Developing a Weather Mask: Two Ponds & GNSS-R

Prepared by:
Weston Hustace

CE 513 GIS in Water Resources
Dr. Tracy Arras - Spring 2023

Project Website:
https://web.engr.oregonstate.edu/~hustacew/ce513/proj/hustacew_ce513_projwebsite.html

Weston Hustace
CE 513 GIS in Water Resources
Final Project - Report
3/16/2023

Introduction:

Surface water is a vital component of the global water cycle which directly influences local drinking water quality and ecosystem health. Recent anthropogenic climate change drives heightened irregularity and intensity of precipitation recharge to these areas, exposing lakes and ponds to longer uninterrupted periods of evaporation, as well as greater likelihood for pollutants from increased occurrence of storm-related surface runoff.

Additional nutrient input (by way of runoff contaminates) introduced to a seasonally stagnant water body can produce ideal conditions for algal growth. Outcomes of such blooms can include: (a). hypoxic eutrophication of the aquatic habitat, (b.) algae containing poisonous toxins that may permeate into drinking water, or (c.) bioaccumulate in fish or shellfish. All of which may impact the quality or availability of food sources for the local community, ecosystem, or industry.

For the reasons stated, monitoring and prediction of harmful algal blooms (HABs) is necessary to inform water resource management decisions. However, it can present as a challenging task. There are many varieties of HABs, and regular in situ sampling over large areas or many surface water bodies can be time-consuming, personnel-intensive, and expensive. A commonly used method to circumvent these limitations includes remote sensing techniques, such as classification or spectral analysis of satellite imagery. This approach can gleam information about the reflectivity of surface water over large areas, while remaining relatively inexpensive. However, publicly accessible satellite imagery such as Sentinel-2 or Landsat are dependent on cloud coverage, and offer relatively coarse spatial resolution (10m, 30m) in relation to some smaller surface water bodies.

GNSS Reflectometry, or GNSS-R, utilizes GNSS multipath signals from a geodetic-grade GNSS receiver to be able to characterize attributes about the surrounding environment. Variations of this case studies have shown this method capable of monitoring: (1) water level fluctuation, (2) vegetation growth, (3) snow depth, (4) water content of soil or vegetation, in addition to other published and developing applications (Larson et al., 2013), (Larson & Small, 2014). It is plausible this method could be used to measure the area and growth rate of vegetation or algal scum on the surface of a freshwater lake. If feasible, this could be a first step towards developing a tool that could contribute towards Harmful Algal Bloom (HAB) monitoring and water resource management.

This course-specific report will focus on a necessary step in the data post-processing workflow for the overall GNSS-R surface vegetation study: developing a precipitation data mask. This weather mask will be used to remove periods of processed data from the results when precipitation occurred, to avoid introducing potential bias.

Site Description:

Stahlbush Pond (Figure 1) is a small water body connected to a larger network of ponds just east of Corvallis, OR. It lies on Knife River Corporation property, located off of HWY 34 and bordering the Willamette River. The pond area of interest (AOI) is approximately 13 acres, center located at 44.546425°, -123.239518°. The antenna setup is located at 44.546326°, -123.240536° (Figure 2).

Bonner Lake (Figure 1) is a small, human-made water body that lies west of Hoskins, OR. The lake AOI is approximately 2.5 acres, center located at 44.680450°, -123.489414°. The antenna setup is located at 44.680504°, -123.489589° (Figure 2).

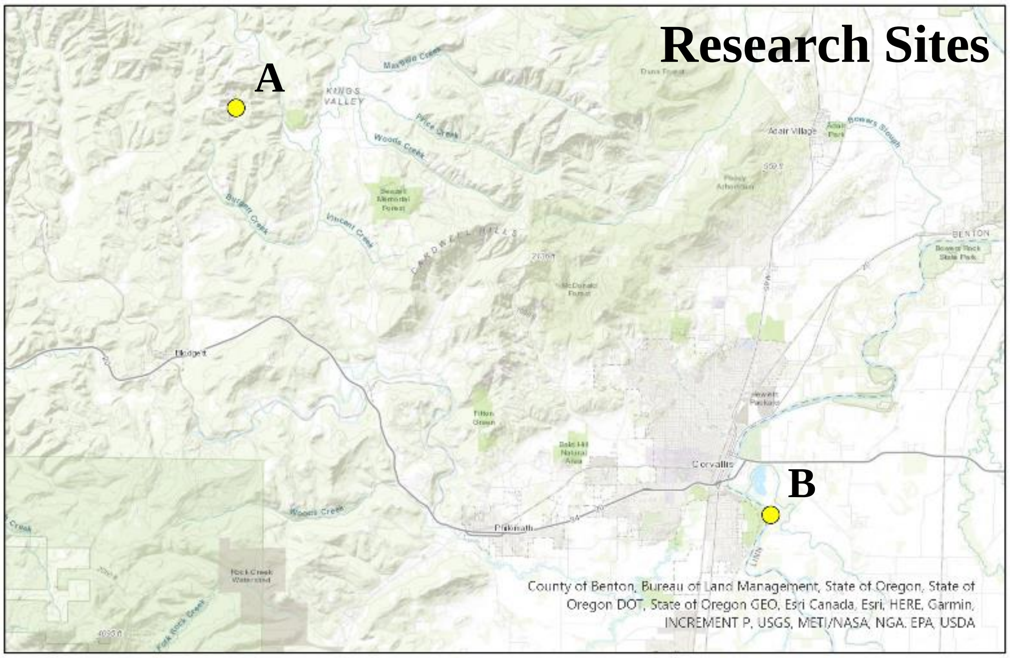


Figure 1: Location of project sites Bonner Lake (A) and Stahlbush Pond (B) in relation to Corvallis, Oregon.

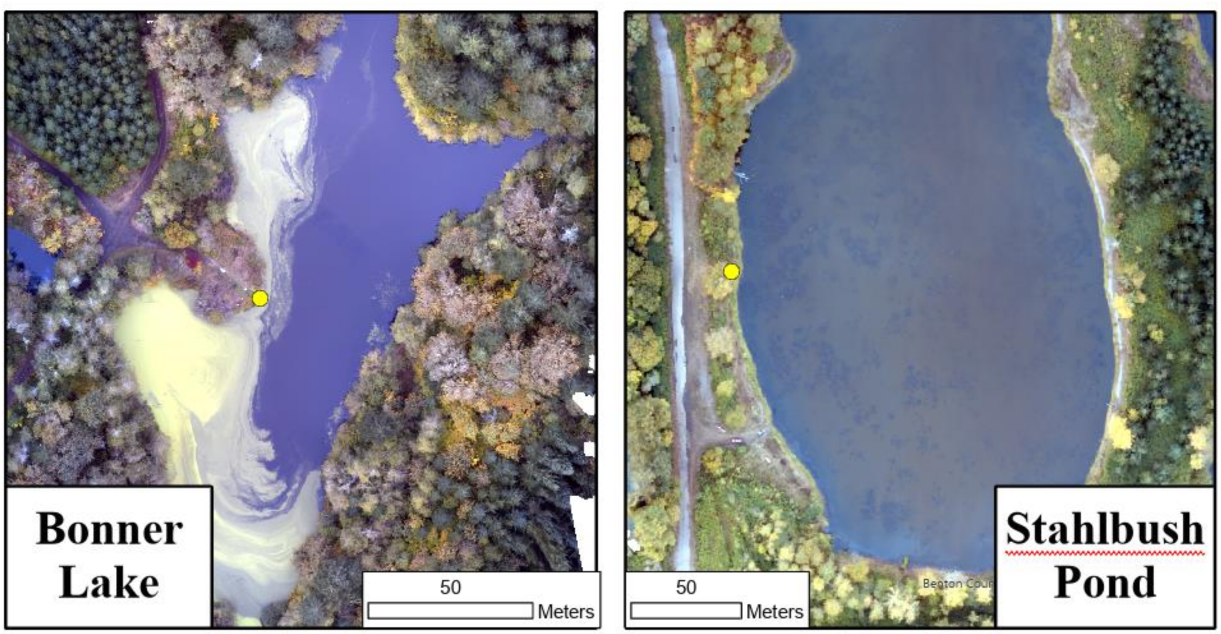


Figure 2: UAV orthoimagery of acquisition sites in Oregon: Bonner Lake (left) and Stahlbush Pond (right). Site-respective GNSS antenna setup locations are displayed in yellow.

Weston Hustace
CE 513 GIS in Water Resources
Final Project - Report
3/16/2023

Data:

Original scope of project focused on summarizing and validating in situ weather gauges with calculated weather mask, which held logged data for duration of every GNSS-R acquisition day in 2022. This collected dataset had quality issues relating to unexpected gaps in tabular records and regular inconsistencies between redundant sensors. As a pivot in focus, the planned weather interpolation methodology was preserved with a change in intent from validation, to calculation of precipitation levels at each site. A pre-processed and available weather dataset was chosen to act as a reference dataset to the interpolation results.

The Global Historical Climatology Network (GHCN) is an international web of more than 100,000 land-based weather gauges. The stations primarily log precipitation data, but subsets of the network track additional climate metrics. The data (accessible through NOAA NCEI) are provided in the form of daily mean values by station, date, or specified region (GHCNd | NCEI, n.d.). The CSV filetype, associated metadata nor literature specify reference information about the stations, so WGS84 was assumed based on nature of the global distribution of the network.

PRISM weather data is processed and made available by the Prism Climate Group based out of Oregon State University. Historical normal and daily summaries of precipitation, temperature, and other climate metrics are available in raster form of the entire contiguous United States (PRISM Climate Group, OSU, n.d.). Native resolution is 800m cell size, but free, publicly available datasets are in the form of 4km resolution, which will be used.

Dataset Title	Reference	Units	Source
Global Historical Climatology Network (GHCN)	WGS84	metric	https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily
PRISM	NAD 83	metric	https://prism.oregonstate.edu/

Table 1: Datasets used in project along with associated metadata and information

Methods:

GHCN data, while accessible through a webGIS portal or ftp server through NOAA NCEI, includes all global climate station data in a single file for the given region or timeframe. Global records for the year of 2022 were acquired for precipitation data. The single CSV file required a programmatic approach utilizing ArcPy and other packages to append station info from a separate metadata text file, and reduce the stations included to only those within a 25km radius of the GNSS-R sites. The files were then split by the day field into daily tabular data and named appropriately.

PRISM climate products were acquired from the Oregon State hosted site, where requested data included daily average precipitation values across the US, downloaded in the form of .bil raster files for the extent of 2022.

Once daily CSV files were built, and the reference climate (2022) acquired, the first day of 2022 (GHCN CSV) was imported into ArcGIS Pro to develop a processing roadmap to be reproduced for the entire year. With the GHCN single day station information as an input, various Kriging interpolation approaches were attempted based off preliminary results and review of model focus. Empirical Bayesian Kriging (EBK) was chosen for the methodology for this project based on the flexibility of input dataset characteristics, specifically allowing non-normally distributed samples, with potential for trends.

A ModelBuilder workflow was developed with the sample daily data as sole input, encompassing the mapping of tabular coordinates in the GHCN CSV files to feature classes and setting parameters of the EBK interpolation step. Once a prediction map is produced for the given precipitation readings, the cell values overlaid at each research site will be appended to the site point data attribute table with the 'Extract Multi-Values to Points' geoprocessing tool (Figure 3).

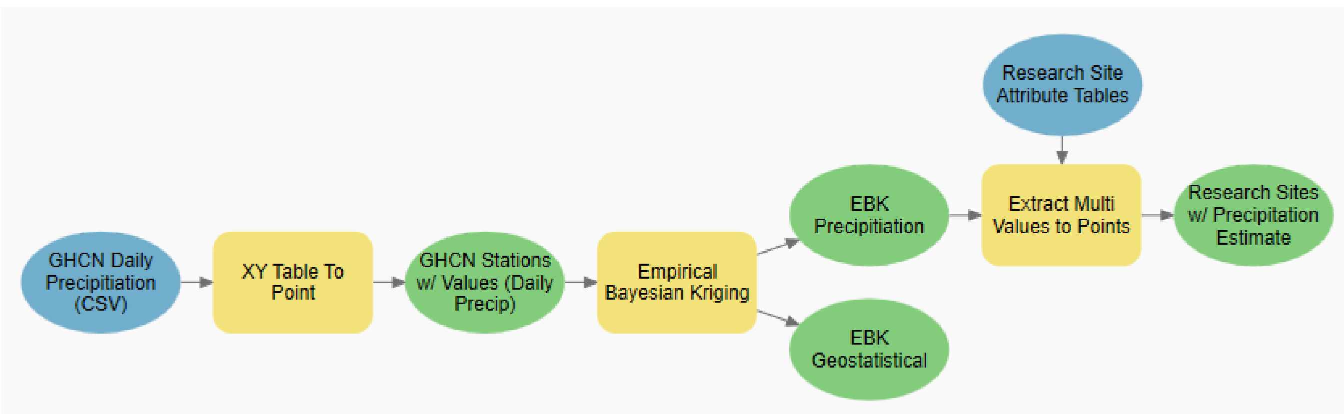


Figure 3: A processing workflow produced in ArcGIS Pro's ModelBuilder, to be converted into code for iteration on daily climate data of 2022.

The finalized ModelBuilder was exported into python code to modify and automate the process from GHCN input to the completion of the EBK step. The Kriging estimation maps were produced for the available days in 2022, then batch imported into the ArcGIS Pro software. In tandem, the full set of GHCN interpolated maps and the reference PRISM precipitation maps were attributed to the two receiver site feature class attribute tables, following the last step laid out in Figure 3.

The collected site precipitation values were exported as tabular data to assess, format appropriately, and align units. The two datasets were summarized and compared.

Results:

Results were summarized in two forms; a time series of actual precipitation values between the GHCN and PRISM data, then a performed binary threshold of 0.005 inches to denote each day with the occurrence of rainfall or lack thereof.

Plotting resulted in good visual alignment of the data (Figures 4 and 5), while comparing agreement of the two mask datasets for each site resulted in an agreement of 96% at the Bonner Lake site and 91% at Stahlbush Pond. One significant observation recorded was a high number of interpolation failures out of the GHCN daily files (77 of 365 days) despite sufficient number of input station values. It's extremely likely this is due to days of no rainfall, when compared to PRISM values, where substituting the lack of data with zero values results in 100% agreement.

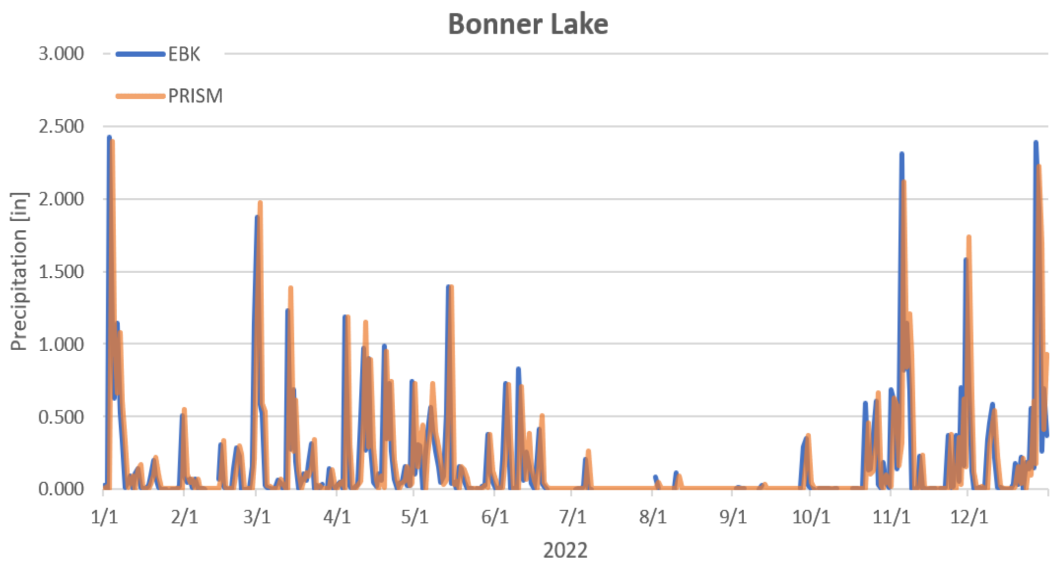


Figure 4: Time series of calculated PRISM and interpolated GHCN precipitation, Bonner Lake.

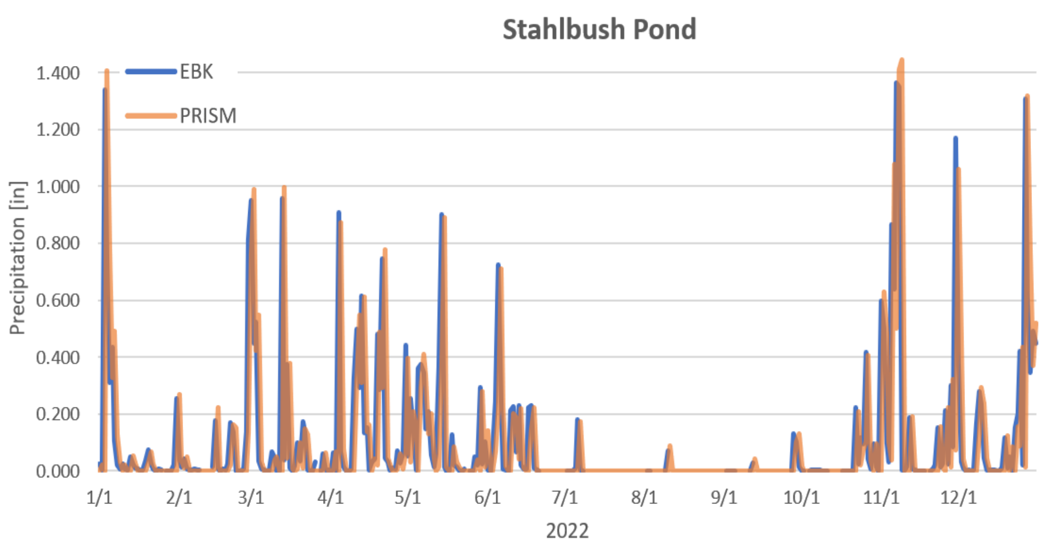


Figure 5: Time series of calculated PRISM, interpolated GHCN precipitation, Stahlbush Pond.

Overall, this course project was a successful exploration and introduction of weather data and the world of interpolation processes. It provided evidence the workflow could be a helpful step in the right direction for simple climate parameters when focused on singular metrics, especially in the framework of the agreement results between the produced prediction values and the truth dataset. It may be most practical to utilize the PRISM data on its own for developing a weather mask, where the climate models take into account many more parameters in addition to local gauges and the geospatial relationship. If revisiting this topic, it would be helpful to continue to experiment with EBK and other Kriging settings, as well as approach other interpolation methods like splines and IDM.

Works Cited

Global Historical Climatology Network daily (GHCNd) | National Centers for Environmental Information (NCEI). (n.d.). Retrieved March 14, 2023, from https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily

Kriging Interpolation—The Prediction Is Strong in this One—GIS Geography. (n.d.). Retrieved March 14, 2023, from https://gisgeography.com/kriging-interpolation-prediction/

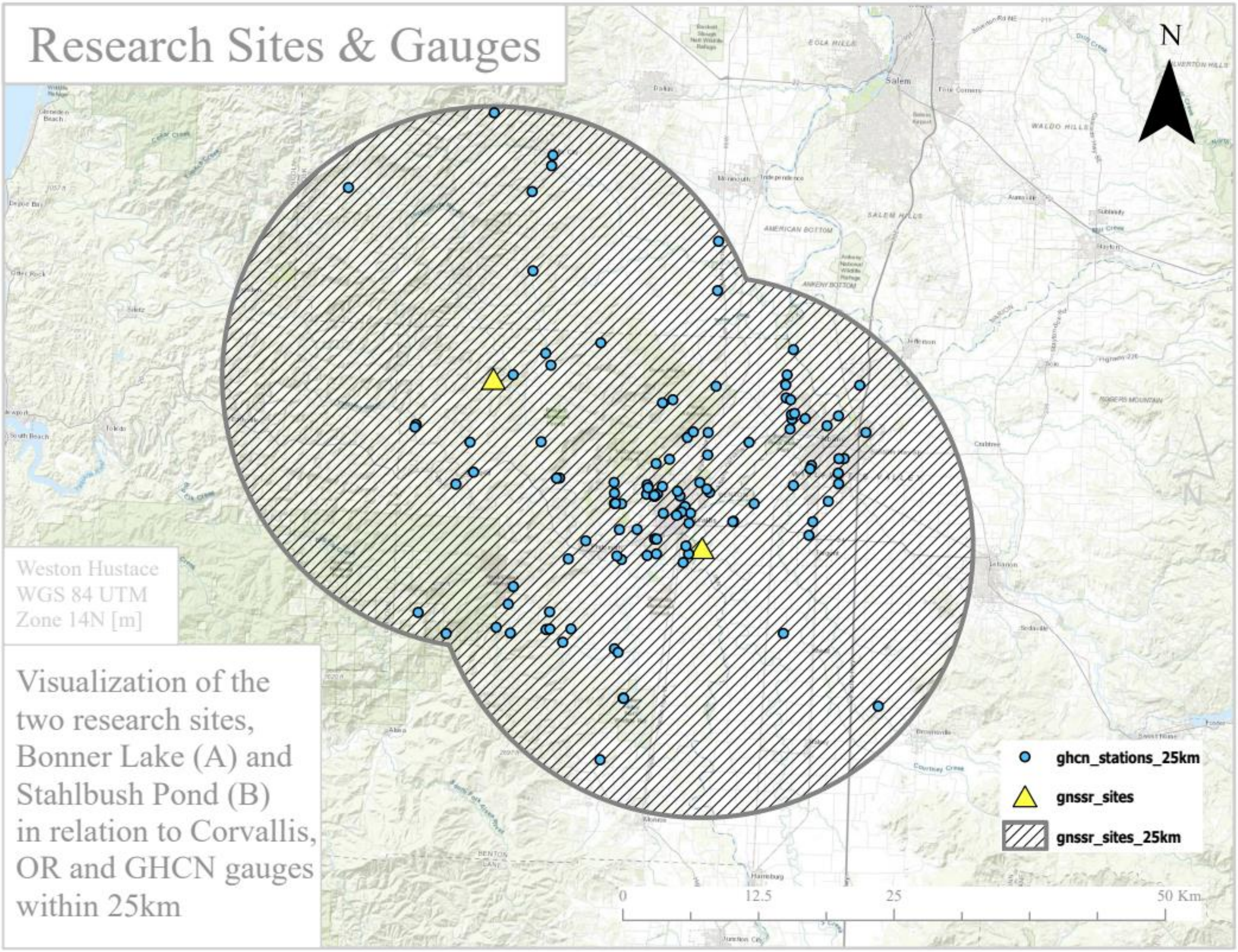
Larson, K. M., Ray, R. D., Nievinski, F. G., & Freymueller, J. T. (2013). The Accidental Tide Gauge: A GPS Reflection Case Study From Kachemak Bay, Alaska. *IEEE Geoscience and Remote Sensing Letters*, *10*(5), 1200–1204. https://doi.org/10.1109/LGRS.2012.2236075

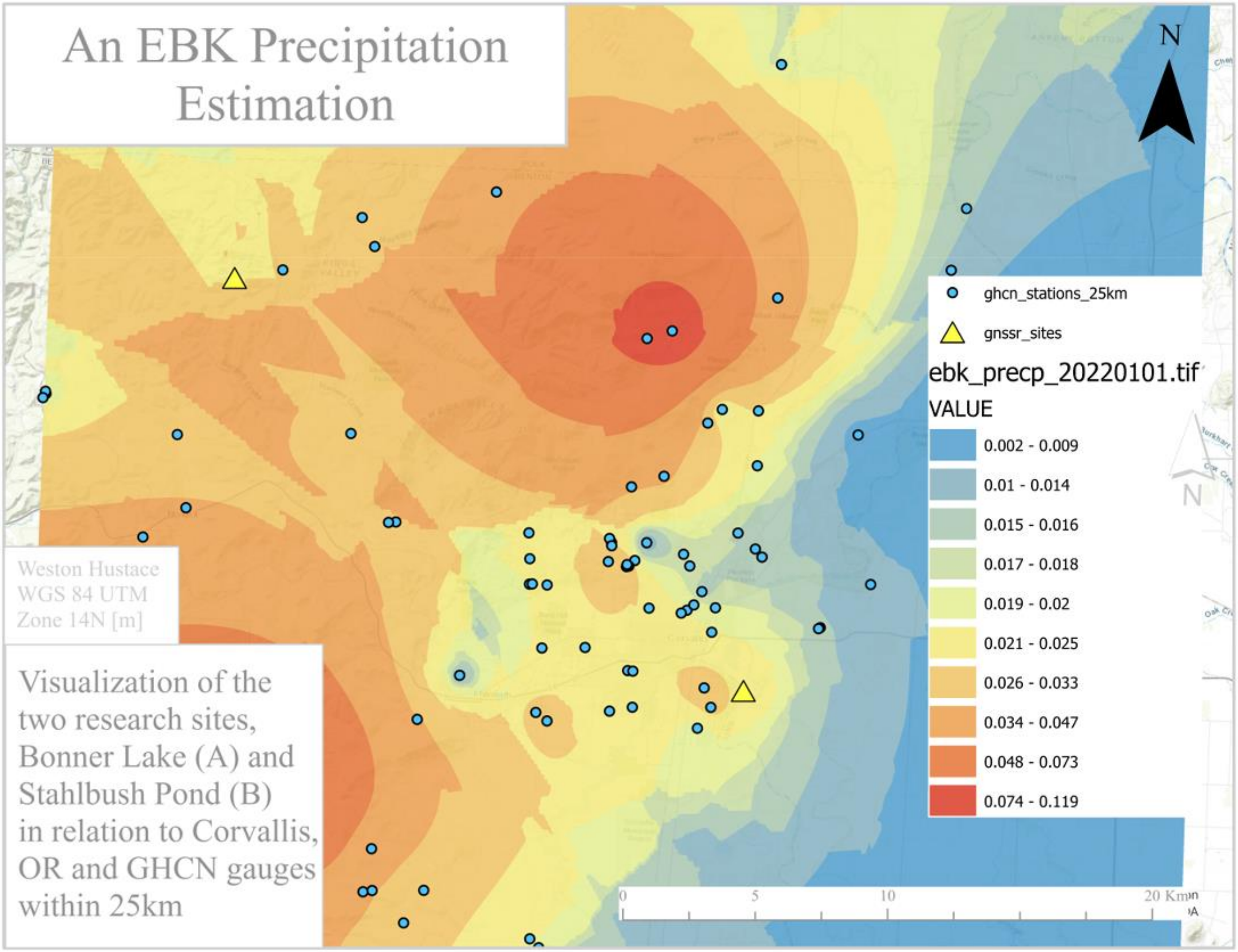
Larson, K. M., & Small, E. E. (2014). Normalized Microwave Reflection Index: A Vegetation Measurement Derived From GPS Networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(5), 1501–1511. https://doi.org/10.1109/JSTARS.2014.2300116

PRISM Climate Group, Oregon State U. (n.d.). Retrieved March 16, 2023, from https://prism.oregonstate.edu/

Weston Hustace
CE 513 GIS in Water Resources
Final Project - Report
3/16/2023

Appendix:




An EBK Precipitation Estimation
Weston Hustace
WGS 84 UTM
Zone 14N [m]
Visualization of the two research sites, Bonner Lake (A) and Stahlbush Pond (B) in relation to Corvallis, OR and GHCN gauges within 25km
N
ghcn_stations_25km
gnssr_sites
ebk_precp_20220101.tif
VALUE
0.002 - 0.009
0.01 - 0.014
0.015 - 0.016
0.017 - 0.018
0.019 - 0.02
0.021 - 0.025
0.026 - 0.033
0.034 - 0.047
0.048 - 0.073
0.074 - 0.119
0
5
10
20 Km