# Dissolved Organic Carbon Trends in Small U.S. Watersheds

Jonathan Gilman

December 12th, 2024

## Contents

# Contents

# List of Tables

## List of Figures

# 1 Rationale and Research Questions

Dissolved organic carbon (DOC) plays a critical role in biogeochemical cycles, influencing water quality, carbon storage, and ecosystem dynamics. Seasonal DOC export reflects interactions between climate, hydrology, and vegetation, such as leaf litter inputs, snowmelt, and rainfall. Additionally, long-term trends in DOC concentrations can indicate environmental changes such as disturbances or climate impacts.

The MacroSheds dataset provides a comprehensive view of DOC dynamics across US watersheds, capturing discharge and DOC measurements at high temporal resolution. This study leverages this dataset to explore: 1. Seasonal Differences in DOC Export: Is there a seasonal difference in DOC export, and how does this vary across the US? What factors might contribute to these patterns, such as snowmelt or vegetation inputs? 2. Long-term DOC Trends and Disturbance Effects: Are DOC exports significantly changing over time? How do experimental (disturbed) watersheds differ from non-experimental watersheds in terms of DOC export trends?

## 2 Dataset Information

The MacroSheds project consolidates all U.S. federally funded watershed ecosystem studies into a unified, open-data platform. It enables researchers to identify differences in watershed functional traits, particularly in the context of a changing climate. Unlike other datasets, MacroSheds emphasizes smaller watersheds that align with the watershed ecosystem concept. This focus minimizes the influence of confounding factors such as complex hydrology and land use, which can obscure key biogeochemical trends and processes.

For this analysis, we utilized MacroSheds data to investigate dissolved organic carbon (DOC) dynamics across the continental United States. The dataset provides high-resolution temporal measurements of stream chemistry and discharge across multiple domains, offering insights into DOC export patterns over time and space. Specifically, DOC concentrations (measured in mg/L) and stream discharge rates (measured in L/s) were analyzed to explore seasonal differences and long-term trends in DOC export.

A critical component of this analysis is the calculation of Volume-Weighted Mean (VWM) DOC concentrations. VWM is calculated as the sum of the product of DOC concentrations and discharge rates divided by the total discharge over a given period. This method integrates DOC and discharge, providing a more representative measure of export than raw concentration values alone. VWM calculations are particularly valuable because they account for variations in flow, which strongly influence DOC transport. For instance, during high-flow events, DOC concentrations might remain steady or even dilute, but total export could increase significantly due to the large volume of water. By incorporating discharge, VWM provides a comprehensive metric for understanding DOC export dynamics and their underlying drivers.

The dataset was carefully processed to ensure quality and relevance. Chemistry and discharge datasets were merged by site and date, and interpolated values were excluded to maintain data accuracy. Only domains located within the continental United States were retained, allowing for consistent spatial analysis. Additional variables, such as hydrological water year (October to September) and season (Fall, Winter, Spring, Summer), were derived to facilitate temporal analyses aligned with hydrological and ecological processes.

By emphasizing smaller, relatively undisturbed watersheds and incorporating advanced calculations like VWM, MacroSheds provides a unique opportunity to study watershed functional traits in response to environmental changes. Its structured format ensures robust, scalable analyses that can inform our understanding of DOC dynamics and their broader implications for biogeochemical cycling in freshwater systems.

# 3 Exploratory Analysis

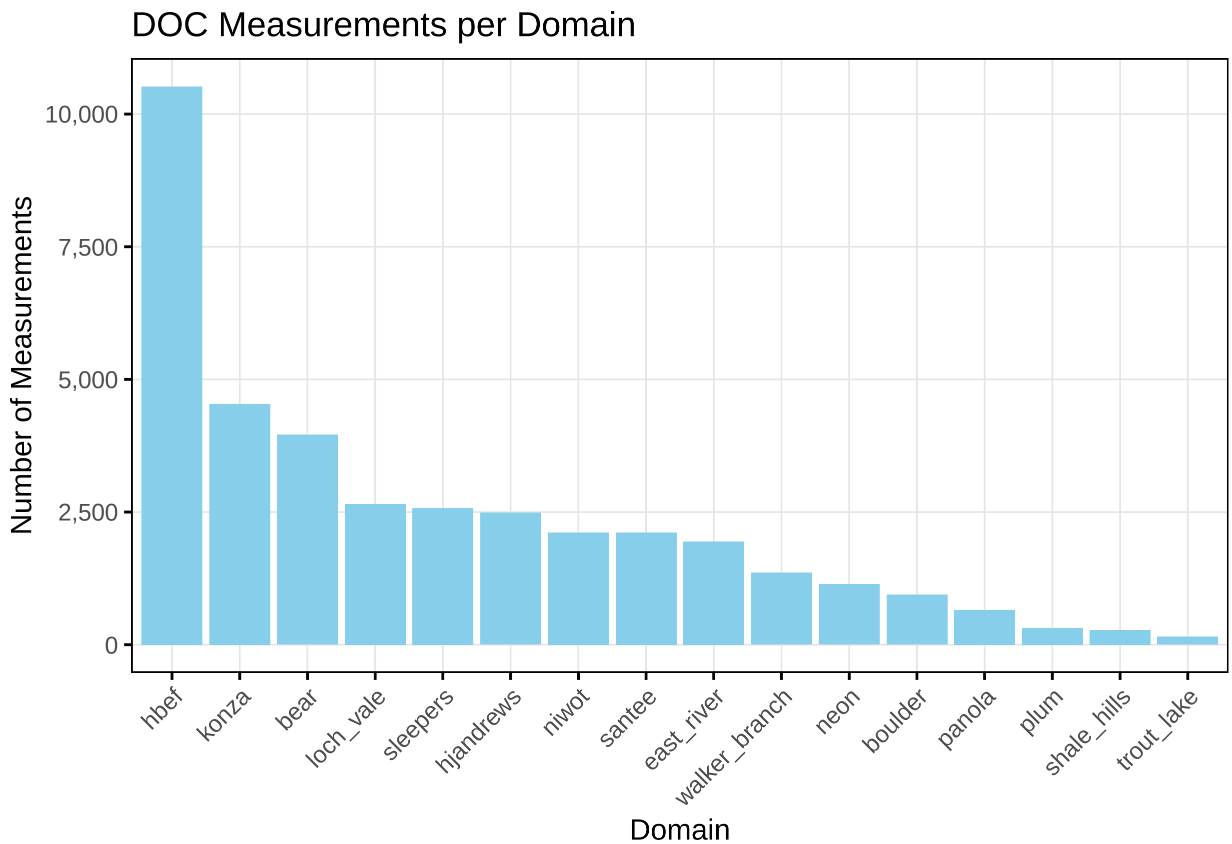


Figure 1: Total DOC observations across domains

Figure 1 summarizes the total number of dissolved organic carbon (DOC) measurements collected for each domain. HBEF has the highest count of measurements, exceeding 10,000, while domains like Trout Lake and Shale Hills have significantly fewer data points. This variation in data density reflects differences in monitoring effort and study focus across sites. Domains with more measurements provide robust temporal coverage, enabling more reliable trend and seasonal analyses, whereas domains with fewer observations may have limited statistical power.

Figure 2 illustrates the number of years of data collection for DOC measurements across different domains. Loch Vale has the longest duration of data collection, spanning nearly 40 years, followed by Panola and HBEF. In contrast, domains such as Neon and Shale Hills have much shorter durations. The length of data collection is critical for analyzing long-term trends, with longer time series offering greater insights into temporal patterns and environmental changes.

Figure 3 shows the temporal extent of data collection for individual sites within each domain. Each horizontal line represents a site, with its start and end points indicating the first and last years of data collection. The annotation at the bottom notes that there are 89 unique sites included in the dataset. The variability in temporal coverage among sites highlights differences in monitoring intensity and duration, which can influence the resolution and scope of trend analyses.

Figure 4 displays the geographic locations of domains within the continental United States. Each red point represents a domain, with labels identifying their names. The spatial distribution of domains spans diverse climates and landscapes, from northern forested regions like HBEF to southern domains such as Santee

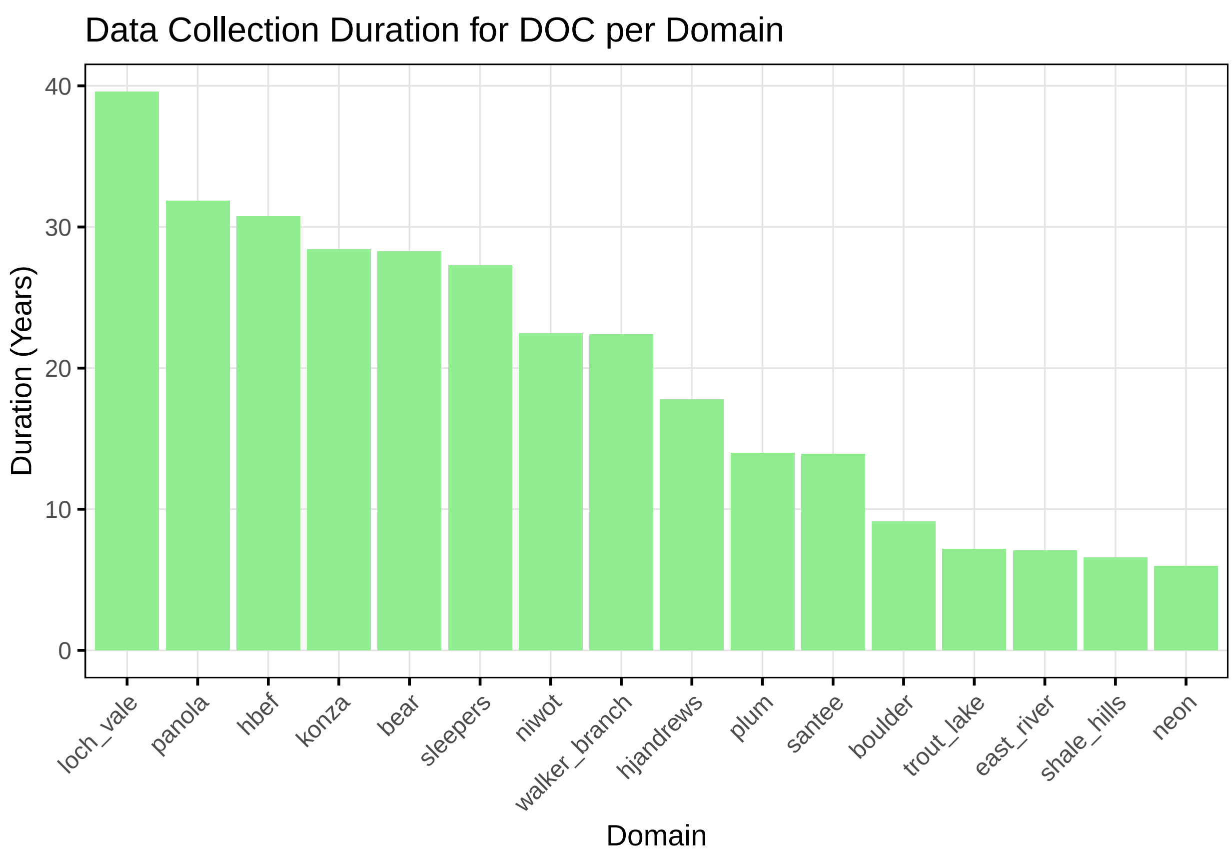


Figure 2: Duration of data collection (in years) across domains.

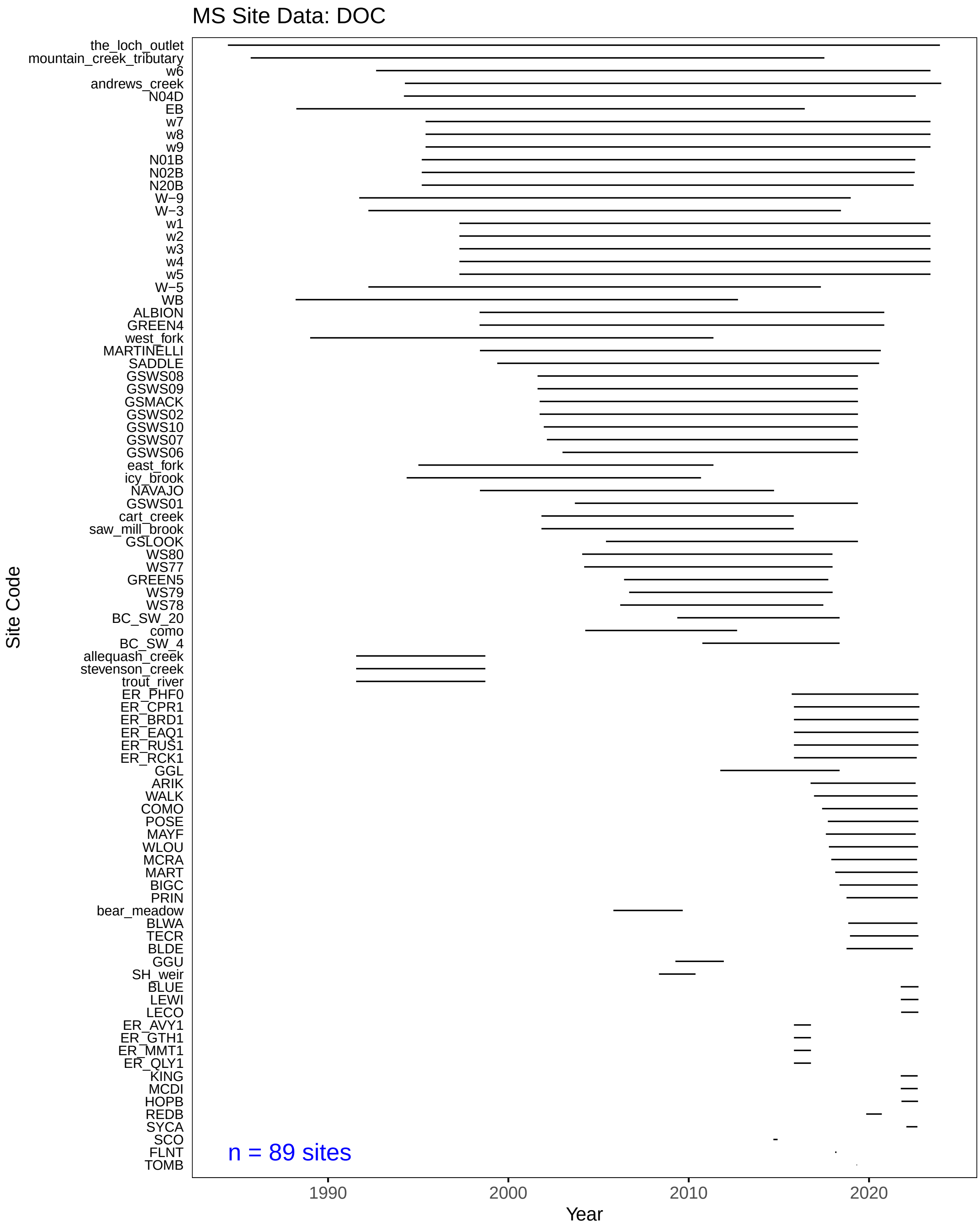


Figure 3: Duration of DOC data collection for each site.

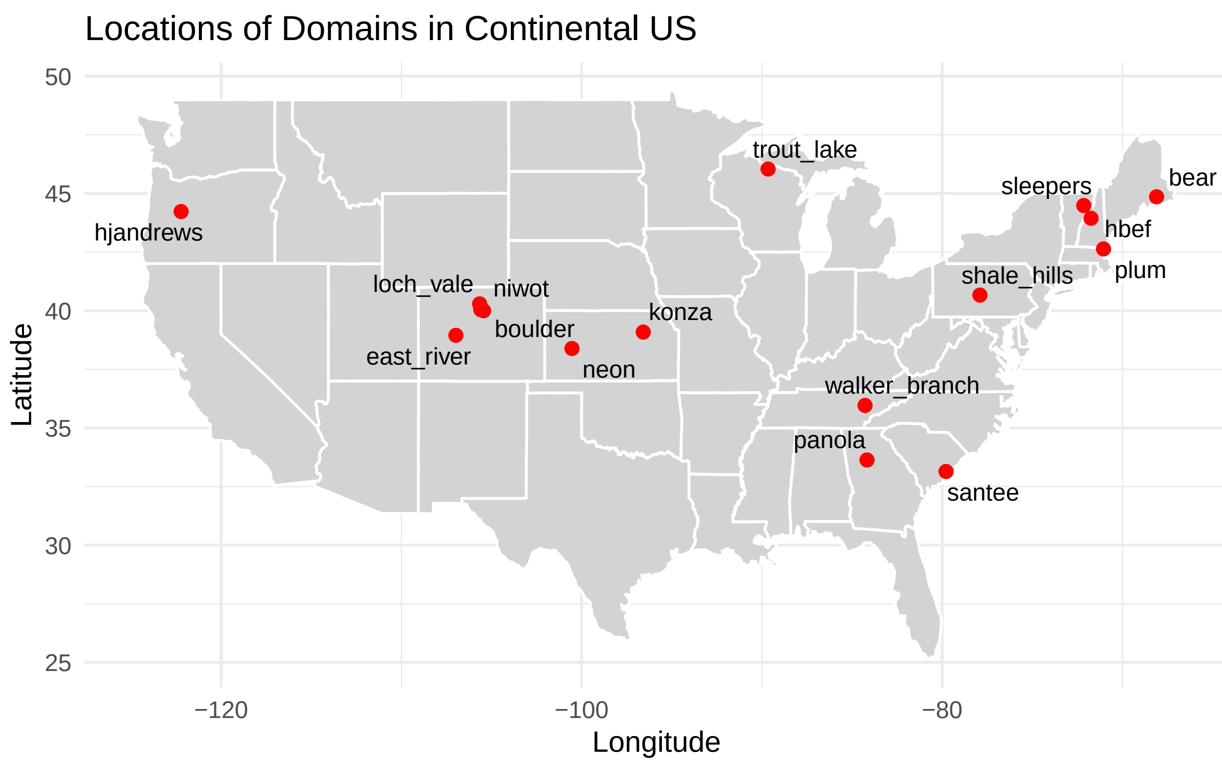


Figure 4: Locations of MacroShed domains in the continental U.S.

and Walker Branch. This geographic diversity allows for comparisons of DOC dynamics across a range of environmental conditions, enhancing the generalizability of the findings.

# 4 Analysis

## 4.1 Question 1: Is there a seasonal difference in DOC export, and how does this vary across the US?

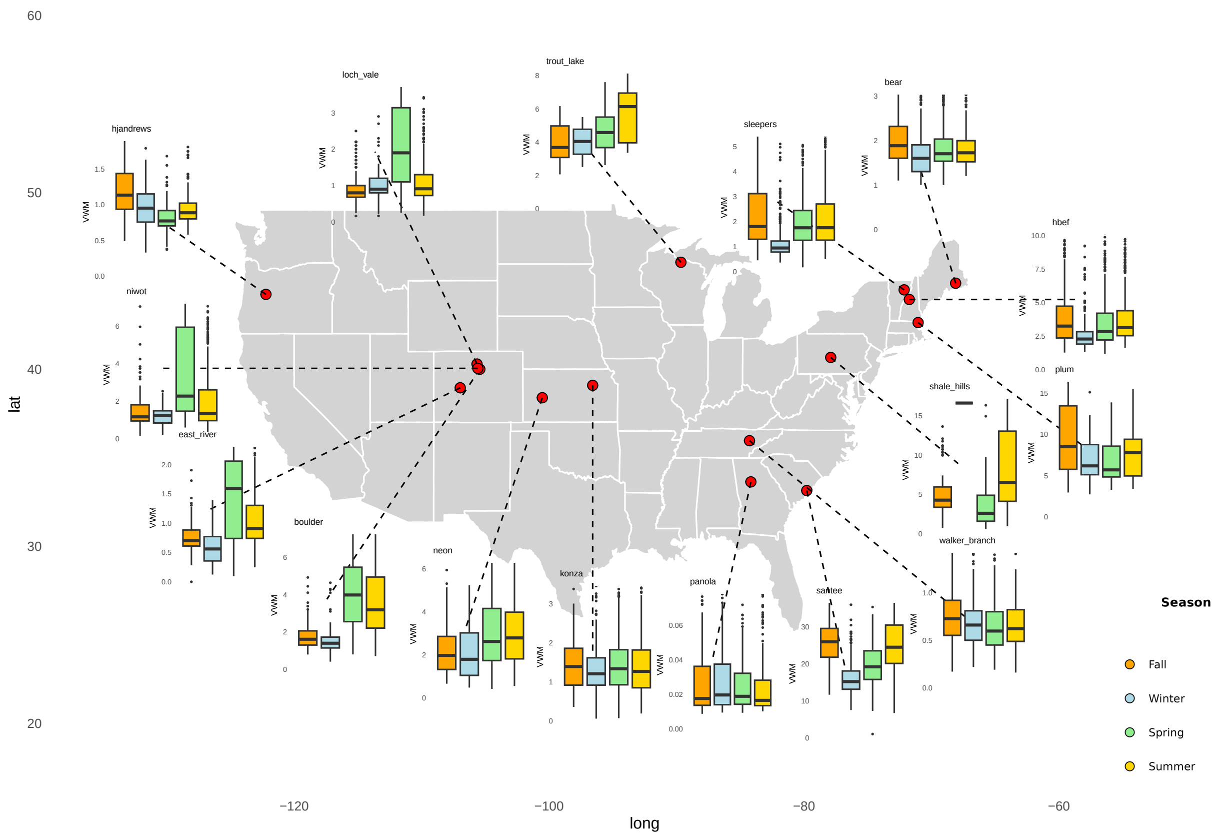


Figure 5: Seasonal DOC export by domain.

Figure 5 showcases seasonal VWM DOC concentrations across U.S. watersheds, with boxplots summarizing variability for each season within domains. Each boxplot represents the interquartile range of VWM DOC values, with the median indicated by the horizontal line and dots capturing outliers. Northern domains like Loch Vale and HJ Andrews display spring peaks, likely influenced by snowmelt, while regions like HBEF exhibit fall peaks, potentially linked to leaf litter decomposition. In contrast, southern domains such as Walker Branch show relatively consistent DOC concentrations year-round, reflecting more stable hydrology. The variation in boxplot sizes and medians highlights regional and seasonal differences in DOC export. Northern forested regions tend to show greater seasonal variability, whereas grassland and southern domains have less pronounced seasonal shifts. These findings emphasize the influence of climate, hydrology, and vegetation on DOC dynamics across diverse U.S. watersheds.

Figure 6 presents seasonal VWM DOC concentrations for each domain, with statistical comparisons highlighted. Boxplots summarize the distribution of VWM DOC values for each season, showing medians (horizontal lines), interquartile ranges (boxes), and variability through whiskers and outliers. Seasonal differences in DOC concentrations were assessed using Dunn's test with Bonferroni correction to control for multiple comparisons.

The letters above each boxplot represent the results of these tests. Matching letters across seasons (e.g., "a") indicate no statistically significant difference between those seasons, while differing letters (e.g., "a" vs. "b")

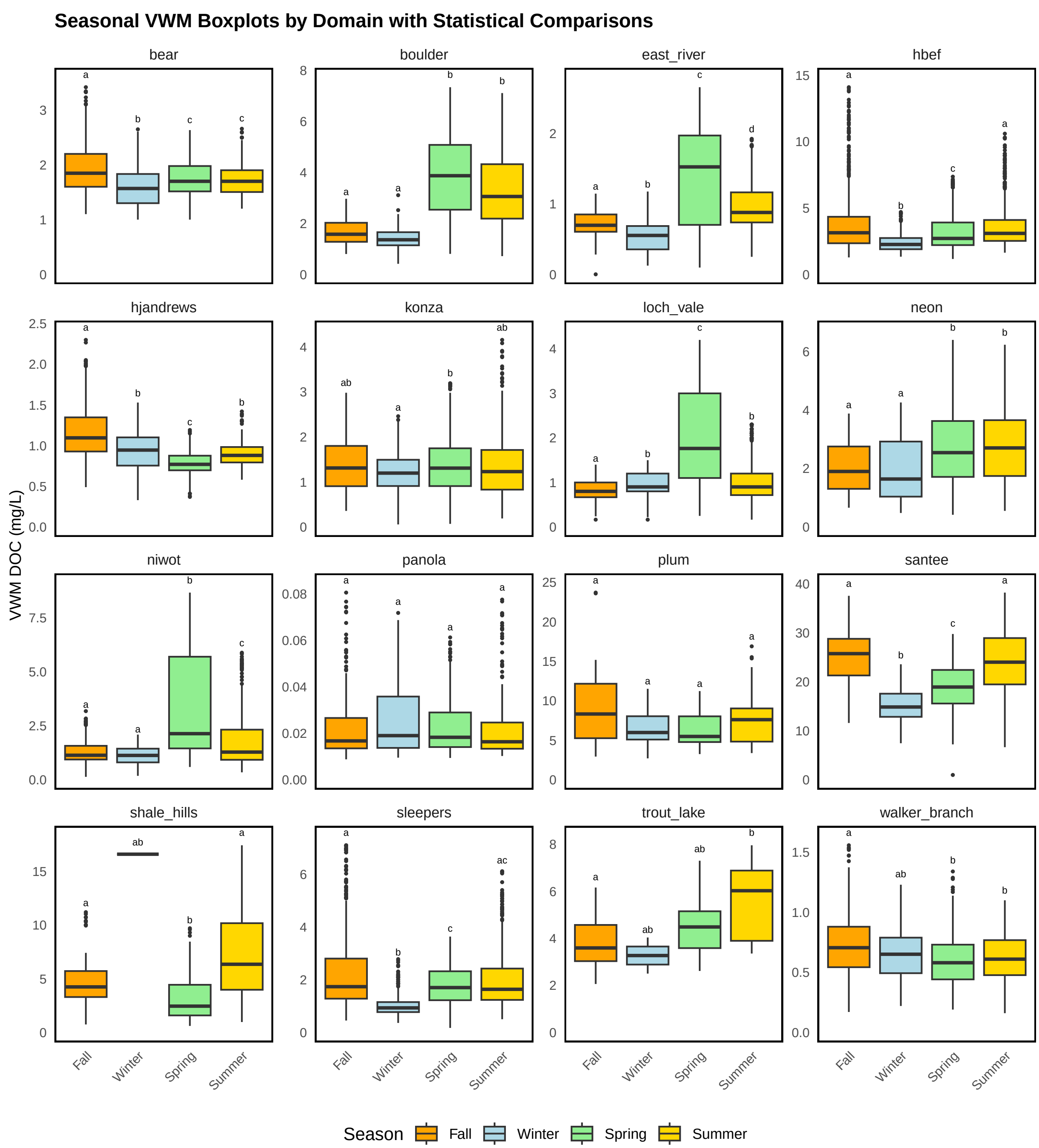


Figure 6: Seasonal VWM DOC concentrations by domain with significant differences indicated by letters.

denote significant differences at a threshold of $p < 0.05$. For example, in the HJ Andrews domain, winter and spring ("b" and "c") show significantly higher VWM DOC concentrations than fall ("a").

Domains like HBEF and Niwot exhibit clear seasonal patterns with distinct statistical differences, potentially driven by factors like snowmelt or leaf litter. In contrast, domains such as Panola and Walker Branch show minimal seasonal variation, reflecting stable hydrological regimes. This analysis highlights the influence of seasonality on DOC export and underscores regional variability in watershed behavior.

## 4.2 Question 2: Are DOC exports significantly changing over time?

Significant DOC Trends Over Time

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 1 row containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_point()`).
```

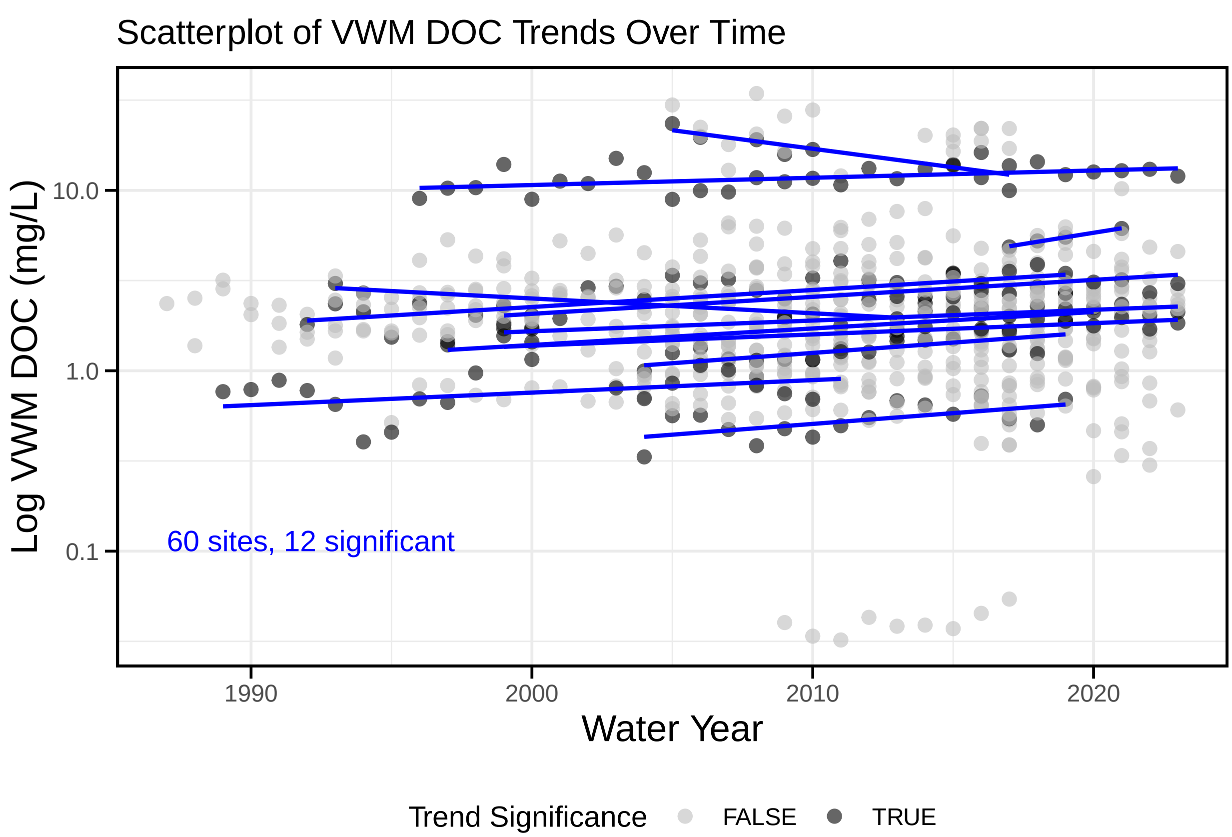


Figure 7: Long-term trends in DOC export by site.

Figure 7 illustrates trends in Volume Weighted Mean (VWM) DOC concentrations over time across 60 sites, with 12 sites showing significant trends. To ensure data quality, we filtered the dataset to include only years with at least 10 months containing at least one observation per month. This criterion minimized potential biases due to incomplete temporal coverage, ensuring more robust trend analysis.

Significance was determined by performing linear regressions of VWM DOC against water year for each site, with a threshold of $p < 0.05$ used to identify significant trends. Sites with significant trends are highlighted with blue regression lines, while non-significant trends are shown as grey points.

While 12 out of 60 sites (~20%) exhibit significant trends, this subset provides valuable insights into long-term DOC export patterns. These trends could reflect changes in watershed processes, land use, or environmental factors, and the notable role of disturbance (e.g., experimental or managed sites) in influencing DOC dynamics warrants further investigation. However, the modest proportion of significant trends also highlights the spatial and temporal variability of DOC exports across the U.S., underscoring the complexity of detecting universal patterns.

Experimental vs. Non-Experimental Watersheds

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 1 row containing non-finite outside the scale range
## (`stat_smooth()`).
```

```
## Warning: Removed 1 row containing missing values or values outside the scale range
## (`geom_point()`).
```

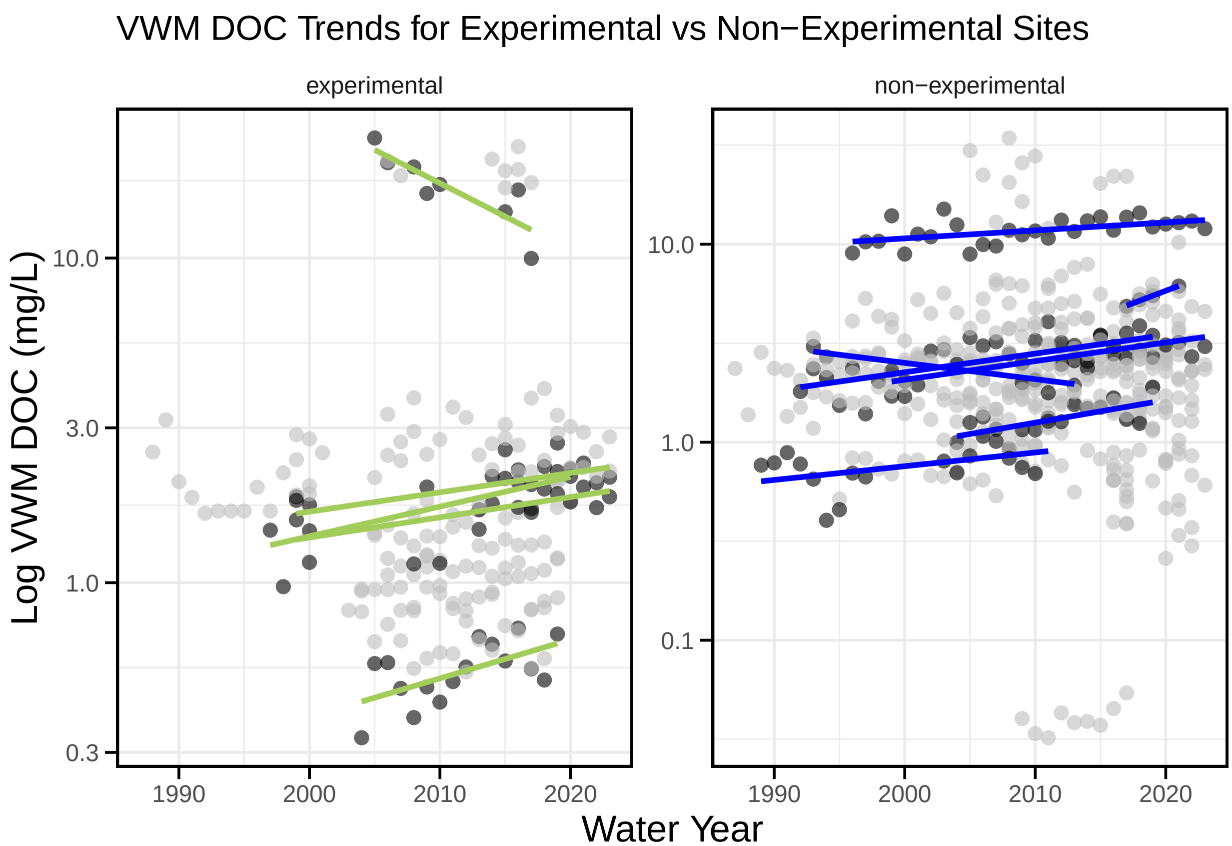


Figure 8: DOC trends in experimental vs. non-experimental watersheds.

Table 1: Table 2. Significant VWM DOC trends by site, categorized by watershed status and domain.

| site_code | slope | p_value | domain | ws_status |
|---|---|---|---|---|
| ARIK | 0.3163595 | 0.0016595 | neon | non-experimental |
| GSWS06 | 0.0141234 | 0.0124852 | hjandrews | experimental |
| GSWS08 | 0.0356349 | 0.0033086 | hjandrews | non-experimental |
| N04D | 0.0372031 | 0.0437061 | konza | experimental |
| W-3 | -0.0456196 | 0.0157762 | sleepers | non-experimental |
| W-9 | 0.0520401 | 0.0001011 | sleepers | non-experimental |
| WS77 | -0.7560523 | 0.0047709 | santee | experimental |
| w2 | 0.0220917 | 0.0096205 | hbef | experimental |
| w3 | 0.0545726 | 0.0099214 | hbef | non-experimental |
| w4 | 0.0253705 | 0.0136306 | hbef | experimental |
| w9 | 0.1022002 | 0.0065993 | hbef | non-experimental |
| west_fork | 0.0127624 | 0.0425567 | walker_branch | non-experimental |

Figure 8 separates the trends in VWM DOC export between experimental and non-experimental sites, providing valuable insights into how disturbances, particularly forest clear-cutting, influence DOC dynamics over time. The experimental sites (left panel) reveal a mix of increasing and decreasing trends, with many showing initial declines followed by steady increases in DOC export. This pattern could indicate the legacy of clear-cutting, where initial disruptions to organic matter cycling lead to reduced DOC export, followed by a gradual recovery as the watershed stabilizes and vegetation regrows, returning DOC levels to a more natural baseline.

In contrast, non-experimental sites (right panel) show a predominance of subtle changes over time, reflecting less external interference and emphasizing natural variability in DOC dynamics. The stark contrast between the two panels underscores how clear-cutting and other experimental manipulations can create significant short- and long-term impacts on watershed biogeochemistry. The observed increases in experimental sites may also suggest a resilience in these systems, where DOC levels eventually rebound despite the initial disturbance. This split approach is particularly useful for teasing apart human-driven versus natural drivers of change in DOC export.

# 5 Summary and Conclusions

This analysis provides valuable insights into the seasonal and temporal dynamics of dissolved organic carbon (DOC) export across diverse watersheds in the continental United States. Seasonal patterns in DOC export reveal distinct peaks in fall and winter, likely driven by processes such as leaf litter decomposition, snowmelt, and enhanced hydrological connectivity. Regions like northeastern forests exhibit particularly high DOC concentrations in the fall, underscoring the role of terrestrial vegetation inputs in shaping seasonal fluxes. The variation across regions highlights the influence of local factors, such as precipitation regimes, land cover, and watershed morphology, in mediating seasonal DOC dynamics.

Over time, DOC export trends illustrate both stability and variability. While most sites show no significant change, the 12 sites with statistically significant trends underscore the potential role of watershed disturbance and climatic factors in altering DOC dynamics. Interestingly, experimental sites, many of which were associated with forest clear-cutting, exhibit greater variability. The increasing trends observed in these sites may reflect DOC levels returning to equilibrium post-disturbance, as organic material stabilizes and hydrological flowpaths recover. Conversely, non-experimental sites with significant trends may point to broader-scale drivers, such as changes in precipitation patterns or rising temperatures due to climate change, influencing DOC mobilization and export.

These findings emphasize the importance of context when interpreting DOC trends. Disturbance events, whether natural or anthropogenic, play a critical role in shaping short-term dynamics, while long-term trends may signal shifts in baseline watershed function. Understanding these dynamics is essential for predicting future DOC fluxes in response to environmental change, informing watershed management strategies, and safeguarding water quality. The nuanced interplay between local processes and large-scale drivers warrants continued investigation, particularly in the face of a changing climate.

GitHub: https://github.com/jonathangilman/Gilman_ENV872_EDA_FinalProject.git