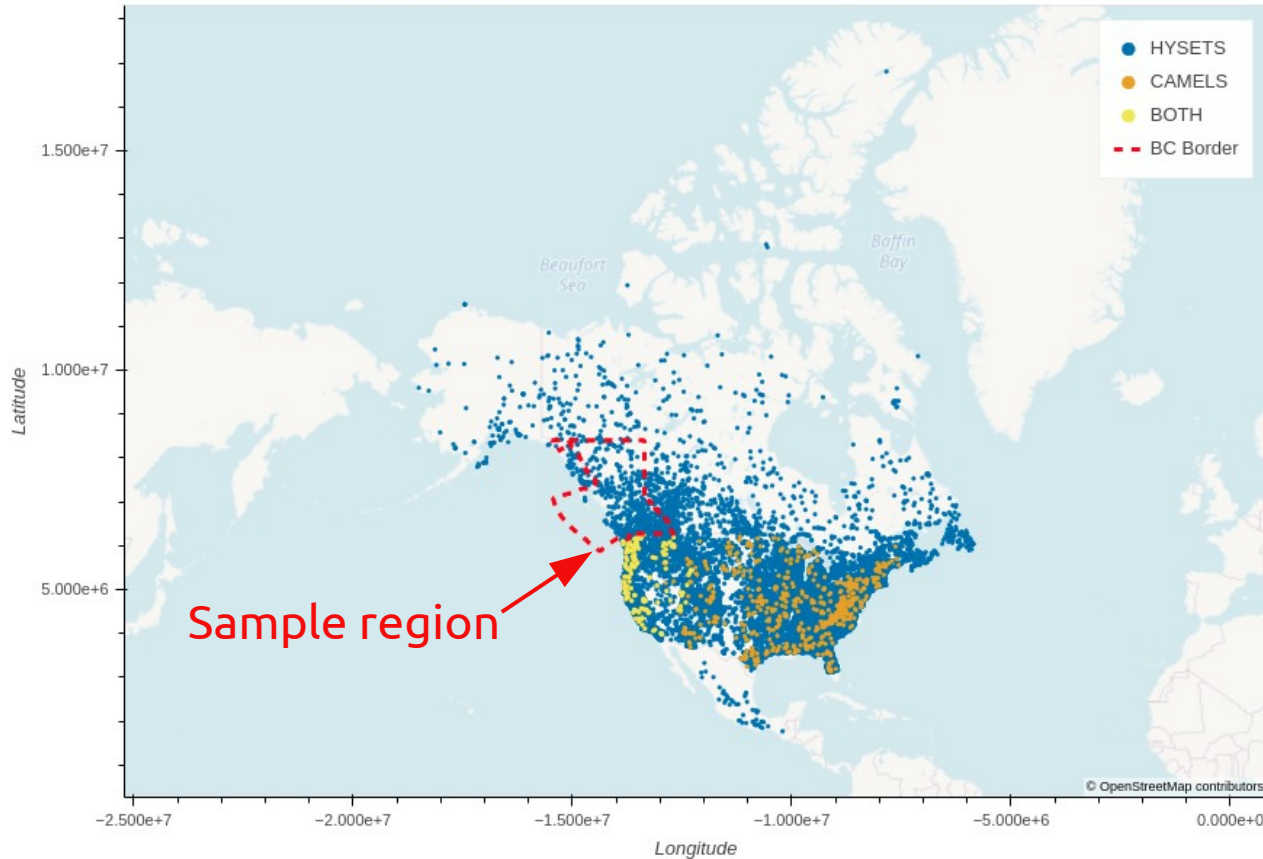


Large Sample Hydrology Dataset Validation



HYSETS

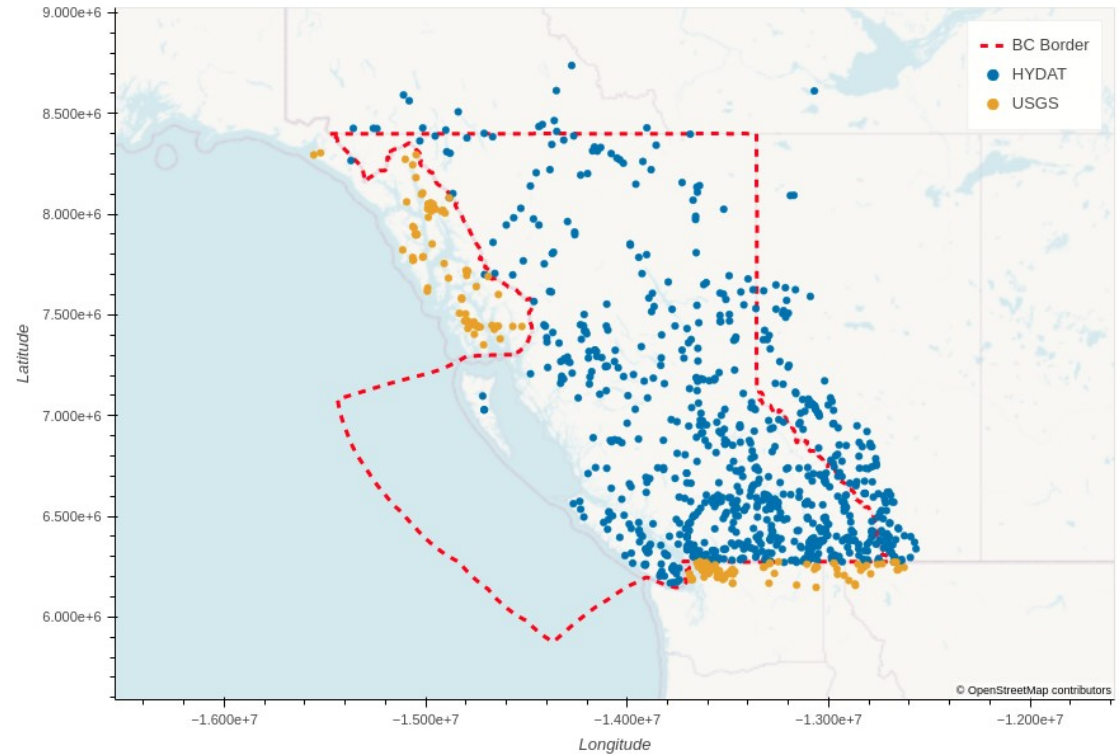
(Arsenault et al. 2019) [1]

+14K basins
streamflow
climate
physiography
geology/soil

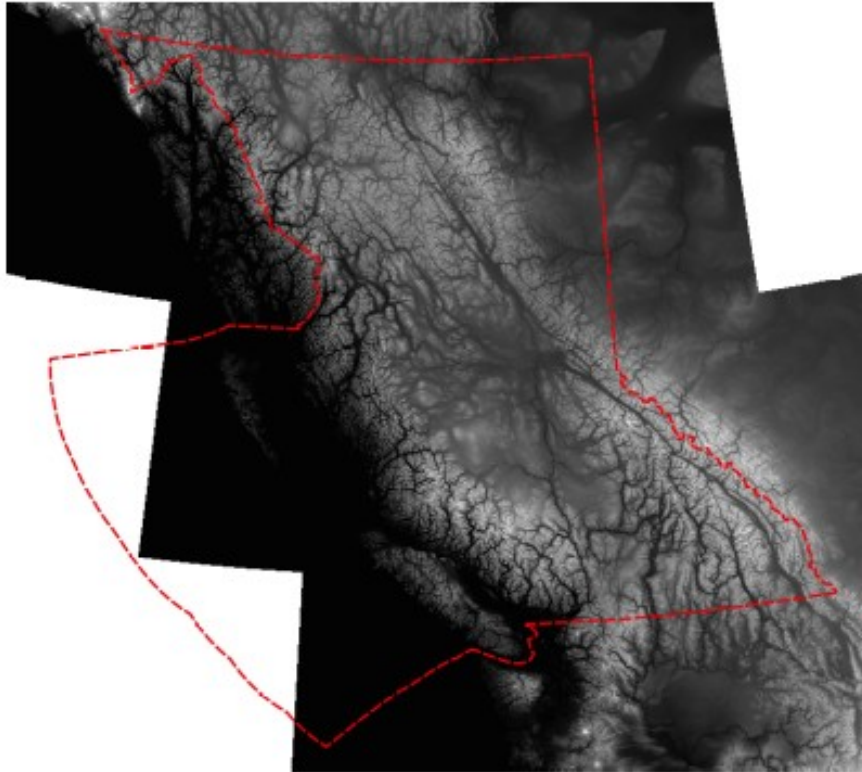
Why?

1. Basin attributes contain information **useful for prediction** [1, 2]
2. Characterize a large region for streamflow monitoring network optimization [3]
3. Convergence towards benchmark datasets [4]

N=986

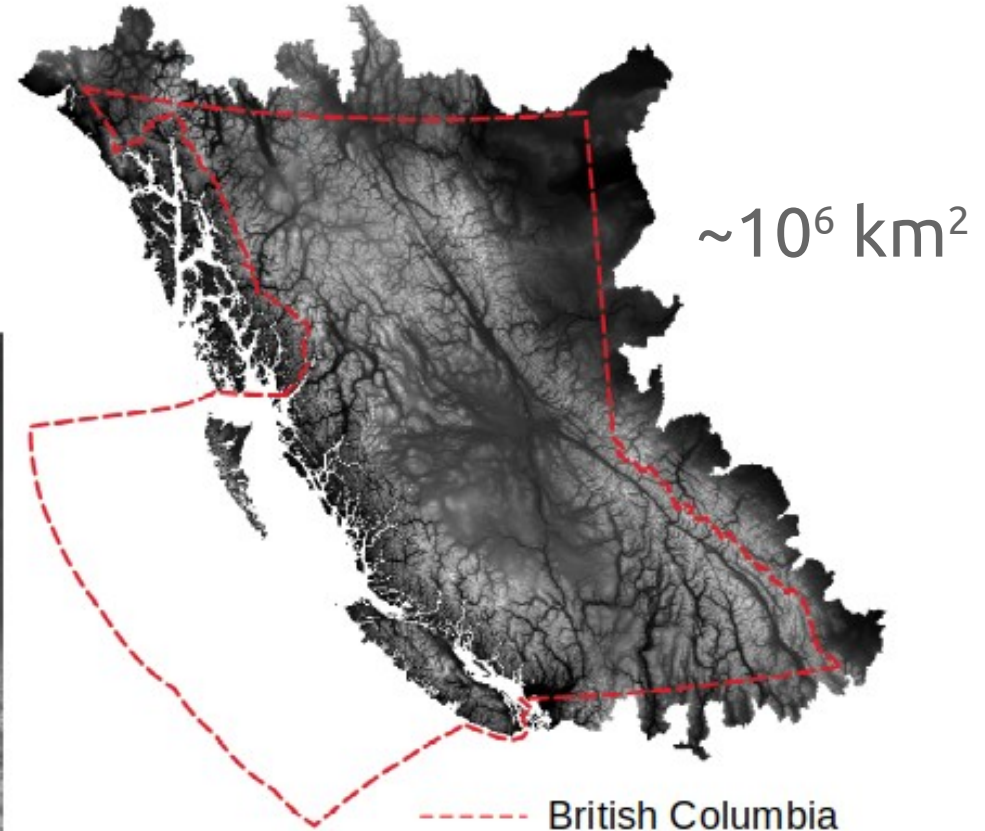


How?



Open Data

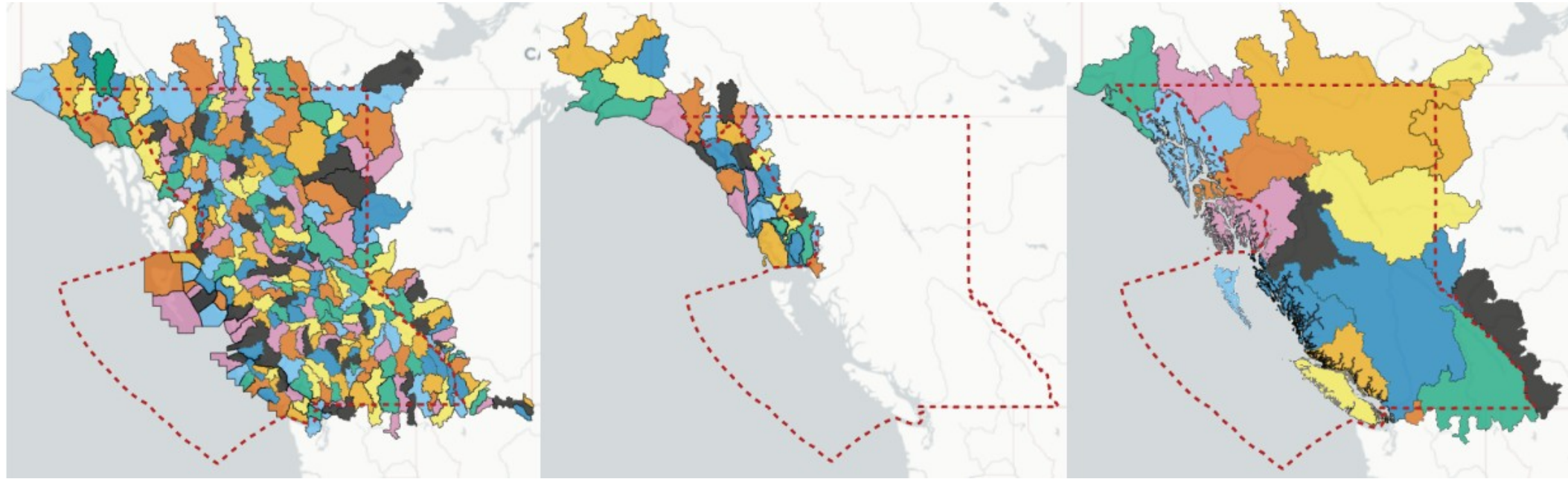
(EarthEnv DEM90 / USGS 3DEP)
(GLHYMPS/NALCMS/SNODAS/ERA5)



Open Tools

(GDAL,QGIS,Whitebox,Pysheds)

How?



WSC

+

USGS

=

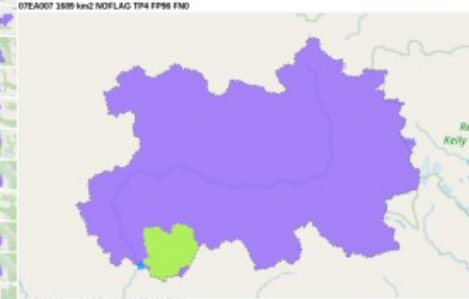
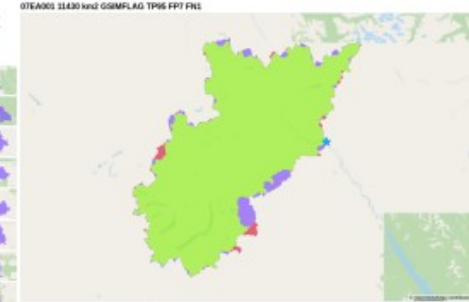
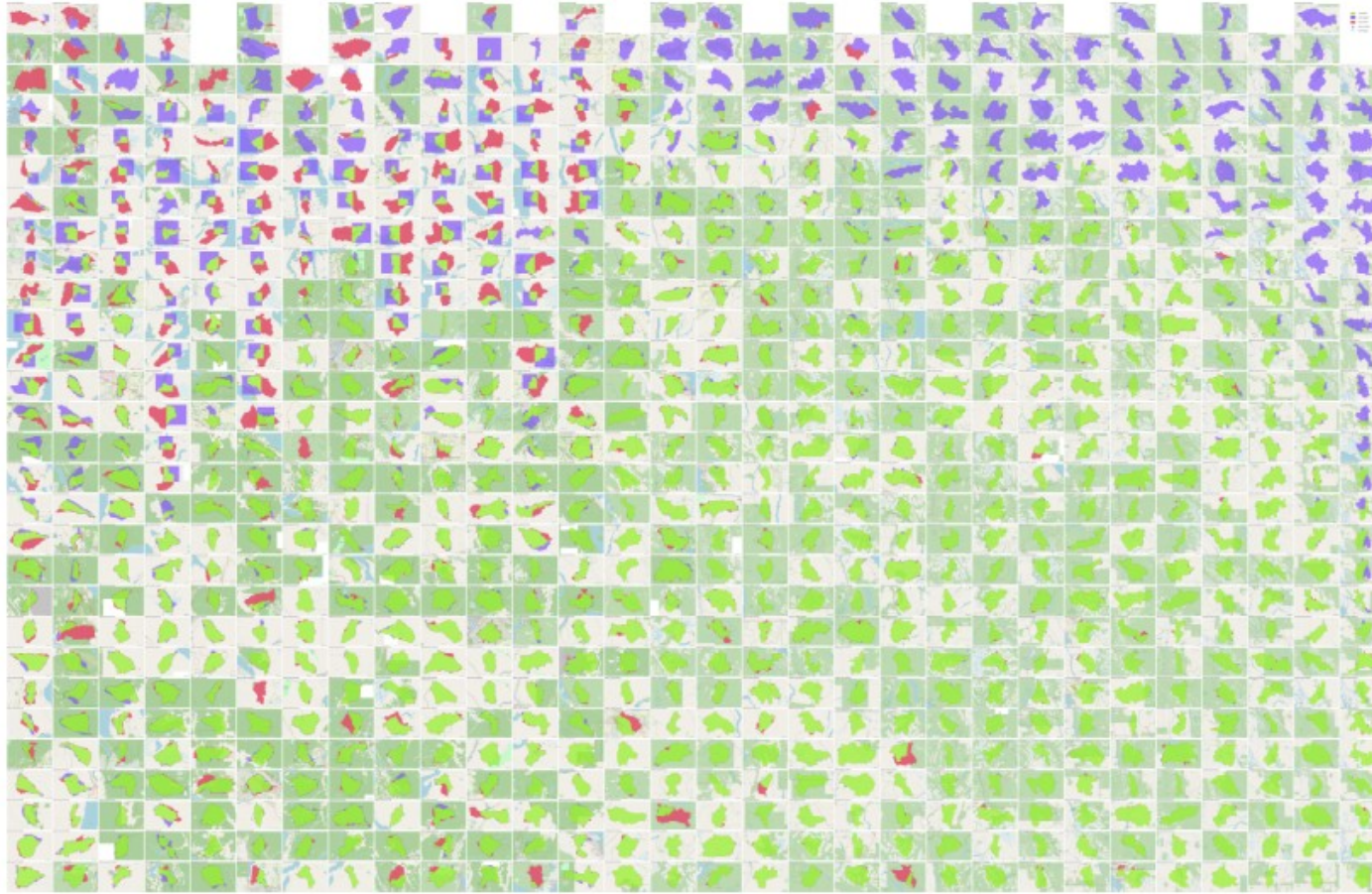
Complete Groups

Consolidation of Open Data Sources

Basin Validation

Large Sample Hydrology Dataset Validation

True positive False Negative False Positive



Bill Watterson

Results & Summary

1. Validation Accuracy

- worst at small and large scales
- manual input required

2. Error Classification

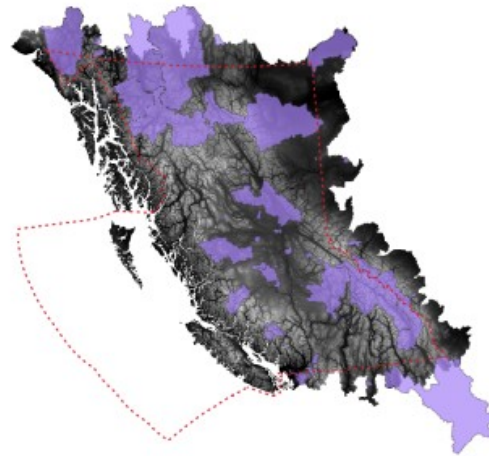
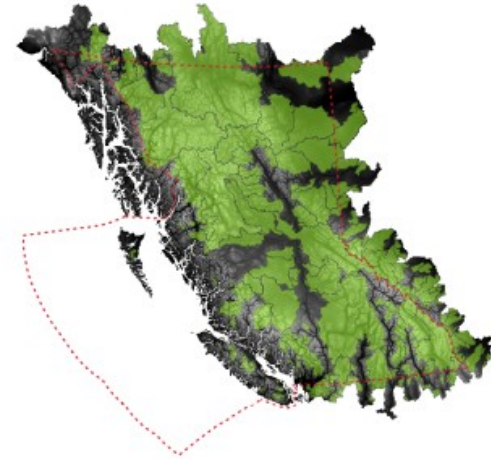
- small basins: low resolution
- large basins: DEM pre-processing

3. Decision Space

- derived streams $\neq$ real world

Open-source code repository:

github.com/dankovacek/hysets_validation



Thank You!

1. Arsenault, R., et al.: A comprehensive, multisource database for hydrometeorological modeling of 14,425 north american watersheds. Scientific Data 7(1), 1–12 (2020).
2. Kratzert, F., et al.: Toward improved predictions in ungauged basins: Exploiting the power of machine learning. Water Resources Research 55(12), 11344–11354 (2019).
3. Werstuck, C., Coulibaly, P.: Hydrometric network design using dual entropy multi-objective optimization in the Ottawa River basin. Hydrology Research 48(6), 1639–1651 (2017)
4. Hrachowitz, M., et al.: A decade of predictions in ungauged basins (pub)—a review. Hydrological sciences journal 58(6), 1198–1255 (2013).

Contact: dkovacek@mail.ubc.ca