003) format and in tab-delimited text format. Each Excel workbook contains a data sheet and a sheet (named README) with variable name definitions and notes.

Description	Downloadable datafile and metadata	Section of report describing data
SAGT ERF1_2 digital segmented network (geospatial dataset)	*erf1_spar.zip* (5.1 MB), *erf1_spar.html*	Hydrologic network of reaches and associated catchments
SAGT ERF1_2 segmented catchments (geospatial dataset)	*shed_cov.zip* (13 MB), *shed_cov.html*	Hydrologic network of reaches and associated catchments
Catchment-level estimates of watershed and reach attributes evaluated for incremental catchments and reaches	Excel version: *SAGT_ERF1_input.xls.zip* (4.1 MB) (metadata included in the README sheet) Textfile version: *SAGT_ERF1_input.txt* (2.1 MB), *README_SAGT_ERF1_input.txt*	Watershed Attributes, Reach Attributes
Estimates of watershed attributes accumulated for the total upstream watershed contributing to the reach segment	Excel version: *SAGT_accumulatedfortotalwatershed.xls.zip* (4.0 MB) (metadata included in the README sheet) Textfile version: *SAGT_accumulatedfortotalwatershed.txt* (2.3 MB), *README_SAGT_ERF1_input.txt*	Accumulation of Catchment-Level Estimates of Watershed Attributes to Estimates for the Total Upstream Watershed
Monitoring sites, station characteristics, and nutrient load estimates	Excel version: *SAGT_monitoredload.xls* (700 KB) (metadata included in the README sheet) Textfile version: *SAGT_monitoredload.txt* (28 KB), *README_SAGT_monitoredload.txt*	Mean Annual Nitrogen and Phosphorus Load at Stream Monitoring Sites

Routines used to modify nutrient-constituent concentration data for load estimation are available to download in SAS (version 9) format and in text format.

Description	Downloadable program file	Section of report describing data
Reformats the water-quality data file from modernized STORET (tilde-delimited) to a SAS datafile in the format used by Fluxmaster (more details provided in paragraphs following this table)	*Reformat_ModSTORET_WQdata.sas* (42 KB) text version: *Reformat_ModSTORET_WQdata.txt* (42 KB)	Review and Revision of Nutrient Concentration Results
Resolves the differences among data sources in the format or convention for recording results (more details provided in paragraphs following this table)	*Convert_remarkcoding_and_otherproblematic.sas* (16 KB) Text version: *Convert_remarkcoding_and_otherproblematic.txt* (16 KB)	Review and Revision of Nutrient Concentration Results
Assigns or calculates a value for a total nitrogen (TN) parameter code, P60000, and for a total phosphorus (TP) parameter code, P66500 (more details provided in paragraphs following this table)	*Combine_nutrient_constituents.sas* (8 KB) Text version : *Combine_nutrient_constituents.txt* (8KB)	Review and Revision of Nutrient Concentration Results

The file "Reformat_ModSTORET_WQdata.sas" reformats the water-quality data file from modernized STORET (tilde-delimited) to a tab-delimited file, interpreting information from several variables (characteristic name, sample fraction, and media) into an assignment of parameter code (pcode) following the convention used in Legacy STORET and in NWIS, and populating the associated remark code variable for results below detection. The tab-delimited file is then converted to a SAS datafile in the format (one line per sample) used by the load estimation program Fluxmaster (Schwarz and others, 2006). Multiple stations may be included in the analysis.

The program "Convert_remarkcoding_and_otherproblematic.sas" resolves the differences among data sources in the format or convention for recording results, by revising the data records retrieved from Legacy and Modernized STORET to match the NWIS format or convention. (The load estimation program, Fluxmaster, is programmed to work with data coded using the NWIS convention.) This routine also corrects cases of obviously erroneous concentration results, such as extremely large values.

Summary of changes for legacy STORET data records:

1. Replace the nonsense numbers (positive and negative) with missing values

2. Replace the zero and negative values that indicate below detection with appropriate detection limit

values, and set remark code to '<'

3. Replace all remark codes that mean '<' (K and U) with '<'

4. Replace remark codes that mean '>' (L) to '>'

Summary of changes for modernized STORET data records:

1. For less than result for which detection-limit was not available in the retrieved data in order to populate the value field (P field) during reformatting: set value field equal to a reasonable estimate of detection limit (75 percentile of all detection limits reported in the SAGT project dataset from STORET, which can be obtained from distribution of values in the detection-limit field, or D field)

2. Replace all remark codes that mean '<' (U) with '<'.

The program “Combine_nutrient_constituents.sas” assigns or calculates a value for a total nitrogen (TN) parameter code, P60000, and for a total phosphorus (TP) parameter code, P66500. The code P60000 is assigned a value equal to P00600; or if P00600 is missing, it is calculated by combining total Kjeldahl nitrogen results and nitrate results (if available), or by combining ammonia nitrogen results, organic nitrogen results, and nitrate results. The code P66500 is assigned a value equal to P00665, or if P00665 is missing, it is calculated by combining dissolved and suspended phosphorus (if available, although this is rarely the case). The rules for combining include how to handle the case of one or more of the constituents having censored values, and how to populate remark code for the calculated parameter.