Upper Missouri River Basin Water Resources

Enhancing Flood and Drought Monitoring through Fractional Available Water Analysis in the Upper Missouri River Basin

Spring 2025 | North Carolina – NCEI
April 4th, 2025

Authors: Townes Ellum (Analytical Mechanics Associates), Chloe Green (Analytical Mechanics Associates), Angelina Herbert (Analytical Mechanics Associates), Valerie Scull (Analytical Mechanics Associates)

Abstract:
Soil moisture is a critical factor in flood and drought forecasting, climate response, and agricultural planning. Situated in the northern region of the Great Plains, the Upper Missouri River Basin is subject to rapid shifts between wet and dry conditions, but effective monitoring in this region remains a challenge due to limited in situ soil moisture data availability. DEVELOP partnered with the NOAA National Weather Service's Weather Forecast Offices in Bismarck, Grand Forks, and Rapid City, the North Dakota State University, North Dakota State Climate Office, and the North Dakota Agricultural Weather Network Center. Our objectives were (1) to validate estimates of fractional available water (FAW) from NASA's Soil Moisture Active Passive (SMAP) satellite against in situ estimates, and (2) to determine the feasibility of incorporating NASA Earth observations in drought and flood monitoring. Our analysis revealed a relatively strong correlation between SMAP-derived FAW and in situ measurements ($R^2 = 0.62$). To explore the role of FAW as a drought indicator, we compared SMAP-derived FAW with streamflow and vegetation health datasets. While the correlations between these datasets were weak, time series analysis showed a relationship between low early growing season SMAP-derived FAW and low streamflow and diminished vegetation health in later months of 2021, a severe drought year. This analysis provided partners with resources in identifying drought signals using FAW, which can advance the use of soil moisture data in hydrological monitoring. These results also suggest that SMAP can be a complementary tool to use with in situ soil moisture monitoring networks.

Key Terms: remote sensing, SMAP, fractional available water, Mesonet, Upper Missouri River Basin, drought, vegetation health, streamflow

Advisors: Ronald D. Leeper (NOAA National Centers for Environmental Information, North Carolina Institute for Climate Studies), Erik Krueger (Oklahoma State University, Department of Plant and Soil Sciences)

Lead: Tallis Monteiro (North Carolina – NCEI)

1. Introduction

The Upper Missouri River Basin (UMRB) is a hydrologic drainage system stretching from the headwaters of the Missouri River to its confluence with the Big Sioux River. The UMRB exists in a natural state of wet and dry seasonality and often experiences rapid shifts between wet and dry conditions (Woodhouse & Wise, 2020). Climate simulations predict an increase in frequency of both drought and flooding events for this region, which can be significantly affected by soil moisture levels (Badger et al., 2018). Soil moisture is a critical factor in flood and drought monitoring, impacting emergency preparedness, climate response, and agricultural planning, but effective flood and drought monitoring for the UMRB remains a challenge due to limited soil moisture data availability. Basin agencies employ several tools to monitor drought conditions, including the U.S. Drought Monitor (USDM) and the U.S. Climate Reference Network (USCRN), but these tools may not be ideal for agricultural drought monitoring because the USDM considers all forms of drought in a convergence of evidence approach and there are too few USCRN stations within this region (Krueger et al., 2024; Zargar et al., 2011). Soil moisture measurements provide a more precise method for monitoring agricultural drought, as soil moisture is the primary variable by which agricultural drought is defined (Krueger et al., 2019). The collection of in situ data, however, is spatially limited by the locations of collection sites. Remote sensing provides adequate spatial coverage, yet it is often temporally limited and can only provide surface data on soil moisture (Leeper et al., 2023). Monitoring hydrological extremes requires, "both a temporal record to adequately standardize or place soil moisture conditions into historical context and sufficient spatial resolution to monitor evolving conditions" (Leeper et al., 2023). For this reason, combining in situ and remotely sensed soil moisture data can be a powerful method for monitoring soil moisture conditions and hydrological extremes.

Recent studies indicate that remotely sensed data products from satellites such as NASA's Soil Moisture Active Passive (SMAP) and the European Space Agency's Active, Passive and Combined satellites can accurately monitor soil moisture levels (Leeper et al., 2023; Xu et al., 2018). In situ measurements are useful for monitoring soil moisture levels over time for a specific location. However, soil moisture levels can vary drastically over short distances due to factors such as elevation, slope, soil texture, land cover, and precipitation patterns (Bell et al., 2010; Brocca et al., 2007; Manns et al., 2014). Satellite datasets alone can account for frequency of flood and drought events, but not necessarily for the magnitude of their extremes. However, remotely sensed standardized soil moisture combined with in situ data has been successful in detecting both the frequency and magnitude of extreme conditions (Leeper et al., 2023).

The National Oceanic and Atmospheric Association (NOAA) National Weather Service's Weather Forecast Offices in Grand Forks and Bismarck, North Dakota and Rapid City, South Dakota provide weather, water, and climate data to agricultural, land, and emergency management stakeholders to advise decisions and weather forecasting in the UMRB (US Department of Commerce, NOAA, National Weather Service, 2020). The National Weather Service seeks to incorporate remotely sensed soil moisture data to improve forecasting in areas that lack extensive Mesonet coverage. We also partnered with two organizations from North Dakota State University: the North Dakota State Climate Office and the North Dakota Agricultural Weather Network (NDAWN) Center. The North Dakota State Climate Office provides Northern Plains weather and climate data to public institutions, government agencies, and citizens. NDAWN manages North Dakota's Mesonet data incorporating over 200 stations across the state and bordering regions (NDAWN, 2025). To inform future drought and flood forecasting, NOAA National Weather Service Weather Forecast, North Dakota State Climate Office, and NDAWN partnered with NASA DEVELOP to validate NASA Earth observations with in situ soil moisture data to provide decision-support products for the UMRB (Figure 1).

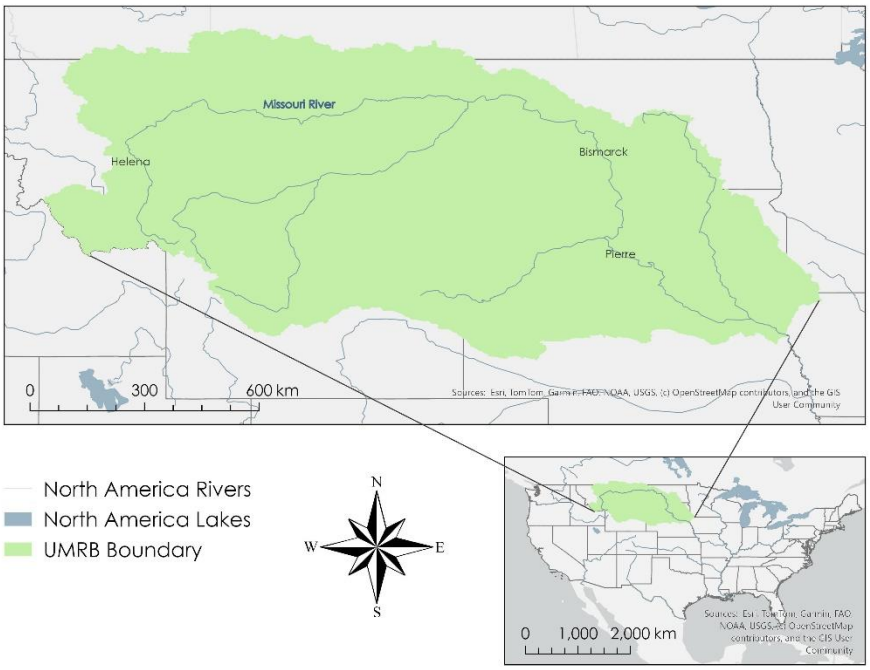

Figure 1. Map of study area containing the Upper Missouri River Basin and major lakes and rivers

This project aimed to create a set of working analyses and visualizations to assess the feasibility of using SMAP satellite data to measure a critical component of soil moisture, fractional available water (FAW). FAW is the fraction of soil moisture that is available to plants (Krueger et al., 2019). For the purposes of this project, the study period ranged from 2020 to 2024 due to the availability of SMAP-derived FAW data and Mesonet data. The primary goal of this project was to validate SMAP-derived FAW values with in situ Mesonet values to determine the feasibility of incorporating NASA Earth observations into drought and flood monitoring. Additionally, we compared SMAP-derived FAW values with two ancillary datasets, streamflow and vegetation health, to analyze the relationship of these data to SMAP-derived FAW and explore the role of FAW as a drought indicator.

2. Methodology

2.1 Data Acquisition

2.1.1 SMAP and Mesonet Soil Moisture

We acquired soil moisture data from NASA's SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture, Version 6 satellite (National Snow and Ice Data Center, 2020). The SMAP satellite utilizes the passive radiometer instrument to collect soil moisture twice daily across the entirety of Earth's surface. Prior to our processing our science advisor, Ronald D. Leeper, downloaded SMAP soil moisture data and pre-processed it by averaging morning and afternoon over passes for each day between 2020 and 2024. We acquired weekly Mesonet station soil moisture data recorded by the UMRB State Mesonets ranging from February 2021 to March 2025. Station data used in this study were part of the UMRB station soil moisture and plains snow monitoring build-out lead by the US Army Core of Engineers (USACE) and the National Integrated Drought Information System. An initiative to increase the number of monitoring stations across the UMRB to 500 stations is being proposed for Montana, North and South Dakota, Wyoming, and Nebraska Mesonet networks (USACE Upper Missouri River Basin Soil Moisture and Plains Snow Monitoring Build-Out, 2020). As monitoring of the region continues to expand, this project intends to review the feasibility of incorporating satellite-derived FAW as supplemental data to stakeholders These data contained fields including date, FAW values for the top 5 centimeters of the soil profile, station identifiers, and coordinate data.

As with the SMAP data, the Mesonet data were downloaded and pre-processed by our science advisor prior to our processing. To ensure each basin contained an adequate number of Mesonet stations for proper data validation, we aggregated FAW data at the Hydrologic Unit Code (HUC) 6 level. HUC refers to the US Geological Survey's classification system for hydrologic areas. We acquired a shapefile containing polygons of all 25 HUC-6 subbasins within the larger UMRB from the US Geological Survey Hydrography Data Downloader (Access National Hydrography Products, n.d.). Afterward, we extracted unique Mesonet station identifiers and coordinate data from the Mesonet FAW dataset to create a shapefile containing all 167 Mesonet stations within the UMRB.

2.1.2 Streamflow and Vegetation Health

To further assess the feasibility of using SMAP data in flood and drought monitoring, we selected four smaller HUC-8 level subbasins for streamflow and vegetation health correlation analysis: James Headwaters, Cedar, Beaver, and Upper Cannonball (Figure 2).

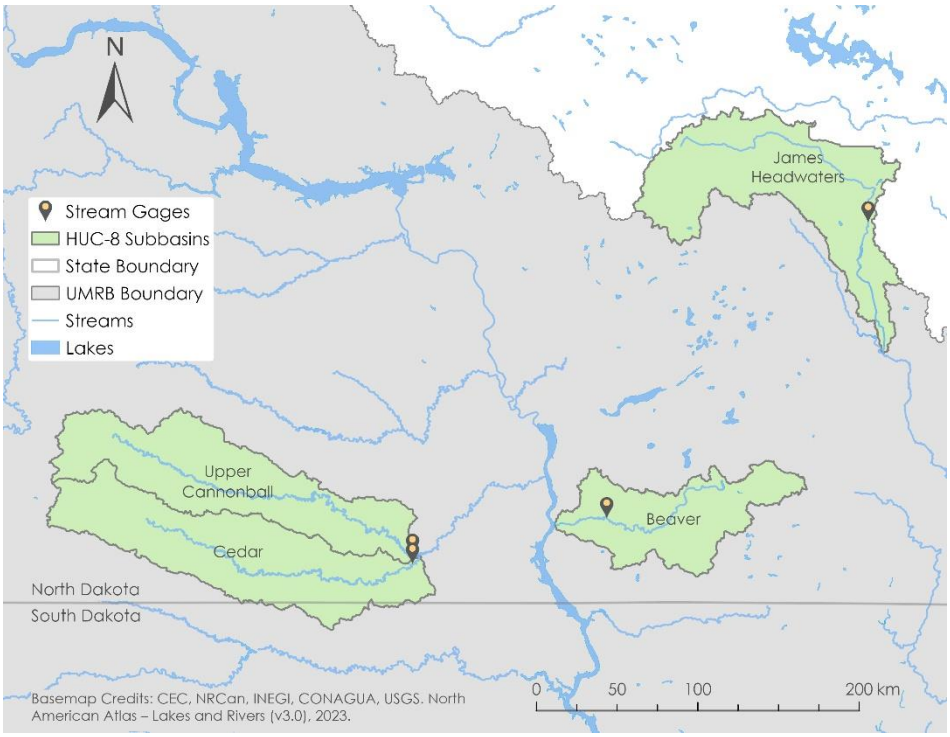

Figure 2. Map of selected HUC-8 subbasins and stream gages within the Upper Missouri River Basin

We used the following criteria to select these subbasins: each contains a headwaters stream, none contain tributaries that originate outside the subbasin, and none contain dams with controlled spillways. We chose one stream gage at the exit of each subbasin to determine the basin's streamflow. This criterion ensured each gage captured the flow of its respective subbasin. We acquired daily streamflow data as discharge in cubic feet per second from the U.S. Geological Survey National Water Information System using the "dataRetrieval" R package (Table A1).

We retrieved the Vegetation Health Index (VHI) dataset from NOAA's National Environmental Satellite, Data, and Information Service (NESDIS) Center for Satellite Applications and Research (STAR) tool for 2020-2024 (Table 1). STAR is a collaborative global and regional Vegetation Health tool. This tool combines readings from countless NOAA satellites using the Advanced Very High Resolution Radiometer (AVHRR) sensor that sits on board their afternoon polar-orbiting satellites: namely NOAA-7, 9, 11, 14, 16, 18 and 19,

and Suomi National Polar-Orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) (National Oceanic Atmospheric Administration, 2019).

We developed R code using the ncdf4 package (Pierce, 2023) to access VHP data, which were NetCDF file outputs for each week starting with the week of January 7, 2020. Before downloading any files, we applied filters to the Blended VH 4 km resolution data (NOAA, 2019) to ensure strictly the vegetation health data were downloaded from the dataset. From the point of access, we uploaded the files to a cloud-based storage bucket for remote reference in R Studio.

Table 1
List of monitors, sensors, and data products used in this project

Data Type	Product	Source	Parameter	Years Used
Earth observation	SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture V006	NASA National Snow and Ice Data Center DAAC	Soil Moisture	2020-2024
In situ	US Army Core of Engineers and National Integrated Drought Information System Build-Out, UMRB State Mesonets	NDAWN, Nebraska Mesonet, Montana Mesonet, Mesonet at SD State, Wyoming Mesonet	Soil Moisture	2023-2024
In situ	USGS National Water Information System	"dataRetrieval" R Package	Stream Discharge	2002-2024
Earth observation	NOAA NESDIS STAR Blended-Vegetation Health Product (VHP) 4 km	NOAA, "ncdf4" R Package	Vegetation Health Index (VHI)	2020-2024

2.2 Data Processing

We used R Studio 2024.12.1+536 (R version 4.4.3), QGIS 3.30.0, and ArcGIS Pro 3.4.2 to manipulate and process FAW, streamflow, and vegetation health data (Pebesma & Bivand, 2023). Similar data processing methods were employed for both the FAW validation portion of the project and the ancillary portion of the project. However, the outcomes each served a different purpose.

2.2.1 SMAP and Mesonet derived FAW

Ronald D. Leeper, conducted the pre-processing of SMAP and Mesonet soil moisture data. To derive FAW from SMAP data, he identified separate thresholds for wilting point (WP) and field capacity (FC) at each SMAP pixel over the continental United States using soil property information (i.e., the fraction of SMAP's pixel represented by sand, silt, and clay) from the U.S. Department of Agriculture's (USDA) Gridded National Soil Survey Geographic Database (gNATSGO) database (Soil Survey Staff, 2023). These values were used to drive the Rosetta version 3 model (Zhang & Schaap, 2017). This model provided van Genuchten parameters to build water retention curves for each SMAP pixel. The SMAP pixel's WP and FC were extracted along the water retention curve that represented the -1500 kPa and -10 kPa, respectively. Using these estimates of WP and FC, our science advisor calculated daily FAW conditions as: (Krueger et al., 2024; Equation 1).

$$FAW = \frac{SMAP_{pixel} - WP_{pixel}}{FC_{pixel} - WP_{pixel}} \quad (1)$$

Our science advisor used nearly the same process to calculate FAW from Mesonet soil moisture data, with two minor differences. Soil characteristic information (percentages of sand, silt, and clay) was taken from the

top 10 centimeters (0 to 10cm) of the gNATSGO nearest grid cell to derive WP and FC using Rossetta v3 and van Genuchten methods.

After this initial pre-processing, we analyzed all SMAP and Mesonet FAW data used for validation within R Studio and QGIS. First, we loaded shapefiles containing our HUC-6 basin polygons and Mesonet station point data into QGIS. An identifier attribute was added to each Mesonet point, identifying which HUC-6 basin the point was within. We selected all HUC-6 basins containing at least four Mesonet stations to ensure each basin had enough data to conduct a robust validation. This eliminated five basins, with 20 basins remaining for validation. We proceeded to select all Mesonet stations located within these 20 basins, eliminating 10 stations, for a total of 157 stations used for validation analysis. We then downloaded the select HUC-6 basins and point data as new shapefiles.

Next, we loaded the raw daily SMAP FAW data from 2020-2024 that was stored in the Amazon Web Services cloud into R Studio. We created a set of 7-day raster stacks of these data, starting with the first week of 2020, and averaged the FAW values across each of these stacks to create a weekly composite. This process was completed for the first 52 weeks of each year; the last day or two days of each year (in the case of leap years) were eliminated from the dataset to neatly aggregate the data into 7-day chunks. We determined that eliminating these days would not impact our analysis, as soils are generally frozen across our study period in December and periods of frozen soil were excluded from the validation analysis. We then created a new raster stack containing the weekly-averaged SMAP FAW data for the years of 2023 and 2024. We chose these years as our validation period because these years had the most complete Mesonet data. We loaded the filtered HUC-6 basin shapefile into R Studio, and used the extract() function in the sf package (Pebesma, 2018) to extract the mean FAW value across each HUC-6 basin for every week in our validation period. We stored these weekly basin averages in a new dataframe.

Next, we loaded the weekly Mesonet FAW data into R Studio, along with our HUC-6 and Mesonet station shapefiles and joined the Mesonet FAW data to the filtered station point data by the station identifier field. We then grouped the Mesonet FAW data by date and HUC-6 basin identifier and calculated the mean across all stations within each basin for every week in our validation period. We stored these weekly basin averages in a new dataframe. Finally, we merged the SMAP and Mesonet weekly basin averages into a single dataframe and exported this as a .csv to be used in our validation analysis.

2.2.2 Streamflow and Vegetation Health Datasets

To obtain a sufficient period of record to calculate streamflow anomalies, we obtained data from each stream gage for the years 2002 to 2024. To align with the vegetation health and SMAP datasets, we aggregated these data into weekly averages and normalized the data with a log transformation. We calculated streamflow anomalies by subtracting the 22-year average value from the observed value for each week (Armstrong et al., 2025). We used anomalies for the years 2020 to 2024 for correlation analysis.

The Vegetation Health dataset presented a time constraint that had implications for the organization of subsequently integrated datasets. STAR vegetation health are weekly data, with one reading every seven days representing the week before. The data were not organized in a Sunday-Sunday format, as the first day of January for the year 2020 fell on a Wednesday. For this dataset and all subsequent datasets, we set Week 1 as Wednesday, January 1 to Wednesday, January 7.

With all raw data stored in the cloud, we developed a script to automate the correct dating and naming of the NetCDF outputs. We created an artificial start date string, starting the year on Wednesday, January 1. The nomenclature for the files was made more specific, including the year and week number that the data represented. The next step was the automated creation of a raster from scratch with the Terra package (Hijmans et al., 2022) for each of the NetCDF files saved to the cloud. For the sake of simplicity, rather than using the project shapefiles, we set the extent to encapsulate the entire world. We saved these new rasters with a similar naming system to the NetCDF files, now as GeoTIFFs, and wrote them back into the cloud.

At this point, the data were pre-processed and ready for manipulation. To begin, we loaded the shapefiles for each of the four subbasins into R Studio using the sf package (Pebesma, 2018). An empty dataframe was created for each basin, those being James Headwaters, Cedar, Beaver, and Upper Cannonball. After ensuring both the basin shapefiles and raster data had matching coordinate reference systems, the VHI and SMAP data within each of these basins were extracted, leaving each dataframe to contain all pixel values for every week of the study period in each basin. We developed a script for the purpose of taking these extracted pixel values and determining the average pixel value for each basin each week. For example, say there are four total pixels in the Cedar basin in week x with values of 1, 2, 3, and 4. The average for this basin in week x would be 2.5. The same process was followed in the other three basins for week x, leaving week x with one average value for each basin. This process was achieved through loops in R Studio that populated the dataframe for each basin with their outputs for the first complete 52 weeks in the years 2020-2024. Finally, we used a table join to join all data frames by their "year" and "week" columns, creating one final encompassing file. The averages in the dataframe accounted for four columns (named respectively with their basin) with additional columns for week number (1-52) and year (2020-2024). Using the output_file function, the data were written to a .csv and saved locally. This process was followed for both SMAP and Vegetation Health. To prepare for the correlation analysis, we combined the main FAW calculation dataset with both the ancillary stream gage anomaly data and the vegetation health data through a table join, making one .csv file containing every weekly average for the entire study period across all basins.

2.3 Data Analysis

2.3.1 SMAP-derived FAW Validation

We first conducted a correlation analysis on the SMAP and Mesonet weekly averages for the entire validation period across the UMRB. We calculated the Pearson correlation coefficient (R-value) between these two datasets using the cor() function in R Studio and proceeded to calculate the R^2 value by squaring this value. We limited this correlation analysis to include only complete observations, i.e. dates in which both the SMAP and Mesonet datasets contained FAW values. We also calculated the mean absolute error (MAE) and bias between SMAP and Mesonet-derived FAW for the entire validation period across the entire UMRB in order to examine the magnitude and direction of error between the two datasets. We plotted these data in a linear regression.

Additionally, we conducted correlation analyses between SMAP- and Mesonet-derived FAW at the individual HUC-6 basin level to examine these statistical relationships on a smaller scale. We calculated the Pearson correlation coefficient, R^2, mean absolute error and bias values across each of our select 20 HUC-6 validation basins. We created a table showing the R^2 values for each validation basin across the validation period and joined these basin-level correlation values to our HUC-6 basin shapefile. We exported this dataframe as a new shapefile, which we loaded into ArcGIS to create a choropleth map that visualized the correlation values at the HUC-6 basin level. We also created two similar choropleth maps displaying mean absolute error and bias at the HUC-6 level, and a boxplot displaying the range of R^2 values across the HUC-6 basins.

2.3.2 Correlation Analysis with Streamflow and Vegetation Health

From the aggregated .csv file, we generated a time series line graph for each subbasin in R Studio to display the relationships between the three variables over time. To plot this time series, the scale for Vegetation Health maps was rescaled from a 0-100 scale to a 0-1 scale, in order to match the scale of the FAW validation. Using the ggplot2 package in R Studio (Wickham et al., 2019), stream flow anomaly, average vegetation health, and SMAP were plotted on a line graph. On the primary Y axis, values of –10 to 10 were chosen to best fit the data ranges of both SMAP and vegetation health. The secondary Y axis displayed values of –7 to 7 to best fit the values for stream flow. The x axis was set to represent the passage of time in weeks, while only having label breaks for the years to keep the visualization as concise as possible. One time series graph was created for each of the four subbasins. In addition to the plotted time series for each subbasin, two spatial correlation maps were generated in ArcGIS Pro to plot the correlation values of FAW to streamflow

and FAW to vegetation health. Correlation values of SMAP-derived FAW to streamflow anomaly and FAW to vegetation health were generated for each subbasin via R Studio.

3. Results

3.1 Analysis of Results

3.1.1 SMAP-derived FAW Validation Results

We calculated a R^2 value of 0.62 for SMAP-derived FAW to Mesonet-derived FAW across the entire UMRB for the years 2023-2024, indicating a strong correlation between these datasets (Figure 3). We also calculated a mean absolute error of 0.15 and a bias value of 0.04.

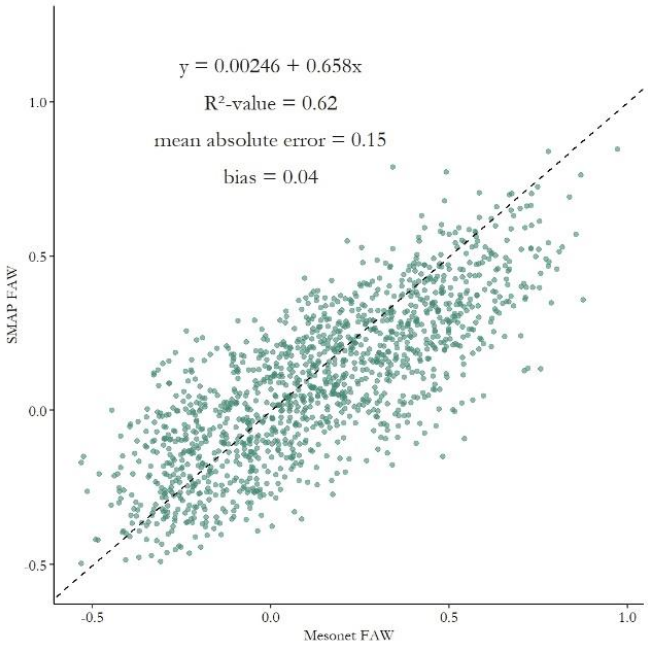

Figure 3. Linear regression model displaying the correlation between SMAP-derived and Mesonet-derived FAW

We also calculated R^2 values ranging from 0.42 to 0.84 across the 20 HUC-6 validation basins with a median value of 0.71, as well as mean absolute error values ranging from 0.10 to 0.25 and bias values ranging from -0.20 to 0.24. These values again indicate a strong correlation at the HUC-6 level. We mapped these R^2 values on a choropleth map and generated a box plot to show the distribution of R^2 values across each HUC-6 validation basin (Figure 4).

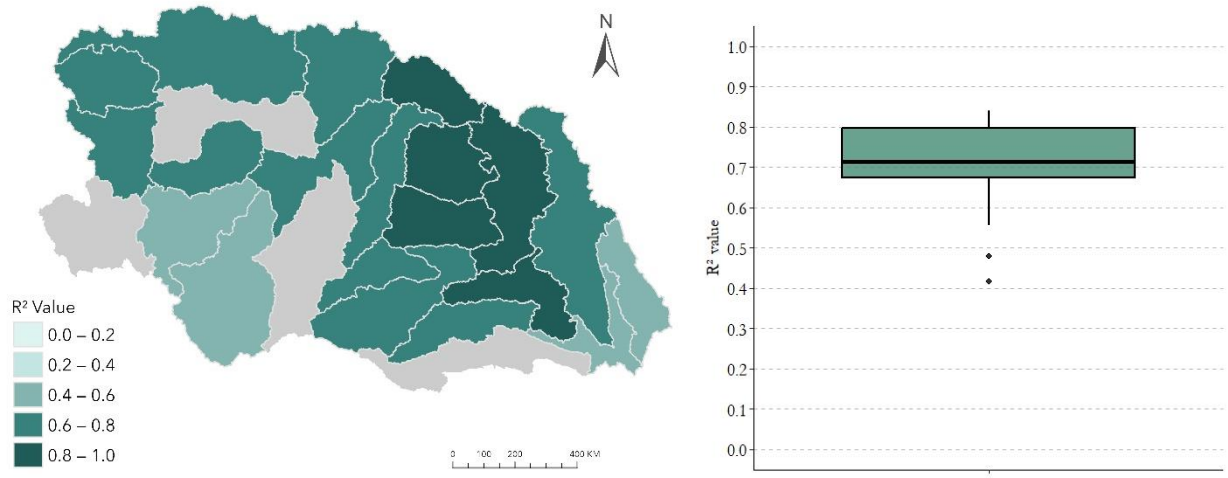

Figure 4. Left: Choropleth map displaying individual HUC-6 R^2 values. Grey basins were not included in analysis. Right: Box plot displaying distribution of individual HUC-6 R^2 values

We mapped the individual HUC-6 mean absolute error values, along with Mesonet station point data, to determine if a spatial pattern existed among mean absolute error values or if station density contributed to the magnitude of error (Figure 5). We observed no pattern in either of these regards. We also mapped the bias values across each HUC-6 basin to examine potential spatial patterns for this statistic (Figure 5). We observed that SMAP tended to display a negative bias in the northern half of the UMRB, and a positive bias in the southern half. In the bias map, areas in blue indicate basins in which bias is close to 0. Areas in green indicate positive bias., areas in purple indicate negative bias.

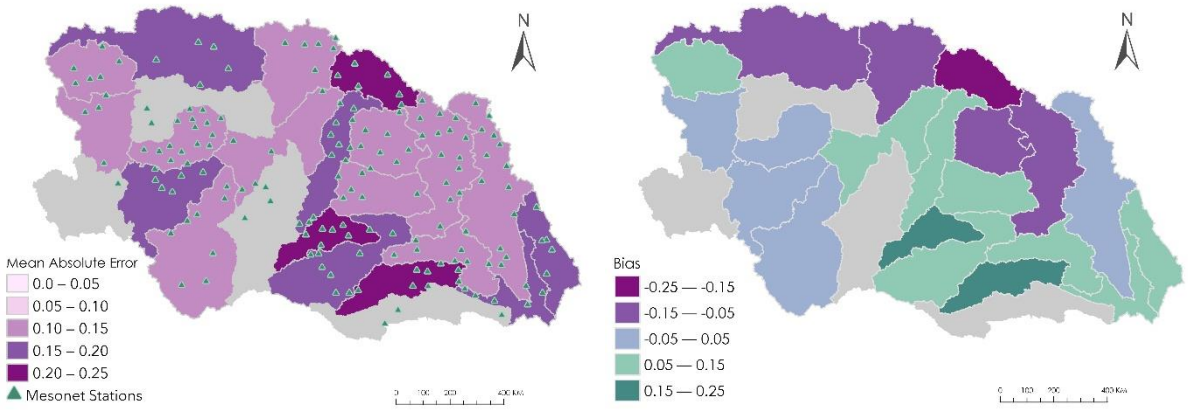

Figure 5. Left: Choropleth map displaying individual HUC-6 MAE values. Mesonet station points are overlayed. Right: Choropleth map displaying individual HUC-6 bias values

3.1.2 Streamflow and Vegetation Health

We produced four time series to understand relationships between SMAP-derived FAW and drought indicators over time for a more encompassing understanding of drought prediction potential. Figure 6 shows is a singular case study for the James Headwaters subbasin, located on an Eastern edge of the greater UMRB. This graph displays the relationship between SMAP-derived FAW, vegetation health, and streamflow over the length of our study period. The remainder of the case studies can be found in the Appendix (Figure A1; Figure A2; Figure A3) and display the same general trend.

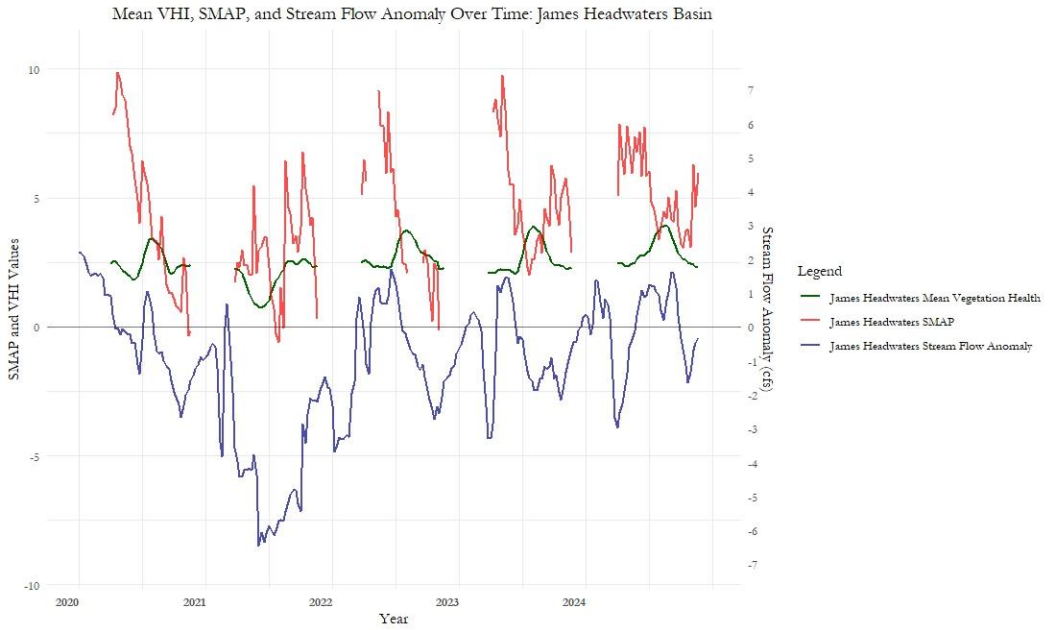

Figure 6. Time Series case study of James Headwaters Subbasin

To further analyze the relationships between SMAP-derived FAW and ancillary datasets, we performed correlation analysis for SMAP-derived FAW to streamflow and SMAP-derived FAW to vegetation health for each of our four HUC-8 subbasins. We used these values to generate two correlation maps (Figure 7). We found weak correlations between the datasets, with vegetation health to SMAP-derived FAW R^2 values ranging between 0 and 0.2 (left) and streamflow to SMAP-derived FAW R^2 values ranging between 0.2 and 0.4 (right).

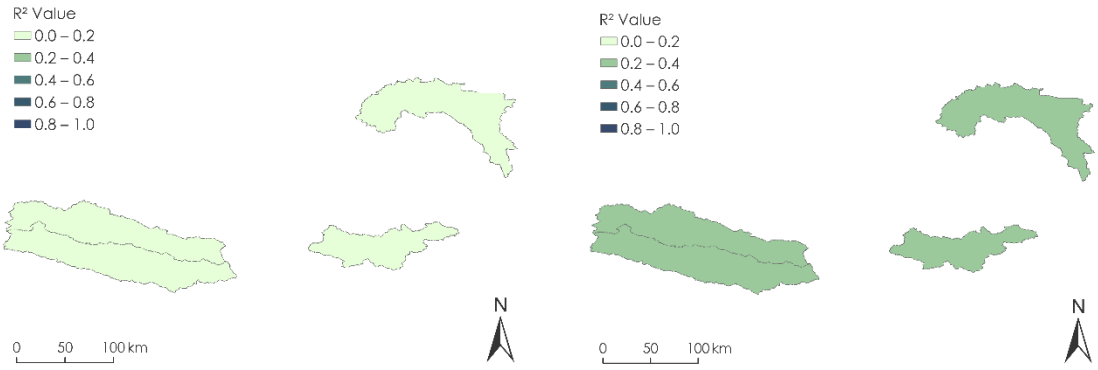

Figure 7. Correlation Maps of Ancillary Data and SMAP-derived FAW. Left: Vegetation health to SMAP-derived FAW R^2 values. Right: Streamflow to SMAP-derived FAW R^2 values

3.1.3 Monthly Analysis Results

We examined the monthly variation and seasonal patterns of FAW for the years 2020-2024. By creating a box plot time series displaying monthly average SMAP-derived FAW values for every SMAP pixel across the UMRB (Figure 8). This time series shows that across this period, FAW peaks in March, and reaches a low point in August.

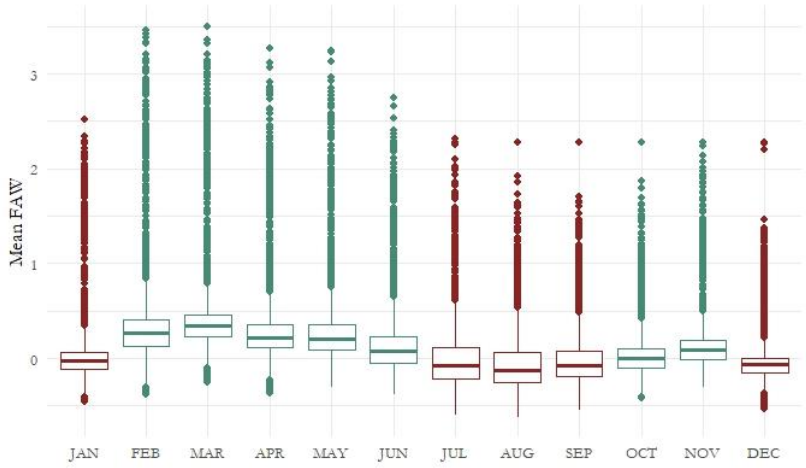

Figure 8. Mean Monthly SMAP-derived FAW 2020-2024

3.2 Errors & Uncertainties

Errors and limitations of this study presented themselves in the data processing and analysis stages. The main issue we ran into was the effect of lag time on correlation strength between soil moisture and streamflow and vegetation health, which may have offset the correlation values with the high-sensitivity satellite data. Both streamflow and vegetation health take additional time to react to changes in soil moisture, and that time would have needed to be accounted for early on in our methodology. Due to a lack of time to investigate these potential lagged relationships, we decided to continue with our original plan of temporal alignment. A second limitation of our study is that FAW exhibited high spatial and temporal variability, and the resolution of our data prevented us from capturing these high-resolution variations. The 9-kilometer resolution of SMAP is too coarse to detect many of these variations, which occur on a far finer, more local scale. Additionally, our analyses were done on a weekly timescale, but FAW can vary greatly even on daily scales, indicating that our work may not capture this high-resolution temporal variance. Accounting for these high-resolution variations would require much higher-resolution data. A final limitation was a lack of FAW data during winter months, as much of the soil in the Upper Missouri River Basin is frozen from late November to April. When soil is frozen, Mesonet stations do not record FAW. SMAP does record measurements during winter months, but we excluded these data as we were unsure of their accuracy. For these reasons, we were only able to run our analyses across the 8 months of the year during which soil is not typically frozen.

We also encountered a mismatch between the observed and expected range of both SMAP- and Mesonet-derived FAW values. In the environment, FAW values generally fall between 0 and 1. However, we observed ranges of –0.54 to 1.22. This range error indicates that the methods used to derive FAW from SMAP and the Mesonet may underestimate FAW and are not an accurate representation of the values that are actually occurring in the environment. Because the mismatch was observed for both the satellite and in situ soil moisture data sets, it is likely that the underlying cause lies with the soil property data that was used to derive these FAW values.

4. Conclusions

4.1 Interpretation of Results

4.1.1 SMAP Validation Results

The observed UMRB-wide R^2 value of 0.62 between SMAP-derived and Mesonet-derived FAW indicates a relatively strong correlation between these two datasets. This indicates that SMAP-derived FAW closely mirrors Mesonet-derived FAW values and can be useful for estimating FAW. Our calculated mean absolute error of 0.15 across the UMRB indicates that on average, SMAP-derived FAW deviates from Mesonet-derived FAW by around 0.15. Given the range of the data, this is a relatively low error value. Our calculated

bias value of 0.04 across the UMRB indicates that SMAP has a slight tendency to overestimate FAW as compared to the Mesonet. This does not necessarily indicate "correctness," but rather a deviation from the typical baseline. The correlation of our individual HUC-6 level R^2 values also indicate a strong correlation with a median R^2 value of 0.71, supporting the validity of SMAP-derived FAW. Likewise, the range of our HUC-6 mean absolute error and bias values indicate that there was relatively low error between SMAP and Mesonet-derived FAW. We found no notable spatial pattern in the distribution of mean absolute error at the HUC-6 scale, but we did notice that basins with negative bias were clustered in the northern half of the UMRB. We hypothesize this may be due to the presence of large quantities of glacial till in the northern half of the basin, but further research would be necessary to further assess causation.

4.1.2 Streamflow and Vegetation Health

There were weak temporal correlations between SMAP and ancillary datasets, which may be explained by complicating factors such as seasonality in the context of vegetation health, and snowmelt in the context of streamflow. One can see that the general pattern of ridge to trough is consistent between the SMAP-derived FAW readings and streamflow readings (Figure 6; Figures A1 - A3), though streamflow slightly lags; this may be attributed to the time taken for streamflow to respond to snowmelt and precipitation events, as well as time for water to travel to the stream gages at the basin exits. The James Headwaters Basin displays a clear example of these lag relationships. We noticed a drought signal in 2021, where soil moisture is extremely low in the early months of the year. High early season SMAP-derived FAW values tend to coincide with a high mid-season peak in vegetation health, visible in 2020, 2022, 2023, and 2024. Identifying drought signals through this type of analysis may help improve drought monitoring in the Upper Missouri River Basin and beyond.

The correlation maps did not show significant spatial variations in R^2 values between the HUC-8 subbasins, potentially because they are similar in size, are near one another, and share similar distances between basin entrances and exit stream gages, which may result in similar lag times following precipitation events. We observed a longer lag time between SMAP-derived FAW and vegetation health than SMAP-derived FAW and streamflow in our time series analysis, which may explain the slightly higher correlation values between SMAP-derived FAW and streamflow in our correlation maps. These datasets would require temporal readjustment to better account for the lag effect on the correlation values. In addition, processing vegetation health data to change map anomaly values may provide a stronger correlation with SMAP-derived FAW.

4.2 Feasibility & Partner Implementation

The National Weather Service Weather Forecast Offices in the Upper Missouri River Basin utilize in-situ Mesonet readings as their primary method for monitoring soil moisture. This method, however, is spatially limited to the Mesonet station locations. The integration of SMAP satellite soil moisture data could provide the spatial capacity to fill gaps in areas without extensive Mesonet coverage. The relatively strong correlation of SMAP-derived FAW to Mesonet-derived FAW supports the implementation of SMAP as a complementary tool to in situ monitoring stations. Additionally, FAW analysis shows promise in identifying drought signals in the Upper Missouri River Basin. Further analysis of early season soil moisture conditions may help scientists better understand vegetation response patterns and could be a beneficial indicator for farmers about the potential productivity of a growing season. Based on the outcomes of this project, we conclude NASA Earth observations can complement the National Weather Service's existing methods for flood and drought forecasting in the Upper Missouri River Basin.

5. Acknowledgements

The Upper Missouri River Basin Water Resources team would like to thank our Lead, Tallis Monteiro, science advisors Ronald D. Leeper and Erik Krueger, and the DEVELOP National Program Office for their guidance and support. Additional thanks go to our Lead Science Advisor, Joel Lisonbee, Project Coordination Fellow Brent Bowler, and project partners, Melinda Beerends and Tommy Grafenauer at the NWS office in Grand Forks, Allen Schlag at the NWS office in Bismarck, Aaron Woodward at the NWS office in Rapid City, and Daryl Ritchison at the North Dakota State University, North Dakota State Climate office and North Dakota Agricultural Weather Network Center.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract 80LARC23FA024.

6. Glossary

Agricultural drought – crop failure resulting from insufficient soil moisture
ArcGIS Pro – a professional desktop GIS application from Esri
CSV – Comma-Separated Values, a text file format using commas to separate values and lines to separate records
Earth observations – satellites and sensors that collect information about the Earth's physical, chemical, and biological systems over space and time
FAW – Fractional Available Water, the measure of water in the soil that is available for plants to intake and use
FC – Field Capacity, the amount of water soil can hold after excess water has drained away
GeoTIFF – A file format for georeferenced raster imagery
HUC – Hydrologic Unit Code, a drainage area classification system used by the US Geological Survey
Mesonet – a network of environmental monitoring stations
NetCDF – a file format and set of libraries used to store and share scientific data
SMAP – Soil Moisture Active Passive, a NASA satellite that measures soil moisture across the globe
RStudio – an integrated development environment for the programming language, R
VHP – Vegetation Health Product
WP – Wilting Point, the minimum amount of water that a plant can tolerate without wilting

7. References

Access National Hydrography Products. (n.d.). USGS; United States Geological Survey. https://www.usgs.gov/national-hydrography/access-national-hydrography-products

Armstrong, D. W., Lange, D. A., & Chase, K. J. (2025). *Methods to determine streamflow statistics based on data through water year 2021 for selected streamgages in or near Wyoming* (No. 2024-5104). United States Geological Survey. https://doi.org/10.3133/sir20245104

Badger, A. M., Livneh, B., Hoerling, M. P., & Eischeid, J. K. (2018). Understanding the 2011 Upper Missouri River Basin floods in the context of a changing climate. *Journal of Hydrology: Regional Studies*, *19*, 110–123. https://doi.org/10.1016/j.ejrh.2018.08.004

Bell, J. E., Sherry, R., & Luo, Y. (2010). Changes in soil water dynamics due to variation in precipitation and temperature: An ecohydrological analysis in a tallgrass prairie. *Water Resources Research*, *46*(3). https://doi.org/10.1029/2009wr007908

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2009). Soil moisture temporal stability over experimental areas in Central Italy. *Geoderma*, *148*(3-4), 364–374. https://doi.org/10.1016/j.geoderma.2008.11.004

Data Products. (n.d.). SMAP Soil Moisture Active Passive; Jet Propulsion Laboratory at California Institute of Technology. Retrieved February 12, 2025. https://smap.jpl.nasa.gov/data/

Hijmans, R. J., Bivand, R., Etten, J. van, Forner, K., Ooms, J., & Pebesma, E. (2022, February 17). terra: Spatial Data Analysis. R-Packages. https://cran.r-project.org/package=terra

Krueger, E. S., & Ochsner, T. E. (2024). Traditional matric potential thresholds underestimate soil moisture at field capacity across Oklahoma. *Soil Science Society of America Journal*, *88*(5), 1678–1690. https://doi.org/10.1002/saj2.20733

Krueger, E. S., Ochsner, T. E., & Quiring, S. M. (2019). Development and evaluation of soil moisture-based indices for agricultural drought monitoring. *Agronomy Journal*, *111*(3), 1392. https://doi.org/10.2134/agronj2018.09.0558

Leeper, R. D., Palecki, M. A., Watts, M., & Diamond, H. (2023). On the detection of remotely sensed soil moisture extremes. *Journal of Applied Meteorology and Climatology*, *62*(11), 1611–1626. https://doi.org/10.1175/jamc-d-23-0059.1

Manns, H. R., Berg, A. A., Bullock, P. R., & McNairn, H. (2014). Impact of soil surface characteristics on soil water content variability in agricultural fields. *Hydrological Processes*, *28*(14), 4340–4351. https://doi.org/10.1002/hyp.10216

National Oceanic Atmospheric Administration. (2019) Center for Satellite Applications and Research - NOAA / NESDIS / STAR. Accessed 4 Mar. 2025. www.star.nesdis.noaa.gov/smcd/emb/vci/VH/index.php

National Snow and Ice Data Center: CIRES at the University of Colorado Boulder. (2020*). SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture* (Version 6).[Dataset]. https://doi.org/10.5067/M20OXIZHY3RJ

National Snow and Ice Data Center (2015). *SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture* (V006) [Data set]. NASA National Snow and Ice Data Center Distributed

Active Archive Center. Retrieved February 20, 2025, from https://doi.org/10.5067/M20OXIZHY3RJ

Pebesma, E. (2018). Simple features for R: standardized support for spatial vector data. *The R Journal*, *10*(1), 439. https://doi.org/10.32614/rj-2018-009

Pebesma, E & Bivand, R. (2023). *Spatial data science: With applications in R*. Taylor and Francis Group, https://doi.org/10.1201/9780429459016

Pierce, D. (2023, November 28). ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. R-Packages. https://cran.r-project.org/package=ncdf4

Soil Survey Staff. Gridded National Soil Survey Geographic (gNATSGO) Database for Missouri. United States Department of Agriculture, Natural Resources Conservation Service. Available online at https://nrcs.app.box.com/v/soils

USACE Upper Missouri River Basin Soil Moisture and Plains Snow Monitoring Build-Out. (2020). Drought.gov; National Integrated Drought Information System. https://www.drought.gov/drought-research/usace-upper-missouri-river-basin-soil-moisture-and-plains-snow-monitoring-build

United States Department of Commerce, National Oceanic Atmospheric Administration, National Weather Service. (2020). *About the NWS*. Weather.gov. https://www.weather.gov/about/

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Takahashi, K. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Woodhouse, C. A., & Wise, E. K. (2020). The changing relationship between the upper and lower Missouri River basins during drought. *International Journal of Climatology*, *40*(11), 5011–5028. https://doi.org/10.1002/joc.6502

Xu, Y., Wang, L., Ross, K., Liu, C., & Berry, K. (2018). Standardized soil moisture index for drought monitoring based on soil moisture active passive observations and 36 years of North American land data assimilation system data: A case study in the Southeast United States. *Remote Sensing*, *10*(3), 301. https://doi.org/10.3390/rs10020301

Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, *19*(NA), 333-349. https://doi.org/10.1139/a11-013

Zhang, Y., & Schaap, M. G. (2017). Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). *Journal of Hydrology*, *547*, 39-53. https://doi.org/10.1016/j.jhydrol.2017.01.004

8. Appendices

Appendix A: *SMAP FAW results*

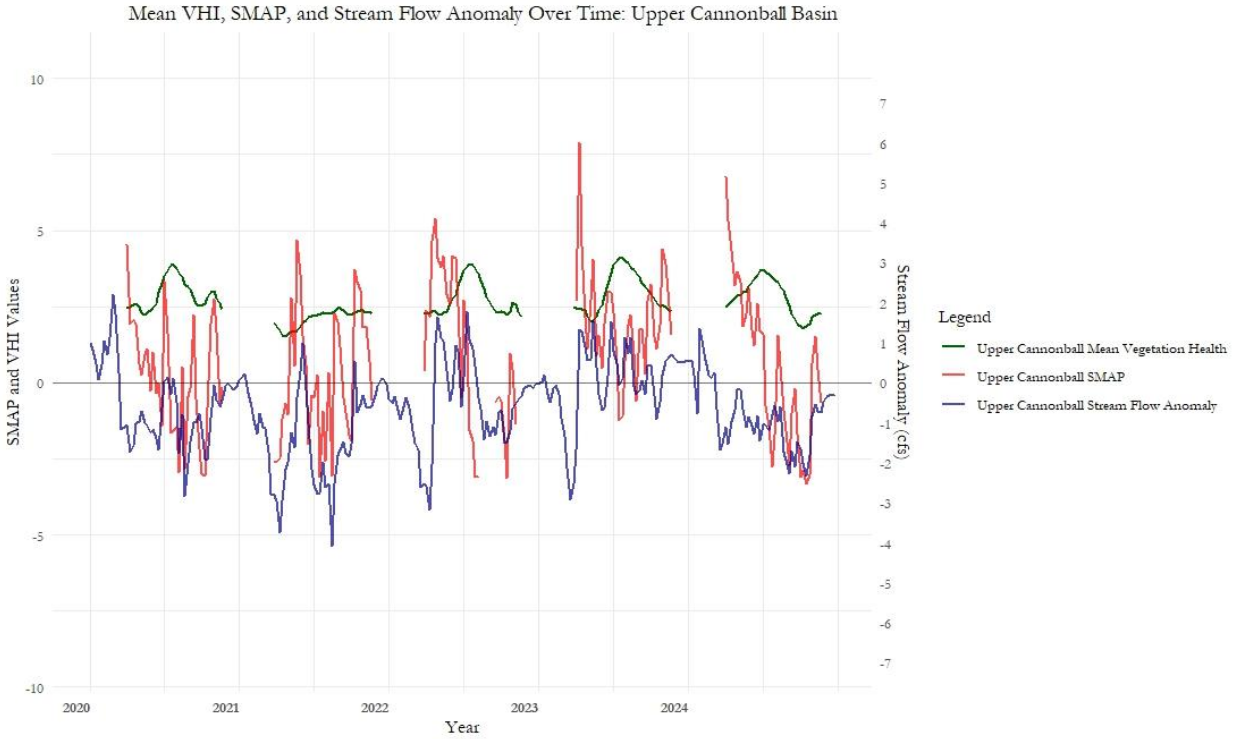

Figure A1. Mean VHI, SMAP, and Stream Flow Anomaly Over Time: Upper Cannonball Basin shows relationships over time and seasonal patterns of SMAP FAW, vegetation health, and streamflow

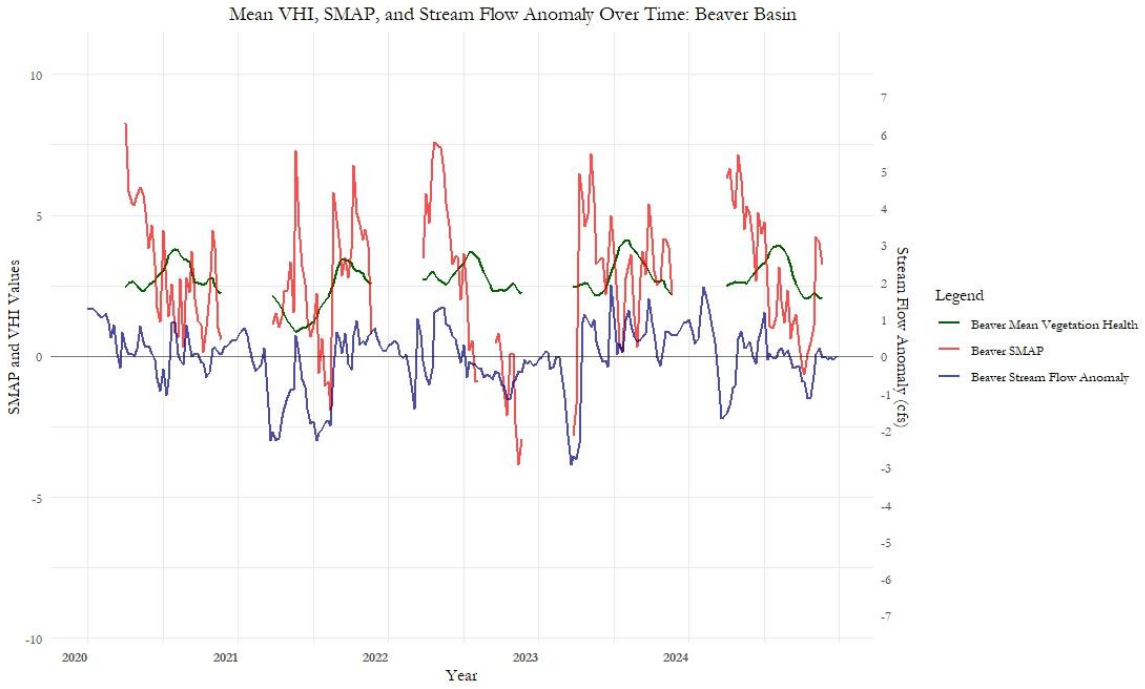

Figure A2. Mean VHI, SMAP, and Stream Flow Anomaly Over Time: Beaver Basin shows relationships over time and seasonal patterns of SMAP FAW, vegetation health, and streamflow

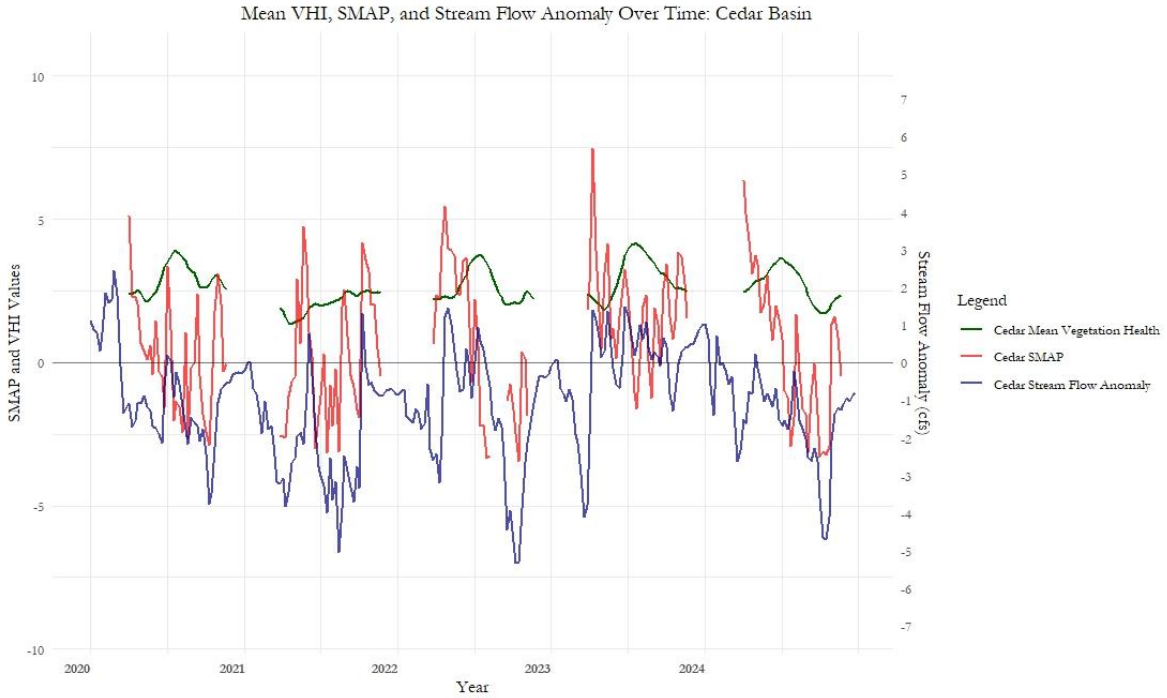

Figure A3. Mean VHI, SMAP, and Stream Flow Anomaly Over Time: Cedar Basin shows relationships over time and seasonal patterns of SMAP FAW, vegetation health, and streamflow

Table A1

Correlation between SMAP FAW and ancillary datasets, vegetation health and streamflow, for each HUC-8 subbasin

Subbasin	**SMAP FAW to Vegetation Health (R^2)**	**SMAP FAW to Streamflow (R^2)**
Beaver	2.32E-04	0.26226446
Cedar	2.62E-05	0.29608814
James Headwaters	2.44E-04	0.26008442
Upper Cannonball	2.20E-04	0.24611315