

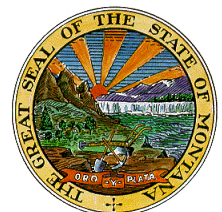


# **Addendum: An Analysis of Daily Patterns of Dissolved Oxygen Change in Flowing Waters of Montana**

**March 16, 2026**

**Prepared by:**

Michael Suplee, PhD  
Montana Department of Environmental Quality  
Water Quality Standards & Modeling Section  
1520 E. Sixth Avenue  
P.O. Box 200901  
Helena, MT 59620-0901



---

**Suggested citation:** Suplee, M.W. 2026. Addendum: An Analysis of Daily Patterns of Dissolved Oxygen Change in Flowing Waters of Montana. Helena, MT: Montana Dept. of Environmental Quality.

---

## Executive Summary

Following the release of *An Analysis of Daily Patterns of Dissolved Oxygen Change in Flowing Waters of Montana* (Suplee, 2023), the Department of Environmental Quality (DEQ) carried out additional analyses pertinent to that report. The results of these additional analyses are provided here.

**For western Montana**, the following topics are addressed:

- An evaluation of possible spatial or temporal autocorrelation in the original regression analysis due to numerous datapoints coming from two rivers (the Big Hole and East Gallatin);
- An evaluation to determine if water surface slope is a principal driver of macroinvertebrate metric patterns in western Montana; and
- Change point determinations for two macroinvertebrate vs. dissolved oxygen (DO) delta regressions in Part I-B of Suplee (2023) that were not completed at that time.

Per the first bullet above, DEQ has concluded that autocorrelation in the regression is not a concern. DEQ would come to the same conclusion regarding a DO delta ( $\Delta$ ) threshold protective of aquatic life regardless of whether all sites or only a single site each from the Big Hole and East Gallatin rivers was included in the regression. Regarding the second bullet, the analytical results showed that whether a stream or medium river was high or low gradient (that is, whether it has greater than or less than 1% water surface slope) is a principal driver of macroinvertebrate metric scores in western Montana streams and medium rivers.

Per the third bullet, DEQ computed the change points for two macroinvertebrate-DO regressions: (a) Beck's Biotic Index ( $v_3$ ) vs. DO  $\Delta$ ; and (b) the Hilsenhoff Biotic Index vs. DO  $\Delta$ . The regressions' change points were 1.96 mg DO/L and 2.54 mg DO/L, respectively. These values are centrally located in the range of candidate DO  $\Delta$ s identified in Suplee (2023). In Suplee (2023), a DO  $\Delta$  threshold of 3.0 mg DO  $\Delta$ /L is recommended. Based on the two new change points, consideration of the error around them, the range of thresholds from the scientific literature, and already-adopted values from other states, DEQ continues to recommend 3.0 mg DO  $\Delta$ /L as the threshold for western Montana.

**For eastern Montana**, DEQ was concerned that it had only identified a DO  $\Delta$  threshold for non-drought periods. This could result in wasted sampling effort if DO data were collected during drought periods. To address this, DEQ has extended its analyses to determine (a) if the same DO  $\Delta$  criterion of 6.0 mg DO/L for non-drought periods is also appropriate for drought periods, and (b) since it is known that drought tends to increase a stream's DO  $\Delta$  value, if a "drought adjustment factor" could be quantified, to apply to DO datasets collected during drought. To quantify the drought adjustment factor, DEQ examined how DO  $\Delta$  changed in plains reference and comparison sites (i.e., sites which have very limited human disturbance) between drought and non-drought periods. DEQ had two independent datasets. The first comprised data from drought and non-drought periods from 2013-2017; this was the "calibration" dataset for the adjustment factor. A second, independent reference stream dataset (2021-2022)—all collected during drought—was used as a "confirmation" dataset. Via a calibration-confirmation approach, DEQ identified a range of candidate adjustment factors that could be used to adjust drought-period datasets prior to comparison to the 6.0 mg/L DO  $\Delta$  criterion. DEQ recommends an adjustment factor of 1.67. The adjustment factor can be applied equally to intermittent and perennial plains streams. An example of applying this adjustment to a plains stream DO dataset (i.e., DO data  $\div$  1.67) is provided in this document.

# TABLE OF CONTENTS

Table of Contents ..... ii

List of Part I Tables ..... iii

List of Part I Figures ..... iii

List of Part II Tables ..... iv

List of Part II Figures ..... iv

Acronyms ..... v

General Introduction: Additional Analyses Related to Dissolved Oxygen Delta ..... 1

Organization of the Rest of this Document ..... 1

Part I Western Montana ..... 2

1.0 Autocorrelation Among Sites which were Used to Derive a DO  $\Delta$  Threshold for Low Gradient Western Montana Streams and Medium Rivers ..... 2

2.0 An evaluation to determine if water surface slope is a principal driver of macroinvertebrate metric patterns in western Montana ..... 7

    2.1 Methods ..... 9

    2.2 Results ..... 10

    2.3 Discussion and Conclusion ..... 12

3.0 Changepoint Analyses on the Beck’s and HBI – DO Delta Regressions ..... 12

Part II Eastern Montana ..... 14

1.0 Problem Definition, Background Information, Project Objectives ..... 14

2.0 Methods, Results ..... 16

    2.1 2013-2017 and 2021-2022 Plains Streams Sampling Sites ..... 16

    2.1 Non-drought and Drought Periods: U.S. Drought Monitor Index ..... 18

    2.2 Relationship between DO  $\Delta$  and DO Minimum in Eastern Montana Streams During Non-Drought and Drought Periods (2013-2017) ..... 20

    2.3 DO  $\Delta$  Patterns Manifested by Reference Streams During Drought and Non-drought Periods (2013-2017) ..... 22

    2.4. Identification and Confirmation of a Drought Adjustment Factor for Application to DO Datasets Collected During Drought ..... 23

3.0 Discussion, Recommendations, and an Example ..... 27

    3.1 Example Application of the Drought Adjustment Factor ..... 27

4.0 Acknowledgements ..... 28

References for Part I, Part II, and Appendices ..... 29

Appendix A. Analyses on Average DO by the Great Lakes Environmental Center (April 2024) ..... 31

Data Dictionary for All Variables Used in Modeling Effort. .... 52

Appendix B. Additional Analysis on DO  $\Delta$  by the Great Lakes Environmental Center (April 2024). Factors near the top and right provide the strongest explanatory power..... 54

Appendix C. Method for Computing a Drought Severity Index ..... 55

    C.1 Determining the Area- and Time-weighted Drought Index..... 55

## LIST OF PART I TABLES

Table 1. Ecoregions Comprising the Geographic Region Addressed in Part I of this Report..... 2

Table 1-1. List of Sites from two Rivers that were Included in Analyses Used to Identify Candidate DO  $\Delta$  Thresholds in Part I-A of Suplee (2023) ..... 3

Table 2-1. Low-gradient Sites Sampled in 2023..... 8

Table 2-2. Ecoregions Comprising Mountain and Low Valleys and Transitional Macroinvertebrate Zones 9

Table 2-3. Mann-Whitney Test p-values for Pair-wise Comparisons of Macroinvertebrate Scores of Different Geographic Zones..... 10

Table 3-1. Changepoint Analysis Results for two Macroinvertebrate Metrics. The analyses were carried out on the datasets presented in Figures 3-1B and 3-2B in Suplee (2023). ..... 13

## LIST OF PART I FIGURES

Figure 1-1. Sites Used in Analyses to Help Support the Derivation of DO  $\Delta$  Thresholds for Low-gradient Western MT Streams and Medium Rivers. Sites with purple triangles provided extant (as of 2022) DO and macroinvertebrate data, while yellow circles were sampled by DEQ in 2023 (numbered, yellow-circle sites are identified by name in Table 2-1 below)..... 4

Figure 1-2. Regression Corresponding to Panel A of Figure 3-1B on Page 29 of Suplee (2023). The blue dotted line is the best-fit logarithmic regression relationship. The black horizontal line is the Low Valleys and Transitional harm threshold for Beck’s Biotic Index (v3) from Schulte and Craine (2023). Orange-filled points are Big Hole River sites, blue-filled are East Gallatin River sites. Triangles are reference sites. .... 5

Figure 1-3. Three Regressions Using only a Single, Randomly-selected Site each from the Big Hole (orange) and East Gallatin (blue) rivers. See text for specific Big Hole or East Gallatin River sites..... 6

Figure 2-2. Box and Whisker Plot of Beck’s Biotic Index (v3) for Reference Site Data from Three Geographic Areas. Each dataset comprises 20 samples. Asterisks indicate outlier values (i.e., greater or less than 1.5 times the interquartile range). ..... 11

Figure 2-3. Box and Whisker Plot of Hilsenhoff Biotic Index (HBI) for Reference Site Data from Three Geographic Areas. Each dataset comprises 20 samples. Asterisk indicates an outlier value (i.e., greater than 1.5 times the interquartile range). ..... 11

## LIST OF PART II TABLES

Table 1. Ecoregions Comprising the Geographic Region Addressed in Part II of this Report..... 14

Table 1-1. Top Five Environmental Variables Affecting Dissolved Oxygen Patterns in Eastern Montana Streams and Medium Rivers, Based on Random Forests Analysis ..... 15

Table 2-1. Reference Sites where Continuous DO was Measured in 2021 and 2022..... 18

Table 2-2. Key Parameters Comprising the US Drought Monitor Index D0 through D4 Categories along with Possible Impacts. “CPC” in the table refers to Climate Prediction Center. .... 19

Table 2-3. Increase in Weekly Average DO  $\Delta$  Due to Drought as Computed for Plains Reference Sites, 2021-2022. See text for methods used to identify the percent increase shown. .... 25

## LIST OF PART II FIGURES

Figure 2-1. Stream Sampling Stations (Black Dots) and their Corresponding Watersheds (in Pink) in the 2013-2017 Plains Streams Study. Major stream segments in each basin are labeled. .... 17

Figure 2-2. Relationships between DO  $\Delta$  and DO Minimum in Montana Plains Streams from the 2013-2017 study. Panel A. During Non-drought Periods. Panel B. During Drought Periods. The gray bands indicate the applicable average weekly DO minimums (per Circular DEQ-7) for streams in the region. ... 21

Figure 2-3. Box and Whisker plot of Nine Reference or Comparison Sites where DO Data were Collected During both Drought and Non-Drought Periods (2013-2017)..... 22

Figure 2-4. Dot Plot of Data from the Nine Reference and Comparison Sites where DO Data were Collected During both Drought and Non-Drought Periods over the 2013 to 2017 Time Period. .... 23

Figure 2-5. Box and Whisker Plot Showing the Percent Increase in DO  $\Delta$  Caused by Drought in Reference and Comparison Sites having Minimal Anthropogenic Disturbance (2013-2017). In the plot, X represents the average, the horizontal line just above it is the median, and the whisker endpoints are 1.5 times the interquartile range. .... 24

Figure 2-6. Comparison of Calibration and Confirmation Datasets for Percent Increase in DO  $\Delta$  Caused by Drought in Reference and Comparison Sites. Panel A. Calibration dataset, comprising eight reference/comparison sites sampled from 2013-2017. Panel B. Confirmation dataset, comprising eight reference sites sampled from 2021-2022 (one high outlier with an increase of 273% is off scale and not shown)..... 26

Figure 3-1. Example of Applying the Adjustment Factor to a DO Dataset Collected During Drought. After adjustment, only the DO  $\Delta$  7-day moving averages highlighted in orange exceed the 6.0 mg/L DO  $\Delta$  threshold. .... 28

## ACRONYMS

<b>CART</b>	Classification and Regression Tree analysis
<b>CPC</b>	Climate Prediction Center
<b>DEQ</b>	Montana Department of Environmental Quality
<b>DO</b>	Dissolved oxygen
<b>DO <math>\Delta</math></b>	Dissolved oxygen delta (daily maximum concentration minus the daily minimum)
<b>EPA</b>	United States Environmental Protection Agency
<b>GLEC</b>	Great Lakes Environmental Center
<b>HBI</b>	Hilsenhoff Biotic Index
<b>HUC</b>	Hydrologic Unit Code (of the United States Geological Survey)
<b>NOAA</b>	National Oceanic and Atmospheric Administration

## GENERAL INTRODUCTION: ADDITIONAL ANALYSES RELATED TO DISSOLVED OXYGEN DELTA

Following the release of *An Analysis of Daily Patterns of Dissolved Oxygen Change in Flowing Waters of Montana* (Suplee, 2023), the Montana Department of Environmental Quality (DEQ) carried out a number of additional analyses in support of the original document. These analyses, along with their results and implications, are provided in this document.

### ORGANIZATION OF THE REST OF THIS DOCUMENT

- **Part I:** Additional analyses for western Montana pertaining to low-gradient streams and medium rivers.
  - **Section 1.0:** Analyses addressing possible spatial autocorrelation effects in regression analyses.
  - **Section 2.0:** Analyses addressing the role of water surface slope in defining patterns in macroinvertebrate metrics in western Montana.
  - **Section 3.0:** Changepoint analysis for the macroinvertebrate metrics vs. DO  $\Delta$  regressions presented in Part I-B of Suplee (2023).
  
- **Part II:** Analyses pertaining to eastern Montana streams and medium rivers for purposes of being able to apply a DO  $\Delta$  threshold protective of aquatic life during both drought and non-drought periods.

## PART I WESTERN MONTANA

**Part I** of this document presents follow-up work pertaining to low-gradient streams and medium rivers in the western part of the state. After release of Suplee (2023), DEQ carried out several additional analyses. Each of these is addressed below in its own section.

**Table 1** below shows the geographic extent of the region under consideration here in **Part I**.

**Table 1. Ecoregions Comprising the Geographic Region Addressed in Part I of this Report**

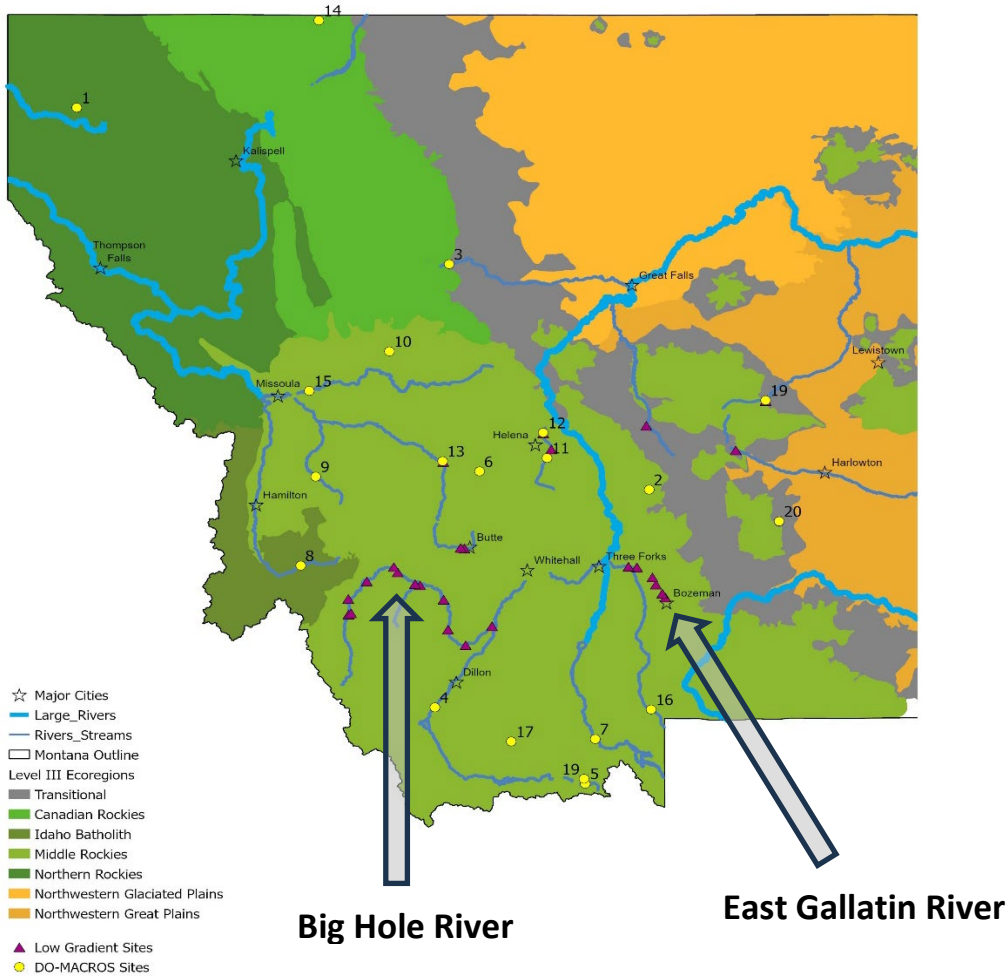
Ecoregions (whole number prefix, level III; number-letter prefix, level IV)
15. Northern Rockies
16. Idaho Batholith
17. Middle Rockies
41. Canadian Rockies
42l. Sweetgrass Uplands
42n. Milk River Pothole Upland
42q. Rocky Mountain Front Foothill Potholes
42r. Foothill Grassland
43s. Non-calcareous Foothill Grassland
43t. Shields-Smith Valleys
43u. Limy Foothill Grassland
43v. Pryor-Bighorn Foothills
43o. Unglaciaded Montana High Plains

### 1.0 AUTOCORRELATION AMONG SITES WHICH WERE USED TO DERIVE A DO $\Delta$ THRESHOLD FOR LOW GRADIENT WESTERN MONTANA STREAMS AND MEDIUM RIVERS

DEQ notes that there is close spatial proximity of sites along the East Gallatin and Big Hole rivers where DO  $\Delta$  and macroinvertebrate data were collected. Sites from these two medium rivers were among the sites used to identify a DO  $\Delta$  threshold for low-gradient streams and medium rivers in Part I of Suplee (2023). The sites in question are in **Table 1-1** and shown in **Figure 1-1** below.

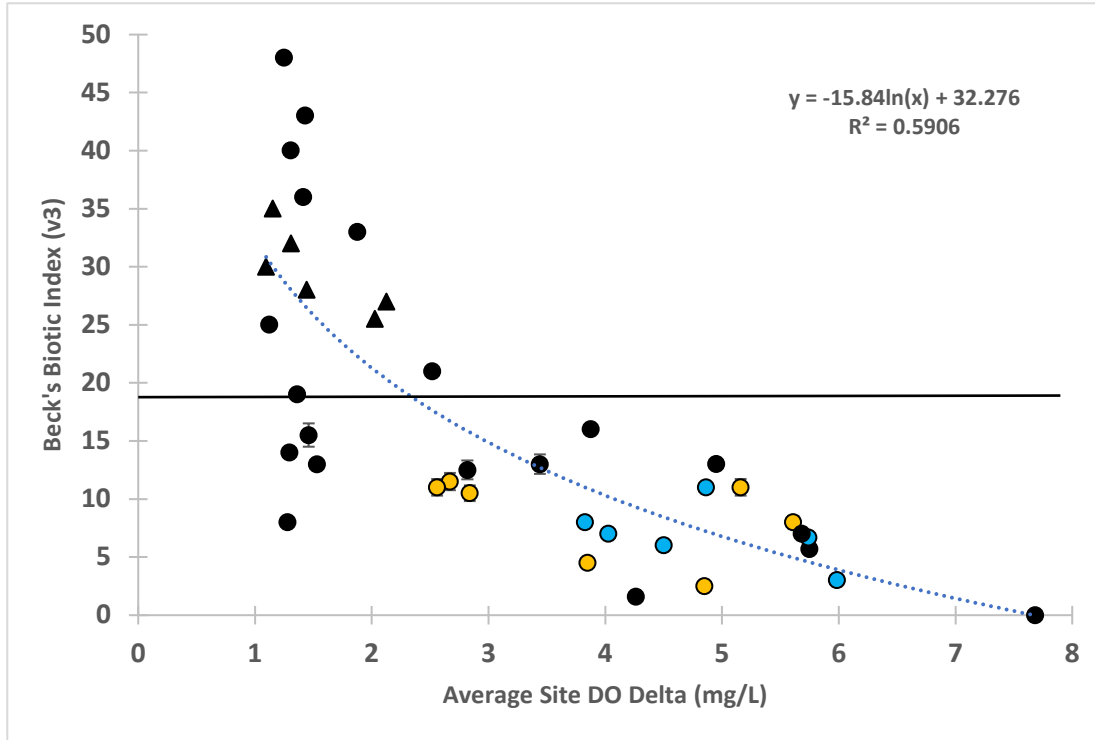
**Table 1-1. List of Sites from two Rivers that were Included in Analyses Used to Identify Candidate DO  $\Delta$  Thresholds in Part I-A of Suplee (2023)**

<b>Site Name</b>	<b>Latitude</b>	<b>Longitude</b>	<b>Water Surface Slope (%)</b>
Big Hole River at Wisdom Bridge	45.61528	-113.45778	0.26
Big Hole River at Mudd Creek Bridge	45.80722	-113.31861	0.22
Big Hole River near Dickie Bridge	45.85972	-113.08361	0.60
Big Hole River at Jerry Creek Bridge	45.78472	-112.91389	0.30
Big Hole River at Maiden Rock	45.70139	-112.73444	0.29
Big Hole River at Kalsta Bridge	45.52667	-112.70083	0.50
Big Hole River at Notchbottom	45.43528	-112.56639	0.22
Big Hole River near Twin Bridges	45.54667	-112.36639	0.01
East Gallatin Site A	45.71410	-111.04760	0.50
East Gallatin Site D	45.73630	-111.07105	0.55
East Gallatin Site G	45.78880	-111.11950	0.54
East Gallatin Site H	45.83059	-111.14617	0.30
East Gallatin Site I	45.88921	-111.26408	0.07
East Gallatin Site J	45.89230	-111.32860	0.15



**Figure 1-1. Sites Used in Analyses to Help Support the Derivation of DO  $\Delta$  Thresholds for Low-gradient Western MT Streams and Medium Rivers. Sites with purple triangles provided extant (as of 2022) DO and macroinvertebrate data, while yellow circles were sampled by DEQ in 2023 (numbered, yellow-circle sites are identified by name in Table 2-1 below).**

The autocorrelation concern revolved around a possible lack of spatial independence of the East Gallatin River and Big Hole River sites which were used (along with other sites) to carry out inferential statistics (regression analysis). A key example of a regression relationship used to identify a DO  $\Delta$  threshold is Panel A of Figure 3-1B in Suplee (2023) and is reproduced below as **Figure 1-2**. A lack of site independence could, potentially, skew the regression relationship and affect any DO  $\Delta$  threshold derived thereof.

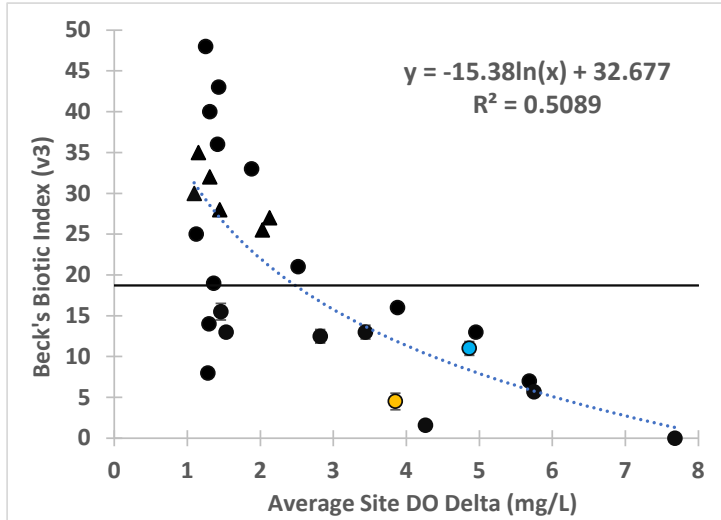


**Figure 1-2. Regression Corresponding to Panel A of Figure 3-1B on Page 29 of Suplee (2023). The blue dotted line is the best-fit logarithmic regression relationship. The black horizontal line is the Low Valleys and Transitional harm threshold for Beck's Biotic Index (v3) from Schulte and Craine (2023). Orange-filled points are Big Hole River sites, blue-filled are East Gallatin River sites. Triangles are reference sites.**

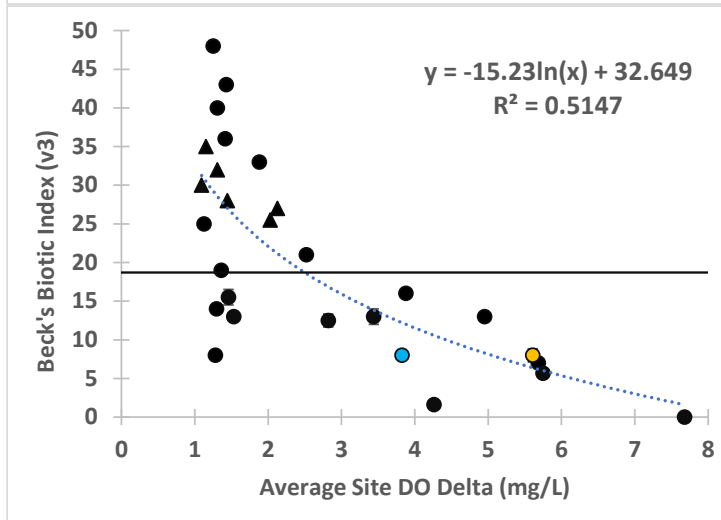
To address the autocorrelation concern, DEQ posed the following question:

*If only a single site from the Big Hole River and a single site from the East Gallatin River were randomly selected and included with the other stream datapoints in the DO  $\Delta$ -macroinvertebrate regression relationship, would DEQ derive a DO  $\Delta$  threshold different from the one it earlier derived from the regression which included all the Big Hole and East Gallatin sites?*

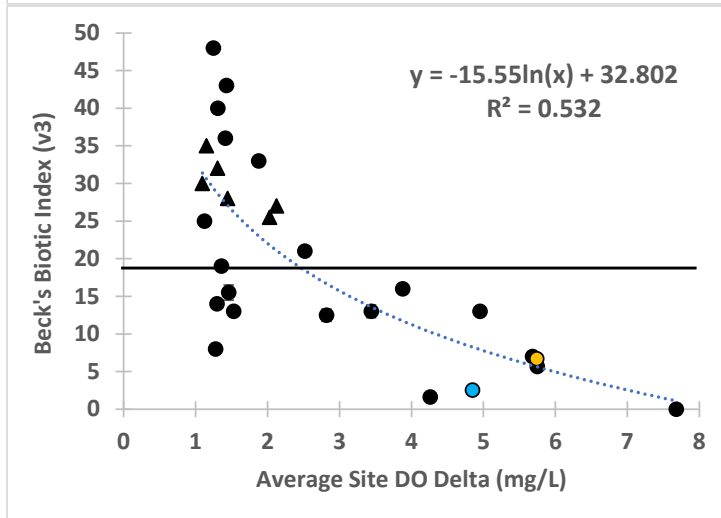
To address this question DEQ randomly selected one Big Hole River and one East Gallatin River site and replotted the regression with those two sites plus the remaining original sites from other streams and rivers. DEQ carried this out six times, each plot containing a single randomly selected data pair (i.e., a DO  $\Delta$  value and an associated macroinvertebrate metric score) from each of the two rivers in question. An example of three of the six regressions is provided in **Figure 1-3**.



A



B



C

Figure 1-3. Three Regressions Using only a Single, Randomly-selected Site each from the Big Hole (orange) and East Gallatin (blue) rivers. See text for specific Big Hole or East Gallatin River sites.

**Figure 1-3, Panel A** contains the Big Hole River at Notch Bottom and the East Gallatin River Site H, **Panel B** includes the Big Hole River near Twin Bridges and the East Gallatin River Site A, and **Panel C** incorporates the Big Hole River at Mudd Cr. Bridge and the East Gallatin River Site G (see **Table 1-1**).

The six reduced-*n* regressions all result in the same finding as was observed from the full dataset—that the best fit regression curve is logarithmic, and it crosses the Becks Biotic Index (v3) threshold (per Schulte and Craine, 2023) between 2 and 3 mg DO  $\Delta$ /L just like in the original analysis (**Figure 1-2**). The  $R^2$  values for the six regressions ranged from 0.501 to 0.532 and were, as expected, a bit lower than in **Figure 1-2** since they are based on less data, but they are very similar overall in their explanatory strength.

From these analyses DEQ has concluded that autocorrelation in the regression is not a concern. DEQ would come to the same conclusion regarding a protective DO  $\Delta$  threshold regardless of whether all sites or only a single site each from the Big Hole and East Gallatin rivers was included in the regression. It should also be borne in mind that this regression was only one line of evidence used to identify the 3.0 mg DO  $\Delta$ /L threshold ultimately recommended in Suplee (2023).

## **2.0 AN EVALUATION TO DETERMINE IF WATER SURFACE SLOPE IS A PRINCIPAL DRIVER OF MACROINVERTEBRATE METRIC PATTERNS IN WESTERN MONTANA**

In 2023, DEQ sampled 20 stream and medium river sites, all low gradient (i.e.,  $\leq 1\%$  water surface slope). The purpose was to provide data to augment the DO  $\Delta$ -macroinvertebrate analytical relationships and threshold identification work in Part I-A of Suplee (2023). The sites sampled in 2023 are shown in **Table 2-1** and **Figure 1-1** (see numbered yellow dots). DEQ's primary criteria for selecting the 2023 sites were that they (a) provided good spatial coverage across a wide swath of the western Montana landscape, (b) they represented different sizes of streams and medium rivers in both reference and non-reference condition, and (c) they all had  $\leq 1\%$  water surface slope.

DEQ notes that most of the 2023-sampled sites are in ecoregions of the Mountains Macroinvertebrate Zone and not in ecoregions of the Low Valleys and Transitional Zone (**Table 2-1**).<sup>1</sup> Ecoregions comprising each of these geographic zones are shown in **Table 2-2**. Sites which were used to support Low Valleys and Transitional Zone regression analyses, but which are located in ecoregions of the Mountains Macroinvertebrate Zone are site numbers 1 and 2, 4 through 10, 14, 15, and 16 (**Table 2-1**); all 12 of these sites have  $\leq 1\%$  water surface slope which DEQ considered to be the key characteristic for sampling them. The remaining sites in **Table 2-1** are from the Low Valleys and Transitional Macroinvertebrate Zone (and also have  $\leq 1\%$  water surface slope).

---

<sup>1</sup> The Mountains Macroinvertebrate Zone and the Low Valleys and Transitional Macroinvertebrate Zone are each comprised of level III and level IV ecoregions; macroinvertebrate characteristics of each zone were analyzed in Schulte and Craine (2023). Basically, the two zones were used as geospatial segregators for differentiating low gradient (Low Valleys and Transitional) from higher gradient (Mountains) streams and medium rivers at a regional scale. **Table 2-2** of this document shows which ecoregions comprise each zone.

Although all 12 of the sites in question sampled in 2023 are low gradient (see “% Water Surface Slope” column in **Table 2-1**), DEQ wanted to be sure that their geographic location in the Mountains Macroinvertebrate Zone was not having a measurable effect on the biological patterns observed at the sites and, as a result, affecting regression analyses in which the 12 sites were included to help derive a DO delta threshold for the Low Valley and Transitional Zone.

**Table 2-1. Low-gradient Sites Sampled in 2023**

Number	Site Name	Site Type	Station ID	Latitude (DD)	Longitude (DD)	Level III Ecoregion	Macroinvertebrate Zone*	% Water Surface Slope
1	Pipe Creek	Stream	K01PIPEC03	48.48895	-115.52419	Northern Rockies	Mountains	0.77
2	Deep Creek	Stream	M09DEEPC10	46.33449	-111.17180	Middle Rockies	Mountains	0.82
3	Sun River	Medium River	M13SUNR64	47.61764	-112.69146	Northwestern Glaciated Plains (Transitional)	Low Valleys and Transitional	0.60
4	Beaverhead River	Medium River	M02BVHDR90	45.06626	-112.80031	Middle Rockies	Mountains	0.29
5	Red Rock Creek	Stream	M01RDRKC01	44.61604	-111.65712	Middle Rockies	Mountains	0.96
6	Little Blackfoot	Medium River	C01LTBLR65	46.43888	-112.46151	Middle Rockies	Mountains	0.45
7	West Fork Madison River	Medium River	M05MDWFR05	44.88117	-111.58234	Middle Rockies	Mountains	0.71
8	East Fork Bitterroot River	Medium River	C05BITER60	45.89515	-113.82223	Idaho Batholith	Mountains	0.79
9	Rock Creek	Stream	C02ROCKC60	46.41035	-113.70605	Middle Rockies	Mountains	0.54
10	Monture Creek	Stream	C03MONTC10	47.12479	-113.14748	Middle Rockies	Mountains	0.96
11	Prickly Pear Creek	Stream	M09PRPEC01	46.51747	-111.94721	Middle Rockies	Low Valleys and Transitional	0.66
12	Prickly Pear Creek at Montana Law Enforcement Academy	Stream	M09PREP02	46.66137	-111.97619	Middle Rockies	Low Valleys and Transitional	0.04
13	Clark Fork River above Little Blackfoot River-Kohrs Bend	Medium River	C01CKFKR03	46.49829	-112.74309	Middle Rockies	Low Valleys and Transitional	0.50
14	Belly River	Medium River	S02BELYR01	48.96806	-113.68263	Canadian Rockies	Mountains	0.30
15	Blackfoot River	Medium River	C03BLACR01	46.89977	-113.75606	Middle Rockies	Mountains	0.09
16	Gallatin River	Medium River	M05GLTNR01	45.05443	-111.15651	Middle Rockies	Mountains	0.50
17	Blacktail Deer Creek East Fork in Robb Creek Wildlife Area	Stream	M02BDEFC01	44.86583	-112.21864	Middle Rockies	Low Valleys and Transitional	1.00
18	Elk Springs Creek	Stream	M01ELKC01	44.64441	-111.6649	Middle Rockies	Low Valleys and Transitional	0.08
19	Middle Fork Judith River	Medium River	M22JUDMF01	46.84653	-110.2860	Northwestern Great Plains (Transitional)	Low Valleys and Transitional	0.44
20	Sweet Grass Creek	Stream	Y03SWTGC07	46.15294	-110.18171	Northwestern Great Plains (Transitional)	Low Valleys and Transitional	0.24

\*Per Schulte and Craine (2023).

**Table 2-2. Ecoregions Comprising Mountain and Low Valleys and Transitional Macroinvertebrate Zones**

Macroinvertebrate Zone	Zone Ecoregions	Ecoregion Level
Mountains	15. Northern Rockies (excl. 15c Flathead Valley)	3
	16. Idaho Batholith	3
	17. Middle Rockies (excl. Level IV Ecoregions in the Low Valleys and Transitional, below)	3
	41. Canadian Rockies	3
Low Valleys and Transitional	15c. Flathead Valley	4
	17s. Bitterroot-Frenchtown Valley	4
	17u. Paradise Valley	4
	17w. Townsend Basin	4
	17aa. Dry Intermontane Sagebrush Valleys	4
	17ac. Big Hole	4
	17ak. Deer Lodge-Philipsburg-Avon Grassy Intermontane Hills and Valleys	4
	42l. Sweetgrass Uplands	4
	42n. Milk River Pothole Upland	4
	42q. Rocky Mountain Front Foothill Potholes	4
	42r. Foothill Grassland	4
	43s. Non-calcareous Foothill Grassland	4
	43t. Shield-Smith Valleys	4
	43u. Limy Foothill Grassland	4
	43v. Pryor-Bighorn Foothills	4
43o. Unglaciaded Montana High Plains	4	

DEQ carried out analyses to determine if the 12 sites were significantly different biologically from their Low Valley and Transitional counterparts. To do this, DEQ analyzed data from reference sites (Suplee et al., 2005) as this approach eliminates anthropogenic impacts that could influence the analysis and cloud the results. DEQ posed the following question:

*Are reference stream Beck's Biotic Index (v3) and Hilsenhoff Biotic Index (HBI) macroinvertebrate scores from the Low Valley and Transitional Zone significantly different from scores from low-gradient ( $\leq 1\%$  water slope) streams located in ecoregions of the Mountains Macroinvertebrate Zone?*

DEQ's hypothesis was that there would be no significant difference between the two zones. DEQ focused on these two macroinvertebrate-based biological metrics because these metrics were important to the identification of a protective DO  $\Delta$  threshold for the Low Valleys and Transitional Zone (see **Figure 1-2** above, and Figures 3-1B through 3-4B in Suplee, 2023).

## 2.1 METHODS

DEQ identified all macroinvertebrate samples collected from reference stream sites between 2005-2023 and segregated them into one of three groups:

1. Sites located in the Low Valley and Transitional Zone
2. Sites located in the Mountains Macroinvertebrate Zone having water surface slope  $\leq 1\%$
3. Sites located in the Mountains Macroinvertebrate Zone having water surface slope  $> 1\%$

The group having the lowest number of samples was the Mountains Macroinvertebrate Zone with  $\leq 1\%$  water surface slope (group 2 above). Only 20 such reference stream macroinvertebrate samples were available for the dataset. To provide balance in the inferential statistical analyses, 20 samples were then randomly drawn (with replacement) from the pool of available samples for the Low Valley and Transitional Zone (group 1), and 20 samples were randomly drawn (with replacement) from the pool of Mountains Macroinvertebrate Zone high-gradient samples (group 3). It is noted here that the Mountains Macroinvertebrate Zone is the largest dataset (211 samples) and is mainly comprised of samples from high gradient sites (i.e.,  $> 90\%$  of all samples in the Mountains Macroinvertebrate Zone are from sites with water slope  $> 1\%$ ).

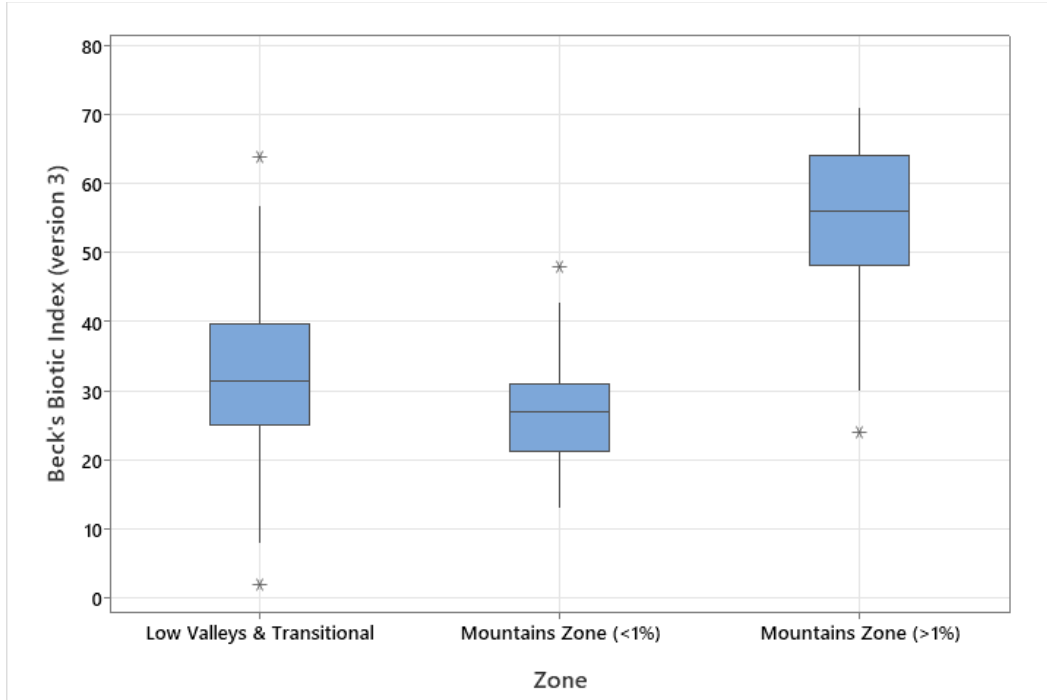
Box and whisker plots were generated for Beck's Biotic Index (v3) and HBI for the three groups, and pair-wise statistical comparisons (Mann-Whitney test, significance level  $\leq 0.01$ , consistent with Suplee, 2023) were carried out in Minitab 21 on the geospatial groups.

## 2.2 RESULTS

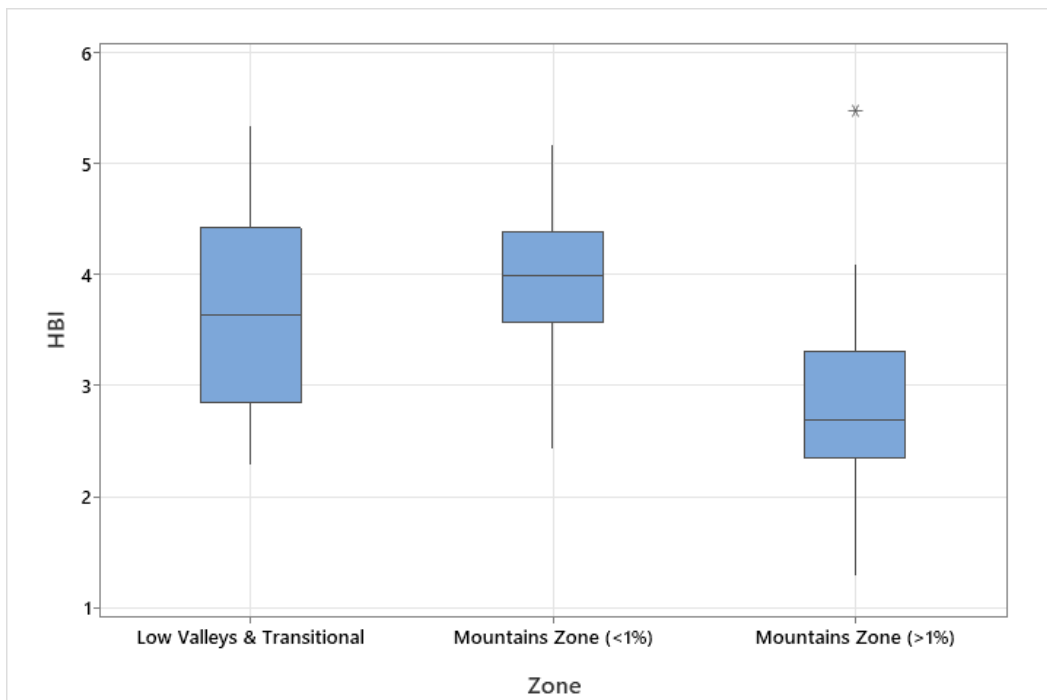
There was no significant difference in Beck's Biotic Index (v3) scores nor HBI scores between reference samples from the Low Valley and Transitional zone (group 1) vs. reference samples from sites in the Mountains Macroinvertebrate Zone having water surface slope  $\leq 1\%$  (group 2; **Table 2-3**). In contrast, there were significant differences ( $p < 0.01$ ) between the Low Valley and Transitional zone (group 1) and the Mountains Macroinvertebrate Zone with water surface slope  $> 1\%$  (group 3), and significant differences ( $p \leq 0.001$ ) between the Mountains Macroinvertebrate Zone with water surface slope  $\leq 1\%$  (group 2) vs. the Mountains Macroinvertebrate Zone with water surface slope  $> 1\%$  (group 3) (**Figures 2-2, 2-3**).

**Table 2-3. Mann-Whitney Test p-values for Pair-wise Comparisons of Macroinvertebrate Scores of Different Geographic Zones**

Mann-Whitney Comparison	Beck's Biotic Index (v3)	HBI
Low Valleys & Transitional <u>vs.</u> Mountains Macroinvertebrate Zone ( $\leq 1\%$ slope)	0.208	0.440
Low Valleys & Transitional <u>vs.</u> Mountains Macroinvertebrate Zone ( $> 1\%$ slope)	$< 0.0001$	0.006
Mountains Macroinvertebrate Zone ( $\leq 1\%$ slope) <u>vs.</u> Mountains Macroinvertebrate Zone ( $> 1\%$ slope)	$< 0.0001$	0.001



**Figure 2-2. Box and Whisker Plot of Beck's Biotic Index (v3) for Reference Site Data from Three Geographic Areas. Each dataset comprises 20 samples. Asterisks indicate outlier values (i.e., greater or less than 1.5 times the interquartile range).**



**Figure 2-3. Box and Whisker Plot of Hilsenhoff Biotic Index (HBI) for Reference Site Data from Three Geographic Areas. Each dataset comprises 20 samples. Asterisk indicates an outlier value (i.e., greater than 1.5 times the interquartile range).**

## 2.3 DISCUSSION AND CONCLUSION

The analytical results presented above indicate that the inclusion, in regression analyses, of a number of low-gradient sites (i.e., those with  $\leq 1\%$  slope) from the Mountains Macroinvertebrate Zone sampled in 2023 should not have had a biasing effect on any DO  $\Delta$  threshold identified for the Low Valleys and Transitional Macroinvertebrate Zone. The biological characteristics of the 12 low-gradient Mountains Macroinvertebrate Zone sites included in the regressions in Figures 3-1B through 3-4B of Suplee (2023) are statistically indistinguishable from their Low Valleys and Transitional counterparts. In contrast, the 12 low-gradient Mountains Macroinvertebrate Zone sites are statistically different from higher gradient sites within the Mountains Macroinvertebrate Zone.

Water surface slope is a key, intrinsic geographic feature of the region's streams and rivers and affects DO patterns (and more). This observation is corroborated by others (Ostermiller et al., 2019). Results in this addendum provide support to the idea that water surface slope, regardless of a site's exact location within a western or transitional Montana ecoregion, is a key characteristic for defining a site's aquatic biological condition.

## 3.0 CHANGEPOINT ANALYSES ON THE BECK'S AND HBI – DO DELTA REGRESSIONS

Part I-B of Suplee (2023) presented, in its figures 3-1B and 3-2B, the regressions between average site DO  $\Delta$  and (a) Beck's Biotic Index ( $v_3$ ) and (b) the HBI macroinvertebrate metric. The regression corresponding to Figure 3-1B Panel A in Suplee (2023) is reproduced in this document in **Figure 1-2** above. The regressions include horizontal lines which represent harm-to-use thresholds for each metric as derived from other sources. However, changepoint analysis was not carried out on these particular regressions<sup>2</sup> and DEQ subsequently concluded that it would be prudent to do so. The changepoint results are provided here.

To determine a changepoint between site average DO  $\Delta$  and a site average macroinvertebrate metric, DEQ used `mvpart` in R (R Core Team, 2022) to run regression tree analysis, setting the tree depth to one (i.e., the root node, which equals the change point; Qian et al., 2003; King and Richardson, 2003).

Results are provided in **Table 3-1** below. The two changepoints (2.54 and 1.96) are centrally located within the range of candidate DO delta values identified in Table 5-1B of Suplee (2023). In Suplee (2023), DEQ recommended a DO  $\Delta$  threshold of 3.0. Based on the two new changepoints, consideration of the error around the estimates, and the range of candidate thresholds from the scientific literature and adopted by other states as shown in Table 5-1B of Suplee (2023), DEQ continues to recommend 3.0 mg DO  $\Delta$ /L as the threshold.

---

<sup>2</sup> Changepoint analysis was carried out on all the macroinvertebrate metric-DO  $\Delta$  regressions in Part I-A of Suplee (2023). See page 14 of that document for a description of changepoint analysis.

**Table 3-1. Changepoint Analysis Results for two Macroinvertebrate Metrics. The analyses were carried out on the datasets presented in Figures 3-1B and 3-2B in Suplee (2023).**

Causal Variable	Response Variable - Code	Response Variable - Description	Predicted Response to Increasing DO $\Delta$	Changepoint Analysis	
				DO $\Delta$ change-point (mg/L)	Relative Error <sup>†</sup>
DO $\Delta$	x_Becks3	Beck's Biotic Index v3	decrease	2.54	0.400
DO $\Delta$	x_HBI	Hilsenhoff Biotic Index (references the TolVal field) using Montana DEQ values	increase	1.96	0.621

<sup>†</sup>Relative error is  $1 - R^2$  root mean square error. This is the error for predictions of the data that were used to estimate the model.

## PART II EASTERN MONTANA

Part II of this document presents follow-up work pertaining to plains streams and medium rivers in the eastern part of the state. See **Table 1** below for the geographic extent of the region under consideration.

**Table 1. Ecoregions Comprising the Geographic Region Addressed in Part II of this Report**

<b>Ecoregions (Whole number prefix: Level III. Number-letter prefix: Level IV)</b>
<b>18. Wyoming Basin</b>
<b>42. Northwestern Glaciated Plains</b> (excluding level IV ecoregions listed below)
42l. Sweetgrass Uplands
42n. Milk River Pothole Upland
42q. Rocky Mountain Front Foothill Potholes
42r. Foothill Grassland
<b>43. Northwestern Great Plains</b> (excluding level IV ecoregions listed below)
43s. Non-calcareous Foothill Grassland
43t. Shield-Smith Valleys
43u. Limy Foothill Grassland
43v. Pryor-Bighorn Foothills
43o. Unglaciated Montana High Plains

### 1.0 PROBLEM DEFINITION, BACKGROUND INFORMATION, PROJECT OBJECTIVES

Suplee (2023) provided a recommended DO  $\Delta$  threshold for plains streams during non-drought periods. After Suplee (2023) was released, DEQ began evaluating methods by which a DO  $\Delta$  threshold could also be applied during drought periods, since drought occurs episodically in eastern Montana. DEQ was concerned that only having a DO  $\Delta$  threshold for non-drought periods would result in wasted sampling effort if DO data happen to have been collected during drought. It would also have left the region's waterbodies without a eutrophication assessment target for substantial periods of time (i.e., whenever drought occurs).

The effect of environmental variables on several DO measurements (e.g., average weekly DO minimum, maximum DO  $\Delta$ ) was examined via Classification and Regression Tree (CART) analysis for eastern Montana streams and medium rivers by the Great Lakes Environmental Center, Inc. (GLEC, 2021); GLEC (2021) in turn informed Suplee (2023). In 2024, GLEC carried out additional work for DEQ (using Random Forests) to again assess environmental causal variables on stream DO and DO  $\Delta$ . Random Forests is a more advanced form of CART analysis and GLEC's 2024 analyses are provided here (**Appendix A, B**). The objective of the updated work was to determine if one or more intrinsic environmental variables—that is, environmental factors having very limited potential to be anthropogenically influenced, like

watershed area—might help better explain stream DO patterns and, potentially, eliminate the need to account for drought effects on DO patterns.

GLEC used the same mix of predictor variables as used in GLEC (2021)<sup>3</sup>; see page 5 of that document for the list (they are also at the end of **Appendix A**). Variables included land use/cover, watershed drainage area above each site, stream category (perennial, intermittent, etc.), slope, drought status, and many other intrinsic and anthropogenically-affected environmental variables. Using Random Forests, GLEC found that it was the anthropogenically-affected environmental variables, not the intrinsic environmental variables, that explained the most variation in average DO and average DO  $\Delta$  (**Table 1-1**). These results are consistent with the earlier findings in GLEC (2021) using CART, where it was found that in eastern Montana “low levels of watershed disturbance and the absence of prolonged drought conditions were the most consistent predictors for optimal DO conditions, expressed as either diel variation or as a minimum.”

**Table 1-1. Top Five Environmental Variables Affecting Dissolved Oxygen Patterns in Eastern Montana Streams and Medium Rivers, Based on Random Forests Analysis**

	Dissolved Oxygen (DO) Response Variable	Intrinsic or Anthropogenically affected	Dissolved Oxygen (DO) Response Variable	Intrinsic or Anthropogenically affected
Rank in Order of Importance for Explaining Variation	Average Site DO	Average Site DO	Average DO $\Delta$	Average DO $\Delta$
1	% disturbed land cover (watershed scale)	Anthropogenically affected	pH	Anthropogenically affected
2	% natural land cover (watershed scale)	Anthropogenically affected	% natural land cover (watershed scale)	Anthropogenically affected
3	% disturbed land cover (near-field; <1 km)	Anthropogenically affected	% disturbed land cover (watershed scale)	Anthropogenically affected
4	% natural land cover (near-field; <1 km)	Anthropogenically affected	Total Nitrogen (mg/L)	Anthropogenically affected
5	specific conductance ( $\mu\text{S}/\text{cm}$ )	Anthropogenically affected	NOAA Palmer Meteorological Drought Index (drought)	Intrinsic*

\*Drought is an intrinsic characteristics of the Montana plains landscape, however its frequency and duration may be affected by anthropogenically-driven climate change (Whitlock et al., 2017).

Based on the findings in GLEC (2021) and the follow-up work in this document using Random Forests, DEQ concluded that drought needs to be given consideration when applying a DO  $\Delta$  target. Drought consistently rises to the top as a key variable affecting plains stream and medium river DO patterns and may be considered an intrinsic environmental variable for purposes herein. As documented in both GLEC

<sup>3</sup> A few of the original variables were dropped, see **Appendix A** for explanation.

(2021) and Suplee (2023), drought alone leads to higher DO  $\Delta$  values in flowing waters of the Montana plains region. The overarching purpose of **Part II** of this document is to identify a process whereby DO  $\Delta$  measurements for the region's streams and medium rivers can be evaluated whether the data were collected during non-drought or drought periods.

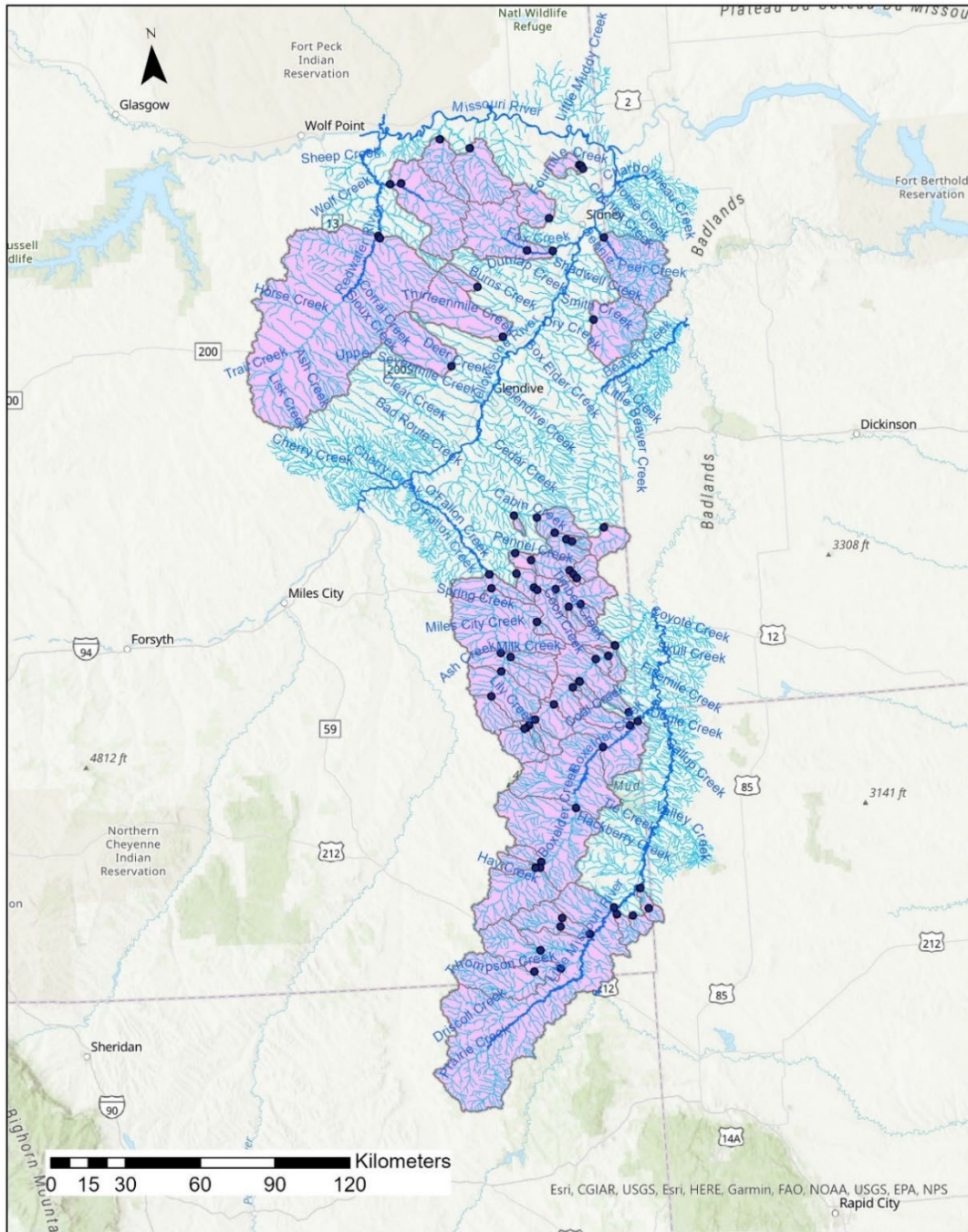
## 2.0 METHODS, RESULTS

The ideas and processes used to address the issue outlined above in **Section 1.0** are presented here. Methods and results are presented together within each section, as applicable.

### 2.1 2013-2017 AND 2021-2022 PLAINS STREAMS SAMPLING SITES

**Figure 2-1** shows Plains sites sampled over a five-year period from 2013-2017. That study's purpose was to characterize DO patterns in plains streams across a wide geographic extent of eastern Montana. Details on the study and its findings are in GLEC (2021) and Section 2.0 of Suplee (2023). Per this addendum, the same sites are being considered.

In 2021 and 2022, DEQ collected continuous DO data in reference streams (i.e., those which have very limited human disturbance; Suplee et al., 2005); see **Table 2-1** below. All the data were collected during a drought period.



**Figure 2-1. Stream Sampling Stations (Black Dots) and their Corresponding Watersheds (in Pink) in the 2013-2017 Plains Streams Study. Major stream segments in each basin are labeled.**

**Table 2-1. Reference Sites where Continuous DO was Measured in 2021 and 2022**

Waterbody Site Name	Station ID	Sampling Year	Latitude	Longitude
Bitter Creek on BLM land	M43BITRC01	2022	48.64890	-106.90250
Little Beaver Creek	Y27LBVRC17	2021	46.067780	-104.335830
Little Beaver Creek	Y27LBVRC01	2021	46.045300	-104.372800
Little Dry Creek	M29LDRYC04	2022	47.34500	-106.36500
Little Dry Creek (Site 2)	M29LDRYC08	2021	47.308370	-106.356770
Rock Creek @ USGS gage near international boundary	M43ROCKC07	2022	48.96940	-106.83890
Rock Cr (BLM land)	M43ROCKC06	2022	48.65416	-107.03250
Rock Creek (Site 1)	REFRC1	2022	48.87583	-106.89670
Rock Creek (Site 2)	M43ROCKC05	2022	48.58613	-106.99804
Snap Creek	M31SNAPC01	2022	47.55660	-106.29230
Poplar River West Fork south of Peerless MT	REFWFPR	2022	48.69695	-105.83190
Poplar River West Fork near Richland MT	M46POWFR01	2022	48.80810	-106.02060
Willow Creek about 3/4 mile downstream Eagles Nest Coulee	M43WILOC01	2022	48.57954	-106.97281
Woody Island Coulee	M38WDYIC01	2022	48.92265	-108.37948

## 2.1 NON-DROUGHT AND DROUGHT PERIODS: U.S. DROUGHT MONITOR INDEX

GLEC (2021) shows that drought affects DO  $\Delta$  in streams and that a useful drought index is the “number of consecutive weeks at a drought severity of  $D_{ZERO}$ ” (first row, **Table 2-2**). GLEC (2021) observes that a given region does not experience a higher intensity drought (e.g.,  $D_3 - D_4$ ) until some duration of lower intensity drought ( $D_0 - D_1$ ) exists. GLEC (2021) shows the break point between drought and non-drought periods to be six consecutive weeks at  $D_{ZERO}$ . That is,  $\leq 6$  consecutive weeks at  $D_{ZERO}$  are non-drought periods, while  $> 6$  consecutive weeks at  $D_{ZERO}$  are drought periods. Instructions on how to compute a drought index value from a public website are provided in **Appendix C**.

**Table 2-2. Key Parameters Comprising the US Drought Monitor Index D0 through D4 Categories along with Possible Impacts. “CPC” in the table refers to Climate Prediction Center.**

Category	Description	Possible Impacts	Ranges				
			Palmer Drought Severity Index (PDSI)	CPC Soil Moisture Model (Percentiles)	USGS Weekly Streamflow (Percentiles)	Standardized Precipitation Index (SPI)	Objective Drought Indicator Blends (Percentiles)
D0	Abnormally Dry	Going into drought: <ul style="list-style-type: none"> <li>■ short-term dryness slowing planting, growth of crops or pastures</li> </ul> Coming out of drought: <ul style="list-style-type: none"> <li>■ some lingering water deficits</li> <li>■ pastures or crops not fully recovered</li> </ul>	-1.0 to -1.9	21 to 30	21 to 30	-0.5 to -0.7	21 to 30
D1	Moderate Drought	<ul style="list-style-type: none"> <li>■ Some damage to crops, pastures</li> <li>■ Streams, reservoirs, or wells low, some water shortages developing or imminent</li> <li>■ Voluntary water-use restrictions requested</li> </ul>	-2.0 to -2.9	11 to 20	11 to 20	-0.8 to -1.2	11 to 20
D2	Severe Drought	<ul style="list-style-type: none"> <li>■ Crop or pasture losses likely</li> <li>■ Water shortages common</li> <li>■ Water restrictions imposed</li> </ul>	-3.0 to -3.9	6 to 10	6 to 10	-1.3 to -1.5	6 to 10
D3	Extreme Drought	<ul style="list-style-type: none"> <li>■ Major crop/pasture losses</li> <li>■ Widespread water shortages or restrictions</li> </ul>	-4.0 to -4.9	3 to 5	3 to 5	-1.6 to -1.9	3 to 5
D4	Exceptional Drought	<ul style="list-style-type: none"> <li>■ Exceptional and widespread crop/pasture losses</li> <li>■ Shortages of water in reservoirs, streams, and wells creating water emergencies</li> </ul>	-5.0 or less	0 to 2	0 to 2	-2.0 or less	0 to 2

## 2.2 RELATIONSHIP BETWEEN DO $\Delta$ AND DO MINIMUM IN EASTERN MONTANA STREAMS DURING NON-DROUGHT AND DROUGHT PERIODS (2013-2017)

In Suplee (2023), a significant regression relationship was observed between DO  $\Delta$  values and their corresponding minimum weekly DO concentrations; see **Figure 2-2A**. This relationship was observed during non-drought periods. Ohio observes a very similar relationship between DO  $\Delta$  and minimum weekly DO concentrations in its low gradient streams and identifies a DO  $\Delta$  of 6.0 mg/L as appropriate to maintain minimum weekly DO standards (Miltner, 2010). Suplee (2023) used the regression in **Figure 2-2A** along with other information (e.g., work by Ohio, Minnesota) to identify a DO  $\Delta$  threshold of 6.0 mg/L.

It results that a similar regression relationship exists during drought periods in Montana plains streams (**Figure 2-2B**). The relationships in **Figure 2-2A** and **Figure 2-2B** are both significant ( $p < 0.001$ ) although during drought the  $R^2$  is lower compared to non-drought. **Figure 2-2B** (drought) indicates that an average DO  $\Delta$  of  $\sim 5.3$  mg/L is associated with the minimum weekly DO concentrations standards applicable in this region (gray horizontal band in the figure).

Taken together, the data indicate the same DO  $\Delta$  threshold (6.0 mg/L) can reasonably be applied to plains streams during drought and non-drought periods.

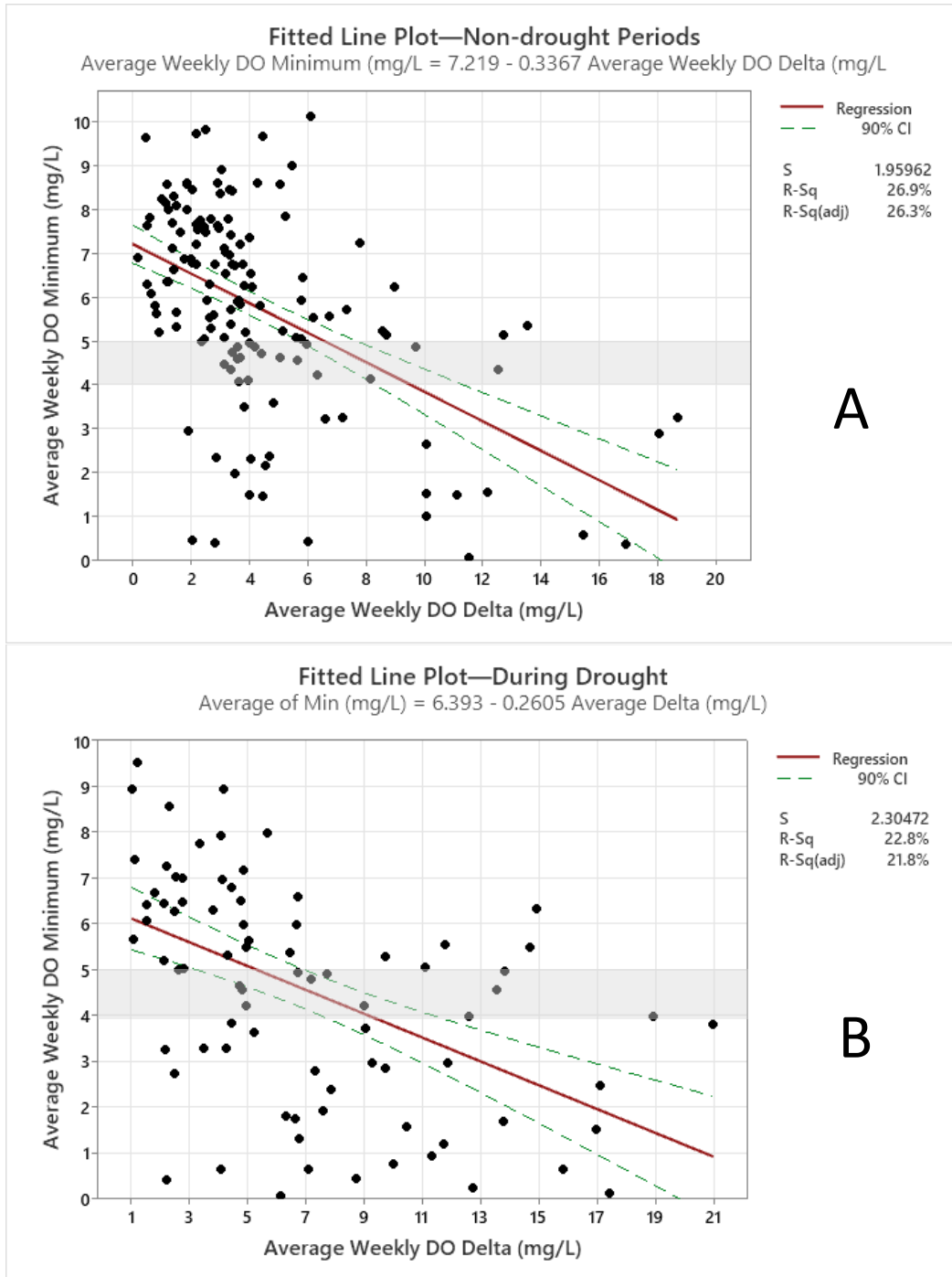


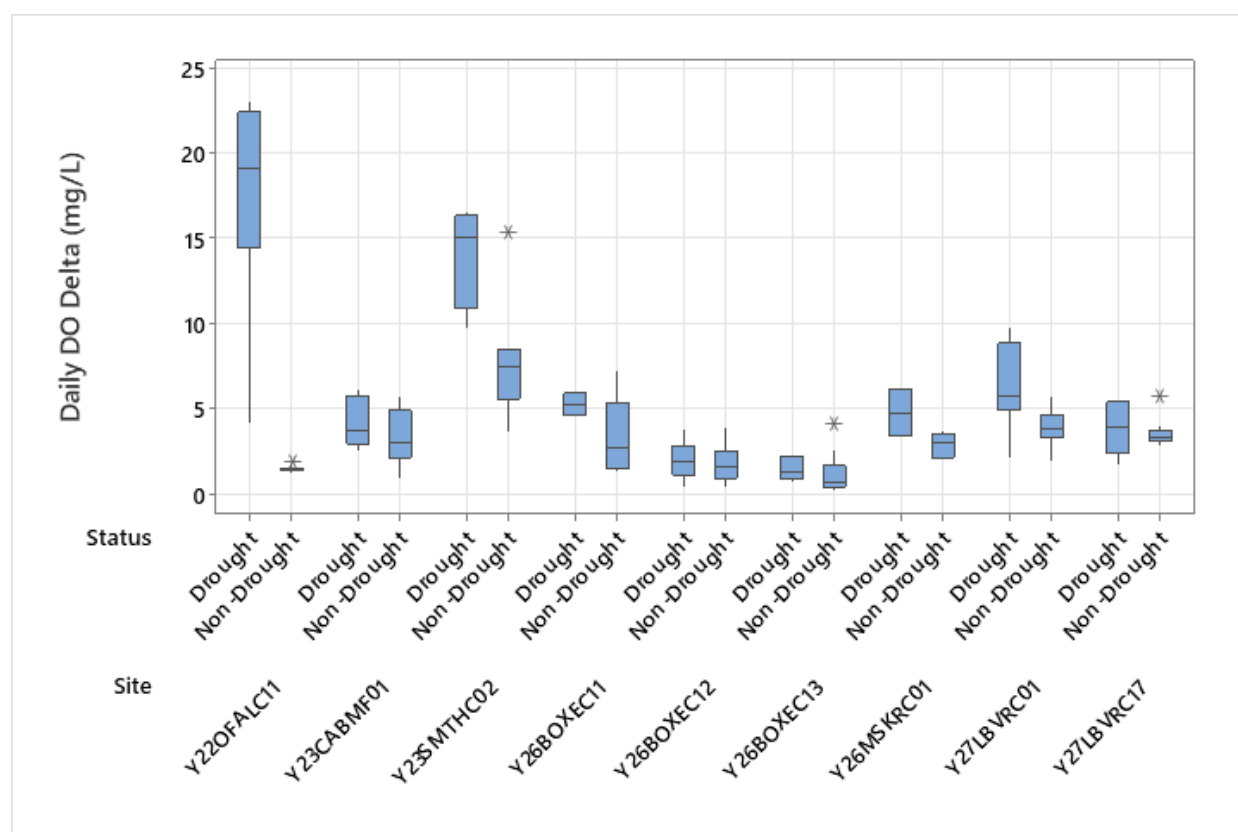
Figure 2-2. Relationships between DO  $\Delta$  and DO Minimum in Montana Plains Streams from the 2013-2017 study. Panel A. During Non-drought Periods. Panel B. During Drought Periods<sup>4</sup>. The gray bands indicate the applicable average weekly DO minimums (per Circular DEQ-7) for streams in the region.

<sup>4</sup> One datapoint with an unusually high DO minimum of 13.4 mg/L (and a corresponding DO  $\Delta$  of 3.5 mg/L) is not shown in order to scale the Y-axis the same as in Panel A. The datapoint is, however, included in computations for regression coefficients, significance levels, and the R<sup>2</sup> value in **Figure 2-2B**.

## 2.3 DO $\Delta$ PATTERNS MANIFESTED BY REFERENCE STREAMS DURING DROUGHT AND NON-DROUGHT PERIODS (2013-2017)

Reference and comparison stream sites are regional sites which have limited human disturbance (Suplee et al., 2005; and Section 2.4.2 in Suplee, 2023). Here, DEQ examines how DO  $\Delta$  patterns changed in reference and comparison sites as a direct result of drought. Because reference and comparison sites have limited anthropogenic influences, they are well suited for an analysis of the baseline effect of drought on regional streams.

During the 2013-2017 plains stream study, nine reference/comparison sites (of 28) had data collected during both drought and non-drought periods (the other 19 sites had datasets exclusively encompassing either drought or non-drought periods). For each of the nine minimally disturbed sites, DO  $\Delta$  was always higher during drought than during non-drought (**Figure 2-3**). Collectively for the nine sites, there was significantly higher DO  $\Delta$  during drought than during non-drought (Mann-Whitney,  $p < 0.001$ )<sup>5</sup>. **Figure 2-4** shows a dot plot for the collective dataset, segregated by drought and non-drought periods.



**Figure 2-3. Box and Whisker plot of Nine Reference or Comparison Sites where DO Data were Collected During both Drought and Non-Drought Periods (2013-2017).**

<sup>5</sup> As can be noted in **Figure 2-3**, O'Fallon Creek (Y22OFALC11) had an unusually high degree of increase in DO  $\Delta$  from drought. If this site's drought and non-drought data are removed from the collective dataset, the Mann-Whitney test p-value is instead 0.07.

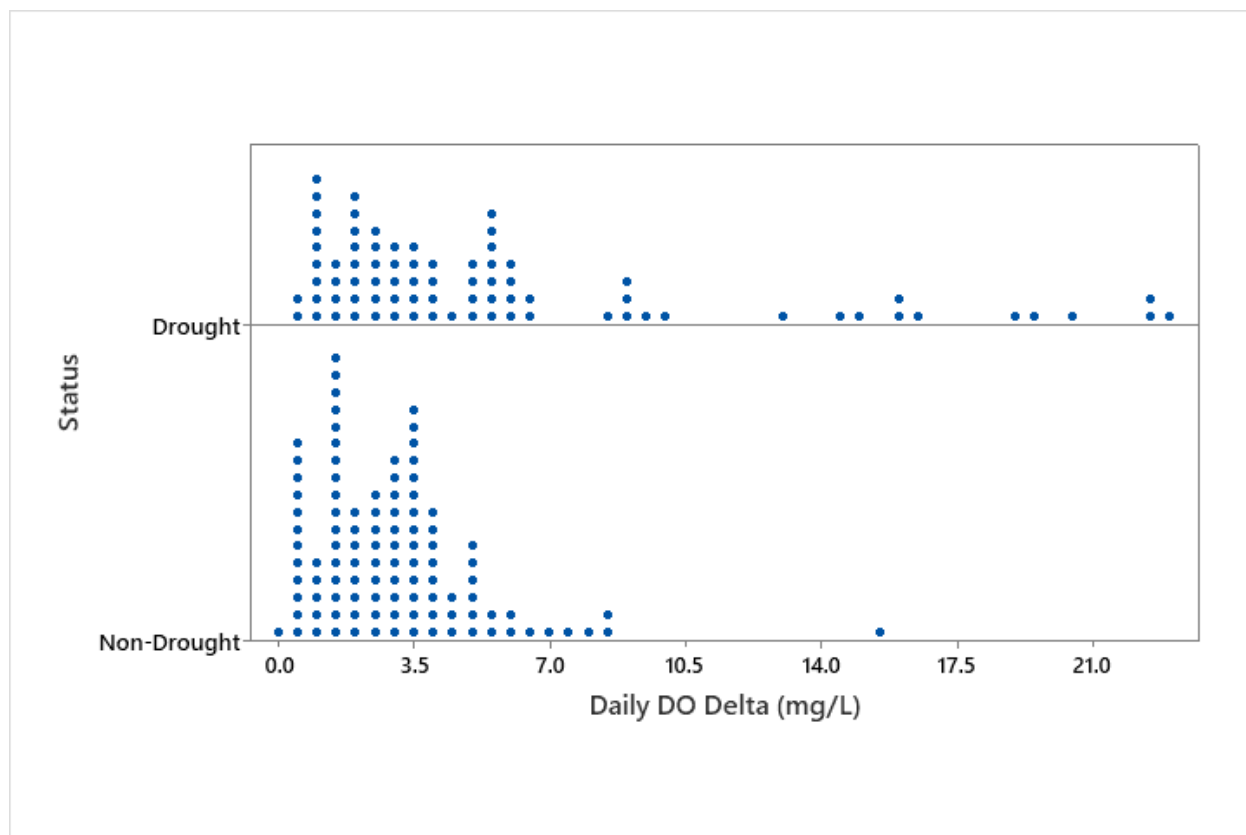


Figure 2-4. Dot Plot of Data from the Nine Reference and Comparison Sites where DO Data were Collected During both Drought and Non-Drought Periods over the 2013 to 2017 Time Period.

## 2.4. IDENTIFICATION AND CONFIRMATION OF A DROUGHT ADJUSTMENT FACTOR FOR APPLICATION TO DO DATASETS COLLECTED DURING DROUGHT

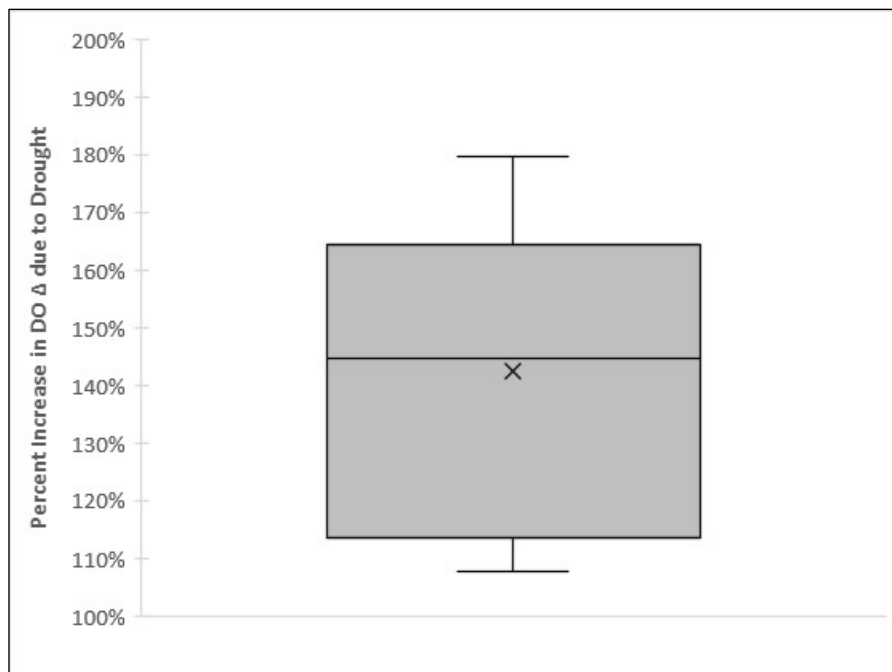
It was shown in **Section 2.3** that the region's least impacted streams (i.e. reference streams) will manifest higher DO  $\Delta$  during drought periods. **Figure 2-5** below shows a box and whisker plot of the percent increase in DO  $\Delta$  induced by drought in plains reference and comparison sites over the 2013-2017 period. **Figure 2-5** is based on a site-by-site analysis comparing DO  $\Delta$  for periods experiencing drought vs. non-drought over the five-year period<sup>6</sup>. One site (Y22OFALC11) was not included in the figure as it is an extreme outlier<sup>7</sup> among the nine reference/comparison sites (it had a >1000% increase in DO  $\Delta$ <sup>8</sup>, see **Figure 2-3**); the likely reasons for this unusually high difference are discussed in **Section 3.0**. The interquartile range of DO  $\Delta$  increase due to drought for the remaining eight sites is 114% to 165% (mean 143%; median 145%). That is, drought causes DO  $\Delta$  measured in plains reference streams to increase on average by about 1.5 times above what it would have been if there had not been drought. This simple empirical model of drought effect on DO  $\Delta$  provides DEQ a range of candidate

<sup>6</sup> All available daily DO  $\Delta$  values for a site during non-drought periods were averaged and then compared to the average of all daily DO  $\Delta$  values from the same site during drought periods. The percent increase was then computed.

<sup>7</sup> An outlier is  $\geq 1.5$  times the interquartile range of a given dataset.

<sup>8</sup> As noted, the extreme outlier (>1000% increase) was excluded from figure. If it were included: Q1=119%; Q3=173%.

adjustment factors which could be used to factor out the known effect of drought. Doing so allows comparison of plains streams datasets to the 6.0 mg/L threshold regardless of whether the data were collected during drought or non-drought periods without concern that a threshold exceedance was driven purely by drought.



**Figure 2-5. Box and Whisker Plot Showing the Percent Increase in DO  $\Delta$  Caused by Drought in Reference and Comparison Sites having Minimal Anthropogenic Disturbance (2013-2017). In the plot, X represents the average, the horizontal line just above it is the median, and the whisker endpoints are 1.5 times the interquartile range.**

To confirm the empirical drought-effect model developed from the 2013-2017 dataset, DEQ used an independent set of data collected from Montana plains reference sites in 2021 and 2022 ( $n=14$  site/year combinations), all collected during drought (i.e., the data were collected during periods when there was  $\geq 6$  consecutive weeks at  $D_{ZERO}$ ). Only two of the 2021-2022 reference sites (Y27LBVRC01, Y27LBVRC 17) are common to the reference/comparison sites from 2013-2017. Thus, the 2021-22 reference dataset provides a high degree of spatial and temporal independence for model confirmation. In theory, adjustment factors that would bring the 2021-2022 reference site datasets into compliance with the target of  $\leq 6$  mg/L average DO  $\Delta$  and  $\leq 15\%$  exceedance rate should fall within the range identified for the 2013-2017 dataset (i.e., **Figure 2-5**). If so, the empirical drought-effect model for DO  $\Delta$  increase will have been independently confirmed.

For the 2021-2022 reference site dataset, an adjustment factor for each site dataset was varied until the dataset's average weekly DO  $\Delta$  could just achieve the  $\leq 6.0$  mg/L and  $\leq 15\%$  exceedance rate target. Eight

reference sites needed adjustment factors to achieve this, four sites already met the target and did not need any adjustment, and two sites were excluded from the analysis for QC reasons<sup>9</sup>.

The results for the eight sites that needed adjustment to meet the target are shown in **Table 2-3**. After adjusting each reference site dataset until it just met the target, it was found that the reference sites manifested drought-induced DO  $\Delta$  increases ranging from 105 to 273% (273% was an outlier). The interquartile range of DO  $\Delta$  increase due to drought was 121% to 166% (mean 156%; median 146%). Whether a stream was intermittent or perennial had no bearing on the degree of effect drought had, as borne out by statistical analysis (Mann-Whitney, perennial vs. intermittent;  $p=0.89$ ).

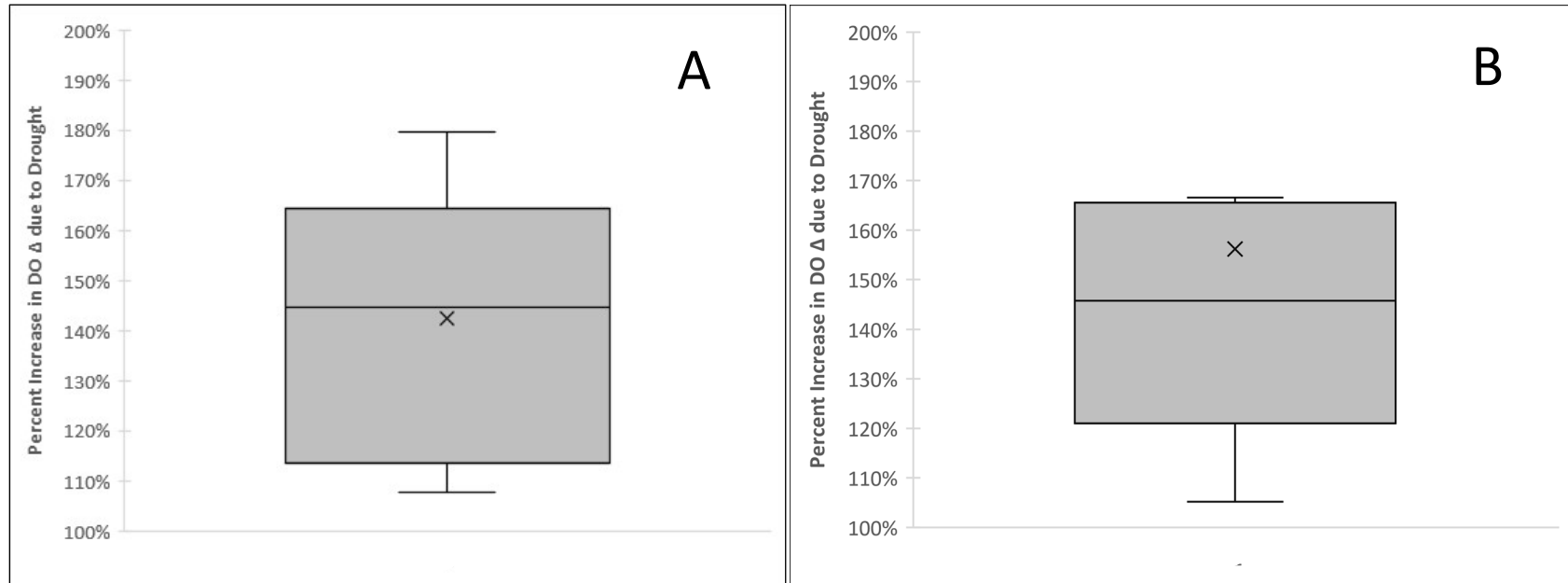
**Table 2-3. Increase in Weekly Average DO  $\Delta$  Due to Drought as Computed for Plains Reference Sites, 2021-2022. See text for methods used to identify the percent increase shown.**

Reference Site	Year	Computed Increase in Weekly Average DO $\Delta$ Due to Drought	Stream Status
M43BITRC01	2022	152%	Intermittent
Y27LBVRC17	2021	117%	Intermittent
Y27LBVRC01	2021	167%	Perennial
M29LDRYC08	2021	140%	Intermittent
M43ROCKC06	2022	105%	Perennial
REFRC1	2022	162%	Perennial
M43ROCKC05	2022	135%	Perennial
M38WDYIC01	2022	273%	Intermittent

**Figure 2-6** compares the box and whisker plots for the 2013-2017 calibration dataset vs. the 2021-2022 confirmation dataset. The confirmation dataset shows a similar interquartile range as the calibration dataset, providing DEQ with a plausible range of drought adjustment factors to consider. Given that all the sites in **Figure 2-6** have limited anthropogenic impacts, DEQ considers it reasonable to establish an adjustment factor that would ensure that most of these sites could meet the target during drought. Given this objective, DEQ considers an adjustment factor of **1.67** to be a good candidate for application to DO  $\Delta$  datasets collected in drought period. The value 1.67 corresponds closely to the upper quartile of the calibration and confirmation datasets (**Figure 2-6**).

Given the above, DEQ considers it reasonable to divide weekly average (or 7-day rolling average) DO  $\Delta$  values from plains sites collected during drought (i.e., >6 consecutive weeks at  $D_{ZERO}$  on the U.S. Drought Monitor Index) by 1.67. The adjusted weekly averages can then be compared to the  $\leq 6.0$  mg/L and  $\leq 15\%$  exceedance rate target. An example of this adjustment process is provided in **Section 3.0**.

<sup>9</sup> These two reference sites have agricultural land use in the upstream watershed (49%) right at the allowable maximum cutoff established in Suplee et al. (2005). They also have moderate local grazing. To be prudent, DEQ concluded it would be best to exclude these sites' DO data from this analysis.



**Figure 2-6. Comparison of Calibration and Confirmation Datasets for Percent Increase in DO  $\Delta$  Caused by Drought in Reference and Comparison Sites. Panel A. Calibration dataset, comprising eight reference/comparison sites sampled from 2013-2017. Panel B. Confirmation dataset, comprising eight reference sites sampled from 2021-2022 (one high outlier with an increase of 273% is off scale and not shown).**

### 3.0 DISCUSSION, RECOMMENDATIONS, AND AN EXAMPLE

Weekly average DO  $\Delta$  values collected during drought in eastern Montana plains streams could be adjusted (that is, divided by) a factor of 1.67 to provide a reasonable estimate of what the DO  $\Delta$  value's magnitude would look like during non-drought conditions. The adjustment factor was derived from an empirical model of drought effects on plains stream sites having limited anthropogenic impacts. The model was then validated using a temporally and (mostly) spatially independent DO  $\Delta$  dataset from plains reference streams. This calibration/confirmation approach provided good confidence in the range of candidate adjustment factors and applies equally to intermittent and perennial plains streams.

As noted in **Section 2.4**, one site showed unusually high DO  $\Delta$  during drought (site Y22OFALC11, 2013-2017 dataset). This site warrants discussion as it was not included in the box and whisker plots in **Figures 2-5** and **2-6**. Site Y22OFALC11 is on O'Fallon Creek. In 2013 (non-drought), water at the site was turbid and flowing at about 5-10 cubic feet per second, and there was no observed attached filamentous algae or macrophytes. Conditions such as these will limit daily DO oscillation—consistent with the small DO  $\Delta$ s observed at the site that year. During drought in 2017, there was only about 1 cubic foot per second of flow (driven mainly by rains in the days prior to sampling), the water was recorded as looking like “pea soup,” and the stream had only sparse attached algae growth and no macrophytes. These observations indicate that in 2017, under drought, the O'Fallon Creek site had switched to a phytoplankton-dominated system. Highly turbid, flashy plains streams lacking macrophytes have the propensity for this type of ecological state change (Suplee, 2004). High phytoplankton biomass and limited flow in 2017 undoubtedly led to the unusually high DO oscillations (and elevated DO  $\Delta$ ) observed during drought.

In short, there are site-specific situations in plains reference sites that can lead to unusually high DO  $\Delta$ s. Building a drought adjustment factor for DO  $\Delta$  to accommodate such high outliers would be imprudent. As such, DEQ recognizes that the recommended adjustment factor (1.67) is appropriate for most—but not all—cases, and is likely conservative (i.e., not set as high as it potentially could be).

#### 3.1 EXAMPLE APPLICATION OF THE DROUGHT ADJUSTMENT FACTOR

DEQ has shown that reference and comparison sites have a central tendency (interquartile range) of increased DO  $\Delta$  due to drought ranging from about 114% to 173%<sup>10</sup>. Based on these findings, DEQ recommends 1.67 as a candidate drought adjustment factor. Datasets collected during drought (>6 consecutive weeks at  $D_{ZERO}$  on the U.S. Drought Monitor Index) could be adjusted as provided in the example outlined below.

In **Figure 3-1** below, DO data have been collected by a deployed instrument for 22 days during drought and, in turn, daily DO  $\Delta$  values were computed (3<sup>rd</sup> column from left). From these, 16 seven-day moving averages are computed (4<sup>th</sup> column from left). Each of these seven-day averages is then divided by the 1.67 drought adjustment factor (top cell, 5<sup>th</sup> column from left) and stored in the 6<sup>th</sup> column from the left. The average DO  $\Delta$  for this dataset (top cell, 2<sup>nd</sup> column from right) is computed from the drought-adjusted values in the 6<sup>th</sup> column from the left. The exceedance rate (top cell, right-most column) is computed by summing the number of adjusted seven-day values that exceed 6.0 mg/L (those highlighted in tan) and dividing by the number of values in the dataset (n=16). In this example, the

<sup>10</sup> This interquartile range includes consideration of the effect of outliers; see also, footnotes 8.

**Figure 3-1** dataset complies with the target (6 mg DO  $\Delta$ /L and 15% allowable exceedance rate) after the drought adjustment has been made, whereas the dataset would not comply without the adjustment.

Station ID	Date	Daily DO delta (mg/L)	Moving 7-day Average DO $\Delta$ (mg/L)	Adjustment Factor	Adjusted 7-day Moving Average DO $\Delta$ (mg/L)	Adjusted Dataset 7-day Average (mg/L)	Exceedence Frequency
Site X	7/29/2022	2.67		<b>1.67</b>		<b>4.00</b>	<b>12.5%</b>
Site X	7/30/2022	3.36					
Site X	7/31/2022	4.05					
Site X	8/1/2022	7.91					
Site X	8/2/2022	4.35					
Site X	8/3/2022	5.72					
Site X	8/4/2022	4.13	4.60		2.75		
Site X	8/5/2022	4.52	4.86		2.91		
Site X	8/6/2022	6.78	5.35		3.20		
Site X	8/7/2022	4.70	5.44		3.26		
Site X	8/8/2022	4.67	4.98		2.98		
Site X	8/9/2022	7.38	5.41		3.24		
Site X	8/10/2022	3.97	5.16		3.09		
Site X	8/11/2022	7.34	5.62		3.37		
Site X	8/12/2022	8.15	6.14		3.68		
Site X	8/13/2022	9.76	6.57		3.93		
Site X	8/14/2022	6.30	6.80		4.07		
Site X	8/15/2022	10.77	7.67		4.59		
Site X	8/16/2022	9.16	7.92		4.74		
Site X	8/17/2022	12.65	9.16		5.49		
Site X	8/18/2022	13.63	10.06		<b>6.02</b>		
Site X	8/19/2022	14.90	11.02		<b>6.60</b>		

**Figure 3-1. Example of Applying the Adjustment Factor to a DO Dataset Collected During Drought.** After adjustment, only the DO  $\Delta$  7-day moving averages highlighted in orange exceed the 6.0 mg/L DO  $\Delta$  threshold.

## 4.0 ACKNOWLEDGEMENTS

Many, many thanks to Rosie Sada (DEQ) for making sure the five-year plains streams study was completed; that study formed the basis of most of the findings in **Part II** of this report. Thanks to the Carter, Little Beaver, Dawson, and Richland County Conservation Districts who assisted us in completing the five-year plains stream study. Many thanks to Dale White (Great Lakes Environmental Center, Inc.) for carrying out additional work on the dataset using Random Forests. Thanks again to Rosie Sada (DEQ) for ensuring continuous DO datasets were collected from plains reference sites in 2021 and 2022. Thanks to Drs. Scott Weir and Erin Eberhard (DEQ) for reviewing the draft of this document.

## REFERENCES FOR PART I, PART II, AND APPENDICES

- Boulesteix, A.L., J. Silke, K. Jochen, and I. Koenig. 2012. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *WIREs Data Mining and Knowledge Discovery* 2: 493–507.
- GLEC (Great Lakes Environmental Center, Inc.). 2021. *Dissolved Oxygen Spatial Analysis, Technical Progress Report —Phase II*. Prepared by Dale White, Principal Research Scientist.
- Heim, R.R. 2002. A review of twentieth-century drought Indices used in the United States. *Bulletin of the American Meteorological Society*, 83(8): 1149-1166.
- Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15: 651–674.
- King, R.S., and C.J. Richardson. 2003. Integrating bioassessment and ecological risk assessment: an approach to developing numeric water quality criteria. *Environmental Management* 31: 795-809.
- Miltner, R.J. 2010. A method and rationale for deriving nutrient criteria for small rivers and streams in Ohio. *Environmental Management* 45: 842-855.
- Ostermiller, J.D., M. Shupryt, M.A. Baker, B. Neilson, E.B. Gaddis, A.J. Hobson, B. Marshall, T. Miller, D. Richards, and N. von Stackelberg. 2019. *Technical Support Document: Utah’s Nutrient Strategy, Scientific Investigations to Support Utah’s Nutrient Reduction Program*. Utah Department of Environmental Quality, Water Quality.
- Qian, S.S., King, R.S., and C.J. Richardson. 2003. Two statistical methods for the detection of environmental thresholds. *Ecological Modelling* 166: 87-97.
- R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> [r-project.org]. Package: mvpart, Version: 1.6-2, Date: 2014-06-22; Author: rpart by Terry M Therneau and Beth Atkinson [atkinson@mayo.edu](mailto:atkinson@mayo.edu), R port of rpart by Brian Ripley [ripley@stats.ox.ac.uk](mailto:ripley@stats.ox.ac.uk), Some routines from vegan -- Jari Oksanen [jari.oksanen@oulu.fi](mailto:jari.oksanen@oulu.fi) ; Extensions and adaptations of rpart to mvpart by Glenn De'ath <[g.death@aims.gov.au](mailto:g.death@aims.gov.au)>. Maintainer: Glenn De'ath [g.death@aims.gov.au](mailto:g.death@aims.gov.au) Description: Multivariate regression trees.
- Schulte, N.O., and J.M. Craine. 2023. *Eutrophication Thresholds Associated with Benthic Macroinvertebrate Conditions in Montana Streams*. Prepared for the MT Dept. of Environmental Quality by Jonah Ventures. October 5, 2023.
- Strobl, C., A.L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*. 8: 25.
- Suplee, M.W. 2023. *An Analysis of Daily Patterns of Dissolved Oxygen Change in Flowing Waters of Montana*. Helena, MT: Montana Dept. of Environmental Quality.

Suplee, M.W., Sada de Suplee, R., Feldman, D., and T. Laidlaw. 2005. *Identification and Assessment of Montana Reference Streams: A Follow-up and Expansion of the 1992 Benchmark Biology Study*. Helena, MT: Montana Dept. of Environmental Quality.

Svoboda, M. 2000. An introduction to the drought monitor. *Drought Network News* 12: 15–20.

Whitlock, C., W.F. Cross, B. Maxwell, N. Silverman, and A.A. Wade. 2017. 2017 Montana Climate Assessment. Bozeman and Missoula, MT: Montana State University and University of Montana, Montana Institute on Ecosystems. 318 p. doi: 10.15788/m2ww8w.

## APPENDIX A. ANALYSES ON AVERAGE DO BY THE GREAT LAKES ENVIRONMENTAL CENTER (APRIL 2024)

### Random Forests for Eastern Montana Rivers and Streams Preliminary Analysis

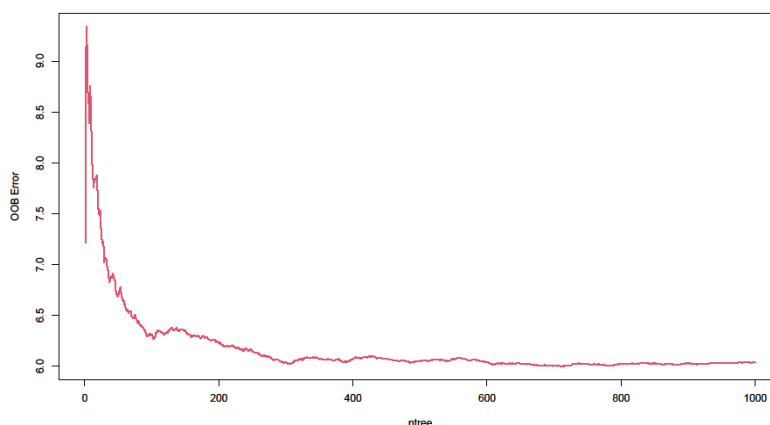
#### Background

Compared to the previous work on classification and regression trees (CART) for eastern Montana rivers and streams, the predictive performance of trees can be substantially improved when aggregating many decision trees. CART decision trees may suffer from high variance. This means that if we split the training data into two parts at random, and fit a decision tree to both halves, the results that we get could be different.

In contrast, a procedure with low variance will yield similar results if applied repeatedly to distinct data sets (in this approach,  $n_{tree} = 1000$  distinct datasets). Bootstrap aggregation, or *bagging*, is a general-purpose procedure for reducing the variance of a statistical learning method. It is particularly useful and frequently used in the context of decision trees. *Random forests* provide an improvement over bagged trees by way of a random small tweak that decorrelates the trees. As in bagging, we build a number (or forest) of decision trees on bootstrapped training samples. But when building these decision trees, each time a split in a tree is considered, a random sample of  $m$  predictors is chosen as split candidates from the full set of  $p$  predictors (this is the tweak to decorrelate the trees). The split is allowed to use only one of those  $m$  predictors. A fresh sample of  $m$  predictors is taken at each split.

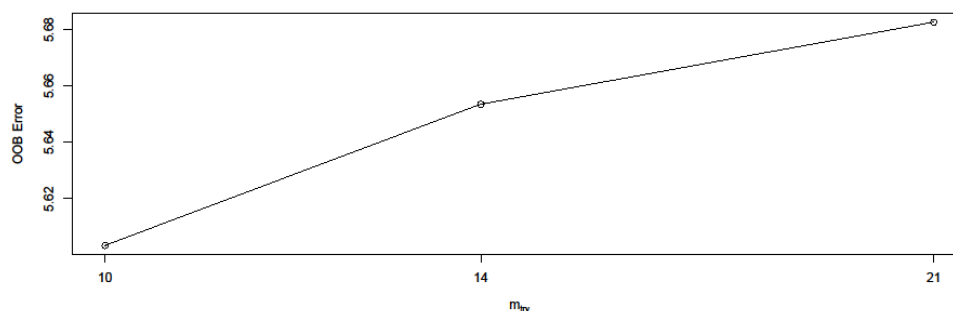
#### Forest Model Parameter Selection

A random forest model was successfully completed with **weekly average of daily average dissolved oxygen (DO)** as the response variable. Daily average DO was computed from the mid-point of the measured daily minimum and daily maximum values. Predictor variables totaled 40 compared to 43 variables used in the CART effort previously (GLEC, 2021). Predictors dropped in the random forest effort included TNe, TPe, and BP. Both TNe and TPe were variables that had missing values imputed for TN and TP. Imputation of single predictors was no longer necessary here as a preliminary run of the forest model imputed missing values for all predictors. BP (barometric pressure) was also eliminated. All variables named in this effort are defined in **Appendix B-GLEC**.



**Figure 1.** Plot of out-of-bag (OOB) error versus number of trees in simulation.

The random sampling of  $\mathbf{m}$  parameters occurred for 1000 tree creations (**Figure 1**). Typically we choose  $\mathbf{m} \approx p/3$  ( $40/3 = 13.33 \rightarrow 14$ ) for numerical predictors. GLEC also ran an optimization scheme to tune for the best subset number and  $\mathbf{m} = 10$  produced the lowest prediction error (**Figure 2**). Hence,  $\mathbf{m} = 10$  was chosen for this model.

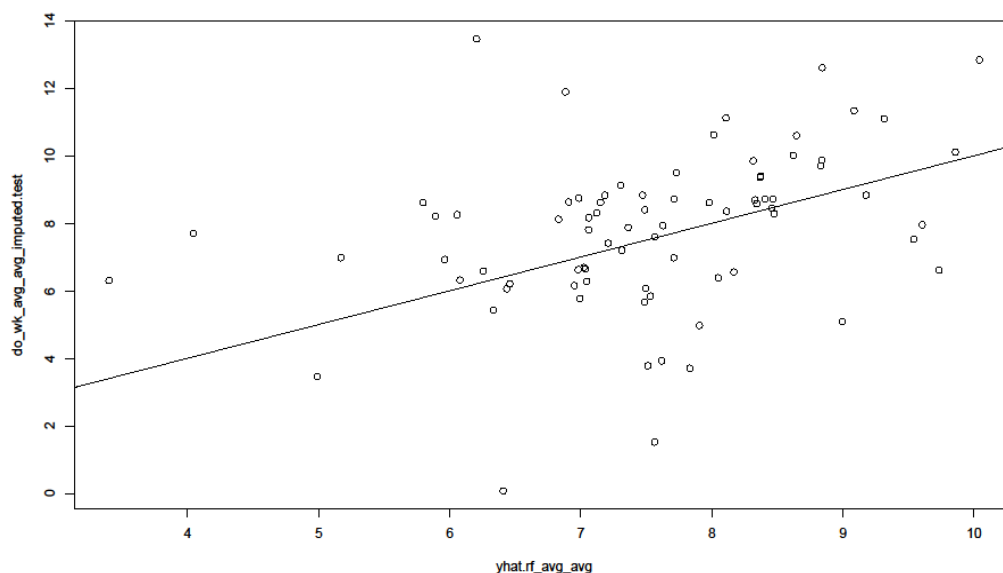


**Figure 2.** Plot of the best values of  $\mathbf{m}$  ( $m_{\text{try}}$ ) versus out-of-bag (OOB) error (prediction error).

Prediction error is found by removing one-third of the total observations ( $1/3$  of  $234 = 78$ ) from building the model and these are called out-of-bag (OOB) observations or the “test” observations. The prediction error is thus called OOB error. The remaining observations ( $2/3$  of  $234 = 156$ ) are called “training” observations. Sampling of observations (via bootstrap) *without* replacement was chosen to avoid a bias from selecting noninformative predictors with many categories more often than informative predictors with few categories (Boulesteix et al. 2012; Strobl et al. 2007).

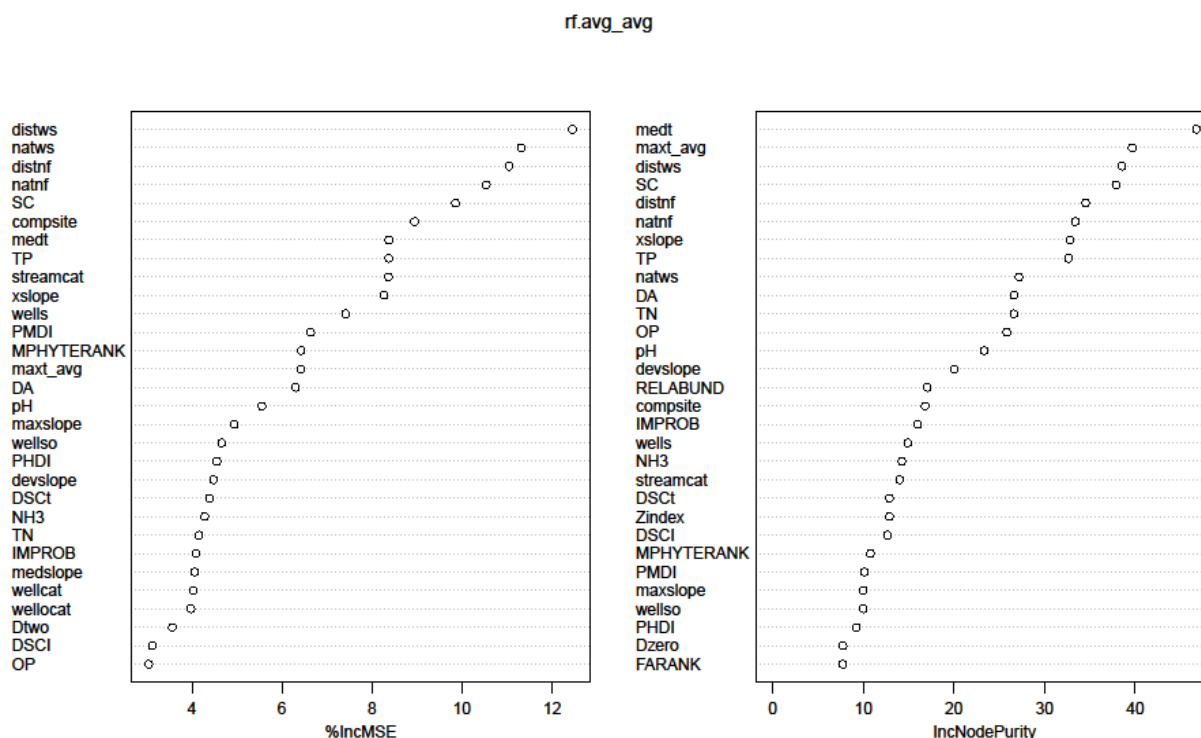
### Forest Model Results

Using the training observations, the forest model had a mean squared error (MSE) of 6.041 (root mean squared error [RMSE] = 2.46 mg/L) and explained 24.86% of the variance. The validation of the model (using the test observations) produced an MSE of 4.999 (RMSE = 2.236 mg/L) (**Figure 3**).



**Figure 3.** Plot of actual average weekly DO (y-axis) versus predicted DO using the OOB observations.

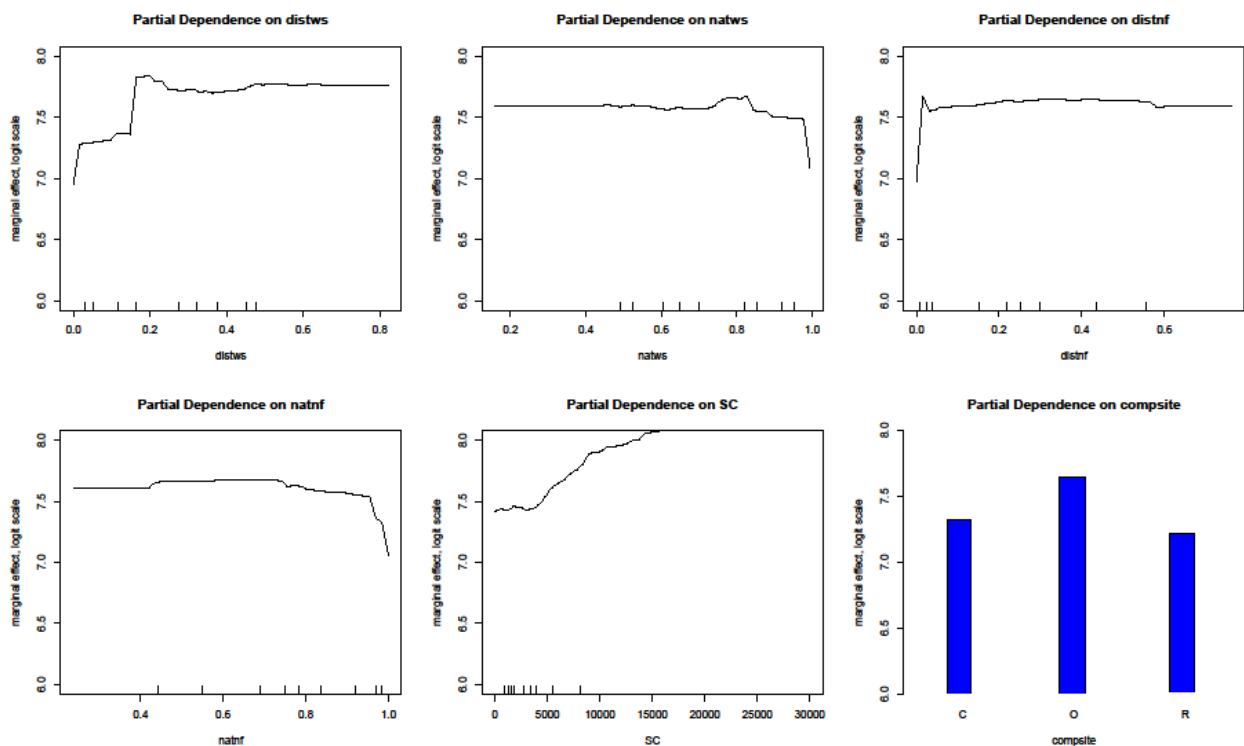
The most important predictor variables used in the model generation are first addressed because subsequent diagnostics are based on the important variables (**Figure 4**). The most variables occur near the top of the y-axis of each plot. The left-hand plot is the percent increase in MSE, a measure of explanatory power, for each predictor variable. This measure is based on the mean *decrease in accuracy in prediction* (of the non-training or OOB observations) when a given variable is excluded from the model. The right-hand plot is the incremental node purity for each predictor variable. This measure is the total *decrease in node impurity* that results from splits over that predictor variable, averaged over all trees. Node impurity for regression trees (the response variable being numerical) is measured by the RSS (residual sum-of-squares) for the training set.



**Figure 4.** Variable importance measures %IncMSE (percent increase in MSE) and IncNodePurity (incremental node purity).

The percent increase in MSE is the most relevant importance measure and distws and natws, SC (specific conductance), and distnf were found to be most important. Distws, natws, and distnf are three land use measures either for disturbed or natural land cover type and whether watershed (ws) or nearfield (nf). For node purity, both medt and maxt\_avg (both temperature predictors) were most important followed by distws and distnf and SC.

A partial dependence plot gives a graphical depiction of the marginal effect of a predictor variable on the response (average weekly of daily average DO). **Figure 5** shows partial dependence plots for each of the first six most important variables valued by reduction in MSE. The DO response increases significantly when watershed disturbed land increases above 20 percent, and near-field disturbed land increases about 5 percent. Contrastingly, DO response declines significantly when natural land increases above 90-95 percent for both watershed and near-field areas. When specific conductance increases above 5000 ( $\mu\text{S}/\text{cm}$ ), DO response increases significantly. Lastly, when the site is either a comparison or reference site, DO response is slightly lower (around 7.3 mg/L) compared to ordinary sites (around 7.7. mg/L). Partial dependence plots for the remaining 34 predictor variables are also shown in **Appendix A-GLEC** (below).

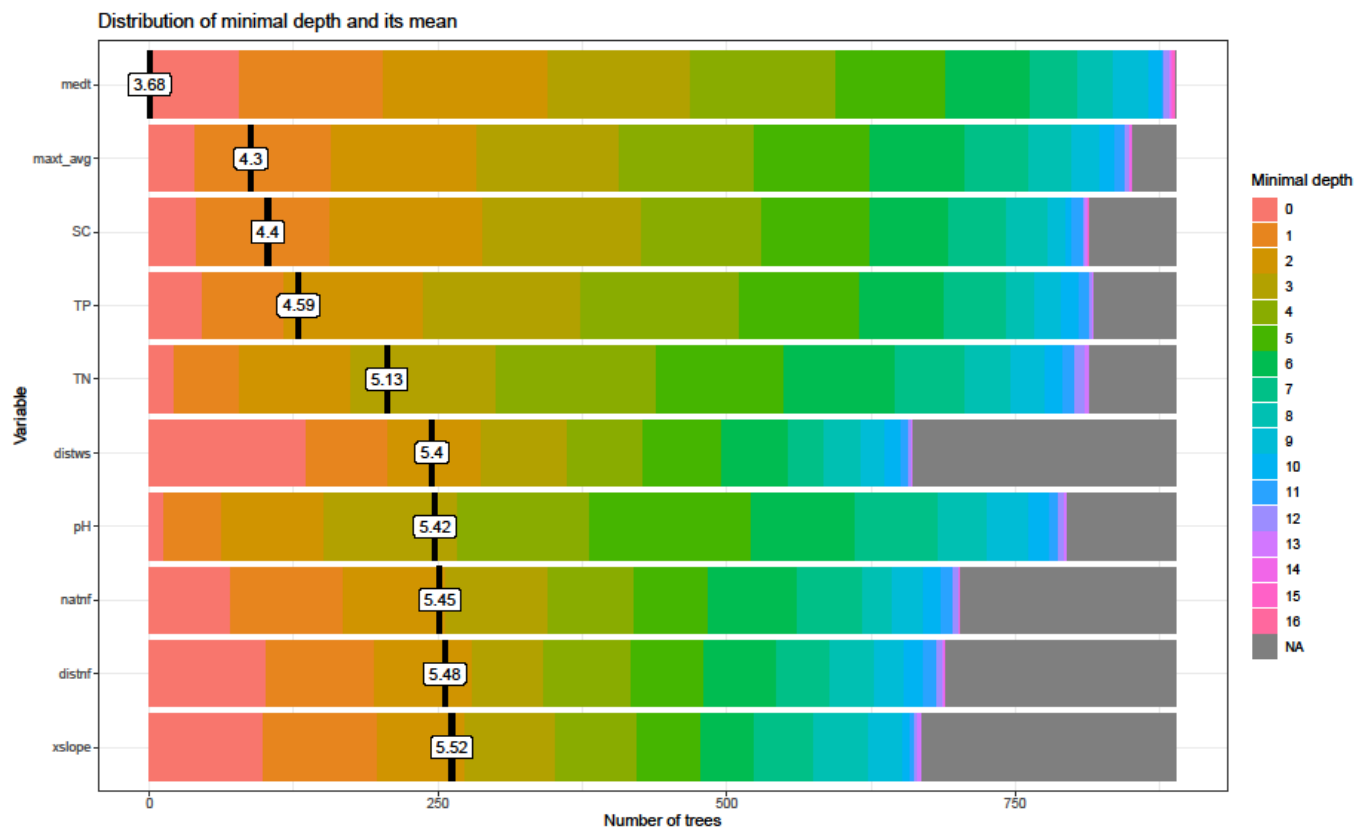


**Figure 5.** Partial dependence plots for each of the first six important variables: distws, natws, distnf, natnf, SC, and compsite.

#### *Distribution of Minimal Depth of the Trees in the Random Forest*

The minimum depth distribution of each predictor variable indicates the number of trees that a predictor participated (**Figure 6**). Note that the depth of a tree is equal to the length of the longest path from root to leaf in this tree. The x-axis ranges from zero trees to the maximum number of trees in which any variable was used for splitting, which is in this case equal to 1000 and is reached by only by the variable medt. The variables are ordered in the plot by their mean minimal depth where medt and happens to participate in all trees. This joint achievement is not always the case.

Examining the whole distribution of minimal depth offers more insight into the role that a predictor plays in a forest in contrast to looking only at the mean, especially as the mean can be calculated in more than one way. **Figure 6** is limited to only those variables that participate in at least 100 (of 1000) trees. Adding this requirement avoids selecting variables that have been by chance used for splitting once at the root.



**Figure 6.** Plot of minimal depth distribution for the top predictor variables. Mean minimal depth value shown by solid black line with mean value inside box. Only variables that participate in at least 100 (of 1000) trees are shown.

*Various Predictor Variable Importance Measures*

**Table 1** presents a summary of several importance measures for each predictor variable used in the random forest model on average weekly of average daily DO. The variables are ranked in **Table 1** by the *mse\_increase* measure. The following measures in **Table 1** are defined here:

**mse\_increase**: mean increase of mean squared error after this predictor variable is permuted (i.e., excluded from the model). *mse\_increase* is a measure of the decrease in predictive accuracy of the forest after perturbation of the predictor variable.

**node\_purity\_increase**: mean node purity increase by splits on this predictor variable, as measured by the decrease in sum of squares.

**mean\_minimal\_depth**: mean minimal depth for this predictor variable (calculated in one of three ways specified by the parameter *mean\_sample*).

**no\_of\_trees**: total number of trees in which a split on this predictor variable occurs.

**no\_of\_nodes**: total number of nodes that use this predictor variable for splitting (it is usually equal to *no\_of\_trees* if trees are shallow).

**times\_a\_root**: total number of trees in which this predictor variable is used for splitting the root node (i.e., the whole sample is divided into two based on the value of the predictor variable).

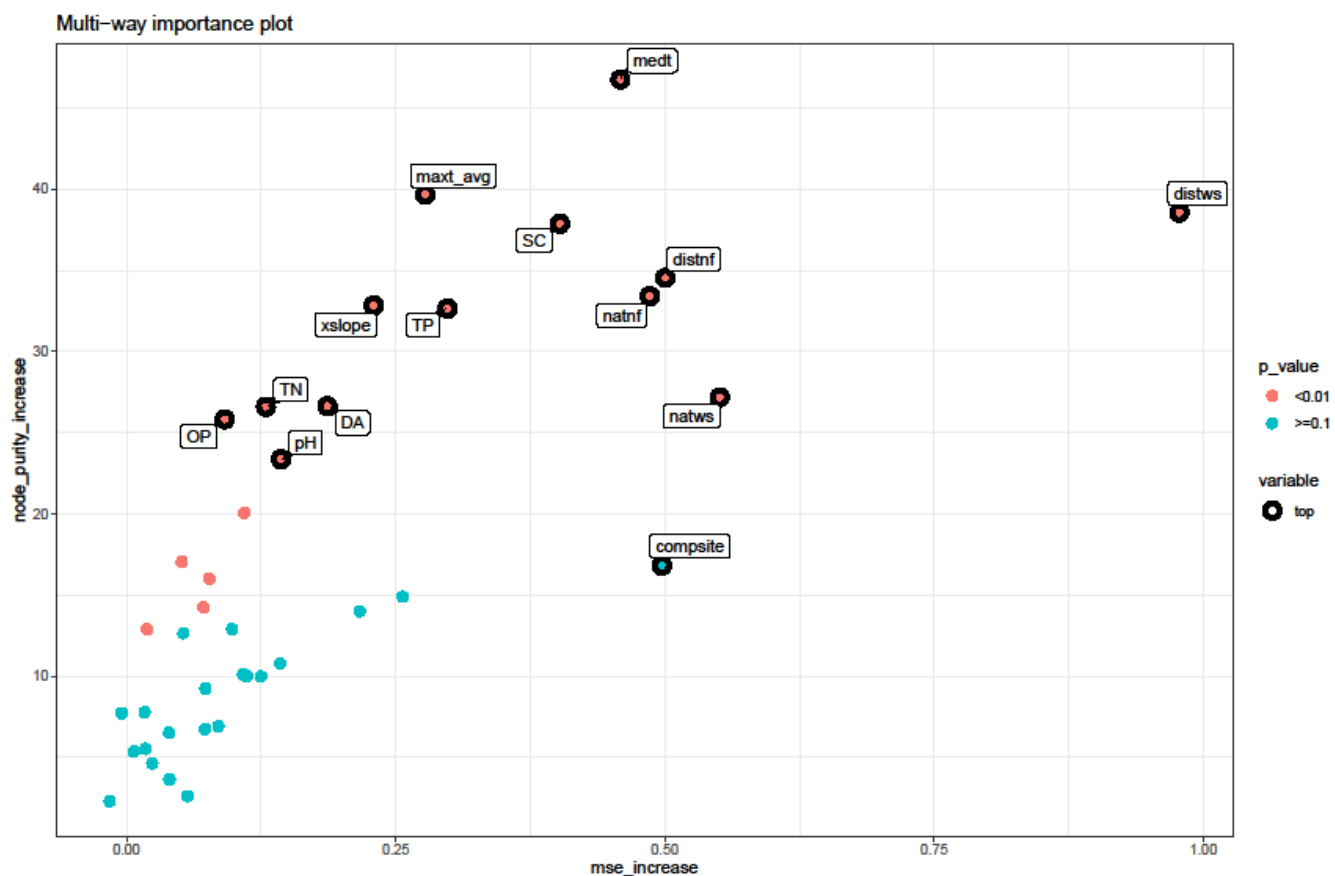
**p\_value**: p-value for the one-sided binomial test; this test indicates whether the observed number of successes (number of nodes in this predictor variable used for splitting) exceeds the theoretical number of successes if they were random (i.e. following the binomial distribution).

A multi-way importance plot (**Figure 7**) can be used to compare two of the above importance measures in a scatterplot. In **Figure 7**, "*node\_purity\_increase*" is included to reflect the structure of the forest and "*mse\_increase*" is included to reflect the accuracy of prediction. By examining importance measures in two dimensions, an improved identification of most important variables is gained and here shows that *distws* ranks high in both node purity and mse dimensions. Also note that five water quality measures (SC, TP, TN, OP, and pH) rank in the top 13 most important variables (of 40 total). Points in red (**Figure 7**) reflect low p-values of a binomial test of whether the predictor variable has been successful node for splitting a tree.

A way to compare all of the importance measures together is shown in **Figure 8**. The most un-related importance measures (i.e., most dispersed and having a low correlation coefficient) are then identified and used in a subsequent (2<sup>nd</sup> version) of a multi-way importance plot to select top predictor variables. Here, in **Figure 8**, we find *no\_of\_nodes* and *mse\_increase* to have low correlation ( $r = 0.445$ ); they are in turn used as importance measures in the multi-way importance plot in **Figure 9a**. In this comparison, all of the land use predictors perform well on both measures as does *medt*. In addition, from **Figure 8**, *number of times the predictor appears as a root* and the *no\_of\_nodes* also show low correlation. These appear in **Figure 9b** and show similar patterns as **Figure 9a**.

**Table 1.** Summary of variable importance measures for each predictor variable used in the random forest model on average weekly of average daily DO.

predictor variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
distws	5.40	1095	0.978	38.56	661	136	5.18E-18
natws	5.86	1017	0.551	27.18	654	76	7.73E-10
distnf	5.48	1121	0.500	34.53	690	101	2.73E-21
compsite	7.75	444	0.497	16.82	404	45	1
natnf	5.45	1198	0.486	33.40	703	71	1.51E-32
medt	3.68	1932	0.459	46.75	889	78	4.13E-236
SC	4.40	1521	0.402	37.87	815	41	1.63E-102
TP	4.59	1595	0.298	32.64	818	46	3.54E-123
maxt_avg	4.30	1745	0.277	39.66	852	40	8.69E-170
wells	7.62	564	0.257	14.89	449	44	1
xslope	5.52	1040	0.230	32.83	669	99	5.52E-12
streamcat	7.45	629	0.217	14.01	500	11	1
DA	5.58	1207	0.187	26.63	710	24	5.19E-34
pH	5.42	1478	0.144	23.37	795	12	2.98E-91
MPHYTERANK	7.72	660	0.143	10.79	500	0	1
TN	5.13	1568	0.130	26.60	815	21	1.88E-115
wellso	8.42	432	0.125	9.98	352	16	1
maxslope	8.16	571	0.112	10.01	450	3	1
devslope	6.53	934	0.109	20.06	606	34	5.27E-04
PMDI	7.58	746	0.108	10.13	546	10	1.00E+00
DSCt	7.81	588	0.098	12.90	440	17	1
OP	5.64	1257	0.091	25.83	730	10	1.12E-42
wellcat	9.11	309	0.085	6.91	278	22	1
IMPROB	6.17	1275	0.077	16.02	738	0	5.10E-46
PHDI	7.79	735	0.073	9.23	523	9	1.00E+00
medslope	8.66	462	0.073	6.74	376	3	1
NH3	6.84	1007	0.071	14.26	645	1	5.63E-09
wellocat	9.93	174	0.057	2.63	165	2	1
DSCI	7.71	555	0.053	12.66	441	10	1
RELABUND	6.02	1376	0.051	17.06	762	0	7.79E-67
Dtwo	9.37	294	0.040	3.64	265	2	1
MARANK	8.30	581	0.040	6.51	445	1	1
Done	9.03	358	0.024	4.65	321	4	1
Zindex	7.05	913	0.019	12.89	610	5	5.17E-03
Dthree	9.02	313	0.018	5.53	287	0	1
Dzero	8.49	495	0.017	7.77	386	5	1
Nox	8.57	597	0.007	5.35	455	0	1
Dfour		0	0.000	0.00	0	0	1
FARANK	8.47	471	-0.005	7.72	388	0	1
MATHICKMM	9.67	276	-0.016	2.32	240	1	1



**Figure 7.** Multi-way importance plot showing the comparison of the *mse increase* and the *node purity increase* for each predictor variable. Plotting points with a surrounding black circle are the top 13 variables as ranked by the importance measures on each axis and the color differentiation of each point reflects its p-value.

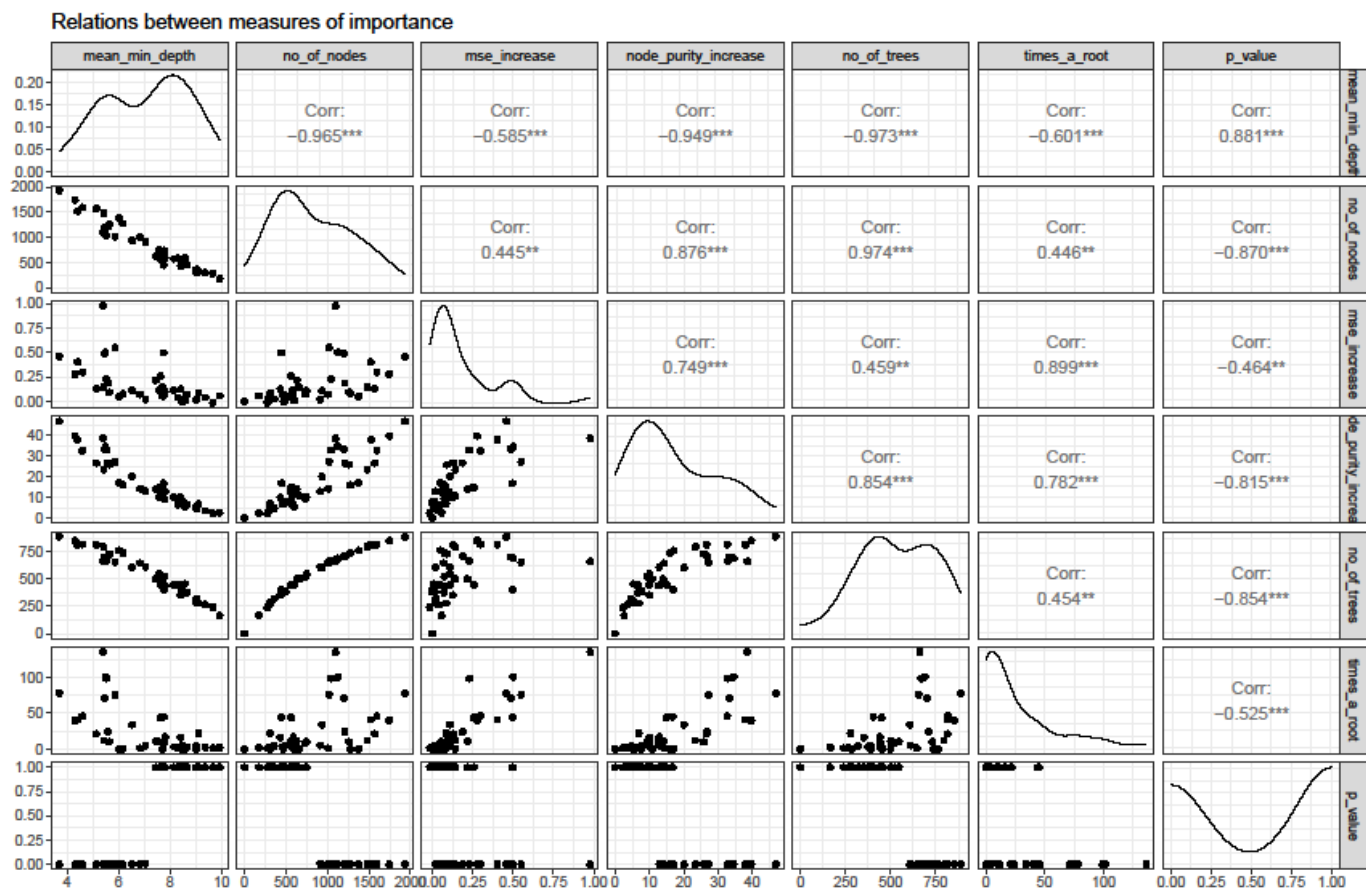
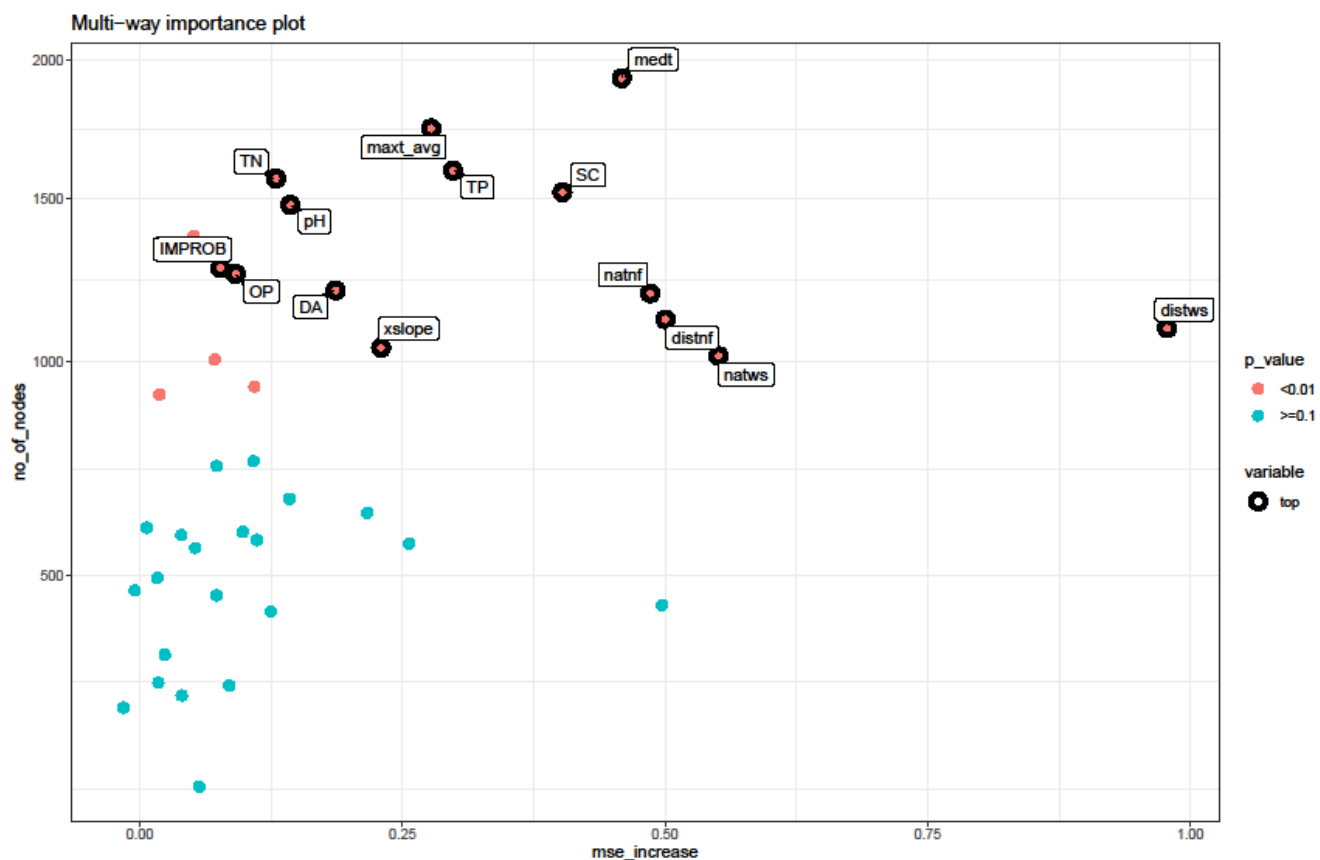
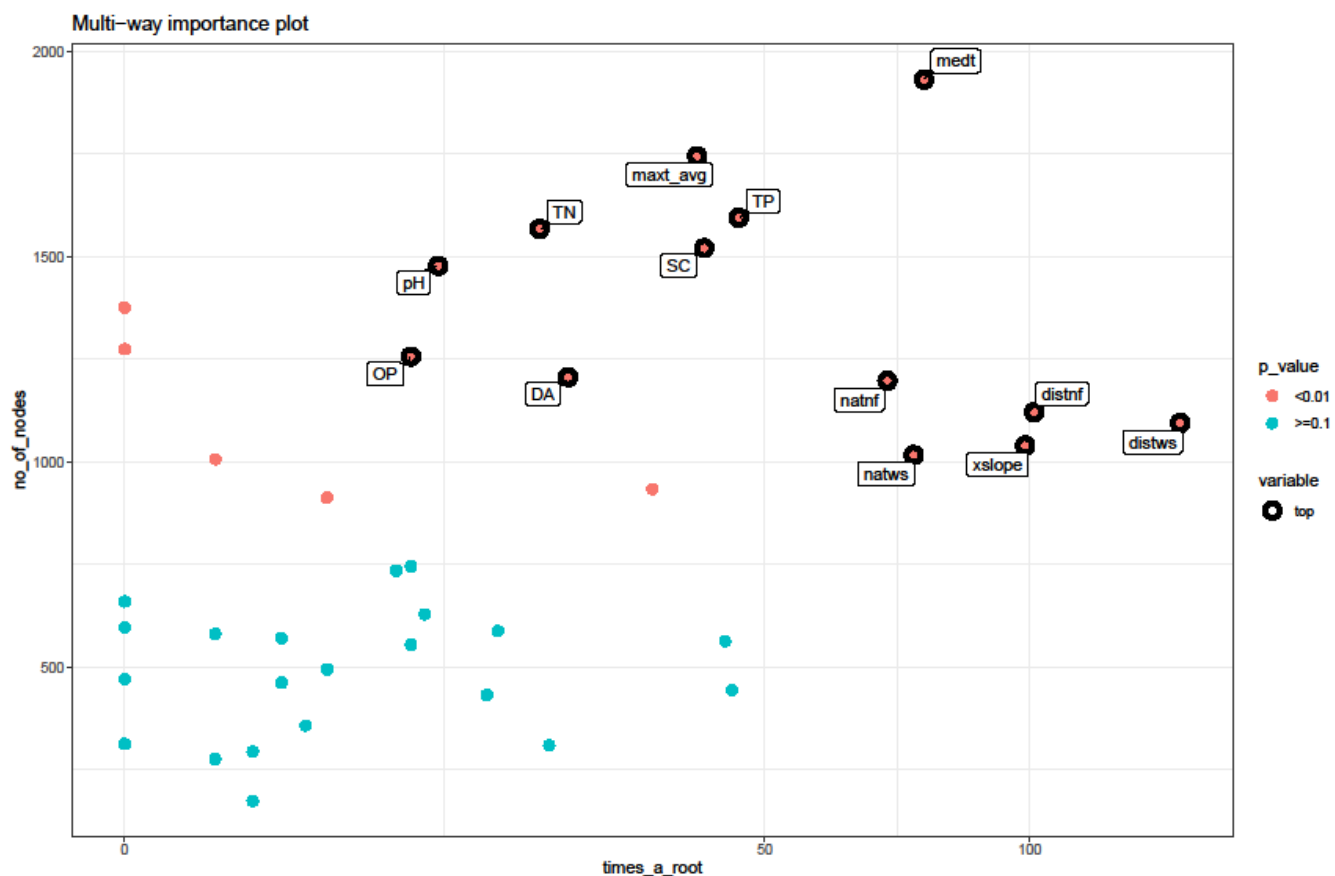


Figure 8. Scatterplot of the seven variable importance measures along with the correlation coefficient between each and the probability density function of each measure on the main diagonal.



**Figure 9a.** A second run of a multi-way importance plot showing the comparison of the *mse increase* and the *number of nodes* for each predictor variable. Plotting points with a surrounding black circle are the top 13 variables as ranked by the importance measures on each axis and the color differentiation of each point reflects its p-value.



**Figure 9b.** A second run of a multi-way importance plot showing the comparison of the *number of times the predictor appears as a root* and the *number of nodes* for each predictor variable. Plotting points with a surrounding black circle are the top 13 variables as ranked by the importance measures on each axis and the color differentiation of each point reflects its p-value.

*Predictor Variable Interactions*

After selecting a set (in this approach 20) of most important predictor variables according to `mse_increase` and `node_purity_increase` (**Figure 7**), the interactions between them can be examined. **Table 2** shows the most frequently occurring interactions along with the mean minimal tree depth and the unconditional mean minimum depth. The 10 most frequently occurring interactions are highlighted in grey and also presented in **Figure 10**. Note that the interaction is ordered (from left-to-right) by decreasing number of occurrences. The most frequent interaction (`medt:RELABUND`) has the second highest mean unconditional depth (**Figure 10**).

One interesting aspect of interactions is that they often can contain predictor variables that are not the most important variables (e.g., RELABUND). Which suggests that variables that are important in maximal subtrees with respect to top predictors are not important in whole trees. This is what we would expect from an interaction: some predictor variable is only important when conditional on another one.

**Table 2.** Summary of predictor variable interactions, ranked by number of occurrences in the forest model.

predictor variable	root variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
RELABUND					
D	medt	2.518	328	medt:RELABUND	6.020
SC	medt	2.457	310	medt:SC	4.404
maxt_avg	medt	2.819	307	medt:maxt_avg	4.301
TN	medt	2.771	307	medt:TN	5.128
pH	medt	2.859	306	medt:pH	5.422
TP	medt	2.655	305	medt:TP	4.592
IMPROB	medt	2.946	300	medt:IMPROB	6.169
medt	medt	2.965	300	medt:medt	3.684
medt	SC	2.952	293	SC:medt	3.684
pH	maxt_avg	2.810	280	maxt_avg:pH	5.422
TP	maxt_avg	2.833	271	maxt_avg:TP	4.592
medt	maxt_avg	3.101	268	maxt_avg:medt	3.684
TN	SC	3.015	268	SC:TN	5.128
RELABUND					
D	maxt_avg	3.197	267	maxt_avg:RELABUND	6.020
maxt_avg	SC	3.370	265	SC:maxt_avg	4.301
TP	SC	3.162	265	SC:TP	4.592
medt	distws	3.336	264	distws:medt	3.684
maxt_avg	distws	3.401	258	distws:maxt_avg	4.301
maxt_avg	maxt_avg	3.097	257	maxt_avg:maxt_avg	4.301
medt	TP	2.957	257	TP:medt	3.684

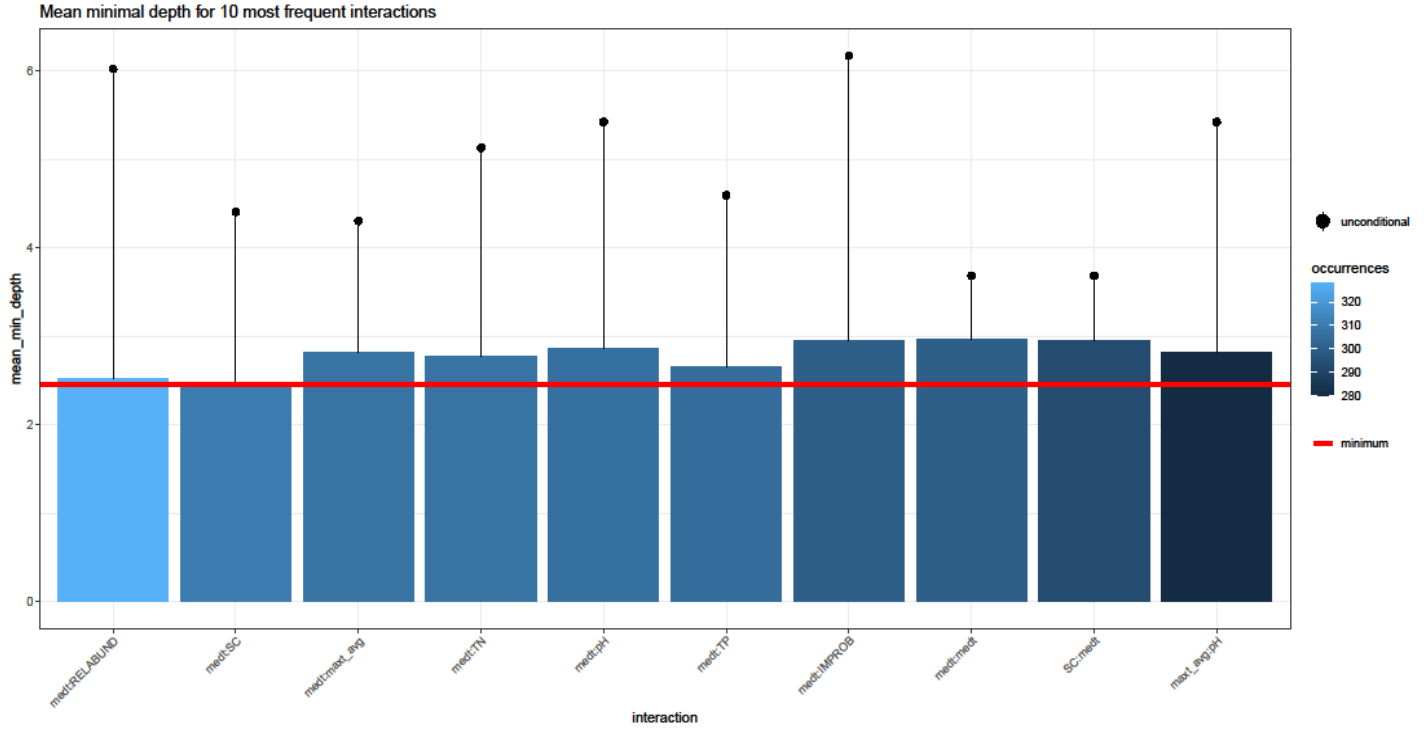
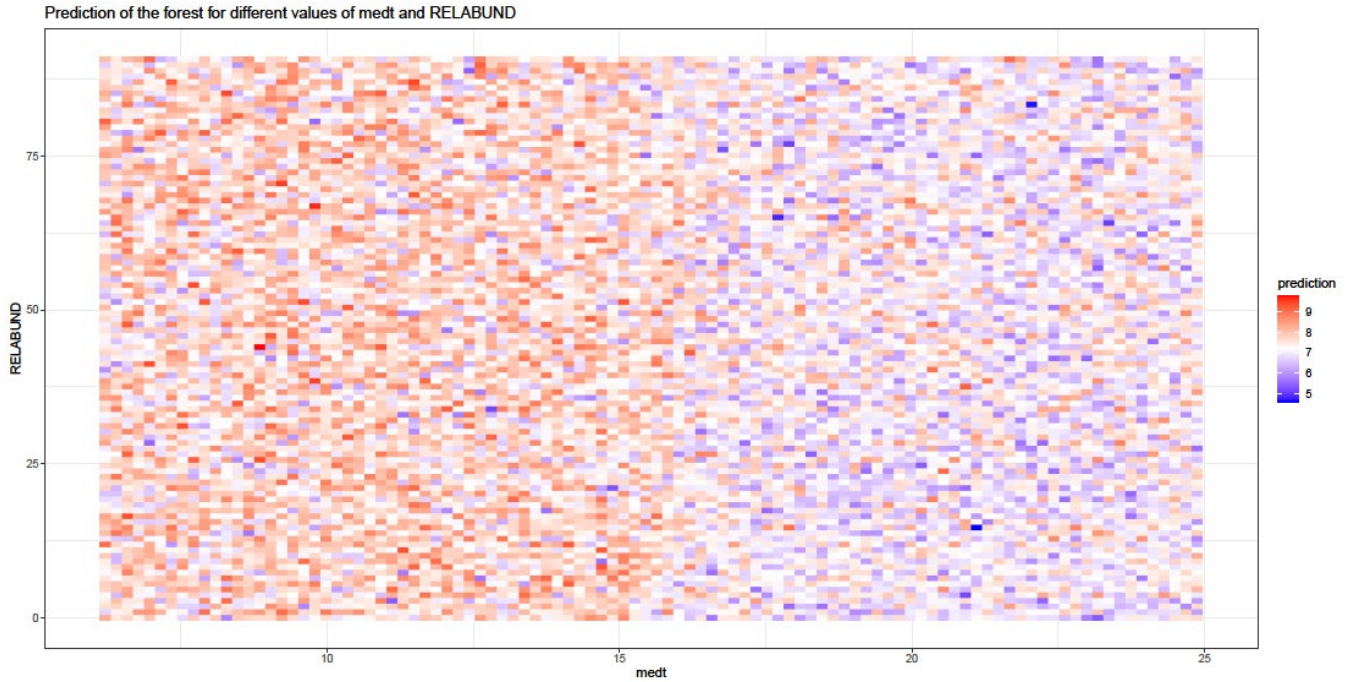


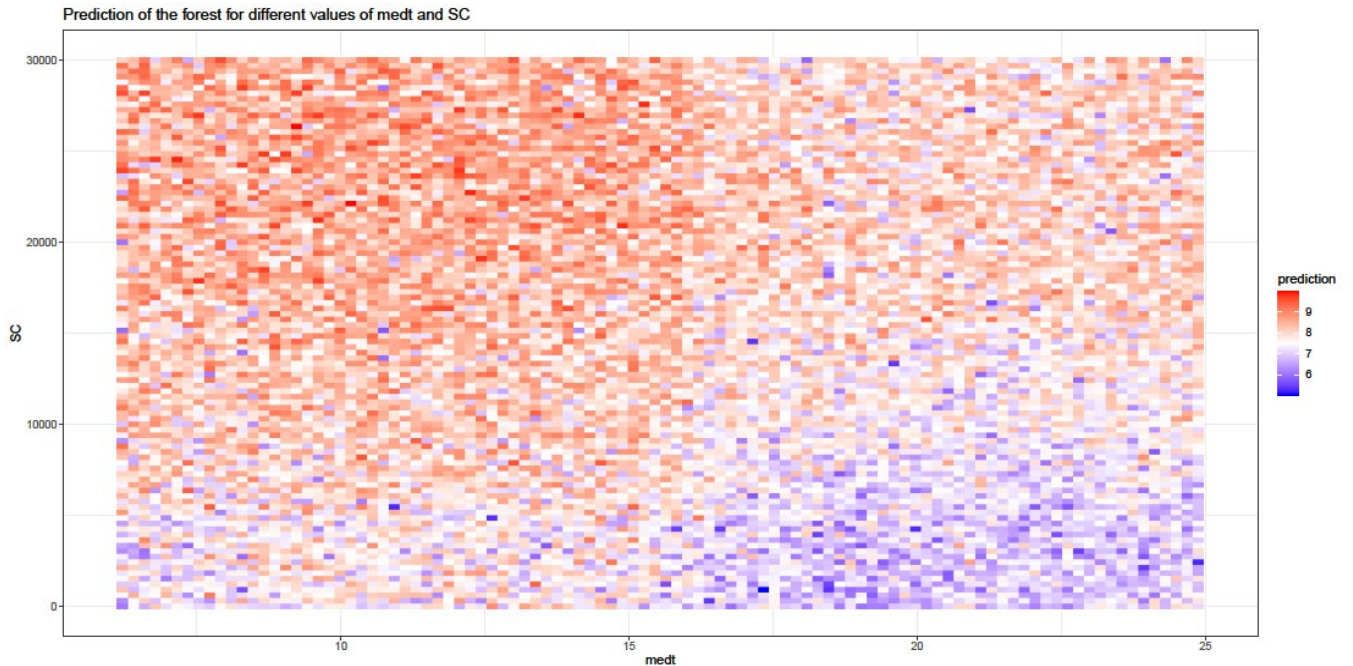
Figure 10. Mean minimal depth for the 10 most frequent interactions.

*Prediction of the Forest on a Grid*

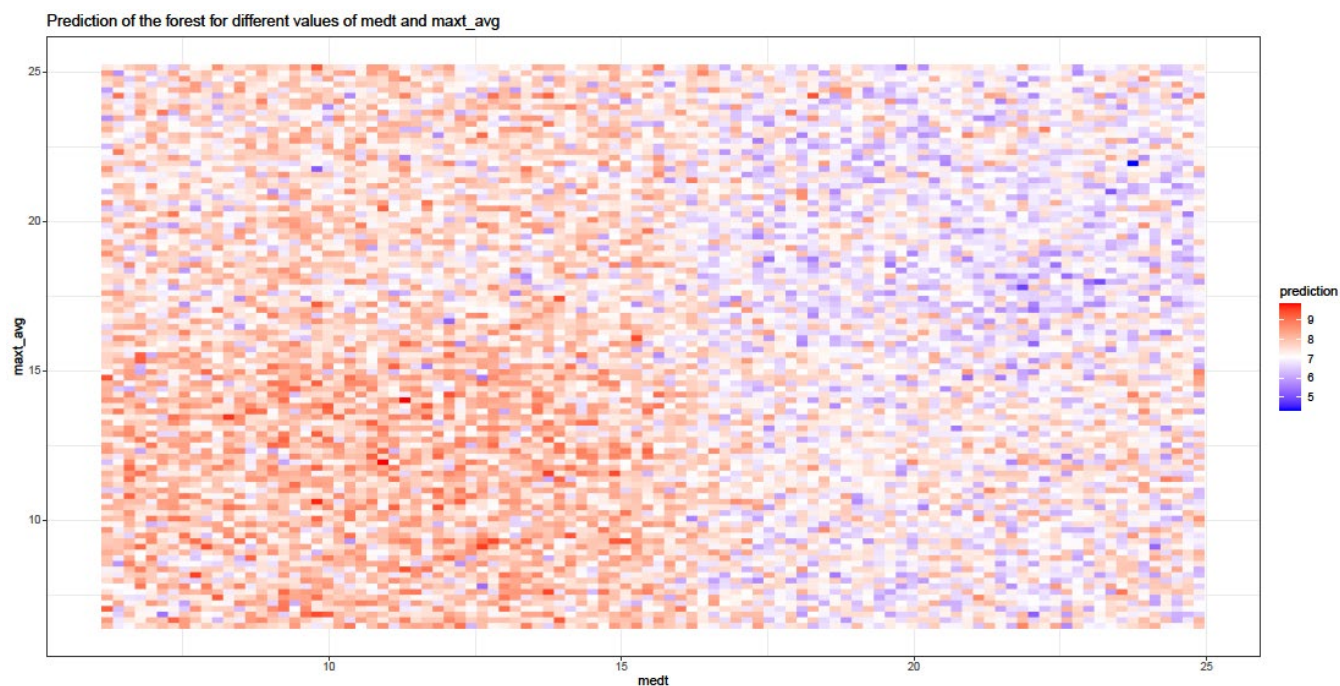
To further investigate the most frequent interactions of predictor variables in the random forest, a prediction grid is presented for pairs having the most frequent interactions (**Figures 11a through 11f**). The prediction grid is plot of the prediction of the response variable (i.e., average weekly DO average) shown as color intensity (red is high DO concentration while blue is low) with values of the pair of variables on each axis. Figure 11b shows the highest differential of predicted DO where concentration is highest with high values of SC and low values of medt.



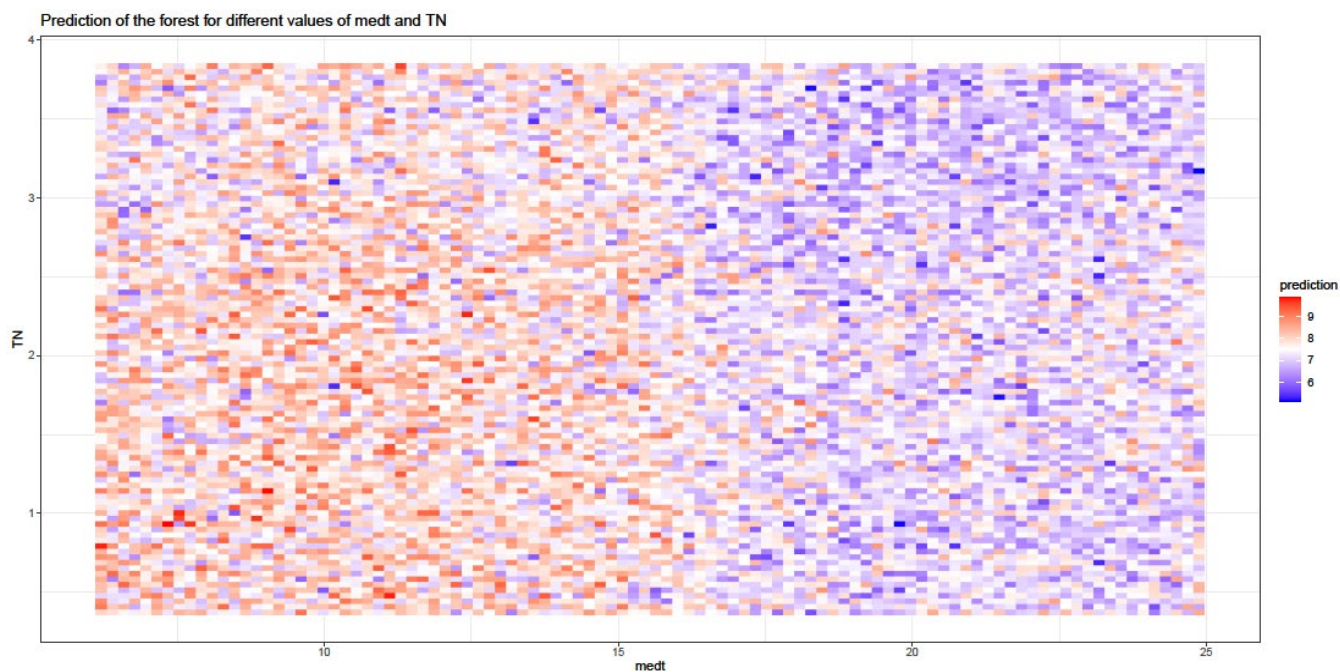
**Figure 11a.** Prediction grid (of average weekly DO daily average) based on the values of RELABUND and medt.



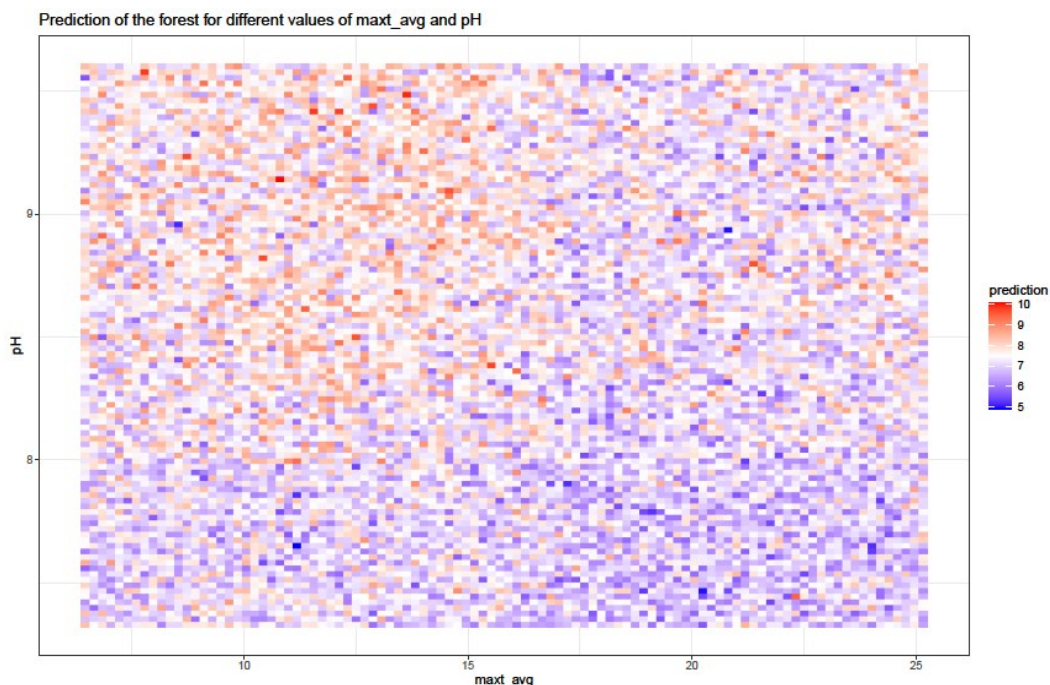
**Figure 11b.** Prediction grid (of average weekly DO daily average) based on the values of SC and medt.



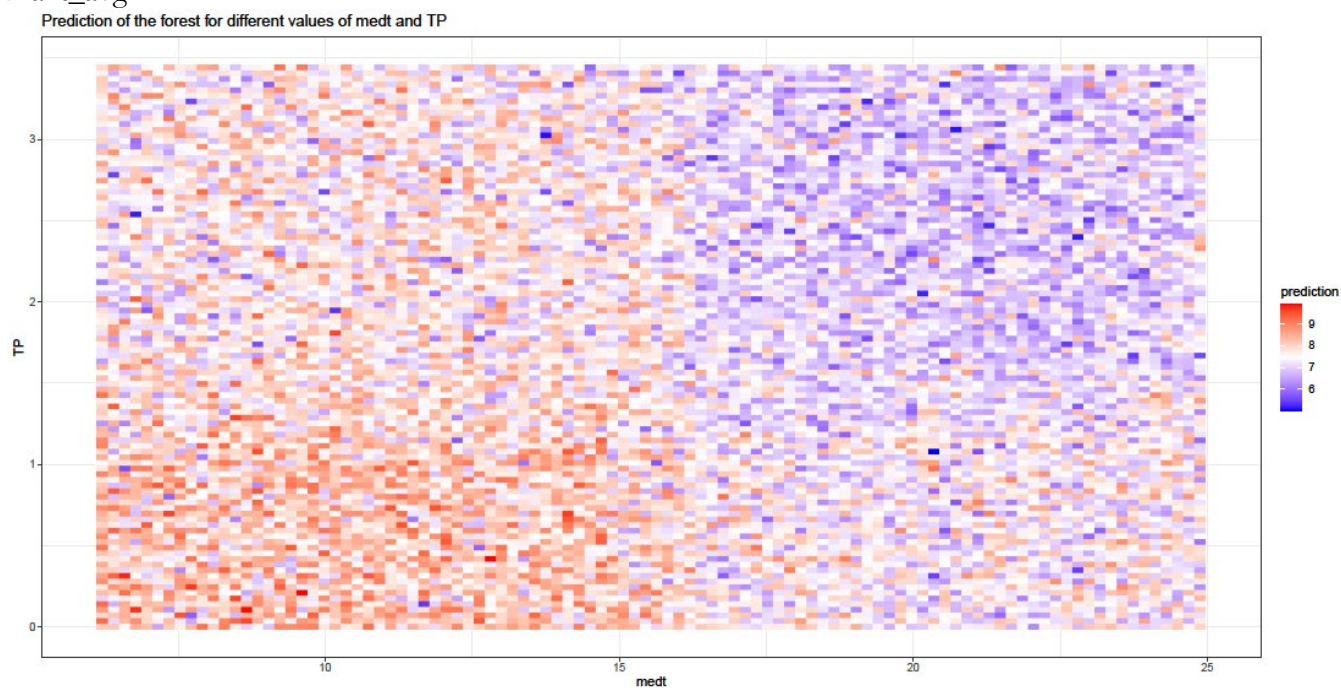
**Figure 11c.** Prediction grid (of average weekly DO daily average) based on the values of maxt\_avg and medt.



**Figure 11d.** Prediction grid (of average weekly DO daily average) based on the values of TN and medt.



**Figure 11e.** Prediction grid (of average weekly DO daily average) based on the values of pH and maxt\_avg.



**Figure 11f.** Prediction grid (of average weekly DO daily average) based on the values of TP and medt.

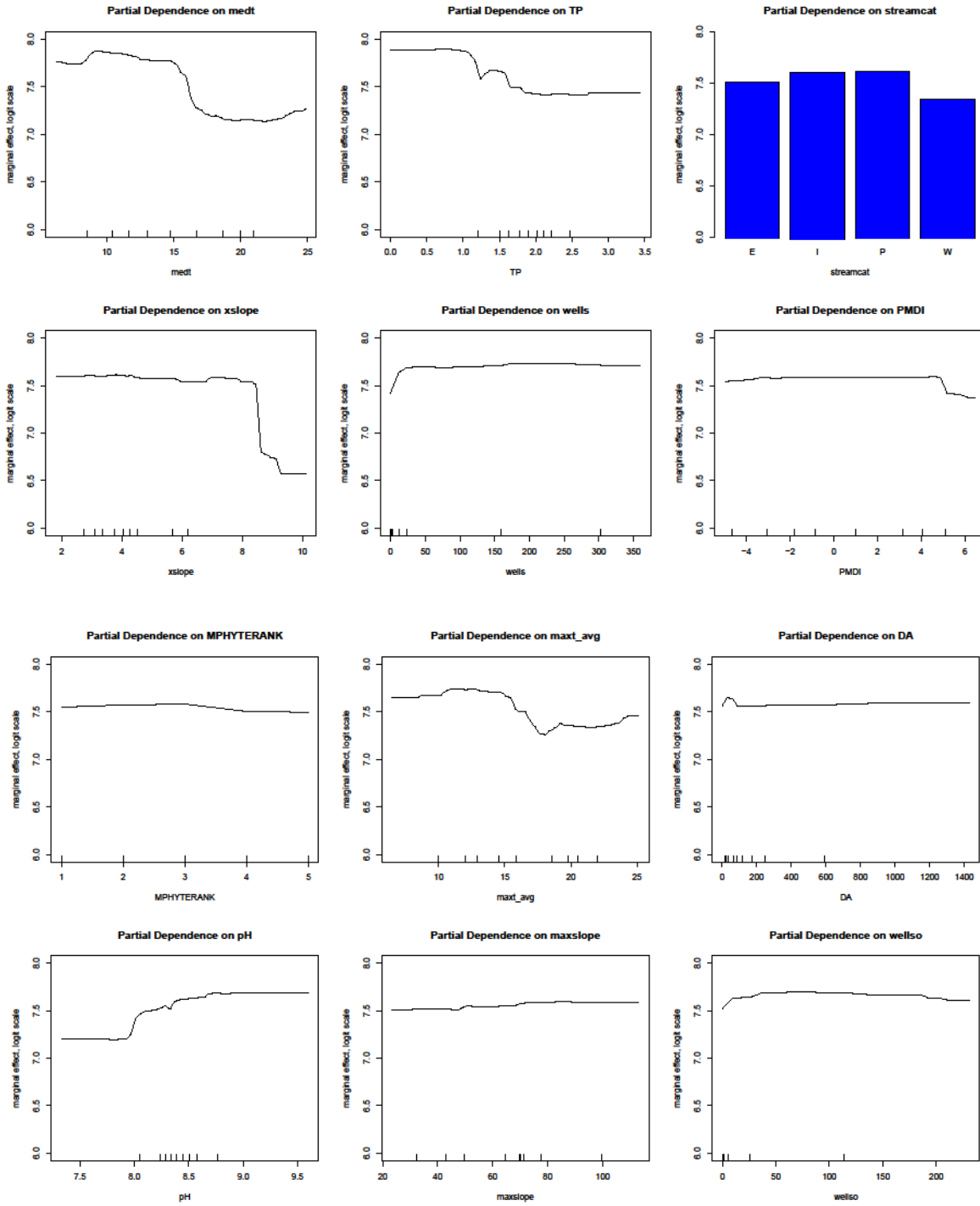
## Conclusions

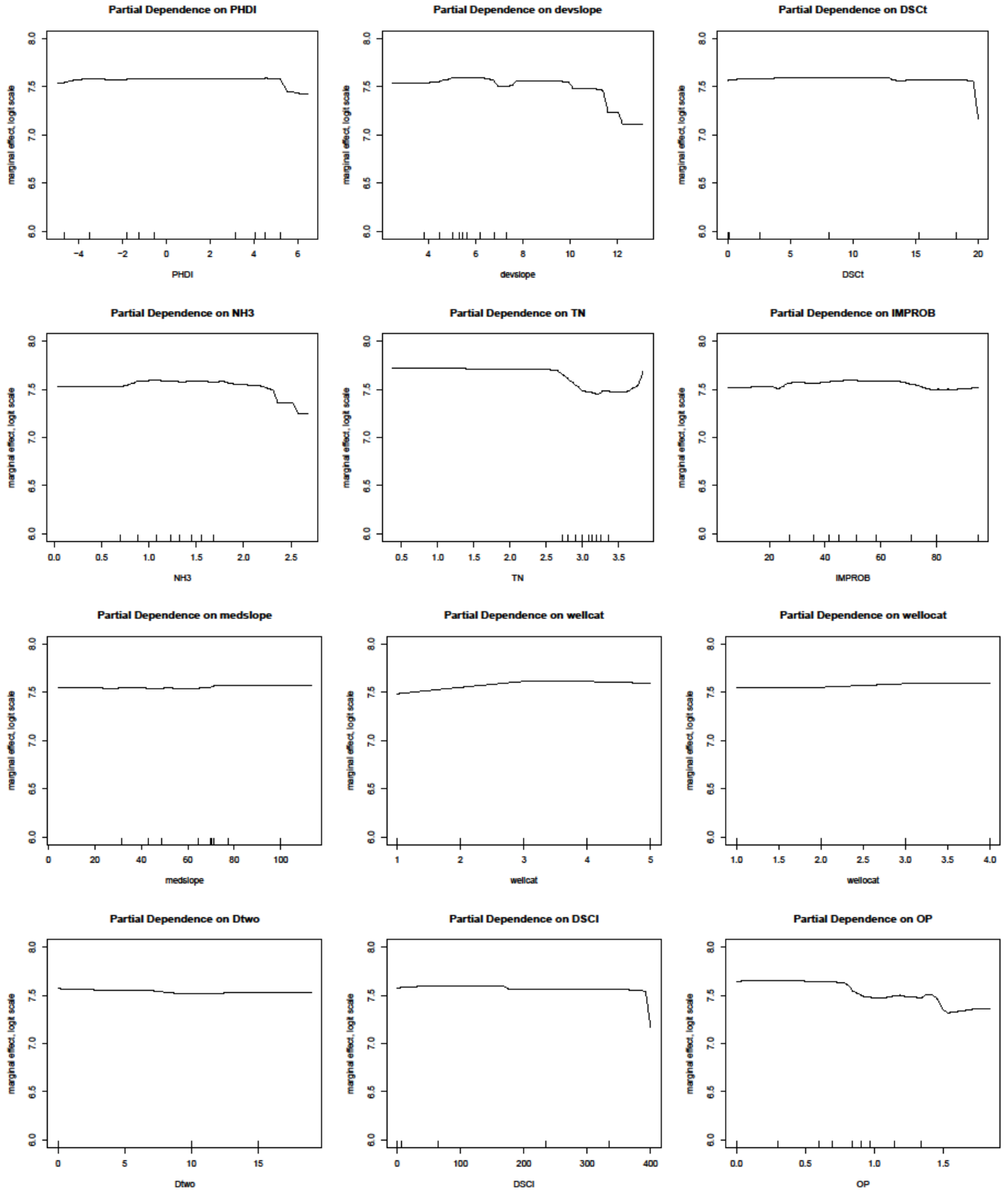
Future effort would consider a smaller set of predictor variables for building the random forest. For example, select only the most important predictors (top 20-25 in both node purity and mse increase) and/or all components of interactions that appeared at least 5-10 times in the forest.

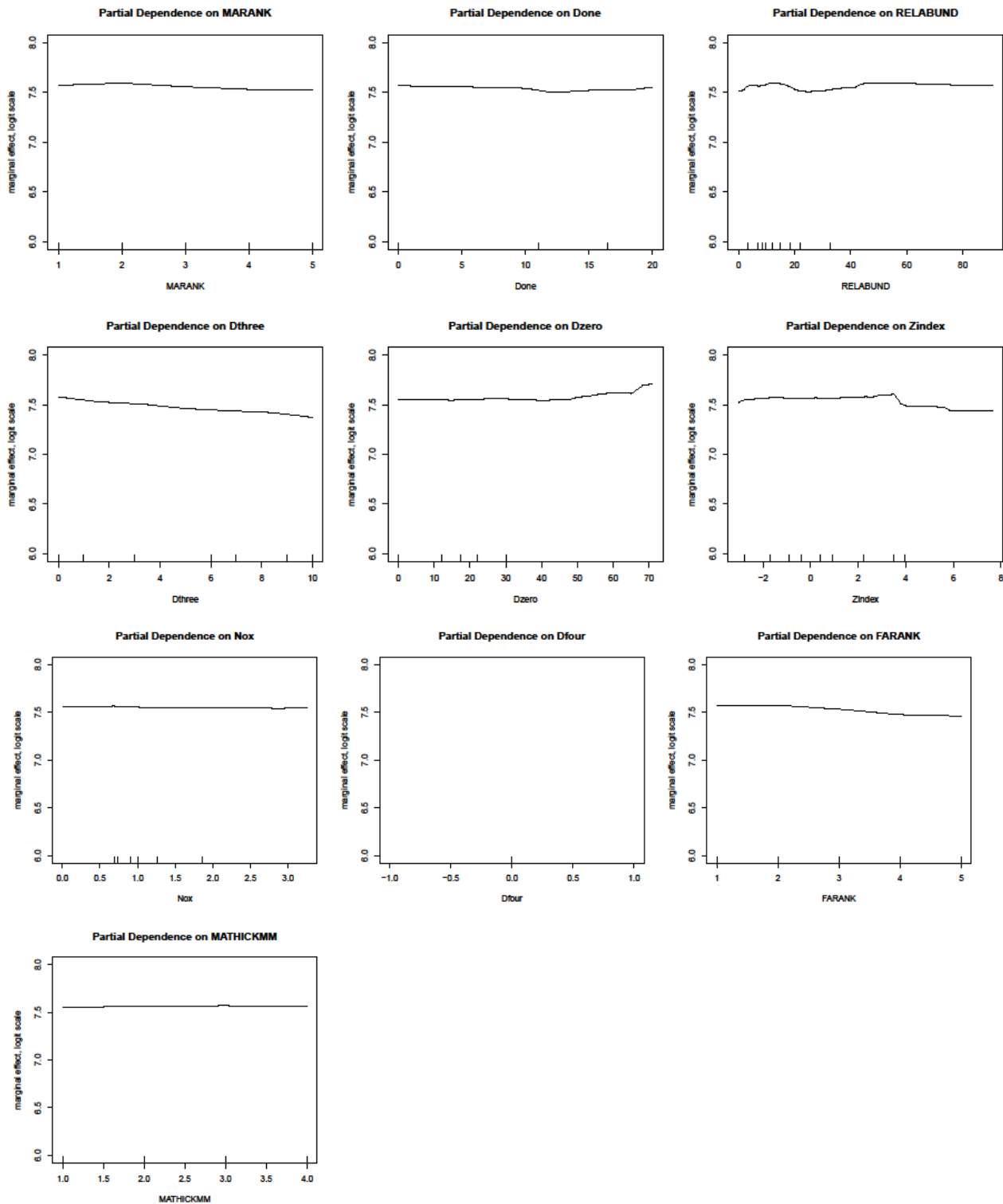
Consider other approaches:

1. Conditional Inference Forests— see Strobl et al. (2007) and Hothorn et al. (2006)
2. Boosted Regression Trees
3. Bayesian Additive Regression Trees

Appendix A-GLEC: Partial Dependence Plots for Remaining Predictor Variables







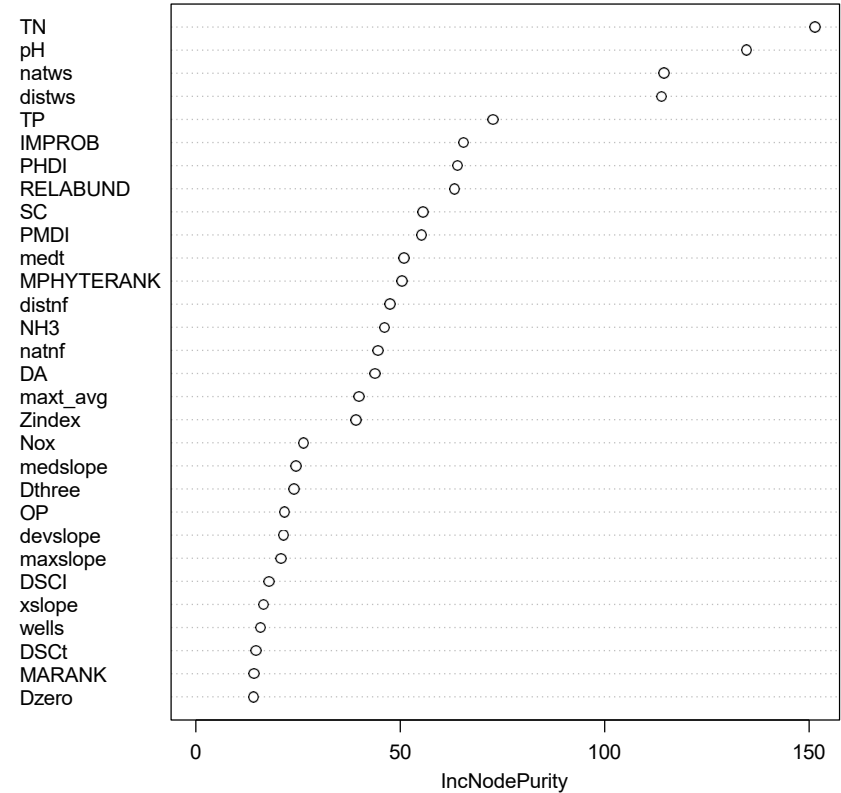
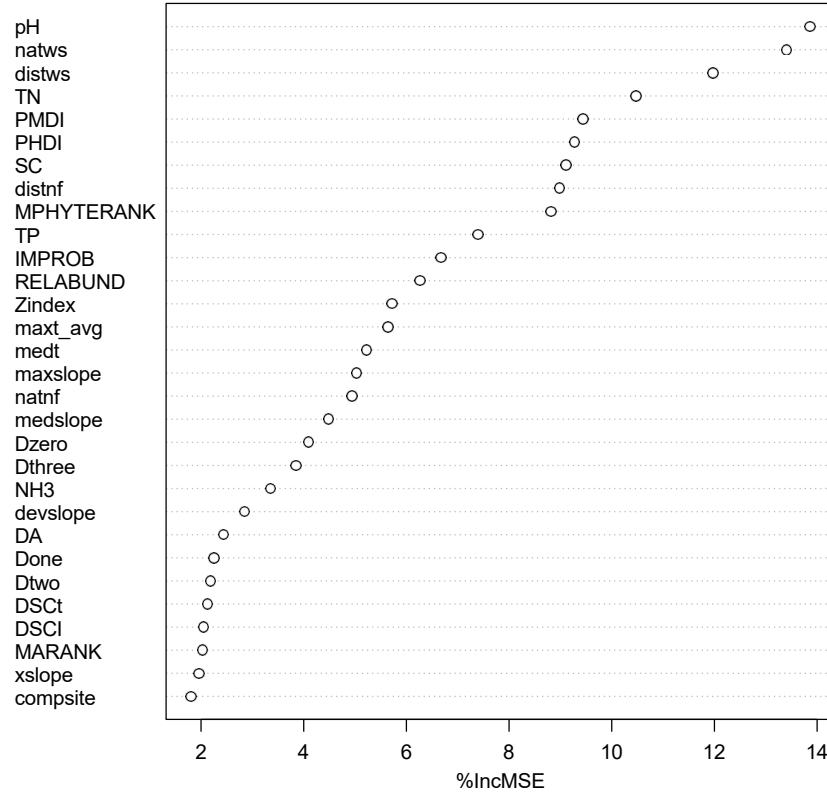
## Data Dictionary for All Variables Used in Modeling Effort.

Variable Name	Definition	Data Type (in R)
<b>Response Variable</b>		
avg_avg	average of weekly suite of daily average DO (mg/L); continuous monitor measurement (with DO & temp)	numeric
<b>Factor Variables (fixed effect)</b>		
DA	drainage area (sq.mi)	numeric
streamcat	stream category: P: perennial, I: intermittent, E: ephemeral, W: wetland	factor
compsite	reference site type: C: comparison site, R: MT DEQ official reference site, O: ordinary site	factor
<b>Aquatic Plant Predictor-Response Variables (random effect)</b>		
MARANK	% cover of micro-algae (as rank)	factor
MATHICKMM	micro-algae thickness (mm)	numeric
FARANK	% cover of filamentous algae (as rank)	factor
MPHYTERANK	% cover of macrophyte (as rank)	factor
RELABUND	% relative abundance of periphyton (nutrient enricher taxa)	numeric
IMPROB	impairment probability (%); indicates nutrient or sediment problem when > 51%	numeric
<b>Water Chemistry Predictor Variables (random effect)</b>		
medt	median of weekly suite of daily averages for water temperature; continuous monitor measurement (with DO & temp)	numeric
maxt_avg	maximum of weekly suite of daily averages for water temperature; continuous monitor measurement (with DO & temp)	numeric
NH3	total ammonia including NH3 and NH4+ (mg/L)	numeric
Nox	nitrite + nitrate (NO23) (mg/L)	numeric
OP	orthophosphate (mg/L)	numeric
pH	pH (international units)	numeric
SC	specific conductance ( $\mu$ S/cm)	numeric
TN	total nitrogen (mg/L)	numeric
TP	total phosphorus (mg/L)	numeric
<b>Drought Indices Predictor Variables (random effect)</b>		
Zindex	NOAA Z-Index (drought)	numeric
PMDI	NOAA Palmer Meteorological Drought Index (drought)	numeric
PHDI	NOAA Palmer Hydrological Drought Index (drought)	numeric
DSCI	NDMC Drought Severity and Cover Index -- weighted sum of D0-D4	numeric
DSCt	transformed (square-root) DSCI	numeric
Dzero	# consecutive weeks at drought severity level D0 (source: NDMC)	integer

<b>Variable Name</b>	<b>Definition</b>	<b>Data Type (in R)</b>
<b>Done</b>	# consecutive weeks at drought severity level D1 (source: NDMC)	integer
<b>Dtwo</b>	# consecutive weeks at drought severity level D2 (source: NDMC)	integer
<b>Dthree</b>	# consecutive weeks at drought severity level D3 (source: NDMC)	integer
<b>Dfour</b>	# consecutive weeks at drought severity level D4 (source: NDMC)	integer
<b>Land Use/Cover Predictor Variables (fixed effect)</b>		
<b>natws</b>	% natural land cover (watershed scale)	numeric
<b>distws</b>	% disturbed land cover (watershed scale)	numeric
<b>natnf</b>	% natural land cover (near-field scale; <1k m)	numeric
<b>distnf</b>	% disturbed land cover (near-field scale; <1k m)	numeric
<b>Well - Oil and Gas - Predictor Variables (fixed effect)</b>		
<b>wells</b>	total well count within watershed	integer
<b>wellso</b>	old (pre-1990) well count within watershed (previously wellsn)	integer
<b>wellcat</b>	total count all wells (as rank)	factor
<b>wellocat</b>	total count of old wells (as rank) (previously wellnew)	factor
<b>Watershed Slope Predictor Variables (fixed effect)</b>		
<b>maxslope</b>	maximum slope in watershed (% slope = tangent x 100)	numeric
<b>medslope</b>	median slope in watershed (% slope = tangent x 100)	numeric
<b>xslope</b>	mean slope in watershed (% slope = tangent x 100)	numeric
<b>devslope</b>	standard deviation (sample, N-1 ?) of slope in watershed (% slope = tangent x 100)	numeric

**APPENDIX B. ADDITIONAL ANALYSIS ON DO Δ BY THE GREAT LAKES ENVIRONMENTAL CENTER (APRIL 2024). FACTORS NEAR THE TOP AND RIGHT PROVIDE THE STRONGEST EXPLANATORY POWER.**

rf.xdelta



## APPENDIX C. METHOD FOR COMPUTING A DROUGHT SEVERITY INDEX

Agencies within National Oceanic and Atmospheric Administration (NOAA) and the US Department of Agriculture (USDA) teamed with the National Drought Monitoring Center (NDMC) to produce a weekly US Drought Monitor Index (DMI) product that incorporates climatic data and professional input from all levels (Svoboda, 2000). Since no single definition of drought works in all circumstances, the DMI authors rely on the analyses of several key indices and ancillary indicators from different agencies to create a final index (Heim, 2002). Key parameters (**Table 2-2**, this document) include the Palmer Drought Index (PMDI), the Crop Moisture Index, soil moisture model percentiles, daily streamflow percentiles, percent of normal precipitation, topsoil moisture (percent short and very short) generated by the USDA, and a satellite-based Vegetation Health Index. The ancillary indicators include the Surface Water Supply Index, the Keetch–Byram Drought Index, the Standardized Precipitation Index, snowpack conditions, reservoir levels, groundwater levels determined from wells, USDA reported crop status, and direct *in situ* soil moisture measurements.

### C.1 DETERMINING THE AREA- AND TIME-WEIGHTED DROUGHT INDEX

As discussed in **Section 2.1** of this document,  $\leq 6$  consecutive weeks at  $D_{ZERO}$  are non-drought periods, while  $>6$  consecutive weeks at  $D_{ZERO}$  are drought periods; this drought criterion should be used when assessing DO  $\Delta$  datasets from eastern Montana.

Drought severity and longevity data can be downloaded by county as a comma-separated Excel file at <https://droughtmonitor.unl.edu/DmData/DataDownload/WeeksInDrought.aspx>. Make sure to select  $D_{ZERO}$  and a time-period corresponding to the period when the DO measuring instrument was deployed instream; all Montana counties will be downloaded for the period you select.

In some cases, the watershed you are evaluating will be contained within one or more counties which are all experiencing the same drought level (i.e., they all have either  $\leq 6$  consecutive weeks at  $D_0$  or  $>6$  consecutive weeks at  $D_0$ ). **In this case no further geospatial analysis is necessary, you can conclude that the prevailing drought conditions for those counties and times apply to your dataset during the time period identified.**

However, it may result that a watershed is split between counties experiencing different drought conditions during the same time period and a more sophisticated GIS method is needed. Data need to be aggregated over different areal extents and time periods relative to the sampling station and their associated drainage areas, as detailed next. An appropriately scaled USGS HUC can be used to delineate the watershed in question; 8 to 12-digit HUCs will be the most appropriate.

The number of consecutive weeks at drought severity level  $D_0$  in a watershed is computed as a weighted sum where the weights represent the percent area of the specified county existing within the drainage basin polygon. This integration is represented as:

$$\# \text{ consecutive weeks at drought severity level } D_0 = [\% \text{ area in } CNTY_a \times \text{number of consecutive weeks at drought severity level } D_{0_{CNTY_a}}] + [\% \text{ area in } CNTY_b \times \text{number of consecutive weeks at drought severity level } D_{0_{CNTY_b}}] + [\% \text{ area in } CNTY_c \times \text{number of consecutive weeks at drought severity level } D_{0_{CNTY_c}}] + [\text{etc.}]$$

where **# consecutive weeks at drought severity level D0** is the weighted drought index for a specific watershed, and CNTY a, b, and c are three counties that intersect the boundary of the watershed. More counties can be included as needed ([etc.]). The % area in CNTY<sub>a</sub> is the percent of the watershed total area in CNTY<sub>a</sub>, % area in CNTY<sub>b</sub> is the percent of the watershed total area in CNTY<sub>b</sub>, etc., for as many counties as may be intersecting with the watershed. Also note that  $\sum \% \text{ area}_{[a,b,c, \text{etc.}]} = 100$ .

The resulting spatial-temporal integrated number of consecutive weeks in the watershed at drought severity level D0 can then be compared to the 6-week cutoff.