

## ESS-DIVE Metadata for grain size methods comparison Data Package

### Title:

Data for “Machine learning photogrammetric analysis of images provides a scalable approach to study riverbed grain size distributions”

### Alternative Identifiers:

### Abstract:

The grain size distribution, often characterized as median grain size (d50) in riverbeds determine hydrologic and biogeochemical function throughout watersheds, but the many methods used to measure or estimate d50 each have their own limitations. Here, we present a new photogrammetric method for estimating d50 from images, and explore how this method may help integrate manual and large-scale modeling d50 methods. This data package includes the data and scripts used in the manuscript entitled Alt+34Machine learning photogrammetric analysis of images provides a scalable approach to study riverbed grain size distributionsAlt+34. Source images can be found at 10.15485/1892052.

This dataset supports a broader study examining the drivers of spatial variability in sediment respiration rates in the Yakima River Basin. We acknowledge the Yakama Nation as owners and caretakers of the lands where we collected the data used in this project. 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.

### Keywords:

Grain size distribution  
Median grain size  
Machine learning  
Computer vision  
Photogrammetry  
Freshwater  
River  
Stream  
Riverbed  
Hydrology  
Biogeochemistry  
Respiration

### Data variables:

Median grain size (d50)  
Stream order  
Urban land cover  
Elevation  
Basin area  
Stream length  
Stream slope  
Precipitation as snow  
Evapotranspiration

Image suitability

**Commented [PR1]:** Aren't presented in <https://data.ess-dive.lbl.gov/view/doi:10.15485/1892052>, should we leave out?

**Pub date:**

**Data usage rights:**

Creative Commons Attribution

**Project:**

River Corridor and Watershed Biogeochemistry SFA

**Funding org:**

U.S. DOE > Office of Science > Biological and Environmental Research (BER)

**DOE Contracts:**

DOE Award #54737

**Related reference:**

[Insert any related references to manuscripts or other data packages.]

**Principal investigator:**

Scheibe

**Contact name:**

Regier

**Contact email:**

peter.regier@pnnl.gov

**Creators:**

Regier, Peter

Chen, Yunxiang

Son, Kyongho

Bao, Jie

Forbes, Brieanne

Goldman, Amy

Kaufman, Matt

Rod, Kenton

Stegen, James

**Start date:**

2021-08-30

**End date:**

2021-09-14

**Location description:**

Yakima River, WA, USA

**Coordinates:**

Refer to metadata spreadsheet

**Commented [RPJ2]:** Need to set up metadata spreadsheet...

**Methods:**

Images of riverbed grains were collected during a 2021 field campaign (171 images across 40 sites in the Yakima River basin, or YRB). A model using the You Only Look Once (YOLO) algorithm was trained to identify and measure individual grains, which was used to calculate the median grain size (d50) for each image. These values were compared to other d50 data sources in the YRB, including USGS manual measurements, and two continental-scale modeling approaches. We used simple spatial statistics and watershed characteristics to understand the similarities and differences between these d50 data sources, and explore if the YOLO approach can help integrate between sparse, accurate manual measurements, and spatially resolved, generalized model estimates.