Title

Geospatial Information, Metadata, and Maps for Global River Corridor Science Focus Area Sites (v2)

Summary

This dataset provides geospatial information, metadata, and maps for the Pacific Northwest National Laboratory (PNNL) River Corridor Science Focus Area (RC SFA; https://www.pnnl.gov/projects/river-corridor) sites. The RC SFA works to transform understanding of spatial and temporal dynamics in river corridor hydrobiogeochemical functions from molecular reaction to watershed and basin scales. The knowledge we gain is used to formulate and test hypotheses and to improve mechanistic representation of river corridor processes and their response to disturbances in multiscale models of integrated hydrobiogeochemical function. The data provided includes Site ID, latitude, longitude, stream name, and common ID (COMID) for sites used across the RC SFA. The COMID can be used to find and download data from NHDPlus (https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus) and other platforms. The sites included are non-exhaustive. Sites (including past sites) will be added to this data package in the future. Data generated from the RC SFA can be accessed at https://data.ess-dive.lbl.gov/portals/PNNLRiverCorridorSFA.

Critical Details

1 – Locations for each site were collected by field teams using a variety of GPS technologies. Each site's location was matched with the nearest **NHDPlus Version 2** (NHDPlusV2; https://www.epa.gov/waterdata/get-nhdplus-national-hydrography-dataset-plus-data) network streamline using the NNJoin nearest-neighbor spatial join plugin (https://plugins.qgis.org/plugins/NNJoin/) for QGIS (QGIS Geographic Information System. QGIS Association. http://www.qgis.org). Stream lines are identified in NHDPlusV2 by a unique common-ID or COMID. The COMID identified by the nearest-neighbor spatial join is referred to here as the raw COMID. Join results (raw COMIDs) were manually reviewed by personnel familiar with the sites. One key review criterion was whether or not the stream name from the site metadata matched the stream name from NHPPlusV2. In situations where the nearest-neighbor spatial join did not identify the correct stream, the correct stream line was manually identified (if possible), and a rectified COMID was extracted.

Data Package Structure

This dataset is comprised of one main data folder. The data folder consists of (1) file-level metadata; (2) data dictionary; (3) readme; (4) methods codes; (5) geospatial information for all RC SFA sites including international geo-sample number (IGSN); (6) maps of all sites and sites in Washington State, USA; and (7) a subfolder with the shapefile of all sites. All files are .csv, .pdf, .shp, .cpg, .dbf, .prj, .qmd, or .shx.

Acknowledgements

We acknowledge the Yakama Nation as owners and caretakers of the lands where we collected some of these data in Washington state. We thank the Confederated Tribes and Bands of the Yakama Nation Tribal Council and Yakama Nation Fisheries for working with us to facilitate sample collection and optimization of data usage according to their values and worldview.

This research was supported by the U.S. Department of Energy (DOE) Biological and Environmental Research (BER) Environmental System Science (ESS) program (https://ess.science.energy.gov/) through the Pacific Northwest National Laboratory River Corridor Science Focus Area (SFA). PNNL is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract No. DE-AC05-76RL01830.

Contact

James Stegen, james.stegen@pnnl.gov

Change History

Version 1	April 2023	Original data package publication
Version 2	June 2023	• Updated coordinates and COMID for site S47R • Added S55N and S56N to distinguish the sites that were visited in 2022 (S55N & S56N) from the sites visited in 2021 (S55 and S56) • Updated file names to reflect new version naming format • Updated map, shapefile, readme, and flmd to reflect changes