Title

Spatial Study 2022: Surface Water Samples, Cotton Strip Degradation, and Hydrologic Sensor Data across the Yakima River Basin, Washington, USA (v3)

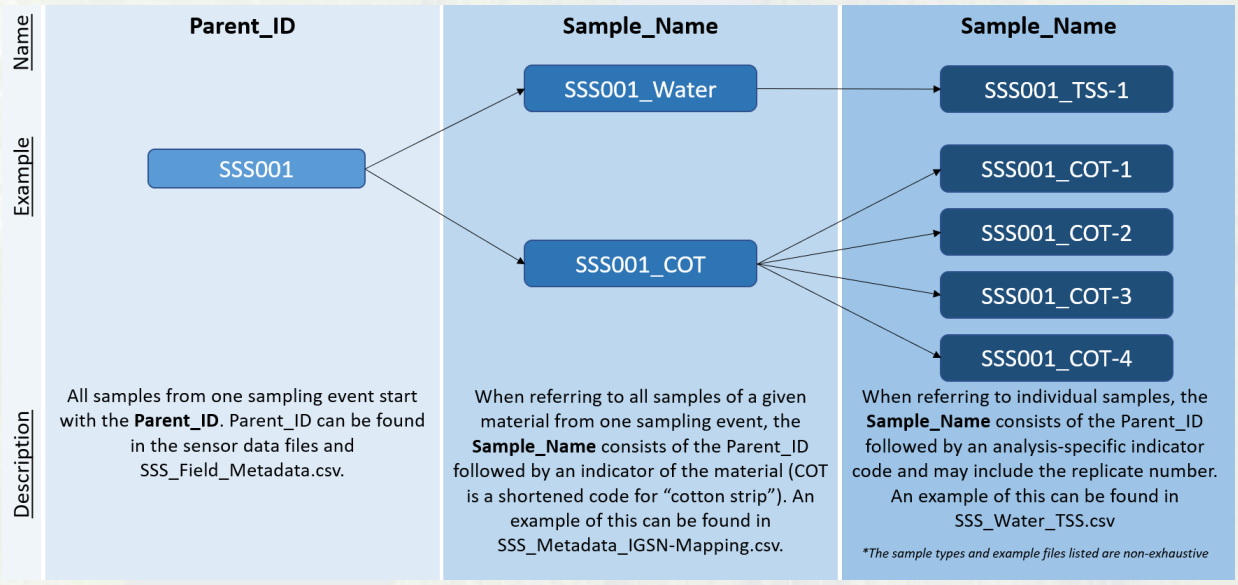
Summary

This dataset supports a broader study examining the drivers of spatial variability in sediment respiration rates in the Yakima River Basin. The dataset provides data and photos generated from sample collection during the same one-week period at 48 sites within multiple rivers throughout the Yakima River Basin in Washington, USA. The contents include surface water geochemistry data; river substrate grain size photos; stream depth data; manual chamber open channel respiration data; and field metadata (including qualitative information on instream and river corridor characteristics). Grain size photos can be used to improve estimates of channel substrate D50 data. The dataset also includes tensile strength and photos from cotton strip field degradation experiments; five-week sensor time series temperature, dissolved oxygen, pressure, pH, specific conductance, chlorophyll A, and turbidity data; plots of the sensor data; and R scripts used to generate the plots.

Samples collected during this study were labeled as "Second Spatial Study" or "SSS." A subset of data from the SSS samples were published in the contiguous United States (CONUS)-Scale Model-Sample (CM) study data package available at https://data.ess-dive.lbl.gov/view/doi:10.15485/1923689 that presents data from across the CONUS. SSS data published in the CM data package were not included in this data package. They include dissolved organic carbon (DOC, measured as non-purgeable organic carbon, NPOC), total nitrogen (TN), grain size, aerobic sediment respiration, dissolved oxygen (DO), and temperature. Parent IDs and Site IDs are consistent between the SSS and CM data packages, and they can be mapped directly so data across packages can be used together. Additionally, sensor data from a similar 2021 spatial study can be found at https://data.ess-dive.lbl.gov/view/doi:10.15485/1892052 and 2021 sample data can be found at https://data.ess-dive.lbl.gov/view/doi:10.15485/1898914. The 2021 spatial study had some sites in common with this 2022 spatial study.

The data package was originally published in April 2023. It was updated in August 2023 (v2; modified files) and September 2024 (v3; modified files). See the change history section below for more details.

Critical Details



1 – Each sampling event has a unique Parent_ID in the format SSS#. The field metadata and data files all contain these unique IDs and can be mapped across each other accordingly. The Parent_ID may have other indicators appended when referring to samples. See figure above for details.

2 – Each physical site has a "Site_ID" and each sampling event in time at that Site has a "Parent_ID" (i.e., in 2021, Site_ID T07 was sampled and the resulting Parent_ID of the samples was SPS_0053. In 2022, Site_ID T07 was sampled again and the resulting Parent_ID of the samples was SSS013). The sensor files indicate both the Site_ID and the Parent_ID to indicate at which Site the sensor was deployed (SiteID) and to which sampling event the data is most relevant (Parent_ID). The identifier most relevant to the data user will depend on the specific analyses being done.

3 – There were 4 sites that were originally mislabeled. All files have these site IDs corrected. However, the quadrat photos of full quadrats include a whiteboard with the original site ID written. To avoid affecting the resolution of these images, the text inside the images was left unaltered. The mislabeled site IDs and the corrected IDs are as follows,

- Site S63 was originally mislabeled as S63P. The correct site ID is S63.
- Site S55N was originally mislabeled as S55. The correct site ID is S55N.
- Site S56N was originally mislabeled as S56. The correct site ID is S56N.
- Site T42 was originally mislabeled as T41. The correct site ID is T42.

There is nothing from the original incorrect site IDs (i.e., S63P, S55, S56, and T41) in this data package.

Data Package Structure

This dataset is comprised of three photo folders and one main data folder with six subfolders. The photo folders contain photographs and videos of cotton strip retrieval and sediment quadrats. The main data folder consists of (1) file-level metadata; (2) data dictionary; (3) field metadata; (4) total suspended solids (TSS) data and cotton strip tensile strength data and averages; (5) field protocol; (6) readme; (7) methods codes; (8) international generic sample number (IGSN) mapping file; (9) sensor installation methods summary; (10) stream depth and averages; and (11) Ultrameter data. The Sonar subfolder consists of Sonar time-series depth data and a processing script. The BarotrollAtm, DepthHOBO, MantaRiver, miniDOT, and miniDOTManualChamber subfolders contain time-series data, plots, and summary files. All files are .csv, .pdf, .txt, .R, .Rmd, .jpg, .jpeg, .AVI, .mp4, or .mov.

Acknowledgements

We acknowledge the Yakama Nation as owners and caretakers of the lands where we collected these data. 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

Approach to change history and versioning:

Updates to **data package** version: When any file within a data package is updated, the data package version number is updated. The data package version number is indicated in the title of the data package, the data package folder name, and in the change history table below. You can access previous versions of the data package by sending a request to ESS-DIVE.

Updates to **individual file** versions: As files are changed, the file version number is also updated. The file version number is indicated in the file name, file level metadata (flmd) file, and the change history table below. The version number on an individual file may not match the version number of the data package. For example, v3 of a data package may include v2 of an individual file.

The change history below describes each file revised during versioning. If you are interested in seeing the exact cells within a file that have changed, you can utilize the daff package in R (https://github.com/edwindj/daff) to compare a previously downloaded file to a newly downloaded file.

In the change history table below, the sub-headers and bullets indicate the type of change in each file:

- New files: Describes new files added that were not present in previous data package versions
- Bulk changes to files: Describes a change to many files within the data package. The indicated superscript with be added to each file name that the change applies to.
- Modified files:
    - Corrected: Describes existing information modified or removed to prevent sharing of incorrect information
    - Added: Describes new information inserted into an existing file (e.g., appending new columns/rows)
    - Updated: Describes modifying existing information to maintain accuracy though version changes. (e.g., changing version number to new version number)

Change history:

Data Package Version	Changes
Version 1 *April 2023*	Original data package publication
Version 2 *August 2023*	**BULK CHANGES TO FILES** [1] indicates parent IGSNs and Site IDs for Site IDs S55N (SSS004), S56N (SSS006), and T42 (SSS016) were corrected. **MODIFIED FILES** **SSS_Water_Temp_SpC_Turb_pH_ChlA_Summary.csv** (v2) [1] • Corrected: Revised chlorophyll A unit description by removing incorrect "per_minute" **SSS_Field_Metadata.csv** (v2) [1] • Corrected: Revised sediment coordinates for Site IDs S15 (change 47.36363, -121.1073 to 47.46361, -121.1073), S47R (change 46.66737, -121.094 to

	46.62141, -121.3027), and S48R (change 46.63824, -121.2631 to 46.64568, -121.2511) **SSS_Metadata_IGSN-Mapping.csv** (v2) [1] • Corrected: Revised coordinates for Site IDs S15 (change from 47.36363, -121.1073 to 47.46361, -121.1073) and S47R (change from 46.6674, -121.094 to 46.62141, -121.3027) • Added: Included the "10.58052/" DOI prefix to IGSNs and Parent_IGSNs in the IGSN Mapping file • Updated: Revised IGSN acronym definition in to align with current guidelines **readme_SSS.pdf** (v2) • Updated: Revised version number in data package title • Updated: Included new versioning information **SSS_flmd.csv** (v2) • Updated: Revised version number to modified files • Updated: Revised IGSN acronym definition in to align with current guidelines **SSS_dd.csv** (v2) • Updated: removed rows that were removed from the data packages and added rows to reflect changes in modified files ***_Air_Press_Temp.csv** (v2) [1] **SSS_Air_Press_Temp_Summary.csv** (v2) [1] **SSS_Air_Press_Temp_TsPlot.pdf** (v2) [1] **SSS_Air_Press_Temp_VPlot.pdf** (v2) [1] ***_Water_Press_Temp.csv** (v2) [1] **SSS_Water_Press_Temp_Summary.csv** (v2) [1] **SSS_ Water_Press_Temp _TsPlot.pdf** (v2) [1] **SSS_ Water_Press_Temp _VPlot.pdf** (v2) [1] ***_Water_Temp_SpC_Turb_pH_ChlA.csv** (v2) [1] **SSS_Water_Temp_SpC_Turb_pH_ChlA_Summary.csv** (v2) [1] **SSS_ Water_Temp_SpC_Turb_pH_ChlA_TsPlot.pdf** (v2) [1] **SSS_ Water_Temp_SpC_Turb_pH_ChlA_VPlot.pdf** (v2) [1] ***_Water_DO_Temp.csv** (v2) [1] **SSS_Water_DO_Temp_Summary.csv** (v2) [1] **SSS_ Water_DO_Temp_TsPlot.pdf** (v2) [1] **SSS_ Water_DO_Temp _VPlot.pdf** (v2) [1] ***_MC_Water_DO_Temp.csv** (v2) [1] **SSS_ MC_Water_DO_Temp_Summary.csv** (v2) [1] **SSS_ MC_ Water_DO_Temp_TsPlot.pdf** (v2) [1] **SSS_ MC_ Water_DO_Temp _VPlot.pdf** (v2) [1] **SSS_Water_Depth_Summary.csv** (v2) [1] **SSS_Water_Transect_Depth.csv** (v2) [1] **SSS_*.jpg** (v2) [1]
Version 3 *September 2024*	**NEW FILES** • SSS_Sonar_Weighted_Avg.R **BULK CHANGES TO FILES** [2] indicates the time series data files, summary file, and plots were reprocessed and corrected. See dd/flmd for details of processing. All summary files were updated to only include datetime start/end, mean, median, min, and max for each variable

	MODIFIED FILES **Folder Structure** • All sample data and related methods information were put into a "Sample_Data" subfolder. All sensor data and related methods information were put into a "Sensor_Data" subfolder. **SSS_Sonar_Depth_Processing.Rmd** (v2) • Updated: Revised to run with files in data package **SSS_Water_Depth_Summary.csv** (v3) • Updated: Revised to reflect outputs of "SSS_Sonar_Depth_Processing" **SSS_Water_Temp_SpC_Turb_pH_ChlA_Summary.csv** (v3) • Updated: Removed summary stats **SSS_Installation_Methods.csv** (v2) • Updated: Revised to include updated details of processing **readme_SSS.pdf** (v3) • Updated: Reformatted change history. • Updated: Revised to reflect new changes in change history. **SSS_flmd.csv** (v3) • Updated: Revised description to include details about reprocessing • Updated: Revised version number to modified files. **SSS_dd.csv** (v3) • Updated: Revised description to include details about reprocessing ***_Air_Press_Temp.csv** (v2 and v3) [2] **SSS_Air_Press_Temp_Summary.csv** (v3) [2] **SSS_Air_Press_Temp_TsPlot.pdf** (v3) [2] ***_Water_Press_Temp.csv**(v2 and v3) [2] **SSS_Water_Press_Temp_Summary.csv** (v3) [2] **SSS_ Water_Press_Temp _TsPlot.pdf** (v3) [2] ***_Water_Temp_SpC_Turb_pH_ChlA.csv** (v2 and v3) [2] **SSS_Water_Temp_SpC_Turb_pH_ChlA_Summary.csv** (v3) [2] **SSS_ Water_Temp_SpC_Turb_pH_ChlA_TsPlot.pdf** (v3) [2] ***_Water_DO_Temp.csv** (v2 and v3) [2] **SSS_Water_DO_Temp_Summary.csv** (v3) [2] **SSS_ Water_DO_Temp_TsPlot.pdf** (v3) [2] ***_MC_Water_DO_Temp.csv** (v2 and v3) [2] **SSS_ MC_Water_DO_Temp_Summary.csv** (v3) [2] **SSS_ MC_ Water_DO_Temp_TsPlot.pdf** (v3) [2]

Note: An asterisks (*) in a file name indicates multiple files with the same file and file naming structure. See the flmd for more details about specific files and the meaning of the asterisks.