Title

Data and scripts associated with “Allometric scaling of hyporheic respiration across basins in the Pacific Northwest USA

Summary

This data package is associated with the publication “Allometric scaling of hyporheic respiration across basins in the Pacific Northwest USA” submitted to JGR-Biogeosciences (Regier et al. 2025).
This study used reach-scale modeled estimates of hyporheic aerobic respiration made by the River Corridor Model (Fang et al. 2020) and watershed characteristics across the Willamette and Yakima River basins to explore potential allometric scaling (i.e., power-law relationships between size and function) of cumulative hyporheic respiration across catchment-to-basin scales. Scaling was explored quantitatively via the R^2, slope, and y-intercept of relationships between cumulative hyporheic respiration and watershed area, divided into hyporheic exchange flux (HEF) quantiles. We also explored relationships between allometric scaling and other watershed characteristics through linear regression, spatial patterns, and mutual information analyses. Our results also suggest variability of hyporheic respiration allometry for middle exchange flux quantiles, and in relation to land-cover. Our findings provide initial evidence that allometric scaling may be useful for predicting hyporheic biogeochemical dynamics across watersheds from reach to basin scales. This data package is associated with the GitHub repository found at https://github.com/peterregier/rc_wrb_yrb_scaling.

Brief Overview of Methods

This data package includes reach-scale modeled outputs for hyporheic respiration and additional watershed hydrobiogeochemical parameters from the River Corridor Model, detailed in Fang et al. (2020) and presented in Son et al. (2022). Our analysis calculated cumulative values for hyporheic respiration and other properties, including watershed area, for every reach within the Yakima and Willamette River basins with outputs written as CSVs to the “data” folder. R scripts in the “scripts" folder were used to prepare model outputs, calculate allometric scaling relationships, interpret these relationships in the context of watershed characteristics, and create the figures in the manuscript. The model was trained using USGS sites from across the Columbia River Basin, including 38 sites for dissolved oxygen (using 4470 measurements), 50 sites for dissolved organic carbon (4842 measurements), and 64 sites for nitrate (3070 measurements). Dissolved organic carbon is the only carbon source for HZ respiration represented in the model. We gap-filled missing vertical/lateral water exchange flux and residence time values for ~13% of reaches using an eXtreme Gradient Boosting (XGBoost, v1.5.0.2) machine learning model. We calculated allometric scaling relationships as cumulative hyporheic aerobic respiration (g CO_2 d^{-1}) against watershed area (km^2). We categorized the scaling behavior of relationships between cumulative hyporheic respiration and watershed area for each combination of HEF bin and basin (n = 20) into four allometry classes: Uncertain, Sublinear, Linear, and Super-linear. We defined the allometry categories based on visual identification of an R^2 threshold of 0.8 where values exhibit a “knee”. To reveal watershed features explaining variation in the spatial distributions of hyporheic respiration across each basin, we used mutual information analysis, which calculates the dependency between two variables based on Shannon's entropy. For more details on the steps of this workflow and our statistical approach, see the associated manuscript.

Data Package Structure

The data package is organized into several key directories. The “data” folder contains multiple CSV files, including landscape heterogeneity, scaling analysis, and watershed boundary data. The “figures” folder has all figure files in both PDF and PNG formats. Core analysis scripts and figure generation scripts are in the “scripts” directory, systematically numbered for sequential execution. The root directory includes essential project files; please see the file ending in “flmd.csv” for a list and description of all files contained in this data package and the file ending in “dd.csv” for data dictionaries used to describe tabular column headers.

Citations and Acknowledgements

The authors thank Dr. Francisco Guerrero for his significant contributions in the conceptualization and analysis of the study's methodology and for generating key results that informed this manuscript; this acknowledgement does not imply any evaluation of the final manuscript.

This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), Environmental System Science (ESS) Program as part of the River Corridor Science Focus Area (SFA) at the Pacific Northwest National Laboratory (PNNL). PNNL is operated by Battelle Memorial Institute for the DOE under Contract No. DE-AC05-76RL01830. WW was supported by the National Science Foundation Macrosystems Biology program under award EF-1926591.

Cite this data package with the appropriate DOI. Cite the associated manuscript in any work that that uses analyses or conclusions presented in the manuscript. To cite the paper:
Regier, P., Son, K., Chen, X, Fang, Y., Jiang, P., Taylor, M., Wollheim, W., Stegen, J. (2025). Allometric Scaling of Hyporheic Respiration across Basins in the Pacific Northwest United States. Journal of Geophysical Research: Biogeosciences. DOI: https://doi.org/10.1029/2024JG008344.

Citations:

- Regier, P., Son, K., Chen, X, Fang, Y., Jiang, P., Taylor, M., Wollheim, W., Stegen, J. (2025). Allometric Scaling of Hyporheic Respiration across Basins in the Pacific Northwest United States. Journal of Geophysical Research: Biogeosciences. DOI: https://doi.org/10.1029/2024JG008344.
- Forbes, B., Barnes, M., Boehnke, B. T., Chen, X., Cornwell, K., Delgado, D., et al. (2023). WHONDRS River Corridor Dissolved Oxygen, Temperature, Sediment Aerobic Respiration, Grain Size, and Water Chemistry from Machine-Learning-Informed Sites across the Contiguous United States (v4). Retrieved from https://data.ess-dive.lbl.gov/view/doi:10.15485/1923689
- Fang, Y., Chen, X., Gomez Velez, J., Zhang, X., Duan, Z., Hammond, G. E., et al. (2020). A multirate mass transfer model to represent the interaction of multicomponent biogeochemical processes between surface water and hyporheic zones (SWAT-MRMT-R 1.0). Geoscientific Model Development, 13(8), 3553–3569.
- Son, K., Fang, Y., Gomez-Velez, J. D., & Chen, X. (2022). Spatial microbial respiration variations in the hyporheic zones within the Columbia River Basin. Journal of Geophysical Research: Biogeosciences, n/a(n/a), e2021JG006654.

Contact

Peter Regier, peter.regier@pnnl.gov

Change History

Change history:

Data Package Version	Changes
Version 1 *April 2025*	Original data package publication