Title

Spatial Study 2021: Sensor-Based Time Series of Surface Water Temperature, Specific Conductance, Total Dissolved Solids, pH, and Dissolved Oxygen from across Multiple Watersheds in 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 two-hour time series hydrological and water chemistry sensor data, manual chamber open channel respiration data, handheld sensor water chemistry data, river substrate grain size photos, general environmental context photos, and field metadata (including qualitative information on instream and river corridor characteristics) collected during the same two-week period at 47 sites within multiple rivers throughout the Yakima River Basin in Washington, USA. Grain size photos can be used to improve estimates of channel substrate D50 data. Related sample-based water chemistry data are published separately at https://data.ess-dive.lbl.gov/datasets/doi:10.15485/1898914.

This data package was originally published September 2022. It was updated January 2023 (modified files) and June 2024 (new and modified files). See the change history section below for more details.

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.

Brief Overview of Methods

Sensor time series data, field metadata, and grain size photographs were collected at 47 sites by 2-person teams in multiple rivers within the Yakima River Basin in late August through mid-September, 2021. Sensors were deployed underwater for 2 hours at each site to collect continuous time series data. Triplicate dark bottle incubations were deployed to collect open channel respiration data. Riverbed sediment grain size data were collected by laying a 80-cm x 80-cm quadrat composed of sixteen individual 20-cm x 20-cm squares at select locations on the exposed sediment of the riverbed and select locations underwater on the riverbed. Site metadata, including sensor deployment information and general environmental information, were manually recorded in the field. For complete information regarding data collection, see the Sampling Protocol for details on water sample and site metadata collection and the Sensor Data Protocol for details on sensor data collection.

Critical Details

1 – Dissolved oxygen data produced by the River Corridor SFA from miniDOT sensors typically include an offset to account for sensor drift. However, data from this study does not include an offset because most sensors were relatively new.

Data Package Structure

This dataset is comprised of four main folders, one containing three sensor-specific subfolders and the others containing photographs. The SFA_SpatialStudy_2021_SensorData main data folder includes file-level metadata (FLMD), data dictionary (dd), installation methods, field metadata, Ultrameter water chemistry data, field data collection protocols, international generic sample number (IGSN) mapping file,

and a readme file. The "Sensor_Manual_Specifications" subfolder contains pdf files from the manufacturer of each sensor with details on the sensor specifications. Each sensor subfolder (BarotrollAtm, MantaRiver, and MinidotManualChamber) contains a sensor data subfolder for timeseries data and a subfolder for plots and summary statistics. The BarotrollAtm Data subfolder contains In Situ Rugged BaroTROLL pressure and temperature data. The MantaRiver Data subfolder contains Eureka Manta+ 35B multisonde temperature, specific conductance, and pH data. The MinidotManualChamber Data subfolder contains PME MiniDOT Logger dissolved oxygen (mg/L and percent saturation) and temperature data. The folder SFA_SpatialStudy_2021_EnvironmentalContextPhotos contains environmental context photographs and videos. The folders SFA_SpatialStudy_2021_SedimentQuadratPhotos_Part1 and SFA_SpatialStudy_2021_SedimentQuadratPhotos_Part2 contain sediment quadrat photographs. All files are .csv, .pdf, .R, .jpg, .jpeg, .mp4, or .mov.

Acknowledgements

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
- 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 *September 2022*	Original data package publication
Version 2 *September 2022*	**MODIFIED FILES** **Minidot_Summary_Statistics.csv** (v2) • Added: Calculated and included normalized root mean square error (NRMSE). **SPS_Sample_dd.csv** (v2) • Added: Reflected additional column headers of modified files. **SPS_Sample_flmd.csv** (v2) • Updated: Reflected changes in version number of modified files. **readme_SPS_Spatial_Sample.**pdf (v2) • Updated: Reflected changes in change history.
Version 3 *June 2024*	**NEW FILES** • *_Manual_Specifications.pdf **BULK CHANGES TO FILES** [1] indicates the summary statistics reported now include mean, median, minimum, and maximum. For manual chamber dissolved oxygen, slope, R-squared, RMSE, and NRMSE are also reported. Battery is no longer included in any summary file. The first and last datetime is now included in the summary file. The summary file has also been updated to reflect the changes in the time series files. [2] indicates the time series plots have been updated to reflect the changes in the time series files. [3] indicates the violin plots have been removed because they were not needed. **MODIFIED FILES** **MantaRiver_*.csv** (v2) • Corrected: Removed data at the beginning of the time series to account for sensor equilibration. • Corrected: Removed turbidity due to inaccurate calibration. **Minidot_*.csv** (v2) • Corrected: Recalculated dissolved oxygen saturation using the streamMetabolizer::calc_DO_sat() function. • Updated: Combined data from all three miniDOTs into one file. • Updated: Trimmed the time series to the 90-minute incubation • Updated: Rounded all data types to three decimal points. • Updated: Removed the serial number from the file name and clarify sampling location included in the file name (“Temporal_Sites” is now “Yakima”). **SPS_Ultrameter_WaterChem.csv** (v2) • Corrected: Corrected the column names in the header rows to match the column headers. • Updated: Removed summary stats and outlier detection. • Updated: Revised to reflect changes in file name. Previously named “SPS_Ultrameter_WaterChem_Summary.csv”. **BarotrollAtm_*.csv** (v2)

	• Updated: Trimmed the time series to match the 90-minute manual chamber incubation. • Updated: Rounded all data types to three decimal points. **SPS_Sensor_Field_Metadata.csv** (v2) • Updated: Removed serial number columns. **SPS_Sensor_IGSN_Mapping.csv** (v2) • Updated: Added DOI to IGSN **SPS_Sample_dd.csv** (v3) • Updated: Revised to reflect changes to column headers of modified files. **SPS_Sample_flmd.csv** (v3) • Updated: Revised to reflect changes in version number of modified files. **readme_SPS_Spatial_Sample.**pdf (v3) • Updated: Reformatted change history. • Updated: Revised to reflect new changes in change history. **BarotrollAtm_Summary_Statistics.csv** (v2) [1] **BarotrollAtm_TsPlot.pdf** (v2) [2] **BarotrollAtm_VPlot.pdf** (v1) [3] **Manta_Summary_Statistics.csv** (v2) [1] **MantaRiver_TsPlot.pdf** (v2) [2] **MantaRiver_VPlot.pdf** (v1) [3] **Minidot_Summary_Statistics.csv** (v3) [1] **Minidot_TsPlot.pdf** (v2) [2] **Minidot_VPlot.pdf** (v1) [3]

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.