DEQ
Montana Department
of Environmental Quality

Comments and Reponses for the Electrical Conductivity (EC) and Sodium Adsorption Ratio (SAR) Assessment Method for Rosebud Creek, the Tongue, Powder, and Little Powder Rivers, and the Tongue River Reservoir

February 2021

Greg Gianforte, Governor
Chris Dorrington, Director DEQ

This document presents comments and responses to the *Electrical Conductivity (EC) and Sodium Adsorption Ratio (SAR) Assessment Method for Rosebud Creek, the Tongue, Powder, and Little Powder Rivers, and the Tongue River Reservoir* developed by Montana Department of Environmental Quality. Results from this method are used in the process of Beneficial Use Assessment for Montana's Surface Waters (Makarowski, 2020). Comments were accepted from July 27th - September 27th, 2020.

Comment 1: For more than 40 years, Decker Coal Company has been the largest major permitted point source discharger to the Tongue River in Montana. Extensive monitoring demonstrates that there are no measurable impacts to the water quality in Tongue River as a result of our operations. The results from our long-term seasonal monitoring of the water quality immediately upstream and downstream of the Tongue River Reservoir confirms our stance that the Decker Coal Company operation has had no negative impact to water quality.

Locally, in proximity to the Tongue River Reservoir, Decker Coal Company has accumulated over 40 years of water quality data that shows the surrounding native topography, soils, stream flows, and groundwater flows will not always meet the same proposed limits DEQ is forcing on Decker Coals MPDS discharge. DEQ's proposed standard is excessive as the native inflow of water will not meet the proposed standards.

There are numerous examples and data that prove the proposed limits are more stringent that the native water quality. DEQ Coal and Opencut Mining Bureau have taken EC samples during the spring and fall and have found that native drainages greatly exceed the proposed 1500 uS/cm. DEQ has not fully evaluated all of the sources of solute influx and outflux within the Tongue River Reservoir. DEQ made statements during a public meeting that the exceedances in the EC and SAR standards downstream in the Tongue River, near its confluence with the Yellowstone River during the identified critical months of March and April, are largely attributed to low flow rates. These values are influenced by seasonal surface water flows and groundwater discharge. The reduced seasonal recharge to, and discharge from, the Tongue River alluvial aquifer, within the lower extents of the river valley in recent years is the results from some of the relatively recent significant changes in longstanding irrigation practices, points of diversion and in the means of conveyance of irrigation water within the lower Tongue River valley, as many water users transitioned from historic flood irrigation application to the use of sprinkler equipment. DEQ has not considered or implemented this data in any evaluation.

Early published studies by the Montana Bureau of Mines and Geology in 1975 (Bulletin 97, Van Voast W.A. and Hedges R.B.) and 1982, (Hydrogeologic Map 5, Van Voast, W.A. & Thompson, K.S.) accurately quantified the natural pre-mining loads that discharged to Tongue River from within Decker Coal's early permitted mine areas, adjacent to the west and east sides of Tongue River Reservoir. Below are the published solute concentrations and groundwater inflow rates to the Tongue River from the mine disturbed D1 and D2 coal seams and their associated permeable residual burn units (scoria):

Inflow 3.4 CFS, TDS 630 mg/L; From West Decker Pit 16 (North) Area Inflow 0.1 CFS, TDS 1,010 mg/L; From West Decker 11 and 12 Pit Areas Inflow 0.1 CFS, 1,530 mg/L; From East Decker 13, 14, and 15 Pit Areas Inflow 0.07CFS, 2,290 mg/L; From Spring Creek Mine

While these studies did quantify the expected minor load contribution from the Spring Creek Mine, they did not appear to fully evaluate the additional solute load contributions from the Dl and D2 coal seams (and associated scoria) that discharged into the reservoir from other areas outside of the permit area, such as: all of the coal subcrop areas near the inlet area and those areas farther north of Decker

Coal's mines, or from the undisturbed D3 coal seam beneath the floor of the reservoir. Also, they did not specifically identify or quantify the concentrations and groundwater discharge rates into the Tongue River Reservoir from the Tongue River alluvium or the Badger Creek alluvium, downstream from the "State Line" USGS Station (06306300). Nor did these studies include the solute load naturally contributed to the reservoir from the Deer Creek alluvium, immediately north of the East Decker mine permit boundary.

As part of the DEQ Tongue River Salinity TMDL Project, on October 10, 2017 in Miles City, MT, the DEQ presented draft results of a surface water modeling effort which was used to statistically evaluate changes in the solute load in Tongue River from the State Line to Miles City. For a portion of the model, the total solute concentrations were estimated from specific conductivity measurements obtained at the USGS station near the Montana state line (about 3.5 miles upstream of the inlet to Tongue River Reservoir and Decker Coal's TR0977 monitoring station). These values were then compared to similarly obtained estimated total solute concentrations at a second USGS station, downstream a short distance below Tongue River Dam, which coincides with Decker Coal's TR1078 monitoring station. The attached Figures 1 and 2 utilize supporting data from the draft surface water model that was provided to Decker Coal by the DEQ. The attached Figure 1 is a comparison of the DEQ's estimated TDS values on days when Decker Coal collected laboratory samples from near the inlet of the reservoir. Figure 2 compares the DEQ estimated TDS concentrations and Decker Coal's analyzed TDS concentration sampled from the same location below the dam.

From Figure 1, Decker Coal believes much of the greater variability and higher solute concentrations observed in the laboratory TDS analytical results versus DEQ's estimated TDS are attributed to solute load contributions from some of the other sources of inflow to the Tongue River between the monitored state line station and the inlet to the reservoir. These additional solute load sources were not accounted for in the draft surface water model and could potentially be attributed to Decker Coal's discharge. As should be expected, plots of the estimated and laboratory analyzed TDS in Figure 2 shows notably less variability because all of these measurements and samples were obtained from the same location.

It is important to note that the associated data and Figures 1 and 2 also show a decrease in laboratory measured TDS concentrations downstream of the Tongue River Dam when compared to those obtained from samples collected near the inlet to the reservoir. This would indicate a net reduction in solute load occurring between the inlet and the dam, even with the additional load contributions from Decker Coal's discharge. Similar observations are evident throughout the life of Decker Coal's mining operations. Some explanations are presented below.

Mining has temporarily interrupted some of the groundwater (and the contributing solute load) that naturally discharged from the pre-mine pit areas at both of Decker Coal's operations.

Additionally, starting in 1999 through approximately 2017, coal bed methane extraction has significantly further depressurized the mined and underlying coal seams. As a result, there has been further reduction in those solute loads naturally contributed to the Tongue River from the Dietz 1 and Dietz 2 coals subcrops near the reservoir inlet, and from the Dietz 3 coal subcrop beneath the northern end of the reservoir. Since the reconstruction of the Tongue River Dam in 1999, the higher water level has significantly increased the amount of stored water in the Tongue River Reservoir that also discharges into the active and reclaimed mined pits, via hydraulic connection through the remaining unmined permeable alluvium, shallow coal and scoria. During nearly all of the years of

mining by Decker Coal, most of the pit inflows were returned back into the Tongue River Reservoir by permitted discharge. A portion of the seasonal inflow (and associated solute load) from the reservoir, however, remains as recharge to the reclaimed pits. These effects are confirmed by the recovering groundwater levels measured in the reclaimed spoils. The fraction of groundwater inflow that is retained in the backfilled pits from Tongue River Reservoir, and the recharge to the current regionally depressurized unmined Dietz 3 coal seam via its subcrop beneath the reservoir represents actual reductions in the solute load in the Tongue River.

Additionally, there are points below the USGS Stateline gauging station and above Decker Coal's TR0977 sampling point where water is diverted from the Tongue River to flood and sprinkler irrigate approximately 300 acres. These withdrawals are sufficient to contribute to a temporary reduction in the solute load during the irrigation season. Conversely, the fraction of surface water return flow to the river from these same irrigated fields and the return from bank storage during the late fall and early spring months, at somewhat higher solute concentrations, would be accounted for as an increase in load.

Finally, a permitted diversion point has recently been upgraded to pump industrial water from Tongue River Reservoir by Spring Creek Mine. The pump/diversion point is situated in El/2 NWl/4 Section 35 TBS R40E and is believed to have been operational in 2016. Its operation would certainly contribute to further reduction in the solute load in Tongue River Reservoir.

There are challenges in accounting for all possible surface water contributions to the reservoir in any given year. Nearly all of the volume of surface water runoff that is collected in pit sumps can be used for dust suppression, or infiltrates as additional recharge to the backfilled pit areas. Finally, the contributing loads from several moderately sized native ephemeral drainages farther north from the Decker Coal mines may contribute runoff occasionally, but are unaccounted for by active monitoring.

The attached Figure 3 compares long term monitored laboratory SAR reported values in Tongue River from Decker Coal's TR0977 and TR1078 monitoring locations. As shown, starting just before 2001, results from both monitoring locations showed notable increases in the monitored SAR. This corresponds with the commencement of Wyoming and Montana permitted wastewater discharge from coal bed methane (CBM) operators, upstream of the Tongue River Reservoir. With the start of a significant decline and eventual termination of gas production within the Tongue River drainage starting in approximately 2009, subsequent monitored SAR at both locations have gradually declined to nearly pre-CBM values by 2016. These graphs confirm there is no significant increase in monitored SAR, downstream of Decker Coal's permitted discharge points. This would support our position that Decker Coal's current discharge limits for SAR are sufficiently protective. We firmly believe that any further reduction in SAR limits for the Tongue River Reservoir would be unduly restrictive to Decker Coal Company and would not benefit downstream users.

The results of long term investigative and compliance related hydrologic monitoring efforts by Decker Coal closely support the predictions presented in The *Tongue River Basin Project*, March 1996 FEIS which was compiled during preparations to reconstruct the Tongue River Dam. The FEIS accurately identifies the outcomes of increasing the capacity of the Tongue River Reservoir resulting from the higher post-reconstructed dam height. The bulleted items below are excerpts from this FEIS, with the assumptions of the model that "conditions during the 44 years after the dam is rebuilt are similar to those that occurred between 1946 and 1989":

- *"Following rehabilitation of the dam, the reservoir water levels typically would increase above current levels. The increased water levels would result in more groundwater seepage into coal mine pits. Measures proposed to mitigate economic effects on mines would be accomplished using additional power for pumping and greater pumping capacity, additional sediment pond capacity, and obtaining permission to increase annual discharge under the MPDES from DEQ."*
- The *"groundwater seepage rates into mine pits are expected to increase. This increased seepage would require increased pumping rates to dewater the pits. Would increase flows to settling ponds, and increase discharge of pit water at permitted discharge stations. Increased discharge flows have the potential to impact the water quality of Tongue River Reservoir and Tongue River downstream of the reservoir*."
- *"This could negatively impact coal recovery and mine operation. The FEIS also states that the elevated water levels would increase the concentration of TDS discharged back to the reservoir."*
- *"The resulting estimate of the increase in the TDS concentration in the reservoir, assuming typical streamflows and reservoir mixing, likely would not be detectable. Since downstream discharges occur after the mixing of the TDS flows within the reservoir, increases in TDS concentrations resulting from pit dewatering may not be detectable. Impacts of increased mine pit discharges on reservoir and downstream water quality would be negligible in the short and long terms."*
- *"increase dissolved-solids concentrations of the Tongue River at Miles City in most months. Under median and dry streamflow conditions, the depletions for irrigation would cause the largest in dissolved-solids during the late irrigation season. Increases would be largest under dry streamflow conditions and water would pose a moderate salinity hazard for irrigation use. The increase in salinity would be the greatest near Miles City and progressively less, upstream towards the dam."*
- *"dissolves solids concentrations in the Tongue River at Miles City would be small in most months. In dry years, dissolved-solids concentrations would increase in June through August and during the winter; increases would create a moderate salinity hazard for irrigation use. Effects on salinity would be greatest near Miles City and progressively less toward the Tongue River Dam." "At some point, additional future consumptive use of water will create salinity problems and limit the suitability of water for irrigation use. Future water allocation decisions will need to account for salinity effects. The development of such a model likely will be part of future environmental compliance activities required under the Settlement Act to be completed when the Tribe proposes specific water uses."*
- *"Decker Coal operations have little impact on the reservoir or the river."*The EPA's 2010 NPDES Permit Writers' Manual states that: *"Many state water quality standards have general provisions allowing some consideration of mixing of effluent and receiving water when determining the need for and calculating WQBELs. Depending on the state's water quality standards and implementation policy, such a mixing consideration could be expressed in the form of a dilution allowance or regulatory mixing zone. A dilution allowance typically is expressed as the flow of a river or stream, or a portion thereof. A regulatory mixing zone generally is expressed as a limited area or volume of water in any type of waterbody where initial dilution of a discharge takes place and within which the water quality standards allow certain water quality criteria to be exceeded*."

For the current and previous MPDES permit terms, Decker Coal Company utilizes a permitted mixing zone allowance at both mining operations. The characteristics of the mixing zone below the East Decker 002 Outfall and West Decker 007 were evaluated by implementing DEQ approved mixing zone studies on two separate occasions. These studies, conducted near typical seasonal low and high reservoir operating stages, clearly demonstrate that the relative higher solute concentrations in mine effluent thoroughly dilutes to background solute concentrations in the reservoir within 200 feet or less, of the discharge inlet. Based upon this information, the current and previous monthly and maximum discharge limits were created by the DEQ. Again, our data continues to show that there is no measurable increase in solute concentrations, or other monitored analytical parameters between Decker Coal's designated long time monitored stations near the inlet to the reservoir, upstream of the mines, and immediately downstream of Tongue River Dam.

Therefore, Decker Coal strongly believes the current permitted discharge limits are sufficiently stringent to prevent measurable deterioration of water quality in the Tongue River and Tongue River Reservoir, as demonstrated by actual laboratory analytical results obtained using approved monitoring programs. Based upon Decker Coal's long history of monitoring and reporting of the water quality results from Tongue River near our mining operations, we fail to see any benefit to water quality in the Tongue River, or to other downstream water users resulting from the proposed discharge limits of 1,000 uS/cm monthly average and 1500 uS/cm maximum discharge limits for Tongue River Reservoir. Given the amount of analytical data that supports our position, we consider the proposed Specific Conductance and SAR limits unduly restrictive to Decker Coal Company, when fully considering the significant number of other known natural and non-mining related influences to the water quality and quantity in Tongue River between the Montana state line and Miles City.

As an additional comment, the DEQ's WQPB 2020 Draft Integrated Water Quality Report *Appendix A: Impaired Waters* does not list salinity in the Tongue River as an impairment until its confluence with Hanging Women Creek. The first reference of Specific Conductance impairment in Tongue River downstream of the Montana State Line, from coal mining, is approximately 90 miles downstream at its confluence with Beaver Creek.

As a result, Decker Coal Company would like to restate from the above FEIS: *"Decker Coal operations have little impact on the reservoir or the river." "At some point, additional future consumptive use of water will create salinity problems and limit the suitability of water for irrigation use."* And *"groundwater seepage rates into mine pits are expected to increase. This increased seepage would require increased pumping rates to dewater the pits."* Knowing those facts, the reservoir water levels were still raised; greatly impacting the amount of water that seeps into the mine and in turn increases the discharge rates from the mine back into the reservoir. Implementing the proposed discharge limits is unduly restrictive to our operation and there is legitimate concern that the proposed option to further reduce point source discharge concentrations and load limits for the Tongue River Reservoir will not translate into the desired improvement in water quality further downstream.

> **Response:** The comment focuses on Water Quality Standards, Permitting and Total Maximum Daily Load program topics. The comment does not directly address the EC/SAR assessment method or assessment outcomes and is therefore considered out of scope for this project. However, the Department believes it is appropriate to provide feedback since the comment addresses the underlying standards for the assessment method and the steps being taken by the Department to help achieve the standards. Note that the EC and SAR standards for the Tongue River have been adopted by the Board of Environmental Review and are codified at

Administrative Rules of Montana (ARM) 17.30.670 and are not "proposed" standards as implied by the Commenter. The EC and SAR standards are set to protect the agricultural use at levels shown to be achievable for the Tongue River Reservoir and segments of the Tongue River based on recent and historical data. Reductions to meet the EC standard in downstream segments of the Tongue River must address cumulative impacts from multiple sources throughout the watershed, including sources to the Tongue River Reservoir and upstream of the reservoir. Reductions from one specific pollutant source alone, through proposed discharge limits, is not expected to achieve the load reduction necessary to meet a downstream standard. That is why Total Maximum Daily Loads developed throughout Montana depend on pollutant load reductions from all identified sources.

Comment 2: Is it correct that the EC and SAR Assessment Method for Rosebud Creek, the Tongue, Powder, and Little Powder Rivers, and Tongue River Reservoir is only effective for the mainstems? If yes, is there a method that is in the works for the tributaries? If no, how does the method allow for extrapolating limits to tributaries?

Response: The Department acknowledges that assessment methods for the tributaries need to be developed. Irrigation practices and timing on the tributaries vary from those on the mainstems and nonanthropogenic conditions may cause or contribute to EC values that do not meet the existing tributary standard. To help address this concern, the Department has developed guidance for an external entity that wants to pursue a nonanthropogenic standard (NAS); located at http://deq.mt.gov/water/Surfacewater/standards; under Supporting Technical Documents.

Comment 3: Please describe DEQ's analysis of available EC or specific conductance (SC) data supporting that EC is more variable and therefore necessitates more stringent data requirements and a more complex assessment methodology than other parameters and documents showing that the methodology is consistent with the EC and SAR criteria approved under CWA § 303(c)

Response: After further analysis of available SC data and review of the Draft Assessment Method, DEQ has reduced the number of months and samples required to perform an assessment for both EC and SAR. To mitigate the complexity of the assessment method, increase feasibility, and still capture monthly variability of SC, DEQ will now require a minimum of four samples to perform an assessment for the monthly standard and one sample to perform an assessment on the do not exceed standard. With a minimum dataset, samples must be collected in April or adjoining months with one year of data collection required. This is the most variable and sensitive timeframe of SC conditions. Data requirements to delist will include 24 samples collected over two years, in each of the two years sampled, samples collected between the months of March through May must be included.

Comment 4: EPA is concerned the level of complexity may make it challenging to apply DEQ's current assessment methodology.

Response: The level of complexity and repeatability of the assessment method has been addressed by revising the minimum data requirements and data analysis guidance. Additional simplifications are provided in the final document.

Comment 5: The draft EC/SAR methodology does not currently describe the minimum dataset for DEQ to have sufficient data to complete an assessment. EPA strongly encourages DEQ to clarify and make transparent to the public its minimum data requirements to complete an assessment in the final version of the methodology.

> **Response:** DEQ has revised the assessment method to clearly describe minimum data requirements. Minimum data requirements are as follows: To perform a new assessment, or to list an assessment unit (AU), one sample must be collected during each calendar week of a full month during high flow and the irrigation season (March-May). Samples must be collected one week apart. See Section 3.5.2. To perform an assessment to delist an AU, a minimum of 24 samples must be collected over two years. Sampling to delist an AU must occur over 3 months of each sampling year, with one month collected during high flow irrigation season (March through May) and one month must be collected during low flow (June through February). During the months sampled, one sample must be collected during each calendar week for the full month. See Section 3.5.2.

Comment 6: Please explain why DEQ did not pursue an approach (similar to the metals assessment methodology1) that establishes a percentage of the dataset that should be collected during high flow conditions, when higher SC concentrations may be observed.

> **Response:** DEQ has adopted a similar approach to the metals assessment by requiring sampling during high flow. See response to comment 5.

Comment 7: Please clarify in Montana's final IR what assessment decisions were made for 2020 that applied these draft methods and whether DEQ reviewed those assessments based on any modifications to the draft assessment method.

> **Response:** DEQ assessed the Tongue River; Beaver Creek to Twelve Mile Dam (MT41C001_014) for the final 2020 IR using the draft assessment method. The record for this segment has been updated for the final 2020 IR.

Comment 8: EPA would like to review DEQ's SOP and QAPP that address collection of continuous data in order to understand how the state determines when it will not use data because of fouling, drift, calibration issues, etc.

> **Response:** DEQ currently has a Small Data Logger Standard Operating Procedure and is documenting our process of data review for SC continuous data loggers. DEQ will share the Small Data Logger Standard Operating Procedure and the quality assurance documents as they are completed.

Comment 9: Please describe DEQ's plan for evaluating TDS data as part of an agricultural use assessment.

> **Response:** DEQ recommends evaluating TDS as part of the parameter suite with EC and SAR.

Comment 10: The assessment methodology should be revised to include the complete definition from ARM: "Sodium adsorption ratio (SAR)" means a value representing the relative amount of sodium ions

to the combined amount of calcium and magnesium ions in water using the following formula: SAR = [Na]/(([Ca]+[Mg])/2)^{½}, where all concentrations are expressed as milliequivalents of charge per liter.

Response: DEQ will make the suggested change to the SAR definition.

Comment 11: Section 3.4 states "Continuous datasets are not subject to temporal independence evaluation unless there is a situation where a partial dataset is being used along with discrete data (**Section 6.0**)." What is DEQ's rational for blending dependent and independent datasets? Should it matter if the discrete data are temporally independent if the vast majority of the dataset may be dependent? Please explain.

Response: The way monthly average EC is calculated in the assessment method makes it feasible to combine time dependent and independent data. The process to combine different types of data (continuous and discrete) from multiple sites has been simplified and updated to provide clarity. See Section 3.4 and 5.3 of the Electrical Conductivity (EC) and Sodium Adsorption Ratio (SAR) Assessment Method for Rosebud Creek, the Tongue, Powder, and Little Powder Rivers, and the Tongue River Reservoir (hereinafter "EC and SAR Assessment Method")

Comment 12: Section 3.4.1 also "an assessment unit exhibits one or more significant shifts in type and intensity of potential SC or SAR sources such that clear breaks could be made to designated new homogenous reaches, sub-segmenting may be justified." Section 3.4.2. further suggests that site-level analyses may occur before data aggregation ("an assessor may compare data from site to site after the monitoring to determine if the segment is homogenous or if sources affect EC and SAR and thus different reaches are assessed. During the assessment process, if data analysis shows that one site's data are not like the other sites in the assessment unit, a further reach break may be called for after data have already been collected."

The process for evaluating data at the monitoring site level is unclear. Please clarify DEQ's process for evaluating data at individual sites.

Recommend defining "reach" versus "assessment unit" so readers understand the distinction between a reach versus assessment unit.

Response: DEQ acknowledges these comments and provided clarification in the final document. See EC and SAR Assessment Method Section 3.4.1 "If an assessment unit exhibits one or more significant shifts in type and intensity of potential EC or SAR sources such that clear breaks could be made to designate new homogenous reaches, a reach break may be justified."

Comment 13: Section 3.5.2 states that "For both EC and SAR, when comparing against the "no sample may exceed" standards, a single sample can indicate impairment, so a single sample can constitute sufficient data if that sample meets data quality requirements (Section 5.0)." We recommend adding a separate section to the beginning of the document that details how discrete and continuous data will be compared to the "no sample may exceed" standard. Additionally, it would be helpful to clarify whether the sampling location and/or assessment unit would be considered impaired if a single value collected using a continuous data logger exceeds the "not to be exceeded" value.

Response: DEQ has added additional information to clarify how continuous and discrete data will be evaluated against the do not exceed criteria and outlines minimum datasets that are

associated with the do not exceed criteria. A minimum data to evaluate the do not exceed standards is set at four samples during high flow and irrigation season, preferably during April. Additional clarification has also been added to indicate one sample that meets data quality objectives is necessary to evaluate the "not to exceed" standards. See Section 3.5.2.

Comment 14: Please explain the rationale for requiring "at least three years (consecutive or non-consecutive) from the most recent ten years" as the applicable timeframe for assessments. If fewer than 3 years of data are available, does DEQ consider that as "sufficient data"? The rationale should align with the State's EC/SAR monthly average criteria. Section 6.1.1 states "An assessment unit will be considered impaired for SC if any single monthly average SC concentration exceeds the monthly average EC standard and/or if one single SC concentration exceeds the no sample may exceed standard." This language suggests that a single exceedance of the monthly average SC concentration is sufficient for listing. If that is the State's interpretation of the water quality standard, we recommend revising the assessment methodology to state that *one month* of data (versus three or two) represents the minimum dataset. Most importantly, the criteria do not allow any exceedances of a monthly average. Therefore, EPA recommends revising the assessment methodology to indicate that a single month of data constitutes the minimum dataset required for an assessment which aligns with the criteria.

Response: The minimum data requirements have been adjusted to align with the EC/SAR criteria. See response to comment 5.

Comment 15: Please explain the rationale for requiring *three* calendar months of data as the minimum requirement for the irrigation season and *two* calendar months for the non-irrigation season (Section 3.5.4).

Response: The minimum data requirements have been adjusted to align with the EC/SAR criteria. See response to comment 5.

Comment 16: EPA offers the following high-level concepts regarding the minimum dataset needed for continuous data.

Continuous data records can be incomplete resulting in missing data points for the duration of a sample statistic calculation. Missing data points may occur because of sonde malfunction (such as dead battery or fouling) or because deployments were shorter than the duration of the criterion (such as a 3-day deployment used to evaluate a 7-day mean). States should describe the decision-making process used to determine adequate data coverage for a sample statistic calculation.

The appropriateness of using incomplete continuous datasets will depend on whether the available data would reasonably be expected to represent conditions during the missing data timeframe. When one discrete data point is used to represent a 4-day duration criterion, assessors assume that the waterbody condition has not changed over that duration. Similar assumptions can be made for continuous datasets. One day of monitoring could likely represent conditions over 3-or 4-days, or 3 days of monitoring could represent conditions over 7-days. Factors that can influence whether collected data could adequately represent unmonitored conditions include natural watershed characteristics, types of anthropogenic stressors in the watershed, and weather conditions and streamflow over the timeframe being evaluated. States can set target data coverage for each criterion duration, though decisions will often be made on an individual sample statistic basis. To inform the duration of monitoring data to target, states can examine existing continuous datasets to determine if shorter collection periods adequately

represent monitored conditions of longer durations. Staff experienced with the waterbody can also use best professional judgement to make decisions on adequacy of data coverage.

The timing of the missing data will also influence if collected data are representative of unmonitored conditions. Dissolved Oxygen and pH can have diel swings and are expected to be lowest in early morning and highest in the afternoon. If missing data occurs in the morning, the resulting distribution of data points may not represent the true population and could therefore result in an inaccurate sample mean and minimum statistics. However, if missing data occurred in the afternoon, the sample mean may be inaccurate, but a criterion sample statistic based on minimum values would be unaffected. Depending on the criterion being assessed, the state will need to decide if the missing continuous data would result in an inaccurate sample statistic. The decision-making process to determine adequate continuous data coverage within a criterion duration should be outlined in the state assessment methodology.

EPA's detailed comments on the minimum dataset proposed in DEQ's EC assessment methodology are included below.

Section 3.5.6 states: "To assess monthly conditions using continuous SC measurements, a preferred sample frequency should include every 24-hour period of an entire calendar month sampled on a frequency no greater than 30-minute intervals; it is acceptable to collect data at a higher frequency (i.e. 15-minute intervals). Other sample frequencies may be acceptable."
The bar for minimum datasets appears to be different for continuous vs discrete data. For example, the document states: "To assess monthly conditions using discrete samples, a preferred sample frequency must include four individual samples within each calendar month with approximately one week between samples. "To assess monthly conditions using continuous SC measurements, a preferred sample frequency should include every 24-hour period of an entire calendar month sampled on a 30-minute frequency." To better align with the data sufficiency requirements for discrete data, would it be reasonable to suggest that "a preferred sample frequency could include four days of 24-hr periods of continuous data separated by three days." This approach would seem comparable to the data sufficiency requirements established for discrete data.

> **Response:** The Department revised the minimum data requirements that apply to any type of data. We will continue to include recommended data collection frequency and duration for continuous data sets. We analyzed percent relative difference of rolling 3-day, 7-day and monthly averages for the Tongue River and found up to 42% RPD for the 3-day to monthly average and 38% for the 7-day to monthly comparisons. Therefore, DEQ will keep the monthly averaging technique from the draft document. We address reduced data requirements by reducing the number of months needed to complete an assessment. This averaging approach works for EC because it is much less variable than DO or pH over short periods of time (example diurnal cycles) but tends to shift seasonally and in response to runoff.

Comment 18: The State's approach also allows for consideration of an "alternate minimum data[est]"; however, few details are provided that describe what constitutes the minimum dataset for an assessment. The assessment methodology currently focuses on describing the "upper" bar or ideal amount of continuous data that would be collected but does not describes what constitutes the "sufficient minimum" amount of continuous data. That decision is left open to interpretation.

Response: Minimum data requirements provide an absolute minimum amount of data that can be considered in the final document. Recommended data collection should be followed when feasible during development of monitoring plans.

Comment 19: We recommend DEQ consider running a number of "simulations" on existing datasets to evaluate whether the monthly average concentration changes if derived using 3 days of continuous data versus 7 days versus 30 days, etc. This information could provide a basis for the selection of the minimum required dataset for SC. Additionally, variability that may be introduced through rain events could be addressed by requiring a certain number of samples be collected during high flow events.

Response: See Response to Comment 16. DEQ revised language in the assessment method to include seasonal based high flow requirements and the variability associated with these events. DEQ completed relative percent difference (RPD) simulations of selected monthly average conditions on the Tongue River to running 3 and 7-day averages. Based on the outcome of the analysis we stand by our approach to calculate monthly averages.

Comment 20: The assessment methodology seems to give equal weight to continuous, grab samples and probe measurements. EPA's understanding is that the primary source of variability in specific conductivity is rainfall or runoff. Continuous data are more likely to capture this variability but will also capture baseflow conditions. Please explain DEQ's reasoning for weighting discrete (grab and meter) data equally to continuous data, when discrete data will likely only capture baseflow conditions and possibly provide redundant information? Instead, we recommend DEQ consider giving more weight to quality continuous data over discrete samples, if available.

Response: DEQ also understands that the primary source of variability is rainfall or snowmelt. The department has added language to clarify the data compilation process. This clarification gives quality continuous data more weight than previously calculated, but gives equal weight to any sample that meets data quality objectives. Previously, data collected (discrete or continuous) on any given calendar day was averaged together. With new guidance, discrete data are averaged with the closest continuous datapoint collected i.e., discrete data collected at 9:00am will be averaged with a continuous datapoint collected at 9:00am on the same day. With this timeframe of averaging, continuous data are weighted appropriately. Additionally, SC is not as variable within a day when compared to many other field meter measurements, so weighting toward a specific part of a day with a discrete sample added to a continuous data set should not skew a daily average.

Comment 21: Please explain DEQ's rationale for blending discrete and continuous data versus treating them as separate datasets?

Response: The department understands and acknowledges that continuous data collection comes with its challenges and data logger failure can occur. Therefore, it is imperative to allow for a blend of discrete and continuous datasets in order to allow for increased likelihood of having sufficient EC and SAR data. This approach also avoids parallel approaches for different data acquisition types (continuous vs discrete) with potentially disparate outcomes for the same parameter.

Comment 22: We recommend DEQ consider modifying the assessment method to apply the following approach: Step 1. DEQ completes a QC review on the available continuous data to determine if the data

are credible and high quality. Step 2. If any data point (continuous or discrete) exceeds the "do not exceed" criterion, the waterbody is listed. Step 2. If there are no exceedances, then the continuous data are used for the assessment decision. Step 3. Review any discrete data to see if they would result in different assessment decisions. Step 4. Complete assessment.

Response: DEQ updated the assessment method step by step guidance. See Section 5.3.2

Comment 23: Section 5.1.1. Step 6 states: "within each calendar month, identify samples that are temporally independent."

Please explain how this step is completed. For example, for grab samples or meter results, is it assumed (based on the previous Durbin-Watson tests) that samples collected more than 3 days apart are "independent"? If so, that assumption should be clearly stated. Or, is a new Durbin-Watson analysis used to make this determination?

The Durbin-Watson analysis was completed using data from another watershed. Does DEQ plan to repeat the analysis for the Tongue and Powder watersheds?

Why is the State concerned with temporal independence with discrete data and not continuous?

Response: For discrete datasets, partial continuous datasets, or any combination of these, determining if data are spaced out to represent each week of the month is necessary to include representation for an entire calendar month. This temporal independence overcomes the likelihood of misrepresenting the variability over a month's time by only collecting data during a fraction of the month (for example: 3 or 7 straight days). Temporal independence is not applicable for full continuous datasets because the whole month is represented on a constant interval and all data points are therefore equally weighted during averaging and the whole month is represented. Although a percent relative difference analysis was completed, new Durbin-Watson tests will not be completed for other watersheds that are identified in this assessment method. The Durbin-Watson test was only completed to determine the minimum spacing between samples that can be applied to the monthly average calculation.

Comment 24: Step 6 also states "Within each calendar month for all data, identify samples that may be spatially independent (i.e., sites that are longitudinally separated along the assessment unit or reach)." Please explain DEQ's process for determining spatial independence and identify the analytical methods used to determine spatial independence.

Section 5.1.1. also states: "If spatial samples are clearly dependent, those values are grouped." Does this mean that those values are averaged? Please explain what is meant by "grouped."
Step 7 states "Review data site by site to see if any spatial patterns emerge (at a given point in time, DEQ would generally expect relatively consistent values throughout an assessment unit); if unexpected spatial patterns emerge (e.g. two adjacent sites within an assessment unit contain noticeably different values), consult the section manager or technical lead for guidance. This step may result in creating reaches within the assessment unit."

This language suggests that results are analyzed for each monitoring site. However, other places in the document refers to analyzing the data at the assessment unit scale. Please clarify.

In addition, please elaborate on DEQ's process for determining if there are "significant shifts in type and intensity of potential SC or SAR sources" and that "sub-segmenting may be justified." From the process described in Section 5.1.1, it is unclear how these shifts would be determined.

> **Response:** DEQ revised the document to clarify how to create reach breaks. All data within an analysis unit (segment, or reaches), would be combined.
>
> Data are first reviewed site by site to ensure the assessment unit EC and SAR conditions are homogeneous and reach breaks are not needed. Reach breaks may be warranted if differences between sites are identified, especially relating to standards comparison. A major factor for creating reaches would be if areas of a waterbody are above a standard or assessment thresholds, while others are not. In this case, averaging a waterbody wide condition may not be appropriate. DEQ assessment methods should set the stage to protect or restore all areas of a waterbody. Factors that are considered during this process are land use changes, major tributaries, isolated weather events or point sources. After this initial spatial analysis, the data are assessed at an appropriate scale, either waterbody wide or at a reach scale, whichever is appropriate. DEQ has updated the document to reflect the reach break approach. The minimum data requirements and data analysis steps have been updated to address how to group and analyze data. Assessment units and reach definitions are reviewed in more detail within DEQ's Beneficial Use Assessment for Montana's Surface Waters document.

Comment 25: Step 8 recommends "group[ing] results that are temporally dependent." Please clarify how dependent results are treated. Are they averaged together? If so, we recommend revising the language to make this clear.

For continuous data, DEQ assumes that the samples are temporally dependent. How important is it to having temporally independent samples for grab samples if continuous data are considered dependent and may represent the majority of the dataset?

If continuous data are temporally dependent, does it make sense to calculate an individual daily average or even a weekly average? Please explain how are those data treated.

> **Response:** The assessment method monthly data aggregation and data analysis assume that some samples may be temporally dependent, and others may not. The averaging over short intervals (days) to moderate intervals (weeks) to the ultimate interval (month) addresses collection interval variation along with the time dependence issues.

Comment 26: Section 5.1.1. currently focuses on the methods used to aggregate the data but does not provide explicit details for comparison to the water quality standard. For example, steps should be added at the end that indicate if a single monthly average sample is exceeded, the assessment unit is considered impaired.

> **Response:** Section 5.1.1 is not intended to provide information on data analysis or assessment decisions. References to the appropriate section(s) where data analysis and assessment decisions are made will be provided.

Comment 27: Table 5.6 provides significant detail on the process for aggregating data for each assessment unit. Adding an overarching description of the process would be helpful. Could the

approach be simplified to state: “daily average values are calculated for each monitoring location. If more than one monitoring location exists in an assessment unit, those station daily averages are aggregated resulting in one daily average per assessment unit. Then, weekly averages are calculated from the assessment unit daily averages, and monthly averages from the resulting weekly averages.”

> **Response:** Additional overarching descriptions will be included into the final EC/SAR assessment method. Table 5.6 was deleted to provide a more streamlined document.

Comment 28: Section 6.1.1. reads “An assessment unit will be delisted for EC if there are no exceedances of either the monthly average EC standard or the no sample may exceed [the] standard over the ten-year period.” Should that be revised to say “no sample may exceed the standard over a 3 year period?”

> **Response:** Section 6.1.1 will be revised to add clarification to delisting decisions. With new consideration about minimum data requirements, the guidance will be, “No sample may exceed the standard over a two-year period, and at least one third of the dataset must include high flow samples.”