UDWQ IR 2022 Draft Response Letter

David C. Richards, Ph.D.
OreoHelix Ecological
Vineyard, UT 84059
406.580.7816

I would like to thank Utah Division of Water Quality (UDWQ) for providing citizens with the opportunity to comment on their 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. UDWQ should be commended for their efforts. However, I do 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 in particular, to address the pitfalls of over reliance on RIVPAC O/E models.

My background

I have been conducting ecological research on biological criteria related to water quality for several decades. My MS Thesis (Richards 1996) was titled, "The use of macroinvertebrates as indicators of water quality in mountain streams of Montana" My Ph.D. dissertation focused on population viability of a sensitive aquatic mollusk and its interactions with an invasive freshwater taxon (Richards 2004). I was employed by one of the leading macroinvertebrate taxonomy labs in the western USA, EcoAnalysts Inc. for approximately 13 years conducting many biological assessments throughout the western USA using and developing a multitude of bioassessment methods and metrics. I contributed extensively to the development of biocriteria programs for the States of Montana, Idaho, and Arizona using a multimetric approach. Along with my colleagues from EcoAnalysts Inc. and Idaho Department of Environmental Quality, I recently published a paper in the journal *Environmental Monitoring and Assessmen*t titled, "Temperature threshold models for benthic macroinvertebrates in Idaho wadeable streams and neighboring ecoregions" (Richards et al. 2018). I have been conducting ecological research on several waterbodies in Utah including the Jordan River, Provo River, Utah Lake, Great Salt Lake, and numerous others, much of which focus on bioassessment evaluations. I bring exceptional expertise in our efforts to develop useful and meaningful assessment criteria for Utah's waterbodies based on the best available science, so that we can continue to protect these valuable resources.

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.

A few specific comments upfront concerning RIVPACs models development.
What was the unit area used to develop models? There are three conflicting units in the draft:

> From Page 53. "Predictions of expected "E" taxa are obtained empirically from ***reference site collections made throughout Utah***. Reference sites represent the ***reference conditions in different biogeographical settings*** throughout the state. "
>
> From Page 52. "Predictions of E are obtained empirically from reference sites that together are assumed to encompass the range of ecological variability observed among ***streams in the region*** where the model was developed."
>
> From page 53. "An extensive list of 74 GIS-based ***watershed*** descriptors is evaluated for potential predictor variables in models that predict the probability of membership within biological groups for sites not used in model construction".
>
> From Page 55. "Utah currently assesses ***watersheds based on established AUs***."

DWQ states the use of different biogeographical settings, region, watershed, and AUs for model development. It can't possibly be all four. Which unit was it? These differences will have major consequences on how the model is applied to streams, particularly E, the expected number of taxa the sole denominator and half of the equation in the so called easy to interpret model.

E, expected number of taxa is irrelevant.
Choice of area is especially important for derivation of E, expected number of taxa. For example, whichever unit was used in DWQ models, the expected number of taxa, E would be the same within that unit. That is how the model is designed. E, is therefore a constant within any area unit be it region, watershed, AU, biogeographical setting, and dividing by that constant for all sample location benthic invertebrate results within that unit would be meaningless unnecessary confusing irrational and nonsensical. Division by a constant E would only result in O, the observed number of taxa in a sample, or as every aquatic ecologist is familiar with, the taxa richness metric.

Website links useless
I also went to the links provided on page 52. The WCMAFE website had no information on how Utah DWQ models were made. Hence, there is no available mechanism to understand the model components and assumptions. For example, I have no way of knowing what reference condition expected taxa were for the Jordan River. As I have stated many, many, times to DWQ there is no reference condition for the Jordan River and no valid E can be used to evaluate. Also, the link to the EPA website was invalid.

Here are more comments on DWQ's RIVPACs models

River Invertebrate Prediction and Classification System Models

Entire section.

Comment: There is no reason to justify using a single measure to describe highly complex biological integrity and report as one numeric index just to summarize into a single, easily interpretable number. Biological integrity/beneficial use is one of the main reasons DWQ conducts biological assessments, determines criteria, and sets regulations. UDWQ is mandated to protect beneficial uses, including aquatic life. To simplify biological integrity into one number just because it is easily interpretable (by who? DWQ trained biologists? Citizens of UT?) is a disservice to citizens of UT and is not the best protection criterion of our waterbodies. I do not know of any other state, federal, tribal, or county agency that relies solely on one biological assessment metric. Utah DWQ is the only one that does this, as far as I know.

This section "*River Invertebrate Prediction and Classification System Models"* in the draft appears to be written primarily to justify the use of RIVPACS models by UDWQ. The draft states that 'Recently, many western states have adopted the RIVPACS model… such as Colorado, Montana, and Wyoming. These States indeed use O/E models, but the O/E metric is just one of many in a multimetric assessment program (see Table 1). To claim that these states also use O/E models leads the public to believe that UDWQ's use of O/E as a stand-alone metric is valid, which it is not.

Table 1. Some metrics used by other states

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

Bioassessment metrics used by Wyoming (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

Bioassessment metrics used by Idaho (IDEQ 2011).
% 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

I don't agree that using a single taxon 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 (Richards 2016). 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. The data and algorithms used in these models are extremely difficult to obtain and often not available, thus not transparent. Other metrics used by other agencies, such as taxa richness, functional feeding group, etc. are very transparent and easily calculable.

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 (Table 1 for example and see Literature Cited). 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.

As I have emphasized to UDWQ on numerous occasions, **RIVPACS O/E models do not quantify loss of predicted taxa**. In the case of UDWQ assessments, O/E quantifies only those taxa that were identified from a single (N = 1) composite sample collected from several types of habitats (including riffles and runs) that can exhibit much variability between the macroinvertebrate assemblages. Samples were also identified in the laboratory using a subsample (typically 600 organisms, with large and rare counts). O/E simply quantifies what was observed in a sample, nothing more. Taxa not identified may have or may not have been lost from the waterbody UDWQ can only conclude that they simply weren't observed.

Probability of Capture > 50%
Again, as I have discussed on numerous occasions, probability of captures (Pc's) >50% preclude those very macroinvertebrate taxa that constitute biological integrity in a water body. As an example, waters in the Bonneville Basin and in some other parts of UT have unique mollusk assemblages found nowhere else in the world. Most of Utah's mollusks, including native mussels, clams, and non pulmonate snails do not occur in UT waters at Pc rates > 50%. By relying on RIVPACS O/E > 50% Pc, UDWQ failed to protect the unique mollusk assemblages in UT and apparently was completely unaware of their declines during the time period when continued molluscan viability may have been protected/ensured. This reliance on a single metric with > 50% Pc to assess biological integrity also likely is not protecting other rare and uncommon macroinvertebrates (< 50% Pc) that are again, by definition biological integrity.

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'.

From the lengthy discussion in the draft, it appears that UDWQ is more interested in the continued reliance on a single metric (O/E) that had good statistical properties (e.g., more sensitive and precise) than incorporating other metrics or using a < 50% Pc that may prevent loss of rare, uncommon, and unique taxa and provide greater insights into the types of impairments that Utah waterbodies experience.

It is my opinion that O/E models may be able to detect large levels of biological stress, but not biological integrity.

RIVPACS O/E precision and predictive ability
The new O/E model in the draft is claimed to be a less precise predictive model than the previous used by UDWQ. A loss of precision in the updated model should be critically reevaluated. Was this updated model selected because in saves time and money?

Several problems in simplifying the model are as follows:

Incorporation of 1rst order and 8th plus order streams and rivers.
All aquatic ecologists know that there is a big difference in macroinvertebrate assemblages in typical 1rst order vs. 2nd to 5th streams and between 8th plus rivers and 2nd to 5th order stream (please review the River Continuum Concept by Vannote et al.).

Taxonomic resolution.
A coarser taxonomic resolution results in a major loss of valuable information provided by individual taxa when 'rolled up' to higher taxonomic level. It also means that some unique or ecologically valuable taxa may be unaccounted for and lost from the AU without knowledge by UDWQ. For example: combining all species of caddisflies in the genus *Rhyacophila* at least 5 species or more could be lost without UDWQs knowledge. Or by combining all species of the mayfly genus *Baetis*, several of the more sensitive species may have been lost. UDWQ is well aware that taxonomic (phylogenetic) similarity has very little predictive power for sensitivity to different types of impairment (Richards 2016, UDWQ 2017).

Seasonality effects
Seasonality also affects macroinvertebrate assemblages. Summer season has fewer taxa in larval stages that are needed for taxonomic identification and O/E derivation. Comparing summer collected vs. late autumn to early spring samples increases variability and thus O/E results (e.g., summer samples likely will have fewer taxa and lower O).

Because of these pitfalls, I caution UDWQ not to try to accommodate broader spatial and temporal data into O/E models simply to cut costs. This will result in loss of predictive power in ability to detect impairment. Remember that all assessments and monitoring efforts will eventually have to be measured at the watershed or site-specific level and a macroinvertebrate assessment program that reduces variability at the onset will be more cost effective in the long run.

UDWQ is in an ideal situation to vastly improve macroinvertebrate biological assessments. UDWQ has a strong working relationship with the USU Bug lab including the leading developers of RIVPACS models at USU and other entities. They should take

full advantage of this opportunity to develop a robust biological assessment program comparable to other federal, state, tribal, and county agencies in the region.

It appears to me that many millions of dollars have been spent developing RIVPACs O/E regional models when it would have been much more prudent to train UDWQ staff to recognize the macroinvertebrate taxa that occur in UT and become proficient in understanding their ecology, natural and life history, examine sample results and easily evaluate which taxa were missing and why at the watershed level.

Model Construction and Performance

Page 49. Table 12.

Comment: These predictor models and variables are mostly watershed based. It is highly commendable that UDWQ is now 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.

As discussed in earlier comment letters; 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. PRISM models rely on historic data (e.g., most of the climate data metrics in Table 12). As an example, "Watershed maximum of mean 1961-1990 annual number of wet days' was 28-year-old past data. Conditions likely have changed substantially in 28 years. 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 12 had nothing to do with environmental conditions during the time when the sampled invertebrates lived. This introduces an unmeasurable and significant error to every Pc 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 as such. Only field measurements should be used when possible.

PRISM data errors are also spatially derived mostly from misuse 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. PRISM data often are not precise, and values can change substantially between small changes in elevation within a watershed and sometimes within a few hundred meters. In addition, PRISM values are model predicted values and subject to error.

Assessments Specific to Lakes, Reservoirs, and Ponds

Starting on page 53

Methods are lacking in the draft to evaluate biological integrity/aquatic beneficial uses. There are no zooplankton, benthic macroinvertebrate, or fish numeric or narrative metrics. Without such metrics, there likely is no possibility of evaluating whether biological beneficial uses are supported or not supported. A program needs to be started by UDWQ to develop robust multimetric biological assessments for lentic waters.

In many instances UDWQ refers to cold-water vs. warm water uses. Temperatures that exceed 20° C do not necessarily mean impaired. It is possible that the water body is naturally a warm water fishery and may have been misclassified or that increased temperatures due to climate change have affected temperatures. This is a problem with stream assessments as well (e.g., Jordan River). There is also no reason for UDWQ to infer that a cold-water fishery is superior to a warm water fishery by stating that cold water uses are a 'higher' use than warm water use. For example, UDWQ states their goal is to meet the highest attainable use. We need to get away from the idea that cold-water mountain streams and lakes have some greater innate value than lower elevation warm-water bodies. Global climate change may insure this, eventually.

My overall conclusion is that the UDWQ 2018 Draft reflects a concerted effort by UDWQ to manage Utah's waters that are protective of biological integrity (and other uses) and is to be commended. However, the draft is heavy on numeric -criteria –based- measures such as DO and weak on how these metrics actually relate to biological integrity, the real measure of water quality as mandated by the Clean Water Act. Finally, there seems to be no clear scientific or otherwise causal link between the numeric based metrics and the 'beneficial uses' particularly biological, that UDWQ is evaluating.

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.

For this comment letter, I include statistical analyses I conducted several years ago.

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.

A RIVPACs Evaluation Example

Table of Contents

Introduction 14
Justification 16
Methods 16
  Dataset 16
  Statistical Analyses 16
Results 17
  Metric Correlations 17
  Summary Statistics 17
  OLS regression 19
  Structural Equation Models 21
  Effects of evenness on taxa richness: hypothetical example 23
  Jordan River O/E Bioassessment Example 26
Discussion 30
  Implications of Evenness on O/E Scores and UDWQ bioassessments 31
  RIVPACS O/E 'Probability of Capture' is Problematic 31
  Misinterpretation of O/E 32
  Additional Bioassessment Metrics in Use 32
  Economics vs. Bioassessment Quality. 35
Conclusion 35
Literature Cited 36
Appendices 40

List of Figures

Figure 1. Relation of O/E scores to taxa richness. Red line is OLS regression linear fit 20
Figure 2. Relation of O/E scores to evenness. Red line is OLS regression linear fit 20
Figure 3. Structural Equation Model (SEM) and path diagram of the direct and indirect effects of richness and evenness metrics on O/E scores. 21

List of Tables

Table 1. Pearson correlations between richness and diversity metrics 17
Table 2. Summary statistics for taxa richness metric. 18
Table 3. Summary statistics for evenness metric. 18
Table 4. Summary statistics for O/E score metric. 18
Table 5. OLS regression results of O/E score as a function of taxa richness and evenness. 19
Table 6. Standardized SEM results of the effects of richness and evenness on O/E scores. 21
Table 7. Standardized SEM direct effects of richness and evenness on O/E scores. 22
Table 8. Standardized SEM indirect effects of richness and evenness on O/E scores. 22
Table 9.Standardized SEM total effects of richness and evenness on O/E scores. 23
Table 10. Hypothetical invertebrate sample with 30,000 individuals and equal proportional abundances. 24
Table 11. Hypothetical invertebrate sample with 30,000 individuals and unequal proportional abundances. 25
Table 12. List of taxa found in Jordan River using BLM/USU Buglab database. 26
Table 13. Jordan River evenness score, Dominant Family, % Dominant Family, Dominant Taxon, and % Dominant Taxon from 32 BLM/USU Buglab samples. 28
Table 14.Four Jordan River sites with O/E scores and evenness values that were compatible between the two datasets. 29
Table 15. Taxa observed in the Jordan River sample 142114.. 29
Table 16. Condensed list of metrics that are routinely generated by BLM/ USU Buglab for UDWQ's bioassessment program. 34
Table 17. Estimated gross state product and population for Utah and surrounding states (2010 data). 35

List of Appendices

Appendix 1. Bioassessment metrics used by Montana 40
Appendix 2. Bioassessment metrics used by Wyomng 40
Appendix 3. Bioassessment metrics used by Idaho 41
Appendix 4. Bioassessment metrics used by Arizona 42
Appendix 5. Bioassessment metrics used by New Mexico 42
Appendix 6. Bioassessment metrics used by Colorado 42

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      Number of obs   =       262
-------------+----------------------------------   F(2, 259)       =    179.95
       Model |  8.28553352         2  4.14276676   Prob > F        =    0.0000
    Residual |  5.96265445       259  .023021832   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
------------------------------------------------------------------------------
```

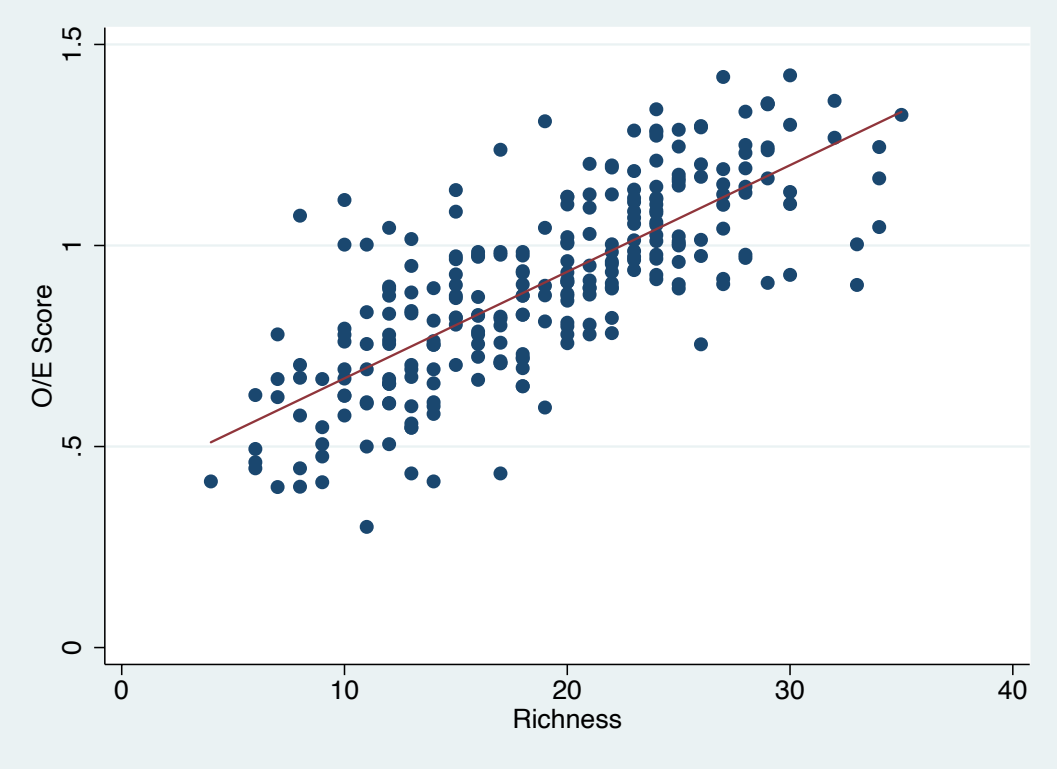


Figure 1. Relation of O/E scores to taxa richness. Red line is OLS regression linear fit. See Table 5 for OLS results.

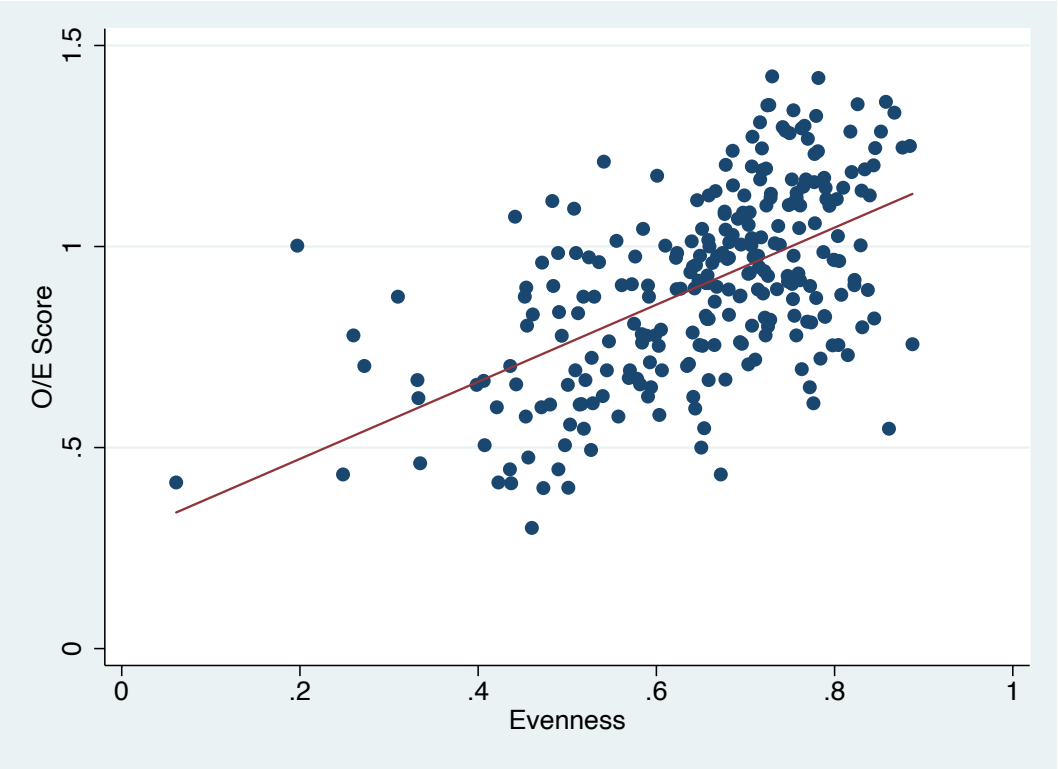
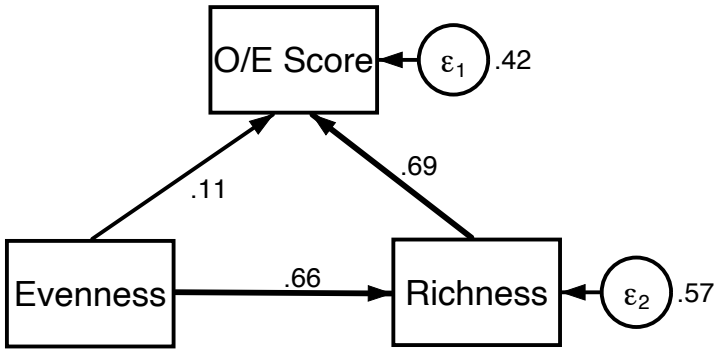


Figure 2. Relation of O/E scores to evenness. Red line is OLS regression linear fit. See Table 5 for OLS results

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

	Coef.	OIM 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.

A few Recommendations and Suggestions

1. UDWQ needs to provide user-friendly public access to RIVPAC O/E and PRISM models. Transparency (repeatability) is a key component of scientifically validity.

2. Macroinvertebrates are the corner stone of biological integrity. As such, UDWQ should put much more effort into developing useful macroinvertebrate metrics in a multimetric assessment program that could include an O/E metric.

3. There is a need to include references or links in the draft to UDWQ field macroinvertebrate sampling protocols or add one or two sentences in the draft that include methods used such as riffle/run habitats, 8 composite samples, 600 organism subsample including large and rare, taxonomic resolution used, etc.

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

Sincerely,

David C. Richards, Ph. D.
OreoHelix Consulting
Vineyard, UT 84059

Literature Cited and Suggested Readings

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.

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.

Montana Department of Environmental Quality. 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.

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

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. 2016. Real and Perceived Macroinvertebrate Assemblage Variability in the Jordan River, Utah can Affect 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. 2016. Does Phylogeny Predict Sensitivity to Ammonia in Freshwater Animals using USEPA Ammonia Criteria Data? Submitted to the Wasatch Front Water Quality Council. Salt Lake City, UT. Oreohelix Consulting, Vineyard, UT.

Richards, D. C. 2016. Is Reliance on a Single Bioassessment Metric for Assessing Water Quality in Utah's Rivers and Streams Prudent? Draft Technical Report to Wasatch Front Water Quality Council. Salt Lake City, UT. Oreohelix Consulting, Vineyard, UT.

Stout III, Ben M. "River Continuum Concept as an Analytical Template for Assessing Watershed Health" Wheeling Jesuit University. 2003.

Thorp J.H., Delong M.D.: “The Riverine Productivity Model: An Heuristic View of Carbon Sources and organic processing in large river ecosystems”. In: Oikos 70 (2) :305-308. Blackwell, Oxford 70 .1994.

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.

UDWQ et al. 2017. Utah and Colorado Water Survey for Mussels and Snails. Final Report. Original Draft-July 1, 2017. Revised Draft-.

Vannote R.L., G. W. Minshall, K. W. Cummins,Can. J. “River Continuum Concept” Fish. Aquatic Science. March 2005.

Vannote R.L., G.W. MINSHALL, K.W. Cummins, J.R. Sedell, C.E. Cushing: “The River Continuum Concept”. Canadian Journal of Fisheries and Aquatic Sciences. 37.1980,1 Ottawa, 130-137.

Ward J.V., J.A. Stanford: The Serial Discontinuity Concept of River Ecosystems. T.D. Fontaine, S.M. Bartell: “Dynamics of Lotic Ecosystems”. Science Publications, Ann Arbor Mich 29-42. 1983.

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.