Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
A	1	303(d)	The North Fork of the Virgin River as assessed at four sites in close proximity upstream of the park (assessment unit ID UT15010008-013) is not supporting the protected uses for E. coli. Our data supports this conclusion and we also support the continued efforts on the part of UDWQ to implement a project to correct this problem.	DWQ thanks you for the comment.
A	2	303(d)	One of the sites (4951265) is also listed for aluminum. We believe that this is in error caused by the failure to link field and laboratory data for the sample date 7/24/2008. The National Park Service collects field parameters and water quality samples at NPS sampling locations, and then submits the water quality samples to the state of Utah for laboratory analysis. Field data are submitted separately, and in some cases, laboratory samples have not been correctly paired with field data. When the field pH of 8.38 was not entered into the database for the sample on that date, the assessment screening incorrectly applied the chronic standard for aluminum when the acute standard should have been applied for waters with a pH greater than 7.0. With the correct pH and standard applied there is no exceedence. The applicable datasheet with a field pH of 8.38 on 7/24/2008 is attached.	The text of the 2012-2014 Integrated Report (IR) Assessment Methodology has been changed in response to comments regarding the frequency of exceedances for acute and chronic criteria. Please refer to Chapter 2, page 23, to see the changes that were made for the number of exceedances required to list a river/stream for Acute and Chronic (with sample sizes of >=4 and <=10) impairments. As a result of this change, the North Fork of the Virgin River is no longer listed as impaired for aluminum.
A	3	303(d)	We were unaware of a temperature problem at this site 4951265 prior to this listing, and look forward to working with the UDWQ staff to evaluate it. Runoff from the irrigated fields upstream is probably a contributing factor, so there may be an improvement concurrent with improving the irrigation system on those fields.	DWQ appreciates your comment and looks forward to working with the NPS on improving conditions on the North Fork of the Virgin River.
A	4	303(d)	The North Fork of the Virgin is also listed for E. coli in the reach downstream of Deep Creek (assessment unit UT15010008-015) based on data from site number 4951199 at the Temple of Sinawava. This may be in error depending on how the analysis was conducted. Our data for this site from 2009 through 2013 shows a geometric mean of 35.8 MPN/100 ml for 57 samples. This is well below the 2A chronic standard. We found that 12 of the 57 individual samples, or 12.3% exceeded the chronic standard. E. coli levels at this site are probably elevated by the downstream persistence of contamination from the pastures upstream of the park, and from water play engaged in by visitors in the more immediate vicinity.	DWQ compiled all publically submitted, cooperator, and DWQ-collected *E. coli* data from 1/1/2003 to 12/31/2012. For the Assessment Unit (AU) UT15010008-015 and Monitoring Location ID (MLID) 4951199, there were 117 samples or data records and 45 data points after quality assurance and quality control checks were performed and the geometric mean of replicates were calculated. Though more data records for the AU and MLID exist, these samples were collected outside of the assessment period outlined in Chapter 2, page 41, of DWQ's 2012-2014 IR Assessment Methodology. After public comment closed, DWQ reassessed and confirmed the *E. coli* listing decision for 4951199 in Chapter 5 of the Draft 2012-2014 IR. As stated on page 43 of Chapter 2, DWQ lists a MLID as Not Supporting for *E. coli* when any MLID has a violation of Rule 1 or 2 during the five most recent recreation seasons. When assessments at a MLID contradict one another (e.g., some years are supporting, others not fully supporting), the presence of a Not Supporting assessment outweighs the Fully Supporting, and the site is considered Not Supporting. In 2011, five rolling geometric means exceeded the Rule 1 30-day geometric mean of 126 MPN/mL for 2A waters, one geometric mean exceeded the Rule 1 30-day geometric mean of 206MPN/mL, and three samples exceeded the Rule 2 maximum of 409 MPN/100mL for 2A waters. The presence of these Not Supporting assessments outweighed the Fully Supporting and insufficient data assessment from 2009, 2010, and 2012. To view the data assessed in the Final 2012-2014 IR, please refer to the data files on the following website (http://www.waterquality.utah.gov/WQAssess/rsdatafiles.htm). The *E. coli* assessment methodology can be accessed on the same website in Chapter 2, pages 40–44 of the Final 2012-2014 IR.
A	5	303(d)	The elevated pH on North Creek (assessment unit UT15010008-014) has been noticeable in our data following the Kolob Fire that occurred in 2007. We expected some recovery as the nutrient flush from the fire progressed through the system and algal growth subsided. Our field observations are that while the algal growth has indeed diminished, we still see some summer samples with a pH exceeding the 9.0 threshold. Our data probably captures the daily peak in pH because samples are usually collected around mid-day when photosynthesis is at its peak.	DWQ thanks you for the comment.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
B	1	303(d)	SF of Sixmile is listed as insufficient data (geomean 292 CFU), yet it was a proper 5 sample geomean	At the Monitoring Location ID 9, Six Mile Creek SF, *E. coli* was assessed and reported as Fully Supporting in the parameter summary reports located on the website (http://www.waterquality.utah.gov/WQAssess/rsdatafiles.htm). During DWQ's public solicitation for data to be considered for the 2012-2014 IR, DWQ also received pH, water temperature, and dissolved oxygen (DO) data for Six Mile Creek SF. For all three parameters, sample counts were insufficient (e.g., <4 samples), and the parameters received a 3A, insufficient data category. (Please refer to the parameter summary reports located on the website.) Where assessments among uses and across parameters differ or contradict one another, a Not Supporting outweighs the other assessments. When a Not Supporting is not present, a 3A, insufficient data category outweighs a No Evidence of Impairment or Fully Supporting assessment category. DWQ will continue to work on improving this level of information to stakeholders and reviewers in future reports to avoid confusion.
B	2	303(d)	Many sites had astronomical turbidity levels Both Sage Cr's and SF Sixmile) but no listings. What is needed for turb listings?	Utah's turbidity standard is a maximum increase of 10 NTUs for recreational uses. Turbidity could also be assessed under the Narrative Standards (R317-2-7.2). However, Utah currently does not have assessment methods for turbidity. Turbidity is anticipated to remain unassessed until an assessment method is developed.
B	3	303(d)	Otter Creek 195 CFU geomean) (and all the creeks data was submitted for) are listed as 2B secondary contact, but many support primary and are in high recreation areas. How does the 2B square with the CWA rebuttable presumption principle? Otter Creek and Big Creek are both major rec areas. I see families in campers out there from May through Oct. I have seen kids playing in both. This may also be a broader issue that has never been dealt with since you are just getting into e coli issue. You are in a different EPA region but Reg 8 disapproved WY's categorical downgrade to secondary	The IR assesses water quality in the context of the uses designated in UAC R317-2, Standards of Quality for Waters of the State. To assess waters for a different designated use would require a change to the water quality standards. Potential changes to the standards are considered during the Triennial Review, and the next Triennial Review is scheduled for 2017.
C	1	303(d)	One of the most problematic issues associated with the IR is the use of one acute or chronic criteria exceedance to place a site on the 303(d) list. This lack of confirmation data is inconsistent with prior listing determinations. Specifically, when a monitoring location has less than 1 0 samples an acute or chronic violation may be just an outlier and not representative of any water quality problem. Below is a graph of monitoring site 5994790- Thompson Creek. As can be seen, there are five data points, with four below the aluminum chronic criteria and one over 30 times greater than the next highest value. This outlier may be a true value, but likely is a sampling or laboratory error. It is our opinion that with a small data set (less than ten samples) one chronic or acute exceedance should not be a cause for inclusion in the 303(d) list but may rather be included in one of the Category 3 brackets.	The text of the 2012-2014 IR Assessment Methodology has been changed in response to this concern. Please refer to Chapter 2, page 24, to see the changes that were made for the number of exceedances required to list a river/stream for Acute and Chronic (with sample sizes of >=4 and <=10) impairments. EPA's assessment method guidance is clear on the relevance of extreme or "outlier" vlues and their place in assessment of toxic parameters: (from Part 4 of EPA's 2006 Integrated Report Guidance (http://water.epa.gov/lawsregs/lawsguidance/cwa/tmdl/upload/2005_08_11_tmdl_2006IRG_report_2006irg-sec4.pdf) "Extreme values or "outliers" can be very relevant when dealing with Water Quality Criteria (WQC) aimed at protecting humans or other life forms against adverse effects of acute (short term) exposure to pollutants. The fact that such values may occur fairly infrequently and are not representative of long term average conditions is unimportant when dealing with WQC expressed as short-term that should occur only rarely, if ever. EPA's WQC addressing acute exposure of freshwater aquatic life to toxic chemicals are an example of WQC expressed in this way – they are one-hour average concentrations that should be surpassed no more than once every three years on average. WQC expressed as instantaneous concentrations never to be surpassed address even more rare, but nevertheless harmful, conditions."
C	2	303(d)	When a larger data set is present with greater than 10 total samples the use of one acute violation may still be simply an outlier and an error in data collection or sample analysis. The sample could be a real value, but the infrequency the acute violation occurs would make it practically impossible to determine a source for the high value. As can be seen from the following graph, there is only one exceedance in about 37 samples. This high number of samples with just one sample greater than 75 times the average of all other samples lends suspect to the one exceedance. The above data is from monitoring location 4992480 - Mill Creek above the confluence with the Jordan River at the USGS Gage Station. Again, it is our opinion that such a single exceedance should not be a cause for listing on the 303(d) list.	The text of the 2012-2014 IR Assessment Methodology has been changed in response to this concern. Please refer to Chapter 2, page 24, to see the changes that were made for the number of exceedances required to list a river/stream for Acute and Chronic (with sample sizes of >=4 and <=10) impairments.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
C	3	303(d)	Confirming this opinion that a single exceedance is insufficient for listing, below is a graph from Monitoring Location ID 4990987 - North Canyon Creek in Legacy Nature Preserve. This data set has one significant exceedance that is inconsistent and much higher than all the other samples in the data set. This again confirms that 303(d) listing is inappropriate with outliers that are likely inaccurate. While we would rather see all such monitoring locations indicated as Category 1 waters as fully supporting for this pollutant at this location, we understand that it may be necessary to include them in Category 3 as lacking sufficient data.	The text of the 2012-2014 IR Assessment Methodology has been changed in response to this concern. Please refer to Chapter 2, page 24, to see the changes that were made for the number of exceedances required to list a river/stream for Acute and Chronic (with sample sizes of >=4 and <=10) impairments.
C	4	303(d)	Finally, the entire assessment of iron exhibits a significant level of problems relating to the accuracy of all iron analysis. Below is a table extracted from the provided data set. As can be seen, all of the above samples were collected on the same date. In addition, all of the samples were significantly higher than all other samples at their individual monitoring site. The samples above were the only samples above the criteria at each location. In this case either laboratory error occurred, the samples were reported in the wrong units, there was a continuous field sampling bias or there was a significant release of iron in each creek on the vary same day. Besides the Columbus-Rexall Mine outfall, all of the other iron acute violations should probably be included in one of the Category 3 brackets. This conclusion is supported by the information from monitoring location 10 5994410- Johnson Canyon. Below is a graph of the data from this location. While two samples exceeded the criteria, the age of the two exceedances is such that to place it on the 303(d) list is inappropriate in our opinion. Again, this sample belongs in category 3 with insufficient data to make a determination.	The change to assessment methodology (from 1 to 2 exceedances for toxic parameters) has resulted in these sites meeting their uses for iron. As discussed elsewhere in DWQ's resposnes to similar comments, DWQ agrees with the comment that a single high exceedance for metals criteria should not be cause for impairment and that a more appropriate recurrence interval of 2 or more exceedances be applied to the assessment methods.
C	5	303(d)	In conclusion, we recommend two amendments to the IR relating to single exceedances or iron analysis: 1. At least two exceedances must occur before being placed on the 303(d) list. Single exceedances may be placed in Category 3 indicating insufficient data. 2. All iron listings, except the Columbus- Rexall Outfall, should be placed in Category 3 until an outlier analysis is done and the vast difference between the typical value and the outlier is explained or validated with additional sampling.	The text of the 2012-2014 IR Assessment Methodology has been changed in response to this concern. Please refer to Chapter 2, page 24, to see the changes that were made for the number of exceedances required to list a river/stream for Acute and Chronic (with sample sizes of >=4 and <=10) impairments.
C	6	303(d)	From the spreadsheets with the data for the integrated report it appears that calcium and magnesium concentrations were used to calculate an applicable hardness for each monitoring event at each location. While we accept that calcium and magnesium hardness (Ca/Mg hardness) normally is a significant portion of the total hardness, other cations such as iron or manganese may also add to the total hardness of water. For pollutants where the toxic concentration is hardness dependent the calculated hardness may be adequate to approximate the hardness for purposes of IR listing. This assumption is inappropriate for aluminum when assessing for chronic violations. As can be seen from the USGS hardness map below, water in Utah is generally considered hard.	DWQ applied hardness corrections where dictated by standards using calcium and magnesium hardness coupled with the respective toxic parameter results. For aluminum, the default criterion was changed to 750 mg/l in cases where we lack hardness values as per responses below.
C	7	303(d)	The Utah Administrative Code in R317-2-14 has a footnote for the chronic criteria for aluminum, which states that "Where the pH is equal to or greater than 7.0 and the hardness is equal to or greater than 50 ppm as CaC03 in the receiving water after mixing, the 87 ug/1 chronic criterion (expressed as total recoverable) will not apply, and aluminum will be regulated based on compliance with the 750 ug/1 acute aluminum criterion (expressed as total recoverable)." It is our contention that the use of a calculated hardness or the lack of hardness data to determine applicability of the 87 ~g/L is inappropriate. We suggest that if actual hardness measurements do not exist demonstrating a true hardness of over 50 ppm the water be included in Category 3 awaiting confirmation of the accurate hardness measurement. This is especially true when the aluminum concentrations just exceed the criteria. Following is a graph of aluminum concentration at monitoring location ID 4990340 - Farmington Creek at USGS Gage Station. Calculated hardness values at this location range from 27 ppm to 461 ppm. Because of the sensitivity to hardness for applicability of the standard, we suggest that before inclusion on the 303(d) list this site be sampled and actual hardness measurements taken. As a comparison, below is a graph of hardness measured at the intake of the Little Cottonwood Treatment plant owned by the Metropolitan Water District of Salt Lake and Sandy located on Little Cottonwood Creek. This would be a snow fed mountain creek similar to Farmington Creek. As can be seen, measured total hardness rarely falls below 50 ppm.	DWQ agrees with the comment, and, in response to this concern, the assessment process was altered to more appropriately assess aluminum in situations where a pH or hardness value was not paired with the aluminum results. Footnote 6 in R317.2 states "The criterion for aluminum will be implemented as follows: Where the pH is equal to or greater than 7.0 and the hardness is equal to or greater than 50 ppm as CaC03 in the receiving water after mixing, the 87 ug/1 chronic criterion (expressed as total recoverable) will not apply, and aluminum will be regulated based on compliance with the 750 ug/1 acute aluminum criterion (expressed as total recoverable)." In the draft IR, the 87 mg/l criterion was adopted in cases where there was no paired hardness and/or pH data to make this determination. We recognize the need to clearly document the instances when paired aluminum, pH, and hardness values require the application of the 87 mg/l. The revised draft therefore defaults to the 750 mg/l aluminum in cases where there is no pH or hardness data to suggest using the 87 mg/l criterion.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
C	8	303(d)	EPA recognized the infrequent times when the 87 !Jg/L standard should be applied in the footnote to the criteria in the National Recommended Water Quality Criteria (http://water.epa.gov/scitech/swguidance/standards/criteria/current/index.cfm). This footnote states: FootnoteS There are three major reasons why the use of Water-Effect Ratios might be appropriate. 1. The value of 87 j.Jg/1 is based on a toxicity test with the striped bass in water with pH = 6.5-6.6 and hardness <1 0 mg/L. Data in "Aluminum Water-Effect Ratio for the 3M Plant Effluent Discharge, Middleway, West Virginia" (May 1994) indicate that aluminum is substantially less toxic at higher pH and hardness, but the effects of pH and hardness are not well quantified at this time. 2. In tests with the brook trout at low pH and hardness, effects increased with increasing concentrations of total aluminum even though the concentration of dissolved aluminum was constant, indicating that total recoverable is a more appropriate measurement than dissolved, at least when particulate aluminum is primarily aluminum hydroxide particles. In surface waters, however, the total recoverable procedure might measure aluminum associated with clay particles, which might be less toxic than aluminum associated with aluminum hydroxide. 3. EPA is aware of field data indicating that many high quality waters in the U.S. contain more than 87 !Jg aluminum/L, when either total recoverable or dissolved is measured. Based on the lack of actual hardness measurements, the sharp cutoff of the criteria at 50 ppm hardness for applicability, and EPA's recommendation on the appropriate use of the water effects ratio for aluminum, we again state that all monitoring locations where hardness was calculated or not available or where hardness values were close to the 87 mg/1 concentration be included in Category 3 waters rather than being placed on the 303(d) list at the present time.	DWQ agrees with the comment, and, in response to this concern, the assessment process was altered to more appropriately assess aluminum in situations where a pH or hardness value was not paired with the aluminum results. Footnote 6 in R317.2 states "The criterion for aluminum will be implemented as follows: Where the pH is equal to or greater than 7.0 and the hardness is equal to or greater than 50 ppm as CaC03 in the receiving water after mixing, the 87 ug/1 chronic criterion (expressed as total recoverable) will not apply, and aluminum will be regulated based on compliance with the 750 ug/1 acute aluminum criterion (expressed as total recoverable)." In the draft IR, the 87 mg/l criterion was adopted in cases where there was no paired hardness and/or pH data to make this determination. We recognize the need to clearly document the instances when paired aluminum, pH, and hardness values require the application of the 87 mg/l. The revised draft therefore defaults to the 750 mg/l aluminum in cases where there is no pH or hardness data to suggest using the 87 mg/l criterion.
C	9	303(d)	It is understood that streams and rivers are affected by multiple factors when a pollutant in-stream concentration is compared to a water quality criteria. We are aware that headwater areas may be impacted by mine drainage, that lower areas are impacted from storm water drainage and point source discharges and that natural occurring pollutants may enter a stream or river at any location. From the standpoint of 303(d) listing what concerns us are locations where a small segment of a river or stream is identified with an impairment when upstream and downstream locations do not show the same increased concentrations. Below is a graph of monitoring location ID 4994170- Jordan River at 7800 South. As can be seen, there are about 7 occurrences where the concentration exceeds the chronic criteria by a relatively small amount. Normally this frequency of exceedances would be justification for concern and probably a rationale for listing. However in this case, the upstream and downstream concentrations do not support this conclusion. Below is a graph of the upstream concentration at monitoring location ID 4994600- Jordan River at Bluffdale Road. While some of the measured selenium concentrations come close to the water quality chronic criteria, none of them exceed the 4.6 ~m/L standard. Thus, for an increase to occur in the stream or river in this section only, some additional source would have to be located between Bluffdale Road (and goth South for that matter although this site had much fewer measurements) and 7800 South. By monitoring location ID 4994090 - Jordan River at 5400 South the in stream concentration is again well below the chronic criteria concentration as shown on the graph following. Although not as robust a data set, there again are no exceedances demonstrated in the river. These three locations are not that far apart and the velocity of the stream is such that any fish, at least, would be through the area of concern well within the four day averaging period for a chronic exceedance. We accept the values at 7800 South as being accurate, we question if, given the lack of exceedances up and down stream this segment should be listed as impaired for selenium. Our opinion is that, at least for selenium, this reach of river should be included in Category 3 until additional monitoring demonstrates a source of the increase or continued exceedances show harm may be occurring.	DWQ assesses each individual site independently of others for the purposes of making determinations of impairment. Once a site is placed in Category 5, DWQ will evaluate potential pollutant sources and make a determination of appropriate next steps. If the cause of the impairment turns out to be the result of hydrologic modification, the site and associated reach could be reclassified to be a 4c water, impairment caused by other factors. However, in the case of a toxic parameter, the sources of the pollutant would need to be identified and all efforts made to reduce pollutant loads. This would most likely occur through a TMDL analysis, though alternatives to TMDLs could be explored.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
C	10	303(d)	From the spreadsheets we were unable to ascertain if the individual pH and temperature were used when assessing a data set for ammonia compliance. Specifically we looked at monitoring location ID's 4990790 and 4990880 both on the State Canal. Our continuous monitoring of the State Canal shows an average pH of 8.1. In addition, the November 2013 to February 2014 time frame averages a pH of 7.9. These pH averages would not produce a chronic or acute violation at a temperature less than 18-degrees C. We recommend that all 303(d) listings for ammonia have the exceedances compared to the criteria based on actual pH and temperature for that specific sample. In addition, we also recommend that this data set be evaluated for outliers. If only one combination of pH temperature, and ammonia exceeds the acute or chronic water quality standard this data point should be treated as an outlier and the site not listed. At least for the State Canal, we believe this water body should not be included on the 303(d) for ammonia.	For the purposes of assessing ammonia, actual paired measurements of temperature and pH were used to evaluate results against the criterion. These two sites (4990790 and 4990880) are no longer indicated as impaired due to the change in the number of exceedances required to determine impairment for toxic parameters. For the revised draft, two or more exceedances are required for impairment.
C	11	303(d)	Frequently copper is found in many locations throughout the USA to exceed the hardness adjusted criteria. However, copper exceedances are often mitigated through the use of the water effects ratio tool EPA has provided. We have seen in reports found on the internet that the use of a water effects ratio study has increased the copper toxicity concentration by as much as a factor of thirteen times the original hardness adjusted. We recommend that before a stream or river which is near a point source be listed for copper, the affected point source be given one IR cycle to produce the water effects ratio study. During that one IR cycle the site would be shown as Category 3 water. Following is a graph from monitoring location 10 4990987- North Canyon Creek in the Legacy Nature Preserve. Even a water effects ratio of two would cause all the exceedances to be within acceptable limits for this sampling location. This action would prevent listing the site for copper and a possible TMDL for this pollutant. We recommend this as a prudent step in the IR process.	DWQ agrees that a water effects ratio would provide a more refined, site-specific criterion, as would using the USEPA biotic ligand model. Using DWQ's current methodology, a site listed as impaired for copper could be delisted during a future IR cycle if new data are collected that supports a finding that the initial listing was incorrect. Unless an external entity chose to provide DWQ with new data, such as a water effects ratio or biotic ligand model, the initial investigations for the TMDL will include determining a more refined criterion using the biotic ligand model. Water effects ratios are based on bioassays, and DWQ is unlikely to have resources to support a water effects ratio determination. DWQ will consider using Category 3 as an alternative to Category 5 for the next IR cycle. The methods used should be consistent across waters; that is, the same methods should be used for waters with and without point sources.
D	1	IR	The Southern Utah Wilderness Alliance (SUWA) respectfully requests an additional two-week comment period extension for the Utah Department of Environmental Quality, Division of Water Quality, Draft 2012-2014 Integrated Report. This additional extension is warranted due to the complexity of the issues; scope and overall breadth of information; and, potential impact on human health and safety. The requested two-week comment period extension will allow SUWA to better ensure that the Integrated Report embodies the state of Utah's policy "to conserve the waters of the state and to protect, maintain and improve the quality thereof for public water supplies, for the propagation of wildlife, fish and aquatic life, and for domestic, agricultural, industrial, recreational and other legitimate beneficial uses." Utah Admin. Code R317-2-1A. In addition, it will help to better ensure compliance with relevant federal law. See, e.g., 33 U.S.C. § 1251 et seq.; 40 C.F.R. §§ 130 et seq. and 131 et. seq.	DWQ is unable to grant an extension to the comment period. However, DWQ is committed to working with stakeholders to address their concerns and will make every attempt to address comments received. Please see Letter s and the responses to the comments received on August 1, 2014.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
E	1	303(d)	Chapter 2 Assessment Methods. Page 18; The document states: Since many of Utah's AUs contain multiple monitoring locations, this approach results in a more accurate analysis of the data on a site specific basis and provides greater resolution across sites for comparison end specification of the water quality impairments within en AU. However, for reporting purposes in populating the USEPAs Assessment Database (ADS), states ere required to report use support at the AU level since it is the smallest unit of measure for individual water bodies. Therefore, if any site with sufficient date within en AU is not meeting water quality standards (i.e. category 5 at the site level) the AU will be listed as category 5. Similarly, all sites must be meeting uses for en AU to be listing as category 1 (fully supporting all uses). Comment: This is an untenable decision and will lead to considerable effort and cost in trying to perform or make sense of a TMDL. For example, the site on the Jordan River at 7800 S is listed for Se. But sites above and below this site are fully supporting. With no known point or nonpoint sources, where are you going to start with a TMDL -without spending lots of time and money? ... after which you will not likely find the source - other than perhaps an unmitigatable groundwater source that is revealed only because of upstream diversion of the river. You should more appropriately put this site in Category 3. DWQ's newfangled "conservatism" is a misnomer and will only lead to confusion and additional costly field evaluations and staff time. The IR/303(d) list should not be such an arbitrary exercise. Rather, only if the majority of sites within the AU, and sites within the downstream AU are in criteria violation, should the AU be listed. Within the context of the river continuum, please explain how you will develop a TMDL on a single site.	This change in the assessment methodology was made to avoid falsely listing an AU where several violations occur singly at several separate monitoring locations within an AU, as previous assessments have been performed. In addition, this approach avoids pooling data from multiple tributaries with potentially very different watershed characteristics, pollutant sources, and hydrology. DWQ believes that assessing at the MLID level results in a defensible dataset capturing the watershed characteristics at the site level. Many TMDLs are developed at specific sites that are representative of a given river segment and become compliance points for assessing whether the impaired water is meeting criteria, and in this respect the assessment methods are comparable. DWQ is exploring other ways of dividing our AUs into more meaningful segments (e.g. reach-based AUs) on which to perform assessments. We are currently evaluating the approach to define our AUs based on the U.S. Geological Survey National Hydrography Dataset reach codes to better refine the extent of impairment in our watersheds.
E	2	Assessment Methods	Page 20, Figure 4: The "diamonds" leading from the diamond "TMDL approved or not needed" are confusing, if not wrong. As I understand, according to your own guidance, if a TMDL is approved OR not needed, the site should be Category 4.	The diagram clearly states that if a TMDL is completed or not needed, the AU is listed as Category 4.
E	3	303(d)	Page 20, Sample size requirement: The relaxation of statistical rigor should not be explained as "more conservative". In fact, the statement " ... .4 samples (toxic parameters) collected at a monitoring location within the most recent five years to make an assessment of designated use support .. .. This rule helps ensure that assessment decisions are not made from small or sporadic data sets. However, because DWQ considers all existing and readily available data when making assessments, smaller numbers of samples may be used along with other sources of data and information to make impairment decisions. In the end, any observation that numeric criteria have been exceeded will be used to either conclude impairment or prioritize the AU for follow-up monitoring •..· Comment: This is an inaccurate statement. Indeed, the use of 4 samples is both a small and very likely sporadic data set. Further, just one sample violation, whether among 4 or 40 samples is used by DWQ to designate not supporting. So the actual range of percent violation may range from 1% to 25%. This is a vast divergence from the past assessment protocol and is not standardized rigorous science. This approach was used without any determination of the possibility of a statistical outlier. At a minimum, DWQ should perform the standard outlier test (1.5 X the median of the upper and lower quartiles of the data set). This new assessment protocol, in addition to using a single sample site data within an AU, which is also scientifically indefensible, has nearly doubled Utah's list of impaired waters compared to the 2010 IR. This will spread limited financial and staff resources much beyond current levels and for poor scientific reasoning. A more robust data set with better statistical rigor is absolutely essential and will be far more scientifically defensible than this new protocol. DWQ needs to explain how it will proceed with TMDL development for these sites. DWQ should, at a minimum, place the newly identified impaired sites in category 3.	Due to public comment and internal review, DWQ adjusted its exceedance rate for toxic parameters to two or more samples compared to one or more samples in the draft IR. This was in consultation with the Environmental Protection Agency (EPA) and their guidance on a recurrence interval of two or more in a 3-year period. DWQ does not use a percent exceedance method for toxic parameters. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
E	4	303(d)	Page 21. Table 2, third column states: Numerous recurrence intervals are listed. Minimum and 30-day averages are used for assessments based on grab samples. Comment: This statement is confusing; DWQ should explain how it defines recurrence intervals. Secondly, the nonscientific use of grab samples to perform an assessment of the 7-day average or 30-day average chronic DO criteria has been the subject of several earlier comments on DWQ IR and TMDL documents for more than five years. EPA has also reviewed these documents and my comments, which criticize DWQ for ignoring EPA guidance (EPA 440/5-86-003, 1986). Therefore, by ignoring EPA recommended assessment methods against chronic DO criteria, EPA and DWQ remain inexcusably remiss in their performance and review of the chronic DO assessment. As I have stated before, with such heavy financial and political responsibility, DWQ and EPA should have followed EPA's own guidelines in establishing 7-day and 30-day average DO values in order to truly assess the water body for DO impairment. Now, 28 years after that criteria document was published, and even after recording data sondes have been available for more than 10 years (making this assessment extremely easy and scientifically accurate) EPA and DWQ continue to ignore this protocol. Accordingly, the following examples demonstrate, once again, the scientific shortfall of DWQ's inappropriate and scientifically indefensible chronic DO assessment protocol.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	5	303(d)	In this exercise, I will review the data set that DWQ used in performing the beneficial use support assessment. 1. It should first be noted that the morning DO minima in the Salt Lake Valley will occur between about 0600 and 0630 for the period of about a week before to a week after summer solstice. Earlier and after this period the minima rapidly grow later to occur at about 0800 to 0830 by the beginning and end of the growing season. This has been observed repeatedly in data recorded on the JRFBWQC ysieconet.com website. Oddly, the afternoon maxima isn't nearly as sensitive to the summer solstice, with the minima usually occurring between 1630 and 1800, depending on riparian vegetation and tree height at the sample location. Also, because the Salt Lake Valley is located toward the western side of the Mountain Time Zone, solar noon occurs at about 1330, rather than at the 1200 hour. As such, it is also important to note that the daily average DO falls very near the solar noon. The facts are, that these minima and maxima data values were easy to collect as grab samples, and as of 10 years ago, there is no excuse not to use this technology for such an important assessment as the Jordan River. Similarly, the fact that DWQ grab sample data were collected without regard for the diel DO pattern, and particularly with a preponderance of data collected during morning hours- even during 2012 - suggests the absence of scientific understanding and motivation necessary to collect appropriate data. Further, this assessment protocol was performed without regard to EPA's own guidelines. To claim the lower Jordan impaired based upon such biased and inadequate is inexcusable and profoundly unacceptable for this or any IR report during the last 25 years. State Canal data: The following chart summarizes the data set used in DWQ's assessment of the State Canal at approximately 400 S. Bountiful (access road to Newstate and Burnham Duck Clubs). Briefly, six of the 116 samples (5.1%) violated the 4.5 acute DO criterion for fish reproduction. Three of the 116 samples (3.4%) violated the 4.0 acute DO criterion for the non fish production period. Conclusion: <1 0% violated =fully supporting. However, for the chronic (7-day and 30-day) criterion, DWQ did not comply with EPA guidelines and tallied the total number of individual measurements that exceeded 5.5 mg/L. Twenty out of the 116 samples (17%) had DO values < 5.5. Yet, all but one of the 20 samples was collected before the solar noon (see figure below). This is a direct violation of EPA's assessment guidelines for assessing 7 -day or 30-day average DO criteria and nullifies DWQ's conclusion that the site is impaired. Nevertheless, these data have some screening value and the AU could logically be placed in Category 3 until an accurate assessment, as described above, is performed.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
E	6	303(d)	Jordan River 1000 ft below South Davis South Plant The following chart summarizes the data the DWQ used to assess this site. Three out of the 40 samples (7 .5%) used in this assessment violated the 4.5 and the 4.0 mg/L acute DO criteria. Conclusion: <1 0% of samples violate the criterion = fully supporting. Seven of the 40 samples (17.5%) of the data points were < 5.5 mg/L. However all of these measurements were performed before the solar noon, severely biasing the data toward the morning minimal values- as opposed to performing the chronic DO assessment according to EPA guidelines. Therefore, DWQ's assessment of this site as impaired only results from a biased, scientifically indefensible method.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	7	303(d)	Jordan River AB Burnham Dam and St Canal The following chart summarizes the data the DWQ used to assess the Jordan River at Burnham Dam. Data were collected between 2010 and 2012. Of the 62 samples, only 1 sample (1.6%) violated the 4.0 or 4.5 acute DO criteria. Conclusion =site is fully supporting. Six measurements (9.7%) were below 5.5 mg/L. Of these, all but 1 were measured before the solar noon. Nevertheless, only 9.7% of the measurements exceeded the criterion. Therefore, regardless of the inaccurate method that DWQ uses to assess chronic criteria, this site is fully supporting.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	8	303(d)	Jordan River at 500 N: The following chart summarizes the data the DWQ used to assess the Jordan River at 500 N. Data were collected between 2010 and 2012. Of the 62 measurements, four samples (6.4%) violated the 4.5 mg/L criterion for the reproductive season. Three measurements (4.8% were below the 4.0 mg/L criterion for the nonreproductive period. Conclusion: <1 0% of samples violated the criterion =fully supporting. Seven measurements (11 %) were <5.5 mg/L. However, four of the seven samples were measured before the solar noon, biasing the data toward the morning minimal values - as opposed to performing the chronic DO assessment according to EPA guidelines. Therefore, DWQ's assessment of this site as impaired only results from a biased, scientifically indefensible method.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
E	9	303(d)	Jordan River at Gadsby Power Plant: The following chart summarizes the data the DWQ used to assess the Jordan River at the Gadsby Plant. Data were collected during June, July and August of 2004. Of the 29 measurements, two (6.9%) were< the 4.5 mg/L criterion and 1 measurement (3.4%) was below the 4.0 criterion. Conclusion: <1 0% of samples violated the criterion = fully supporting. A total of 13 measurements (44%) were below 5.5 mg/L. However, all of these measurements were performed prior to the solar noon, severely biasing the data toward the morning minimal values- as opposed to determining the true daily and 7- and 30- day average values according to EPA guidelines. Therefore, DWQ's assessment of this site as impaired only results from a biased, scientifically indefensible method.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	10	303(d)	Jordan River at 1800 N: The following chart summarizes the data that DWQ used to assess the Jordan River at 1800 N. Of the 51 measurements collected during this 8-year period, five measurements (9.8%) were in violation of the 4.5 mg/L reproduction season criterion and the 4.0 mg/L non reproduction period. Conclusion: <1 0% of samples violated the criterion =fully supporting. Eleven measurements (21.6%) were< 5.5 mg/L. However, eight of the 11 measurements (73%) were performed before the solar noon, severely biasing the data toward the morning minimal values- as opposed to determining the true daily and 7- and 30- day average values according to EPA guidelines. Therefore, DWQ's assessment of this site as impaired only results from a biased, scientifically indefensible method.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	11	303(d)	Jordan River at 1300 S: The following chart summarizes the data the DWQ used to assess the Jordan River at 1300 S. As with the data collected at the Gadsby Plant, all data used at 1300 S werecollected during the summer of 2004. 2004 was a unique year in that it represents one of the most severe drought years in recent history. This resulted in both lower-thanaverage flows (see USGS chart below) and consequently warmer than average temperature (data not shown, but within DWQ's database). This represents another primary issue with listing the lower Jordan River as impaired for DO. Specifically, DWQ knows that the river is vastly dewatered as of result of diversion to the Surplus Canal. This diversion results in a loss of 50 to -98 % of river flow with an average of about 60% during normal summer flow. This dewatering results in drastically reduced flow and velocity and an increase in temperature (as compared to that in the Surplus Canal). In turn, the river transitions to nearly exclusively a depositional zone, dominated by silt, clay and organic debris, making it subject to frequent dredging, all of which also results in severe habitat loss (Fiuckiger and Miller 201 0; Biowest 1990, 1992). Hydrological modification and habitat loss are two specific "g" factors identified in 40 CFR 131.1 O(g) that qualify an AU for the performance of a Use Attainability Analysis as an off ramp to meeting normal or appropriately applied and achievable water quality criteria or biological integrity goals. I admonish DWQ for ignoring these absolute limiting factors in that they will forever prevent the lower Jordan from obtaining habitat quality goals as defined in the Federal Clean Water Act. Further, this dewatering and warming reduces DO solubility and atmospheric reaeration that would normally occur in the Jordan River under natural flows; (i.e. the Surplus Canal (the same water) has never experience a DO criteria violation). Consequently, measurement values during the summer of 2004 represent the most severe conditions since humans settled in Salt Lake Valley. From the data, two of the 14 samples (14%) violated the 4.5 mg/L DO criterion and one measurement (7%) violated the 4.0 non-reproductive period criterion. It could be concluded that the two violations constitute impairment. However, as discussed above, serious ecological conditions are created as a result of dewatering and indeed during the sampling event of July 14, 2004 the River was exceptionally dewatered to about 2 CFS (see USGS figure below), in anticipation of a storm event. The river would generally be composed of several nearly-stagnant pools linked by a relative trickle of very warm water. This constitutes a very real example of why the Lower Jordan should undergo a UAA to document such severe habitat destruction as result human interference. Finally, the entire data set was collected before about 1015 hours. Consequently, the nine samples (64%) were severely biased toward the morning minimal values- as opposed to determining the true daily and 7- and 30- day average values according to EPA guidelines. Therefore, DWQ's assessment of this site as impaired onlyresults from a biased, scientifically indefensible method.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
E	12	303(d)	Jordan River at 700 South: The following chart summarizes the data the DWQ used to assess the Jordan River at 700 S. As with the data collected at the Gadsby Plant and 1300 S, all data used at 700 S were collected during the drought summer of 2004. All critical comments listed above also apply to the 700 S data. Particularly, the data were collected during one of the hottest and lowest-flow summers in recorded history. The only difference is that, of the 18 DO measurements, there were three afternoon measurements. Three of the 18samples (17%) violated the 4.5 mg/L DO criterion, including one measurement made during the July 14 severely dewatered period. Two samples (11%) violated the 4.0 standard. The use of this 1 0-year-old data, resulting in just barely exceeding the assessment criterion should have prompted DWQ to place this site, the Gadsby Plantsite, and 1300 S site in category 3 - insufficient (recent) data to make an assessment decision. This is particularly true when two of these sites constitute the only sites and sampling periods where violations occurred in the lower Jordan River. This information,along with the consistent misapplication of data for chronic DO criteria assessment suggests that DWQ is overly anxious to "list" sites as impaired.	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	13	303(d)	In summary, DWQ sampling and assessment protocols continue use only the same grab sample data used for the instantaneous acute criteria assessment for assessing the chronic DO criteria by merely comparing the instantaneous data to the 7-day and 30-day criteria (5 mg/L) - requiring the same 10% of "violations" to claim non supporting. As commented many times and for many years earlier, this treats the 7-day and 30-day average DO criteria as de facto acute criteria. Again, this is unnecessarily stringent, scientifically inappropriate and violates EPA's own guidelines for collecting and assessing data against the 7 -day and 30-day average DO criteria. EPA guidelines instruct: measurements must capture the morning DO minimum and the afternoon DO maximum concentration. The daily mean is the average of these two values. This is to be repeated for 7 consecutive days (for the 7-day average criterion) or 30 consecutive days (for the 30-day average criterion). The average DO for the 7 (or 30) consecutive days is then determined and compared to the respective criteria (5.5 mg/L). This should be the guiding protocol. As these assessment guidelines were recommended by EPA in its DO criteria document (EPA440/5-86-003 Ambient Water quality Criteria for Dissolved Oxygen, April1986), EPA is also negligent in not correcting Utah's 305(b) assessment method for chronic DO criteria throughout these many years of performing beneficial use assessments. Both EPA and DWQ need to respond to this comment. :	Phased TMDLs to address the lower Jordan River's DO impairment have been approved. We look forward to working with stakeholders to devise a method to evaluate continuous DO data to accurately assess continuous water quality data. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.
E	14	303(d)	The assessment of ammonia values has also been inappropriately performed at two sites on the State Canal. The following graph includes all the data used by DWQ for assessing the site at approximately 400 S in Bountiful. The chart simply graphs the data against the ammonia criteria generated using the mathematical adjustments for temperature and pH values at that site and time of sample collection. The graphs clearly show that there are no acute or chronic criteria violations - even during the fish reproduction season. Therefore, the assessment of this site should be changed to "fully supporting". State Canal Above the South Davis South Plant: The following graph includes all the data used by DWQ for assessing this site. The chart graphs the data against the ammonia criteria generated using the mathematical adjustments for temperature and pH values at that site and time of sample collection. In addition, the pH is graphed in order to demonstrate to occurrence of an outlier value on January 8, 2008. This value is nearly 2 orders of magnitude greater than typical winter values, which hover very near 7.9. Use of the statistical outlier calculation, clearly demonstrates that this is an outlier. Consequently this site should be listed in category 3 until additional seasonal data are collected to demonstrate whether the ammonia criteria are truly violated. This is another example of why the change in assessment criteria where only one sample violation is necessary to list a site is inappropriate and allows the assessment to be subject to outlier data. This assessment procedure should be abandoned.	For the purposes of assessing ammonia, actual paired measurements of temperature and pH were used to evaluate results against the criterion. These two sites (4990790 and 4990880) are no longer indicated as impaired due to the change to the number of exceedances required to determine impairment for toxic parameters. For this revised draft, two or more exceedances are required for impairment. DWQ has identified a number of assessment method improvements and will take these comments into consideration in that process for the 2016 IR. DWQ plans on engaging stakeholders and issuing revised methods for public comment in January 2015.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
E	15	303(d)	Assessment of Mill Creek near the Jordan River: The conclusion of impairment of Mill Creek due to 0/E calculation provides a scientific/assessment policy review. For example, no chemical parameters exceed their respective criteria. Therefore the reason for listing this site in category 5 is questionable. An alternative and recommended assessment would be to place the site in category 4 - due to hydrologic modification (dewatering) or physical habitat loss (channelization, riparian destruction, sedimentation)- all of which are the truthful and directly affect 0/E. Overall, the use of 0/E to list in Category 5, when there are no accompanying chemical violations, is too premature and may be an inaccurate assumption. Category 4 or Category 3 are the logical assessment endpoints when the essential accompanying data is not available or is not used.	Following EPA guidance on independent applicability (http://www.epa.gov/npdes/pubs/owm0232.pdf), DWQ assesses biological criteria independently of chemical criteria. Once placed in Category 5, DWQ's Watershed Protection staff will follow up to confirm the impairment, identify stressors, and begin the process of restoration. Next steps will depend on the stressors identified. If the stressor is determined to be a pollutant, DWQ will pursue development of a TMDL. If the stressor is determined not to be a pollutant, the waterbody could be recategorized to 4c in a future assessment cycle.
F/G	1	303(d)	Many of the criteria/analyses used in the 2014 Narrative were developed specifically as first round 'screening' tools, particularly 0 /E bioassessment. These assessments should not be used for anything other than their intended use; initial screening. They were never intended for monitoring or any scientific analysis. Assessments by their very nature are a critical link between science and managers. Indeed managers often depend on scientists to develop easily understandable measures of ecosystem health and scientists concerned with the ecosystems that they study are civilly obliged to produce the most meaningful, state-of-the-science assessments to managers. However, assessments are not science and should not be used as a substitute, although assessment results can sometimes be used to develop scientific hypotheses for further studies. In addition, and contrary to current management agencies' agendas, assessments are not valid monitoring tools, primarily because of their poor discriminatory power and lack of ability to measure anything less than very large changes in ecosystem health and they should not be used as such. Example of Bioassessment tool that is too simplified for assessment. monitoring or scientific investigation. For example, the 0 /E biocriteria method produces a single value that was designed for a single group of organisms (macroinvertebrates) and their assumed response to a combination of many environmental stressors. 0 /E models first determine which group a 'test' stream is to be included in prior to 0/E calculation (AU grouping methods are also subject to scrutiny and will be discussed further). Once the grouping of the stream is assigned, then 0 /E is calculated. Again, macroinvertebrate taxa in 0 /E are responding to generalized, accumulative types of stressors. However, it is well known that individual macro invertebrate taxa respond differently to a wide array of environmental stressors. That is one reason they have evolved as separate species. Macroinvertebrate taxa and certainly entire aquatic ecosystems are not simple, single- dose response interactions. Aquatic ecosystems are extremely complex with interactions occurring at multiple levels, starting from individual species up to interaction with the entire ecosystem. Each of the hundreds of species in an aquatic ecosystem has a niche that can be defined as a Hutchinsonian N-dimensional hyper- volume, not just one, two or three dimensions, but many. Multiply these multidimensional niches of each species by interactions with all the other species (competition, predation, parasitism, mutualism, facilitation, etc.), trophic level interactions, functional feeding groups, interactions between aquatic and riparian ecosystems, stream connectivity (which include meta population dynamics and genetic interactions and which were part of Karr's definition of 'biological integrity' but are now completely ignored by water quality management agencies), natural spatial and temporal variability, anthropomorphic impacts, etc. etc. and any stream ecologist would shake their head from side- to- side in complete disbelief that the health of a stream could be summed up to one subjectively determined number (score), even if it means making things simple for managers to easily comprehend. One single 0 /E score cannot possibly hope to capture the health of an entire aquatic ecosystem. It is just not possible. Thus the reliance of 0 /E to assess water quality condition and whether a water body supports or doesn't support its designated use often results in an injustice to water users and regulators, particularly those who are responsible and dedicated to maintaining water quality. It also does a disservice to the very waters of UT, themselves and the biota that reside in them. Utah's rivers and streams are much too valuable and a treasure to all its residents to use flawed, highly simplified, or limited number of metrics. The reliance of 0 /E by UTDWQ also can reflect poorly on the very agency that citizens of UT have entrusted to protect their waters.	The O/E biological assessment method is commonly used across the United States, Europe, and Australia to assess the quality of their waters. In Utah, this method has been in use for nearly 10 years and improves with each assessment cycle. Also, please be aware that DWQ uses multiple lines of evidence as discussed throughout the 2014 IR, and DWQ follows the Consolidated Assessment and Listing Methodology (CALM) guidance referenced in the Biological Assessments of Rivers and Streams section of Chapter 2 in the report. Further, DWQ is required by R317-2-7.3.c. to use quantitative biological assessment methods that have "documented methods that have been subject to technical review and produce consistent, objective and repeatable results that account for methodological uncertainty and natural environmental variability." Alternative biological assessment methods would require the same level of technical review and documentation that has been completed for the currently employed methods. DWQ is not aware of other methods applied by other states or recommended by the EPA that incorporate n-dimensional variables or explicit statistical analysis of Hutchinsonian niche theory. DWQ would welcome documentation of how these and other ecological theories could be applied on a state-wide basis to biological assessments.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
F/G	2	303(d)	Throughout the Narrative, DWQ uses Type I error rates to support its decisions concerning water quality conditions, i.e. "Is there a problem?" and "How extensive is the problem". Type I, a classical frequentist's statistic, tests the null hypothesis of no effect vs. the alternative hypothesis of an effect, in this case; no impairment vs. impairment. Type I error, measured as alpha or 'p' (1-alpha), occurs when the null hypothesis is actually true but was rejected as false. That is, there truly was no impairment but the conclusion was that there was impairment (also known as a false positive). Type I error is often illustrated by the story of' crying wolf when there actually wasn't a wolf. If one cries wolf too many times, no one would believe them when the wolf actually showed up. By focusing on Type I error, particularly with small sample sizes, too often UTDWQ may have ended up reporting impairment when there really was no impairment This could add a considerable workload and expense down the road for UTDWQ or other managers trying to determine if indeed there was impairment which then must commit to an expensive TMDL or other restorative efforts or commit additional resources necessary for delisting a water body. It can also result in doubt of DWQ's ability to truly detect impairment. It has been suggested that this approach may be "proactive, precautionary, or erring on protection, etc." This may be true, particularly given the small sample sizes used in the assessments. The number of samples used in an assessment determines the level of ability to detect changes. Small sample sizes typically only allow for detection of large differences, whereas large sample sizes can detect small differences in assessment tests, although in some cases small differences may be ecologically irrelevant. However, by limiting sample size to very small levels, individual data points will have undo influence, particularly outliers. These few data points could prompt UTDWQ to conclude impairment when in fact there was no impairment If more data were used, a single data point would have less of an effect on the conclusion. An example would be if three samples were collected in the same location and a metal such as aluminum was found to be in exceedance of standards in one of those three samples (33% of the samples) vs. the same sample was found to be in exceedance but one thousand samples were collected (0.01% of the samples). Type II error, on the other hand, is a more precautionary type of error that may have been of better use by UTDWQ for determinations. Type II error (B) is the probability of concluding there wasn't impairment when in fact there was (not crying wolf, when it actually it was there). This is of major concern because DWQ may be evaluating UT waters incorrectly by not detecting true impairment and possibly allowing impairment to continue. Significantly and ecologically meaningful Type II error levels often require substantially more data points. However, by collecting additional data and using existing data from all available sources, assessments that incorporate Type II error could save time and money in the long run. The reliance on Type I error evaluation allows DWQ to say, "oops we thought there was an impact but on closer evaluation it looks like there wasn't". "Guess all that money spent to fix things wasn't necessary after all". In addition by using only a few samples, Type II error would likely result in the inability of the assessment to detect a true impact, which could be more harmful than Type I error. Again, more samples should be included in an assessment.	DWQ's assessment methods balance the need to identify impaired waters with making maximum use of available data. Based on other comments received on the IR, a large number of waters listed as impaired in the draft IR have been moved to Category 3A. For example, toxic listings are based on at least two exceedances in the final IR. DWQ concurs with the commenter that additional data collection will be necessary to confirm impairment before developing a TMDL or reassessing a water body in Category 3 (insufficient data). DWQ refers the commenter to the Sample Size Requirement section in Chapter 2 of the IR. A minimum of four samples must be available for a site to be assessed. DWQ's monitoring budget is limited, and the agency has been creative in the use of funds to monitor the state using a rotating basin and probability sampling protocol. DWQ welcomes additional data provided by cooperators.
F/G	3	303(d)	The term "Biological Composition" used throughout the Narrative. Not sure what 'biological composition' means, but however UTDWQ defines 'biological composition'; 0/E does not measure it. 0/E purportedly only measures taxa richness, not composition. Community or assemblage composition of course is not the same as taxa richness.	The two instances where this wording occurred were corrected in the document. Thank you for the comment.
F/G	4	303(d)	Comment X. Introduction Page 6: First paragraph last sentence: "Even completely subjective .... (fish kills ..) Not sure how a fish kill can be considered less subjective than the other morepreferred methods of evaluation used in the narrative. I guess if one was to say, "all those fishies look dead to me" (i.e. subjective) as opposed to going over and kicking a few and smelling them (objective) and concluding, "yep, they are dead" would count. In reality all of the methods used and decisions made in this narrative are subjective, including decisions based on statistical tests. This is the inherent nature of every conclusion based on statistical inference. A decision to use a Type I error alpha level of 0.05, 0.10, or any value is subjective. Why not alpha= 0.06? or 0.04? The decision to be consistent with the choice of an alpha level for every test (e.g. 0.05) and for every circumstance is also subjective. An alternative would be to examine each test (criterion) and its resultant alpha level and then making a conclusion based on their own merits or importance. This is particularly critical due the economic and ecological importance of a decision to list an AU as impaired or not. Fish kill observations are perhaps the most objective of all the	The comment appears to be editorial because no actions are recommended; therefore, no response is necessary.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
			decisions used in this narrative. If a fish kill is observed, then the stream is impaired. However, determining the cause of the kill would require further investigation.	
F/G	5	303(d)	Category 4C page 8, first sentence "Category 4C: The impairment is not caused by a pollutant: Assessment units are listed in this subcategory if the impairment is not caused by a pollutant (e.g., habitat alteration, hydromodification)." Comment: Many of the streams listed as "not-supporting" in the Narrative should have been in this category and not in the not-supporting category.	Without specific examples of situations where waters should be assessed as 4C and no justification provided, DWQ cannot address this comment.
F/G	6	303(d)	"·Quantitative biological assessment results for streams and rivers are statistically different than the reference site conditions." Comment: Assessments could be statistically different than reference conditions because of many factors other than impairment (see example below).	As with any assessment, there are a number of variables that must be quantified to determine how much of the difference is attributed to a "signal" (e.g. human-caused stress) and/or "noise" (e.g., various sampling errors). DWQ develops thresholds specific for each assessment method—in this case the O/E model, which incorporates the variability associated with sampling. This information is explained in the Biological Assessments of Rivers and Streams section of Chapter 2 of the report.
F/G	7	303(d)	Page 34, First Sentence: "E is then calculated as the sum of all taxa Pc s that had a greater than 50% chance of occurring at a site given the site's specific environmental characteristics." It appears that 0/E development requires that taxa have a probability of occurrence of> 0.50 in (reference) streams to be part of the model. If this interpretation is correct, then an unknown number of taxa are automatically removed from consideration in 0 /E. This unknown number of omitted taxa could be very large depending on which group the reference streams are placed under. It appears that only the ubiquitous, common, cosmopolitan, tramp species are used as 'observed' taxa; taxa that are likely to be tolerant of a wide range of environmental variables and not likely responsive to stressors. For example Baetis sp. (mayflies) have a probability of occurrence for a reference group of 0.95 (expected to occur). Obviously, certain Baetis sp. are cosmopolitan (e.g. Baetis bicaudatus) and which have a wide range of environmental tolerances. B. bicaudatus also occur in streams in less than reference condition (e.g. Mill Creek, SLC at confluence with Jordan River). Thus baetid mayflies as a group and particularly B. bicaudatus are very poor indicators of water quality. As far as a test sample observed inclusion or exclusion of B. bicaudatus, or any other taxon, it would be entirely dependent on whether a complete census of the entire reach of stream under consideration was conducted or not. If a census of the entire reach of concern was not conducted and only composite sampling and then laboratory subsampling was conducted, it would be impossible to know if that taxon was truly present or absent. Low abundance of a taxon and hence its omission from observed status does not mean extinction of that taxon from a water body. It simply means it was not observed. In addition, a taxon could occur at 90% of the total assemblage abundance in many streams (i.e. be locally abundant within some streams), but have a probability of occurrence of< 0.50 in all reference streams and therefore, excluded from the 0/E model. The 0/E model assumes that macroinvertebrate assemblages in all streams in a group are similar and individual taxa occur at equal abundances and none are unique or whose populations are not dynamic. However, the IR mentions that samples should be conducted or evalutated every three years or so to account for natural variabilityin taxa abundances. Again, all of these assumptions are not likely. 1) A test stream may fit poorly into its designated reference group and its macroinvertebrate assemblage may be quite a bit different than reference, 2) taxa don't occur at equal abundances within any stream and therefore have unequal probabilities of detection using DWQ methods, and 3) a major factor in the population dynamics of a taxon is its generation time. Taxa with short generation times (e.g. midges; several weeks to a few months) have greater variability in abundances due to environmental conditions than do taxa with longer generation times (e.g. large stoneflies, 1- 3 years). If test samples aren't conducted during a similar point in a taxon's natural population abundance cycle, erroneous conclusions about 'observed' status will be made. As stated in theIR, 0/E ignores many taxa; many of which may be rare, uncommon, cryptic, or even may be very common in many streams. This could include federally listed threatened and endangered species which most U.S. citizens support the protection of. UTDWQ appears to be ignoring an unknown number of taxa, not assessing their status (i.e. biological integrity), and differing to USFWS to deal with them under the Endangered Species Act (if they ever make it to that list). Also, rare and uncommon taxa are much more likely to be indicators of water quality and should be the focus of water quality assessments rather than cosmopolitan taxa. In addition, ecologists are now well aware that rare and uncommon taxa often have disproportionally greater influence on ecosystem function than common taxa. For example, the salmonfly, Pteronarcys californica may have a	The O/E model uses a >0.50 probability of taxa occurrence as a threshold because repetitive studies have shown that these models are more sensitive and precise at detecting biological stress. Additional information about these O/E model results has been added to the Biological Assessments of Rivers and Streams section of Chapter 2 of the report. DWQ is required to use quantitative biological assessment methods that have "documented methods that have been subject to technical review and produce consistent, objective and repeatable results that account for methodological uncertainty and natural environmental variability." Alternative biological assessment methods would require the same level of technical review and documentation that has been completed for the currently employed methods, which have been in use in Utah for nearly 10 years.

Letter	Comment #	Section	Comment	Response
			probability of occurrence <0.50 in the stream group but when it occurs in large abundances in some streams it has a disproportionately large influence on the entire ecosystem, from reducing CPOM to FPOM all the way to a being a critical food item for breeding birds when it occurs as an aerial adult. Other less common taxa such as the stonefly Yoroperlla sp. or caddisfly Helicopsyche borealis, etc. are likely excluded from the models. These examples and many other 'uncommon' taxa are uncommon for several reasons: because they have limited geographic distributions or distributions within a region or streams, limited range of environmental tolerances and conditions, or are less tolerant to human disturbance than the cosmopolitan taxa used in the model. The IR acknowledges the problems with 'rare' and 'uncommon' taxa but it is likely that far too many taxa are considered 'rare' or 'uncommon' in the 0/E models and are done so if favor of making the 0 /E models function. How many or what proportion of taxa can occur in < 50% of the streams and were excluded from the 0/E model? IfO/E represents local extinctions then the status of these taxa were not included in the estimate and 0 /E could be grossly underestimating local taxa extinction. Again, 0 /E would not be quantifying loss of biodiversity except in the crudest sense and if it does it will be only for ubiquitous taxa.	
F/G	8	303(d)	Section: RIVER INVERTEBRATE PREDICTION AND CLASSIFICATION SYSTEM (RIVPACS) MODELS\Page 32 3rd paragraph second sentence: "In essence, 0/E quantifies loss of biodiversity." Comment: No it does not. 0/E does not quantify loss of biodiversity. It may on occasion, but it is unknown from the assessment ifloss of biodiversity (taxa richness) actually caused the change in 0 /E values. Particularly when ftxed count subsampling methods are used in the taxonomy lab. It could very well be that biodiversity hasn't changed but that one or more taxa may have happened to become more or less abundant (i.e. change in evenness). If used at all, 0 /E should more appropriately be used as a rudimentary measure of 'evenness'. Please see hypothetical example below.	Thank you for the comment. The sentence has been revised to read "O/E quantifies loss of predicted taxa."
F/G	9	303(d)	"Despite the mathematical complexities of model development, 0 /E is easily interpreted as it simply represents the extent to which taxa have become locally extinct as a result of human activities. For example, an 0 /E ratio of 0.40 implies that, on average, 60% of the taxa have become locally extinct as a result of humancaused alterations to the stream" Comment: Again, this is likely not true (see example below). In addition, the statistical methods (models) that went into the development of RIVPACS 0 /E model have associated error or variability. For example, cluster analyses that were used to develop reference groups have associated error rates and there are many cluster analysis methods available, each potentially resulting in a different set of reference groups. DWQ likely used the most appropriate cluster method based on either cluster model comparisons or best professional judgment or both. However, there still are error rates associated with the best method used. The probability of occurrence of a taxon in a reference group also has inherent uncertainty or error. RIVPACS 0/E models are 'models within models' each of which contributes uncertainty either additively or multiplicatively. These error rates need to be taken into account and reported in the IR.	The O/E biological assessment method is commonly used across the United States, Europe, and Australia to assess the quality of their waters. As with any assessment, there are a number of variables that must be quantified to determine how much of the difference is attributed to a "signal" (e.g.. human-caused stress) and/or "noise" (e.g., various sampling errors). DWQ develops thresholds specific for each assessment method—in this case the O/E model, which incorporates the variability associated with sampling and natural variability. This information is explained in the Biological Assessments of Rivers and Streams section of Chapter 2 of the report. The thresholds are updated with each assessment cycle to accommodate new data and provide more accurate assessments.
F/G	10	303(d)	Natural Variability and Sampling Error 1) Natural variability (e.g. annual, seasonal, and year- to-year variability in physical conditions and macroinvertebrate abundances), 2) within stream variability (riffle to riffle or riffle to other type of habitat e.g. riffles tend to have more taxa than pools), 3) field sampling error (e.g. estimating the 1 sq. ft. area needed to be sampled), 4) sample processing error (e.g. proper preservation, handling, and storage), and 5) laboratory error (rolling up of taxa, different levels of taxonomic QA/QC, etc.); while hopefully kept to a minimum can add up to the likelihood of erroneous 0/E scores far greater than the 0.01level used by UTDWQ to conclude fully supporting vs. non- supporting. For example, the UTDWQ 0 /E scores of 0.83 or 0.78. It appears that in an 0/E model with for example, 100 taxa, the assumed loss of only one taxon could result in a change in use support status even though it could have been due to natural variability or sampling or modeling error.	DWQ recognizes that all models have error. DWQ strives to quantify those errors and incorporate them into assessment thresholds with each new assessment cycle. For example, the current model incorporates macroinvertebrate samples that were collected year-round, thus incorporating seasonal variability into the model. The remaining points identified by the commenter are all accounted for in the model because reference samples are held to the same quality standards as every sample. This is so that if these potential errors occur, they occur across all samples.
F/G	11	303(d)	Fixed Count Subsampling Error: Composite samples of eight, 1-sq. ft. kick samples recommended/endorsed by UTDWQ can often have large number of individual organisms, sometimes > 10,000 individuals. To reduce the amount of cost and effort in processing this large number of organisms and to standardize samples across regions, UTDWQ and 0 /E models typically use laboratory produced 500 organism subsamples. Data from these processed samples are then entered into 0 /E resulting in scores that are assumed to represent taxa richness and/ or "the percentage of taxa that have become locally extinct as a result of human- caused alterations to the stream". Again, this is most likely an incorrect conclusion as illustrated by an example	DWQ contracts with taxonomy laboratories that have EPA- and DWQ-approved quality assurance and quality control practices and procedures to ensure the highest level of consistent, quality data. In particular, the current laboratory requires subsample counts that exceed the number recommended in the literature. Please review Ostermiller, J.D. and C.P. Hawkins. 2004. Effects of sampling error on bioassessments of stream ecosystems: application to RIVPACS-type models. Journal of the North American Benthological Society 23(2):363–382.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
			below. Hypothetical example of UT DWQ 0 /E miscalculation and false conclusion The following is a hypothetical example of the 0 /E fixed count subsampling problem. Methods: Two com posited samples from eight, 1-sq ft. kick samples were collected at the same location but in different years. Sample 1 was collected the year prior to Sample 2. Both samples were collected in riffles and 'other' acceptable habitats (e.g. runs, banks, etc.). There were ten taxa collected in both samples; two mayfly (Ephemeroptera) taxa, two stonefly (Plecoptera) taxa, one caddisfly (Trichoptera) taxon, two midge (Chironomidae) taxa, and one taxon from the following groups, snails (Gastropoda), scuds (Crustacea), and segmented worms (Oligochaeta). In the first sample, the number of individuals for each of the ten taxa was 1000. In the second sample there were substantially more mayflies and stoneflies than the first sample and less individuals of the other six taxa (Table 1). In both samples the total number of individuals was equal; 10,000. This number of individuals is not unusual for composited samples from stream systems in UT. Results of the mean numbers of individuals of each taxon for each of the two samples using a 500-organism subsample method are in Table 1. Any mean values< 1.0 in the 500 count subsample (Table 1) indicates that on average the taxon occurred less than once in the 500 count subsample and was therefore, never observed or counted. Any mean values > 1.0 indicates the taxon occurred in the 500-count subsample and was observed and counted. Results: The mean total number of taxa counted and reported from the 500-count subsample in Sample 1 was ten and the mean total number of taxa in sample 2 was five. This represents a 50% difference in total taxa reported, even though there was actually the same number (1 0) of taxa collected in the original samples. Conclusions: The conclusion using UTDWQ 0/E criteria would be that: 50% of taxa became extinct from when Sample 1 was taken to when Sample 2 was taken due to human impact Therefore, the stream is not supporting its designated use. The conclusion of a stream ecologist would be that: Biodiversity may not have changed from year- to -year and production (total number of individuals) may not have changed, as well, but cannot tell using results from a fixed subsample method. If the entire samples were analyzed, then biodiversity and production did not change. Indicator taxa that often represent good water quality (mayflies and stoneflies) increased by 2 to 3 times, and those typically considered poor water quality indicators decreased by almost 67 times. However, most stream ecologists disagree with bioassessment programs that suggest that all midges, snails, crustacean, and worms should be classified as poor water quality indicators and caution should be applied to this statement. The conclusion of the stream ecologist would also be that it appears that water quality improved and could have been due to natural variability or given additional data, likely improved because of decreased water temperature or conditions that favored mayflies and stoneflies, particularly stoneflies in the functional feeding group, shredders. Increased shredder abundance was likely due to increased riparian cover. The scraper snail taxon Physa sp. may have decreased in abundance due to less light from increased riparian cover. Whatever those humans did (e.g. increased riparian cover which may have decreased temperature, increased allochthonous production, and decreased autochthonous production); Keep up the great work! Using fixed count subsample method could have resulted in just one single taxon not being counted if it occurred at low abundances in the stream and thus lowering the 0 /E score from supporting to not supporting even if it was present in the stream.	
F/G	12	303(d)	Jordan River Question: Was 0/E conducted for all sites on Jordan River or just those that resulted in notsupporting? If 0 /E wasn't conducted in these A Us then decisions were based on only one type of measure (line of evidence) (e.g. chemical, etc.). Also, was there a reference stream to compare the Jordan River to and what was it?	Biological data collection on the Jordan River is collected as part of probabilistic sampling and through specific studies within DWQ and with our water quality protection partners. Eight sites located on the Jordan River were assessed using O/E scores for the 2012-2014 IR. All sites scored poorly except one located near 9000 South. The EPA requires independent application of biological- and chemical-specific assessment approaches. O/E is not derived comparing a site to a particular reference site but through a summation of Pc derived through taxa presence/environmental gradient combinations of reference sites. The model building methodology is explained in the Biological Assessments of Rivers and Streams section of Chapter 2 of the report.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
F/G	13	303(d)	Conclusion: 1. Assessments are a simplified tool to aid managers in their decisions, nothing more. Assessment methods should not be used to monitor water quality and their results should not be interpreted as scientific evidence. As~essments should also not be considered as a substitute for good science. 2. Results of assessments that rely on very small sample sizes should be interpreted with extreme caution. Many streams listed as 'not- supporting' should likely be relisted as Category 3, insufficient data and others should be listed as 4C. 3. Biological integrity and ecosystem function, as defined in the Clean Water Act, can not justifiably be summarized into one score because these concepts are extremely complex, there is prolific natural variability, and error rates associated with sample collection, sample processing, and final score calculation may not simply be additive but are likely compounded with every step in the assessment procedure. UTDWQ should not claim that 0 /E is a measure of change in biodiversity (taxa richness) because other factors including those discussed in these comments, likely effect scores. 0 /E scores should be interpreted with extreme caution and not used as a primary tool in assessments.	The O/E biological assessment method is commonly used across the United States, Europe, and Australia to assess the quality of their waters. DWQ is required to use quantitative biological assessment methods that have "documented methods that have been subject to technical review and produce consistent, objective and repeatable results that account for methodological uncertainty and natural environmental variability." Alternative biological assessment methods would require the same level of technical review and documentation that has been completed for the currently employed methods, which have been in use in Utah for nearly 10 years. In Utah, the method has been vetted with diverse stakeholders during water quality standards development, and DWQ encourages you to participate in these workgroups.
H	1	303(d)	In Chapter 5 of the draft 2014 IR, UDWQ proposes to include five sites in Arches, Canyonlands, and Natural Bridges on the 303(d) list of impaired waters (Table 1). All of these sites are associated with small, groundwater-fed systems that occur in dry wash settings. In all cases, the extent of perennial surface water is less than one mile in length, and surface water often is stagnant for much of the year. Their limited spatial extents, isolated hydrologic settings, and exposure to environmental extremes naturally result in water quality conditions that are highly variable and not reasonably assessed on the basis of water quality standards designed for flowing streams and rivers. In particular, evaporation of groundwater-fed pools lacking surface water inputs would be expected to naturally elevate total dissolved solids (TDS) and dissolved metal concentrations, as well as naturally increase water temperatures and decrease dissolved oxygen concentrations. After considering these and other factors, in 2010 NPS SEUG began to shift the emphasis of its monitoring efforts for small spring-fed systems from water quality to spring flow. We ask that UDWQ consider removing these sites from the 303(d) list in the final 2014IR. Further information about these sites can be found in several NPS reports that are available online.	DWQ agrees to the evaluation provided by the commenter and has made requisite changes in Chapter 5. DWQ staff will work with NPS staff to design and implement defensible monitoring plans that are comparable to those of DWQ to ensure that they are appropriately sampling their rivers and streams.
H	2	303(d)	In Chapter 2 (p. 21), the draft 20 14 IR indicates that "For each parameter, if >=10 samples are available for a monitoring location within the most recent 5-years, then the AU is considered to be supporting its designated use(s) if <10% of the samples exceed the numeric criterion .. .. In circumstances where insufficient observations exist in the 5 year dataset to make a determination, l 0 years of data is evaluated following the same assessment rule." Based on assessment datasets provided by the UDWQ, this rule does not appear to have been correctly applied for one proposed listing parameter at two sites in Canyonlands (Table 2). In both cases, exceedance values for the parameter are less than 10 percent. Based on the assessment rule, the assessment units would be considered to be meeting their designated uses and therefore should not be included on the 303(d) list.	DWQ does not apply a percent exceedance method for determining attainment of water quality standards for toxic parameters such as Cd and Fe, as the commenter is suggesting. The assessment is performed as outlined in Chapter 2, which states that a site is not meeting its designated uses if >2 exceedance of the applicable standards occur on a sample size of four or more samples. In the case of Salt Creek, there are no impairments indicated. However, two of the sites indicated in the previous comment are also included in the same AU, and the final determination is Category 3A requiring further study. In the case of the Colorado River AB Cnfl / Green R, there is no impairment for Fe, as per DWQ methods; however, this site is impaired for selenium.
H	3	303(d)	In three cases (Table 3), UDWQ appears to have based 303(d) assessments of the aluminum (Al) standard on 10 years of data even though sufficient data are available for the recent 5-year period. In addition, the Al standard is pH-dependent. Although NPS does not regularly collect pH at these sites, a review of existing data indicates that pH within the Green and Colorado Rivers is regularly above 7.00. Accordingly, the Al standard should be assessed on the basis of the acute 750 mg/L criterion rather than the chronic 87 mg/L criterion that appears to have been used by UDWQ in preparing the draft 2014 IR. For the chronic criterion, the 5-year exceedance rate was less than 10 percent for all three sites, and no exceedances of the acute criterion occurred at any of the listed sites in the last 10 years. Therefore these sites should not be included on the 303(d) list on the basis of the Al standard.	During the public comment process, DWQ made an adjustment to the assessment methods for aluminum. In the draft report, the default value of 87 mg/l was used if there was no accompanying hardness or ph. Given the conditions that prevail in Utah, it was determined to be more appropriate to default to the 750 mg/l value in these situations. These sites are no longer exceeding the criterion in the revised assessment.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
I	1	GSL	Chapter 2, page4; chapter 2 page 5, table 1; chapter 3, page 3; Chapter 7, page 2: Great Salt Lake is a Unique Ecosystem and requires its own unique set of regulatory criteria. We fully agree that the Great Salt Lake is a highly unique ecosystem in which regulatory decisions need to be tailored to the particular biological, ecological and limnological characteristics of the lake and its various bays. It is necessary for resource regulators to take into account the unique aspects of GSL in terms of examining and evaluating water quality issues. Clearly GSL differs from both fresh water and marine systems in nutrient cycling, response to toxins, threshold levels for impairment, and in individual and population scale responses to environmental change or perturbation. DWQ has done an admirable job of acknowledging the unique aspects of GSL and outlining the process through which the features of GSL will be documented and evaluated. We support DWQ in its comment that: "Numeric criteria that are broadly applied to other water bodies are generally not applicable to the lake because of its unique saline ecology, biogeochemistry, and hydrology." Clearly GSL and its various bays require site-specific assessments for establishing water quality criteria.	DWQ thanks you for the comment and appreciates the support of our efforts to develop site-specific standards for Great Salt Lake.
I	2	GSL	Chapter 7, page 3 and 17: "bay-by-bay assessment of GSL" and the need to recognize the interconnectivity of the bays. We support the bay-by-bay approach identified by DWQ on page 3. On page 17 the linkages between the bays are briefly discussed. We agree with DWQ that the nutrient loads and limitations differ substantially among the bays. We also fully agree that the linkages between the bays need to be further studied and should always be taken into account in the management of each bay and in the overall assessment of the health of GSL. We believe more detailed research needs to be done on the linkages between the bays and the significance that this may have in the overall integrity of the GSL ecosystem. The characteristics of each bay need to be assessed, yet management decisions should take into account the influence of each bay on the entire GSL ecosystem viewed as an entire body of water. This is particularly true of nutrients—transient elevations in nutrient levels may be desirable in one bay (for example Farmington Bay) in order to ensure that other, larger bays, such as Gilbert Bay and Gunnison Bay are not severely depleted of nutrients. We are concerned that Gilbert Bay is nitrogen limited and that Farmington Bay serves as an important source of nitrogen and other nutrients. A strong research effort needs to be done to fully assess these linkages and to determine the optimal way to improve nutrient conditions in Gilbert Bay while simultaneously addressing the impacts of transient elevated nutrient levels that are observed in Farmington Bay.	DWQ thanks you for the comment. DWQ concurs that nutrient cycling within the lake, including connectivity between bays, deserves additional study. DWQ plans to conduct research related to these issues in the future as part of the overall Great Salt Lake strategy.
I	3	GSL	Chapter 7, page 7: Methyl mercury concentrations in excess of total mercury. GSLBSC strongly supports monitoring programs in Gilbert Bay, Farmington Bay, Ogden Bay, Bear River Bay, and to a lesser extent, in Gunnison Bay. The mercury data discussed on page 7 is somewhat disconcerting as it is reported that methyl mercury levels exceeded total mercury. GSLBSC supports the rejection of such samples as indicated in the report. Some discussion of how this happens and how such spurious results can be avoided would be helpful. GSLBSC would appreciate a comment or list of the most trusted analytical laboratories for high saline samples.	DWQ informed Brooks Rand Labs that the methylmercury concentrations were greater than the total mercury concentrations without qualification for the six samples in 2012. The laboratory reviewed the samples and reported that everything was done correctly. The laboratory was unable to reanalyze the samples to further rectify the results so the results were rejected by DWQ. This situation reaffirms that quality assurance and quality control of data is critical. DWQ will continue the data validation process per the requirements of the Great Salt Lake Baseline Sampling Quality Assurance Program Plan and will report the findings every 2 years.
I	4	GSL	Chapter 2, page 5, table 1; Chapter 7, page 5, section "Background and Purpose": Brine shrimp needs to be recognized as a "Species of Protected Aquatic Wildlife". In table 1 on page 5 of chapter 2, Great Salt Lake, Gilbert Bay's designated use is described as follows: "Protected for infrequent primary and secondary contact recreation, waterfowl, shore birds and other water oriented wild life including their necessary food chain". In the section "background and Purpose" of chapter 7 (page 5) it states that "Brine shrimp tissue samples are collected to evaluate dietary exposure to birds" Both chapters fail to account for the very important fact that brine shrimp from Great Salt Lake are considered protected aquatic wildlife and, as such, are tightly monitored and controlled by the Utah Division of Wildlife Resources. Administrative Rule R657-52-11 specifically identifies brine shrimp as a "Species of Protected Aquatic Wildlife". Brine shrimp need to be recognized as fundamentally essential to the ecosystem of the Great Salt Lake and that their protection be based on the fact that they are already designated as protected aquatic wildlife, in addition to, and independent of, their critical ecological functions within the food chain.	DWQ recognizes that brine shrimp are the keystone species of Great Salt Lake. Currently, brine shrimp are protected aquatic wildlife under the designated use of "…waterfowl, shore birds and other water-oriented wildlife including their necessary food chain." (R317-2-6). DWQ supports the current general description of brine shrimp as part of the necessary food chain as opposed to specifying each species that serve as dietary food items for birds. From DWQ's perspective, the general category is all inclusive and avoids the potential of inadvertently omitting any important dietary items. However, should the commenter wish to pursue this further, we recommend that it be brought to DWQ's Water Quality Standards Workgroup.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
I	5	GSL	Chapter 7, Page 8: Speculation on the exposure risk of brine shrimp to DBL contaminants as a result of mixing. In Paragraph 2, there is a brief discussion of the deep brine layer (DBL) and speculation about mixing between the DBL and the epilimnion and consequent exposure of brine shrimp to contaminants in the DBL. This type of comment is unnecessary and misleading because the monitoring programs are measuring contaminants and nutrients in the epilimnion—therefore there is no need to speculate on contributions from the DBL to the epilimnion because it has been directly measured. Additionally, the reference to Belovsky et al., 2011 as a definitive comment on mixing between the DBL and the epilimnion is exaggerated; their study did not conduct detailed assessments of the hydrochemical linkages between these two layers. While they did conduct some preliminary work on the topic, there remains a great deal of highly sophisticated work that needs to be done to state anything conclusive about the chemical exchanges between these layers.	Further research is needed to assess the mechanisms that could transfer methylmercury from the deep brine layer to the epilimnion where aquatic organisms reside. The monitoring data only indicate concentrations at a given time but do not speak to cycling including transfer mechanisms between the deep brine layer and the epilimnion. DWQ will continue to sample methylmercury and total mercury concentrations in both layers to assess biotic exposure.
I	6	GSL	Chapter 7, page 11: Challenges pertaining to the use of existing water quality criteria for fresh or marine systems when studying GSL. The discussion on Page 11 reveals the challenges of using existing water quality criteria for fresh or marine systems. We applaud the comments of DWQ in which it is recognized that these existing criteria are really only useful in the framework of a benchmark. While this is useful it remains absolutely necessary to establish site specific criteria for GSL.	DWQ thanks you for the comment and appreciates the support of our efforts to develop site-specific standards for Great Salt Lake.
I	7	GSL	Chapter 7, page 11-12: Comments regarding mercury contamination in GSL. We agree with the comments on page 12 with regard to mercury contamination in the biota and water of the GSL—current results do not support the earlier dramatic reports of widespread and highly elevated levels of mercury in the GSL ecosystem. Total mercury levels are far below present EPA regulations of 940 ng/L for both the epilimnion and the DBL, whereas mercury levels in much of the biota of the lake levels are below actionable thresholds.	DWQ thanks you for the comment.
I	8	GSL	Chapter 7, page 14: References to concentrations. In the table and discussion on page 14 please check all references to concentrations. It looks as though 0.05 was incorrectly reported as 0.5. Also, some discussion about detection limits would be helpful.	All units, thresholds, and detection limits have been reviewed and corrected as necessary in the final IR.
I	9	GSL	Chapter 7, page 18: Important clarification made regarding Carlson's Trophic Index. On page 18 a very important clarification with regard to Carlson's Trophic Index was made: DWQ correctly stated that the chlorophyll levels associated with the index are an assessment of productivity and not water quality per se. This is a very important point and must be kept at the forefront of discussions about chlorophyll levels, primary productivity and the associated water quality assessments linked to such data. GSL is highly unique and the biota is characterized by boom and bust cycles that at certain times can be quite elevated but such cycles are all part of a "typical" biological response to environmental conditions and should not be viewed as an adverse water quality impact without taking into account the long-term temporal and spatial patterns.	DWQ thanks you for the comment. DWQ will continue to pursue appropriate indicators of health for Great Salt Lake, including those related to trophic state.
I	10	GSL	Chapter 7, page 20: "ASSESSMENTS AND DATA GAPS" Suggestions regarding data gaps. With regard to data gaps the GSLBSC suggests that lake volume, bay inter-connectivity and its influence on water quality outcomes should be taken into consideration with all water quality criteria evaluations. The volume, circulation, and linkage across bays will have a huge influence on many of these measurements and in defining harm.	DWQ thanks you for the comment. Lake volume, bay inter-connectivity, and circulation have been added to the data gaps section of the final IR.
J	1	Wetlands	Page 3, paragraph 1, line 5 – "These wetlands…": It would probably be a good idea to make a stronger statement of the importance of habitat function here particularly in light of your charge to assess support of beneficial use as currently defined.	Thank you for your comment. For the purposes of the IR, the description of habitat function is sufficient. Additional detail on habitat function is available in DWQ's other wetland-related reports. References to these reports are included in Chapter 4.
J	2	Wetlands	Page 3, paragraph 2: Since the initial focus of the State's MMI is on IW's it would make sense to develop more background on the ecological contribution of IW's, which in turn would support the primary emphasis on them. As written, they amount to wetlands that are constructed for hunting waterfowl and no mention is made of any ecosystem services they provide. Although managed, they too provide the many wetland ecosystem services while carrying the burden of treating incoming waters and assimilating ecologically detrimental pollutants as waters drain toward the	Thank you for your comment. DWQ recognizes the value of applying the concept of ecosystem services to Utah's waters. The concept is a central principal in DWQ's ongoing work to develop wetland-specific use classes. However, discussion of this concept in the IR is unnecessary and premature.

Letter	Comment #	Section	Comment	Response
			Lake.	
J	3	Wetlands	Page 7, paragraph 3: This short list of metrics runs the risk of highly over-simplifying the complexity of biological response even in these managed systems. Our work shows the need for more detailed metrics and tighter sampling windows to capture IW's at risk of degradation (Hoven et al. 2014).	Thank you for your comment. The text briefly describing DWQs approach to wetland assessment has been revised to reflect the list of metrics as potential indicators. As presented, the text describes three major groups of biological response indicators. These are the major groups identified from the 2009 (DWQ) report (as well as Miller and Hoven, 2007). Subsequent work (WSpur projects; FBJRWC 2010 compendium) has not yet shown the need for additional indicators or that specific indicators should be excluded. This section of the work describes what was done; adding more potential complexity to the Methods would not accurately describe how data evaluation and analysis were carried out. DWQ could not locate a publically available copy of the referenced document: "Hoven et al. 2014." DWQ would be pleased to examine this work and incorporate those results into our understanding as part of our overall wetland program.
J	4	Wetlands	Page 7, Table 1, Cultural Eutrophication metrics: Please define acceptable or unacceptable thresholds used, otherwise they are just potential stressors.	For the purposes of this chapter, eutrophication metrics are used as indicators of potential stress.
J	5	Wetlands	Page 7, Table 1, Physical Habitat Degradation: Water depth is specifically related to management objectives. It may be a stressor, however, it is important to note that it is generally not due to natural fluctuations but rather management prescriptions. How (and why) the wetland is managed could covary with other stressors and should be considered. Also, why is size of a wetland considered a metric? Size can be used as a categorical descriptor, but is not a metric. What time of day was temperature recorded? Was the temperature data normalized with depth? What was the acceptable percentage or coverage of emergent marsh?	Thank you for your comment. Table 1 represents a partial list of potential stressors. As such, there are no known benchmarks for this wetland class (except for toxic constituents with numeric criteria that apply to aquatic life uses); Table 1 lists some aspects of wetland stress that DWQ is currently investigating. Further evaluation of wetland stress and response patterns (if any) will be presented in the 2016 IR as well as other, project-specific reports. Overtime, loss of wetland area could be an indicator of system-wide physical habitat degradation.
J	6	Wetlands	Page 8, Table 1, Toxic Constituent metrics: Please define acceptable or unacceptable thresholds used, otherwise they are just potential stressors.	For the purposes of this chapter, toxic constituent metrics are used as indicators of potential stress.
J	7	Wetlands	Page 8, paragraph 3: State makes assumption that 25 - 100 cm is optimal based on what data and what part of the growing season? May be excluding high functioning, quality habitat that exists at depths more shallow than 25cm. Sampling windows span too much time. Our reports show major changes between months (e.g., June vs. July; August vs. Sept.) (Hoven et al. 2011, 2014).	Thank you for your comment. This text has been revised to more accurately reflect (1) the project goals and objectives, based on active wetland management area plans as well as minimum depth requirements for routine water sample collection and (2) the breadth and timing of the sampling Index Periods. These are described in more detail in Appendix 1.
J	8	Wetlands	Page 8, paragraph 4: What extent of surface mat qualifies as indicative of a degraded system? Our research shows that the development of extensive surface mats, which we define as greater than 75% cover, is rare and short term. What if the survey window misses the timing of surface mat development? Do you have statistical evidence that indicates strong correlation between surface mats and a metric that reflects biological response (presumable percent cover SAV)? Our research suggests otherwise (Hoven et al. 2011, 2014).	Thank you for your comment. As described in the section on Surface Mats in the "Indicators of Wetland Condition" section, all benchmarks for biological responses were determined empirically, as upper and lower quartiles of the data. In this analysis, DWQ is examining surface mats as an indicator of biological response, not stress. As such, a strong correlation between surface mats and other responses would require examination of the degree of association among response variables. DWQ would be pleased to examine the references cited in the comment; however publically available copies of those works could not be located.
J	9	Wetlands	What macroinvertebrate community composition (although I believe you mean assemblage) connotes ecological health in a highly managed system?	This is a good question, and one that DWQ hopes will be addressed for Great Salt Lake ponded wetlands through continued monitoring efforts and directed studies that target potential reference standard sites. The term "community" is commonly used to describe a group of distinct organisms that co-occur in the same area and over similar time period (sensu Fauth et al., 1996 [Am. Nat. Vol 147, pp 282–286]. This definition has been added to the chapter as a footnote. DWQ uses this working definition and considers "community" to be more easily understandable to the public than "assemblage."
J	10	Wetlands	Page 9, paragraph 1: How do you make the connection between the listed metrics (biological indicators) and land use when the IW water source is affected by practices far upstream? How is site-specific impact from adjacent land use determined and included in the MMI without being too generalized, i.e., subwatershed level?	DWQ recognizes the importance of the questions posed by the commenter. However, they extend beyond the scope of the IR chapter. Projects have been initiated in collaboration with program partners to better understand these and similar issues, including an improved dataset of flowpaths among Great Salt Lake wetlands. For the case at hand, a cursory land use assessment could be implemented (and is being implemented) where potential reference standard sites would be located in areas that lack urban and industrial development, in comparison to the Great Salt Lake wetlands along the Wasatch Front.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
J	11	Wetlands	Page 9, paragraph 2: Willard Spur is not an IW and not managed. Setting water quality standards based on a natural system that receives high flushing flows during spring runoff is probably not comparable or relevant.	Thank you for your comment. Willard Spur is occasionally impounded, but it is not currently managed with any hydrologic goals. This paragraph describes how data and project goals for several wetland classes associated with Great Salt Lake are intertwined and contribute to DWQ water quality goals.
J	12	Wetlands	Page 12, paragraph 2: SAV cover and condition could be biased from waterfowl grazing - they arrive mid-August and reach a maximum density by mid-September. If different ponds were sampled one month apart, the grazing pressure could be unequal.	While the dates of waterfowl arrival vary among species and from year to year, it is true that unknown elements could affect the results we obtained from our sampling efforts. DWQ has added a note to the chapter acknowledging variation in natural ecosystems.
J	13	Wetlands	Page 14, paragraph 1: What data support greater than 25% surface mats as indicative of poor condition? What evidence do you have that supports the notion that an impounded wetland must essentially not have any algae (less that 1%) to qualify as good condition? Algae are a natural component of aquatic systems and important contributors to a balanced ecosystem. Extensive, healthy beds of SAV can and do coexist with well over 50% cover of surface mat. It is quite possible there are other stressors that have detrimental impacts on the ecological condition of the ponds. If there are negative impacts on the condition of waterfowl habitat and hunting related to algal mats, the data should support it. Our data show that surface mat cover and SAV cover are not correlated (Hoven at al. 2011, 2014). There is no indication of negative impacts from surface mat cover in your summary other than a broadly applied assumption with no statistical evidence.	Please see the description of assessment methods in the "Development of an Impounded Wetland Assessment" section. Benchmarks for GOOD and POOR wetland condition classes were determined empirically from the data as the upper and lower quartiles. Revisions to assessment methods are beyond the scope of the IR but could be considered as part of ongoing wetland program development. DWQ would be pleased to consider data and rationale supporting alternative metrics of ecological condition.
J	14	Wetlands	Page 14, Figure 4: Not many ponds show extensive surface mats (above 60%, our conservative definition, Hoven et al. 2014). Surface mat cover is probably not a good predictor of ecological condition.	The selection of indicators for ecological health is not dependent on the probability of occurrence but rather the functional importance in the system. DWQ would be pleased to consider data and rationale supporting the commenter's assertion that extensive surface mats are not relevant to ecological condition.
J	15	Wetlands	Page 15, paragraph 1: Your PMI metric shows that phytophillic macroinvertebrates are associating with sites that have SAV as they should. What is the ecological significance relative to wetland condition that differs from that implied by SAV percent cover? Isn't this a covariable?	An apparent association between measurements of PMI and SAV Index scores is illustrated in Figure 4-6. The relationship is far from tight, suggesting that there is additional variation that may need to be explored before determining whether one of these metrics is redundant.
J	16	Wetlands	Page 16, paragraph 2: Lack of SAV index sensitivity is possibly due to the broad sampling windows. IW's that host a high number of waterfowl during the first half of September receive high grazing pressure from the waterfowl and could bias your results and interpretation. Best to consider an alternate sampling window as recommended by Hoven at all. 2014. Also, the State should consider incorporating additional metrics that have been shown to reflect good biological response to stressors (Hoven et al. 2011; Hoven et al. 2014).	It is difficult to ascribe causes to a "lack of sensitivity," generally. The purpose of this report is to describe our biological response measures for IWs among three Great Salt Lake subwatersheds; inferences into the mechanisms of specific responses are beyond the scope of this chapter. DWQs wetland assessment goals involve the monitoring of a large number of sites; as such, our measurements must necessarily be simple yet robust, so alternative measures supported by peer-reviewed literature are welcomed. DWQ could not locate a publically available copy of the referenced document, "Hoven et al., 2014."
J	17	Wetlands	Were the "good" sites measured for DO and pH at the same time of day as the "poor" sites? Both measures vary highly throughout a 24hr cycle in healthy and degraded wetlands.	Yes. Briefly, a field crew could typically sample two sites in a day. Depending on travel time, samples were collected from the first site in late morning (10:00–11:00 am), whereas samples from the second site were collected in early afternoon (1:30—3:00 pm). Both DO and pH measurements would be expected to increase from dawn to 5:00 pm or 6:00 pm. For these reasons, variability is less than if sampling could occur during an entire 24-hour period. The key signal measurement available for this wetland type is whether respiration exceeds photosynthesis during the main portion of the "day" (sunrise ranges from 6:00 am to 7:15 am in mid-September).
J	18	Wetlands	Page 17, paragraph 1: This observation illustrates our findings of no significant correlation between degradation of a site as reflected by SAV cover and surface mat cover (Hoven et al. 2010; Hoven et al. 2014). Again, why is this considered an important biological response and included in the State's MMI?	Thank you for your comment. DWQ is working through the proposed indicators from previous work (Miller and Hoven, 2007; DWQ, 2009). A strong correlation between SAV and surface mat cover would require an examination of the degree of association between response variables, to check for potential redundancy. This effort is beyond the scope of the IR.
J	19	Wetlands	Page 17, paragraph 2: DO and pH data are practically irrelevant when measured at different times of day and used for comparison among sites.	Briefly, a field crew could typically sample two sites in a day. Depending on travel time, samples were collected from the first site in late morning (10:00–11:00 am), whereas samples from the second site were collected in early afternoon (1:30—3:00 pm). Both DO and pH measurements would be expected to increase from dawn to 5:00 pm or 6:00 pm. For these reasons, variability is less than if sampling could occur during an entire 24-hour period. The key signal measurement available for this wetland type is whether respiration exceeds photosynthesis during the main portion of the "day" (sunrise ranges from 6:00 am to 7:15 am in mid-September).

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
J	20	Wetlands	Page 17, paragraph 3: What is the value in re-scaling and combining the metrics into one final score? How does it provide guidance toward the development of specific water quality goals when the only useful metric is either SAV cover or PMI? I suppose it warrants further investigation, however, our comparison of State selected metrics versus use of additional vegetative metrics show a likelihood of overlooking some potentially degraded sites if only State selected vegetative metrics are used (Hoven et al. 2014).	Thank you for your comment. In order to provide an assessment of the "beneficial use support for waters of the state" (of which wetlands are one) (UAC R317-2-7.1), it is necessary and appropriate for DWQ to develop and use tools that integrate and summarize results from our monitoring and assessment efforts. The MMI approach is one of many statistical tools used to compare the relative "health" (or condition) of ecosystems, including wetlands (Karr and Chu, 2007; Biological Monitoring and Assessment: Using Multimetric Indexes Effectively. EPA 235-R97-001. Seattle, Univ. Washington). DWQ welcomes the incorporation of data from partners and collaborators into additional tools and modified indicators. For more information please see the section called "Submission of Data from Outside Sources" in Chapter 2. DWQ could not locate a publically available copy of the referenced document, "Hoven et al. 2014."
J	21	Wetlands	Page 18, paragraph 2: There is no indication that surface mat cover is a key measure of biological integrity. What data does the State use to back the inclusion of this metric?	Thank you for your comment. The rationale and goals behind biological indicator selection is described in the "Wetland Health and Assessment Methods" and "Indicators of Wetland Condition" sections.
J	22	Wetlands	Page 18, parapgraph 3: What frame of reference does the State use for DO and pH that are collected at different times of day as indicators of good ecological health?	Field measures of pH and DO were not used as biological response indicators. While there are numeric criteria for aquatic life use (3B, 3D) for these parameters, the criteria do not apply to IWs, by rule (footnote 2a to Table 2.14.2). Thank you for your comment.
J	23	Wetlands	Page 19, Figure 8: Not very good separation of 95% confidence intervals. The metrics used are oversimplified and probably inadequate as indicators of biological response. Our comparison of State-applied metrics versus additional vegetative metrics shows a lack of sensitivity for using percent cover SAV, percent cover surface mat and SAV condition index as the sole assessment of SAV condition from a vegetative perspective (Hoven et al. 2014). We show that in addition to using a more appropriate sampling window and need for seasonal reference for natural changes in development and establishment of a SAV community, other metrics related to physiological queues from the SAV are important to include. Those are branch density and productivity of drupelets and tubers. Other biological responses related to the environmental condition of a site such as percent cover of algae on SAV and other coating on the leaves formed by biofilm, diatoms and / or sediment, make a more direct link between stressors (elevated metals and nutrients) and negative impacts on biological response (Hoven et al. 2011; Hoven et al. 2014).	DWQ could not locate a publically available copy of the referenced document, "Hoven et al., 2014." DWQ would greatly appreciate a link for the cited reference (Hoven et al., 2014) to examine these patterns as described by the commenter and to help DWQ identify appropriate monitoring metrics for this and other wetland classes. DWQ considers wetlands broadly. SAV health represents one of several potential metrics of condition. Thank you for your comment.
J	24	Wetlands	Page 19, paragraph 1: Stream assessment protocols are not applicable in wetland settings. Impounded wetlands do not operate similarly to flowing streams on a biogeochemical, biological, or ecological stand point. What basis does the State provide that they can use stream assessment protocols in IW's?	DWQ has not applied stream assessment protocols to IW. The discussion in the referenced paragraph relates to the need to assess the health of streams in the contributing watersheds to Great Salt Lake and IWs.
J	25	Wetlands	Appendix 2: Please provide units in your graphs.	Thank you for your comment. The figure captions have been revised.
K	1	303(d)	In Chapter 2 of the Integrated Report, it is stated that the State utilizes the FDA's mercury screening value for assessment of water quality and potential listing of a waterway on the State's 303(d) list as compared to the EPA's mercury screening value for issuing fish consumption advisories, which can be used to list a waterway on the 303(d) list. We believe this use of the FDA's mercury screening level leads to confusion and inconsistency in assessing impaired waterways. We further believe that the EPA mercury screening level should be used for both the issuance of fish consumption advisories and assessment of impairment and listing of waterways in the 303(d) list for consistency and providing a more protective standard for the public.	This comment is beyond the scope of the IR because it recommends a change to the Water Quality Standards. Utah's Water Quality Standards do not currently include a fish-tissue criterion for mercury. Suggested changes to the Standards can be submitted during the next Triennial Review (UAC R317-2-1C) in 2017. The potential adoption of a tissue-based methylmercury criterion is currently being evaluated by the Water Quality Standards Workgroup for the 2014 Triennial Review.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
K	2	303(d)	Another aspect of the mercury testing and analysis that we would like to comment on that is not expressly stated in the Integrated Report is the methodology used for collecting and analyzing fish tissue samples for contaminants. Currently, the State collects fish samples through an opportunistic partnership arrangement with the Division of Wildlife Resources, which limits and skews the samples collected for analysis. We recommend that the Division improve the fish tissue collection protocol by collecting predetermined sample sizes of specific age and size classes of carefully targeted fish species, rather than the uncertain samples sizes of unspecified age and size classes of random fish species currently collected from DWR's collection efforts. By focusing on carefully planned sample sizes of specified age and size classes of targeted fish species, the Division can improve its data analysis and provide better assessment of water quality and protection of public health. The Division should revise its sampling rules and methodology to more closely resemble the rule for collecting water column samples, especially with respect to minimum sample sizes. Also, we recommend that the Division not use averaged or blended samples to provide more specific data on bioaccumulation, which is essential with mercury contamination. Finally, we recommend the Division amend its mathematical analysis methodologies to provide better data regarding contamination by age and size class of fish.	Thank you for the suggestions on improving Utah's mercury monitoring program, but this comment was presented as a brief request and lacks sufficient detail on which to base an analytical change in methodology. The current assessment methodology for issuing fish consumption advisories was developed jointly by the Utah Departments of Health, Environmental Quality, and Natural Resources. This methodology has also been vetted with the Utah Mercury Workgroup. The approach is scientifically defensible and protective of human health by rigorous control of both false positive and false negative decision errors (erroneously concluding that a consumption advisory should be issued and erroneously concluding that a consumption advisory should not be issued) within the constraints of the available resources.
K	3	303(d)	Our second comment involves the State's enforcement of the narrative water quality standards. Currently, the State declares that it is unlawful for anyone to "discharge or place any waste or other substance in such a way as will be or may become offensive such as unnatural deposits, floating debris, oil, scum or other nuisances such as color, odor or taste; or cause conditions which produce undesirable aquatic life or which produce objectionable tastes in edible aquatic organisms …" (Utah Administrative Code, R317-2-7.2). For several years, our organization and others have complained to the Division that it is not enforcing the narrative standards, especially with respect to the discharge of floating debris and off-flavor in fish resulting from the discharge of pollutants. This lack of enforcement is particularly evident along the Jordan River, where floating debris has remained a chronic problem, and where fish are not edible due to the off-flavor resulting from chemicals discharged into the river. We believe the Division is in violation of its own rule and the Clean Water Act by not establishing criteria for these parameters and not enforcing the law. We insist that the Division take all necessary actions and steps to address these problems and comply with the law.	DWQ disagrees that the Narrative Standards are being violated in the Jordan River. No credible data or analyses were provided to support that 1) the existence of discharges of floating debris or off flavors in fish and 2) if the floating debris and off-flavors exist, that the Narrative Standards are not being met. As such, the descriptions provided by the commenter are of a general nature and do not provide sufficient information to use in an impairment determination.
L	1	303(d)	The 2014 Integrated Report, Chapter 5 – 303(d) Rivers and Streams includes a number of new listings in Rich County. The new listings include North Fork of Sage Creek, Sage Creek, and Duck Creek. An unspecified cooperator established new sites and submitted their sampling data (identified as "Public data" in the report) to Utah Division of Water Quality – Division of Water Quality (UDEQ-DWQ). The collected data identifies exceedances in E. coli standards for the above streams. BLM has concerns about the quality of the public data being incorporated into the current Integrated Report for identifying Impaired Waters. We understand the importance incorporating all pertinent data into the Integrated Report, but assert that analysis is only as good as the data it is based on. Below are our comments and questions for your consideration regarding the data collected by public data contributors: Please include the Sample Analysis Plan(s) submitted to UDEQ-DWQ that outlines adequate monitoring protocols, QA/QC methods, site locations, and other pertinent information used in sampling as described in the2014 Integrated Report Chapter 2 – Methods. Identify how UDEQ verified protocols were followed, methods were appropriate, sample locations were appropriate, and approved sampling equipment was used. Were sample and trip blanks taken to identify other possible sources of contamination? Were samples repeated to ensure accuracy? Explain how the samples were analyzed including hold times for samples. If there is any uncertainty of the methods, data collection, analysis, or QA/QC process, we recommend the data not be incorporated into the Integrated Report. As stated in the Integrated Report, all physical, biological, and chemical water quality data has a utility, but data also has limitations. If UDEQ-DWQ does decide to include data where the quality of a cooperator's or other water sampling partner's data is unkown or differs in methods or QA/QC, we recommend that the data be identified so users of the report understand the differences in data collected and analysis. We recommend data and analysis to determine which waters are included on the 303(d) Impaired Status list only be of the highest caliber. The Utah BLM/ Salt Lake Field Office and	DWQ thoroughly reviewed the Sampling and Analysis Plan provided by the group collecting the referenced data. The plan included quality assurance and quality control procedures found to be compatible with our own monitoring program, including well documented sampling procedures and methods.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
			UDEQ-DWQ have a strong and mutually beneficial relationship through the cooperative water quality monitoring program. We work hard to ensure that we meet UDEQ-DWQ requirements so that data we supply is of high quality and meets the needs of both agencies' missions. We ask that all other public data contributors sampling on public lands managed by BLM are subject to the same requirements. This is particularly important for BLM as our land management activities and obligations can be substantially affected by waters with impaired status.	
M	1	303(d)	SUWA is concerned that DWQ did not sufficiently pursue all reasonably available sources of data in compiling the IR. The Clean Water Act requires the State of Utah to submit its 303d list to EPA by April 1 of every even numbered year for its review and approval or disapproval. 33 U.S.C. § 131(d)(2); 40 C.F.R. §§ 130.7, 130.10, and 131.21. The State of Utah, through DWQ, did not submit a 303(d) in 2012 or 2014 as required. In addition, SUWA's review of DWQ's "Public Notice Archive" indicates that DWQ did not issue a Notice of Data Request in either 2011 or 2012. See http://www.waterquality.utah.gov/PublicNotices/pnarchive2011.htm and http://www.waterquality.utah.gov/PublicNotices/pnarchive2012.htm. Indeed, it was not until 2013 that DWQ issued a "Notice of Data Request for Clean Water Act Section 305(b) State Water Quality Assessment Integrated Report 2014. See http://www.waterquality.utah.gov/PublicNotices/docs/2013/Misc/2014_Public_Notice_Request_For_Data_Website.pdf. The deadline for submitting data for the 2014 IR was March 22, 2013.1. Based on its review of the IR and conversations with DWQ, BLM, and DOGM staff SUWA is concerned that DWQ did not obtain relevant and readily available data for use in the IR. See IR Chapter 2, at 10 ("Whenever possible, the aim of DWQ is to obtain all data and information with sufficient time to compile the information by April of odd years."). For example, though the IR states that "DWQ routinely obtains and analyzes data collected and processed by the United States Geological Survey," id., the IR in fact does not reflect USGS's data for the period of 2010-2013. This is a significant omission.	In response to the lack of a Notice of Data Request in 2011 or 2012: this was due to the fact that DWQ did not issue an IR in the 2012 cycle. Instead, DWQ issued the Notice in March 2012 in preparation for compiling a combined 2012 and 2014 report. With respect to the omission of USGS data in the 2012-2014 IR, DWQ staff made a decision in December 2013, after compiling and reviewing USGS data, that the time and effort to format and evaluate the data for comparability with other data sources would prevent DWQ from completing the IR in a timely fashion as specified on the timeline agreed upon with EPA. DWQ staff is committed to including this data in the 2016 IR. On October 1, 2014, DWQ issued another Notice for Data Request to compile all readily available data and is fully committed to integrating USGS and other readily available data sources in the 2016 IR.
M	2	303(d)	In addition, SUWA believes that in preparing the IR DWQ did not obtain or consider relevant water quality data required to be compiled by oil and gas operators in Uintah, Duchesne, and Carbon Counties. For example, BLM's 2010 West Tavaputs full field development environmental impact statement called for the establishment of five (5) new surface water quality monitoring sites in the Nine Mile Canyon drainage, in addition to the five existing Utah STORET locations. See West Tavaputs full field development record of decision2, Attachment 7 (Long Term Monitoring for Water Resources), at unpaginated 2 (attached hereto). Likewise, BLM's 2012 Greater Natural Buttes Infill environmental impact statement called for the installation of ten (10) new monitoring sites in the Uinta Basin (White River, Bitter Creek, Bitter Creek, Willow Creek, Cottonwood Wash, and Coyote Wash). See Greater Natural Buttes record of decision3, Appendix C (Long-Term Monitoring Plan for Water Resources), at C-7 to -8 (attached hereto). Appendix C also identifies several USGS monitors in the Greater Natural Buttes project area, none of which are referenced in the IR. See id. at C-3 to -4; see also id. at C-5 (map depicting Greater Natural Buttes project area and USGS monitoring locations). It is critically important that before DWQ sends the IR to EPA for review, it assess what data was collected through the 2013 Notice of Data Request and what other information should have been obtained, but was not. This would include insuring that USGS's data is fully incorporated into this IR and not deferred for a later IR. DWQ's refusal to do so would call the accuracy and integrity of the IR into question. For example, waterbodies that DWQ has identified as a "category 3" – insufficient data – may instead be accurately classified as supporting designated uses or not. Likewise a DWQ determination that a certain waterbody is not meeting designated uses but a TMDL is not required ("category 4") may instead be classified as requiring a TMDL. DWQ is already late in submitting the 2012 and 2014 IRs to EPA; it should not hurry to submit these reports without insuring that they are using all readily available data. DWQ should also insure that it has the most current data produced pursuant to various federal agency approved projects, including Greater Natural Buttes and West Tavaputs.	DWQ and BLM regularly cooperate on the collection of water quality data. DWQ did not receive data associated with the long-term monitoring efforts described by the commenter. DWQ will actively pursue additional data from BLM as part of the 2016 IR process. As mentioned in previous comments, in order to meet agreed-upon deadlines with EPA, DWQ will not be able to integrate the USGS data due to the time and effort involved in formatting and comparing to other data sources. In response to the concerns regarding classification, this same analysis can be performed in the next IR cycle, and any missed impairments can be documented and insufficient datasets can be augmented where feasible to perform a complete analysis. DWQ has been working with EPA staff to ensure a defensible IR is developed in a timely fashion in direct response to a Notice of Intent filed against EPA by SUWA citing DWQ's failure to develop Reports in 2012 and 2014 in a timely fashion (DWQ-2014-004785).

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
M	3	303(d)	SUWA is particularly interested in seeing improved water quality in streams within several watersheds, including: Colorado River Southeast, Lower Colorado, Colorado River West, and Uinta. Despite the fact that numerous streams have been identified as not supporting their designated uses in these watersheds, the overwhelming majority of these streams have been identified as low or medium priority for completion of a TMDL. We would appreciate the opportunity to meet with you and your staff and better understand the methodology used to prioritize the preparation of certain TMDLs over others.	We welcome the opportunity to work with all stakeholders in setting TMDL priorities.
N	1	303(d)	In general the report seems very well done. We are happy to see the changes in monitoring and assessment you have made. We are very pleased to be able to access the original source data that was used in making the assessments, as well as all the summary data and tables. The narratives are quite clear. We are concerned that the tables did not include previously listed impaired waters (waters not supporting designated uses) in the table for Category 5 (Not Supporting). We understand this was done to simplify the tables, but that was not made clear in the document. This led to confusion in our review, initially leading us to believe that four stream segments that had been listed in 2008 for benthic macroinvertebrate habitat had been removed, along with another stream segment listed in 2008 for temperature. The first group was confusing because they did not appear on the delisting table, the last because it was actually listed in the table for Category 1 (Supporting). We were able to clarify this by talking with your staff.	DWQ thanks BLM for the feedback on the 2012-2014 IR and apologizes for the confusion. In the final 2012-2014 IR, DWQ has included the previous listing decisions from past assessment reports. These previous listings are AUs where no new data were collected since the 2010 IR (e.g., data collected after 12/31/2008) or submitted to DWQ during the formal request for publicly submitted data. Because no new information was assessed, the previous listing holds. DWQ encourages BLM and other stakeholders and reviewers to continue to provide feedback on how to improve the transparency of the IR and to provide information in a clear way to help streamline the review process.
O	1	GSL	For a broader understanding by stakeholders, we suggest that UDWQ identify the existing data collected through other efforts and by other agencies that could be considered in assessing Great Salt Lake water quality. This information could be included in table form. Reasons for not considering some datasets in this Chapter should be discussed.	For the 2012-2014 IR, DWQ incorporated others' research by reference in Chapter 7. DWQ agrees with the suggestion that we should identify existing data collected through others' efforts listed in table form and will consider this suggestion for the 2016 IR.
O	2	GSL	Regarding the reported data from the 2012 Cavitt paper, we suggest that the Chapter include a description of the issues with samples exceeding holding times/temperatures due to a shipment error. Results impacted by quality control issues should be reported as estimated values due to reliability concerns.	DWQ reviewed the 2012 Cavitt report and as the commenter pointed out, the holding time requirement for the 2012 egg data was not met due to a shipping error. DWQ corrected the reported 2012 egg data in Tables 7-14 and 7-15 by noting that the results are qualified as estimated.
P	1	Wetlands	In addition to SAV, emergent vegetation (e.g. bulrush, cattails, etc.) are extremely important habitat for waterfowl and shorebirds. Metrics that address emergent vegetation need to be included in an MMI. SAV and macroinvertebrate metrics don't reflect emergent vegetation condition.	A landscape-scale metric of emergent wetland cover is included in the impounded wetland dataset as a supplemental indicator. However, these data have not yet shown to be useful in describing wetland condition for this wetland class. In addition, the analyses presented here are based on the previously developed assessment methodology for IWs; evaluation of potential supplemental metrics is beyond the scope of this chapter. Emergent vegetation is not currently considered relevant to the health of shallow ponds for waterfowl use, but potential supplemental metrics that involve the fraction of emergent vegetation associated with pond boundaries are being evaluated by DWQ.
P	2	Wetlands	Invasive species metrics need to be included in the MMI. For example; Phragmites is rampant. The snail Radix auricularia now exists in some of the IW and its impacts are unknown. Of course carp completely alter IW ecosystems wherever they occur.	Evaluation of potential supplemental indicators is beyond the scope of this chapter but will be considered in ongoing wetland program development. DWQ welcomes data and rationale that support additional or alternative indicators of health.
P	3	Wetlands	If IW are completely man-made, then water quality standards can be set to whatever level depending on their ability to meet the beneficial use of "waterfowl, shorebirds , and associated biota".	Water quality standards are developed to be protective of the most sensitive designated use. Changes to standards, use designations, or numeric criteria to less protective levels would require substantial justification and approval by EPA.
P	4	Wetlands	"Probabilistic' surveys are simply randomly selected sites. Why do water quality agencies need to use the term 'probabilistic' other than to sound like they know something about survey design	Thank you for your comment. DWQ uses the term probabilistic instead of randomly selected to be consistent with EPA guidance.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
P	5	Wetlands	The correct term would be 'assemblages' not communities. Communities include other biota such as fish, and plants.	The term "community" is commonly used to describe a group of distinct organisms that co-occur in the same area and over a similar time period (sensu Fauth et al., 1996 [Am. Nat. Vol 147, pages 282–286]. This definition has been added to the chapter as a footnote. DWQ uses this working definition and considers "community" to be more easily understandable for the public than "assemblage."
P	6	Wetlands	Professional Ecologists haven't been using the term 'food chain' for several decades. It is well known that ecological 'links' are extremely complex and not linear (i.e. chain) and for those interested are now called 'food webs' even in simplified ecosystems such as IWs. Food chains as far as I know are now only used in children's introductions to ecology.	DWQ uses the term food chain in the assessment because the aquatic life use designations in Utah Rule use the term "…associated food chains.' Please see Utah Administrative Code R317-2-6, "Standards of Quality for Waters of the State, Use Designations." Changes to use designations require rule change which is outside the scope of the integrated report.
P	7	Wetlands	From my analyses and personal experience and this report, extensive beds of SAV often are associated with high incidence of surface mats. Conditions promoted by extensive beds of SAV could actually be a primary cause of increased surface mats. For example, extensive SAV tends to stagnate water flow which in turn alters other chemical, biological, and physical conditions that are likely favorable to surface mats.	Thank you for your comment.
P	8	Wetlands	Though out the document and appendices the three watersheds are used entirely as descriptors but no map has been provided that outlines these watersheds.	A map of the Great Salt Lake basin including the four primary watersheds has been added to the final version of the chapter (see Figure 4-2).
P	9	Wetlands	If PMIs are strongly associated with SAV then this would be a redundant metric and add extra work throughout the water quality analysis and be an inefficient use of limited DWQ resources.	The PMI metric has been built up over time, based largely on the work of Dr. Larry Gray identifying and describing the taxa from macroinvertebrate samples. As such, the PMI metric aggregates the abundance of taxa expected to be closely associated with healthy SAV growth (whether due to life-cycle requirements, food preferences, or prior observations from other aquatic systems). As described later in the results section, the distribution of values for PMI and SAV cover appear to be associated (see Figure 4-6), as predicted, but a substantial amount of variation remains to be explained.
P	10	Wetlands	Plant associated macroinvertebrates (i.e. phytophilous) may be inappropriate or not the most beneficial metric. It appears the L. Gray based phytophilous taxa from studies conducted in Montreal Canada and those species in GSL IW although somewhat taxonomically similar may not behave ecologically similar to those taxa in Canada. In addition, many taxa that are beneficial to waterfowl and shorebirds occur in more open water (e.g. benthic profundal, basal) including waterboatman (Corixids) and backswimmers (Notonectids)(not to mention larval and juvenile fish that are important to fish eating birds). Also, physa snails occur in very large abundances in SAV and lay their eggs on SAV in GSL IW (personal observation). Work done by Larry Gray is an extremely important starting point for developing methods to monitor water quality in IW but much more research is needed on the life histories and ecologies of these simple macroinvertebrate assemblages. Until then, I do not recommend using this metric.	Thank you for your comment. DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. DWQ welcomes data and rationale that support additional or alternative indicators of health.
P	11	Wetlands	Gyraulus has limited distribution in IWs and reasons for its presence/absence may be due to other factors.	*Gyraulus* sp. was observed in more than 33% of samples. DWQ acknowledges that the causal factors for the presence/absence of species can be complex. This complexity is one reason that biological indicators are useful.
P	12	Wetlands	If the SAV Index score is based on percent cover, I assume 100% cover receives the highest score? An IW that is completely covered with SAV is probably not ecologically 'healthy' or the best beneficial use for birdies. A diversity of habitats is needed for birds and macroinvertebrates. Many bird species avoid thick SAV areas, need emergent vegetation for rearing chicks, open water to navigate, etc. In addition, many macroinvertebrates and fish including those used as bird food resources require more diverse habitats other than 100% SAV. I do not know what an ideal diversity of habitats is, but it probably is not 100% SAV.	Thank you for your comment. DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. DWQ welcomes data and rationale that support additional or alternative indicators of health.
P	13	Wetlands	Unless you re evaluate more closely. It may not be a 1:1 relationship but there are obviously relationships between SAV in early vs. late summer.	Thank you for your comment. DWQ will continue to evaluate the best use of SAV metrics in assessment wetland health. DWQ welcomes data and rationale that support additional or alternative indicators of health.

Letter	Comment #	Section	Comment	Response
P	14	Wetlands	Is it early vs. late summer or summer vs. early autumn?	Consistency in the description of the two index periods has been clarified. Thank you for your comment.
P	15	Wetlands	I am assuming that these CDFs are based on 50 samples. I don't see how many samples were from each subpopulation so roughly about 17 per subpopulation would suggest that percentages such as proportion of data below value of 50% or > 75% could actually only be two or three samples? CDFs are a good way to interpret this kind of data but sample size is so low it may not be useful for this report.	Thank you for your comment. A statement describing the sample sizes for subwatersheds and upper and lower quartiles has been added to the final version of the chapter.
P	16	Wetlands	Jordan River subwatershed or Lower Weber?	Text is correct as written (referring to Figure 4- 5 (left)); sites with GOOD condition occurred in both the Lower Bear-Malad and the Lower Weber subwatersheds.
P	17	Wetlands	The only conclusion I can make from PMI increasing with SAV Index is that PMIs are associated with SAVs and that taxa a priori determined to be 'plant associated' actually were. This relationship doesn't relate to multiple trophic levels what so ever.	While it is true that an association between PMI and SAV cover was expected (and therefore not surprising), there are multiple trophic levels represented by the macroinvertebrates alone, plus SAV, such that this assessment effort does not rely on one trophic level. However, the broader point—that higher PMI values tended to also have greater SAV scores (as total cover or as an index score)—is better shown in Appendix 2, Figures A2-7 or A2-30. The scatterplot in Figure 4-6 illustrates a considerable amount of variation between the PMI and SAV indices that remains to be explained by future work.
P	18	Wetlands	In figure A2-21 water depths have no units. Looks like no measurement were made at < 30 (feet, inches, micrometers?). If samples were collected in shallower water a relationship may have been found ergo salinity	Thank you for your comment. Units have been clarified and made more explicit in the final version. The site selection criteria are described in the Wetland Health and Assessment Methods section that clarify the depth range of IWs of interest.
Q	1	Wetlands	This is just not true, I challenge you to cite the paper or [expert] person that has seen cyanobacterial mats in FB wetlands. We haven't seen this in 10 years in intensive monitoring. Either cite a legitimate reference or remove the sentence. You are trying to excite the reader without cause.	Observations of cyanobacterial mats in Farmington Bay wetlands are described in Miller and Hoven, 2007. However, the context of this paragraph, and the statement that the commenter apparently objects to, is of a description of the history of DWQs Wetlands Program, including the observations that serve as a basis for DWQs programmatic interest in wetlands since 2004.
Q	2	Wetlands	What is a sward?	A sward is an expanse of short grass. A definition has been added to the final version of the chapter. Thank you for your comment.
Q	3	Wetlands	You are ignoring Hoven and Miller's work. Particularly our 2012 report. Is this on purpose?	Thank you for your comment. DWQ would be pleased to evaluate relevant work on Great Salt Lake wetlands, but we require additional information to locate the report referenced by the comment.
Q	4	Wetlands	You need to include the dates of sampling each pond - at least in the appendices. This is a huge variable that you have ignored.	Thank you for your comment. This variable—sampling date—has not been ignored because it is represented by the two index periods (referred to in the text as the "summer" and "early autumn" sampling periods) repeatedly in the chapter.
Q	5	Wetlands	This is not human caused stress - other than constructing the artificial impoundments - rather it appears to be just age (Natural eutrophication see the other comments).	Thank you for your comment. The term "human caused stress" has been removed from the text.
Q	6	Wetlands	None of these are associated solely with cultural eutrophication - Rather, you need to discuss the age and be aware of and discuss the history of the way the ponds are managed (to produce luxuriant growth of SAV) and water availability for flushing (i.e. if not substantially flushed, the OM (including nutrients) will settle and accumulate. This is the source of virtually all of your metrics listed here.	Thank you for your comment. The term "cultural" has been removed from the table. Nutrients and productivity are well recognized as indicators of eutrophication in aquatic systems.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
Q	7	Wetlands	Max and Min water depth are set by the pond managers. Only in rare cases would they not have enough water to set the depth where they want. Also, the relationship of these habitat degradation metrics to the biological metrics is not addressed. Is there a true relationship	Thank you for your comment. DWQ has not yet identified or characterized a potential relationship between biological response metrics and water depths, but these data allow us to ensure that the sampled sites remain within the defined boundaries of the sample frame. As an example, three of 53 sites went dry between the first and second index periods (i.e., water depths were zero), representing a severe case of physical habitat degradation (lack of water).
Q	8	Wetlands	So, why are these listed if you have not done any correlations to the biological indicators?	Thank you for your comment. They are listed in the final version as potential stressor metrics.
Q	9	Wetlands	How did you perform this validation, and I suggest it be included here. This is to serve as the 305(b) section of the IR, You should not skimp on methods, especially for program development.	Thank you for your comment. The validation report is still being revised, but will be released as it is finalized. Please see the reference (CH2MHill, 2014) on DWQ's website for more information.
Q	10	Wetlands	Percent?	Thank you for your comment. Relative cover (as the proportion of cover attributed to a specific plant species) is commonly discussed in terms of percentages, while some data analysis techniques use the fractional value.
Q	11	Wetlands	Water availability/stagnation?	Thank you for your comment. DWQ cannot speculate as to the cause or drivers of the observed lower SAV Index Scores within the Lower Weber subwatershed (relative to the other subwatersheds) with the available data.
Q	12	Wetlands	When is early autumn? Didn't this vary by several weeks - depending on the order of pond visits? The timing of these events is important,whether its mid to late august vs mid to late september - but also depends on water quality - as related to flushing. As shown for the Willard Spur, this may simply be related to pond stagnation. You should point this possibility out here in. You should include this important factor before you do any more probably surveys.	Thank you for your comment. The description of the sampling windows for both summer and early-autumn IPs has been clarified.
Q	13	Wetlands	This is absolutely NOT TRUE. We have spent 10 years trying correlate surface mats to degraded conditions/loss of SAV or Macroinvertebrate degradation. Read our work! Actually you admit to such below - so you just need to change the statement.	Thank you for your comment. Whether surface mat cover relates to SAV cover or macroinvertebrate community composition is irrelevant to the assessment described in this chapter because the cover of surface mats, as described in section Indicators of Wetland Condition in the subsection Surface Mats (see also comment Q10). Extensive surface mats may indicate a degradation to the recreation designated use (associated with waterfowl hunting).
Q	14	Wetlands	As compared to what. Your assumption is that 25% cover is poor, but you have no supporting data for such a conclusion. You are trying to add a fabricated reason for support to one of our early hypotheses - that surface mats MAY be responsible for other ecological degradation - but after 7 more years of careful research - we can find no linkage to ecological degradation. And apparently, you can't either! Although we all think that mats are important, we can't find a reason. You need to discard this metric from you analysis. It is meaningless.	Thank you for your comment. Please see the last paragraph of section, Development of an Impounded Wetland Assessment, for the text describing how the GOOD and POOR assessment classes were derived.
Q	15	Wetlands	Well, I would hope that this is true, but there is still huge variability. Can you address possible sources of variability?	Thank you for your comment. DWQ agrees with the commenter about the size of apparent variability for the PMI vs. SAV index scatterplot, however, we cannot describe the effect size of other factors with the data that are currently available.
Q	16	Wetlands	This is what I'm talking about. There is no known linkage to other ecological indicators.	Thank you for your comment. DWQ acknowledges that linkages between SAV and ecological indicators deserve additional investigation.
Q	17	Wetlands	OK, you should discuss the distinct possibility that the bugs are just responding to more SAV - i.e. it co-varies.	Thank you for your comment. A short discussion of potential influence of co-variables has been added to the final version of the chapter.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
Q	18	Wetlands	Once again, this supports our conclusions as well. Surface mats are simply a red herring for "bad" conditions. Also, consider that healthy SAV reaches the surface, hence sometimes providing an "anchor point" for drifting mats. Also, if you spend enough time in the wetlands, you would notice that mats are very ephemeral. It's hit or miss whether you catch a pond with any surface mat. AND, I wonder if the questionare to duck hunters about surface wasn't a "loaded" or leading question as to poor conditions. How would they possible know or care about surface mats?	Thank you for your comment. DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. This includes metrics associated with surface mats. DWQ welcomes data and rationale that support additional or alternative indicators of health.
Q	19	Wetlands	You need to explain how this was validated.	Thank you for your comment. The rationale, methods, and results of the MMI validation are available in the validation report (CH2MHill, 2014), which will be available shortly on DWQs website.
Q	20	Wetlands	Obviously, the surface mat metric is of little value for predicting wetland health. Also, the Macroinvertebrate metric co-varies with surface mat coverage and both of these co-vary with DO and pH - because healthy SAV will produce lots of oxygen and consume lots of CO2 -raising the pH. Therefore, the metrics you have chosen are not sensitive enough to the stressors that do drive changes and are very confusing.	Thank you for your comment. DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. This includes metrics associated with surface mats. DWQ welcomes data and rationale that support additional or alternative indicators of health. Unfortunately, there is insufficient evidence to support the points that the commenter brings up
Q	22	Wetlands	Only in your own mind - because there is no link to other biological degradation. Also, you have no idea how much of the P is recycled (from sediments) vs watershed P. It may primarily be a result of pond management history i.e. if it flushes P will be low and there will be a net sink to the sediments - if it doesnt flush, there will likely be a large sediment source - or at least neutral (i.e. P in = p out of the pond). So this is also not directly linked to cultural eutrophication.	Thank you for your comment. However, as stated in the methods section, extensive cover of surface mats was defined as an apparent degradation to the recreation use for these wetlands, and cultural eutrophication is not mentioned in the paragraph the commenter is responding to.
Q	23	Wetlands	You should comment (pr at least speculate) on what stressors are unique to the Weber watershed.	Thank you for your comment. DWQ prefers to refrain from speculation beyond the scope of our dataset, although the evaluation of stressors remains an active pursuit for the wetlands program.
Q	24	Wetlands	Although this seems to be a worthwhile excercise in probabilistic statistics, you have not ventured at all into the aspects of pond management and internal biogeochemistry that are responsible for the differences in pond condition. Consequently, the metrics, especially for the SAV, are very simplistic and general and really have quite poor relationship to the the stressors that you have identified here. i.e. the range of DO, the range of pH, the range of EC the range of water depth, have not related to any of the actual biological indicators identified by Miller et al. 2012, Hoven, et al 2010 or Hoven et al. 2014 or Carling et al. 2013. I suggest you carefully read these reports and publications before you continue along this line. They might motivate you in altering the next steps listed here and assist you in where you place future resources and efforts.	Thank you for your comment and the references. DWQ has obtained the paper described by the Carling et al., 2013 reference; however DWQ would appreciate a more complete citation for the other works so that we can examine those works. DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. DWQ recognizes that pond management is an important aspect of wetland function.
R	1	Wetlands	are depths measured in feet or miles?	Units have been clarified for the figures. Thank you for your comment.
R	2	Wetlands	Actually it suggests that for some strange reason plant associated invertebrates are associated with submerged aquatic plants.	Thank you for your comment. DWQ acknowledges that the linkages between invertebrates and other wetland characteristics are complex.
R	3	Wetlands	why would it be considered a stressor at all if the beneficial use of IWs is to support waterfowl	This sentence has been clarified by simplifying the statement. The comparisons presented in the boxplots show how a dependent variable (in the case of Figure A2-8, as the comment refers, Total invertebrate biomass) differs among sites that represent the upper and lower quartiles of the biological response metric (in this case, SAV Index). Any notions of stressor versus response have been removed.
R	4	Wetlands	This relationship needs to be remedied throughout analysis and conclusions. Looks like surface mats are not a 'stressor' or indicator metric of poor condition. Suggest further analyses or dropping surface mats all together	Thank you for your comment. A significant portion of the analysis presented in this chapter was to examine how well the previously defined metrics (via Miller and Hoven, 2007; DWQ, 2009) work when applied to a larger dataset of randomly selected sites. However, as described in the methods section, this metric (occurrence of algal mats in early autumn) does not need to relate to SAV or other biological aspects of condition, because it was identified independently as an indicator of degraded conditions from a recreational use perspective (Duffield et al., 2011).

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
R	5	Wetlands	Not sure how surface mats can respond to macroinvertebrates. How do macroinvertebrates affect the surface mats?	These figures simply show how the analysis of GOOD / POOR conditions, as defined in the main text, correspond among variables (or metrics). They are not intended to show causality, and clarification has been added to the chapter (Box 1).
R	6	Wetlands	Were PMs collected in the surface mats? Are some of the taxa associated more with surface mats than SAV? If so the taxa richness and SI results suggest surface mats increase diversity, which is supposedly a good thing.	Thank you for your comment. When surface mats were encountered at a sampling location, they were a component of the "sweep" for macroinvertebrates, particularly because the ponds were generally quite shallow (less 100 cm).
R	7	Wetlands	Again, increased taxa richness probably indicates better conditions and a wider selection of ducky food.	Thank you for your comment. The Simpson's Index (SI) aspect of the benthic macroinvertebrate metric reflects taxa richness.
R	8	Wetlands	If you didnt measure bug biomas from surface mats, this is irrelevant	Thank you for your comment. When surface mats were encountered at a sampling location, they were a component of the "sweep" for macroinvertebrates, particularly because the ponds were generally quite shallow (less 100 cm).
R	9	Wetlands	This is probably because there is more productivity in sites with greater SAV and surface mats.	Thank you for your comment. DWQ agrees. Sites that contain greater cover of both SAV and surface mats likely have great primary productivity.
R	10	Wetlands	In general, PMI and SAV index are strongly correlated and one can be substituted for the other. Pick one please.	Thank you for your comment; however, DWQ has found that the correlation between PMI and SAV index is not strong enough, based on the available data, to justify removing a metric at this time.
R	11	Wetlands	Please do not use 'response' in the captions. You dont know if these were true responses or covariates. Maybe use 'relationships"	Thank you for your comment. Descriptive text for these figures has been revised in Box 1, where these box plots are used to represent comparisons between a series of characteristic variables versus upper and lower quartiles of the three main biological response variables.
R	12	Wetlands	than' not 'that'. Please review this manuscript and remove grammatical, spelling, errors.	Thank you for your comment. The editorial change has been made.
R	13	Wetlands	Not surprising. I bet there was a very strong correlation between PMI scores and diversity metrics especially taxa richness. If so, chose one or the other.	Thank you for your comment. DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. This includes metrics associated with PMI and diversity. DWQ welcomes data and rationale that support additional or alternative indicators of health.
R	14	Wetlands	What?	Thank you for your comment, although the intention of the comment is unclear.
R	15	Wetlands	The following captions and discussions make no sense. Looks like you didnt complete the report and is in rough draft status	Thank you for your comment. Figure captions have been revised in the final version of the chapter.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
R	16	Wetlands	This would logically follow that good sites (higher SAV scores etc.) were more productive than poor sites. Hence greater bug bioass.	Thank you for your comment. However, it is not necessarily the case. If the composition of macroinvertebrate communities shifted among condition classes, then differences in predation, individual size (biomass per organism), or biomass turnover could counteract a simple productivity vs. biomass relationship (see Gray, 2013; report on macroinvertebrates in Willard Spur). DWQ will continue to evaluate the best wetland assessment methods and metrics as part of our ongoing wetland program development. This includes metrics associated with macroinvertebrate biomass, productivity, and composition. DWQ welcomes data and rationale that support additional or alternative indicators of health.
S	1	303(d)	SUWA is concerned that DWQ did not sufficiently pursue all reasonably available sources of data in compiling the IR. The IR does not incorporate or use reasonably available information and data collected by the National Park Service (NPS) for waters in or near national parks and national monuments in Utah, including Arches, Bryce Canyon, Canyonlands, Capitol Reef, and Zion National Parks and Hovenweep and Natural Bridges National Monuments. This information was collected between October 1, 2009 and September 30, 2012. In fact, DWQ staff helped collect, analyze or otherwise compile some of the information. However, much (if not all) of the information collected by NPS is not reflected in the IR. The following is a non-exhaustive list of information/data not incorporated into the IR by DWQ.	This response applies in general to the commenter's assertions that the accompanying National Park Service report "Water Quality in the Northern Colorado Plateau Network, Water Years 2010–2012" represents a legitimate beneficial use assessment of those waters in question. The document itself cites in section 2.4.3 that "The goal is to provide advance warning of an impending problem before it becomes severe, rather than to meet regulatory definitions of impairment." It goes on to describe their assessment methods as using a 10% exceedance approach, which is only consistent with the methods used by DWQ for conventional parameters such as DO, temperature, pH, and total dissolved solids. The report also states in section 2.4.5 that "this report compares water quality data to designated beneficial use criteria without stating whether a designated beneficial use was attained. Those designations are left to the states." DWQ has no direct evidence that their methods are comparable with our own, given the level of information provided in this report. We interpret their definition of "reportable exceedances" as an internal data evaluation for NPS management purposes and not as an independent 303(d) assessment. Furthermore, we have been working with NPS for many years providing them with laboratory allocation at the Utah Public Health Laboratory (UPHL). This is the same analytical laboratory used by DWQ, and all sample results for cooperative monitoring agencies such as the NPS come directly to DWQ, are incorporated into our assessment database, and are disseminated to the various agencies. Our assumption is that all these data are contained in the 2012-2014 IR, and the discrepancies between their results and DWQ's are due to different assessment methods. We cannot view the report as comparable to the assessment performed as part of the 2012-2014 IR. Site-specific comments are addressed individually below.
S	2	303(d)	Arches National Park The NPS conducted 2,882 designated beneficial-use evaluations for water quality results at nine sites in or near Arches National Park between October 1, 2009 and September 30, 2012. Six of the monitored sites recorded exceedances of respective water quality standards. This information was not incorporated into the IR. For example, Salt Wash reported exceedances for dissolved aluminum, dissolved oxygen, total phosphorus, and total dissolved solids (TDS). However, the IR lists Salt Wash as impaired for TDS only. The IR must list all impaired waters and for all impairments, not just TDS. Courthouse Wash is another example. The segment above the confluence with the Colorado River reported exceedances for total phosphorus while the segment referred to as "Upper Courthouse Wash" reported exceedances for dissolved arsenic, dissolved oxygen, E. coli, total phosphorus and temperature. The IR lists dissolved oxygen and arsenic only.	Salt Wash illustrates a significant difference in assessment method for aluminum between the NPS report and the 2012-2014 IR. For toxics such as aluminum, percent exceedance is not used as an exceedance method; rather results are compared to the criterion if >1 samples exceed the value, the site is considered as impaired. Furthermore, R317.2 contains a number of equations to adjust for mitigating conditions such as hardness, temperature, and pH which can affect a metal's toxicity to aquatic life. In the case of aluminum, the standards specify which criterion applies depending on the hardness and pH. It states "The criterion for aluminum will be implemented as follows: Where the pH is equal to or greater than 7.0 and the hardness is equal to or greater than 50 ppm as CaC03 in the receiving water after mixing, the 87 ug/1 chronic criterion (expressed as total recoverable) will not apply, and aluminum will be regulated based on compliance with the 750 ug/1 acute aluminum criterion (expressed as total recoverable)." For Salt Wash, all the hardness and pH data indicate that the 750 mg/l be applied as per standards. Therefore, the IR does not list this site as impaired for aluminum because there were no exceedances of the 750 mg/l criterion. Similarly, the NPS report does not apply equivalent assessment methods at the Courthouse Wash. Chapter 2 contains our assessment methods for all the parameters listed.
S	3	303(d)	Bryce Canyon National Park The NPS conducted 1,122 designated beneficial-use evaluations for water quality results at four sites in Bryce Canyon National Park between October 1, 2009 and September 30, 2012. This information was not incorporated into the IR. Sheep Creek, a tributary to the Paria River, reported exceedances for total phosphorus but this is not reflected in the IR. Similarly, Yellow Creek is not listed in the IR.	Chapter 2 describes the methods for assessing sites for total phosphorus. In the case of the 2012-2014 IR, nutrient assessment methods are currently under development, and, until such time as numeric criteria are established, DWQ has placed these waterbodies in Category 3D. Once criteria are developed, we will assess in future IR cycles.

Response to Comments on the Utah Division of Water Quality's 2012-2014 Integrated Report

Letter	Comment #	Section	Comment	Response
S	4	303(d)	Canyonlands National Park The NPS conducted 5,678 designated beneficial-use evaluations for water quality results at fifteen sites in or near Canyonlands National Park between October 1, 2009 and September 30, 2012. This information was not incorporated into the IR. For example, the Colorado River above its confluence with the Green River (referred to in the IR as "Colorado River-3") reported impairments for total phosphorus, dissolved selenium, and temperature. However, selenium is the only impairment of these three listed in the IR. Moreover, the Colorado River below "Big Drop #3 Rapids" (referred to in the IR as "Colorado River-2") reported exceedances for total phosphorus and temperature but is listed in the IR for aluminum only. Finally, Salt Creek, in southeastern Canyonlands near Peekaboo Spring and Crescent Arch, reported exceedances for temperature and dissolved oxygen, dissolved mercury, and total phosphorus, respectively. The IR lists cadmium and selenium as the only impairments. Capitol Reef National Park The NPS conducted 1,570 designated beneficial-use evaluations for water quality results at six sites in or near Capitol Reef National Park between October 1, 2009 and September 30, 2012. This information was not incorporated into the IR. Oak Creek reported exceedances for pH and total phosphorus and separate segments of Sulphur Creek reported exceedances for E. coli, total phosphorus, temperature, and TDS, respectively. The IR does not list all of these impairments. Zion National Park The NPS conducted 1,865 designated beneficial-use evaluations for water quality results at five sites in or near Zion National Park between October 1, 2009 and September 30, 2012. This information was not incorporated into the IR. According to NPS, La Verkin Creek at Lee Pass Trail reported exceedances of relevant water quality standards but the IR lists this water segment as "Supporting." Hovenweep National Monument The NPS conducted 853 designated beneficial-use evaluations for water quality results at three sites in Hovenweep National Monument between October 1, 2009 and September 30, 2012. This information was not incorporated into the IR. Cajon Spring and Square Tower Spring are not listed in the IR despite reported exceedances. Natural Bridges National Monument The NPS conducted 624 designated beneficial-use evaluations for water quality results at three sites in Natural Bridges National Monument between October 1, 2009 and September 30, 2011. This information was not incorporated into the IR. Armstrong Canyon Creek reported exceedances for dissolved aluminum, dissolved oxygen, dissolved mercury, and total phosphorus but is listed on the IR for dissolved oxygen only. Moreover, Owachomo Bridge Spring and Sipapu Bridge Spring do not appear on the IR despite recorded exceedances. It is arbitrary and capricious for DWQ to ignore the water quality information/data in the Water Quality Report, especially when DWQ helped compile, analyze, or otherwise interpret much of it. This information must be incorporated into the IR. Federal regulations require DWQ to examine all existing and readily available data when making assessment decisions, which includes consideration of data collected by DWQ and others such as NPS.	As per previous responses, DWQ and the NPS do not have equivalent assessment methods for their different programs. In addition, DWQ uses the same raw data source as the NPS report but arrives at different conclusions based on differing quality control measures, assessment methods, and appropriate application of water quality standards. DWQ will continue to work with their cooperators to assist in the design of sampling analysis plans to include proper sampling locations and methods. Furthermore, DWQ will ensure that cooperating agencies are properly interpreting WQ standards to be consistent with the results of the IR.