

## Title

Spatial Study 2022: Surface Water Samples, Cotton Strip Degradation, and Hydrologic Sensor Data across the Yakima River Basin, Washington, USA (v2)

## Summary

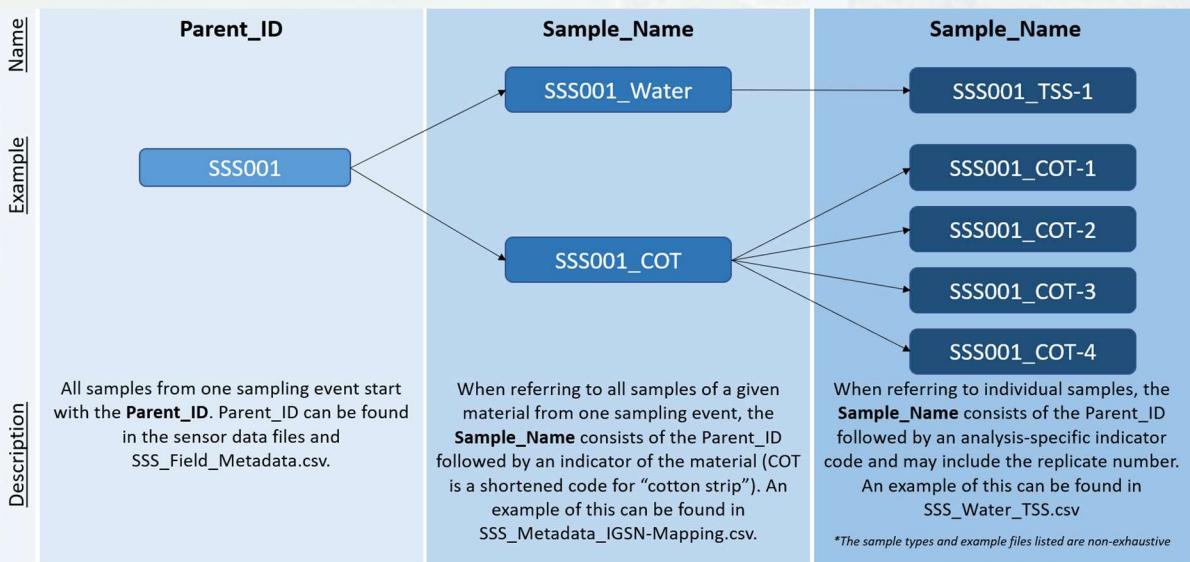
This dataset supports a broader study examining the drivers of spatial variability in sediment respiration rates in the Yakima River Basin. The dataset provides data and photos generated from sample collection during the same one-week period at 48 sites within multiple rivers throughout the Yakima River 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. 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 below 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 and averages. 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](mailto:james.stegen@pnnl.gov)

## Change History

Version 1	April 2023	Original data package publication
Version 2	August 2023	<ul style="list-style-type: none"><li>Corrected parent IGSNs and Site IDs in field metadata, IGSN mapping file, depth files, photo file names, and sensor data (including data, plots, and summary files) for Site IDs S55N (SSS004), S56N (SSS006), and T42 (SSS016).</li><li>Corrected parent IGSN for S55N, S56N, and T42 in SESAR IGSN database online.</li><li>Corrected coordinates in IGSN mapping file 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).</li><li>Corrected sediment coordinates in field metadata file 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).</li><li>Added the "10.58052/" DOI prefix to IGSNs and Parent_IGSNs in the IGSN Mapping file.</li><li>Updated IGSN acronym definition in IGSN mapping file, dd, and flmd to align with current guidelines.</li><li>Updated date columns headers to align with current guidelines in flmd file.</li><li>Removed incorrect "per_minute" chlorophyll A units from the unit description in SSS_Water_Temp_SpC_Turb_pH_ChLA_Summary.csv</li><li>Updated title, flmd, dd, and readme to reflect changes.</li></ul>