Examining the Hydrologic Properties of the Missouri River Basin

https://github.com/cwatson1013/Hydrologic_Data_Analysis_Final_Proj

Rachel Bash, Keqi He, Caroline Watson, and Haoyu Zhang

1 Rationale and Research Questions

The Missouri River is the largest river in North America (2,540 miles) and has the second-largest watershed spanning 529,000 mi^2 (Kammerer 1990). Its watershed covers portions of ten states, which account for approximately one-sixth of the continental United States, as well as a small part of Canada (U.S. Army Corps of Engineers Missouri River Basin Water Management Division 2018). The headwater is located in the Bitterroot Mountains River of northwestern Wyoming and southwestern Montana (Committee on Missouri River Ecosystem Science 2002). The basin is primarily rural but also has several large and medium-sized cities. In 2012, the basin was home to approximately 15.5 million people, and many large cities in the region are in proximity to the Missouri River (U.S. Army Corps of Engineers Missouri River Basin Water Management Division 2018). Demands for managing the river to benefit human livelihoods has resulted in a drastic modification of the river and floodplains. Now, the river is used intensively in multiple ways, including municipal, agricultural, hydropower, flood control, and recreation (Bureau of Reclamation 2016b).

Within the 328 million acres of the basin's total area in the United States, 64.2% are for agricultural use, while the rest are for recreation, fish and wildlife, and urban development (Bureau of Reclamation 2016b). 37.1% of the total basin area is pasture and range grassland primarily for grazing, and cropland consists of almost 92 million acres (U.S. Army Corps of Engineers Missouri River Basin Water Management Division 2018). As of 2012, irrigated land comprises of 14.2 million acres, and require delivery of approximately 13 million acre-feet of surface water annually (U.S. Army Corps of Engineers Missouri River Basin Water Management Division 2018). Wetlands and water bodies, on the other hand, cover 3.7 and 1.8 million acres, respectively (U.S. Army Corps of Engineers Missouri River Basin Water Management Division 2018). Despite of the low proportion of water areas (2.3%), they are the pivotal foundation for agricultural or other usages, and thus critical to the whole region's economy.

The agricultural, urban, and industrial development in the region inevitably causes nutrient loading and enrichment in water bodies, especially for nitrogen (N) and phosphorus (P). Agricultural input through fertilizer is the predominant anthropogenic source for nutrients in water bodies in the basin (Committee on Missouri River Ecosystem Science 2002). Furthermore, nutrient enrichment is considered nationally as one of the leading factors for water quality impairment. According to the Clean Water Act 303(d) list (2015), more than 160 stream reaches or waterbodies were reported by the United States Environmental Protectin Agency (US EPA)to suffer nutrient-related impairment in 2015.

In addition to the fluctuations in nutrient concentration, discharge appears to be highly variable in the basin, and both severe drought and flooding events have occurred in the basin regularly. For example, in the spring and summer of 2011, an unprecedented flooding event caused over $2 billion damage and five fatalities, leading to FEMA disaster declaration made in all states along the Missouri River (National Oceanic and Atmospheric Administration 2012). During the flooding event, around 11,000 people evacuated from Minot, North Dakota, owing to the high water level in the Souris River, which flooded 4,000 homes (National Oceanic and Atmospheric Administration 2012). In 2012, a drought struck the Central Great

Plain, including the basin, and inflicted at least $12 billion of agricultural loss before July 2012 (NOAA Drought Task Force 2013). Recently, another flooding event occurred in the spring of 2018, due to January through May being the wettest period on record for the U.S.

Given all the background information, we would like to know the current state of Missouri River and its tributaries, with a focus on the changing patterns in discharge and nutrient levels. Water bodies along the downstream are more likely to be impaired by nutrients accumulated from the upstream. Croplands and pastures, which can further load nutrients into streams, are distributed throughout the lower basin (Figure 1). Therefore, study sites for this analysis were concentrated in the southeast of the Missouri River Basin. By analyzing data retrieved from these sites, we first revealed the general yearly discharge pattern and changes in variability over the years. Then we investigated how the dramatic change in discharge (i.e., water quantity) could potentially interact with nutrient enrichment (i.e., water quality). We examined a few specific flooding and drought events, during which changes in both water quality and quantity were well recorded so that inferences on the interplay between quantity and quality could be made. The effect of population on nutrient enrichment was also examined to illustrate potential non-agricultural impact from human activities. Finally, based on the pattern in the past and the best model we could fit, we attempted to predict the likely future conditions and trends in the Missouri River Basin. Our research questions and accompanying hypotheses are below:

1. How have changes in discharge (i.e. water quantity) interacted with nutrient enrichment (i.e. water quality) in the Missouri River Basin?
    a) Discharge has become more variable over time
    b) Nutrient levels have increased over time
    c) Nutrient levels increase with discharge
2. What effects do specific flood and drought events have on the water quality and quantity of rivers in the Missouri River Basin areas of interest?
    a) Rivers will exhibit a flushing behavior due to the land use and type of flow during storms
    b) Discharge will decrease during drought due to decreased overland flow.
3. What factors contribute to the variability of total nitrogen in the rivers?
    a) Land use, year, discharge, phosphorus, and HUC region will contribute to the variability of total nitrogen across sites
4. Given past and current data, what can we predict about the future state of water in the Missouri River Basin?
    a) Total flow in the Missouri River Basin is increasing (non-stationary) over time
    b) The future situation of the river basin will see the continuation of current trends of increasing overall volume of flow.

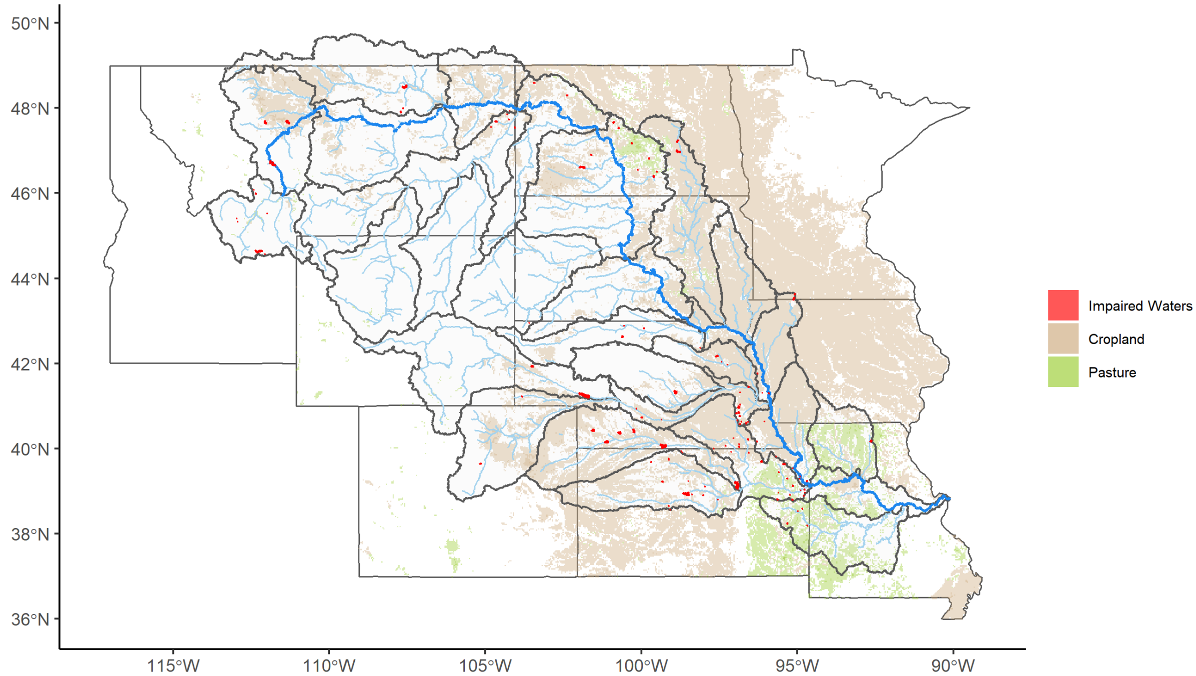

Figure 1: Map of agricultural lands and impaired waters in the Missouri River Basin. Row and Close Grain Crop Cultural Formation shown by tan-colored areas, pasture and hay field shown by yellowgreen areas, imparied water bodies according to Clean Water Act Section 303(d) (EPA 2015) denoted by red areas, and all HUC4 watersheds in the Missouri River Basin (1001-1030) delineated by gray polygons. The thin, light-blue lines outline all streams of a Strahler order higher than 2, and the thick blue line represents the mainstem of Missouri River.

2 Dataset Information

The data we are analyzing comes from the United States Geological Survey (USGS) database called the National Water Information System interface, or NWIS. We pulled data from the interface using the R package `dataRetrieval`. Because we are interested in the southern part of the Missouri River basin, we pulled sites from each HUC4 subbasin from 1020 to 1030. We chose these subbasins because they had a variety of tributaries that all flowed into the Missouri River, representing a variety of river sizes and lengths. We filtered the subbasin queries to only show sites that contained discharge, nitrogen, and phosphorus data. Once we found the sites containing all the data needed, we chose two sites from each HUC subbasin for a total of 22. We chose the two sites from each HUC subbasin by comparing the periods of records for each of our chosen variables and finding the sites with the longest periods of records. We chose to look at two sites per HUC region (for a total of 22) in order to maintain a digestible scope (Figure 2). We retrieved data on water quantity and water quality (N, P concentrations) (Table 1).

Only six sites within our HUC subbasin boundary contained high-frequency nitrogen data. Therefore, we looked at these six sites to analyze floods and answer our research question about floods. Since flooding and droughts can be thought of as opposites, we looked at the same six sites for droughts as we did for flooding (Table x - add table??).

The data we are analyzing comes from the United States Geological Survey (USGS) database called the National Water Information System interface, or NWIS. We pulled data from the interface using the R package `dataRetrieval`. Because we are interested in the southern part of the Missouri River basin, we pulled sites from each HUC4 subbasin from 1020 to 1030. We chose these subbasins because they had a variety of tributaries that all flowed into the Missouri River, representing a variety of river sizes and lengths. We filtered these subbasin queries to only show us sites that contained discharge, nitrogen, and phosphorus data. Once we found the sites with all of this data, we chose 2 sites from each HUC sub basin for a total of 22. We chose the two sites from each HUC sub basin by comparing the periods of records for each of our chosen variables and finding the sites with the longest periods of records. We chose to look at two sites per HUC region (for a total of 22) in order to maintain a digestible scope (Figure 2; Table 1). From NWIS, we retrieved data on water quantity and water quality (Table 2). From R's `noncensus` package, we retrieved data on county population for counties that our sites are located.

Only six sites within our HUC subbasin boundary contained any high frequency nitrogen data. Therefore, we also looked at these six sites in order to do analyses and answer our research question about flooding. Since flooding and droughts can be thought of as opposites, we looked at the same six sites for droughts as we did for flooding.

We have two main datasets:

- The daily values dataset containing discharge, nitrogen, and phosphorus data for 22 sites.
- The high-frequency dataset containing high-frequency data for nitrogen and discharge

for 6 sites.

Table 1: A list of the 22 chosen sites in the lower Missouri River Basin

No.	Site Number	Site Name	HUC 4 Region	State	County Name	County Population
1	6768000	Platte River Near Overton, NE	1020	NE	Dawson County	24326
2	6805500	Platte River At Louisville, NE	1020	NE	Sarpy County	158840
3	6775900	Dismal River Near Thedford, NE	1021	NE	Thomas County	647
4	6794000	Beaver Creek At Genoa, NE	1021	NE	Nance County	3735
5	6800000	Maple Creek Near Nickerson, NE	1022	NE	Dodge County	36691
6	6800500	Elkhorn River At Waterloo, NE	1022	NE	Douglas County	517110
7	6610000	Missouri River At Omaha, NE	1023	NE	Douglas County	517110
8	6609500	Boyer River At Logan, IA	1023	IA	Harrison County	14928
9	6810000	Nishnabotna River Above Hamburg, IA	1024	MO	Atchison County	5685
10	6818000	Missouri River At St. Joseph, MO	1024	MO	Buchanan County	89201
11	6844500	Republican River Near Orleans, NE	1025	NE	Harlan County	3423
12	6856600	Republican R At Clay Center, KS	1025	KS	Clay County	8535
13	6877600	Smoky Hill R At Enterprise, KS	1026	KS	Dickinson County	19754
14	6874000	Sf Solomon R At Osborne, KS	1026	KS	Osborne County	3858
15	6887500	Kansas R At Wamego, KS	1027	KS	Pottawatomie County	21604
16	6892350	Kansas R At Desoto, KS	1027	KS	Johnson County	544179
17	6902000	Grand River Near Sumner, MO	1028	MO	Livingston County	15195
18	6905500	Chariton River Near Prairie Hill, MO	1028	MO	Chariton County	7831
19	6921070	Pomme De Terre River Near Polk, MO	1029	MO	Polk County	31137
20	6926510	Osage River Below St. Thomas, MO	1029	MO	Cole County	75990
21	6894000	Little Blue River Near Lake City, MO	1030	MO	Jackson County	674158
22	6934500	Missouri River At Hermann, MO	1030	MO	Gasconade County	15222

Table 2: Range and Count of Relevant Data

Parameter	Unit	First Date	Last Date	Minimum Value	Maximum Value	Total Count	Type of Data
Discharge	cfs	1899-10-01	2019-11-06	0.00	739000.0	645422	daily and high frequency
Nitrogen	mg/L	1967-06-13	2019-10-16	0.02	100.0	6207	daily and high frequency
Phosphorus	mg/L	1969-07-23	2019-10-16	0.00	28.5	7021	daily

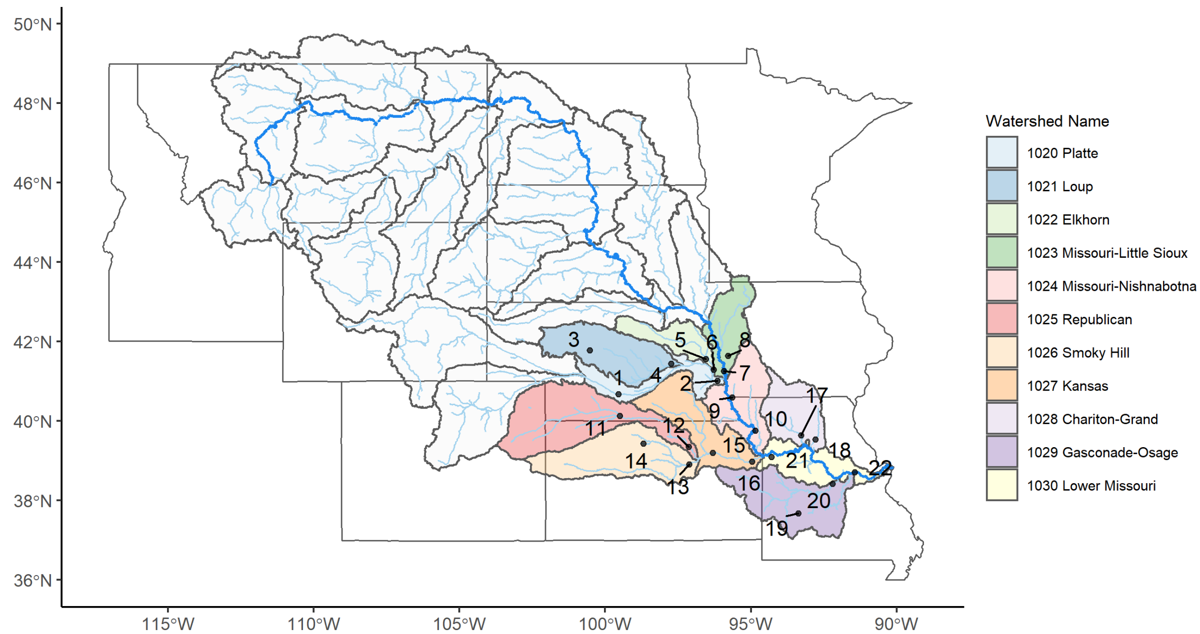

Figure 2: Map of USGS sites used for long term analysis.

3 Exploratory Analysis

3.1 Exploration of variables

Basic data wrangling was needed in order to obtain all variables of interest. After obtaining all pertinent data, each variable was visualized in order to see the shape of the data and the range of values (Figure 3). Any necessary transformation or cleaning was completed after this step. Additionally, discharge, nitrogen, and phosphorus were plotted for each site and examined together in order to see if there were any obvious patterns or trends (Figure 4).

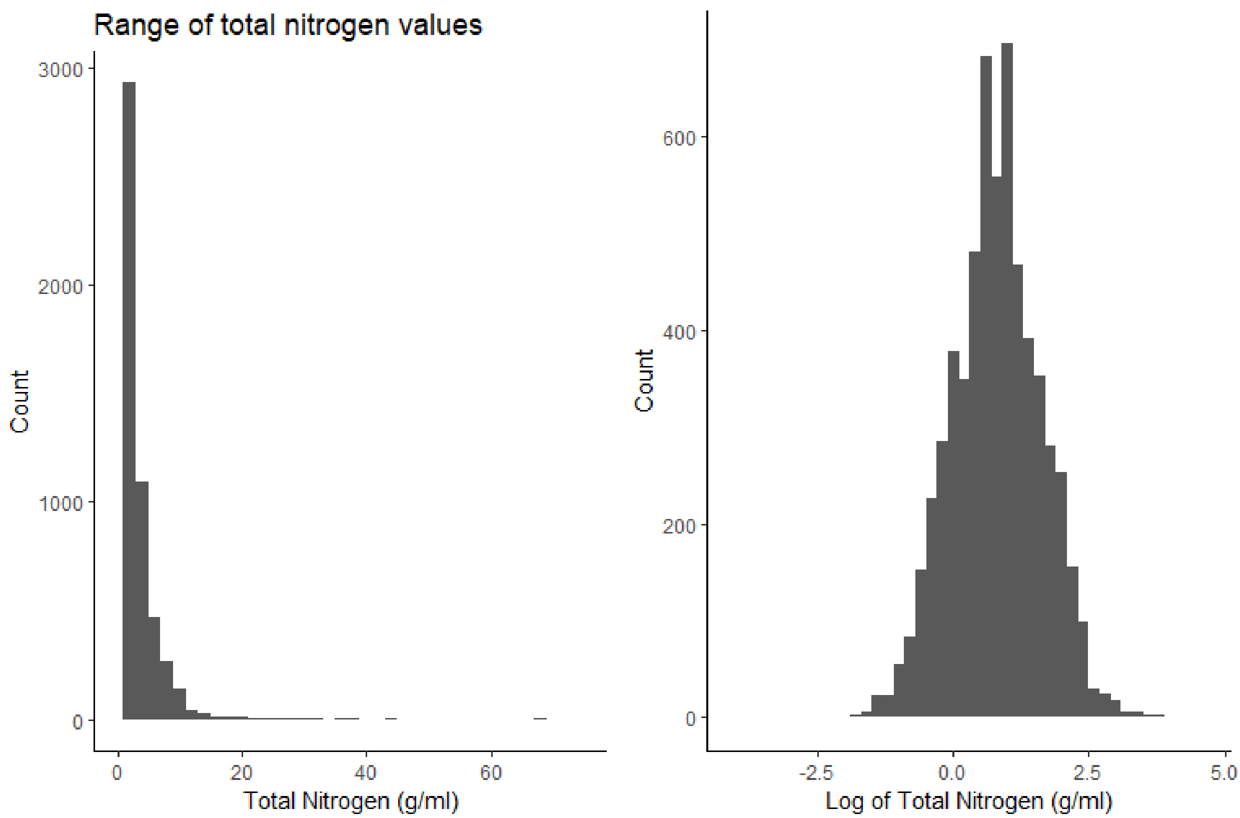

Figure 3: Histograms showing the range of total nitrogen values obtained from all sites during the total period of record. Raw (left) and logged (right) nitrogen data are shown to show that any analysis requiring normally distributed data will use the log transformed nitrogen data.

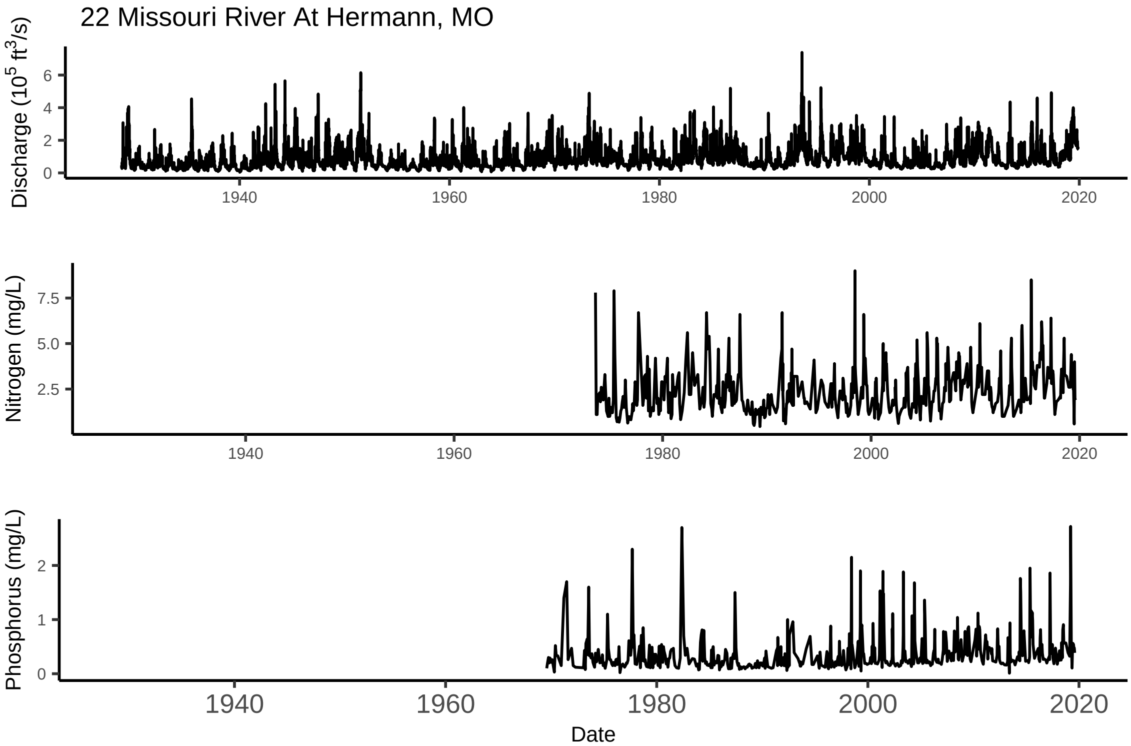

Figure 4: Discharge, Nitrogen, Phosphorus daily data of Site No. 22, the Missouri River At Hermann, MO.

3.2 Yearly Discharge Pattern

Typical discharge patterns over the course of a year for each HUC 4 watershed from 1020 to 1030 were generated by compiling all available discharge data at the 22 selected USGS sites (Figure 5). Each site was graphed separately, and then compared with the over site in its HUC 4 region. Generally, discharge reaches its peak during the summer and falls to minima during the winter, and most of the sites exhibit rather high variations across years, as indicated by the large difference between the medians and the first or the third quartiles. Furthermore, highest variations in discharge appear to occur in the summer, whereas discharge in the winter varies less among years. The large variability within a year and the seasonal pattern is only obvious in larger streams and rivers but not in small creeks (e.g. site 5 Maple Creek near Nickerson, site 8 Boyer River, site 21 Small Blue River).

Exceptions to the generalization are three sites with medium discharge (site 1 Platte River, NE; site 3 Dismal River, NE; site 11 Republican River, NE). They have fairly constant discharge and variation that doesn't change over the course of a year, and the the first and the third rivers mentioned even have slightly lower discharge during the summer. Platte River forms by the confluence of North Platte and South Platte Rivers, and both stems have snowmelt as their water source, resulting in the observed higher discharge during the spring and low in the summer. Dismal River forms by two forks that arise from groundwater (Ogallala Aquifer), which is supposed to be more stable than precipitation. The majority of Republican River basin is underlain by Ogallala Aquifer (Bureau of Reclamation 2016a), and a plausible high proportion of water source from groundwater could explain the little variation across years.

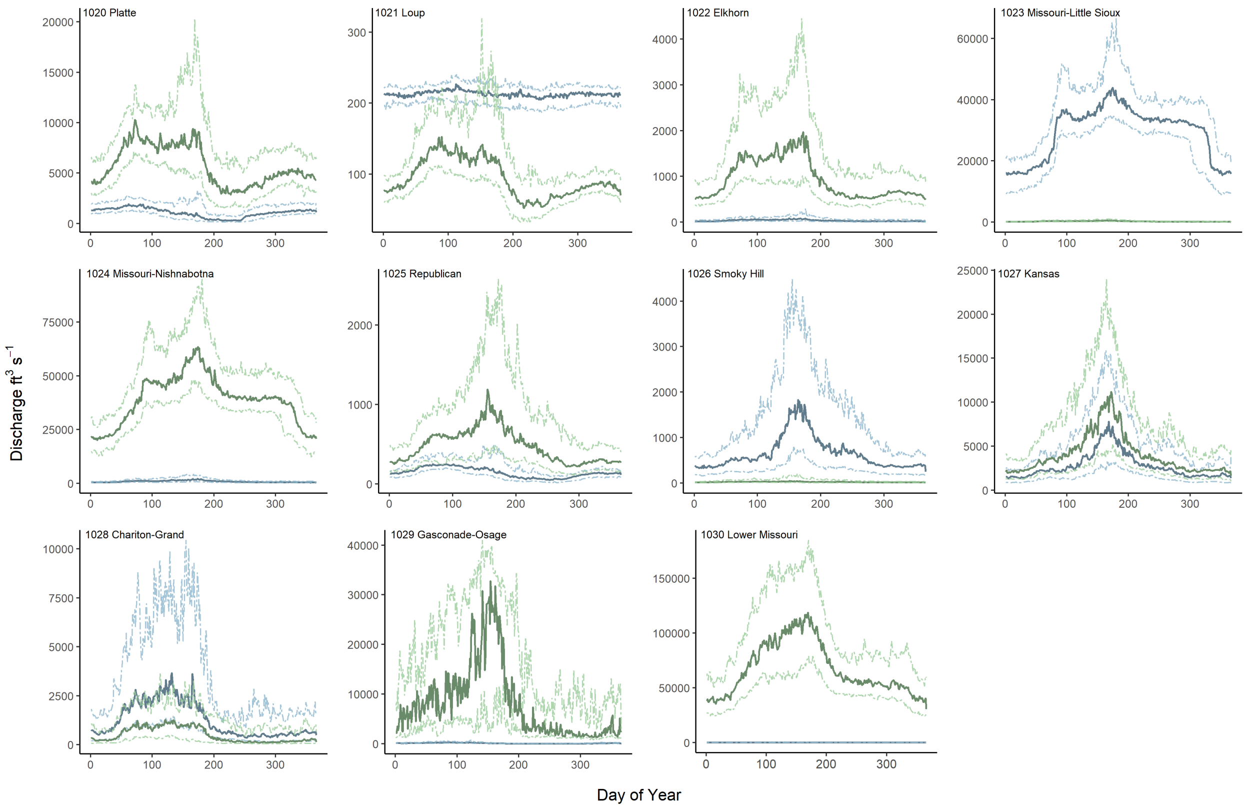

Figure 5: Yearly discharge pattern for the lower Missouri HUC 4 watersheds. Thick, solid, and dark lines are the median of all discharge records on a day of year, while thin, dashed, and light lines are the 25th and 75th discharge quantiles on that day. The blue lines represent the downstream site in a HUC4 watershed, while the green lines represent the upstream site.

3.3 Change in Discharge Variations over Years

To reveal whether discharge has become more variable over time, we first used linear model to analyze the relationship between standard deviation (SD) of discharge within a year and across years (Figure 7; Table 3). Among all the 22 sites, 14 sites seemed to increase in SD over time, and four of them show statistically significant increase. By contrast, only eight sites had declining SD trends, and two of them exhibited statistically significant increased SD, or more absolute variability. Despite that only data from six sites (around 27%) yielded conclusive results, there were twice as many streams with increasing SD than with decreasing SD. Thus, our results suggest that the whole Missouri Basin appear to have increased in absolute variability over years since around the mid-19th century. The low percentage of significant results could result from the small effect size, and/or the short time span of available data for some sites.

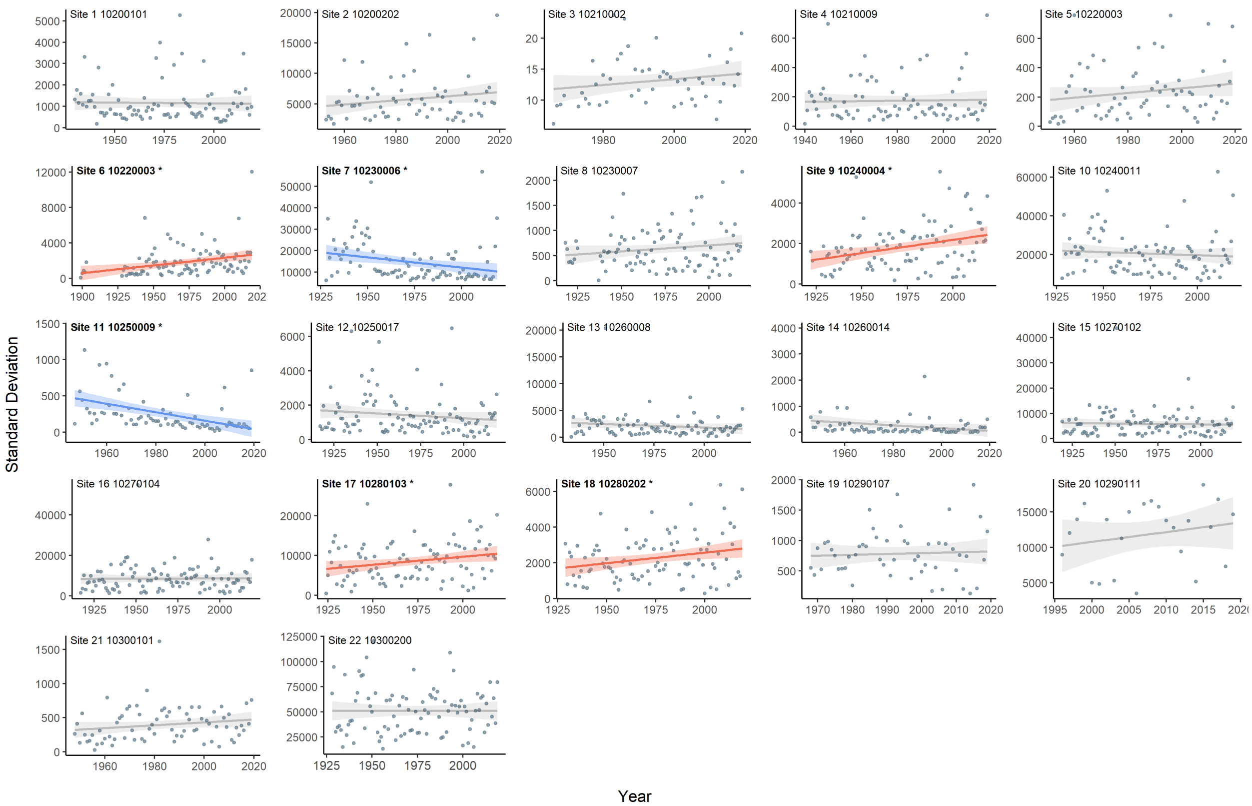

Figure 6: Exploratory analysis examining the changes in the standard deviations (SD) of discharge within a year at 22 sites over the whole time span of available data. Red regression lines and confidence intervals represent increasing SD, whereas blue for decreasing SD over time. Bolded site labels and darker regression lines represent significant results, whereas regression lines in gray are non-significant ($\alpha = 0.05$).

Another statistic that is commonly used in academic literature for comparing variability of streams is coefficient of variation (CV) (e.g. Jowett and Duncan 1990; Coe and Birkett 2004; Déry 2005). It is defined as the ratio of standard deviation to the mean, and therefore accounts for difference in the mean. Linear regression was also performed for the relationship between CV and years (Figure 7; Table 3). Unlike inconsistent results for SD, 12 out of 22 sites produced significant results, and 11 sites exhibited declining trends in CV over years, whereas only one site had a rising trend. 9 sites had non-significant trends for SD while declining trends for CV (site 1, 2, 5, 8, 10, 15, 16, 21, 22), one site had increased SD but decreased CV (site 9), and three sites had increasing SD and non-significant changes in CV (site 6, 17, 18). These contrasts suggest their mean discharges appear to rise over time. The inference accords with our later findings on discharge over years examined by time series analysis (see Question 4). Site 7 had both decreasing SD and CV but with a more negative slope for CV, and site 11 had decreasing SD but non-significant CV. Non-significant change in SD but rising CV was observed at site 12, whose discharge was later shown to diminish (see Question 4). No site had declining SD and increasing CV, and all the remaining sites produced non-significant results for both variability statistics.

Overall, our results indicate that the absolute values of variability represented by SD seem to increase at more sites, but after considering mean discharge, the majority of streams and rivers investigated showed less relative variability. It is uncertain, however, if such trends in relative variability would persist in the future. More monitoring on discharge in the basin are required in the future to reach a more definitive conclusion. Moreover, with regards to causing casualties and damage to properties, absolute variability may be more informative than relative values. Hence, the choice between the two statistics depends on the context in which the rivers are analyzed and findings are interpreted (e.g. academic research or water management).

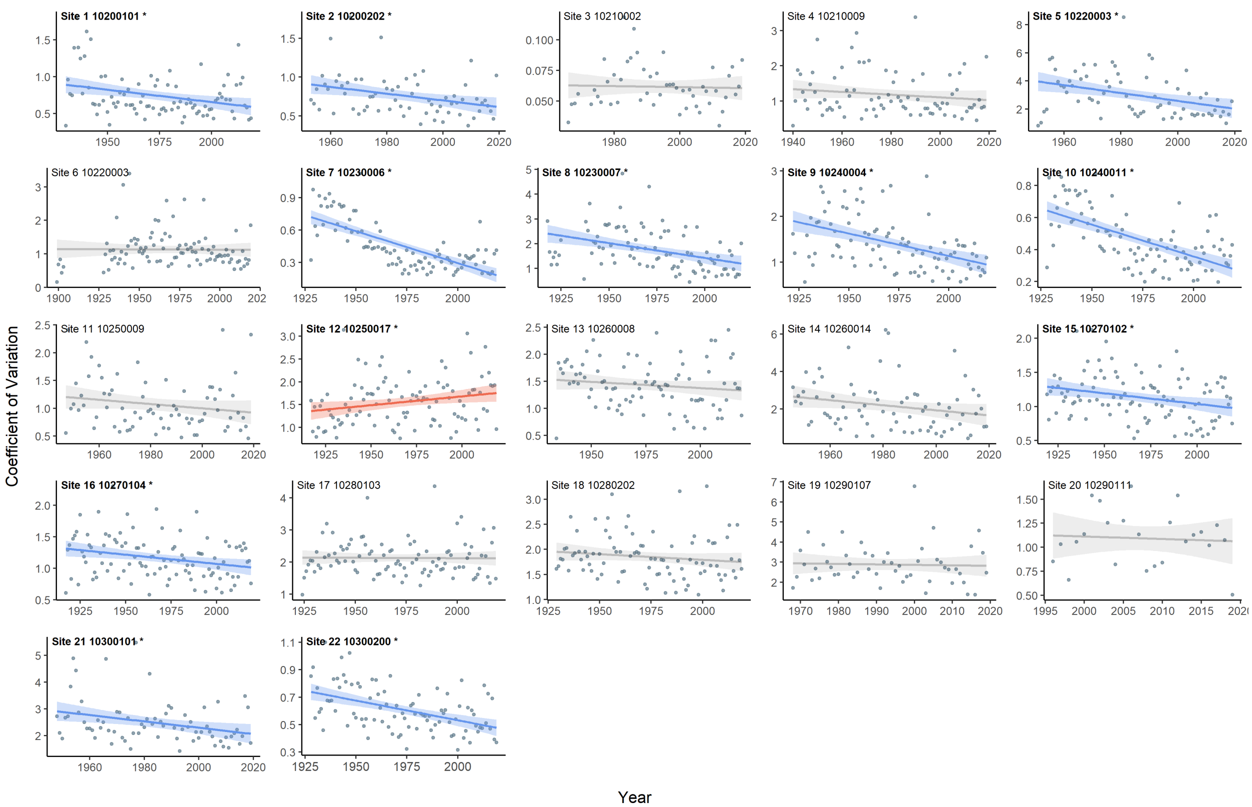

Figure 7: Exploratory analysis examining the changes in the coefficient of variation (CV) of discharge within a year at 22 sites over the whole time span of available data. Red regression lines and confidence intervals represent increasing CV, whereas blue lines for decreasing CV over time. Bolded site labels and HUC8 with asterisk represent significant results, while regression lines in gray color are non-significant ($\alpha = 0.05$).

Table 3: Slopes of linear regression between year and standard deviation or coefficient of variation of discharge at 22 sites.

			Standard Deviation		Coefficient of Variation	
No.	Site Name	HUC4	Slope	*P*-value	Slope	*P*-value
1	Platte River Near Overton, NE	1020	-0.805	0.830	-0.003	0.004**
2	Platte River At Louisville, NE	1020	33.691	0.144	-0.004	0.008**
3	Dismal River Near Thedford, NE	1021	0.047	0.196	0.000	0.805
4	Beaver Creek At Genoa, NE	1021	0.168	0.812	-0.004	0.198
5	Maple Creek Near Nickerson, NE	1022	1.613	0.145	-0.028	0.002**
6	Elkhorn River At Waterloo, NE	1022	17.400	0.002**	0.000	0.911
7	Missouri River At Omaha, NE	1023	-95.171	0.009**	-0.006	0.000***
8	Boyer River At Logan, IA	1023	2.404	0.134	-0.012	0.000***
9	Nishnabotna River Above Hamburg, IA	1024	13.001	0.002**	-0.010	0.000***
10	Missouri River At St. Joseph, MO	1024	-32.123	0.452	-0.004	0.000***
11	Republican River Near Orleans, NE	1025	-5.861	0.000***	-0.004	0.139
12	Republican R At Clay Center, KS	1025	-5.579	0.157	0.004	0.020*
13	Smoky Hill R At Enterprise, KS	1026	-12.875	0.226	-0.002	0.227
14	Sf Solomon R At Osborne, KS	1026	-5.271	0.072	-0.014	0.056
15	Kansas R At Wamego, KS	1027	-6.676	0.714	-0.003	0.006**
16	Kansas R At Desoto, KS	1027	2.556	0.910	-0.003	0.007**
17	Grand River Near Sumner, MO	1028	40.400	0.026*	0.000	0.916
18	Chariton River Near Prairie Hill, MO	1028	11.861	0.020*	-0.002	0.200
19	Pomme De Terre River Near Polk, MO	1029	1.389	0.704	-0.002	0.798
20	Osage River Below St. Thomas, MO	1029	138.275	0.311	-0.003	0.769
21	Little Blue River Near Lake City, MO	1030	2.096	0.142	-0.012	0.009**
22	Missouri River At Hermann, MO	1030	0.657	0.994	-0.003	0.000***

Note:
Significance level
* <0.05;
** <0.01;
*** <0.001

4 Analysis

4.1 Question 1: How have changes in discharge interacted with nutrient enrichment in the Missouri River Basin?

We analyzed nitrogen and phosphorus trends at 22 sites using long-term daily value data and the Seasonal Mann-Kendall Trend Test. The Seasonal Mann Kendall Trend Test is used to analyze seasonal data collected over time for consistently increasing or decreasing trends (monotonic) in Y values. It is suitable for analyzing the data with seasonality, non-parametric, no temporal autocorelation, identical distribution.

The trend of nitrogen and phosphorus concentration of 22 sites are quite different across the

22 sites (Table 4). Among all 22 sites, 7 sites show significantly increasing trends in nitrogen concentration, while 7 show decreasing trends. For phosphorus, concentrations at 7 sites have increased over time, while decreased at other 5 sites. But generally, the trend of nitrogen and the trend of phosphorus are similar, which indicates that there is a high possibility that nitrogen and phosphorus come from the same sources. Despite inconsistent results at all sites, those on the mainstem show higher nutrient concentration (No. 7, Missouri River at Omaha, NE; No. 10, Missouri River at St. Joseph, MO; No. 22, Missouri River at Hermann, MO), we can see a significant increasing trend of nitrogen and phosphorus. Our results suggest that water quality in the Missouri River is still deteriorating and protection is urgently needed.

Table 4: Nitrogen and phosphorus trend over time for each site

No.	Site Number	Nitrogen Trend	Nitrogen Trend Significance	Phosphorus Trend	Phosphorus Trend Significance
1	06768000	Decrease	*	Decrease	***
2	06805500	Increase	***	Increase	***
3	06775900	Decrease	***	Decrease	
4	06794000	Increase		Increase	
5	06800000	Increase	***	Increase	***
6	06800500	Increase	*	Increase	
7	06610000	Increase	**	Increase	***
8	06609500	Decrease		Increase	
9	06810000	Decrease		Increase	
10	06818000	Increase	·	Increase	
11	06844500	Increase	·	Decrease	**
12	06856600	Decrease	**	Decrease	***
13	06877600	Decrease	·	Increase	
14	06874000	Decrease	***	Decrease	***
15	06887500	Increase	***	Increase	***
16	06892350	Increase	**	Increase	***
17	06902000	Decrease	***	Increase	
18	06905500	Decrease		Increase	
19	06921070	Decrease	·	Decrease	
20	06926510	Decrease	***	Increase	***
21	06894000	Decrease	***	Decrease	***
22	06934500	Increase	***	Increase	***

Note: $***$ means p-value < 0.001; $**$ means $0.001 \leq$ p-value < 0.01;$*$ $0.01 \leq$ p-value < 0.05; "·" means $0.05 \leq$ p-value < 0.1; otherwise $0.1 \leq$ p-value < 1

In order to determine whether nutrient levels increase with discharge, C-Q (concentration - discharge) plots were created for each site (Figure 8). At sites with high frequency nitrogen data, high frequency nitrogen and discharge values were used. When sites did not have high frequency nitrogen data, daily values were used. A linear model with total nitrogen as the response variable and discharge as the independent variable was created for every site to determine if there was a linear relationship between the two. In 18 out of 22 total sites, there was a significant relationship where as discharge increased, nitrogen also increased ($p < 0.05$). The four sites that showed either a nonsignificant change in discharge or a decrease, were all located in HUCs 1025 and 1026, the farthest HUCs away from the maintstem of the Missouri River.

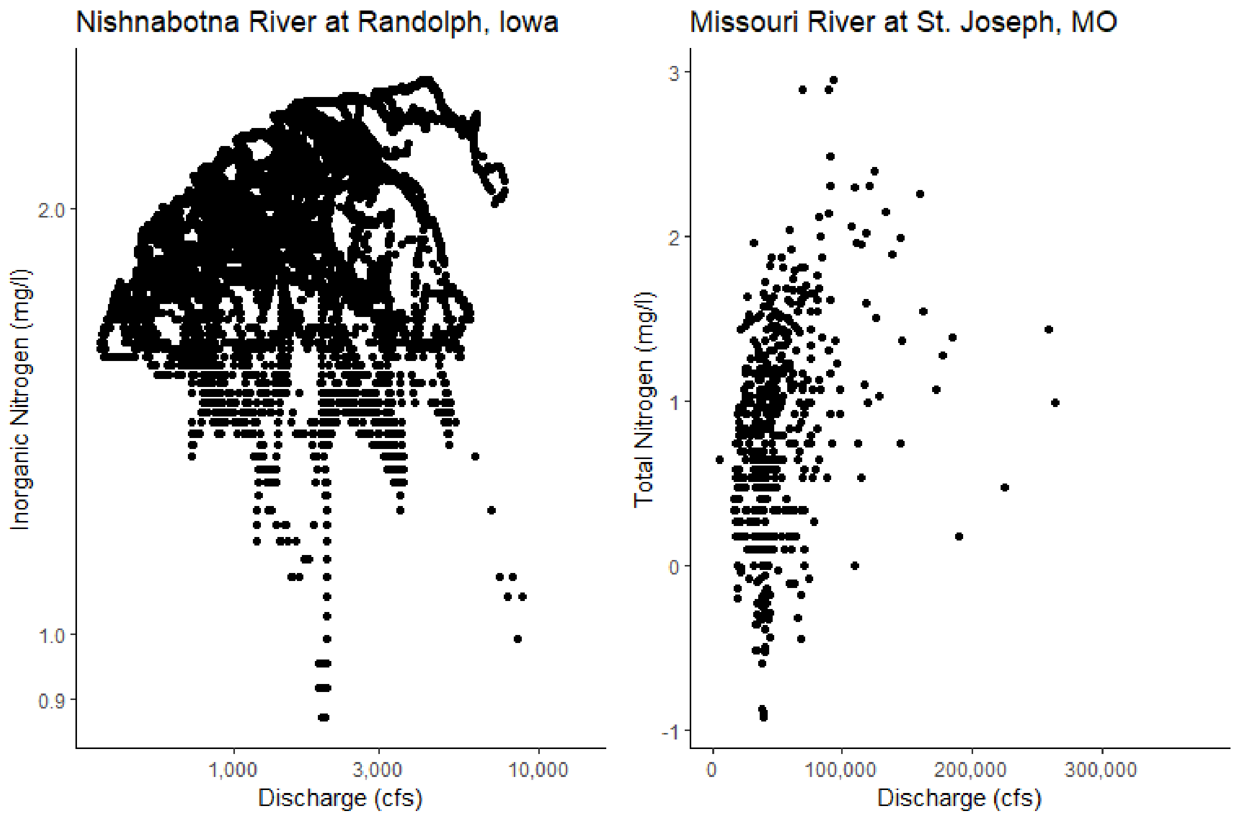

Figure 8: C-Q plots of two sites in the Missouri River Basin. High frequency (left) and daily values (right) shown for two different sites. Nitrogen values are log transformed.

4.2 Question 2: What effects do specific flood and drought events have on the water quality and quantity of rivers in the Missouri River Basin areas of interest?

4.2.1 Flooding

Population of each county (from the 2010 census) in the four states that make up our region area of interest (Kansas, Nebraska, Missouri, and Iowa) were found using the “counties” database from R’s `noncensus` package. We decided that population could be used as a proxy for land cover, as a greater population would indicate more development and fewer agricultural fields or open spaces.

Baseflow and quickflow from each site were determined with the `EcoHydRology` package. After linearly interpolating the instantaneous discharge data in order to account for gaps, total baseflow volume was found and the percent of discharge exported as baseflow was calculated. We predicted that a site within a county with a large population would have a lower percent of its discharge exported as baseflow, because quickflow would be more common in areas with a lot of development. Similarly, we also predicted that a site within a county with a small population would have a greater percentage of its discharge coming from baseflow. More developed areas often have flashier floods, and so we were curious to see whether we can relate population to an element of flashiness - the percentage of discharge exported as quickflow.

Contrary to our hypotheses, greater county population does not contribute to a decrease in percent of discharge as baseflow ($p = 0.4199$, $F = 0.8067$) in our sites of interest. This may be due to our small sample size of sites, or perhaps the size of the rivers in our study.

High frequency nitrogen data was only available for six sites within our region. In order to better understand the behavior of rivers during floods, we examined dygraph plots of discharge and nitrogen, and created hysteresis plots. Storms from the fall of 2018 (September - December) were examined for each river site with data from that time period. We chose to only look at storms that occurred in the second half of the year in order to avoid conflating snowpack melt and precipitation affects.

We predicted that most rivers in the area would behave the same way, and that rivers would exhibit flushing behavior. We thought that flushing rivers would be more prevalent because of the many agricultural fields in our region of study, and any overland flow to the rivers would bring with it a high concentration of nutrients (nitrogen being one of them) from the fertilized fields. Our results say otherwise (Table 5). Our sites of interest had both positive and negative slopes in the hysteresis plots, and also exhibited both clockwise and counter clockwise directions of flow (Figure 9). The same river was analyzed multiple times throughout the year, and even the same river showed different slopes and directions in the hysteresis plots for different storm events.

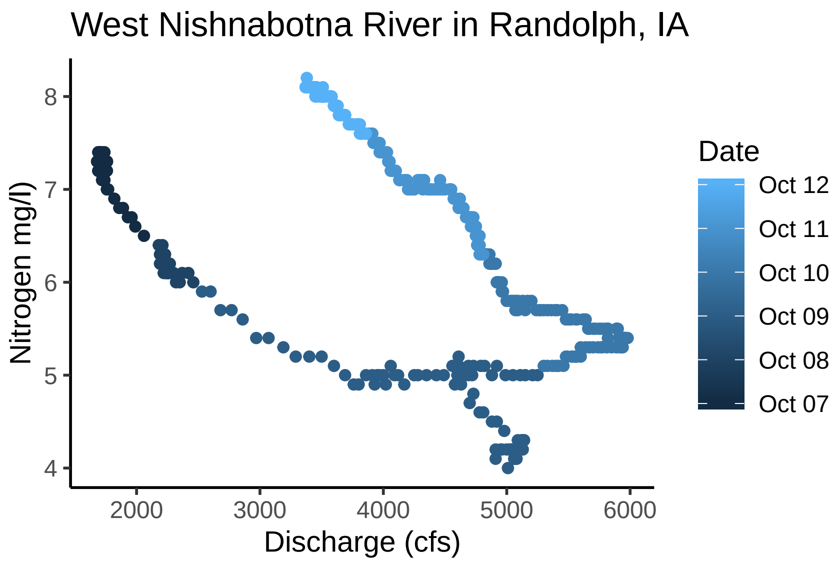

Figure 9: Hyteresis plot for site number 9 in Randolph, IA that shows a negative slope (diluting behavior) in a counter clockwise direction.

Table: 5 Hysteresis results for all sites in area of interest with high frequency nitrogen data.

Site Name	Site Number	Time Period	Direction	Slope
West Nishnabotna River in Randolph, IA	06808500	Oct 7 - 13, 2018	counter clockwise	negative (Figure 9)
Nodaway River at Clarinda, IA	06817000	Oct 8 - 12, 2018	clockwise	negative
Kansas River in Desoto, KS	06892350	Nov 30 - Dec 5, 2018	counter clockwise	positive
Missouri River at Hermann, MO	06934500	Oct 7 - 20, 2018	counter clockwise	negative
Mill C at Johnson Drive, Shawnee, KS	06892513	Nov 27 - Dec 4, 2018	clockwise	negative
Grand River, Sumner MO	06902000	Sep 6 - 10, 2018	clockwise	positive

4.2.2 Droughts

Not only have massive floods characterized years and seasons throughout the Missouri River Basin, but also droughts have had a major impact on the rivers as well. The six sites that were analyzed for floods were also examined for droughts, in order to make direct comparisons and maintain consistency.

Drought has many technical definitions, but we decided on using the 7Q10, which is the minimum 7-day average flow that will occur every 10 years (with a probability of 10%) (United States Environmental Protection Agency 2019). The 7Q10 measurement is used to determine streamflow limits, allocate water resources, dilution rates in watersheds, and more (United States Geological Survey 2019). The 7Q10 is calculated by determining the drought recurrence interval, which is the minimum 7-day average discharge. The recurrence interval can then be used to calculate the probability, which would be the probability or likelihood of having a minimum 7-day average flow of this magnitude or greater. In this case, the minimum 7-day average flow that has a 10% chance of going below that value in a given year.

The recurrence interval is calculated using the following equation:

$$T = (n + 1)/m$$

where T is the recurrence interval, n is the number of years of data, and m is the ranking of the event within the years of data.

Exceedance probability is calculated to determine the probability of having a discharge event of a given size or greater in a year. The equation is simply the inverse of the recurrence interval, or:

$$P = 1/T$$

where P is the exceedance probability, and T is the recurrence interval.

The 7Q10 was calculated for each of the six sites by first filtering the discharge data to only include the selected site. The 'rollmean' function was then used to calculate the average flow every 7 days for said site. Once the 7-day average flow was calculated, the data was grouped by the year so that it could then be summarized to obtain the minimum 7-day average flow. After grouping and summarizing the data, there was one value, the minimum 7-day average flow, for each year of data obtained. From here, the recurrence interval and probability were calculated in each data frame.

The miniumum 7 day average flow (cfs) varied for each river and between years among rivers and in the same river (Figure 10).

The 7Q10 value was then determined from the probability column by filtering the data to show the minimum 7-day average discharge with probability less than or equal to 0.1 (Table 6).

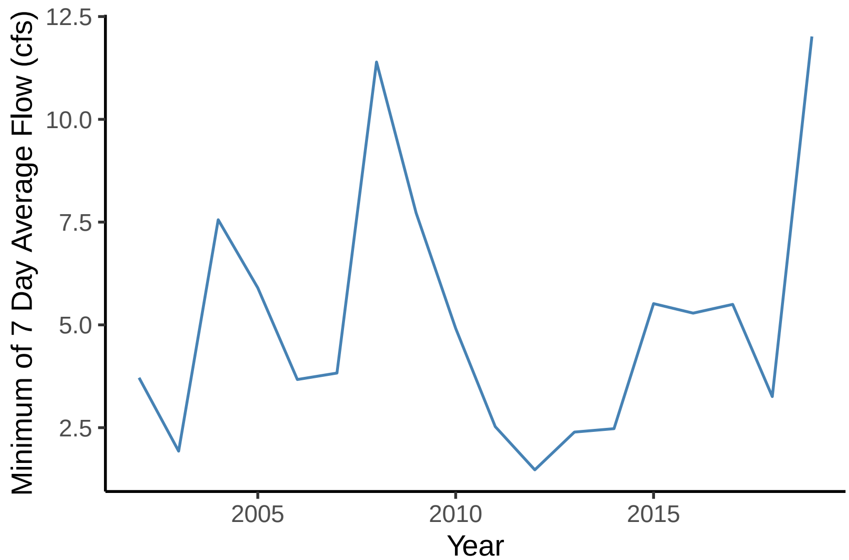

Figure 10: Plot showing the minimum 7 day average flow (cfs) at Mill Creek for the period of record (October 2002 - November 2019). The miniumum 7 day average flow was very low in 2012 and decreased again in 2018.

Table 6: 7Q10 values for the six streams that were analyzed for droughts. The 7Q10 represents the minimum 7-day average flow that has a 10% probability of occuring in a given year. Mill Creek has the lowest 7Q10, whle the Missouri River has the highest 7Q10. This isn't surprising given that the Missouri River is the main river and Mill Creek is a smaller tributary flowing into the Missouri River.

Site Name	Site Number	7Q10 Minimum Discharge (cfs)
West Nishnabotna River in Randolph, IA	06808500	41.3
Nodaway River at Clarinda, IA	06817000	5.93
Kansas River in Desoto, KS	06892350	562
Missouri River at Hermann, MO	06934500	12414
Mill C at Johnson Drive, Shawnee, KS	06892513	1.47
Grand River, Sumner MO	06902000	39.1

In addition to calculating the 7Q10 values for the 6 sites, the sites were analyzed to determine the 30-day moving average flow over a period of years. This was done using code from USGS to calculate moving averages and historical quantiles (DeCicco 2016). The 25th, 50th, and 75th quantiles were calculated to determine the minimum, median, and maximum values for discharge in a given year to determine historical flow. The 30-day moving average flow over a period of years was then plotted with Normal (25th to 75th percentile), Drought Watch (10th to 25th percentile), Drought Warning (5th to 10th percentile), and Drought Emergency (0 to 5th percentile) (Figure 11). This information about what flow would constitute a drought is extremely useful for water managers.

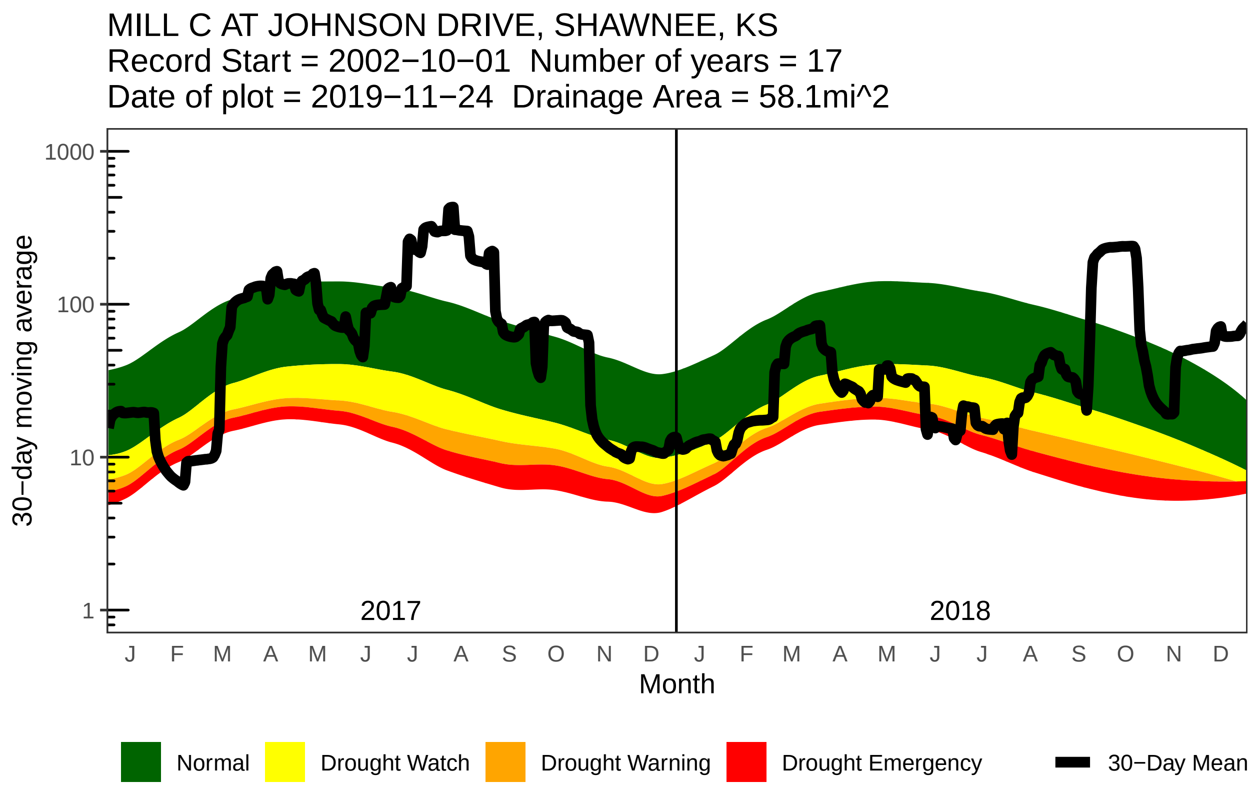

Figure 11: Plot showing the 30-day moving average of Mill Creek in 2017 and 2018 and whether it falls within a normal condition, drought watach condition, drought warning condition, or drought emergency condition. In March and April of 2017 and June, July, and August of 2018, Mill Creek was in a drought warning or drought emergency.

4.3 Question 3: What factors contribute to the variability of total nitrogen in the rivers?

In order to better understand what contributes to the variability of nitrogen concentration in our study area of interest, we created a linear model with fixed effects and random effects. Fixed effects that were considered were year, population, phosphorus, and discharge. By considering the HUC 4 region each site is in as a random effect, we can estimate the amount of variation of nitrogen between each HUC 4. In order to better model the interaction, we chose to divide the year variable by 10 and the population by 1000 so that rescaling is not an issue. The model chosen is below:

```
lmer(data, log(total.nitrogen) ~ year + population  + phosphorus +
(1|huc4))
```

Year, population, and total phosphorus are fixed effects while the grouping of the HUC 4 region is a random effect. This means we take the HUC 4 region variance into account when modeling the effect of year and population on total nitrogen. Using log transformed nitrogen data, we find that there is a significant difference in total nitrogen across each decade and every 1000 people ($p < 2e-16$) and with changing phosphorus amounts. The regression equation is below:

$$Nitrogen = 0.28 + 1.01(year) + 0.99(population) + 1.44(phosphorus)$$

Because we used a log transformed data set, we have to exponentiate the coefficients in the model to interpret them. The nonsensical intercept of the regression equation is 0.28, which is the mean of total nitrogen when year is 0, population is 0, and phosphorus is 0. When exponentiating the year coefficient, we can conclude that every decade has a 1.01 multiplicative effect on nitrogen. When exponentiating the population coefficient, we can conclude that every 1000 people have a 0.99 multiplicative effect on nitrogen. Lastly, we can conclude that every unit increase in phosphorus has a 1.44 multiplicative effect (or 44% increase) of nitrogen. The model demonstrates the the level of decade, population of the county, and phosphorus levels affected the total nitrogen in each HUC 4 region. Residuals are evenly distributed, and the Rˆ2 value = 0.567, indicating that 56.7% of the variation in nitrogen is explained by this model.

4.4 Question 4: Given past and current data, what can we predict about the future state of water in the Missouri River Basin?

4.4.1 Discharge Trends Analysis

We also analyzed discharge trend at 22 sites using daily value data and Mann-Kandall Trend Test (Table 7). In our analysis, discharge of 18 sites increases while 4 sites show the decreasing trend of discharge. However, these four sites all locate in HUC 1025 and 1026, away from the main stem of the Missouri River. Therefore, overall, the discharge in Missouri River is increasing.

Table 7: Discharge Trend over Time

No.	Site Number	Discharge Trend	Discharge Trend Significance
1	06768000	Increase	
2	06805500	Increase	***
3	06775900	Increase	***
4	06794000	Increase	
5	06800000	Increase	***
6	06800500	Increase	***
7	06610000	Increase	***
8	06609500	Increase	***
9	06810000	Increase	***
10	06818000	Increase	***
11	06844500	Decrease	***
12	06856600	Decrease	**
13	06877600	Decrease	
14	06874000	Decrease	·
15	06887500	Increase	
16	06892350	Increase	·
17	06902000	Increase	*
18	06905500	Increase	*
19	06921070	Increase	
20	06926510	Increase	
21	06894000	Increase	**
22	06934500	Increase	***

Note: *** means p-value < 0.001; ** means $0.001 \leq$ p-value < 0.01;* $0.01 \leq$ p-value < 0.05; · means $0.05 \leq$ p-value < 0.1; otherwise $0.1 \leq$ p-value < 1

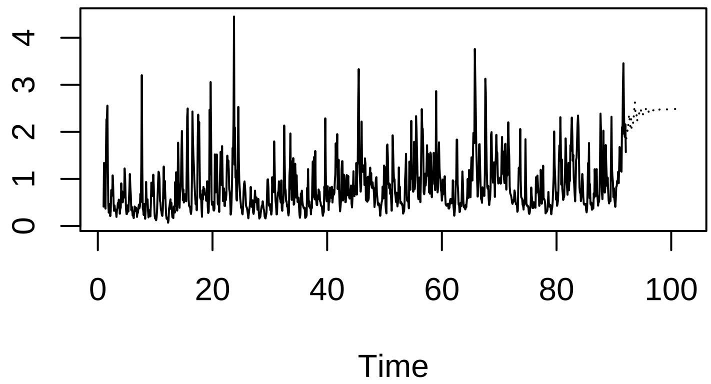

Figure 12: Historic and 10-year prediction of discharge at Site No. 22, the Missouri River At Hermann, MO. (Solid Line: Historic Measurement of Discharge; Dash Line: 10-year Prediction of Discharge)

4.4.2 Discharge Prediction

As discussed in Question 2, both severe drought and flooding events occurred in the basin in the past, which caused fatal damage to local agriculture. Therefore, it is important to make accurate forecasts of discharge, providing data support to policymakers.

We employed Autoregressive and Moving Average Models (ARMA) to effectively forecast into the future. The ARMA model is a tool for understanding and predicting future values in a time series. The AR part involves regressing the variable on its own lagged (i.e., past) values. The MA part involves modeling the error term as a linear combination of error terms occurring contemporaneously and at various times in the past. By using the discharge measurement data in the past and the ARMA model, we make discharge prediction for every site (Figure 12 as an example). From Figure 12, we can see that only the first few years of discharge simulation is reasonable. Therefore, we decided only to make one-year prediction of discharge for every site (Table 8).

Table: 8 One-year Discharge Prediction for Every Site

No.	Site Number	Dec. 2019	Jan. 2020	Feb. 2020	Mar. 2020	Apr. 2020	May 2020
1	06768000	2019.992	1942.129	1929.816	2040.176	2032.186	1990.873
2	06805500	14060.7	12802.25	12845.87	23729.49	15056.78	16585.16
3	06775900	239.2089	238.9801	238.3116	239.5916	240.5232	242.5456
4	06794000	148.0349	129.4563	129.6965	327.3795	165.0577	190.15
5	06800000	173.1581	150.6924	141.7451	313.545	159.0779	199.5598
6	06800500	2607.337	2598.611	2672.196	3208.066	3341.665	3977.717
7	06610000	70976.66	60591.76	55874.04	73723.59	81113.53	83595.71
8	06609500	708.6843	691.4817	645.5966	1176.8041	714.8702	814.9989
9	06810000	2422.32	2391.721	2377.854	3204.073	2417.139	2731.861
10	06818000	90011.3	77760.16	70549.03	116938.26	110927.07	130362.46
11	06844500	154.9544	145.2851	147.9408	228.2876	171.8784	166.0301

No.	Site Number	Dec. 2019	Jan. 2020	Feb. 2020	Mar. 2020	Apr. 2020	May 2020
12	06856600	2212.686	2012.946	1966.35	2354.776	2052.399	2615.569
13	06877600	1999.598	1814.097	1688.58	1807.765	1634.919	2308.793
14	06874000	174.3205	140.6582	123.4546	114.6624	110.1691	107.8727
15	06887500	14398.77	10484.501	7784.95	6938.432	5800.966	7571.298
16	06892350	17433.744	13064.978	10272.795	9334.909	7730.841	11531.101
17	06902000	9634.729	9456.489	10440.135	10753.103	9392.489	12326.771
18	06905500	1946.371	2523.677	2663.863	2701.267	2556.851	3142.349
19	06921070	302.1977	290.1968	335.3464	317.8665	267.6631	404.104
20	06926510	12184.39	13730.31	13039.62	12564.56	14026.34	14174.23
21	06894000	181.6666	206.4976	203.7249	206.7256	214.5459	320.4555
22	06934500	189461.2	194893.2	194212.8	215398.5	212735.9	230589.3

No.	Site Number	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020	Nov. 2020
1	06768000	2023.824	2008.55	1928.585	1996.895	1979.38	2026.699
2	06805500	18430.86	15168.09	15216.99	14719.46	13635.78	12780.16
3	06775900	242.0369	241.6409	242.0486	242.2389	241.9548	240.2764
4	06794000	189.022	172.7318	157.0474	146.7917	151.0374	136.0562
5	06800000	187.1167	149.6114	148.6135	149.6721	146.2639	136.3859
6	06800500	3785.948	2912.286	2434.121	3091.355	2733.932	2520.393
7	06610000	90800.91	85219.7	78593.88	80467.92	78227.7	70854.33
8	06609500	800.516	733.6069	650.3337	655.3034	724.2999	689.9273
9	06810000	3004.86	2650.792	2460.292	2777.403	3156.369	2595.649
10	06818000	131671.16	110560.74	99105.45	103960.49	121688.94	96803.96
11	06844500	184.0507	321.3987	164.3744	150.4602	149.7053	158.701
12	06856600	2571.786	2483.559	2404.832	2283.906	2196.924	2058.036
13	06877600	2034.465	1885.749	1831.551	1777.794	1663.053	1556.348
14	06874000	106.6991	106.0993	105.7928	105.6361	105.556	105.5151
15	06887500	8326.411	9052.989	6796.397	6676.547	6316.088	7646.2
16	06892350	11710.275	12100.027	10099.463	9499.502	8509.621	9945.916
17	06902000	11756.818	9390.933	9517.493	10274.174	12340.875	9446.777
18	06905500	2904.042	2597.213	2511.393	2678.037	2898.101	2629.494
19	06921070	262.339	255.1629	247.3837	250.4061	269.6108	277.8386
20	06926510	16191.64	16801.49	15671.6	14122.26	13155.48	12236.42
21	06894000	220.6817	228.4371	209.8449	231.759	198.792	190.647
22	06934500	240828.8	227848.5	213285.9	216789.6	230480	210515

5 Summary and Conclusions

In this project, we aimed to take a close look at discharge and nutrient data at 22 sites to see if we could spot any trends and better understand the hydrologic properties of the Missouri River Basin.

5.1 Discharge and Nutrient Variation

Our research question #1 focused on changes in nutrient enrichment and changes in discharge and whether those changes affected their interactions. Most rivers and streams in the Missouri River Basin have higher discharge as well as larger variation during the summer than the winter, confirming hypothesis 1a. A few rivers that originate from aquifer or groundwater display more stable discharge pattern throughout a year.

As for variation in discharge across different years, conclusion varies based on two statistics, standard deviation (SD) and coefficient of variation (CV). Absolute magnitude of variation generally increases in the whole basin, while relative variation has diminished at almost all sites examined here. The disparity between the conclusions drawn from two variation statistics stems primary from higher mean discharge in more recent years. CV accounts for mean discharge and is therefore more accurate when comparing rivers with different discharge or comparing rivers across different time periods. However, information provided by SD is also informative and should not be disregarded, as damage caused by flooding is better explained by absolute magnitude of variation than relative variation. Both statistics should be considered when formulating conclusions on variability in discharge, and the emphasis on either one depends on the context where interpretation is made (e.g. for academic management purposes) as well as the audience to which information is offered (e.g. researchers or managers).

Our hypothesis, 1b, that nutrient levels have increased over time, was not proven. There was so much variability across sites that any prediction cannot hold true across the entire lower basin. A little more conclusive evidence was found for our hypothesis 1c, that nutrient levels increase discharge, since 18 of the total 22 sites were found to have significant increases.

5.2 Floods and Droughts

Analyzing nutrient content in flooding events was a lot harder than originally anticipated. We wanted to analyze flooding across all sites, but were unable to do so because of a lack of high frequency nitrogen data. In addition, the data we did have was inconsistent and unclear. Our hypothesis 2b, that rivers will exhibit flushing events, was not correct. It was wrong to assume that all rivers in the area acted the same way during floods, simply because this basin is mostly agricultural land. The sites analyzed for floods exhibited a variety of behaviors depending on the storm and the time of year of the storm.

From the drought analysis of the six tributaries and rivers, it is evident that minimum discharge amounts vary greatly (range: 1.47 cfs – 12414) (Table 6). Hypothesis 2b predicted that discharge would decrease during periods of drought, and it does, but the 7-day minimum average flow clearly varies among sites. Although the analysis shows that droughts decrease flow, it is difficult to compare droughts and flow between the sites because the rivers are of different sizes and magnitudes, and thus discharge amounts vary greatly. For example, three sites were analyzed that are located on the Missouri River, which is much larger than some of the tributaries that were analyzed. Thus, the discharge patterns are not able to be compared because of the difference in sizes. If we were to do this analysis again, we would likely choose not located on the Missouri River and compare them, or choose only sites that are on the Missouri River to have a better comparison.

5.3 Linear model predicting nitrogen

The linear model we created suggested that nitrogen variability could be explained by the variability of year, population, and phosphorus. Our hypothesis 3a was mostly correct, however the model failed to include discharge as a significant factor. We think this is odd, and are not sure why this is the case, since in hypothesis 1c, we found significant increases in nitrogen with increases in discharge. We speculate that perhaps the model, which took out the variance of the HUC 4 region as a random effect, did not account for the variability across regions, when in fact it should have. If time allowed, further research into this question would have been helpful.

5.4 Discharge Trends Analysis and Discharge Prediction

From the discharge trend analysis, we found that out of all 22 sites, only 4 sites that are located in HUC 1025 and 1026, away from the main stem of the Missouri River show decreasing trends of discharge. Overall, the discharge in Missouri River is increasing, confirming our hypothesis 4a.

In addition, we employed Autoregressive and Moving Average Models (ARMA) to forecast the discharge at all 22 sites into the future. Due to that only the first few years of discharge simulation is reasonable from the ARMA, we only made one-year prediction of discharge for every site. Our simulation confirms our hypothesis 4b, that the future will see a continuation of current trends.

5.5 Trends of the Concentration of Nitrogen and Phosphorus

Although we saw varying results from our analysis of long-term daily nitrogen and phosphorus data across all sites, we do see a significant increasing trend of nitrogen and phosphorus at the mainstem of the Missouri River. Corresponding with the linear model we created for question 3, we can also see that the concentration of nitrogen will increase with the concentration of phosphorus and year, which also indicates that nitrogen concentration will continue to trend upward over time.

In conclusion, although it was difficult to make sweeping conclusions across all our sites due to the size of the river and location variability, we were able to find out a lot about the Missouri River Basin and its water quality and quantity. We determined that the mainstem of the Missouri River did have increasing discharge, nitrogen, and phosphorus levels over time, while the tributaries had more varied characteristics.

References

Bureau of Reclamation. 2016a. “Republican River Basin Study.” Government Document. https://www.usbr.gov/watersmart/bsp/docs/finalreport/republican/republican-river-basin-study-final-report.pdf.

———. 2016b. “SECURE Water Act Section 9503(c) Report to Congress.” Report. https://www.usbr.gov/climate/secure/docs/2016secure/2016SECUREReport-chapter6.pdf.

Coe, Michael T., and Charon M. Birkett. 2004. “Calculation of River Discharge and Prediction of Lake Height from Satellite Radar Altimetry: Example for the Lake Chad Basin.” Journal Article 40 (10): n/a–n/a. https://doi.org/10.1029/2003wr002543.

Committee on Missouri River Ecosystem Science. 2002. *The Missouri River Ecosystem: Exploring the Prospects for Recovery.* Book. Washington, DC: The National Academies Press. https://doi.org/doi:10.17226/10277.

DeCicco, L. 2016. “Calculating Moving Averages and Historical Flow Quantiles.” Website. https://owi.usgs.gov/blog/moving-averages/.

Déry, Stephen J. 2005. “Decreasing River Discharge in Northern Canada.” Journal Article. *Geophysical Research Letters* 32 (10). https://doi.org/10.1029/2005gl022845.

Jowett, Ian G., and Maurice J. Duncan. 1990. “Flow Variability in New Zealand Rivers and Its Relationship to in-stream Habitat and Biota.” Journal Article 24 (3): 305–17. https://doi.org/10.1080/00288330.1990.9516427.

Kammerer, J.C. 1990. “Largest Rivers in the United States.” Report. U.S. Geological Survey, Department of the Interior.

National Oceanic and Atmospheric Administration. 2012. “Service Assessment the Missouri/Souris River Floods of May – August 2011.” Government Document.

NOAA Drought Task Force. 2013. “An Interpretation of the Origins of the 2012 Central Great Plains Drought.” Government Document. https://cpo.noaa.gov/Meet-the-Divisions/Earth-System-Science-and-Modeling/MAPP/MAPP-Task-Forces/Drought/Drought-Task-Force-II/An-Interpretation-of-the-Origins-of-the-2012-Central-Great-Plains-Drough

United States Environmental Protection Agency. 2019. “Definitions and Characteristics of Low Flows.” Website. https://www.epa.gov/ceam/definition-and-characteristics-low-flows#1Q10.

United States Geological Survey. 2019. “Low-Flow (7Q10) Analysis.” Website. https://www.sciencebase.gov/catalog/item/5ad8a1c3e4b0e2c2dd279fc5.

U.S. Army Corps of Engineers Missouri River Basin Water Management Division. 2018. “Master Water Control Manual Missouri River Basin.” Government Document.