Montana Water Resources II

Enhancing a Composite Moisture Index for Drought and Flood Monitoring in the Missouri River Basin

 **Technical Report**

Final Draft – April 1st, 2021

Chloe Schneider (Project Lead)

Bevan Pearson

Madelyn Savan

Catherine Buczek

***Advisors:***

Ronald Leeper, NOAA National Centers for Environmental Information, North Carolina Institute for Climate Studies (Science Advisor)

***Previous Contributors:***

Dean Berkowitz

Egla Ochoa-Madrid

Julie Sorfleet

# 1. Abstract

The Missouri River Basin is a major global breadbasket, containing large amounts of agricultural land. Recent devastating weather events have motivated regional organizations to dedicate efforts to drought and flood monitoring and early warning systems. The DEVELOP team partnered with the following organizations: Montana Climate Office, NOAA National Weather Service (NWS) Missouri Basin River Forecast Center, NOAA Regional Climate Services of the Central Region, NOAA Physical Sciences Laboratory, and the US Army Corps of Engineers' Missouri River Basin Water Management Division. The team collaborated with the partners in their efforts to monitor flood and drought by enhancing a composite moisture index (CMI) for the Missouri River Basin. The CMI leverages NASA Earth observations to derive snow cover data from the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) mission and snow water equivalent and snow depth datasets from the NOAA NWS National Operational Hydrologic Remote Sensing Center’s Snow Data Assimilation System (SNODAS). Building upon this framework, the team added groundwater storage data from the Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On missions, soil moisture data from Soil Moisture Active Passive (SMAP), and United States Geological Survey in situ streamflow data to the CMI. To test the tool’s validity, the team compared the CMI results to historically extreme dry and extreme wet years, March 2017 and March 2019, respectively. The CMI accurately reflects both 2017 and 2019 climate conditions in the Missouri River Basin. The refined CMI enhances the understanding of antecedent soil moisture conditions and improves flood and drought forecasting in the Missouri River Basin before the growing season.

**Key Terms**

# *Key Terms:* composite moisture index, flood, drought, MODIS, NDSI, SNODAS, SWE, GRACE

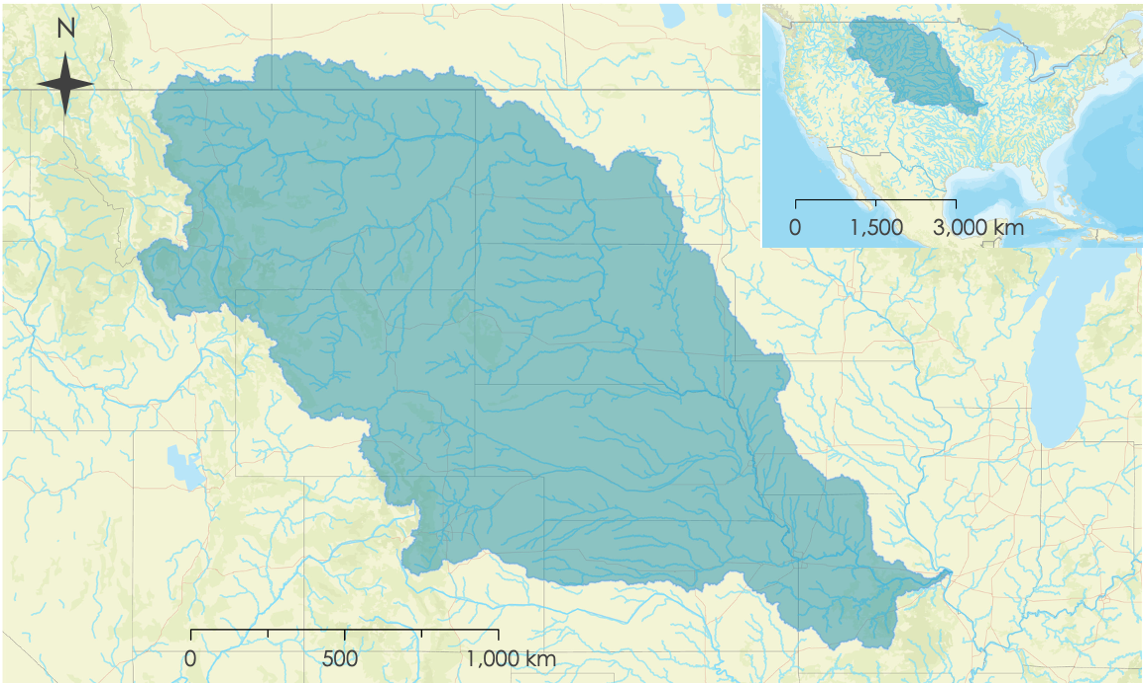
# 2. Introduction

***2.1 Background Information***

The Missouri River Basin (MRB) is the largest river basin in the United States, encompassing 500,000 square miles of drainage from all of Nebraska and parts of Colorado, Iowa, Kansas, Minnesota, Missouri, Montana, North Dakota, South Dakota, and Wyoming. The Missouri River Basin supports much of the United States agricultural industry and produces approximately 46% of the wheat, 22% of the grain corn and supports 34% of the cattle found in the United States (Mehta et al., 2013). In 2017, an unexpected flash drought devastated the region, causing agricultural losses exceeding $2.6 billion and record wildfires (Hoell et al., 2019). Regional organizations are seeking to improve drought early warning systems and communication with stakeholders in order to better prepare for extreme weather events (Hoell et al., 2020).

State agencies and regional organizations use a variety of drought indicators including the Palmer Drought Severity Index and the Standardized Precipitation Index to evaluate drought conditions (Hoell et al., 2020). The Surface Water Supply Index, developed by Dezman and Shafer (1982) for drought monitoring in Colorado, focuses on water supply from melting snow, as snowmelt accounts for a large portion of annual streamflow in the MRB. While state agencies employ various drought indicators in their monitoring efforts, there is currently no single drought index that holistically considers various moisture conditions that can be applied across the entire forecast region (Hoell et al., 2020). In fall 2020, the Montana Water Resources I team worked with regional partners to develop an index that consolidates multiple hydrological and climate indicators into a single metric to better address flood and drought risk across the Missouri River Basin. The Montana Water Resources I team developed a Composite Moisture Index (CMI) that incorporates soil moisture, snow cover, snow depth and snow water equivalent measurements derived from NASA Earth observations and ancillary datasets. The spring 2021 Montana Water Resources II team expanded the functionality of the CMI by incorporating additional data sets to increase both the temporal and spatial resolution of the CMI.

This term’s project expanded the functionality and predictive abilities of the monthly CMI developed by the previous team for the Missouri River Basin (Figure 1) from January 2004 to December 2020. Our team incorporated *in situ* streamflow data and groundwater storage estimates in the CMI computations in order to account for water storage conditions in the basin and further improve the CMI estimates of antecedent moisture conditions. We further strengthened the CMI as a drought monitoring tool by leveraging NASA’s Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (FO) for groundwater storage estimates which have been strongly correlated to *in situ* groundwater storage data in various regions and climates around the world (Li et al., 2019). The term GRACE will describe both the GRACE and GRACE-FO satellites hereafter. Our team integrated NASA Earth observations, *in situ* observations and modeled data products to develop a holistic drought index that accounts for regional basin characteristics and can be utilized in drought monitoring efforts throughout the Missouri River Basin.



*Figure 1*. Extent of the Missouri River Basin (Hydrologic Unit Code [HUC] 2, Region 10) in teal, created in ESRI’s ArcGIS Pro.

***2.2 Project Partners & Objectives***

The team partnered with the Montana Climate Office, the National Oceanic and Atmospheric Association (NOAA) National Weather Service (NWS) Missouri Basin River Forecast Center, NOAA Regional Climate Services of the Central Region, US Army Corps of Engineers, Northwestern Division, and NOAA, Physical Sciences Laboratory. The Montana Climate Office provides timely climate information on their online portal, the Upper Missouri River Basin drought indicators dashboard. In the future, the Montana Climate Office will incorporate the CMI on their drought indicators dashboard and further expand the state’s Drought Early Warning System. NOAA National Weather Service, Missouri Basin River Forecast Center will use the index to inform decision-makers of flood and drought indicators for the months following snowmelt season. This project developed a monthly CMI that holistically and quantitatively assesses the moisture state of the Missouri River Basin and can be used by our partners to communicate hydrological and climate conditions to stakeholders and the public. Our partners will use the CMI to gain a better understanding of soil moisture conditions and will help assist local decision makers with flood and drought forecasting.

# 3. Methodology

***3.1 Data Acquisition***

This project combined NASA Earth observations, modelled data products, and *in situ* measurements to generate the CMI. The DEVELOP team collected data from NASA online databases and Google Earth Engine (GEE). The data download process was scripted in order to enable automation in the future. Table 1 outlines all data sources used in this project.

Table 1

*Data sources for the six parameters needed to calculate the CMI*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Type** | **Sensor / Source** | **Data Product** | **Acquisition Method** | **Parameter** | **Years Used** |
| Earth observation | Terra MODIS NDSI | Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Snow Index (NDSI) Snow Cover Daily L3 Global 500m SIN Grid, Version 6 | Google Earth Engine Python API | Snow cover | 2000 – present |
| Earth observation | GRACE; GRACE FO | Groundwater and Soil Moisture Conditions from GRACE Data Assimilation L4 7-days 0.125 x 0.125 degree, Version 4.0 | University of Nebraska-Lincoln https server | Groundwater storage | 2002 – present |
| Model | SNODAS | Snow Data Assimilation System (SNODAS) Data Products at NSIDC, Version 1 | NOAA FTP Server | Snow water equivalent, snow depth | 2003 – present |
| Earth observation | SMAP | SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 4 | Google Earth Engine Python API | Soil moisture | 2015 – present |
| *In situ* | USGS Stream Gages | USGS Water Data for the Nation | ‘waterData’ package in RStudio | Stream discharge | 2000 – present |

The DEVELOP team first acquired Level 3 1-km gridded Terra MODIS NDSI snow cover data by modifying the input parameters (e.g. timeframe) of an existing GEE Python API script in Jupyter Notebook. Next, the team clipped the data to the study area using a shapefile of the Missouri River Basin (HUC 2, Region 10) from the National Resources Conservation Service Watershed Boundary Dataset. After clipping, the team applied the GEE reducer mean function to get a monthly average snow cover value. In addition to snow cover, the team used two modeled estimates of snowpack—snow water equivalent (SWE) and snow depth—from the NOAA Snow Data Assimilation System (SNODAS). These data were downloaded from a NOAA FTP server at daily 1-km resolution for the continental United States, via an R script that utilized the ‘httr’ package for downloading zipped data.

Next, the team downloaded groundwater and soil moisture indicators. Using a Python script in Jupyter Notebook, the team downloaded GRACE shallow groundwater observations from an online https server curated by University of Nebraska-Lincoln. The Python script used the selenium package to open a Mozilla Firefox browser and download a file for each weekly GRACE groundwater .tiff file on the server.

Similar to Terra MODIS NDSI, the team used a GEE Python API script to download data from the NASA-USDA Soil Moisture Active Passive (SMAP). The team created a URL for the desired raster image, which was downloaded as a zipped archive and unzipped in an R script using the ‘httr’ package. The team accessed preprocessed surface soil moisture anomaly (SSMA) data, clipped to the study area, and applied the GEE reducer mean function to derive monthly rasters for the available period of record.

The team downloaded mean daily stream discharge readings from select USGS stream gages from the USGS Water Data for the Nation website using the ‘waterData’ package in R Studio. The study area was delineated into HUC-6 level basins within the HUC-2 Missouri River Region (see Figure A1). Partners at the Missouri Basin River Forecast Center helped the team to select the farthest downstream USGS stream gage in each HUC-6 basin in order to approximate streamflow for the entire basin (Figure A2). After importing the list of gage locations into R, the team averaged ft3/sec stream discharge (variable code 00060) over fall (September-December) and spring (April-August) for each year of the study period, excluding all non-applicable readings.

***3.2 Data Processing***

Each snow, groundwater, streamflow, and soil moisture dataset was converted from daily or weekly temporal resolution to monthly means. During data acquisition, the team clipped all datasets to a Missouri River Basin (Hydrologic Unit Code [HUC] 2, Region 10) shapefile sourced from the National Resources Conservation Service Watershed Boundary Dataset. All datasets except SMAP then needed to be normalized into a common unit relative to their respective climatic history for CMI computations. The normalization enabled each dataset to be integrated into a composite index, despite the varying native units and periods of record. The complete CMI workflow is pictured in Figure A3 and further detailed in the remainder of section 3.2 below.

***3.2.1 Normalizing Data into Anomalies***

Drought or flood events stem from exceptional moisture states that deviate from the long-term average, or ‘normal’ conditions of an area. To detect how observations deviate from historic conditions, one must establish relative climate normals: the average climatological conditions for a weather variable for a given time frame. The team produced monthly deviations from climate normals by calculating standardized anomalies for each year of a given month utilizing the empirical cumulative distribution function (ECDF) and qnorm functions in R 4.0.3 through RStudio 1.3.1093 (R Core Team 2020). This was done for each dataset except for SMAP, as it was already available in standardized anomalies, and GRACE, which was available in percentiles and solely went through the qnorm function. The team used the ‘stats’ package (v3.6.2; R Core Team 2019) to compute an ECDF for a given month across each historical dataset. The ECDF ranks all monthly pixel observations from 0 to 1, or driest to wettest for the SNODAS, MODIS, and stream gage data. With support from our scientific advisor, the team developed a custom function as an input to the calc function from the ‘raster’ package (v3.3-13, Robert Hijmans 2020) so as to generate an ECDF cell-wise across a raster surface for a given month while subsequently assigning the corresponding percentile value to each pixel. Next, these percentile values were used as inputs to the ‘qnorm’ function in R utilized within the calc function, which returns anomaly values for each pixel. The outputs of all of our datasets were in standard anomalies comparing monthly mean observations to that month’s climatic history, standardly used in drought communication. The calculated anomaly value for each parameter at a pixel is between -2 and +2. The value “0” denotes that the parameter is equal to the climate normal (mean), a value of +q means that the parameter is q standard deviation above the mean, and a value of -q means that it is q standard deviation below the mean.

***3.2.2 Spatial Aggregation & CMI Calculation***

The team took two approaches to spatially aggregate the disparate native resolutions of the four datasets for compositing. Firstly, each anomaly dataset was resampled to the coarsest resolution, that being SMAP at 27km, using the nearest neighbor assignment method via the ‘raster’ package in R. The second approach involved spatially aggregating input datasets at their native resolutions to the HUC-6 level in order to compute CMI values for each HUC-6 level basin within the study area. To do so, the team extracted the mean pixel value for each HUC-6 polygon using the extract function in the ‘raster’ package in R. Following the spatial aggregation, a CMI was calculated by averaging six anomaly values for each of the climate variables. Therefore, the team produced two CMI maps, one at a 27km resolution and one at a HUC-6 level resolution, both computed by averaging the indicators. After exporting the CMI results as .csv tables, the team visualized the CMIs for select years in ArcGIS Pro.

***3.3 Data Analysis***

***3.3.1 Principal Component Analysis***

The team used two approaches to analyze the CMI results: (1) comparison and weighting of each input parameter and (2) correlation with spring stream discharge measurements. The team conducted a Principal Component Analysis (PCA) to assess the suitability of the datasets for CMI classification and explore which variables in the CMI drive the most variance in the dataset. PCA is a method to reduce the dimensionality of datasets while preserving statistical information, therefore allowing us to better interpret how each of the six moisture indicators influenced our index. We used PCA to project our six-dimensional dataset down to the directions of the first two principal components. This projection provides us with 2-dimensional data that we can plot and visualize. For our analysis, we combined all monthly mean data from March 2015 through December 2020 at the HUC 6 resolution into a single data frame. We then imported the data frame into R. We ran the PCA using the prcomp function (native to R) and plotted the data using the “factoextra” package.

***3.3.2 Correlation Analysis***

To assess the relative accuracy of CMI results, the team looked for correlation between hydrologic basin level (HUC-6) CMI outputs with *in situ* cumulative streamflow data. As outlined in Section *3.1*, the team generated one cumulative discharge value per season in every basin of the Missouri River Watershed, listed by year. For the correlation analysis, spring to summer (April-July) accumulated discharge values were used to estimate actual moisture conditions across each HUC-6 basin for the period during which droughts or floods impacts the growing season. The team compared average spring and summer discharge values in each HUC-6 basin with available stream data to each winter month’s (December-April) CMI output using the lm (linear model) function in the R ‘stats’ package. The team used the linear regression to identify a coefficient of determination (R2) for each winter month’s CMI and spring and summer cumulative discharge from January 2004 - March 2020, where sufficient data were available. The data were graphed in order to visualize the relationships between the two datasets.

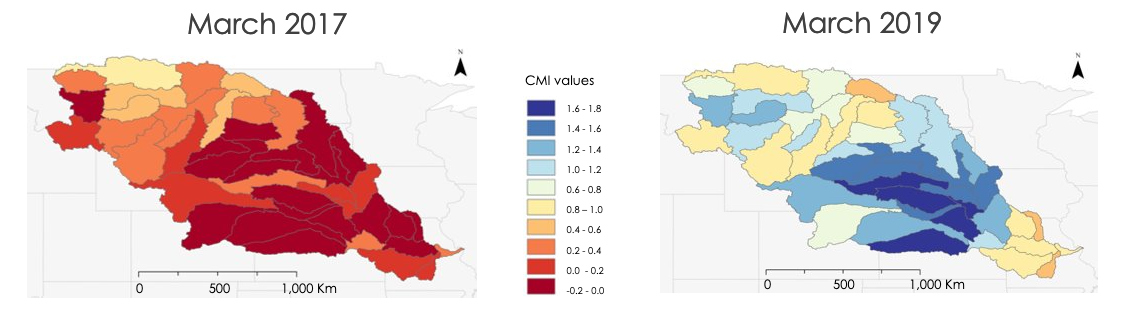
# 4. Results & Discussion

***4.1 Analysis of Improved CMI***

***4.1.1 CMI results***

The principal results for this feasibility study were composite moisture indices for each winter month (December-April) from 2000-2021, at both pixel-based and HUC-6-based spatial resolution. The CMI values range from -2 to +2, where -2 corresponds to drier than average conditions, 0 corresponds to average conditions based on the two-decade study period, and +2 corresponds to wetter than average conditions. We chose to map these results from two sample periods, March 2017 and March 2019, in Figures 2 and 3 below. We mapped results from March because the Montana Water Resources I team found that March CMI outputs corresponded best with subsequent spring and summer streamflow. We chose 2017 and 2019 as example years because all datasets were available for these years, and 2017 was a known drought year, whereas 2019 was a major flood year in the region. Both the HUC-based and pixel-based resolutions of the CMI outputs performed well and were consistent with reported climate conditions (Figure 2 and 3).

*Figure 2.* CMI outputs at the HUC-6 spatial resolution for March 2017 and March 2019

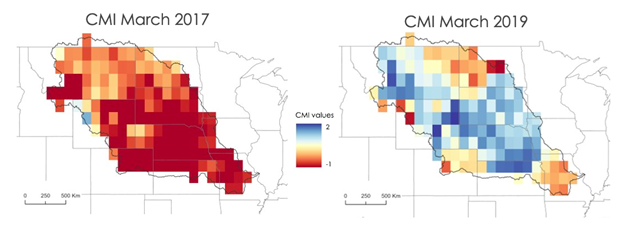


**HUC-based CMI Output for March 2017 and 2019**

March 2017

March 2019

**Pixel-based CMI Output for March 2017 and 2019**



March 2019

March 2017

*Figure 3*. CMI outputs at the pixel-based resolution for March 2017 and March 2019

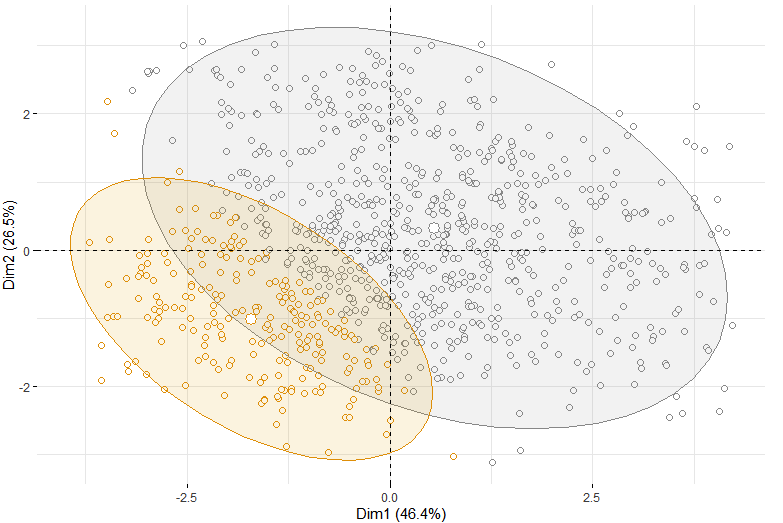
Figure 3 shows CMI results for the pixel based spatial aggregation approach where all values were resampled to the resolution of the coarsest variable (SMAP) at ~25km. Note that the pixel-based moisture index excludes the antecedent fall streamflow parameter. However, the hydrologic unit-based CMI maps shown in Figure 2 include antecedent streamflow for all but 8 HUC-6 basins. In both figures, the left-hand map of 2017 values show redder colors, which correspond to lower CMI values and indicate an aggerate of drier than normal snowpack and soil moisture conditions. This matched expectations, as 2017 was a drought year. In contrast, the right-hand maps show bluer colors, corresponding to higher CMI values that indicate an aggregate of wetter than normal conditions. This confirmed expectations as 2019 was a major flood year. Interestingly, when comparing March 2017 with March 2019, the CMI results appear to be more sensitive to interannual moisture fluctuations in the lower basin than in the upper basin. This is illustrated by more changes in color in the southern pixels and HUCs between the 2017 and 2019 maps.

***4.1.2. Principal Component Analysis***

Principal component analysis results provided our team with a two-dimensional view of the six-dimensional dataset. Figure 4 shows all data points plotted on the first two principal components which combined, account for 72.9% of all variance in the data. The data points are colored by the CMI value associated with a given data point. Positive CMI values are displayed in grey circles and negative CMI values are displayed in yellow circles. Confidence ellipses are displayed in the shaded areas around the two categories.

Plotting on the first two components broadly distinguishes between negative and positive CMI conditions. The clustering of the data is an indication that the selected parameters are well suited for the CMI classification. Figure 4 illustrates that the variance is related to CMI values as lower (drier) CMI values are aggregated in the bottom left and higher (wetter) CMI values are aggregated in the right of the plot.

**2- Dimensional Principal Component Analysis of Positive and Negative CMI Values**



Dimension 1 (46.4%)

Dimension 2 (26.5%)

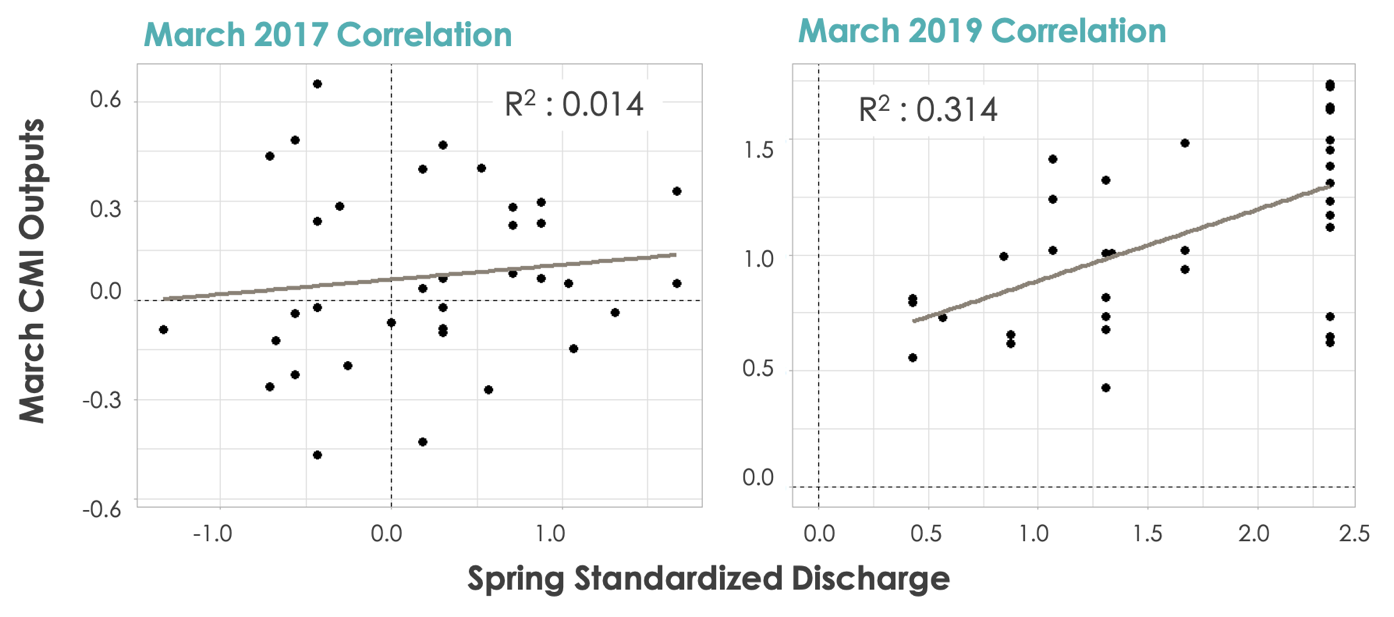
*Figure 4.* 2-dimensional principal component plot showing the clustering of positive CMI values (colored grey) and negative CMI values (colored yellow) across six moisture indicators.

After exploring the suitability of our moisture indicators for CMI classification, we investigated the variance of the individual indicators. The first principal component points in the direction of the most variance of our data and is therefore of interest to identify which indicators contribute most to the first principal component. A bar graph of the contributions of each indicator allows us to visualize the contributions of each moisture indicator to the first principal component (Appendix A5). In this analysis, a red dashed line at approximately 17% contribution is a reference line that corresponds to the expected value if the contributions of each indicator were equal. Indicators with a contribution above the reference line are considered an important contributor to the principal component. The first principal component accounts for 46.4% of all the variance in our data. Indicators that are considered important contributors of the first principal indicate which variables are driving the most variance in the data. Based on the percent contributions, snow water equivalence, snow depth and snow cover monthly mean measurements are the main contributors to the variance (appendix A5).

***4.1.3 Validating Winter CMI with Spring Stream Gauge Output***

The relationship between the CMI and spring discharge is useful in understanding the overall strength of the CMI and its use as an indicator of spring snowmelt. In analyzing the correlations between winter CMIs and spring cumulative discharge for 35 different HUC regions over a 20-year time period, we identified many different spatial and temporal patterns. We analyzed relationships between CMI and spring discharge using R2 values, which indicate the percentage of variance the independent variable (spring discharge) explains collectively in the dependent variable (CMI).

Looking at correlations between CMI results and standardized spring discharge for each HUC in March 2017 and 2019 in Figure 5, the data did not show a strong relationship between CMI values and spring discharge for March 2017 (left), with a R2 of 0.014. For March 2019, the data showed a higher correlation of 0.314. Additionally, all HUCs experienced greater-than-normal spring to summer streamflow in 2019. This is in line with the CMI results, which indicated wetter than normal conditions across the MRB. We found that these relationships were weakened by spatial discrepancies in the correlation of the CMI. When analyzing CMI versus spring discharge for each HUC independently, we found a spread of R2 values ranging from 0.01 to 0.8. This reflects that the CMI was more accurate in some areas than others.



**March 2019 Correlation**

**March 2017 Correlation**

March CMI Outputs

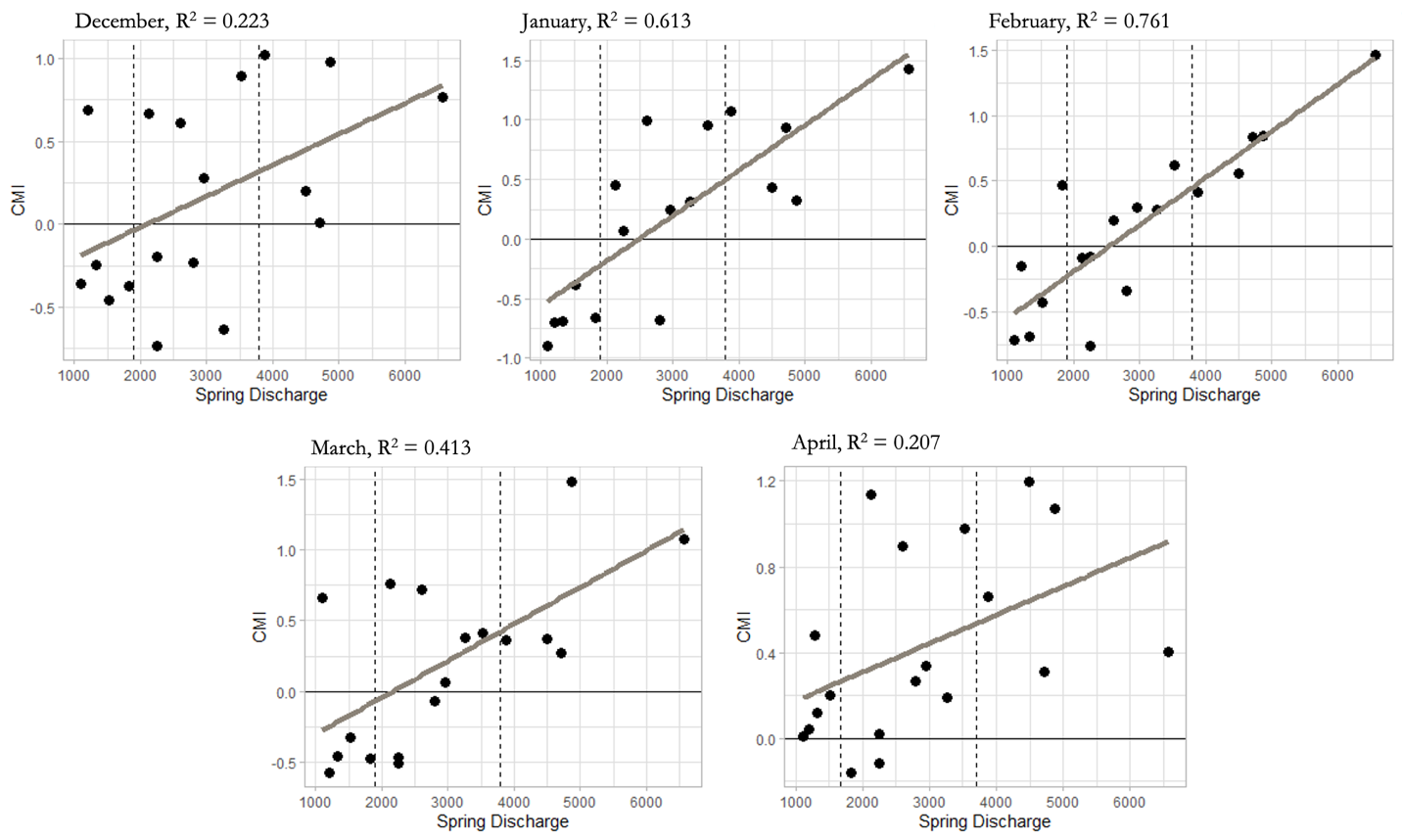
Spring Standardized Discharge

*Figure 5.* CMI versus spring discharge put into anomalies compared to historical records. Each dot represents a HUC-6 basin in the watershed. Left is March 2017 with an R2 value of .014, right is March 2019 with an R2 value of .314.

The team used the lower 25th and upper 75th percentiles of spring cumulative discharge for each subbasin to assess whether CMI values accurately captured extreme moisture states as indicated by anomalously high and low runoff (Figure 6). The team referred to CMI values that accurately described extreme discharge events as ‘hits,’ which were CMI values below zero for the lower 25th percentile of cumulative discharge data or above zero for the upper 75th percentile of cumulative discharge data. CMI values that did not accurately describe extreme discharge events were considered ‘misses’ when values in the lower 25th and upper 75th percentiles of cumulative discharge did not satisfy either of the previously stated conditions. To assess the strength of the CMI in reflecting high and low discharge values, we report the percent of hits out of all hit and missed data points. For example, in HUC 102200, December had an R2 value of 0.223 and a 90% hit percentage, which increased in January (0.613 and 100%) and February (0.761 and 90%), with a slight reduction in R2 in March (0.413 and 90%) and April (0.207 and 50%) (Figure 6). This analysis was conducted over each HUC.

*Figure 6.* CMI vs spring discharge for December through April for HUC 102200, found in the central region of our study area (see figure A4) and comprising primarily of headwater streams. Each dot represents the given month for each year from 2004 to February 2020. Linear relationships between spring discharge and CMI values are plotted in each graph, with the R2 value of these relationships indicated in the plot titles.

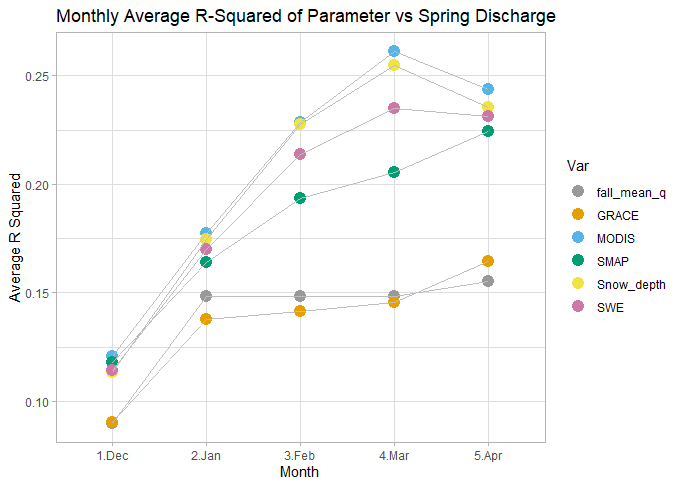
**CMI vs. Spring Discharge in HUC 102200 for December through April**



The distribution of R2 values between HUC-based CMI with spring discharge values varied spatially and temporally. The R2 relationship between the HUC-based CMI with spring discharge seems to vary spatially, but reflects an overall increase from December to March, then a decrease into April, as seen in Figure A6. The results of the summed R2 values for each HUC-6 sub-basin over all months indicate which HUC-6 sub-basins consistently have the highest R2 values (see Figure A7). The spatial distribution of high R2 values seem to lie within subbasins comprising of higher-elevations (see Figure A4) and headwaters, which are the HUCs along the perimeters of the MRB, specifically on the Western border (see Figure A7 and Figure A4). This analysis is useful to understand spatial and temporal strengths and limitations of the CMI.

We found that March CMI values had the best relationship with spring discharge across the entire MRB. The R2 value between CMI vs spring discharge averaged across all subbasins was very low (0.14) in December, increasing consistently to 0.26 in January and then 0.33 in February, peaking at 0.42 in March, before dipping down to 0.27 in April, as seen in Figure A8. This indicates that the CMI is able to best forecast spring to summer discharge rates in March, with some forecasting ability in the prior months. The increasing correlation values from December to March likely reflect building snowpack across the headwaters that peaks in March before the spring melt, which ultimately contributes to runoff.

In evaluating how each independent variable correlated with spring discharge, the team found that MODIS, SNODAS snow depth, and SWE had the respective highest correlations with spring discharge, all with increasing correlations from December to March, then decreasing in April (Figure 7). This trend was expected as all variables are snow-related. As snow accumulates in December through March, the CMI is more accurate, until snowmelt begins in April and the amount of moisture stored in snow decreases while spring discharge increases due to snowmelt runoff. SMAP has the next highest correlations, then correlations drop with fall discharge (referred to as Fall Mean Q in the graph) and GRACE. These non-snow variables see overall increases from December to April. Note that for fall discharge, one would expect the same R2 value across December to April as fall and spring discharge values are constant seasonally; however, this is not portrayed in Figure 7 as data availability varied across the basin. This analysis can aid in determining which indicators relay spring moisture conditions most effectively and can also be utilized to determine weighting schemes of the parameters used as input into the CMI.



**Monthly Average R2 of Each Parameter**

**vs. Spring Discharge**

Month

Average R2

Fall Mean Q

GRACE

MODIS

SMAP

Snow Depth

SWE

Parameter

January

February

March

April

December

*Figure 7.* Average R2 of each parameter vs spring discharge of all HUCs for December through April for each year from 2004- February 2020.

***4.2 Future Work***

The methodology used in this research was written with the intention for future automation through bash scripting. An automated process will act similarly to a forecast model where data is pulled, calculated, and displayed daily, weekly, or monthly. Upon completion of automation, the team suggests adding functionality that would allow users to select how the parameters are calculated (i.e., percentiles versus standardized anomalies).

Stream discharge calculations can be simplified in the future by using all stream gages within the MRB and sorting the data by Na values. Our method of selecting gages by hand does not make the CMI reproducible in other forecast regions if the CMI were to be adopted in another river basin. In addition, stream discharge values are influenced by upstream flow which is made up of lateral and vertical groundwater water percolation within the HUC-6 basin. To consider only local hydrology, upstream influences should be removed by subtracting the sum of all upstream HUC-6 discharge values from the from the discharge at the local HUC-6 basin. Similarly, a sensitivity analysis of how the input parameters were binned into categories and weighted in the CMI summation would help demonstrate which methods produce the strongest correspondence with spring and summer runoff. Conducting further analyses of how additional climate variables and drought indicators contribute to runoff would help in developing a dynamic weighting scheme tailored to specific subbasins and adjusted for time of year.

Analysis comparing winter CMI values to spring cumulative discharge could be improved by identifying the start of snowmelt in each sub-watershed. Cumulative spring discharge was generalized through spring and early summer months, which may not have captured the realistic snowmelt timeframe, and was likely influenced by liquid precipitation noise. To accurately capture changes in cumulative discharge as a result of winter snowmelt, future teams could conduct a time-lag analysis between CMI and runoff, calculate changes in snow cover or SWE to identify periods when snow is melting, and add a liquid precipitation component to the analysis using wet-bulb temperature to sort between solid and liquid precipitation.

# 5. Conclusions

Our project partners were interested in a product that could express a holistic understanding of soil moisture conditions in the Missouri River Basin. The conditions antecedent to spring were of particular interest as snowmelt runoff plays a key role in soil moisture conditions and flood and drought conditions thereafter. The initial NASA DEVELOP Montana Water Resources team developed the framework for a Composite Moisture Index that utilized NASA Earth observations to provide information on these winter conditions, specifically snow cover, snow depth, snow water equivalent, and soil moisture. This term, our team improved the CMI product by incorporating streamflow and groundwater measurements to more accurately describe moisture storage states during fall and winter months in the basin. We found that the CMI results are useful in reflecting general wet and dry conditions but are spatially and temporally limited. The CMI results and correlation analysis indicate that the CMI is most reliable in headwater watersheds and higher elevations, and is most accurate in March.

Our team also completed a sensitivity analysis and a principal component analysis of the CMI input parameters to understand how each metric contributes to the CMI in relation to spring streamflow conditions. This will allow our partners to adjust the weights of the input parameters to more accurately describe the impact of each variable on moisture state in the basin. In this way, the updated CMI was able to provide information on winter conditions that contributed to spring moisture states, addressing the primary need expressed by our partners. The CMI will aid in analyzing and indicating seasonal moisture conditions that can inform drought and flood planning efforts. The CMI methodology that our team developed with the guidance of our partners will be adopted by the Montana Climate Office to serve as an additional forecast tool for their drought monitoring dashboard and NOAA NWS Missouri Basin River Forecast Center will consider including the CMI as part of their operations.

# 

# 6. Acknowledgments

This project was made possible through the support and mentorship of select individuals and organizations. We extend a big thanks to the following:

* Zachary Hoylman, Research Hydrologist at the Montana Climate Office
* Kelsey Jencso, Montana State Climatologist at the Montana Climate Office
* Doug Kluck, NOAA, Regional Climate Services Director, Central Region
* Ronald Leeper, Researcher at NOAA, National Centers for Environmental Information, North Carolina Institute for Climate Studies
* Kevin Low, Hydrologist at NOAA, National Weather Service, Missouri Basin River Forecast Center
* Andrew Hoell, NOAA Physical Sciences Lab, Meteorologist
* Britt Parker, National Integrated Drought Information System, Regional Drought Information Coordinator
* Kevin Stamm, Senior Hydraulic Engineer, and Kevin Grode, Missouri River Basin Water Management Reservoir Regulation Team Lead, of the US Army Corp of Engineers Northwestern Division
* Katie Lange, NASA DEVELOP Fellow/Center Lead at NOAA National Centers for Environmental Information
* Andrew Shannon, former NASA DEVELOP Fellow/Center Lead at NOAA National Centers for Environmental Information

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 NNL16AA05C.

# 7. Glossary

**API –** Application Programming Interface

**CMI** – Composite Moisture Index

**ECDF** – Empirical Cumulative Distribution Function

**Earth Observations** – Satellites and instruments that collect remotely-sensed information about the Earth’s physical, chemical, and biological systems over space and time

**FTP** – File Transfer Protocol

**GEE** – Google Earth Engine: a cloud-based geospatial processing platform

**GRACE** –NASA's twin satellite Gravity Recovery and Climate Experiment

**HUC** – Hydrologic Unit Code, a system of classifying levels of a drainage area

**MODIS** – Moderate resolution Imaging Spectroradiometer

**NOAA** – National Atmospheric and Oceanic Administration

**SMAP** – Soil Moisture Active Passive

**SNODAS** – Snow Data Assimilation System

**SWE** – Snow Water Equivalent

# 8. References

Hall, D. K., Salomonson, V.V., and Riggs, G.A. 2016. MODIS/Terra Snow Cover Daily L3 Global 500m Grid. Version 6. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/MODIS/MOD10A1.006> [February 2021]

Hoell, A., Parker, B., Downey, M., Umphlett N., Jensco, K., Akyuz, F., Peck, D., Hadwen, T., Fuchs, B., Kluck, D., Edwards, L., Perlwitz, J., Eischeid, J., Deheza, V., Pulwarty, R., & Bevington, K. (2020). Lessons learned from the 2017 flash drought across the U.S Northern Great Plains and Canadian Prairies. *Bulletin of the American Meteorological Society, 101(12)*, E2171-E2185. <https://doi.org/10.1175/BAMS-D-19-0272.1>

Hoell, A., Perlwitz, J., & Eischeid, J. (2019). *The causes, predictability, and historical context of the 2017 U.S. Northern Great Plains drought*. Retrieved from <https://www.drought.gov/documents/causes-predictability-and-historical-context-2017-us-northern-great-plains-drought>

Li, B., H. Beaudoing, and M. Rodell, NASA/GSFC/HSL (2020), GLDAS Catchment Land Surface Model L4 daily 0.25 x 0.25 degree GRACE-DA1 V2.2, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: 24 February 2021,

<https://doi.org/10.5067/TXBMLX370XX8> [February 2021].

Li, B., Rodell, M., Kumar, S., Beaudoing H.K., Getirana, A., Zaitchik, B.F., Goncalves, L.G., Cossetin, C., Bhanja, S., Mukherjee, A., Tian, S., Tangdamrongsub, N., Long, D., Nanteza, J., Lee, J., Policelli, F., Goni, I.B., Daira, D., Bila, M., … Bettadpur, S. (2019) Global GRACE data assimilation for groundwater and drought monitoring: Advances and challenges. *Water Resources Research, 55(9)*, 7564-7586. <https://doi.org/10.1029/2018WR024618>

Mehta, V. M., Knutson, C. L., Rosenberg, N.J., Olsen, J. R., Wall, N.A., Bernadt, T.K., & Hayes, M. J. (2013) Decadal climate information needs of stakeholders for decision support in water and agriculture production sectors: A case study of the Missouri River Basin. *Weather, Climate and Society, 5(1),* 27-42. <https://doi.org/10.1175/WCAS-D-11-00063.1>

National Operational Hydrologic Remote Sensing Center. 2004. Snow Data Assimilation System (SNODAS) Data Products at NSIDC, Version 1. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. <https://doi.org/10.7265/N5TB14TC> [September 2020].

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, and R. Bindlish (2016). SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 4. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/ZX7YX2Y2LHEB> [September 2020].

Ryberg, K. R. and Aldo V. Vecchia (2017). waterData: Retrieval, Analysis, and Anomaly Calculation of Daily Hydrologic Time Series Data. R package version 1.0.8. <https://CRAN.R-project.org/package=waterData>

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

Shafer, B.A. & Dezman L.E. (1982). Development of a surface water supply index (SWSI) to assess the severity of drought conditions in snowpack runoff areas. *Proceedings of the* *50th Annual Western Snow Conference, USA.*

U.S. Geological Survey. (2021). U.S. Geological Survey (USGS) daily hydrologic data from USGS web ser-vices (USGS Water Data for the Nation), accessed 24 February 2021. <https://doi.org/10.5066/F7P55KJN>

Wickham, H. (2020). httr: Tools for Working with URLs and HTTP. R package version 1.4.2. <https://CRAN.R-project.org/package=httr>

# 9. Appendix

Diagram

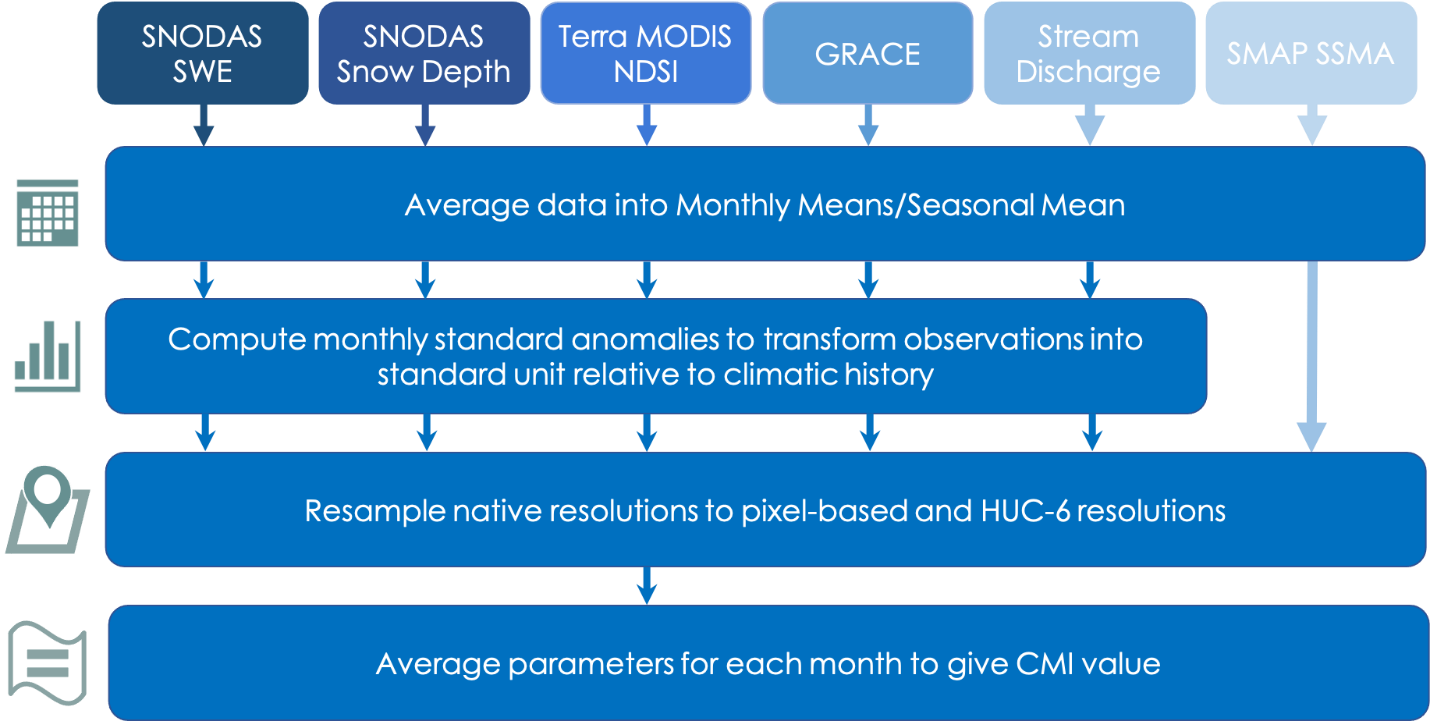
Description automatically generated

*Figure A1.* Respective HUC-6 labels within the MRB used to delineate the HUC-base CMI.

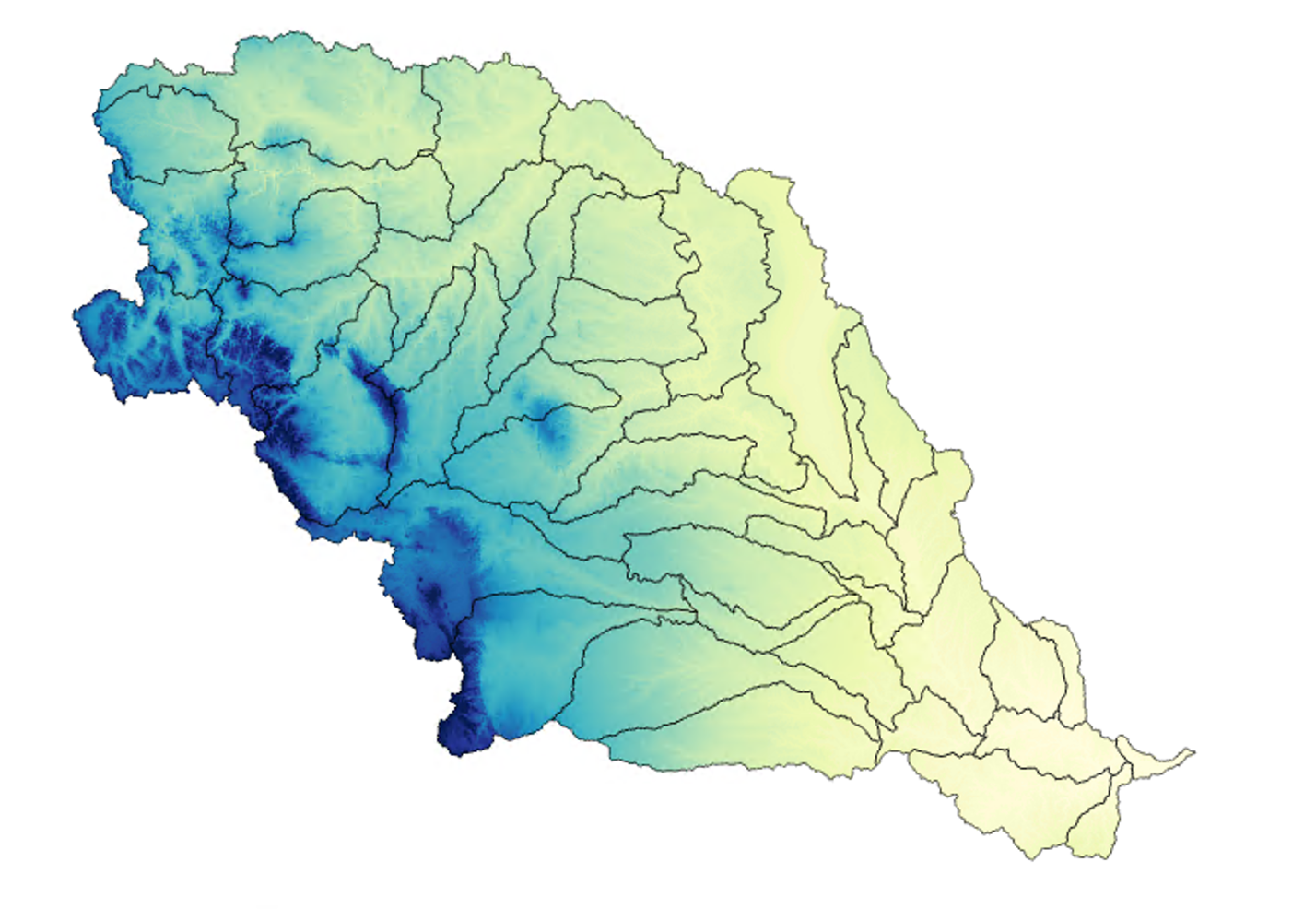
Diagram, map

Description automatically generated

*Figure A2.* Missouri River Basin extent with HUC-6 sub-basins delineated in grey. The red dots indicate the location of the chosen USGS stream gages for this study.



*Figure A3.* Data processing workflow to calculate CMI, read from top to bottom. On the top row are the different datasets used to calculate the final CMI.



Elevation

High

Low

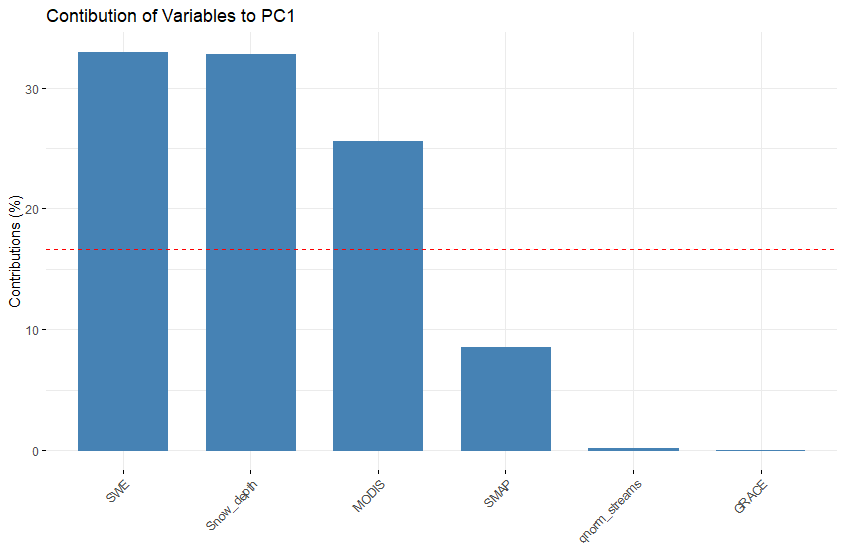
**Missouri River Basin Elevation Simplified**

*Figure A4.* Elevation in the MRB with yellow to green representing lowland regions and green to dark blue representing highland regions.

*Figure A5.* Percent contribution to the first principal component of each of the six moisture indicators. The red dashed line is a reference line that corresponds to the expected value if the contributions of each indicator were equal.

**Percent Contribution of Each Moisture Indicator**

**to Principal Component 1**



SWE

Snow Depth

MODIS

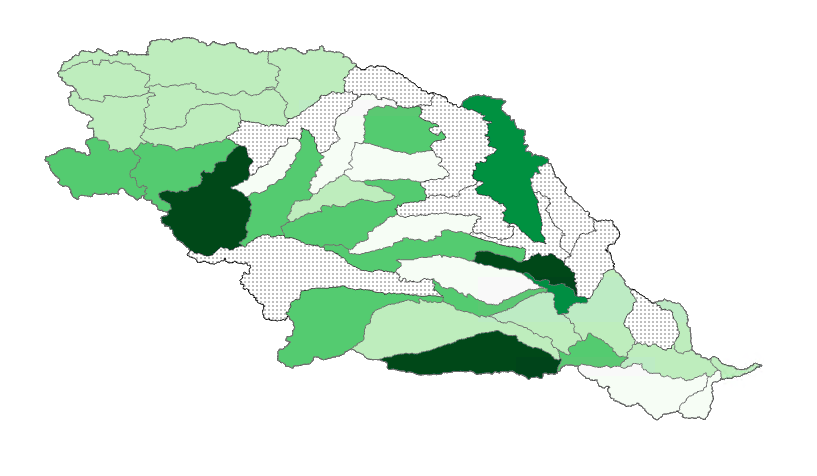
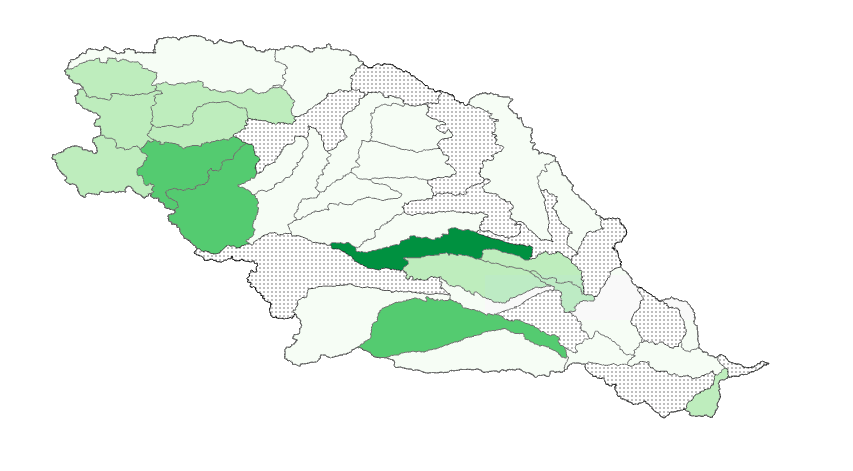
SMAP

Stream Discharge

GRACE

Moisture Indicators

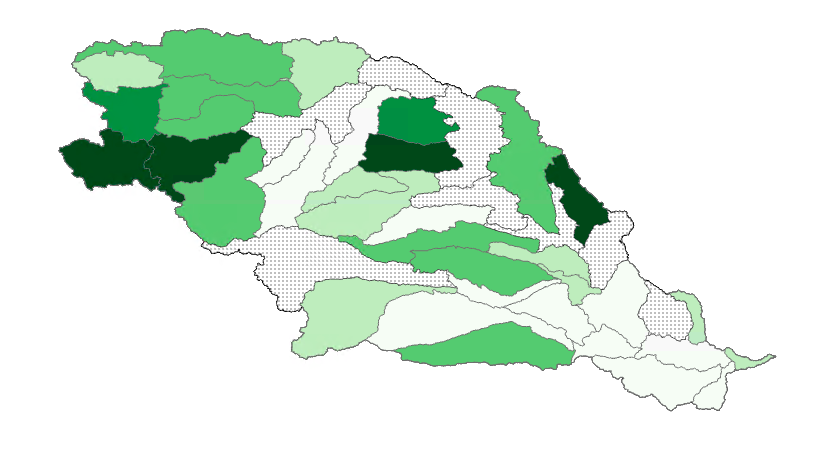
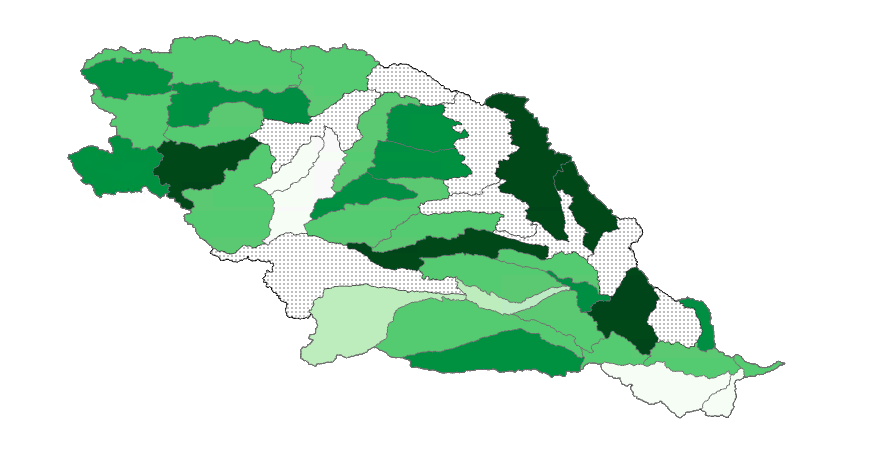
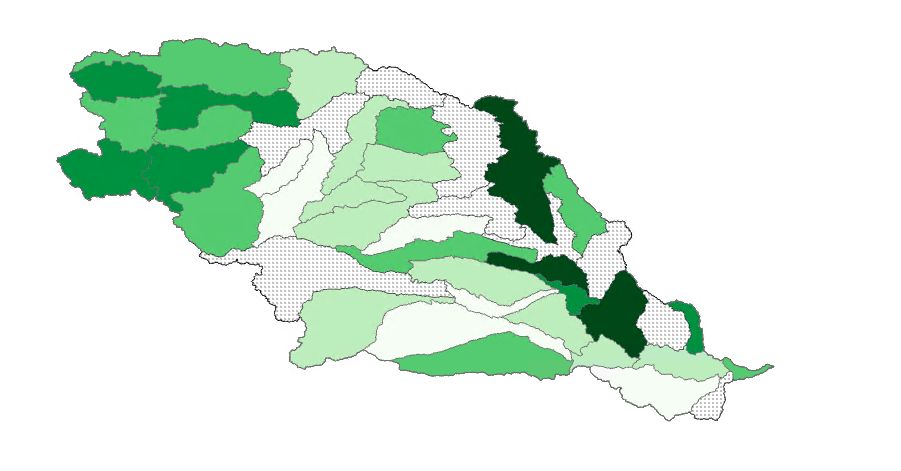
Contributions (%)



December

January

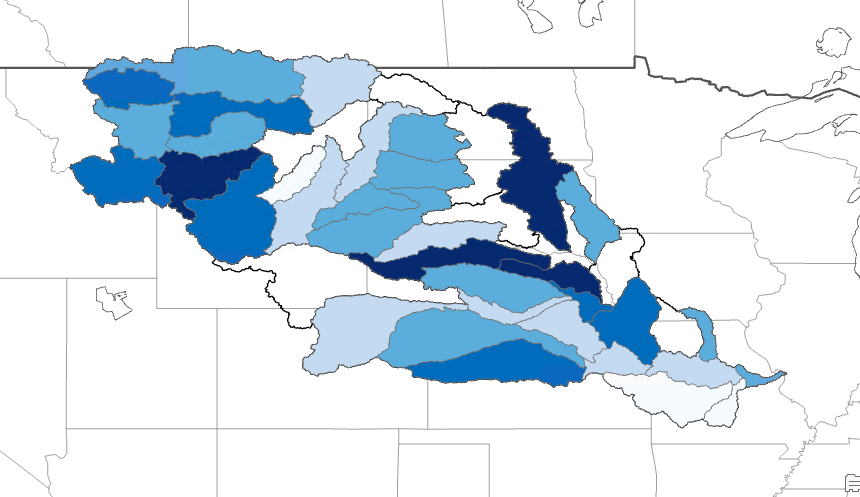
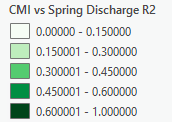
February



March

April

**HUC-based CMI vs. Spring Discharge R2 Values**



No Data

0.00 – 0.15

0.15 – 0.30

0.30 – 0.45

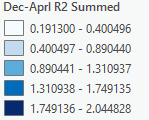
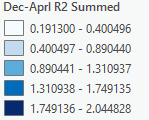
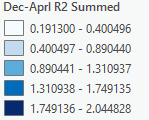
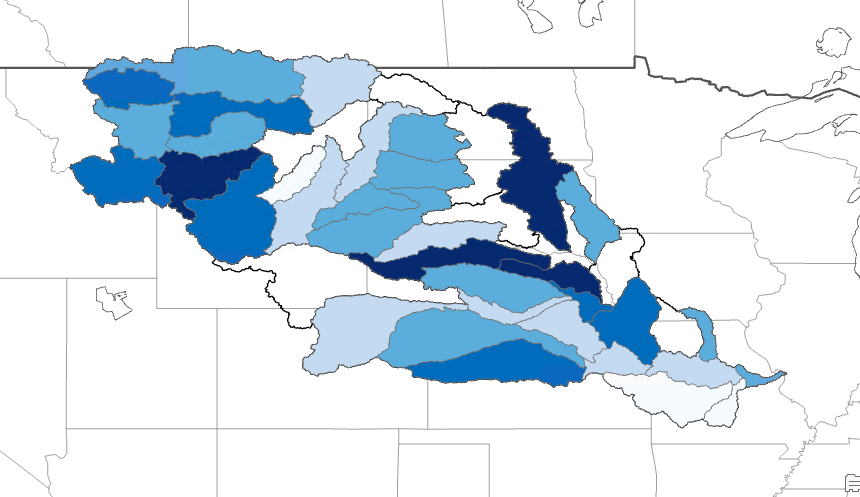
0.45 – 0.60

0.60 – 1.00

R2 Values

*Figure A6.* The spatial distribution of HUC-based CMI results compared to spring streamflow discharge values in the MRB. Darker green hues indicate areas where the total R2 value is higher, while lighter green hues indicate areas where total R2 are lower. The grey-dotted sub-basins represent areas where there is no stream gage data.

CMI vs. Spring Discharge R2 Summed for all Months and Years



Total R2 Summed

0.19 – 0.40

0.40 – 0.89

0.89 – 1.31

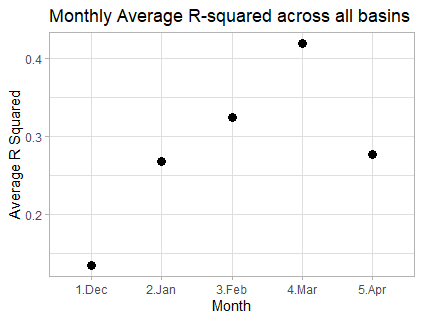
1.31 – 1.75

1.75 – 2.04

No Data

**CMI vs. Spring Discharge R2 Summer for all Months and Years**

*Figure A7.* The sum of R2 values in each sub-basin for all years and months from the period of record. Darker blues indicate areas where the total R2 value is higher, while lighter blues indicate areas where total R2 are lower. White or transparent represents areas where there is no stream gage data.



**Monthly Average R2 Across All Basins**

Month

Average R2

January

February

March

April

December

*Figure A8.* Average R2 of CMI vs spring discharge of all HUCs for December through April each year from 2004- February 2020.