

Examining the Hydrologic Properties of the Missouri River Basin

https://github.com/cwatson1013/Hydrologic_Data_Analysis_Final_Proj

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

Abstract

The Missouri River provides critical water resources that drives the region's agriculture, industry, and ecosystems. This is a region that experiences surface water variability, characterized by damaging floods and severe droughts, greatly impacting the agricultural production of the area. This project highlights the changes in streamflow and water quality over time, and identifies key characteristics of the river. . . . Twenty six sites across the lower Missouri River Basin were examined in order to get a fuller picture of the Missouri River and its tributaries over time.

<Information in these brackets are used for annotating the RMarkdown file. They will not appear in the final version of the PDF document>

Contents

1	Research Question and Rationale	5
2	Dataset Information	6
3	Exploratory Data Analysis and Wrangling	9
3.0.1	High Frequency Nitrogen and Discharge	10
4	Analysis	12
5	Summary and Conclusions	13
5.1	Example for autoreferencing	13

List of Tables

1	Summary of Daily Value Data at 22 sites in the Missouri River Basin	7
2	Summary of Water Quality Data in the Missouri River Basin	8

List of Figures

1	Hysteresis plots	11
2	Absorbance frequency	13

<Note: set up autoreferencing for figures and tables in your document>

1 Research Question and Rationale

<Paragraph detailing the rationale for your analysis. What is the significant application and/or interest in this topic? Connect to environmental topic(s)/challenge(s).>

<Paragraph detailing your research question(s) and goals. What do you want to find out? Include a sentence (or a few) on the dataset you are using to answer this question - just enough to give your reader an idea of where you are going with the analysis.>

2 Dataset Information

<Information on how the dataset for this analysis were collected, the data contained in the dataset, and any important pieces of information that are relevant to your analyses. This section should contain much of same information as the README file for the dataset but formatted in a way that is more narrative.>

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 lower Missouri River basin, we pulled sites from each HUC4 subbasin from 1020 to 1030 (see Figure below). We chose these subbasins because they had a variety of tributaries that all flowed into the Missouri River, and we wanted a variety of river sizes and lengths. We filtered these subbasin queries to only show us sites that had discharge, nitrogen, and phosphorus data. Once we found the sites with all of this data, we chose 2 sites from each HUC sub basin as our 22 “best sites”. Our best sites had the overall best time period range for all of our “must have” variables.

Only seven sites within our HUC subbasin boundary contained any high frequency discharge and nitrogen data. Therefore, we also looked at these 7 sites in order to do analyses and answer our research question about flooding.

After doing initial data wrangling and analysis on our 22 “best sites”, we decided to pare it down further and only do subsequent analyses on **10** sites. While we initially wanted to look at many sites that were varied in size and location, we determined that it was too many to look at and draw relevant conclusions from.

We have three main datasets:

- The daily values dataset with our 22 “best sites”
- The water quality dataset with our 22 “best sites”, with only six sites that had total coliform data.
- The high frequency dataset with 7 sites that contain both high frequency discharge and high frequency nitrogen data.

<Add a table that summarizes your data structure. This table can be made in markdown text or inserted as a `kable` function in an R chunk. If the latter, do not include the code used to generate your table.>

<C will do data table for water quality and daily values, R will do for high freq>

Table 1: Summary of Daily Value Data at 22 sites in the Missouri River Basin

[illegible]

Table 2: Summary of Water Quality Data in the Missouri River Basin

[illegible]

3 Exploratory Data Analysis and Wrangling

<Include R chunks for 5+ lines of summary code (display code and output), 3+ exploratory graphs (display graphs only), and any wrangling you do to your dataset(s).>

<Include text sections to accompany these R chunks to explain the reasoning behind your workflow, and the rationale for your approach.>

```
#high frequency data wrangling
```

```
highfreqsite2019 <- highfreqsiteinfo %>%  
  filter(end_date > "2019-03-31"); head(highfreqsite2019)
```

```
## Warning in Ops.factor(end_date, "2019-03-31"): '>' not meaningful for  
## factors
```

```
## [1] X                agency_cd          site_no  
## [4] station_nm         site_tp_cd        dec_lat_va  
## [7] dec_long_va        coord_acy_cd      dec_coord_datum_cd  
## [10] alt_va             alt_acy_va        alt_datum_cd  
## [13] huc_cd             data_type_cd      parm_cd  
## [16] stat_cd            ts_id             loc_web_ds  
## [19] medium_grp_cd      parm_grp_cd       srs_id  
## [22] access_cd          begin_date        end_date  
## [25] count_nu  
## <0 rows> (or 0-length row.names)
```

```
highfreqsites.DN <- readNWISuv(site = c("06808500", "06817000", "06892350", "06934500"),  
                                parameterCd = c("00060", "99133"),  
                                # Discharge in cfs & Nitrate in mg/l NO3-N  
                                startDate = "2019-01-01",  
                                endDate = "2019-11-01") %>%  
  renameNWISColumns() %>%  
  rename(Nitrate_mgl = 6)
```

```
#individual sites
```

```
Hermann <- highfreqsites.DN %>%  
  filter(site_no=="06934500")  
Desoto <- highfreqsites.DN %>%  
  filter(site_no=="06892350")  
Clarinda <- highfreqsites.DN %>%  
  filter(site_no=="06817000")  
Randolph <- highfreqsites.DN %>%  
  filter(site_no=="06808500")
```

3.0.1 High Frequency Nitrogen and Discharge

There were 7 sites in our region of interest that had high freq N data, and only 4 sites had high freq N data during the floods of 2019. The sites looked at in depth are:

- West Nishnabotna River in Randolph, IA
- Nodaway River at Clarinda, IA
- Kansas River in Desoto, KS
- Missouri River at Hermann, MO

The Missouri River is the biggest river, with an average of 214693 cfs discharge rate during the year 2019, and the Nodaway River is the smallest river, with an average of 1185 cfs discharge rate for 2019.

In March of 2019, a bomb cyclone hit the midwest. Our initial research question, what effect did the March 2019 storm have on water quality, attempted to look into the behavior of nitrogen in the discharge of the rivers. Unfortunately, instantaneous Nitrogen values stopped recording during the peak of the storm events in March, so it was hard to create hysteresis plots that exhibited the type of storm and its effects on nitrogen concentration.

Even though Nitrogen concentrations were not recorded in March, they were recorded in other times of the year. 2019 was a wet year and many large storm events occurred.

Warning: Removed 9 rows containing missing values (geom_point).

The Figure 1 shows Hysteresis plots for two storm events in the Missouri River Basin. The storm event on the West Nishnabotna River exhibits an oddly-shaped plot that has a negative slope, indicating it is a diluting storm. The Kansas River experienced a storm in late February that has a counter-clockwise motion and a positive slope, indicating a flushing storm. These two plots illustrate that two rivers near each other can have very different behaviors.

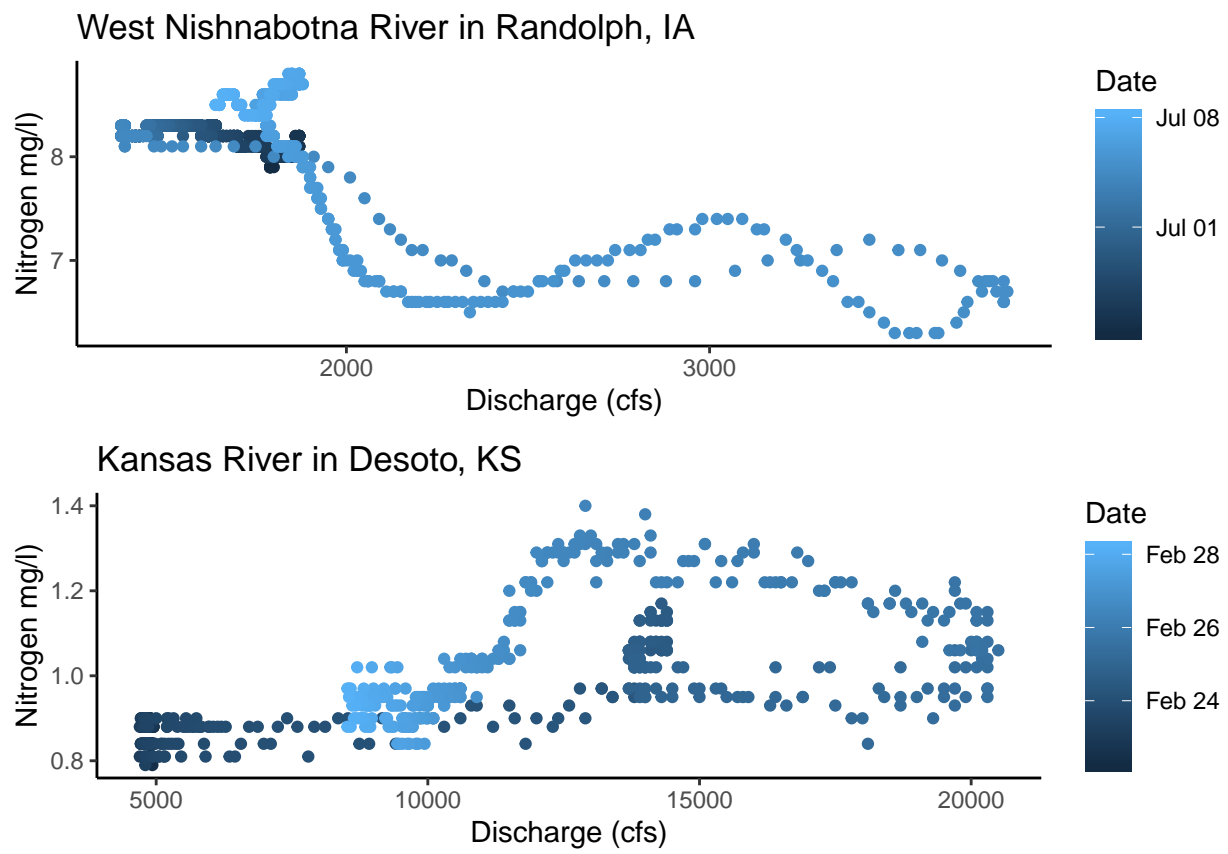


Figure 1: Hysteresis plots

4 Analysis

<Include R chunks for 3+ statistical tests (display code and output) and 3+ final visualization graphs (display graphs only).>

<Include text sections to accompany these R chunks to explain the reasoning behind your workflow, rationale for your approach, and the justification of meeting or failing to meet assumptions of tests.>

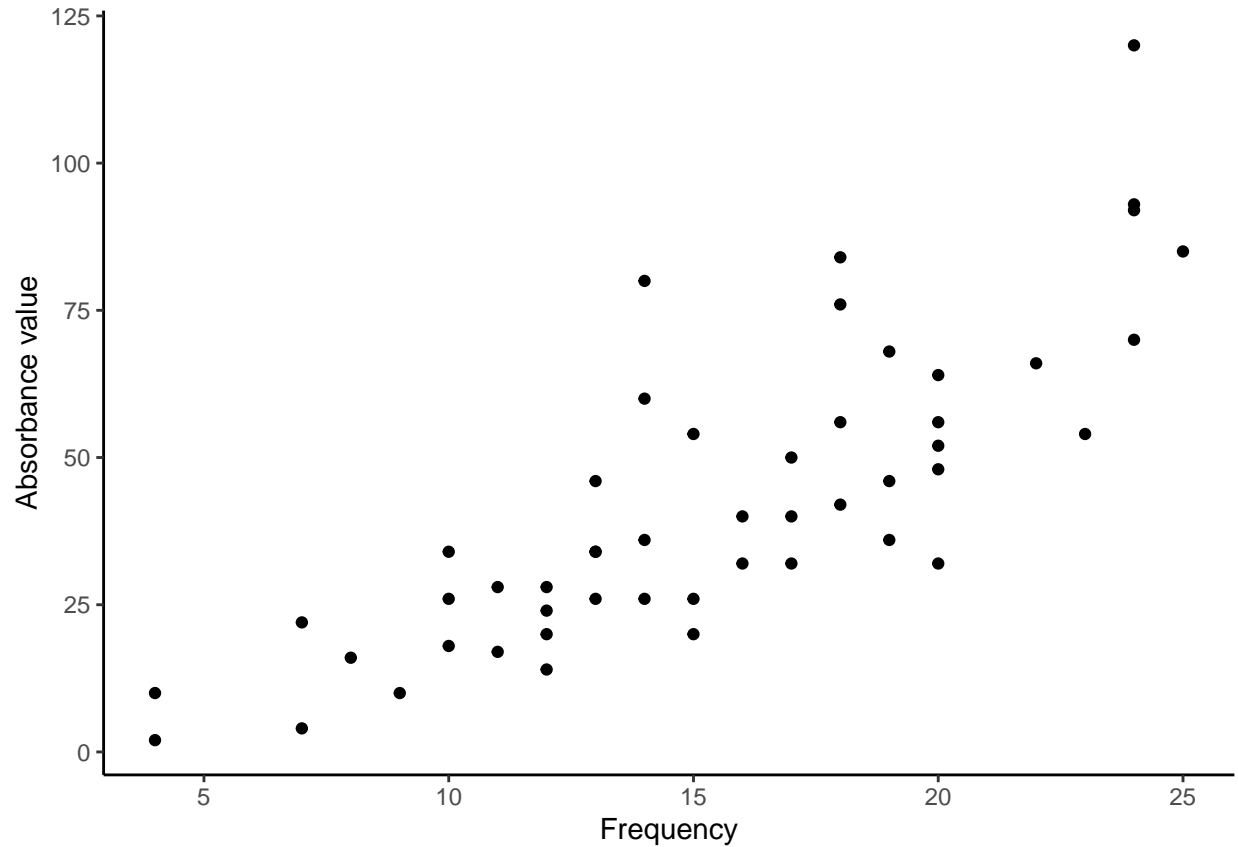


Figure 2: Absorbance frequency

5 Summary and Conclusions

<Summarize your major findings from your analyses. What conclusions do you draw from your findings? Make sure to apply this to a broader application for the research question you have answered.>

5.1 Example for autoreferencing

As seen by Figure 2, Absorbance values are not normally distributed. This is expected, as we are dealing with ecological data.