Utah Department of Water Quality 2016 Draft Integrative Report Comment Letter

To:
Utah Department of Water Quality
Salt Lake City, UT

From:
David C. Richards, Ph. D.
OreoHelix Consulting
Vineyard, UT
Phone: 406.580.7816
Email: oreohelix@icloud.com

Date:
August 28, 2016

RE: Utah Department of Water Quality 2016 Draft Integrative Report Response Letter

I would like to thank UDWQ for extending the timeline for written comments on their 2016 IR draft. UDWQ has done a tremendous job in trying to evaluate and protect Utah's valuable water resources and it is reflected in this draft. However, I have some comments that may prove helpful in the next revision of the draft and in particular on how biological evaluations are presently being conducted. Hopefully UDWQ is in the process of revising their biological assessment program to better reflect the state of science and address the pitfalls of reliance on RIVPAC O/E models.

In addition, I am enclosing two draft technical reports that I submitted to the Jordan River/Farmington Bay Water Quality Council:

1. "Is reliance on a single bioassessment metric for assessing water quality in Utah's rivers and streams prudent?".
2. "Real and Perceived Macroinvertebrate Assemblage Variability in the Jordan River, Utah can Effect Water Quality Assessments".

Please consider these two attachments as more detailed, integral parts of my comments on the draft.

The following are my general comments:

Chapter 1. Introduction.

Figure 4. Page 14.

Comment: Eliminate the colored boxes at the bottom of the figure. They distract the reader into thinking the columns are assessment categories and not the rows.

Chapter 2: 2016 303(d) Assessment Methods.

Page 37. Addressing Nitrogen and Phosphorus.

Comment: Too little nutrients in a waterbody particularly phosphorus can lead to nuisance algal blooms such as *Didymosphenia geminata* (Didymo = 'rock snot'). There is a large amount of literature supporting this and UDWQ has been informed of this at several technical committee meetings.

In addition, many beneficial uses such as waterfowl, fisheries, phytoplankton, zooplankton, benthic invertebrates, and even the resource extractive brine shrimp industry in GSL depend on an ample supply of nutrients in the system. Too few nutrients could easily reduce primary and secondary productivity in the "food chain". More nutrients may be needed under certain conditions and waterbodies.

Comment: In the third full paragraph you cite Ostermiller et al. 2014 but it is not in the Literature Cited section of Chapter 2. This appears to be an important reference. Perhaps it is the document you are directing readers to on the DWQ website?

Page 50. Biological Assessments. First sentence.

Comment: In addition to protecting cold and warm water fish species, Utah's beneficial use also requires the protection of non-game fish and other aquatic life not just those necessary in the food chain. There is a major difference between "other aquatic life" and "those necessary in the food chain". In addition, the term food chain has not been used in ecology for several decades. The correct term is 'food web'. This misuse of terminology reflects the antiquated unrevised definitions of beneficial use in light of advances in our understanding of ecology and its continued use could negatively reflect on the departments understanding of modern ecological concepts.

Utah's beneficial use also protects waterfowl, shore birds, other water-oriented wildlife and the organisms on which they depend. I did not see a discussion on these groups of animals in the Biological Assessments section, nor on any group of fishes including, cold water, warm water or non-game. Other states incorporate fish IBIs. Is UDWQ planning on developing fish or bird IBIs?

Page 51. First sentence in 3rd paragraph.

Comment: You introduce a term, "biological integrity" without a definition. There are many definitions of biological integrity most of which UDWQ is cursorily familiar with.

From previous conversations that I have had with UDWQ it appears UDWQ personnel are using a very simplified definition to fit agenda needs, i.e. bioassessment output. Biological integrity is not a measurable attribute but an abstract idea, similar to "human health". There is no one measure of biological integrity (particularly O/E) just as there is no one measure of human health. I often use the analogy of visiting a doctor and the only measure the doctor uses to assess my health is body temperature. If the physician only used this one measure to assess my health I would immediately seek another more qualified one, and eventually in all likelihood the physician would lose his/her license. Just a reminder; bioassessments do not quantify integrity, they are only an indicator.

Page 51. Last sentence in 3rd paragraph.

Comment: I don't think using a single taxa richness based metric, RIVPACS O/E would constitute a robust index of biological integrity. It is only one metric that does not address anything other than richness and apparently does not do an adequate job of that (See Attachments). There is also no reason to make a 'robust IBI' easily interpretable. Ecological interactions between dozens of organisms and their responses to human caused impairment are anything but easily interpretable. RIVPACS O/E models themselves are not easily interpretable. By using the term 'robust' you are misleading the public.

Page 51. River Invertebrate Prediction and Classification System Models. 2nd paragraph.

Comment: The three western state the IR uses to support their use of RIVPACs, Colorado, Montana, and Wyoming also include dozens of other metrics needed to address ecological complexities. These states use RIVPACS as just one in their suite of metrics with no special weight given to RIVPACS. Thus there is no justification for only using RIVPACS O/E in UDWQ bioassessment program. Please see the attached reports that discusses this further.

Page 52. First paragraph.

Comment: O/E absolutely does not quantify loss or local extinction of taxa. It only quantifies the failure to observe predicted taxa using limited sampling effort. In many cases, taxa were not lost, they just weren't found. These statements suggesting that local extinctions have occurred are highly misleading to Utah's citizens and suggest that UDWQ personnel do not have a full understanding of the RIVPACs models. Please review the attached reports for additional comments of this critically important concern, particularly the section, "Misinterpretation of O/E' in the Discussion.

Page 52. 2nd paragraph

Comment: Although O/E may have an intuitive biological meaning, there are so many assumptions, generalizations, and errors associated with derivation of results that its accuracy in assessing loss of taxa and impairment is highly questionable. There are several other diversity metrics in use throughout the world that are much simpler to derive and interpret than RIVPACS O/E. These metrics can easily substitute for O/E or at least supplement it. For example, richness and evenness are better indicators than O/E

for several reasons, 1) they are not confounded with other models (e.g. PRISM, a costly and proprietary model that is not transparent except for those who can afford to pay for its use), 2) they are independently verifiable, and 3) they allow assessment of change at local-scale due to point source impacts. Please see section, "Additional Bioassessment Metrics in Use" in the Discussion in the attached draft report.

Page 53. First complete paragraph.

Comment: There apparently are no direct, real world, reference site(s) to compare with Jordan River, Green River, Colorado Rivers, or any stream or river in UT. Only generalized, regionwide, summary, and averaged hypothetical reference sites. For example, the Jordan River's source is Utah Lake, a shallow remnant of Lake Bonneville and its terminus is the Great Salt Lake. Historically the Jordan River had a wide meandering braided channel that migrated across its valley. These conditions make the Jordan River a truly unique river and I assume there is no real world reference river in the state only reference conditions based on averaged watershed values. The Green River downstream of Flaming Gorge Reservoir should not be considered a reference site if UDWQ has chosen to do so. The Green River is a highly regulated river and does not resemble its condition prior to construction of the dam. Of course, the Colorado River does not have any other river(s) to compare with in Utah and no hypothetical reference rivers and "E" scores should be used on such a national treasure.

Also, was the same "E" in the O/E model used for the entire length of the Jordan River? Hopefully not. Obviously, the Jordan River habitat changes from its upstream sections to downstream and the macroinvertebrates reflect this change. Using the same 'E' for the entire Jordan River would be cause for concern.

It would be helpful if the final IR included a table of reference streams used to develop O/E and an appendix with additional model values including "E" taxa.

Page 53. 2nd complete paragraph.

Comment: Calculating 'E' using a probability of capture (P_c) of $\geq$50% is extremely problematic and results in a poor assessment of biological integrity. Taxa with P_cs < 50% are likely the most sensitive taxa and the very taxa that respond to impairment more that those with Pc > 50%. The statement that "Using a Pc limit set at greater than 50% typically results in models that are more sensitive and precise, which results in a better ability to detect biological stress" is based on two relatively limited studies that evaluated precision using their own methods, i.e. circular reasoning and these were hardly typical. UDWQ is setting a precedent by using Pc > 50% based on results that are not solidly supported in the literature and not established scientific fact but based on a vague ill-defined term in the two studies, 'sensitivity'. Please review section, RIVPACS O/E 'Probability of Capture' is Problematic' in the Discussion in the attached report for more discussion on P_c.

Page 54. Table 8.

Comment: These predictor models are mostly watershed based. It is highly commendable that UDWQ is assessing biological integrity at the watershed level rather than at the region wide level which it has done in the past. By assessing biological integrity at the watershed level more accurate and precise conclusions will be made. However, watershed averages are just that, averages. Macroinvertebrate assemblages can easily change from the top of a watershed to the bottom and an average value likely will not capture those responses.

In addition, I solicited comments from Mr. Brett Marshall, River Continuum Concepts, Manhatten, MT a leading authority on bioassessments. To summarize his comments:

a. As I discussed earlier, PRISM models are proprietary black box and as such are not independently verifiable and thus are scientifically invalid. The scientific method requires the possibility of independent validations. PRISM models are not reproducible or transparent, which as we all agree, is what we are all striving for.

b. PRISM models rely on historic data (e.g. most of the climate data metrics in Table 8). "Clearly the past has absolutely nothing to do with the macroinvertebrates collected next year. Similarly, the average of multiple years has nothing to do with invertebrate assemblages that are mostly multivoltine or univoltine. Their lives are shaped only by the conditions in the years during which they lived… not over multiyear averages. Variables in Table 8 had nothing to do with environmental conditions during the time when the sampled invertebrates lived. This introduces an unmeasurable and significant error to every Pcs calculated and prevents the use of field data, which would be site specific. It may have been useful in developing regional models… but it has no place in continued assessment/monitoring and should never be used. Only field measurements should be used".

c. PRISM data errors are also spatially derived mostly for miss use of regional models to monitor local scale changes. These models will complicate every O/E assessment conducted anywhere that there are natural gradients, introducing error in every local assessment including all of the assessments included in the IR.

Page 55. First paragraph.

Comment: Using updated models that accept data from first to eighth plus order rivers and stream at all seasons and a coarser taxonomic resolution can only reduce UDWQ's ability to detect impairment. Macroinvertebrate assemblage composition changes from season to season. An example of coarser taxonomic resolution effects would be the genus Baetis and family Baetidae (mayfly family). Both phylogenetic levels have species that can occur from the coldest headwaters to warmer lowland rivers and even in wetland ponds. Also, the primary goal is to improve biological integrity or the 'full suite of naturally occurring taxa that occur in a site'. Coarser taxonomic resolution eliminates this ability. Member of invertebrate families occur in almost all streams, from headwaters to valleys and often across all of North America. It is not possible to measure the integrity of a stream based on coarse taxonomic resolution.

Page 56. Last sentence.

Comment: The use of the 10^{th} and 5^{th} percentiles of reference site thresholds is completely arbitrary. The assessment categories need to be based on actual field measures of beneficial use and then those field derived percentiles used. For example, if the designated beneficial use is to support foraging waterfowl, then that threshold should be used. It appears that the 10^{th} and 5^{th} percentiles were not 'devised' by UDWQ but arbitrarily chosen.

Page 58. Assessment of Lakes and Reservoirs

Comment: It is well known that lakes and reservoirs are ecologically dissimilar. They should not be combined and compared using the same assessment criteria.

Page 71. Last paragraph. Last sentence.

Comment: All lakes eventually evolve into 'eutrophic conditions'. This is called lake succession and is inevitable. There is a big difference between eutrophic condition and eutrophication. The sentence should state, "……cyanobacteria may be indicative of 'eutrophication', not 'eutrophic condition'.

Chapter 5. Narrative Standard Assessment….. Application to Utah Lake.

Harmful Algal Bloom Assessment. Page 12. First sentence.

Comment: Utah Lake is no longer a functioning natural lake. It is considered a shallow, eutrophic, irrigation reservoir. Therefore, assessments and in particular, bioassessments should treat it as such. The same assessment methods and standards that UDWQ applies to natural lakes should not apply to Utah Lake.

Page 12. Recreational Uses in Utah Lake. First paragraph.

Comment: "….more recently, swimming and wading". If UDWQ reviews the Utah Lake Legacy book and video on YouTube, it will be clear that Utah Lake was historically used for swimming and not just recently.

Page 12. Recreational Uses in Utah Lake. 2nd paragraph.

Comment: The average number of visitors to Utah Lake State Park is not 253,599. That is the average number of visits. The park does not count the number of people in a vehicle and does not count how many visits a visitor comes to the park.

In "Figure 2. Number of visitors to Utah Lake State Park….", it appears that there was a sharp decline in visitors starting in 2013. However, in the sentence above the figure the report states that 'the number of people recreating on Utah Lake is expected to increase". These are two differing interpretations of recreational trends on Utah Lake. This needs to be reconciled.

My overall conclusion is that the UDWQ 2016 Draft IR is heavy on numeric -criteria –based-measures such as DO, but very weak on how these metrics actually relate to biological integrity the real measure of water quality as mandated by the Clean Water Act or even to recreational

use. Finally, there seems to be no clear scientific or otherwise causal link between the numeric based metrics and the 'beneficial uses' particularly biological, that they are supposed to evaluate.

If you have questions concerning my comments, please feel free to contact me at anytime.

Sincerely,

David C. Richards, Ph. D.
OreoHelix Consulting

Version 1.1

Real and Perceived Macroinvertebrate Assemblage Variability in the Jordan River, Utah can Effect Water Quality Assessments

Draft Technical Report

August 27, 2016

To:
Jordan River/Farmington Bay Water Quality Council
Salt Lake City, UT

By
David C. Richards, Ph. D.
OreoHelix Consulting
Phone: 406.580.7816
Email: oreohelix@icloud.com

Summary

Seasonality, field sampling error, subsampling, taxonomic resolution, and the river continuum affects our understanding of macroinvertebrate assemblage relationships in the Jordan River and confounds generalized water quality assessment models and conclusions.

Introduction

The Jordan River flows from Utah Lake and into the Great Salt Lake through the most densely populated area of Utah. By any measure; this river has been compromised. Its headwater source, Utah Lake, has undergone a catastrophic ecosystem shift in the last century and is now a shallow, eutrophic, highly regulated, agricultural-centric-use reservoir. More often than not, water needs to be physically pumped out of Utah Lake and into the Jordan River to maintain baseline flows. Tributaries to the Jordan River also primarily consist of Utah Lake water conveyed through a convoluted and difficult to resolve system of canals and dams. Most clean cold waters from its tributaries originating in the Wasatch Range no longer reach the Jordan River but instead are diverted for agricultural or culinary uses. Several spring tributaries that occur alongside the Jordan River are rapidly being appropriated and contaminated by residential and industrial users.

Historically, the Jordan River regularly flooded its impermanent banks and meandered across its valley forming many braided channels along its course terminating in the Great Salt Lake. The Jordan River naturally followed the river continuum with steeper gradients and larger substrata upstream and more meandering channels, lower gradients, and finer sediments along its downstream reaches. Physical, chemical, and biological conditions reflected this natural gradient. Although not formally quantified and akin to Utah Lake; the Jordan River has also undergone a catastrophic ecosystem shift and will continue to function as a highly regulated and restrained urban stream into the foreseeable future.

Macroinvertebrates are perhaps the most useful biological indicators of conditions in rivers and streams, including the Jordan River and are the cornerstone of most water quality assessment programs throughout the world. The Utah Department of Water Quality (UDWQ) relies on macroinvertebrates as their sole determinant of the biological component of their water quality assessments of the Jordan River. UDWQ has delineated the Jordan River into eight Assessment Units (AU) for water quality assessment purposes. Six out of the eight Jordan River AUs are considered impaired based on a macroinvertebrate metric, O/E (UDWQ 2016). However, it goes without saying that many natural factors can contribute to differences in macroinvertebrate assemblages including seasonal shifts and changes in the river continuum from upstream to down. In addition, it is well known that many other ‘sampling errors’ can affect interpretation of assemblages including the level of taxonomic resolution and associated error, subsampling effects, and actual field sampling error.

In this preliminary analysis, I highlight differences in macroinvertebrate assemblages in the Jordan River based on spatial and temporal factors, and sampling error, and then show how these differences can effect water quality assessments.

Methods

Data set

Available data were required that used comparable macroinvertebrate sampling methods and standardized taxonomic resolution and which focused on water quality assessment in the Jordan River. Twenty-five samples were chosen from the USU/USGS BugLab Mapit website: http://wmc6.bluezone.usu.edu. These samples were collected by UDWQ for water quality assessment from 2000 until 2006. Several of the samples were collected from the same site at different years (Table 1).

Table 1.UDWQ Macroinvertebrate dataset used in this report. Data were obtained from USU/USGS BugLab Mapit website. Note: Latitude/Longitude for samples 114429, 115117, 117487, 118510, 121480, 126843, 127668, and 129968 (Bluffdale Road crossing) are incorrect in the USU/USGS website. The correct latitude/longitude for these samples is 40.48717, -111.93626.

UDWQ Sample ID	Station	Location	Lat	Long	Month	Day	Year
114429	499460	Bluffdale Road crossing	40.48638916	-111.0852814	May	2	2000
114433	499088	Newstate Canal Road crossing	40.9056015	-111.9336014	May	3	2000
114442	499417	7800 South	40.6094017	-111.9203033	May	13	2000
115117	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Oct	12	2000
117487	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Mar	30	2001
118510	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Oct	18	2001
121480	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Apr	3	2003
124961	499088	Newstate Canal Road crossing	40.9056015	-111.9336014	Nov	24	2003
126843	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Oct	20	2004
127346	499088	Newstate Canal Road crossing	40.9056015	-111.9336014	Dec	1	2004
127666	4992290	1700 South	40.73361206	-111.9227753	Nov	10	2005
127667	4994101	6800 South	40.62333298	-111.9199982	Nov	10	2005
127668	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Oct	4	2005
127669	499088	Newstate Canal Road crossing	40.9056015	-111.9336014	Nov	10	2005
129968	499460	Bluffdale Road crossing	40.48638916	-111.0852814	Oct	19	2006
140272	4992890	3900/4100 South Crossing	40.68611145	-111.9202805	Sep	26	2007
140273	4991890	500 North Crossing	40.78027725	-111.9377747	Sep	25	2007
140274	4994270	9000 South Crossing	40.58750153	-111.9119415	Sep	27	2007
140275	4994520	Bangerter Highway Crossing	40.52338028	-111.9210205	Sep	28	2007
141615	4994100	6400 South	40.71722412	-111.5163879	Oct	30	2008
142102	4992480	Mill Creek above conflulence with Jordan River at USGS Gage Station	40.70861053	-111.9196701	Nov	9	2009
142111	4991820	Cudahy Lane above Davis S WWTP	40.84116745	-111.9499969	Nov	16	2009
142112	4991800	1000 ft below South Davis WWTP	40.84500122	-111.9524994	Nov	16	2009
142113	4992880	3300 S Crossing above confluence with Mill Creek	40.71611023	-111.9255524	Nov	9	2009

142114	4992320	1100 W 2100 S below confluence with Mill Creek	40.72499847	-111.9250031	Nov	9	2009

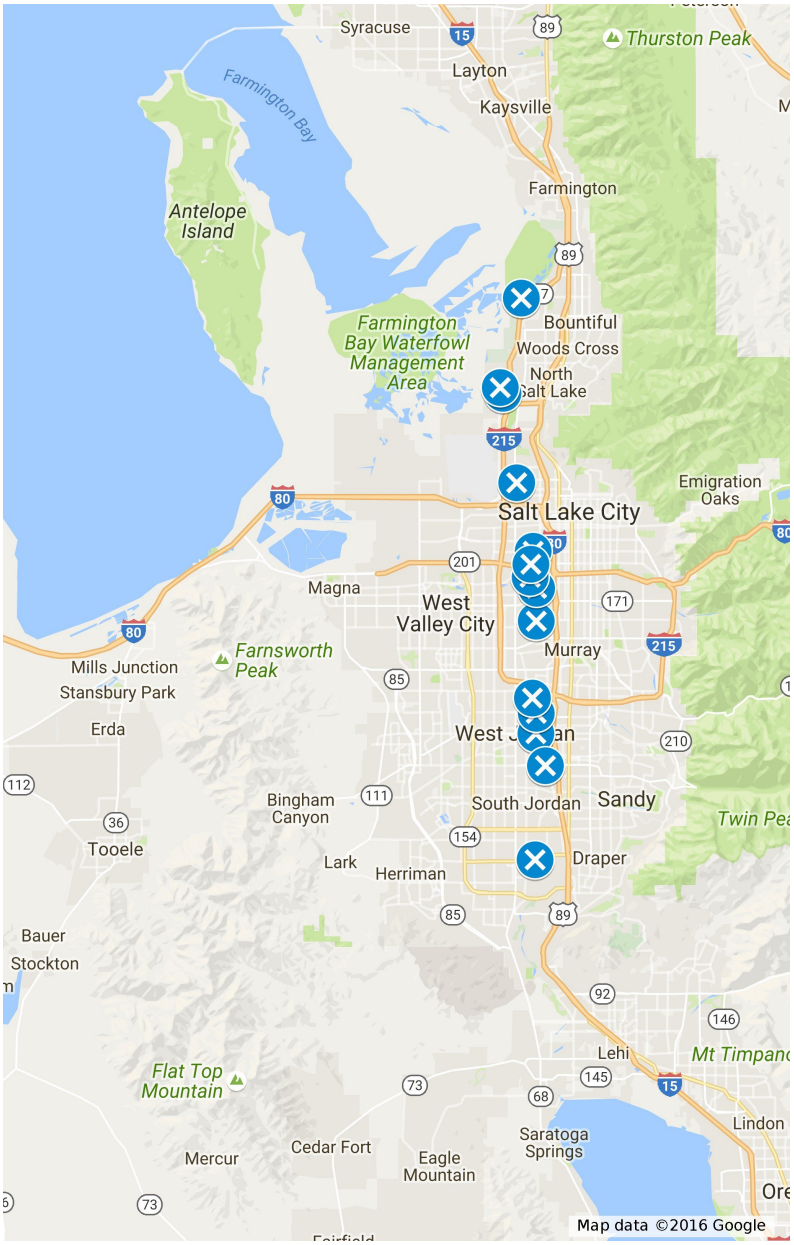


Figure 1. Map of macroinvertebrate sample locations on the Jordan River used in this analysis. Refer to UDWQ website for maps of their Jordan River assessment units.

Statistical Methods

Non metric multidimensional scaling (NMS), multiple response permutation procedure (MRPP), and Indicator Species Analyses (ISA) were performed on the macroinvertebrate dataset using PC-ORD Version 6.0 (McCune and Mefford 2011). A Sorensen Bray distance measure on log + 1 transformed macroinvertebrate abundances was used. Taxa were filtered and 'rolled up' to lower taxonomic resolution when appropriate and taxa that occurred in less than 2 of the 25 samples were eliminated. NMS and MRPP were evaluated by; month, year, UDWQ assessment

unit, and % laboratory subsampled (lab split). Simple box plots of taxa abundances at each of the UDWQ assessment units were also produced.

Results

Taxa List for Jordan River

The following is a taxa list for the Jordan River based on the samples analyzed (Table 2). Richards 2016 includes a more robust list for the Jordan River.

Table 2. Complete taxa list from all samples analyzed in this report.

Aeshna
Aeshnidae
Ambrysus
Amphipoda
Anax walsinghami
Antocha monticola
Argia
Argia emma
Baetidae
Baetis
Bivalvia
Caecidotea
Caenis
Centrarchidae
Chironomidae
Chironominae
Clitellata
Coenagrionidae
Collembola
Corbicula fluminea
Corixidae
Curculionidae
Dina parva
Diptera
Dubiraphia
Elmidae
Ephemerellidae
Ephemeroptera
Erpobdellidae
Fallceon quilleri
Ferrissia

Ferrissia rivularis
Fluminicola coloradoensis
Gammarus
Gastropoda
Glossiphonia complanata
Gyraulus
Gyrinus
Helobdella stagnalis
Hemerodromia
Heptageniidae
Hetaerina
Hetaerina americana
Hyalella
Hyalella azteca
Hydrobiidae
Hydropsyche
Hydropsychidae
Hydroptila
Hydroptilidae
Isoperla
Lebertia
Lepidostoma
Leptoceridae
Leptohyphidae
Leptophlebiidae
Microcylloepus pusillus
Microcylloepus similis
Nemata
Nephelopsis obscura
Oligochaeta
Ophiogomphus
Optioservus
Optioservus quadrimaculatus
Orconectes virilis
Ordobrevia nubifera
Orthocladiinae
Physa
Pisidium
Planorbidae
Plecoptera

Potamopyrgus antipodarum
Probezzia
Psychoda
Sigara
Simuliidae
Simulium
Simulium vittatum group
Sperchon
Sperchonidae
Stenelmis
Tanypodinae
Trichoptera
Tricorythodes
Tricorythodes minutus
Trombidiformes
Turbellaria

Macroinvertebrate Assemblage Differences in the Jordan River

The best NMS ordination using all 25 samples for the entire Jordan River had a final stress of 10.17 for a 3-D solution, a final instability of < 0.00001 for 91 iterations. R^2 values for Axis 1 =0.51, Axis 2 = 0.28, and Axis 3 = 0.12 for a cumulative R^2 of 0.90 (Figure 2 and 3). There was an observable statistically significant difference in macroinvertebrate assemblages between months (Figure 2 and 3, MRPP; A = 0.10, p = 0.002). For additional MRPP comparisons between months see Appendix 2.

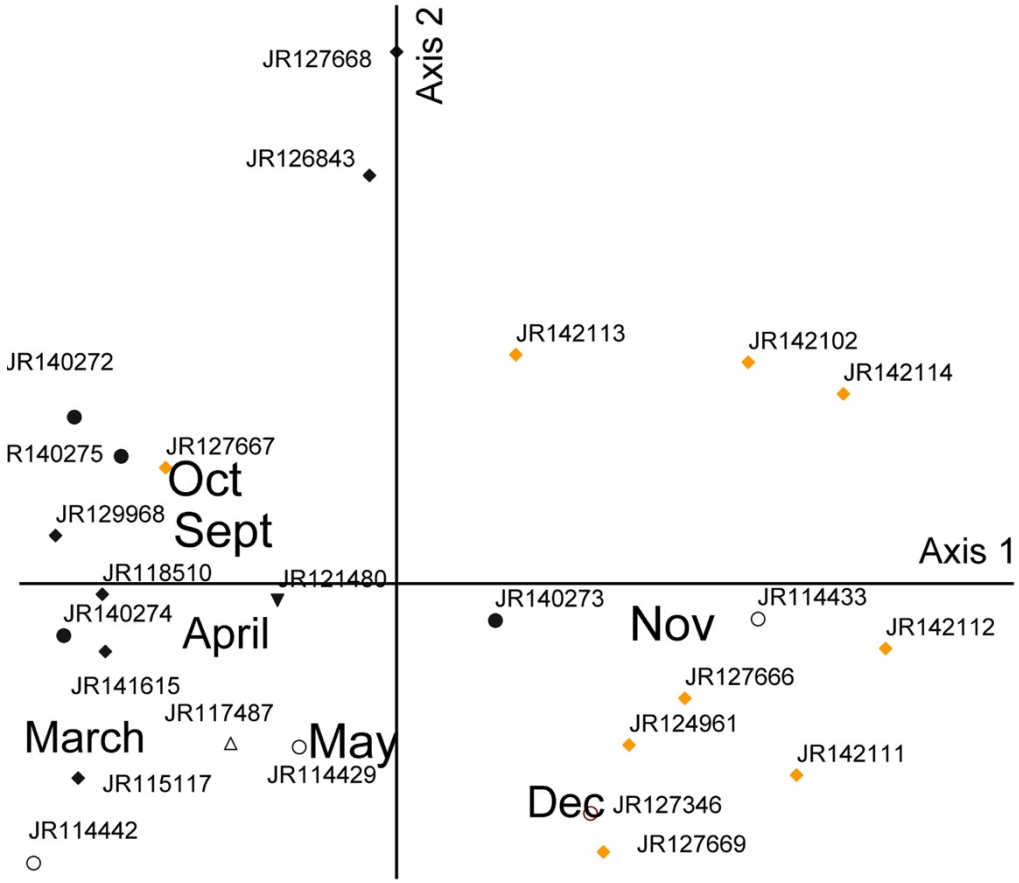


Figure 2. NMS ordination axes 1and 2 with month centroids added.

JR127668 (Jordan River at Bluffdale Road crossing October 2005) and JR126843 (Jordan River at Bluffdale Road crossing October 2004) separated out by themselves along Axis 2 in the upper left quadrant (Figure 2)(See following NMS/MRPP sections for addition comparisons).

JR142102 was the sample collected from Mill Creek near the confluence with the Jordan River and was similar to JR142114 the Jordan River at 1100W 2100 S below the confluence with Mill Creek in Figure 2 but the Mill Creek sample separated away from other samples in Figure 3 (axis 1 vs. axis 3) except for JR142113 the Jordan River sample collected at 3300 South just upstream of confluence with Mill Creek.

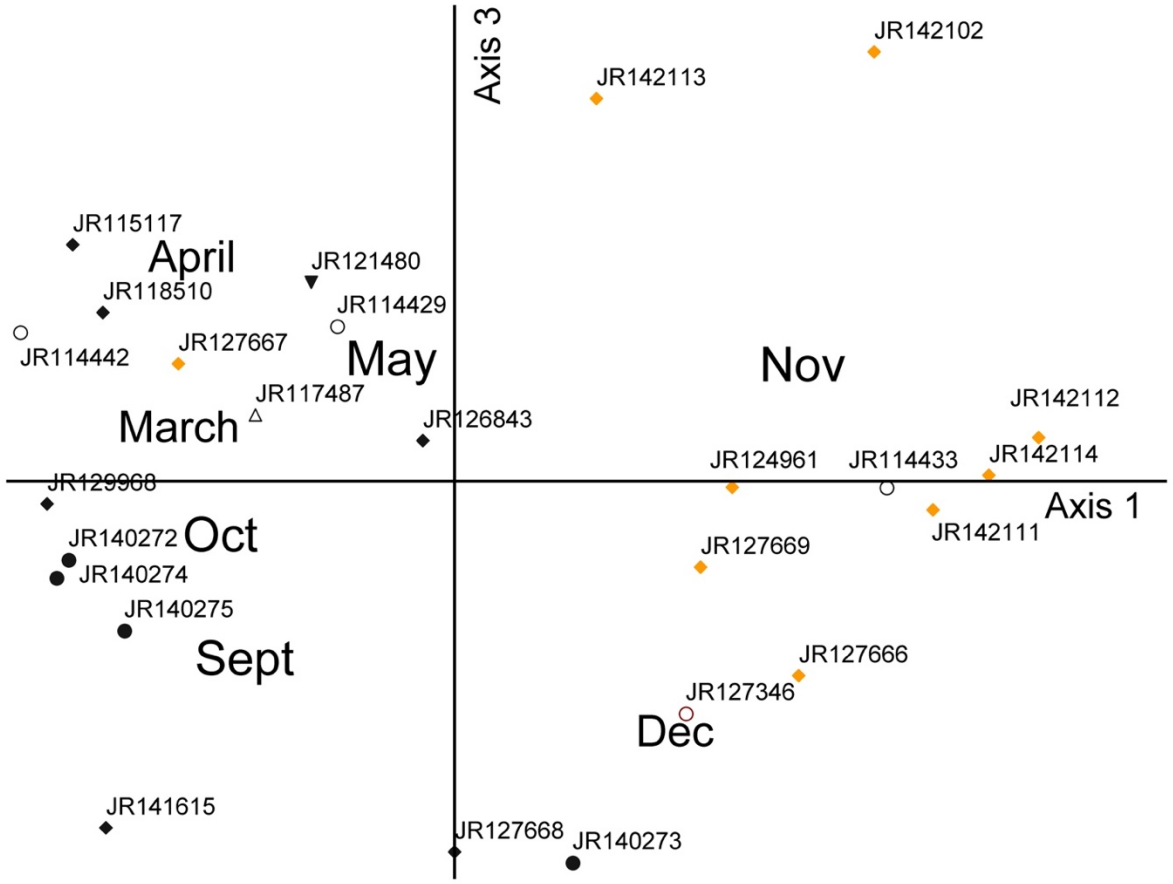


Figure 3. NMS ordination axes 1and 3 with month centroids added.

Macroinvertebrate assemblages also clearly differed between UDWQ Assessment Units, with Unit 1 (furthest downstream unit) and Unit 2 significantly differing than the other units (Figure 4, MRPP; $A = 0.15$, $p < 0.001$). The largest variability in assemblages was Assessment Unit 7, the upstream unit. For addition MRPP comparisons see Appendix 2.

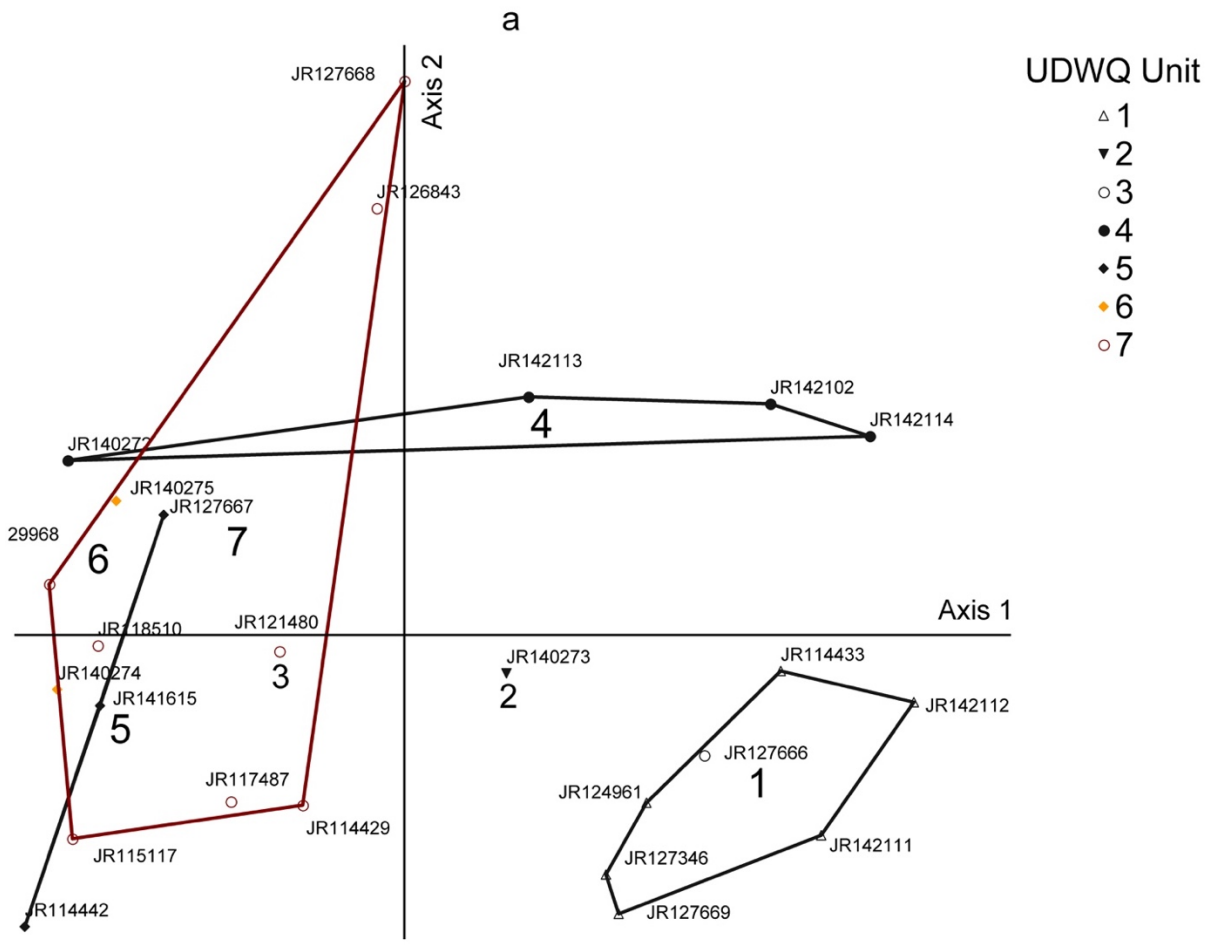


Figure 4. NMS ordination axes 1and 3 with UDWQ Assessment Unit centroids added. UDWQ Units 1, 2, 3, 4, 6, and 7 were listed as impaired (UDWQ 2016).

UDWQ listed Units, 1, 2, 3, 4, 6, and 7 as impaired but not Unit 5 however, the NMS results (Figure 4) suggest that Unit 5 macroinvertebrate assemblage was not much different than Units 3, 6, and 7.

Macroinvertebrate assemblages were also significantly affected by lab split (% subsampled)(Figure 5).

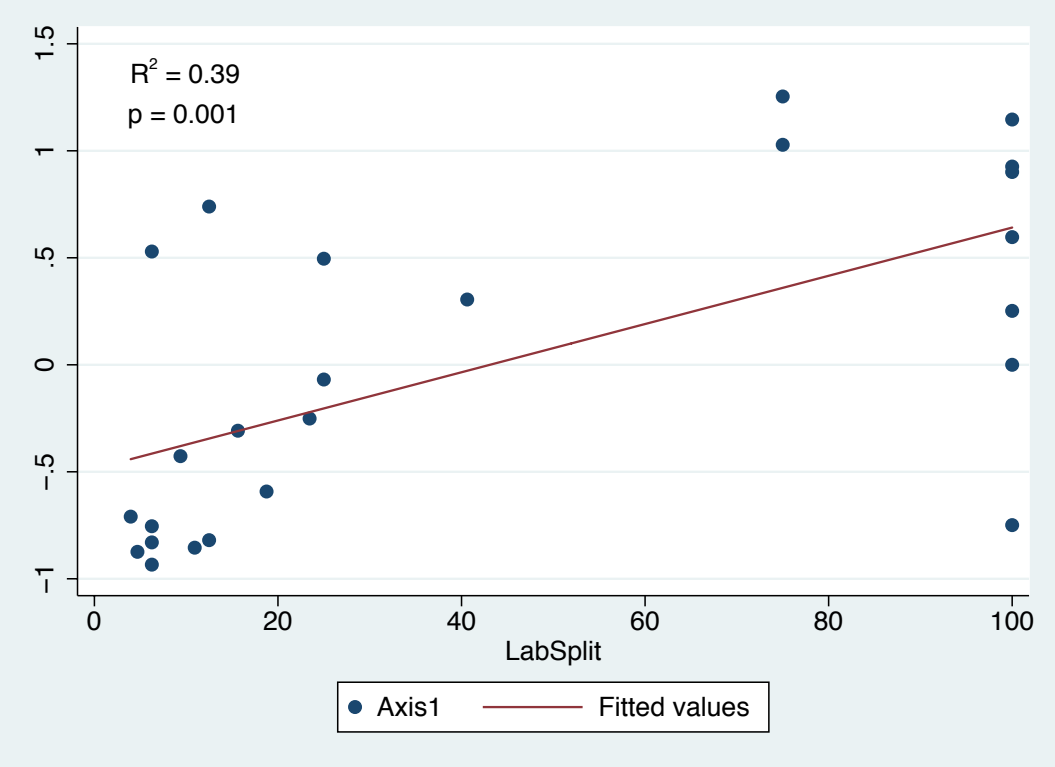
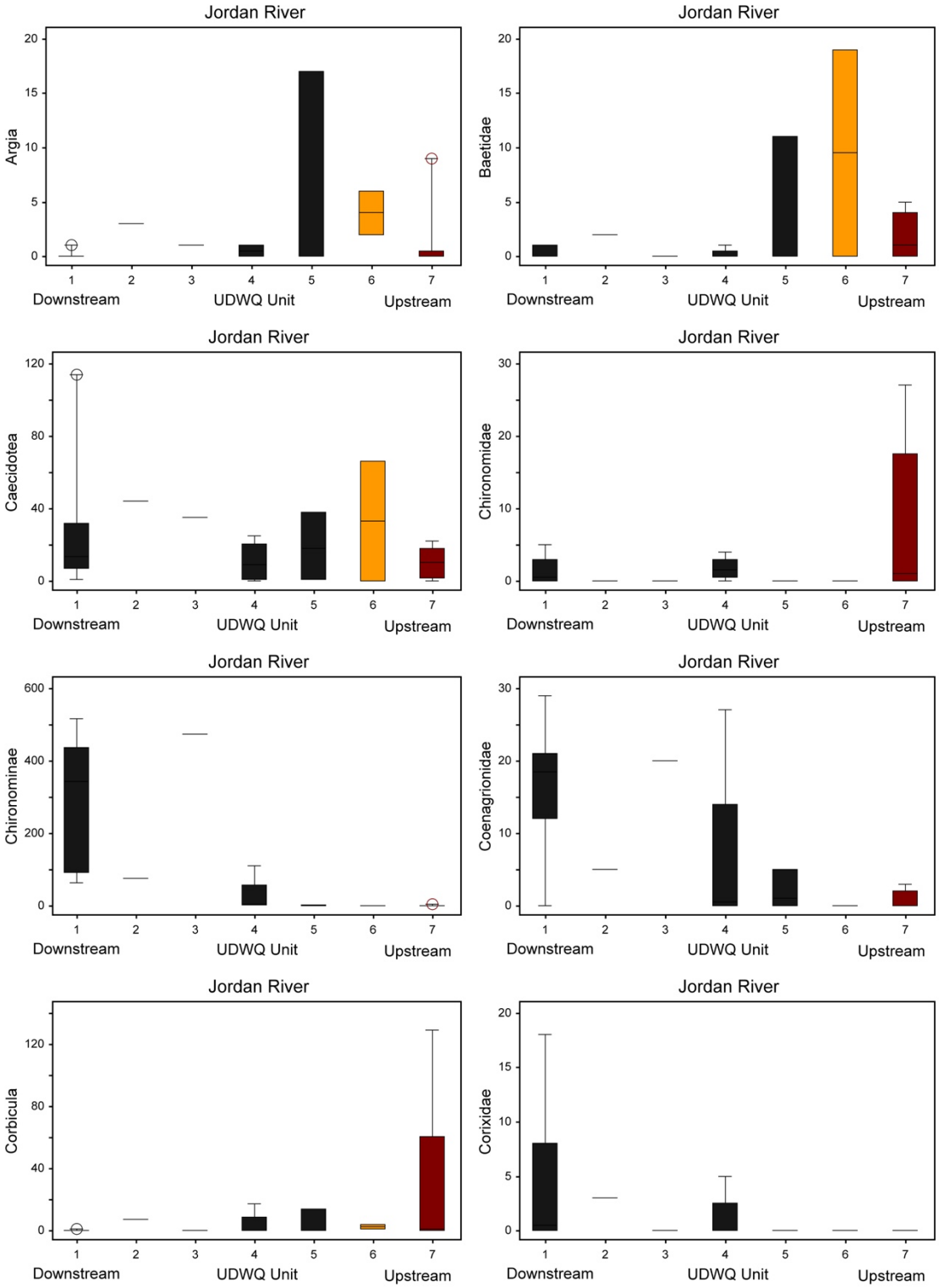
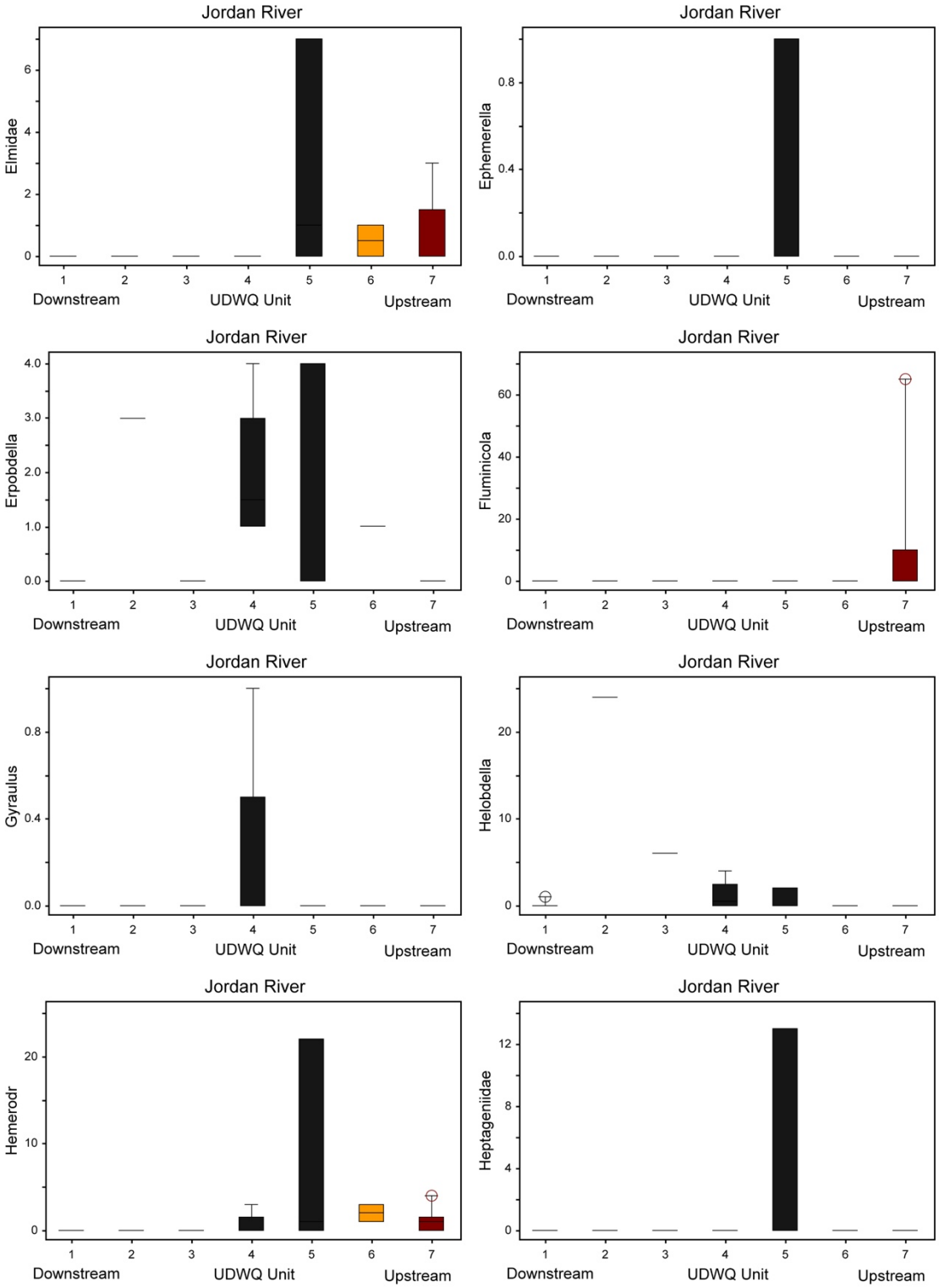


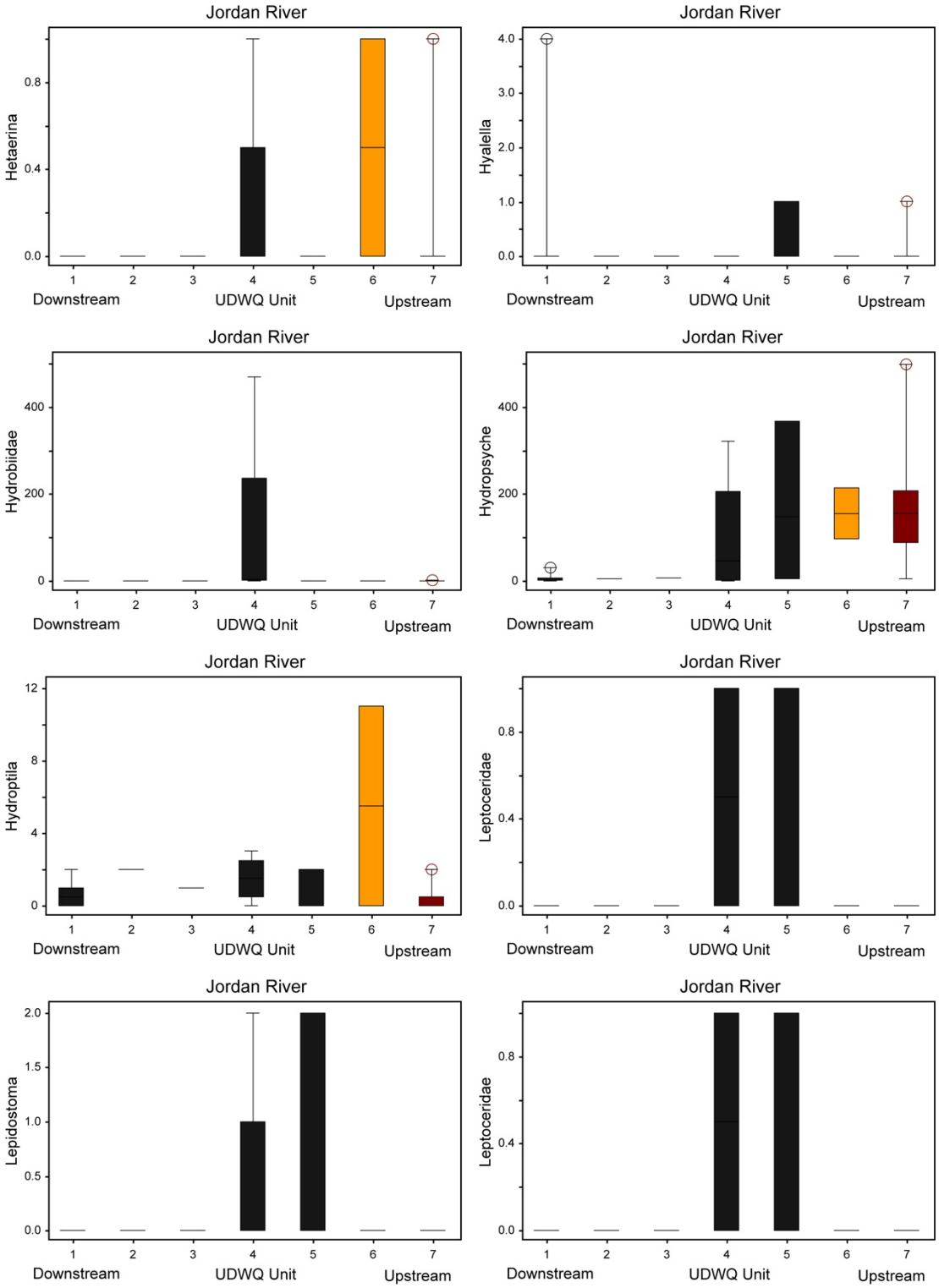
Figure 5. Linear fit model of NMS axis 1 vs. Lab Split.

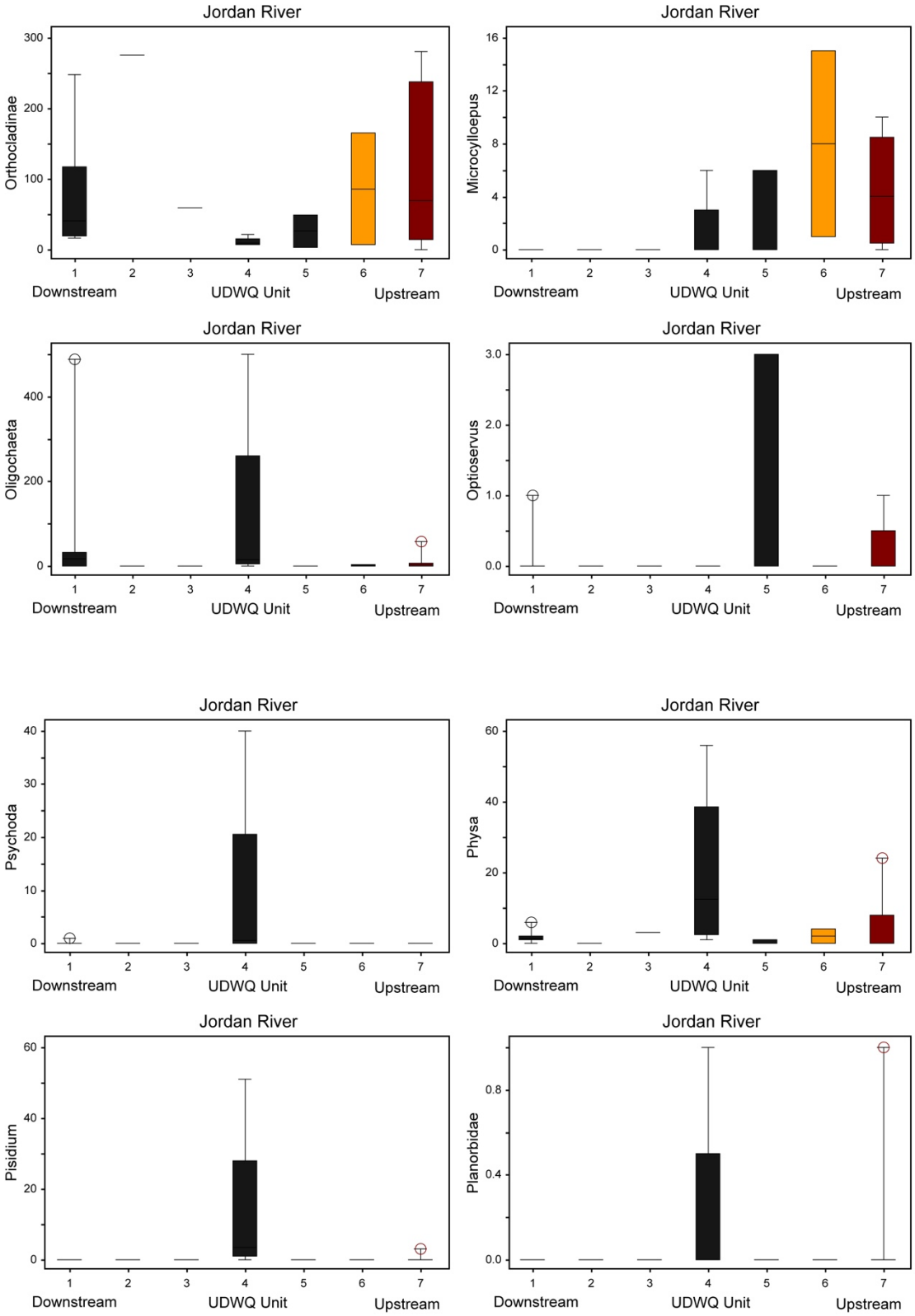
Taxa Abundances in Assessment Units

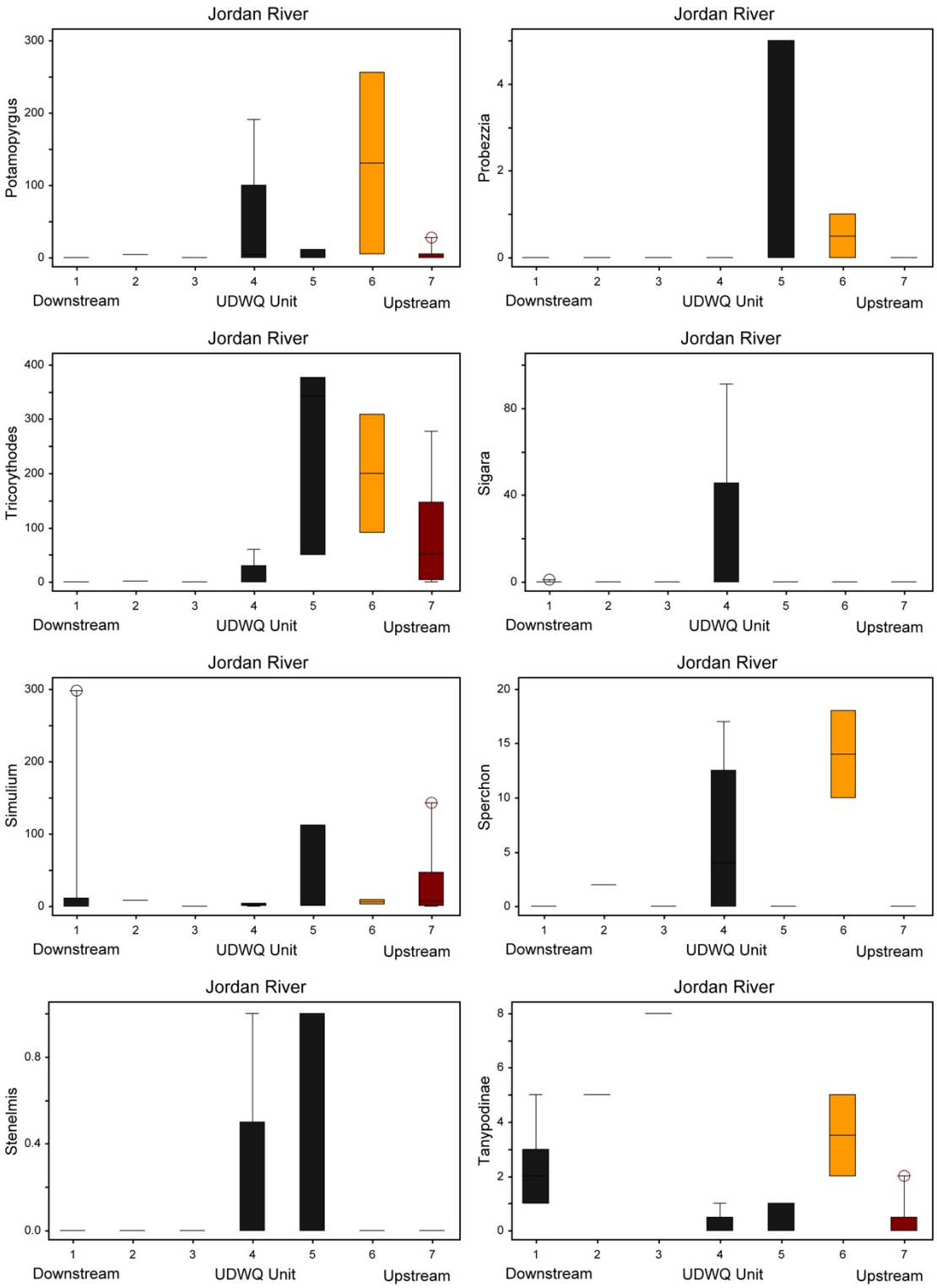
As expected and predicted by the River Continuum Concept, individual taxa varied in abundances between UDWQ Assessment Units (Figure 6).


Jordan River
Argia
Baetidae
Caecidotea
Chironomidae
Chironominae
Coenagrionidae
Corbicula
Corixidae
Downstream
UDWQ Unit
Upstream


Jordan River
Elmidae
Ephemerella
Erpobdella
Fluminicola
Gyraulus
Helobdella
Hemerodr
Heptageniidae
Downstream
UDWQ Unit
Upstream


Jordan River
Hetaerina
0.0
0.4
0.8
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Hyalella
0.0
1.0
2.0
3.0
4.0
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Hydrobiidae
0
200
400
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Hydropsyche
0
200
400
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Hydroptila
0
4
8
12
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Leptoceridae
0.0
0.4
0.8
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Lepidostoma
0.0
0.5
1.0
1.5
2.0
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream
Jordan River
Leptoceridae
0.0
0.4
0.8
1 2 3 4 5 6 7
Downstream
UDWQ Unit
Upstream


Jordan River
Orthocladinae
Microcylloepus
Oligochaeta
Optioservus
Psychoda
Physa
Pisidium
Planorbidae
Downstream
UDWQ Unit
Upstream


Jordan River
Potamopyrgus
Probezzia
Tricorythodes
Sigara
Simulium
Sperchon
Stenelmis
Tanypodinae
Downstream
UDWQ Unit
Upstream

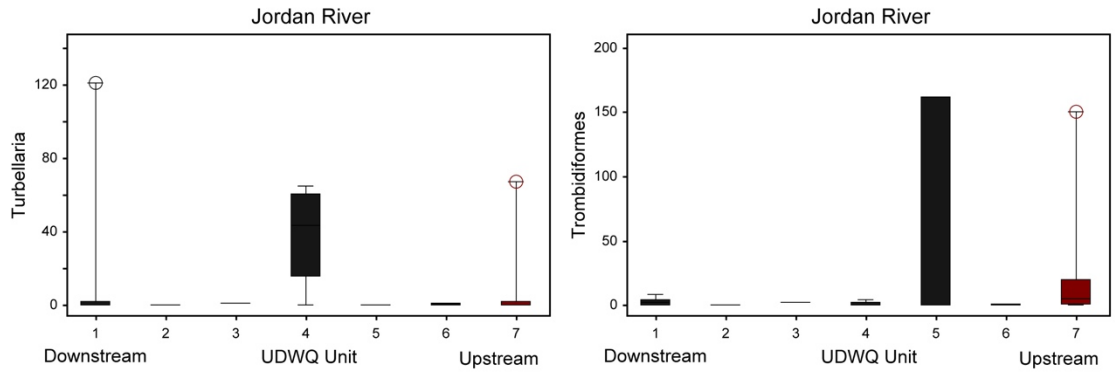


Figure 6. Macroinvertebrate taxa abundances for each of the seven UDWQ assessment units. Box plots are 25th to 75th centiles, range, and outliers.

Jordan River Assemblages Upstream vs. Downstream

Jordan River macroinvertebrate assemblages clearly and significantly differed between upstream Unit 7 and downstream Unit 1 and within these sections due to sampling, seasonal, and annual effects and water quality effects, particularly the major water diversion at 2100 South (Surplus Canal), which could have affected assemblages in the furthest downstream Units 1 and 2. The best fit NMS ordination resulted in a very good final stress = 8.02, final instability < 0.001, using 49 iterations for a 2-D solution (Figure 7). Axis 1 $R^2 = 0.41$ and Axis 2 $R^2 = 0.51$ for a cumulative $R^2 = 0.92$. MRPP A = 0.2, $p < 0.001$.

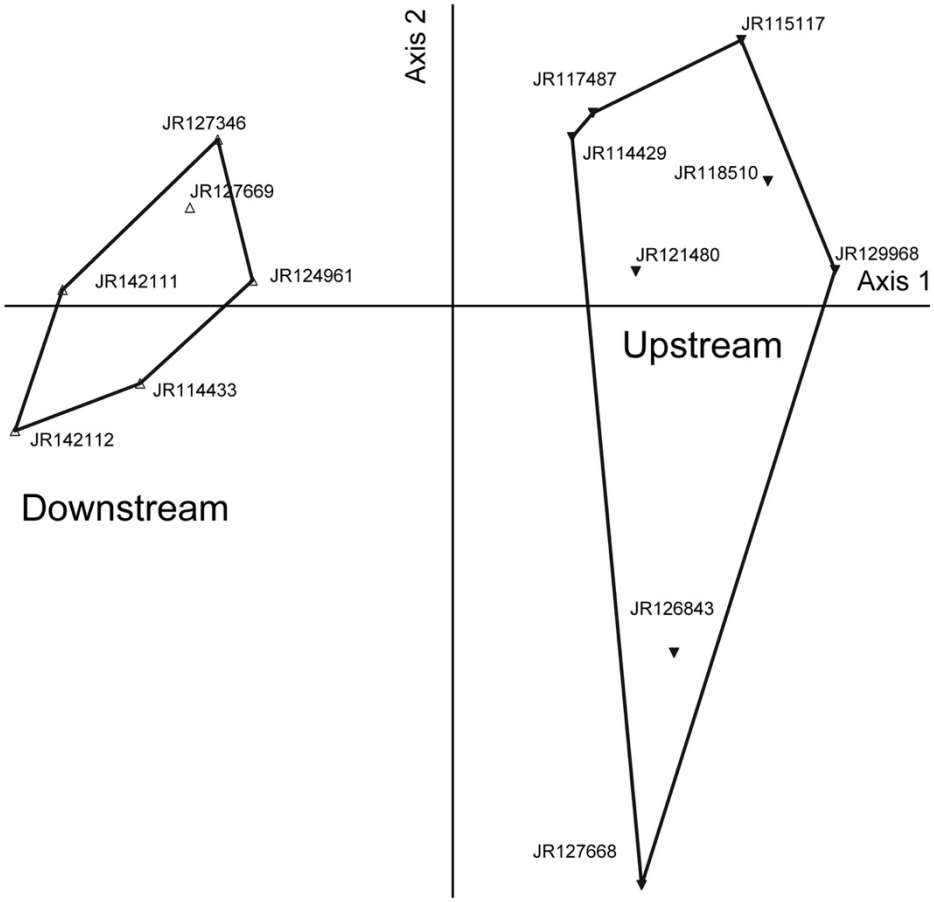


Figure 7. NMS Axis 1 and 2 for Jordan River macroinvertebrate samples from UDWQ Assessment Unit 1 (furthest downstream) and Assessment Unit 7 (furthest upstream).

One of the reasons upstream (Unit 1) and downstream (Unit 7) assemblages differed so much was because a large proportion of the Jordan River is diverted into Surplus Canal at 2100 South and flows are greatly diminished downstream. However, the Jordan River also changes character naturally from upstream to downstream and the taxa clearly showed this (Figure 6).

Indicator Taxa

Several taxa were significantly more abundant either upstream or downstream and are useful indicators using Indicator Taxa Analysis (Table 3).

Table 3. Taxa that were significant indicators of either upstream or downstream macroinvertebrate assemblages using Indicator Taxa Analysis.

Downstream				
Taxon	Observed Indicator Value (IV)	IV Mean	IV Std Dev	p
Chironominae	93.6	45.6	11	0.0018

Tanypodinae	83.4	42	10.75	0.0048
Coenagrionidae	70.7	42	10.94	0.0204
Corixidae	50	23.6	10.99	0.0566

Upstream				
Taxon	Observed Indicator Value (IV)	IV Mean	IV Std Dev	p
Hydropsyche	78.2	55.3	6.51	0.0004
Microcylloepus	75	35.3	11.33	0.0138
Tricorythodes	75	35.5	11.21	0.016
Hemerodromia	62.5	31.6	11.89	0.0316

As shown earlier in Figure 2, the most variability in assemblages occurred in the Upstream section (assessment unit 7). This was primarily due to two samples, JR126843 collected October 2004 and JR127668 collected October 2005. This variability prompted the following multivariate analysis.

Assemblages Upstream in October

The macroinvertebrate samples collected from Assessment Unit 7 in October were also clearly and statistically different between years. 2004 and 2005 October samples were much different than 2000, 2001, and 2006 samples. The best fit NMS ordination had an extremely low and highly accurate final stress of 0.0001, < 0.0001 instability, using 26 iterations for a 1-D solution. Axis 1 R^2 was 0.67.

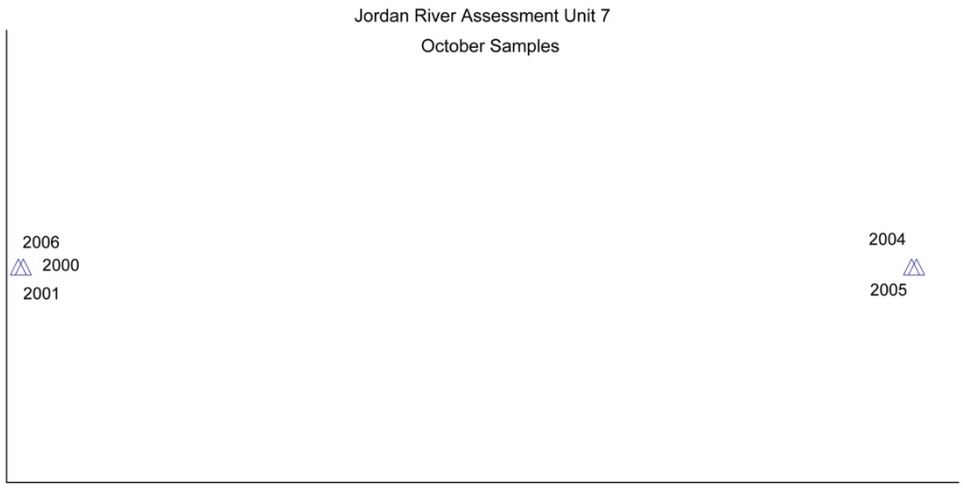


Figure 8. NMS ordination Axis 1 for macroinvertebrate samples collected by UDWQ at Assessment Unit 7 in October at five different years.

Taxa responsible for these differences were mostly more Tricorythodes, Hydropsyche, Orthocladinae, Argia, and Oligochaetes in the 2000, 2001, and 2006 samples vs. more Physa, Corbicula, Fluminicola, Caecidotea, and Coenagrionidae found in the 2004 and 2005 samples (Table 4).

Table 4. Correlations (r)between macroinvertebrate taxa and Axis 1 of the NMS ordination of UDWQ Assessment Unit 7 October samples.

Taxon	r
Tricoryt	-0.988
Hydropsy	-0.881
Orthocla	-0.831
Oligocha	-0.666
Argia	-0.545
Hemerodr	-0.433
Dina	-0.412
Elmidae	-0.406
Hydrobii	-0.406
Hydropti	-0.406
Optioser	-0.406
Simulium	-0.192
Chrnmida	-0.166
Baetidae	-0.096
Trombidi	0.243
Potamopy	0.304
Turbella	0.49
Microcyl	0.549
Coenagri	0.72
Caecidot	0.882
Fluminic	0.98
Corbicul	0.993
Physa	0.996

The NMS axis was also strongly correlated with percent lab split (r = 0.75) which was a function of the number of individuals encountered in a sample. The more individuals in a sample, the greater the percentage needed to be subsampled (labsplit) to meet standardized count criteria. Richards 2016 showed the relationship between labsplit, evenness and taxa richness and how these can negatively affect O/E scores.

Table 5. Sample ID, year, and % subsampled UDWQ macroinvertebrate samples from Assessment Unit7 collected in October.

Sample ID	Year	Lab-Split

		(% subsampled)
JR115117	2000	12.50
JR118510	2001	6.25
JR126843	2004	25.00
JR127668	2005	100.00
JR129968	2006	4.68

Discussion and Conclusion

Although the Jordan River is obviously impaired; much of the differences in macroinvertebrate samples in the Jordan River can be explained by seasonality, sampling error, subsampling effects, and the river continuum. It is more difficult to explain why there were more Physa, Corbicula, Fluminicola, Caecidotea, and Coenagrionidae found in the 2004 and 2005 upstream samples and more Tricorythodes, Hydropsyche, Orthocladinae, Argia, and Oligochaetes in the 2000, 2001, and 2006 upstream samples even though all of these samples were collected in October. Tricorythodes, Orthocladinae, and Argia are short lived taxa and their abundances can change annual more so than the other taxa. Physa are more typically found in the slower, shoreline sections. Perhaps samples were taken along the shoreline in 2004/2005. Corbicula are now abundant throughout the Jordan River and because they are relatively long lived they should have been collected in all samples particularly in 2006 if they recently became established. Fluminicola also have overlapping generations and should have been collected in all years. The most likely explanation for these differences is sampling error, including the effects of laboratory subsampling.

Many of the samples were collected at assessment unit boundaries. Samples collected at boundaries need to be consistently included in the unit immediately upstream because macroinvertebrate assemblages are determined by upstream conditions not downstream conditions. Future samples should be centered within the unit not at the boundaries. UDWQ Assessment Unit boundaries should be adjusted based on changes in macroinvertebrate assemblages and not on other less informative variables (e.g. fish assemblages, DO, convenient landmarks, etc.) and after more samples are collected and analyzed. O/E assessments and other water quality assessments need to reflect this, which can substantially improve assessment conclusions.

Differences in Unit 7 October samples could easily have been due to annual variation in taxa abundances and were also likely due to the percentage subsampled (lab- split). Other factors could have been involved such as sampling error (e.g. location differences, uneven compositing from non-riffle habitat, or taxonomic resolution). It is also possible that some type of impairment affected these assemblages in differing years, however I am not aware of any changes in Unit 7 during these years, particularly 2004 and 2005. Further investigation into any events such as high flow years, dewatering, habitat alterations needs to be conducted. The very discernable differences in assemblages reported from samples collected in the same assessment unit and the same month certainly affected any water quality assessments even if no perturbation occurred during the years sampled. At this time, it is unknown if UDWQ used the same O/E scoring criteria for all assessment units to conclude that most units in the Jordan River were impaired. However, the most recent UDWQ Integrative Report (2016) states that O/E models are based on

at least eleven 'watershed' based climate/environmental variables without defining 'watershed'. The Jordan River watershed drains over 3800 square miles with elevations ranging from 11,900 ft. to 4200 ft. (http://www.utahcleanwater.org/jordan-river-watershed.html). Assigning impaired status to UDWQ Jordan River Assessment Units could have been inappropriate if UDWQ used the entire Jordan River watershed, as was detailed in their 2016 IR, to develop O/E models and averaged watershed "E" (expected) taxa probabilities of capture for final O/E scores.

Finally, highly regulated urban rivers such as the Jordan River cannot be compared to reference rivers or hypothetical 'average watershed' macroinvertebrate assemblages. Highly regulated, urban rivers and their macroinvertebrate assemblages will never be able to achieve reference condition or expected (E) values because they have been irreversibly altered and compromised. New assessment methods need to be developed to assess and monitor the desired unnatural conditions of urban rivers, including the Jordan River.

Recommendations

Because this was a preliminary analysis, more documentation and analyses of the physical and chemical conditions of the Jordan River is highly recommended. Physical and chemical data were likely collected by UDWQ at the same time macroinvertebrate samples were collected. Additional macroinvertebrate samples and physical and chemical data are also needed to assess conditions, particularly if no samples have been collected in the last several years. Regulatory metrics such as the O/E metric need to be adjusted to reflect sampling error, natural variability, and 'watershed' representativeness in the assemblages. O/E metrics also need to be revised to account for the now inescapable permanent condition of the Jordan River; a highly regulated and exceedingly managed urban river. UDWQ Assessment Units should be adjusted based on the best bio-predictor of water quality, macroinvertebrates.

Literature Cited

McCune, B. and J. Mefford. 2011. PC-ORD. Multivariate Analysis of Ecological Data. Version. 6.22. MjM Software, Gleneden Beach, Oregon.

Richards, D. C. 2016. Is reliance on RIVPACS O/E models for monitoring water quality in Utah prudent? Draft Technical Report to: Jordan River/Farmington Bay Water Quality Council. OreoHelix Consulting. August 14. 32 pages.

UDWQ. 2016. Utah's Draft 2016 Integrated Report. Salt Lake City, Utah. Utah Department of Environmental Quality.

Appendices

Appendix 1. NMS sample Axes 1,2, and 3

JR114429	-0.25146	-0.41853	0.33293
JR114433	0.92628	-0.08856	-0.01144
JR114442	-0.93414	-0.71781	0.32096
JR115117	-0.81923	-0.50179	0.50770
JR117487	-0.42682	-0.41176	0.14353
JR118510	-0.75473	-0.02641	0.36261
JR121480	-0.30773	-0.04072	0.42941
JR124961	0.59670	-0.41396	-0.01281
JR126843	-0.06890	1.05133	0.08723
JR127346	0.49564	-0.58994	-0.49939
JR127666	0.73958	-0.29589	-0.41694
JR127667	-0.59258	0.29737	0.25318
JR127668	0.00020	1.36699	-0.79838
JR127669	0.52903	-0.68851	-0.18530
JR129968	-0.87418	0.12582	-0.05000
JR140272	-0.83012	0.43177	-0.16770
JR140273	0.25204	-0.09307	-0.81922
JR140274	-0.85499	-0.13335	-0.20668
JR140275	-0.70967	0.33032	-0.32007
JR141615	-0.74947	-0.17336	-0.74556
JR142102	0.90119	0.57110	0.92485
JR142111	1.02803	-0.49350	-0.06050
JR142112	1.25422	-0.16619	0.09322
JR142113	0.30475	0.58850	0.82399
JR142114	1.14637	0.49017	0.01440

Appendix 2. NMS macroinvertebrate Axes 1, 2 and 3.

Aeshnida	0.85117	-0.37234	-0.27905
Ambrysus	-0.59258	0.29737	0.25318
Argia	-0.41477	0.14812	-0.07331
Baetidae	-0.39807	0.02388	-0.24702
Caecidot	0.14818	-0.00292	-0.04674
Chrnmida	0.13199	-0.08282	0.30684
Chrnmina	0.57648	-0.22309	-0.15906
Clitella	0.64614	-0.41472	-0.32886
Coenagri	0.57951	-0.06881	-0.08094
Corbicul	-0.31674	0.50076	-0.31858
Corixida	0.80561	-0.10376	0.05252
Dina par	-0.62135	-0.32342	0.47741
Diptera	0.43318	-0.28181	-0.01784
Elmidae	-0.63837	-0.03233	0.09106
Ephemere	-0.74947	-0.17336	-0.74556
Ephemero	-0.25146	-0.41853	0.33293
Erpobdel	-0.11132	0.22246	-0.10709
Ferrissi	0.29696	-0.13489	-0.71412
Fluminic	-0.02888	1.23414	-0.42567
Gyraulus	1.14637	0.49017	0.01440
Helobdel	0.49781	-0.02410	-0.39367
Hemerodr	-0.64541	0.16762	0.03120
Heptagen	-0.93414	-0.71781	0.32096

```
Hetaerin  -0.42637   0.34821   0.15131
Hyalella  -0.09146  -0.38090  -0.28363
Hydrobii   0.34928   0.45914   0.58030
Hydropsy  -0.37914  -0.00201   0.07180
Hydropti  -0.03496   0.06211  -0.03480
Lebertia   0.47221   0.27085  -0.31673
Lepidost  -0.22236   0.20757   0.03922
Leptocer   0.02727   0.30067   0.08650
Leptohyp  -0.75473  -0.02641   0.36261
Leptophl  -0.42682  -0.41176   0.14353
Microcyl  -0.53569   0.40419  -0.03814
Nemata    0.08958   0.19893   0.63402
Oligocha   0.48947   0.00855   0.23847
Optioser  -0.41315  -0.20076  -0.23970
Orthocla  -0.00670  -0.13473  -0.00156
Physa    0.48292   0.44196   0.03948
Pisidium   0.10703   0.61922   0.37580
Planorbi   0.44746   0.03582   0.17367
Potamopy  -0.43404   0.37628  -0.36245
Probezzi  -0.66578   0.17722   0.12491
Psychoda   0.98250   0.45989   0.68808
Sigara    1.16071   0.40293   0.02488
Simulium  -0.21914  -0.04498  -0.01282
Sperchon  -0.48511   0.26817  -0.08780
Stenelmi  -0.14391   0.44293   0.53858
Tanypodi   0.31493  -0.22715  -0.22922
Tricoryt  -0.69888  -0.12650   0.02404
Trombidi  -0.20189   0.12506   0.10621
Turbella   0.49961   0.37988   0.28728
```

Appendix 3.MRPP results by month.

Month

```
Chance-corrected within-group agreement, A =  0.10072408
     A = 1 - (observed delta/expected delta)
     Amax = 1 when all items are identical within groups (delta=0)
     A = 0 when heterogeneity within groups equals expectation by chance
     A < 0 with more heterogeneity within groups than expected by chance

    Probability of a smaller or equal delta, p =  0.00244195
```

PAIRWISE COMPARISONS
Note: p values not corrected for multiple comparisons.

	T	A	p
May vs. October	0.25213146	0.00884007	0.35708513
May vs. November	0.05271432	0.00152494	0.42022471
May vs.September	1.55764143	0.06474327	0.06907140
October vs.November	4.64619396	0.10481584	0.00154419
October vs. September	0.86654643	0.02600541	0.18612781
November vs.September	4.08158040	0.12323943	0.00353617

Appendix 4. MRPP by UDWQ Assessment Unit

```
Identifiers for excluded groups:
         2
         3

    Groups were defined by values of: UDWQ Uni
    Input data has:   23 Samples by   52 Taxa
    Weighting option: C(I) = n(I)/sum(n(I))
    Distance measure: Sorensen (Bray-Curtis)

   GROUP:   1
Identifier:   7
   Size:   8   0.52998185  = Average distance
Members:
 JR114429 JR115117 JR117487 JR118510 JR121480 JR126843 JR127668 JR129968

   GROUP:   2
Identifier:   1
   Size:   6   0.42141123  = Average distance
Members:
 JR114433 JR124961 JR127346 JR127669 JR142111 JR142112

   GROUP:   3
Identifier:   5
   Size:   3   0.59540733  = Average distance
Members:
 JR114442 JR127667 JR141615

   GROUP:   4
Identifier:   4
   Size:   4   0.61634908  = Average distance
Members:
 JR140272 JR142102 JR142113 JR142114

   GROUP:   5
Identifier:   6
   Size:   2   0.42817768  = Average distance
Members:
 JR140274 JR140275

    Test statistic: T =   -4.9146274
      Observed delta =   0.51636069
      Expected delta =   0.60872260
    Variance of delta =   0.35318691E-03
    Skewness of delta =   -0.50886650

    Chance-corrected within-group agreement, A =  0.15173070
     A = 1 - (observed delta/expected delta)
     Amax = 1 when all items are identical within groups (delta=0)
     A = 0 when heterogeneity within groups equals expectation by chance
     A < 0 with more heterogeneity within groups than expected by chance
```

```
    Probability of a smaller or equal delta, p =  0.00008412


------------------------------------------------------------------
PAIRWISE COMPARISONS
Note: p values not corrected for multiple comparisons.

   Groups (identifiers)
    Compared       T       A       p
    7 vs.   1   -6.69399161  0.17355259  0.00012495
    7 vs.   5   0.66847261  -0.01898252  0.72674698
    7 vs.   4   -2.38142907  0.06656315  0.02730797
    7 vs.   6   -0.47524985  0.01958275  0.24143208
    1 vs.   5   -4.14904785  0.17225330  0.00240940
    1 vs.   4   -3.84048188  0.11366991  0.00274305
    1 vs.   6   -3.76897841  0.22797295  0.00809950
    5 vs.   4   -1.48355332  0.07606936  0.08465916
    5 vs.   6   0.40392234  -0.02181514      NaN
    4 vs.   6   -0.75705657  0.07703657  0.21191159
------------------------------------------------------------------

****************************** MRPP finished ******************************
10 Aug 2016, 16:59:52
```

Appendix 5. MRPP by Year

```
 Identifiers for excluded groups:
       2006
       2008

    Groups were defined by values of: Year
    Input data has:   23 Samples by   52 Taxa
    Weighting option: C(I) = n(I)/sum(n(I))
    Distance measure: Sorensen (Bray-Curtis)

   GROUP:   1
Identifier: 2000
   Size:   4   0.57616850  = Average distance
Members:
 JR114429 JR114433 JR114442 JR115117

   GROUP:   2
Identifier: 2001
   Size:   2   0.36443168  = Average distance
Members:
 JR117487 JR118510

   GROUP:   3
Identifier: 2003
   Size:   2   0.49666446  = Average distance
Members:
 JR121480 JR124961

   GROUP:   4
Identifier: 2004
```

```
   Size:   2   0.69450498  = Average distance
Members:
 JR126843 JR127346

   GROUP:   5
Identifier: 2005
   Size:   4   0.59811219  = Average distance
Members:
 JR127666 JR127667 JR127668 JR127669

   GROUP:   6
Identifier: 2007
   Size:   4   0.45090043  = Average distance
Members:
 JR140272 JR140273 JR140274 JR140275

   GROUP:   7
Identifier: 2009
   Size:   5   0.52126362  = Average distance
Members:
 JR142102 JR142111 JR142112 JR142113 JR142114

    Test statistic: T =   -3.0722786
      Observed delta =   0.53122804
      Expected delta =   0.60577188
    Variance of delta =   0.58871135E-03
    Skewness of delta =   -0.31823663

    Chance-corrected within-group agreement, A =  0.12305597
     A = 1 - (observed delta/expected delta)
     Amax = 1 when all items are identical within groups (delta=0)
     A = 0 when heterogeneity within groups equals expectation by chance
     A < 0 with more heterogeneity within groups than expected by chance

    Probability of a smaller or equal delta, p =  0.00316615


-----------------------------------------------------------------------
PAIRWISE COMPARISONS
Note: p values not corrected for multiple comparisons.

   Groups (identifiers)
    Compared       T       A       p
  2000 vs.  2001   0.03763096  -0.00161876  0.45489719
  2000 vs.  2003   0.03227199  -0.00273200  0.43564372
  2000 vs.  2004   -0.05221852  0.00378331  0.42477378
  2000 vs.  2005   -0.68636818  0.03125192  0.19663127
  2000 vs.  2007   -2.15130341  0.08303194  0.02093223
  2000 vs.  2009   -2.30615495  0.10649360  0.03155842
  2001 vs.  2003   -1.14584772  0.08272332      NaN
  2001 vs.  2004   -0.76802794  0.08283086      NaN
  2001 vs.  2005   -1.43872743  0.12195933  0.08751927
  2001 vs.  2007   -1.85908602  0.10849158  0.04403912
  2001 vs.  2009   -2.78681255  0.20321745  0.01496853
  2003 vs.  2004   0.78856413  -0.10042125      NaN
  2003 vs.  2005   -0.24282238  0.01758170  0.39029419
  2003 vs.  2007   -1.65612660  0.12478133  0.06245641
```

```
  2003 vs.  2009   -1.55599105  0.07758326  0.07148734
  2004 vs.  2005   0.81794716  -0.08024633  0.78591231
  2004 vs.  2007   -0.74756948  0.04883706  0.21328981
  2004 vs.  2009   -0.63125946  0.03532870  0.25122520
  2005 vs.  2007   -1.72772094  0.08789296  0.06336632
  2005 vs.  2009   -2.75126544  0.09534575  0.01382705
  2007 vs.  2009   -3.88369488  0.19817894  0.00538565
------------------------------------------------------------------

****************************** MRPP finished ******************************
10 Aug 2016, 17:02:10
```

Appendix 6. NMS axis 1 and 2 Jordan River UDWQ Asssessment Unit 7

```
JR114429   0.33746   0.47649
JR114433  -0.88169  -0.22025
JR115117   0.81374   0.75000
JR117487   0.39611   0.54699
JR118510   0.88902   0.35158
JR121480   0.51490   0.09774
JR124961  -0.56378   0.07215
JR126843   0.62465  -0.98037
JR127346  -0.66254   0.46975
JR127668   0.53183  -1.63792
JR127669  -0.74071   0.27884
JR129968   1.07759   0.10265
JR142111  -1.10054   0.04588
JR142112  -1.23603  -0.35351
```

Appendix 7. NMS axis 1 and 2 Jordan River UDWQ Asssessment Unit 7 October Samples

```
Samples      Axis
Number Name       1
   1 JR115117  -0.8240
   2 JR118510  -0.8127
   3 JR126843   1.2191
   4 JR127668   1.2304
   5 JR129968  -0.8127
```

Version 1.4

Is Reliance on a Single Bioassessment Metric for Assessing Water Quality in Utah's Rivers and Streams Prudent?

Draft Technical Report

August 28, 2016

To:
Jordan River/Farmington Bay Water Quality Council
Salt Lake City, UT

By
David C. Richards, Ph. D.
OreoHelix Consulting
P. O. Box 996
Moab, UT 84532
Phone: 406.580.7816
Email: oreohelix@icloud.com

SUMMARY

Utah is blessed with many irreplaceable rivers and streams despite being the 2nd driest state in the USA. Utah's human population, water demands, and booming economy are growing exponentially throughout much of the state and are completely depend on increasingly limited clean water supplies. Evaluating and protecting the health of Utah's rivers and streams is now crucial and will become even more so into the foreseeable future and is reliant on whose citizens and economy are well positioned to appreciate and fund protection.

The State of Utah Department of Water Quality (UDWQ) is responsible for assessing, monitoring, and protecting the 'physical, chemical, and biological integrity' of its waters based on the Clean Water Act (CWA) and by UDWQ's designated 'beneficial uses' under state law. Biological integrity is the cornerstone upon which the health of a river or stream is measured and biological assessments are one of the most important and useful management tools available for restoring and maintaining biological integrity. Bioassessments have been developed for many years and are widely used by management agencies for wadeable waters throughout the world, however, Utah is the only state in the western USA that entrusts its river and stream bioassessments entirely to a single taxa richness based metric, "River Invertebrate Prediction and Classification System" (RIVPACS O/E). All other western state water quality programs in the region integrate multimetric methods. O/E models are complex and are based on many assumptions and generalizations; some of which lead to a poor evaluation of biological integrity. An impaired listing based on O/E can have significant economic penalties on water users. Consequently, the reliance on any single metric such as O/E in a bioassessment program may not be prudent.

A statistical evaluation of O/E as it relates to evenness and other metrics and the effects of subsampling on these metrics was conducted. A discussion of the consequences of a > 50% probability of capture criterion in O/E models and their ability to actually monitor biological integrity is also discussed, as well as some other concerns including a comparison between bioassessment programs in UT and surrounding states.

Macroinvertebrate datasets were obtained from the Bureau of Land Management/Utah State University Buglab database and the Utah Department of Water Quality data that were used in their 2016 draft Integrated Report. Compatible data were merged and filtered to reduce spatial variability. Several metrics reported by the Buglab were examined; O/E score, Taxa Richness, % Labsplit, Abundance, Shannon Diversity, Simpson Diversity and Evenness. Pairwise correlations, linear and quadratic Ordinary Least Squares (OLS) regressions, simultaneous quantile

regressions at the 25th, 50th, and 75th quantiles and Path Diagrams and Structural Equation Models (SEM) were developed.

Evenness and taxa richness were the most important metrics directly and indirectly effecting O/E scores. SEM results suggest that a 1 standard deviation change in evenness (0.14) equaled a 0.96 standard deviation change in O/E scores = 0.22 (0.18 to 0.26, 95% CIs). As little of a change in evenness of approximately 5% can lead to a change from an O/E score of 0.76 (fully supporting) to 0.69 (not supporting) and unrelated to impairment.

A hypothetical but realistic example of the effects of evenness and subsampling on taxa richness resulted in a detection of all taxa in the completely even sample compared to a detection of < 50% of the taxa in an uneven sample when in fact all the same taxa occurred in the original uneven and even samples. Thus natural fluctuations in evenness in a river or stream without a loss or extinction event resulting from human caused impairment could trigger an unjustified management response from 'fully supporting' to 'not supporting'. A real world example is the Jordan River, listed as impaired by UDWQ. Analysis showed that O/E scores should have been rated higher if the effects of subsampling and evenness were considered.

Reliance on a complicated, computationally expensive, generalized, non-site specific metric such as that produced by a RIVPACS O/E model may not be prudent. Replacing the O/E metric with one or several of the other correlated metrics should be considered. At the minimum, these metrics should also be included in a bioassessment program. The decision to use a probability of capture $\geq$ 50% in an O/E model has very strong negative consequences for assessing the biological integrity of Utah's river or streams. Uncommon and rare taxa should always be included in ecological assessments. Detection of impacts will be enhanced by including these taxa because they are often the first to become extinct due to human disturbance. Uncommon and rare taxa have also been shown to disproportionally contribute to ecosystem function and integrity. Their unmeasured loss could fail to warn of an impending ecological shift.

Many RIVPAC O/E users continue to insist that a reduction in O/E scores reflects the extent to which taxa have become locally extinct due to human activities. This is clearly not the case. In many instances, taxa weren't lost; they just weren't found. To continue to assume that native taxa have become locally extinct because O/E scores have decreased reflects a gross misinterpretation of RIVPACS O/E models. There is also no shortage of additional informative metrics used by other state water quality management agencies, including those with fewer resources and human populations than Utah. Utah should follow suit, otherwise it will lag far behind.

Even though a RIVPACS O/E model has the potential to be a useful summary metric: its use as a stand-alone metric is not recommended. O/E relies on too many assumptions, constraints, and inherent errors that necessitates its inclusion into a more comprehensive macroinvertebrate multimetric program. Fewer incorrect assessments of impairment will be made by incorporating the O/E metric into a multimetric program than if used alone. Unfortunately, all metrics are affected by the evenness of a sample and subsampling. This phenomenon needs to be considered in any bioassessment program. The O/E probability of capture < 50% constraint results in a poor evaluation of macroinvertebrate assemblages and thus fails to measure true biological integrity. With Utah's booming economy and exponentially growing population, UDWQ now has the opportunity to build a bioassessment program worthy of its unique rivers and streams.

Table of Contents

List of Figures

List of Tables

List of Appendices

Introduction

Utah is blessed with many irreplaceable rivers and streams, including well known rivers such as the Provo, Bear, Weber, Green, Virgin, San Juan, and Colorado Rivers. Utah is also the 2nd driest state in the USA with human population and water demands increasing exponentially throughout much of the state, particularly along the Wasatch Front in the Greater Salt Lake City metropolitan area. At the same time, Utah's booming economy driven by high tech, high paying jobs has been called "the new economic Zion" (Newsweek 2010). Evaluating and protecting the health of Utah's rivers and streams is vital and will become ever more important into the foreseeable future in a state whose citizens and economy are well positioned to appreciate and fund protection.

"The most direct and effective measure of integrity of a water body
is the status of its living systems".
(Karr and Chu 1997)

The State of Utah Department of Water Quality (UDWQ) is responsible for assessing, monitoring, and protecting the 'physical, chemical, and biological integrity of its waters based on the Clean Water Act (CWA) and by UDWQ's designated 'beneficial uses' under state law. Physical and chemical integrity are manifested in biological integrity. The natural biotic community can only be maintained when physical and chemical conditions are suitable and in good condition. Biological integrity is the cornerstone upon which the health of a river or stream is measured. Although physical and chemical integrity have not been well expressed by regulatory agencies; the definition and understanding of biological integrity has been conferred at length by aquatic ecologists and subsequently simplified for adoption by regulators. One of the most widely recognized definitions of biological integrity is from Karr and Dudley (1981) (adapted from Frey 1977), "the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region". This definition implies that aquatic ecosystems operate on several levels. These parts that sustain and contribute to an aquatic ecosystem's functioning are quantifiable (Karr 1991) and need to be understood in the context of their surrounding environments and evolutionary history (Wikipedia 2014). Of course, the definition of biological integrity presented here is a condensed version taken from Karr and Dudley (1981) and there are many other aspects and definitions of biological integrity that are often ignored by management agencies but should be considered including for example, genetics and metapopulation dynamics.

Biological assessments and biocriteria are one of the most important and useful management tools available for restoring and maintaining the biological integrity of rivers and streams.

Bioassessments rely on empirical knowledge of how a wide range of biological attributes responds to varying degrees of human influence (Karr 1993; Karr and Chu 1997). The most useful bioassessments explicitly embrace several attributes of the biotic assemblages including: taxa richness, indicator taxa (e.g., tolerant and intolerant groups), and assessments of processes such as trophic structure, feeding strategies and other taxa traits. Simply stated, the goal of bioassessments is to measure and evaluate the consequences of human actions on biological systems (Karr 1993; Karr and Chu 1997).

Bioassessments have a long history and are widely used by management agencies primarily for wadeable waters (i.e. streams and small rivers) worldwide. However, Utah is the only state in the western USA that entrusts its river and stream bioassessments entirely to a single taxa richness based metric, "River Invertebrate Prediction and Classification System" (RIVPACS O/E). All other western state water quality programs understand that river and stream ecosystems operate on several complex ecological levels and understand the importance of combining and utilizing a suite of metrics, which typically include several richness, diversity, trait, and functional metrics but may or may not include a RIVPACS O/E model.

The O/E metric is simply the relationship between the observed (O) taxa and the expected (E) taxa in a river or stream. If the number of observed taxa is less than the number of expected taxa, managers often conclude a loss of taxa and diversity and hence a loss of biological integrity and compromised water quality. However, the RIVPACS O/E model is mathematically complex and relies on several summary and averaged watershed descriptors in model construction to predict "E", under least- impaired, reference conditions. In contrast, other commonly used taxa richness, diversity, and evenness metrics are straight forward, easy to calculate, and do not rely on average watershed descriptors for development.

RIVPACs O/E models also integrate a 'probability of capture' in the development of the "E" component. UDWQ uses a 50% probability of capture level (UDWQ 2016), which effectively eliminates invertebrate taxa that occur in < 50% of its 'reference' streams. This has important consequences and can severely misjudge levels of impairment and eliminate the ability to monitor taxa that may be unique to a river or stream and which are a fundamental part of its biological integrity; i.e. taxa that are not cosmopolitan and ubiquitous. In addition, because costs of taxonomic identification are purportedly large, invertebrate samples collected by management agencies are subsampled before metrics are calculated, including O/E scores. The effects of subsampling on O/E and other metrics may be substantial and can be affected by how evenly taxa abundances occur in a river or stream and represented in a sample (evenness). Effects of subsampling and evenness can then contribute to a misinterpretation of these metrics potentially resulting in unreliable assessments of water quality.

Justification

UDWQ is the only state in the western USA that relies on a single metric, O/E, for evaluating the complex biological integrity of Utah's many diverse rivers and streams. UDWQ uses O/E to determine whether to list a river or stream as biologically impaired or not. An impaired listing based on O/E can have significant economic penalties on water users, therefore the reliance on any single metric such as O/E in a bioassessment may not be prudent. A statistical evaluation of O/E as it relates to evenness and other metrics and the effects that subsampling has on these metrics was needed. A discussion of the consequences of a > 50% probability of capture criterion on actually monitoring biological integrity also needed to be addressed, as well as a comparison between bioassessment programs in UT and surrounding states.

Methods

Dataset

Macroinvertebrate datasets were obtained from the Bureau of Land Management/Utah State University Buglab database and the Utah Department of Water Quality data that were used in their 2016 draft Integrated Report. The Buglab dataset included all samples that were analyzed in their lab for the UDWQ bioassessment program from May 11, 1998 to October 16, 2014 (N = 1341 samples). The UDWQ dataset included samples collected from May 12, 2009 to October 16, 2014 (N = 797 samples). The two datasets were merged for a total of 513 samples. UDWQ has determined that 'mountain' and 'desert' expected number of taxa (E) are substantially different and developed their O/E scores accordingly. To help eliminate the effects of this bias, the dataset was filtered to include only the following 'mountain' management units: Bear River, Jordan River, Uinta Basin, Utah Lake, and Weber River. The dataset was then filtered to remove all samples that were not sub sorted to help understand the effects of subsampling and resulted for a final sample size of 262. BLM/USUS Buglab randomly splits a percentage of the sample to obtain 600 random individuals. They then computationally resample 300 of these organisms before calculating their metrics.

Statistical Analyses

Histograms were examined for normality for the following metrics that were *a priori* expected to be redundant and correlated:

1. O/E score
2. Taxa Richness
3. % Labsplit
4. Abundance
5. Shannon Diversity
6. Simpson Diversity and,
7. Evenness

All of the metrics were approximately normally distributed, except two, % Labsplit and Abundance. Transformations were not considered necessary. Pearson product-moment pairwise correlations with p-values were then calculated for the seven metrics.

Linear and quadratic Ordinary Least Squares (OLS) regressions were then computed for combinations of the metrics (predictor variables) that were most correlated with O/E scores (dependent variable) (%Labsplit and Abundance omitted). Simultaneous quantile regressions were also performed at the 25th, 50th, and 75th quantiles to examine if relationships between the dependent variable O/E score varied at different predictor values. Because standard regression analyses such as OLS cannot adequately evaluate indirect effects of a predictor variable on a response variable (e.g. O/E), Path Diagrams and Structural Equation Models (SEM) were developed for the most promising combinations of predictors. Both OLS and SEM are confirmatory statistical models and were used as such. All statistical analyses were conducted using Stata 14.1 for Mac (StataCorp 2016).

Results

Metric Correlations

There were strong significant correlations between many of the subsample derived metrics including those with O/E scores (Table 1). Taxa richness, Shannon and Simpson diversity and evenness metrics were strongly correlated with O/E (Table 1). The two diversity metrics are based on taxa richness and evenness and as expected were strongly correlated.

Table 1. Pearson correlations between richness and diversity metrics (= significant at p < 0.05)*

	O/E score	Richness	LabSplit	Abundance	ShannonD	Simpson
Richness	0.7581*					
LabSplit	-0.1295*	-0.1425*				
Abundance	-0.0896	-0.1095	-0.3895*			
ShannonD	0.7029*	0.8678*	-0.1152	-0.0965		
Simpson	0.6007*	0.724*	-0.0874*	-0.072	0.9453*	
Evenness	0.5612*	0.6581*	-0.071	-0.0566	0.9408*	0.9661*

Summary Statistics

The following three tables (Tables 2-4) are summary statistics for taxa richness, evenness, and O/E scores.

Table 2. Summary statistics for taxa richness metric.

Richness*

	Percentiles	Smallest		
1%	6	4		
5%	8	6		
10%	10	6	Obs	262
25%	13	6	Sum of Wgt.	262
50%	19		Mean	18.9542
		Largest	Std. Dev.	6.681435
75%	24	34		
90%	28	34	Variance	44.64157
95%	29	34	Skewness	.0542729
99%	34	35	Kurtosis	2.274394

Table 3. Summary statistics for evenness metric.

Evenness*

	Percentiles	Smallest		
1%	.2484093	.061079		
5%	.4228132	.1969372		
10%	.4709411	.2484093	Obs	262
25%	.5750272	.2599899	Sum of Wgt.	262
50%	.6802697		Mean	.6538558
		Largest	Std. Dev.	.1367357
75%	.754016	.86718		
90%	.8037984	.8762848	Variance	.0186967
95%	.8310941	.8844925	Skewness	-.9792063
99%	.8762848	.8874369	Kurtosis	4.237731

Table 4. Summary statistics for O/E score metric.

OE_SCORE

	Percentiles	Smallest		
1%	.4	.3		
5%	.5	.399		
10%	.607	.4	Obs	262
25%	.754	.411	Sum of Wgt.	262
50%	.908		Mean	.9069275
		Largest	Std. Dev.	.2336467
75%	1.084	1.354		
90%	1.211	1.36	Variance	.0545908
95%	1.288	1.419	Skewness	-.1137977
99%	1.36	1.423	Kurtosis	2.504571

The BLM/USU Buglab estimated total abundances in the filtered dataset ranged between 965 and 211,604 individuals/m^2 and the proportion of the samples used for subsampling ranged between 1.14 and 87.5%.

OLS regression

The best OLS regression model at predicting O/E scores included two predictors, richness and evenness and resulted in an R^2 of 0.58 (Table 1, Figures 1 and 2). There was very little difference in either a linear or quadratic model, therefore only the more easily interpretable linear model results are included in Table 5. OLS models were also computed with no constant (forcing the fit through the origin) because it was assumed that when richness and evenness were zero, so were O/E scores. Models without constants produced substantially lower coefficient standard errors indicating an improvement of OLS models over those that included constants although models with or without constants produced slopes that were significantly greater than zero. However, interpretation of OLS models without a constant is more difficult compared with models with constants that produce statistically relevant R^2 values. Therefore, only the OLS model that included the estimated constant is reported (Table 5). Simultaneous quantile regression coefficients were not significantly different than the final OLS coefficients indicating relatively consistent prediction of the OLS model throughout the range of data.

Table 5. OLS regression results of O/E score as a function of taxa richness and evenness.

```
. regress oe_score richness  evenness
```

Source	SS	df	MS		
Model	8.28553352	2	4.14276676	Number of obs	= 262
Residual	5.96265445	259	.023021832	F(2, 259)	= 179.95
				Prob > F	= 0.0000
				R-squared	= 0.5815
				Adj R-squared	= 0.5783
Total	14.248188	261	.054590758	Root MSE	= .15173

oe_score	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
richness	.0239824	.001867	12.85	0.000	.0203059	.0276589
evenness	.1876308	.0912297	2.06	0.041	.0079845	.3672772
_cons	.329677	.0460416	7.16	0.000	.2390135	.4203405

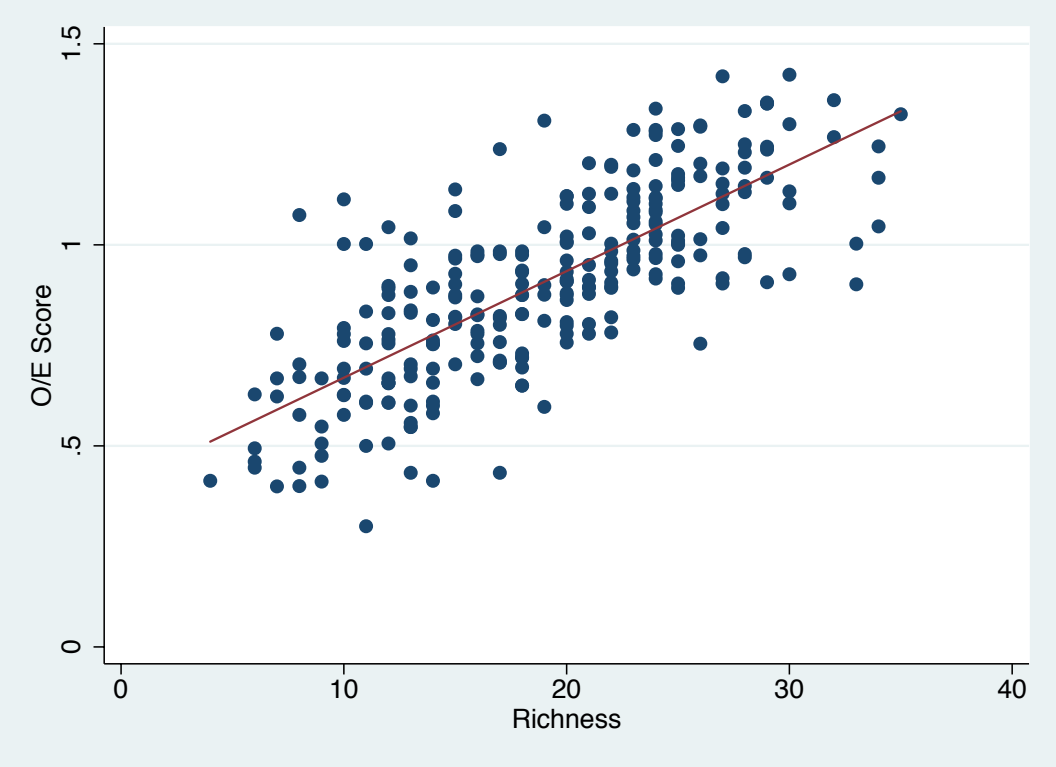


Figure 1. Relation of O/E scores to taxa richness. Red line is OLS regression linear fit. See Table 5 for OLS results.

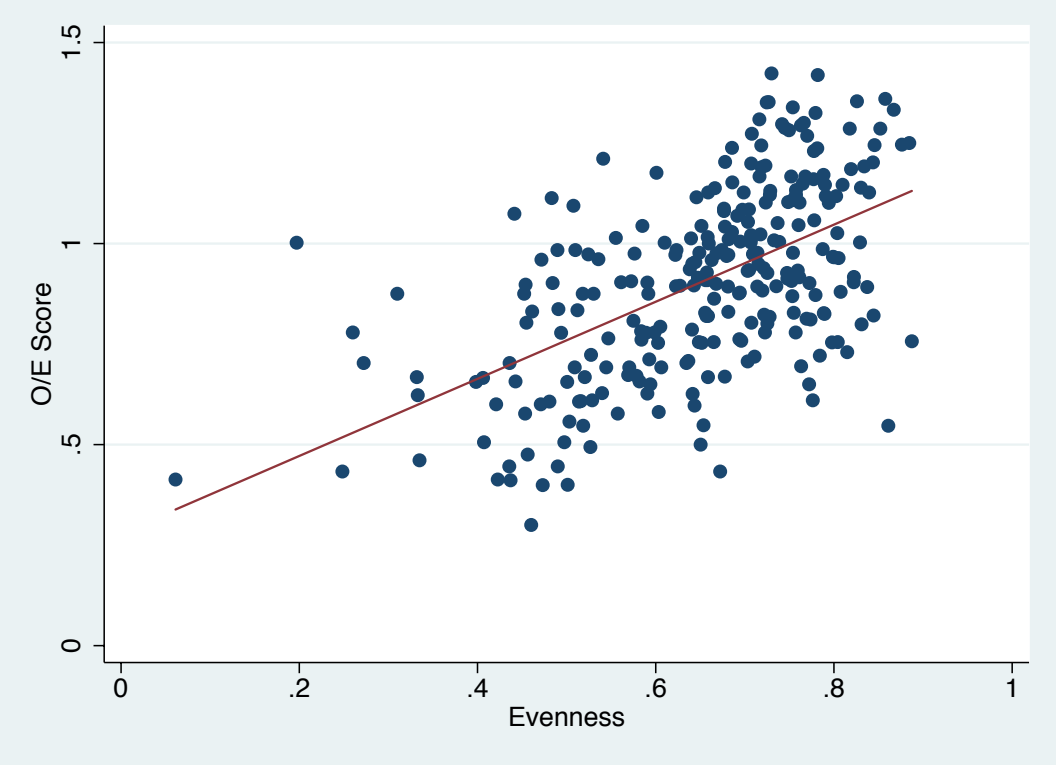
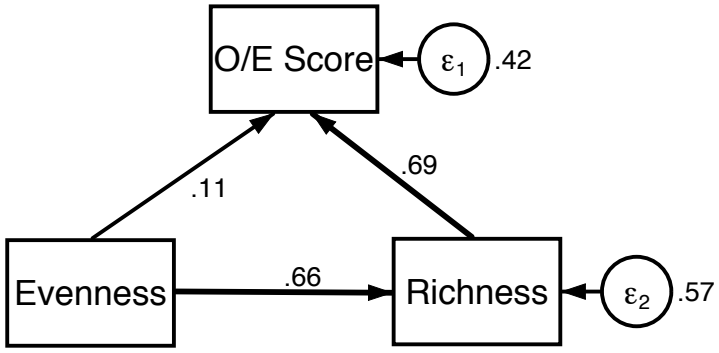


Figure 2. Relation of O/E scores to evenness. Red line is OLS regression linear fit. See Table 5 for OLS results

The linear relationship between O/E and taxa richness was much less variable than the relation between O/E and evenness as shown in Table 5 (standard errors and 95% CIs of p-values) and Figures 1 and 2.

Structural Equation Models

SEMs and path diagrams were conducted using all seven of the predictor metrics of O/E scores. The best SEM also included only richness and evenness as predictors of O/E scores (Figure 3 and Tables 3 - 6). This SEM model was a good predictor of O/E scores and also produced the same as the OLS model above, $R^2 = 0.58$. Richness had a strong direct effect on O/E (0.69) and evenness had significant but lesser direct effect (0.11). However, evenness had a direct effect on richness (0.66) and therefore an indirect effect on O/E scores for a total effect on O/E of 0.56 similar to the effect richness had on O/E scores (0.69) (Figure 3, Tables 6-9).



O/E: $R^2 = 0.58$
Richness: $R^2 = 0.43$
Overall: $R^2 = 0.44$

O/E Total Effects
Richness = 0.69
Evenness = 0.56

Figure 3. Structural Equation Model (SEM) and path diagram of the direct and indirect effects of richness and evenness metrics on O/E scores (values in the figure are standardized).

Table 6. Standardized SEM results of the effects of richness and evenness on O/E scores.

Structural equation model Number of obs = **262**
Estimation method = **ml**
Log likelihood = **-520.69055**

Standardized	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
oe_score <-						
richness	**.6858081**	**.0454566**	**15.09**	**0.000**	**.5967148**	**.7749015**
evenness	**.1098061**	**.0529155**	**2.08**	**0.038**	**.0060937**	**.2135186**
_cons	**1.413707**	**.2305396**	**6.13**	**0.000**	**.961858**	**1.865557**
richness <-						
evenness	**.6581465**	**.0309963**	**21.23**	**0.000**	**.5973949**	**.7188982**
_cons	**-.3109353**	**.2184783**	**-1.42**	**0.155**	**-.7391449**	**.1172743**
var(e.oe_score)	**.4184851**	**.0377126**			**.3507295**	**.4993301**
var(e.richness)	**.5668432**	**.0408002**			**.4922607**	**.6527257**

Table 7. Standardized SEM direct effects of richness and evenness on O/E scores.

Direct effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
oe_score <-						
richness	**.0239824**	**.0018563**	**12.92**	**0.000**	**.0203441**	**.0276207**
evenness	**.1876308**	**.0907059**	**2.07**	**0.039**	**.0098506**	**.3654111**
richness <-						
evenness	**32.15957**	**2.272838**	**14.15**	**0.000**	**27.70489**	**36.61425**

Table 8. Standardized SEM indirect effects of richness and evenness on O/E scores.

Indirect effects

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
oe_score <-						
richness	**0**	(no path)				
evenness	**.7712634**	**.0808391**	**9.54**	**0.000**	**.6128218**	**.9297051**
richness <-						
evenness	**0**	(no path)				

Table 9.Standardized SEM total effects of richness and evenness on O/E scores.

Total effects

	OIM Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Structural						
oe_score <-						
richness	**.0239824**	**.0018563**	**12.92**	**0.000**	**.0203441**	**.0276207**
evenness	**.9588942**	**.0873777**	**10.97**	**0.000**	**.7876371**	**1.130151**
richness <-						
evenness	**32.15957**	**2.272838**	**14.15**	**0.000**	**27.70489**	**36.61425**

The interpretation of the SEM standardized loadings (coefficients) for total effects on O/E scores is fairly straight forward. A 1 standard deviation change in taxa richness results in a 0.23 standard deviation change in O/E score (Table 9). Likewise, a 1 standard deviation change in evenness (0.14) (Table 3) equals a 0.96 standard deviation change in O/E scores = 0.22 (Table 4 and Table 9). The 95% CIs of evenness coefficients are quite wide; 0.79 to 1.13 (Table 9). This suggests that if evenness changes by 1 std. dev., O/E scores could change between 0.18 and 0.26.

The relationship between O/E and taxa richness was much less variable than the relation between O/E and evenness in SEM as shown in Table 6 (standard errors and 95% CIs of p-values).

Effects of evenness on taxa richness: hypothetical example

The following tables, Table 7 and Table 8 are hypothetical invertebrate samples that both have the same number of individuals, N = 30,000, a typical number of individuals that occur in a UDWQ river or stream sample. The abundance and proportional abundance of each of the 30 taxa is the same (evenness = 1.0) in idealized Table 1. In Table 8, abundances and proportion abundances are exaggerated with one taxon, ‘dd’ having substantially more individuals.

On average a random 300 count subsample from the sample taxa pool in Table 7 would result in 30 observed taxa, which is the actual true number of taxa in the taxa pool. In contrast, the number of taxa observed from a random 300 count subsample taken from the sample taxa pool in Table 8 would only be 14. This would be a > 50% reduction in taxa observed even though there were still 30 taxa in the original sample and there was no loss of taxa from the site.

Table 10. Hypothetical invertebrate sample with 30,000 individuals and equal proportional abundances. A fixed 330 count subsample would result in 10 individuals counted per taxon and the total number of taxa observed = 30.

Taxon	Abundance	Proportion Abundance	300 Count	Taxa Counted
a	1000	0.033	10	1
b	1000	0.033	10	1
c	1000	0.033	10	1
d	1000	0.033	10	1
e	1000	0.033	10	1
f	1000	0.033	10	1
g	1000	0.033	10	1
h	1000	0.033	10	1
i	1000	0.033	10	1
j	1000	0.033	10	1
k	1000	0.033	10	1
l	1000	0.033	10	1
m	1000	0.033	10	1
n	1000	0.033	10	1
o	1000	0.033	10	1
p	1000	0.033	10	1
q	1000	0.033	10	1
r	1000	0.033	10	1
s	1000	0.033	10	1
t	1000	0.033	10	1
u	1000	0.033	10	1
v	1000	0.033	10	1
w	1000	0.033	10	1
z	1000	0.033	10	1
y	1000	0.033	10	1
z	1000	0.033	10	1
aa	1000	0.033	10	1
bb	1000	0.033	10	1
cc	1000	0.033	10	1
dd	1000	0.033	10	1
Total Taxa Counted				**30**

Table 11. Hypothetical invertebrate sample with 30,000 individuals and unequal proportional abundances[1]. A fixed 300 count subsample would result in a variable number of individuals counted per taxon and the total number of taxa observed would equal 14.

Taxon	Abundance	Proportion Abundance	300 count	Taxa Counted
a	10	< 0.000	0.1	0
b	20	0.001	0.2	0
c	30	0.001	0.3	0
d	40	0.001	0.4	0
e	49	0.002	0.49	0
f	60	0.002	0.6	1
g	70	0.002	0.7	1
h	80	0.003	0.8	1
i	90	0.003	0.9	1
j	100	0.003	1	1
k	10	0.000	0.1	0
l	20	0.001	0.2	0
m	30	0.001	0.3	0
n	40	0.001	0.4	0
o	50	0.002	0.5	0
p	60	0.002	0.6	1
q	70	0.002	0.7	1
r	80	0.003	0.8	1
s	90	0.003	0.9	1
t	100	0.003	1	1
u	10	0.000	0.1	0
v	20	0.001	0.2	0
w	30	0.001	0.3	0
z	40	0.001	0.4	0
y	50	0.002	0.5	1
z	50	0.002	0.5	0
aa	50	0.002	0.5	1
bb	50	0.002	0.5	0
cc	50	0.002	0.5	1
dd	28551	0.952	285.51	1
Total Taxa Counted				**14**

[1] Although the proportion abundances in Table 11 are exaggerated, real samples are virtually never completely even (e.g. Table 10). This is because individual taxa abundances vary widely both spatially and temporally, sometimes biweekly in the case of short lived taxa. A high turnover taxon can be absent from a sample either because they are in the egg stage or aerial adults, while one to two weeks later their early instar abundances can dominate the sample. Younger, smaller instars are always more abundant than older, larger instars or adults and there are always both inter and intraspecific population abundance dynamics occurring.

Jordan River O/E Bioassessment Example

The Jordan River flows from Utah Lake and into the Great Salt Lake through the most densely populated area of Utah. By any measure, this river has been compromised. Subsequently, six out of the eight Jordan River UDWQ Assessment Units (Units 1, 2, 3, 4, 6, and 7) were considered impaired due to O/E scores in the UDWQ 2016 Integrated Report. However, a synthesis of the BLM/USU Buglab database revealed there were potentially >> 200 macroinvertebrate taxa in the Jordan River (Table 12). These taxa occurred at various probability of captures (see Discussion section on probability of capture problems). Some taxa were highly abundant and widely distributed throughout the length of the river, while other taxa were rare and uncommon spatially and abundance based (Table 13).

Table 12. List of taxa found in Jordan River using BLM/USU Buglab database. Accessed July 11, 2016.

Taxon	**Taxon**	**Taxon**
Ablabesmyia	Cheumatopsyche	Ephemera
Acari	Chironomidae	Ephemerella
Aeshna	Chironominae	Ephemerellidae
Aeshnidae	Chironomini	Ephydridae
Ambrysus	Chrysops	Erpobdella punctata
Amphipoda	Cinygmula	Erpobdellidae
Anagapetus	Cladotanytarsus	Eukiefferiella
Anax	Cleptelmis addenda	Fallceon quilleri
Anax walsinghami	Clitellata	Ferrissia
Ancyronyx	Coenagrionidae	Ferrissia rivularis
Antocha	Coleoptera	Fluminicola coloradoensis
Antocha monticola	Collembola	Fossaria
Apedilum	Corbicula	Gammarus
Apsectrotanypus	Corisella	Gastropoda
Argia	Corixidae	Glossiphonia complanata
Argia emma	Corticacarus	Glossiphoniidae
Asellidae	Corydalus	Glossosoma
Atherix lantha	Corynoneura	Gomphidae
Atrichopogon	Crangonyx	Gyraulus
Baetidae	Cricotopus	Gyrinus
Baetis	Cricotopus bicinctus group	Hagenius
Baetis tricaudatus	Cricotopus trifascia group	Haliplidae
Baetisca	Cricotopus/Orthocladius	Harnischia
Berosus	Cryptochironomus	Helisoma
Bezzia/Palpomyia	Cryptotendipes	Helobdella stagnalis
Bivalvia	Curculionidae	Helodon
Boyeria	Diamesa	Hemerodromia
Brachycentrus	Dicrotendipes	Heptagenia elegantula/solitaria
Buenoa	Didymops	Heptageniidae
Caecidotea	Dina dubia	Hesperocorixa
Caenis	Diptera	Hesperoperla pacifica
Callibaetis	Dolichopodidae	Hetaerina

Calopteryx	Drunella coloradensis/flavilinea	Hetaerina americana
Cambarinae	Drunella doddsii	Hetaerina vulnerata
Centroptilum/Procloeon	Dubiraphia	Heterlimnius corpulentus
Ceratopogonidae	Elmidae	Hyalella azteca
Ceratopsyche	Enchytraeidae	Hydrobiidae
Chauliodes	Enochrus	Hydrophilidae
Chelifera	Epeorus	Hydrophilus

Hydropsychidae	Nephelopsis obscura	Probezzia
Hydroptila	Nigronia	Procladius
Hydroptilidae	Nilothauma	Promoresia
Hygrobates	Notonectidae	Prostoma
Hygrobatidae	Odonata	Protzia
Ischnura	Oecetis	Psectrocladius
Isonychia	Oligochaeta	Pseudocloeon
Isoperla	Oligostomis	Pseudosmittia
Kiefferulus	Ophiogomphus	Psychoda
Labrundinia	Optioservus	Pteronarcys
Laccophilus	Optioservus quadrimaculatus	Pyrgulopsis
Laccophilus maculosus	Orconectes virilis	Pyrgulopsis kolobensis
Lepidostoma	Ordobrevia nubifera	Pyrgulopsis pilsbryana
Leptoceridae	Orthocladiinae	Rhantus
Leptohyphidae	Orthocladius	Rheopelopia
Leptophlebiidae	Ostracoda	Rheosmittia
Leuctridae	Oxyethira	Rheotanytarsus
Libellula	Pagastia	Rhithrogena
Libellulidae	Pantala hymenaea	Rhyacophila
Lopescladius	Parakiefferiella	Rhyacophila brunnea/vemna group
Lumbriculidae	Paraleptophlebia	Rhyacophila vofixa group
Lymnaeidae	Parametriocnemus	Robackia
Macronychus	Parapsyche elsis	Saetheria
Macrostemum	Paratanytarsus	Sepedon
Micrasema	Pentaneurini	Sialis
Microcylloepus	Perlesta	Sigara
Microcylloepus pusillus	Phaenopsectra	Simuliidae
Microcylloepus similis	Philopotamidae	Simulium
Micropsectra	Phylocentropus	Simulium vittatum group
Microtendipes	Physa	Sperchon
Muscomorpha	Pisidiidae	Sperchonidae
Naididae	Pisidium	Stagnicola
Nanocladius	Planariidae	Stempellinella
Nectopsyche	Planorbidae	Stenelmis
Nemata	Plauditus	Stenochironomus
Nemouridae	Polypedilum	Stenonema
Neoplasta	Pomacea bridgesi	Stictochironomus
Neothremma	Potamopyrgus antipodarum	Stratiomyidae
Neozavrelia	Potthastia	Sublettea

Synorthocladius		
Tabanidae		

Taeniopterygidae		
Taeniopteryx		
Tanypodinae		
Tanytarsini		
Tanytarsini		
Tanytarsus		
Thienemanniella		
Thienemannimyia		
Tipulidae		
Torrenticola		
Tribelos		
Trichoptera		
Tricorythodes		
Tricorythodes minutus		
Trombidiformes		
Tropisternus		
Tubificidae		
Turbellaria		
Tvetenia		
Wiedemannia		
Wiedemannia		
Xylotopus		
Zaitzevia		
Zapada columbiana		
Zapada columbiana/oregonensis group		
Zygoptera		

The family Chironomidae accounted for 53% of the Dominant Family entries (N = 17/32) and Hydropsychidae and Leptohyphidae each accounted for 16% (5/32)(Table 13). These three families accounted for a total of 84% of the entries. Percent dominance by abundance ranged from 13% to 95% (Table 13).

Table 13. Jordan River evenness score, Dominant Family, % Dominant Family, Dominant Taxon, and % Dominant Taxon from 32 BLM/USU Buglab samples.

Evenness	Dominant Family	% Dominant Family	Dominant Taxon	% Dominant Taxon
0.44	Chironomidae	17.51	Oligochaeta	75.06
0.62	Chironomidae	13.03	Oligochaeta	54.38
0.38	Leptohyphidae	74.65	Tricorythodes	74.65
0.38	Chironomidae	88.35	Chironominae	77.37
0.50	Hydrobiidae	63.29	Hydrobiidae	63.29
0.58	Hydropsychidae	48.03	Hydropsyche	48.03
0.60	Chironomidae	61.95	Orthocladiinae	47.94
0.47	Leptohyphidae	64.84	Tricorythodes	64.84
0.66	Hydropsychidae	50.12	Hydropsyche	35.95
0.59	Hydropsychidae	37.52	Tricorythodes	32.90

0.53	Leptohyphidae	66.74	Tricorythodes	66.74
0.65	Hydrobiidae	39.10	P.antipodarum	39.10
0.57	Hydropsychidae	74.93	Hydropsyche	62.47
0.56	Chironomidae	53.49	Orthocladiinae	50.18
0.69	Leptohyphidae	39.71	Tricorythodes	39.71
0.59	Chironomidae	44.75	Orthocladiinae	39.97
0.53	Leptohyphidae	47.84	Tricorythodes	46.40
0.57	Chironomidae	57.02	Orthocladiinae	55.18
0.73	Chironomidae	42.35	Orthocladiinae	38.45
0.78	Simuliidae	34.11	Simulium	34.11
0.73	Corbiculidae	46.39	Corbicula fluminea	46.39
0.46	Hydropsychidae	73.32	Hydropsyche	65.47
0.49	Chironomidae	89.56	Chironominae	53.09
0.60	Chironomidae	53.83	Simulium	41.35
0.68	Asellidae	26.89	Turbellaria	28.52
0.63	Chironomidae	48.06	Chironominae	38.69
0.40	Chironomidae	94.49	Chironominae	69.01
0.57	Chironomidae	61.47	Chironominae	52.91
0.67	Chironomidae	72.51	Chironominae	37.31
0.57	Chironomidae	78.40	Orthocladiinae	39.77
0.44	Chironomidae	49.49	Chironominae	46.27
0.39	Chironomidae	89.02	Chironominae	69.93

Only four samples were comparable between the two data sets. O/E scores and evenness are in Table 14.

Table 14.Four Jordan River sites with O/E scores and evenness values that were compatible between the two datasets.

Sample	Site	O/E score	Evenness
142111	Jordan River at Cudahy Lane	0.446	na
142112	Jordan River 1000 ft below South Davis Treatment Plant	0.446	0.44
142113	Jordan River at 3300 S Crossing	0.557	0.50
142114	Jordan River 1100 W 2100 S below confluence with Mill Creek	0.445	0.62

The % Dominant taxon was Oligochaeta (segmented worms) at 54% of the total abundance in the Jordan River Sample 142114(collected 9 November 2009)(Table 15). There were at least fifteen taxa in the subsampled results (Table 15) and likely more in the entire sample but were not counted due to dominance by Oligochaeta.

Table 15. Taxa observed in the Jordan River sample 142114. Oligochaeta comprised 54% of total abundances.

Taxon	Taxon

Bivalvia	Lebertia
Caecidotea	Leptoceridae
Chironomidae	Oligochaeta
Chironominae	Orthocladiinae
Coenagrionidae	Physa
Erpobdellidae	Pisidium
Gammarus	Planorbidae
Gyraulus	Potamopyrgus antipodarum
Helobdella stagnalis	Psychoda
Hydrobiidae	Sigara
Hydropsyche	Tanypodinae
Hydroptilidae	Turbellaria

Results from the OLS and SEM models showed that evenness had a strong effect on taxa richness and O/E scores, thus there were likely more taxa in the complete original Jordan River samples than in the subsamples. Even though the Jordan River is obviously impaired (i.e. it is a highly regulated, dewatered urban/industrial system); O/E scores should have been higher in most of the Jordan River samples if the effects of subsampling and evenness were considered.

Discussion

There were strong effects of evenness and richness metrics on O/E scores, which apparently often affect biological assessments. Taxa richness obviously effects O/E scores because the O/E model is mostly based on this metric. Evenness directly effects taxa richness in a subsample and consequently directly and indirectly effects O/E scores. These effects need to be accounted for by water quality agencies before assigning an assessment score.

The unexplained variability in richness (E_2 = 0.57 in Figure 3) due to its relationship with evenness in the SEM and hence some of the unexplained variability in O/E scores was likely in part due to: 1) natural variability in richness in the different stream types and conditions, 2) varying levels of impairment, and more concerning, 3) the assumptions and variables that went into development of the O/E models. Taxa richness is often greater in mid elevation streams compared to headwaters or lower elevation streams (i.e. the river continuum). Richness is also greater in reference streams than impaired streams, which is why richness is the most widely used metric in bioassessments. A larger data set than the one used in these analyses would have allowed for the separation of stream types and varying levels of impairment and there likely would have been a much stronger relation between evenness and richness and these two variables with O/E. Nothing can be done about the assumptions and subsequent effects of the PRISM variables on the O/E scores evaluated in this analysis except to completely redo the models. RIVPAC O/E models as used by UDWQ rely on at least eleven ‘watershed’ based climate/environmental variables without defining ‘watershed’. For example, the Jordan River

watershed drains over 3800 square miles with elevations ranging from 11,900 ft. to 4200 ft. (http://www.utahcleanwater.org/jordan-river-watershed.html). Assigning impaired status to UDWQ Jordan River Assessment Units could have been inappropriate if UDWQ used the entire Jordan River watershed, as was detailed in their 2016 IR, to develop O/E models and averaged watershed "E" (expected) taxa probabilities of capture for final O/E scores (see Richards 2016a for more discussion on Jordan River macroinvertebrate assemblages in relation to water quality assessments using O/E).

Simple correlations showed that several commonly used and easily calculable metrics; taxa richness, Shannon's and Simpsons diversity indices, and evenness were significantly and strongly correlated with O/E scores (Table 1). This suggests that reliance on a complicated and computationally expensive, non- transparent, metric such as that produced by a RIVPACS O/E model may not be prudent and that replacing the O/E metric with one or several of the other correlated metrics should be considered. At the minimum, these metrics should be included in a bioassessment program and used to supplement O/E scores.

Implications of Evenness on O/E Scores and UDWQ bioassessments

UDWQ uses a mean O/E score of $\geq$ 0.76 as 'fully supporting' and in general, a score of $\leq$ 0.69 as 'not supporting' (UDWQ Integrated Report 2016). If the SEM standardized loadings (coefficients) for the total effects of evenness on O/E scores in Table 9 are reasonable, then that would suggest that a 0.07 decrease in O/E score from 0.76 (fully supporting) to 0.69 (not supporting) would only require a decrease in evenness of about 0.044 (0.037 to 0.053). As discussed in footnote 2, page …, taxa abundances in macroinvertebrate samples are rarely if ever even, and this relatively small change in evenness could easily trigger an assessment from 'fully supporting' to 'not supporting'.

RIVPACS O/E 'Probability of Capture' is Problematic

RIVPACS O/E models include a 'probability of capture' (P_c) component. P_c is the probability that a taxon occurs at a reference site and is used in the development of the "E" expected taxa list. To reduce 'noise' in results and to ease interpretation, many users, including UDWQ, use a $P_C \geq$ 50%. That is, the probability of a taxon occurring at a site is estimated to be greater than 50%. The decision to use a $P_c \geq$ 50% has very strong negative implications for assessing the biological integrity of a river or stream in UT. Many ecologists agree that uncommon and rare taxa should be included in ecological assessments and by including these taxa detection of impacts is improved (Turak and Koop 2003; Nijboer and Schmidt-Kloiber 2004). It is also widely recognized that rare taxa are the first to become extinct due to human disturbance (Leitao 2016). Uncommon and rare taxa have also been shown to disproportionally contribute to ecosystem function and integrity (Leitao 2016). For example, native bivalves are extremely important for

maintaining water quality via their filter feeding activity and of much concern for developing NH_3 criteria. However, bivalves do not occur in >50% of Utah's reference sites and unionids are likely on the brink of extinction in UT (Richards 2016b). A $P_C \geq 50\%$ may easily overlook many, many, taxa that are unique to Utah's rivers and streams including threatened and endangered species, important ecosystems providers, or simply an unknown number of taxa that occur in < 50% of reference streams. These taxa are the true measure of biological integrity and without which will result in a homogenous, biodiversity -limited condition lacking integrity. These taxa are also the most likely to be most sensitive to impacts because their niche breadth is much narrower that taxa that have P_cs > 50%. There is a well-known saying in ecology; 'rare is common, and common is rare' (Pimm et al. 2014). Modifications to RIVPACS O/E models have allowed researchers and managers in England to monitor rare species and to flag Red Data Book threatened species (http://www.ceh.ac.uk/services/rivpacs-reference-database), however they use much lower P_cs. Utah should consider the same.

Misinterpretation of O/E

Many RIVPAC O/E users continue to insist that a reduction in O/E scores reflects the extent to which taxa have become locally extinct due to human activities (UDWQ Integrate Report 2016). This is clearly not the case. The analyses included in this report highlight the fact that subsampling and evenness have significant effects on the number of taxa observed, especially the more uneven a sample and subsample. Taxa weren't lost; they just weren't found. They may not have even decreased in abundance. It is possible that other taxa could have disproportionally increased in abundance for whatever reason and that the 'lost' taxa simply weren't counted. To continue to assume that native taxa have become locally extinct because O/E scores have decreased reflects a gross misinterpretation of RIVPACS O/E models.

Additional Bioassessment Metrics in Use

There is no shortage of metrics in use by water quality management agencies throughout the USA including; richness, diversity, trait, and functional metrics. Each of these metrics addresses different aspects of biological integrity and combined into a suite can be highly useful in water quality assessments. Utah is the only state in the western USA that relies solely on a single metric, O/E. This can be analogous to a physician relying solely on body temperature to assess a person's health. Although measuring body temperature is highly useful, used alone, it cannot assess other ailments (e.g. broken leg, gunshot wound, cancer, etc.). BLM/USU Buglab processes the vast majority of UDWQ invertebrate samples and in addition to calculating O/E scores, automatically provides UDWQ with several dozen potentially useful metrics (Table 12). Surrounding states also include a suite of metrics in their bioassessment programs (Appendices 1-6). By not incorporating simple, easy to use and pertinent metrics, it appears that UDWQ now

lags far behind surrounding states in its bioassessment program. Several of these states also include separate multimetric indices using diatoms and fish metrics. At this time, UT does not use either.

Table 16. Condensed list of metrics that are routinely generated by BLM/ USU Buglab for UDWQ's bioassessment program.

Richness (metrics summarizing all unique taxa in a sample)	Richness
	Abundance
	Shannon's Diversity
	Simpson's Diversity
	Evenness
	# of EPT Taxa
	EPT Taxa Abundance
Dominance Metrics (metrics summarizing all most abundant taxa in a sample)	Dominant Family
	Abundance of Dominant Family
	Dominant Taxa
	Abundance of Dominant Taxa
Tolerance Indices (indices based on the indicator species concept in which taxa are assigned tolerance values)	Hilsenhoff Biotic Index
	# of Intolerant Taxa
	Intolerant Taxa abundance
	# of Tolerant Taxa
	Tolerant Taxa abundance
	USFS Community Tolerance Quotient (d)
Functional Feeding Groups (classification of organisms based on morphological or behavioral adaptations for where and how food is acquired)	# of shredder taxa
	Shredder Abundance
	# of scraper taxa
	Scraper abundance
	# of collector-filterer taxa
	Collector-filterer abundance
	# of collector-gatherer taxa
	Collector-gatherer abundance
	# of predator taxa
	Predator abundance
Functional Traits	# of clinger taxa
	"# of" Long-lived Taxa
Compositional Metrics (richness and abundance of various taxonomic groups)	# of Ephemeroptera taxa
	Ephemeroptera abundance
	# of Plecoptera taxa
	Plecoptera abundance
	# of Trichoptera taxa
	Trichoptera abundance
	# of Coleoptera taxa
	Coleoptera abundance
	# of Elmidae taxa
	Elmidae abundance
	# of Megaloptera taxa
	Megaloptera abundance
	# of Diptera taxa
	Diptera abundance
	# of Chironomidae taxa
	Chironomidae abundance
	# of Crustacea taxa
	Crustacea abundance
	# of Oligochaete taxa
	Oligochaete abundance
	# of Mollusca taxa
	Mollusca abundance
	# of Insect taxa
	Insect abundance

Economics vs. Bioassessment Quality.

All western USA states near UT had the same time frame allotted by EPA for developing their bioassessment programs. It does not appear that economic hardship or small population (taxpayer base) were factors in UDWQs decision to rely on a RIVPACs O/E metric as its sole measure of biological integrity (Table 13 and Appendices). Contrarily, Utah now seems economically poised to lead other states in the region in developing relevant and useful bioassessments. “According to the 2007 State New Economy Index, Utah is ranked the top state in the nation for Economic Dynamism, determined by "the degree to which state economies are knowledge-based, globalized, entrepreneurial, information technology-driven and innovation-based". In 2010, Utah was ranked number one in Forbes' list of "Best States For Business"(Badenhausen 2010). A November 2010 article in Newsweek magazine highlighted Utah and particularly the Salt Lake City area's economic outlook, calling it "the new economic Zion", and examined how the area has been able to bring in high-paying jobs and attract high-tech corporations to the area during a recession (Dokoupil 2010). As mentioned in the introduction and based on the U.S. Census Bureau statistics, Utah has one of the fastest growing populations of any U.S. state, 2nd in 2013. Table 13 compares surrounding states estimated gross state product and estimated population.

Table 17. Estimated gross state product and population for Utah and surrounding states (2010 data).

State	Gross State Product ($ billions)	Rank	Population (millions)	Rank
Utah	130.5	3	3.0	3
Colorado	257.6	2	5.5	2
Wyoming	38.4	8	0.6	8
Arizona	259.0	1	6.3	1
New Mexico	79.7	5	2.1	5
Idaho	58.2	6	1.7	6
Montana	44.3	7	1.0	7
Nevada	126	4	2.9	4

Conclusion

Even though RIVPACS O/E models have the potential to be useful summary metrics, their use as a stand-alone metric is not recommended. O/E models rely on far too many assumptions, constraints, and inherent errors that necessitates their inclusion into a more comprehensive and informative macroinvertebrate multimetric based program. By incorporating the O/E metric into a multimetric program fewer incorrect assessments of impairment will be made

than if it used alone. Unfortunately, all metrics are affected by the evenness of a sample and subsampling. This phenomenon needs to be considered in any bioassessment program. O/E probability of capture < 50% results in a poor evaluation of macroinvertebrate assemblages and thus fails to measure biological integrity. All states in the region other than Utah incorporate multimetric indices and several include the O/E metric, even states with fewer citizens and less resources. With Utah's booming economy and exponentially growing population, UDWQ now has the opportunity to build a bioassessment program worthy of its unique rivers and streams.

Literature Cited

Badenhausen, Kurt (October 13, 2010). "The Best States For Business And Careers". Forbes.

Dokoupil, Tony (November 8, 2010). "How Utah Became an Economic Zion". Newsweek.

Leitao, R. P. et al. 2016. Rare species contribute disproportionately to the functional structure of species assemblages. Proceedings of the Royal Society B: Biological Sciences. Vol. 283. Issue 1828.

Nijboer, R. C. and A. Schmidt-Kloiber. 2004. The effect of excluding taxa with low abundances or taxa with small distribution ranges on ecological assessment. Hydrobiologia. Vol. 515 1:347-363.

Pimm, S. L. et al. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science. Vol. 344. Issue 6187. Review.

Richards, D. C. 2016a. Real and Perceived Macroinvertebrate Assemblage Variability in the Jordan River, Utah can Effect Water Quality Assessments. Draft Technical Report. Submitted to the Jordan River/Farmington Bay Water Quality Council. Salt Lake City, UT. Oreohelix Consulting, Vineyard, UT.

Richards, D. C. 2016b. Recalculation of Ammonia Criteria for Central Valley Water Reclamation Facility's Discharge into Mill Creek, Salt Lake County, UT based on Native Unionoida Surveys and Metapopulation Dynamics. Final Report. Submitted to the Jordan River/Farmington Bay Water Quality Council. Salt Lake City, UT. Oreohelix Consulting, Moab, UT.

Turak, E. and K. Koop. 2003. Use of rare macroinvertebrate taxa and multiple-year data to detect low-level impacts in rivers. Aquatic ecosystem health and management. 167-175

Appendices

Metrics used by other states

Appendix 1. Bioassessment metrics used by Montana (MDEQ 2016)

- Ephemeroptera taxa
- Plecoptera taxa
- % EPT
- % Non-insect
- % Predator
- Burrower taxa %
- Hilsenhoff Biotic Index
- % EPT excluding Hydropsychidae and Baetidae % Chironomidae
- % Crustacea and Mollusca
- Shredder Taxa
- % Predator
- EPT taxa
- % Tanypodinae
- % Orthocladiinae of Chironomidae
- Predator taxa
- % Filterers and Collectors
- O/E

Appendix 2. Bioassessment metrics used by Wyomng (Hargett 2011)

Richness and Diversity Metrics
- % Chironomidae Taxa of Total Taxa
- % Diptera Taxa of Total TaxaX
- % Ephemeroptera Taxa of EPT Taxa
- % Ephemeroptera Taxa of Total Taxa
- No. Ephemeroptera Taxa
- No. EPT
- No. EPT Taxa (less Arctopsychidae and Hydropychidae)
- No. EPT Taxa (less Baetidae, Arctopsychidae, Hydropychidae and Tricorythodes)
- No. EPT Taxa (less Baetidae and Tricorythodes)
- Shannon Diversity (E)

Composition Metrics
- % Ephemeroptera (less Baetidae and Tricorythodes)
- % EPT (less Arctopsychidae and Hydropsychidae)
- % EPT (less Baetidae and Tricorythodes)
- % Tricorythodes of Ephemeroptera

Life History Metrics
- No. Semivoltine Taxa
- No. Univoltine Taxa
- Ratio of Multivoltine Taxa to Unvoltine Taxa +Semivoltine Taxa

Functional Feeding Group/Habitat Metrics

- % Clinger
- % Collector-gatherer
- % Filterer Taxa of Total Taxa
- % Scraper
- % Scraper Taxa of Total Taxa
- No. Burrower Taxa
- No. Predator Taxa
- No. Scraper Taxa

Tolerance Metrics

- BCICTQa
- HBI

Appendix 3. Bioassessment metrics used by Idaho(IDEQ 2011). In addition, IDEQ is developing and intermittent stream index.

- % Chironomidae
- % clingers
- % Ephemeroptera
- % Ephemeroptera and Plecoptera % filterers
- % EPT
- % EPT, excl. Hydropsychidae
- % filterers (adjusted)
- % Multivoltine
- % non-insects
- % Predators
- % Scrapers
- % Tolerant
- % tolerant (adjusted)
- Becks Biotic index
- Clinger taxa (adjusted)
- EPT Taxa
- EPT taxa (adjusted)
- HBI (adjusted)
- Insect Taxa
- Non-insect % of taxa
- Non-insect % of taxa (adjusted)
- Scraper taxa
- Semi-voltine taxa
- Simpson's index
- Sprawler taxa
- Sprawler taxa (adjusted)
- Swimmer & Climber Taxa

Tolerant taxa
O/E

Appendix 4. Some Bioassessment metrics used by Arizona(Jones and Woods 2010). In addition, ADEQ is developing and intermittent stream index.

Total taxa
Diptera taxa
HBI
% Stoneflies
% Scrapers
Scraper taxa
Caddisfly taxa
Mayfly taxa
%Mayflies
%Dominant taxa

Appendix 5. Bioassessment metrics used by New Mexico (NMED 2006)

Clinger Taxa
Coleoptera %
Ephemeroptera Taxa
EPT Taxa
Evenness
Intolerant Percent
Plecoptera %
Plecoptera Taxa
Scraper %
Scraper Taxa
Sensitive EPT %
Shannon DI (log2)
Sprawler Taxa
Swimmer Taxa
Taxonomic Richness
Trichoptera Taxa
O/E

Appendix 6. Bioassessment metrics used by Colorado(Jessup 2009)

Numerous including O/E. See Jessup (2009)

Appendix Literature Cited

Hargett, E. G. 2011. The Wyoming stream integrity index (WSII) multimetric indices for assessment of wadeable streams and large rivers in Wyoming. Wyoming Department of Environmental Quality. Cheyenne, WY.

Idaho Department of Environmental Quality. 2011. Biological assessment frameworks and index development for rivers and streams in Idaho. IDEQ. Boise, Idaho.

Jessup, B. Recalibration of the macroinvertebrate multi-metric index for Colorado. Colorado Department of Public Health and Environment. Water Quality Control Division. Denver, CO.

Jones, J. and J. Woods 2007 to 2010. A statewide assessment of Arizona's streams. Arizona Department of Environmental Quality.

MDEQ 2008. An assessment of the ecological conditions of the streams and rivers of Montana using environmental monitoring and assessment program (EMAP) method. Montana Department of Environmental Quality. Helena MT.

NMED. 2006. Benthic macroinvertebrate stream condition indices for New Mexico wadeable streams. New Mexico Environmental Department. Santa Fe, New Mexico.

Critique of Pilotto et al 1997, "Health effects of exposure to cyanobacteria (blue-green algae) during recreational water-related activities."

Technical Memo

September 7, 2016

To:
Jordan River/Farmington Bay Water Quality Council
Salt Lake City, UT

By
David C. Richards, Ph. D.
OreoHelix Consulting
Vineyard, UT
Phone: 406.580.7816
Email: oreohelix@icloud.com

Introduction

UDWQ 2016 draft IR cites several studies to support their conclusion that cyanobacteria exposure is harmful even at low doses (e.g. @5000 cells/mL). For example, one of the studies that UDWQ cited was Pilotto et al. 1997, "Health effects of exposure to cyanobacteria (blue-green algae) during recreational water-related activities." I reviewed the Picotto et al 1997 paper and have some major misgivings concerning their conclusions and which certainly need to be addressed in the forthcoming revision of the UDWQ 2016 IR.

Review

Symptoms vs. exposure (Table 1)

No analyses were conducted by Pilotto et al. 1997 to determine statistical significance between the different symptoms in their Table 1, page 564. They only reported percentages, hence no conclusions can be made from this table regarding exposure effects. Statistical analyses that included randomized sampling of the data could have easily provided error rates (e.g. CIs) and allowed for statistical examination of effects. One can unwisely postulate that exposure to cyanobacteria may actually reduce eye irritation without a statistical test given random error in any of the eight symptoms and the very uneven sample sizes of the unexposed group compared with the exposed groups before and after exclusion (Table1, Pilotto et al 1997). However, a simple Kruskal-Wallis rank test on the percent unexposed after exclusion vs the percent exposed after exclusion would have resulted in a chi-squared = 2.68 with 1 d.f. and a p = 0.10 (with ties). These results would suggest that there may not have been a very strong exposure effect signal and less confidence in concluding an effect.

As an example, ear irritation, as with all of the types of symptoms listed in Table 1 (Picotto et al 1997), can also be caused by other factors. Swimming in the ocean or a swimming pool often causes ear infections regardless of whether cyanobacteria are present or not. Obviously someone who didn't swim would likely have not gotten ear irritation as opposed to someone who went swimming anywhere prior. A more useful study would have included participants who had recreational water contact in non-cyanobacteria waters (e.g. swimming pools, ocean, etc.) In all likelihood a significant proportion of the population that went swimming in a public pool free of cyanobacteria or in the ocean would have exhibited some of these symptoms. Here is a quote from WebMD (www.webmd.com) titled, Beware of Recreational Water Illnesses:

> "Recreational water illnesses refers to any illness or infection caused by organisms that contaminate water in pools, lakes, hot tubs, and oceans, resulting in diarrhea, skin rashes, swimmer's ear, and other conditions. And they are on the rise. The rate has more than doubled in the past 10 years, according to data from the CDC. Infection-producing germs that can lurk in water include Pseudomonas aeruginosa, which causes swimmer's ear (an infection of the outer ear canal, known medically as otitis externa) and skin rash (dermatitis). Others include cryptosporidium, Giardia lamblia, shigella, and E. coli, which can cause diarrhea. Each year, 10,000 RWI cases of diarrhea and 6.2 million cases of swimmer's ear occur, according to the CDC. "You can catch respiratory illnesses and colds but by far, skin rashes, swimmer's ear, and gastrointestinal bugs are the most common," he says. Diarrhea may occur when contaminated water is swallowed and driven into the mouth or nose, Greene explains. It may not begin immediately after a swim; sometimes it comes on one to two weeks later".

Symptoms listed in Pilotto et al 1997, Table 1 need to be carefully reevaluated by UDWQ for usefulness and relevance to cyanobacteria blooms.

Odds Ratios[1]

There was an apparent error in the odds ratios reported for Model 1: exposure, after exclusion (Table 2 Picotto et al 1997). The upper 95% CI value was reported as 1.54 which was less than the mean value reported of 1.87 (Figure 1 in this critique). This is likely an editorial error not corrected by the editor(s) or by any of the twelve authors during any of the drafts or before accepting the final proof of the manuscript.

Predictor and response variability are so great that none of the 'treatment' odd ratios in any of the 4 models was significantly different than the control, 'unexposed' group or each other, except for one. All of the odds ratio's 95% CIs include values < 1.00, except for the Model 4; '> 60 minutes, >5000', 'after exclusion' treatment. Thus, there is no evidence that any of the categories in any of the models except one, differed in exposure effects including the control (Figures 1 -3 in this critique)(Table 2, Picotto et al. 1997).

Figure 1. Mean and 95% CIs of odds rations for Model 1: unexposed vs. exposed. There was No statistically significant differences between unexposed and expose (i.e. 95% CIs of exposed overlapped the unexposed odds ratio of 1.00. Also, there was an editorial error of the mean and upper 95% CI for exposed group that was not adjusted for in this figure. The odds ratio scale was modified to scale for values < 1.00.

[1] Odds ratio = (exposed with symptoms/not exposed with symptoms)/(exposed without symptoms/not exposed without symptoms)

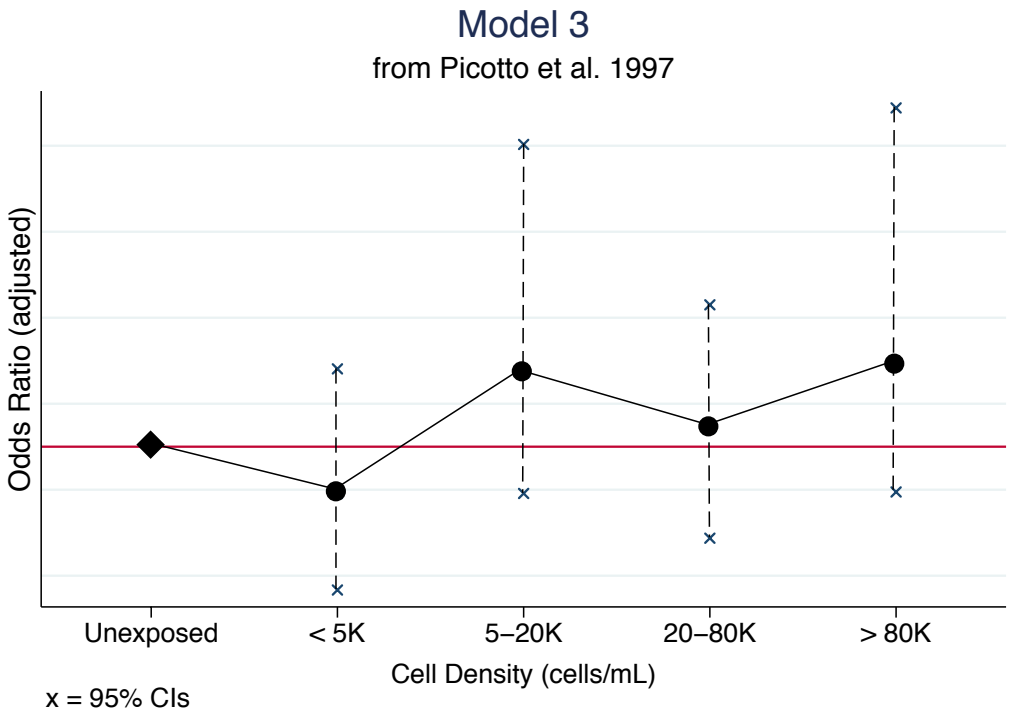


Figure 2. Mean and 95% CIs of odds rations for Model 3: unexposed vs. exposed to four different cyanobacteria cell count densities. There were no statistically significant differences between the unexposed group and any of the exposed treatments or between any of the exposed treatments (i.e. 95% CIs of all the exposed treatments overlapped the unexposed odds ratio of 1.00). The odds ratio scale was modified to scale for values < 1.00.

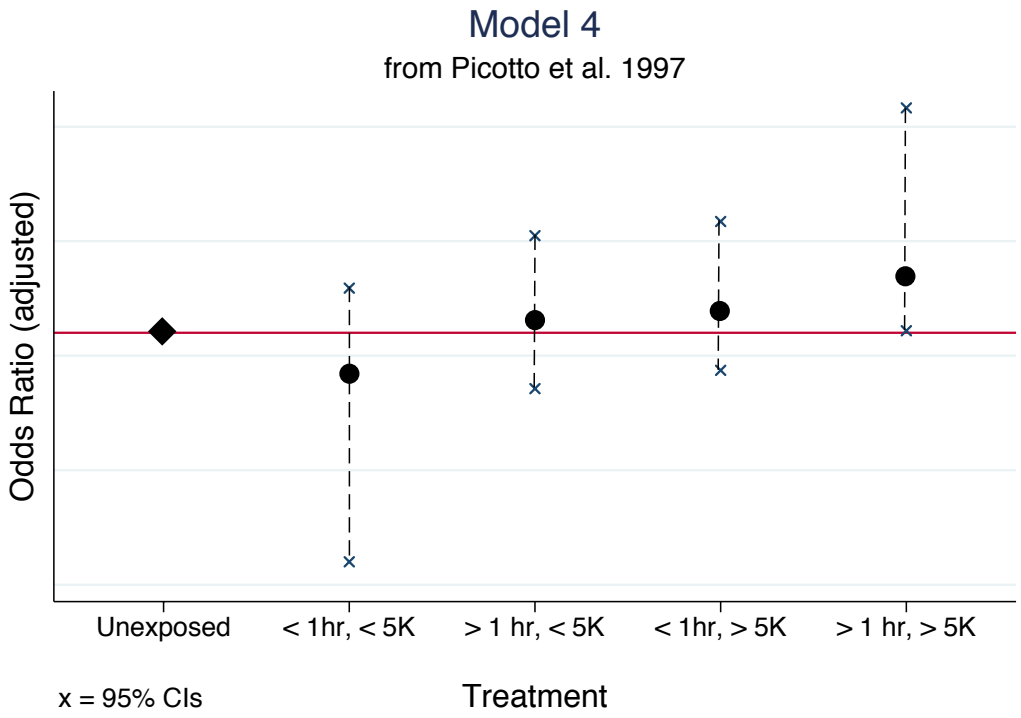


Figure 3. Mean and 95% CIs of odds rations for Model 4: unexposed vs. exposed to four different cyanobacteria cell count densities and durations. There were no statistically significant differences between the unexposed group and any of the exposed treatments or between any of the exposed treatments, (i.e. 95% CIs overlapped the unexposed odds ratio of 1.00), except for the treatment: > 1 hr., > 5K cells. No trend line was added to this figure because all treatments were categorical and not ordinal. The odds ratio scale was modified to scale for values < 1.00.

Trends

Even though the 'after exclusion' analysis for Models 2, 3, and 4 (Table 2, Picotto et al 1997) showed statistically significant trends ($P < 0.05$), their medical exposure relevance is questionable in Models 2 and 3 and possibly irrelevant in Model 4. For example, in linear regression models, the well-known and easily interpretable R^2 value indicates how well the model fits that data and how well the predictor affects the response variable. However, R^2 values can be very small and still be statistically significant. As an example, an R^2 value of 0.1 or even less can be statistically significant but have very poor predictability or of no use determining exposure/medical relevance. P- values associated with R^2 values are a measure of whether the linear prediction line (slope) differs than zero slope. Even a slight deviation from zero slope can result in a significant p-value. Error rates of odds ratio trend p-values can also be calculated via Monte Carlo simulation but were not done in the Picotto et al. 1997 report. It is safe to assume that some percentage of Monte Carlo type simulations of their data would result in statistically significant decreasing symptom trends with increased exposure in contrast to their results that showed increased risk. Additional analyses such as ANOVA type treatment effects models might also be appropriate and help determine the magnitude of the effects.

Model 4 (Table 2, Picotto et al 1997) reported trend is mostly irrelevant because the combination of duration and cell density response categories, as listed, have not been determined to be increasing exposure risk, particularly the two categories; '> 60 minutes, < 5000 cells' and '< 60 minutes, > 5000 cells'. There was no *a priori* knowledge of whether duration (< 60 minutes, > 60 minutes) was a greater risk than cell density (< 5000, > 5000 cells). Trend analyses of odds ratios are dependent on increasing risk categories (ordinal). The four categories used in Model 4 were not entirely ordinal categories; that is there was no clear ordering of two of the five variables. If the two categories in question are kept in the analysis, then the odds ratio analysis for Model 4 technically should have been based on categorical response values not ordinal. Odds ratio trend analysis is not a viable method for categorical values. This is somewhat of a minor criticism but the over reliance on trend p-values to make conclusions and ignoring the magnitudes or lack of significance within and between the treatments themselves is important and discussed in the paragraph above.

Also, no interaction effects between exposure duration and cell density were reported by Picotto et al. 1997. It is unclear from Table 2 if there were interaction effects.

If we follow Picotto et al 1997 lead based on odds ratio analysis presented in Table 2 (illustrated in Figures 2 and 3); we can conclude that cyanobacteria exposure for > 0 but < 1 hour and concentrations > 0 but < 5000 cells/mL are beneficial and can reduce illness, almost 2 X better than without exposure. From the Pilottoe et al. 1997 paper we could conclude that these low dose results are consistent with the idea that cyanobacteria supplements are a health benefit.

In addition, if we ignore confidence intervals, the greatest difference in adjacent odds ratios in Model 3 occurs between < 5K and 5-20K cells/mL (Figure 2 this critique). At < 5K cells/mL cyanobacteria seems to be beneficial and reduces symptoms but at 5-20K there is a > 2X negative affect of cyanobacteria. This phenomenon occurs in Model 4 also. As stated in the preceding paragraph, < 5K cell/mL for < 60 minutes improves health by almost 2X but at the same concentration has negative health effects after 60 minutes of exposure.

Conclusion

The experimental design and categorization of predictor and response variables resulted in such high variability as to make any conclusions as to the negative health effect of cyanobacteria exposure highly suspect and mostly just speculation, other than the Picotto et al. 1997 study showed no significant differences between non exposure and exposure treatments (categories). Improved experimental design may or may not have resulted in detection of significant cyanobacteria exposure effect as could have additional statistical analyses.

There is real cause for concern when management agencies rely on a single study or several studies to determine cyanobacteria exposure policy to develop assessment criteria based on a limited 'lines of evidence' narrative approach. A more appropriate approach would be to conduct a formal meta-analyses based on data obtained from a more exhaustive literature review and by critically evaluating those studies prior for use in meta-analysis or in criteria

development. The Pilotto et al. 1997 publication should be assigned a very low influence value if and when a formal meta-analysis is conducted