

## METADATA EXAMPLE...

Following, we include an excerpt of a [metadata document](#) created by the US Geological Survey (USGS), using the FGDC CSDGM standard, which describes a *shapefile* that contains over 9,000 gage stations around the US. The excerpt includes citation, description (abstract and purpose), and time period of content.

FGDC Metadata Form	DATASET 1
<b>Identification</b>	
Citation	
Originator	James Falcone
Publication_Date	20111012
Title	GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow
Online_Linkage	<a href="http://water.usgs.gov/lookup/getspatial?gagesII_Sept2011">http://water.usgs.gov/lookup/getspatial?gagesII_Sept2011</a>
Network_Resource_Address	
Geospatial_Data_Presentation_Form	vector digital data
<b>Publication_Information</b>	
Publication_Place	Reston, Virginia
Publisher	U.S. Geological Survey
Abstract	<p>This dataset, termed “GAGES II”, an acronym for Geospatial Attributes of Gages for Evaluating Streamflow, version II, provides geospatial data and classifications for 9,322 stream gages maintained by the U.S. Geological Survey (USGS). It is an update to the original GAGES, which was published as a Data Paper on the journal Ecology’s website (Falcone and others, 2010b) in 2010. The GAGES II dataset consists of gages which have had either 20+ complete years (not necessarily continuous) of discharge record since 1950, or are currently active, as of water year 2009, and whose watersheds lie within the United States, including Alaska, Hawaii, and Puerto Rico. Reference gages were identified based on indicators that they were the least-disturbed watersheds within the framework of broad regions, based on 12 major ecoregions across the United States. Of the</p>

9,322 total sites, 2,057 are classified as reference, and 7,265 as non-reference. Of the 2,057 reference sites, 1,633 have (through 2009) 20+ years of record since 1950. Some sites have very long flow records: a number of gages have been in continuous service since 1900 (at least), and have 110 years of complete record (1900-2009) to date.

The geospatial data include several hundred watershed characteristics compiled from national data sources, including environmental features (e.g. climate – including historical precipitation, geology, soils, topography) and anthropogenic influences (e.g. land use, road density, presence of dams, canals, or power plants). The dataset also includes comments from local USGS Water Science Centers, based on Annual Data Reports, pertinent to hydrologic modifications and influences. The data posted also include watershed boundaries in GIS format.

This overall dataset is different in nature to the USGS Hydro-Climatic Data Network (HCDN; Slack and Landwehr 1992), whose data evaluation ended with water year 1988. The HCDN identifies stream gages which at some point in their history had periods which represented natural flow, and the years in which those natural flows occurred were identified (i.e. not all HCDN sites were in reference condition even in 1988, for example, 02353500). The HCDN remains a valuable indication of historic natural streamflow data. However, the goal of this dataset was to identify watersheds which currently have near-natural flow conditions, and the 2,057 reference sites identified here were derived independently of the HCDN. A subset, however, noted in the BasinID worksheet as “HCDN-2009”, has been identified as an updated list of 743 sites for potential hydro-climatic study. The HCDN-2009 sites fulfill all of the following criteria: (a) have 20 years of complete and continuous flow record in the last 20 years (water years 1990-2009), and were thus also currently active as of 2009, (b) are identified as being in current reference condition according to the GAGES-II classification, (c) have less than 5 percent imperviousness as measured from the NLCD 2006, and (d) were not eliminated by a review from participating state Water Science Center evaluators.

---

The data posted here consist of the following items:- This point shapefile, with summary data for the 9,322 gages.- A zip file containing basin characteristics, variable definitions, and a more detailed report.- A zip file containing shapefiles of basin boundaries, organized by classification and aggregated ecoregion.- A zip file containing mainstem stream lines (Arc line coverages) for each gage.

Purpose

This dataset has two purposes: (1) to provide users with a comprehensive set of geospatial characteristics for a large number of gaged watersheds, particularly for gages with long flow record, and (2) to provide a determination of which of those watersheds represent hydrologic conditions which are least disturbed by human influences (“reference gages”), compared to other watersheds within 12 major ecoregions. Identifying reference gages serves important research goals: for example identifying conditions or goals for stream restoration, climate change studies, and more.

...

---

#### Sources cited:

- Extensible Markup Language (XML): <https://www.w3.org/XML/>
- Digital Curation Centre (DCC): <http://www.dcc.ac.uk/>
  - DCC Curation Lifecycle Model: <http://www.dcc.ac.uk/resources/curation-lifecycle-model>
- FGDC CSDGM: <http://www.fgdc.gov/metadata/geospatial-metadata-standards>
- ISO: <http://www.iso.org/iso/home/standards.htm>
  - NOAA resources: <http://www.ncddc.noaa.gov/metadata-standards/>
- DCMI: <http://dublincore.org/documents/dcmi-terms/>
- USGS metadata document:[http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII\\_Sept2011.xml](http://water.usgs.gov/GIS/metadata/usgswrd/XML/gagesII_Sept2011.xml)