FINAL TECHNICAL MEMORANDUM

TO: Dan Heilig, Wyoming Outdoor Council and
Jill Morrison, Powder River Basin Resource Council
FROM: Jean Marie Boyer, PhD, PE, Hydros Consulting Inc.
SUBJECT: Review of ERM Water-Quality Modeling Study of Boysen Reservoir
DATE: July 1, 2018

BACKGROUND

Per your request, I have reviewed the report entitled "Water Quality Compliance Analysis for the Long Range Development Plan at Moneta Divide, Wyoming. A Hydrologic, Hydrodynamic and Water Quality Study of the Boysen Reservoir Watershed" written by Environmental Resources Management (ERM), dated April 23, 2018 (Report). My review focused on the development of the Boysen Reservoir Water-Quality Model developed using GEMSS and the analysis of results. I did not focus on the SWAT modeling, which was conducted to develop daily flows for use in the reservoir model.

My staff and I also briefly reviewed reservoir modeling files, sent to us by ERM. These files provided more detail than what was described in the Report. Given the lack of model documentation and time / resource constraints, model files have not been thoroughly reviewed. However, the review resulted in the identification of several severe and alarming issues, and there may be more.

My comments are summarized in this memorandum and organized under two broad categories:

- The Reservoir Model Cannot be Used for Decision Making; and
- The Compliance Analysis Methods and Findings are Incorrect.

Many of the comments are supported with detailed examples and they are not in order of most important to least important. An overall summary can be found at the end of this document, where the most important concerns are highlighted.

MAJOR POINT: RESERVOIR MODEL CANNOT BE USED FOR DECISION MAKING

It is very clear that the model developed by ERM cannot be used for decision making. Several comments are made below and they are divided into two categories of reasoning.

1. The Model Was Not Developed Properly; and
2. Model Performance Was Not Appropriately Evaluated.

Reason 1: Model Was Not Developed Properly

Numerical reservoir water-quality models require numerous types of detailed inputs. This is especially true if one uses a 3-dimensional (3-D) representation of the reservoir, as was chosen by ERM. Issues associated with data inputs, assumptions made, and "adjustments" used in model development are highlighted below and grouped by type of assumption.

Water Balance Assumptions

A complete and representative water balance for a reservoir is important when modeling its water quality. Boysen Reservoir outflow records are good and the best inflow records are for Wind River above the reservoir and Five-Mile Creek. Distributing the inflows correctly is a critical aspect of modeling water quality in Boysen Reservoir, given the wide range of inflow water quality characteristics in the watersheds of this very large reservoir. A tributary with a low flow and poor concentrations can add a significant load to the reservoir, relative to other sources.

1. Little Data, Yet No Flow Data Collection

Aethon spent 5+ years collecting data to support the analyses needed for project approval, yet chose to focus water-quality data collection at a location with a significant amount of data (below Boysen Reservoir)[1]. Aethon did not collect any flow data to ground-truth the distribution of flows among the 9 simulated tributaries. Therefore, many of the tributaries represented by SWAT-generated flows were uncalibrated and highly uncertain. This could have been avoided.

2. Reservoir Evaporation was Ignored in the Water Balance

Evaporation is an important component of the water balance, especially for Boysen Reservoir. Given its surface area and location, evaporation is significant (on the order of 50,000 AF/year[2]). Correctly accounting for evaporation is important when modeling reservoir water quality in that the process of evaporation tends to increase in-reservoir concentrations (constituents are not removed with the water that is evaporated). If a modeler lumps this into other outflows, the model will unrealistically remove constituents with the outflow. The model as delivered to Hydros by ERM is set up to not include evaporation, as indicated by the model setting in Figure 1.

[1] Aethon did take tributary water quality samples on one day in April 2017

[2] Based on reservoir surface area and NOAA (1982)

Figure 1. Screen Capture from Model Setup Interface Sent to Hydros

3. <u>Flow Adjustments Made to Badwater Creek</u>

Because the flows simulated by the SWAT model for Badwater Creek (above the produced water discharges) were so poor, they were decreased and re-distributed (described in Appendix D of the Report). This redistribution was based on comparisons made to historic data and a basin-wide water resources planning model. The differences with respect to the planning model were added to four other tributaries (Birdseye, Cottonwood, Tough, and Unnamed Creeks) – apparently selected since "they have the greatest uncertainty compared to larger creeks that were previously well-calibrated with reliable flow."

- Note that there are two other tributaries with no flow data - Poison Creek and Muddy Creek. ERM chose not to re-distribute flow to these tributaries, yet they also have the same level of uncertainty.
- The four tributaries chosen for flow increases as a result of this adjustment have the best water quality (using ERM assumed concentrations.

Also, ERM notes "the amount of flow redistributed and load increases were considered small."

- If the redistributed loads were "considered small", they would not have had the effect mentioned in Appendix D of the Report. ERM notes "These changes highly benefitted the overall water quality calibration of Wind River Below Boysen Reservoir."

In addition, simulated flows from other ungaged tributaries, were not compared to the planning model and adjusted in the same manner. Thus, tributary flows were treated inconsistently.

4. <u>"Adjustments" Made to Reservoir Inflows from Wind River</u>

The model was set up to "auto-calibrate" the water balance to user-provided surface water elevations (SWEs). Thus, model inflows were adjusted so that the observed SWEs were simulated. When flow adjustments were needed to complete the water balance, flows were from the Wind River above the reservoir were adjusted. This is the site with the most certainty for inflows (along with Five Mile Creek), yet ERM made adjustments at this location.

Meteorology Assumptions

Meteorological model inputs are important for correctly simulating reservoir hydrodynamics and mixing. Of the several meteorological inputs to the model, wind plays a particularly key role and needs to be characterized correctly.

5. <u>Wind Speeds Were Significantly and Unrealistically Capped</u>

The Wind River basin experiences high wind conditions, as displayed in Figure 2 where wind speed is reported in knots. The model as delivered to Hydros by ERM is set up to “cap” wind speeds to a maximum of 5 m/s (9.7 knots) during the simulation, as indicated by the model setting in Figure 1 above.

Artificially reducing the wind speed serves to reduce mixing and increase stratification (something the model has troubles simulating). It appears that the modeler used this cap to make up for other important model development problems. This significant adjustment was not described anywhere in the Report and the reader is led to believe that the values shown in Figure 2 were used.

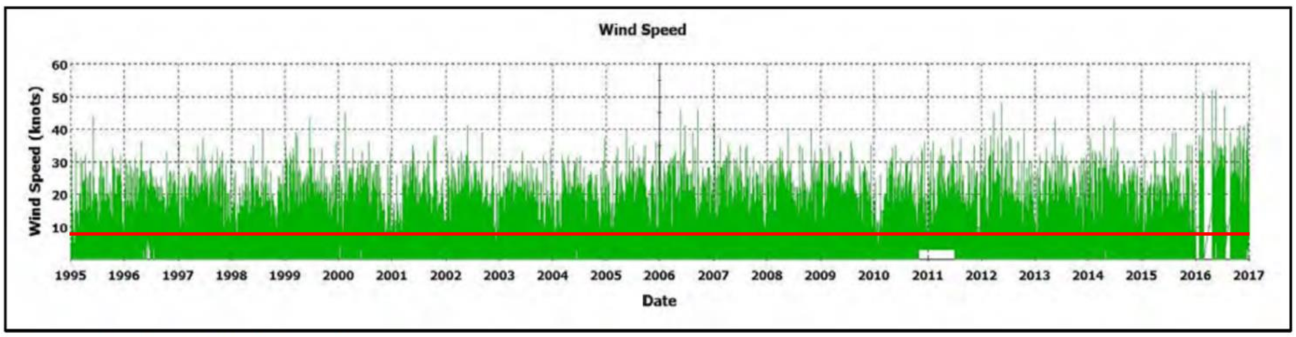

Figure 4-18 Wind Speed

Figure 2. Wind Speed Figure from the ERM Report. Red line added at 5 m/s (9.7 knots)

Inflow Water-Quality Assumptions

It is clear that the water quality of the various inflow sources varies considerably. Thus, it is important to base inflow water-quality assumptions on the best available data. Sometimes, additional data collection is necessary to develop a useable model. This should have occurred for this effort. Some of the assumptions made to make up for the lack of data are described below.

6. <u>Surrogate for Badwater Creek and Lack of Data Collection</u>

Water-quality characteristics of water flowing into the reservoir from Badwater Creek are obviously critical for this effort. It is surprising to know that Aethon spent 5+ years collecting data to support the analyses needed for project approval, yet only collected tributary inflow water-quality data on one day in April of 2017. These often-single data points are the basis for many of the inflow WQ assumptions. And yet, in the case of Badwater Creek, ERM used the one sample from Tough Creek, as a surrogate for conditions upstream of produced water discharges. There is no reason to believe that the water quality in Tough Creek is similar to that of Badwater Creek and no reason is provided as to why sampling did not occur at such a critical location. Again, this could have been avoided. Inflow water quality at numerous locations over time needs to be measured to be able to consider the impacts of the project. Current available data are insufficient.

7. <u>Questionable Use of Method for Quantifying Inflow Concentrations</u>

For some constituents, ERM used the WRTDS (Weighted Regressions on Time, Discharge, and Season) method to describe inflow concentrations for Wind River above the reservoir, 5-Mile Creek, and Muddy Creek (see Appendix F of the Report). The results from using this methodology are questionable and unrealistic in some cases. ERM subjectively capped what was determined to be excessively high concentrations. In addition, odd results sometimes occurred due to extreme values in a single or few data points and certain trends were created that are not described or justified. Examples are shown in Figures 3 and 4. Note that Wind River provides the majority of the inflow into the reservoir (over 70% according to Appendix C of the Report) and its water quality is an important driver of in-reservoir dynamics.

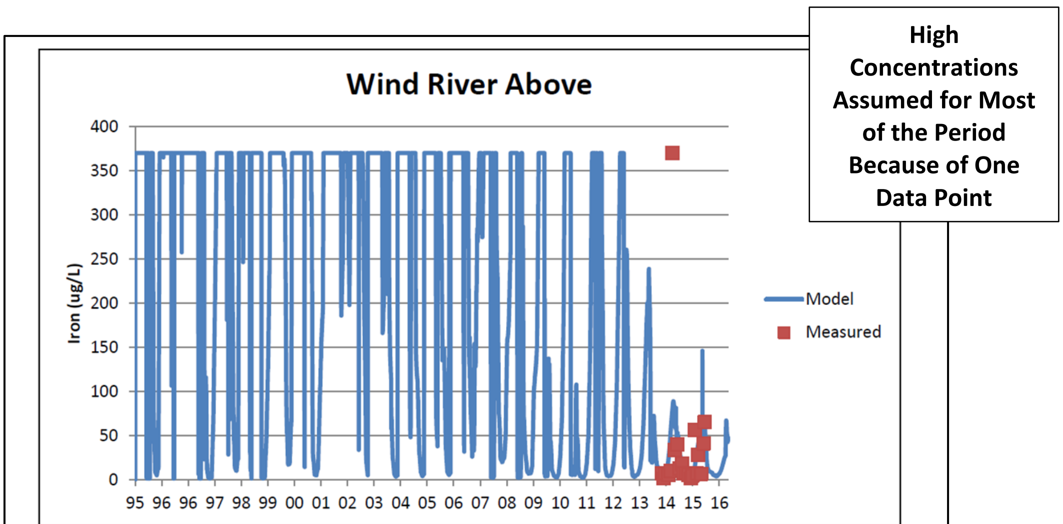

Figure F 12 WRTDS Model Compared to Measured Values for Iron in Wind River Above

Figure 3. Assumed Iron Concentrations at the Wind River Above Boysen Reservoir (from Report)

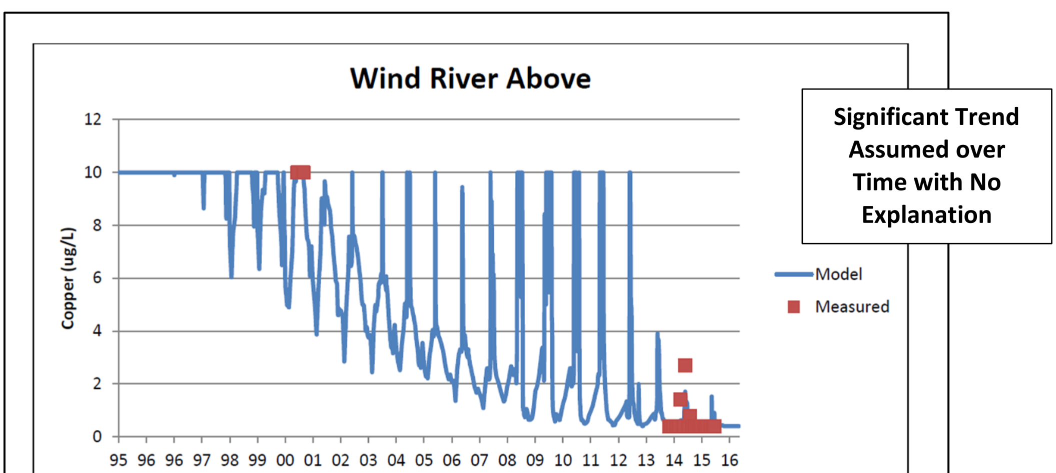

Figure F 9 WRTDS Model Compared to Measured Values for Copper in Wind River Above

Figure 4. Assumed Copper Concentrations at the Wind River Above Boysen Reservoir (from Report)

8. Assumption of Permit Concentrations

In several instances, ERM assumed that the water quality of the produced water was at permit limits for the calibration and validation period. This could be far from actual conditions during the 22-year period. The purpose of calibration and validation is to recreate what actually happened. Using permit limits for calibration needs justification.

9. Dividing up Badwater Creek into Four Sources

It is very odd that ERM chose to separate the flows into Badwater Bay into four distinct sources (Badwater Creek above Alkali Creek, Burlington, Aethon, and Neptune) and have them all entering the same location of the model grid. It is even more confusing to know that some level of treatment at Neptune has been occurring historically but treatment details and flow amounts over time are not described in the Report. Nor can this information be inferred from the model input files. In addition, samples exist for Badwater Creek ~ 5 miles above the reservoir (where the sources are already mixed and is more representative of what is actually flowing into the reservoir) and these samples are not considered by ERM. For calibration, it is important to capture the blended source of water entering the reservoir at this location. It is unclear why ERM developed the model in this manner, when it could have been considered in a more straight-forward way.

Inflow Placement into the Reservoir

Tributaries can enter a particular reservoir differently depending on the density of the inflowing water (Figure 5). The higher the salinity, the higher the density of the inflowing water. Since produced water has very high salinity, it is important to capture inflow placement dynamics correctly for this effort, given:

- the increase in density of the inflow water at Badwater Creek Bay with the proposed project;
- the potential for the diving of inflows as an underflow; and
- the low-level outlet works at the dam.

<u>Thus, there is the potential for impacts to the releases downstream that exceed average impacts in the reservoir.</u> Most commonly-used hydrodynamic reservoir water-quality models simulate these types of dynamics.

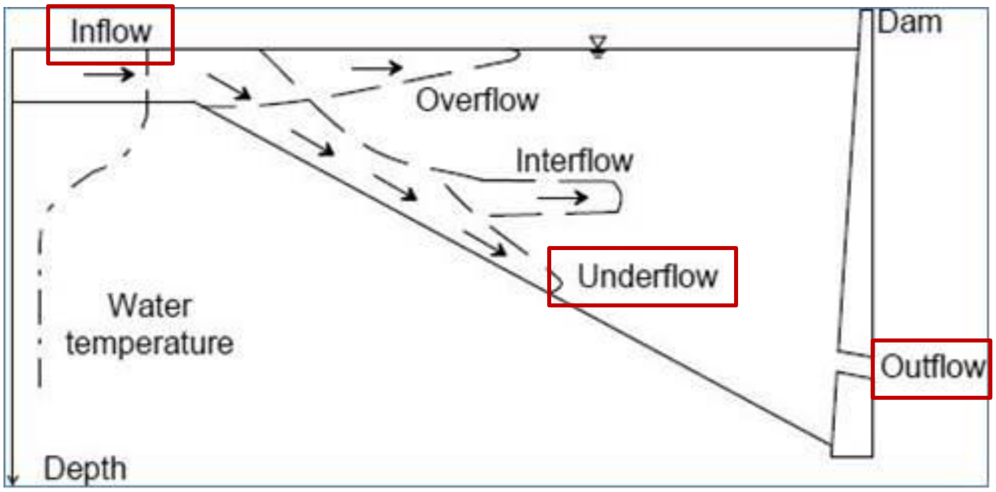

Figure 5. Generic Reservoir Graphic Showing Density Currents and Possible Inflow Patterns

10. <u>Mischaracterization of Inflow Placement</u>

Although ERM describes the importance of water density and transport processes, the model was not set up to distribute inflows vertically based on the density of the inflow and the density profile of the water in the reservoir. Instead, tributary inflows enter each 2-foot layer of the model grid uniformly[3]. Thus, changes to the density of the inflows (through salinity and temperature changes) from the project <u>do not</u> correspondingly change the vertical distribution of the inflows in the model. This is a serious flaw to the model as flows into Badwater Bay will tend to enter the reservoir lower in the reservoir with the project. This may affect water released at the dam through the low-level outlet differently than it has historically. Also, inflow placement assumptions made by ERM are not described in the Report (as they should have) and were only determined based on review of model files.

Representation of Reservoir Releases

Releases from Boysen Reservoir to the Wind River occur at two different locations. Flow through the low-level outletworks (at 4,657 feet; USDOI, 1981) provides water to the penstocks for power production. This is the dominant means of withdrawal due to the potential to generate power. Spilling of water near the top of the reservoir can occur if the SWE is above 4,700 feet.

Water leaving through the outletworks (OLW) can have very different characteristics from water leaving via the spillway, due to vertical variations in water quality characteristics, especially during the stratified period (Figure 6). Thus, outlet operations have a direct impact on water quality in the Wind River below Boysen Reservoir (Class I). Most 2-dimensional (2D) and 3-dimensional (3D) reservoir models (including

[3] The control file specifies that for each inflow, the vertical limits are the bottom of the reservoir at the location of the inflow and the water surface. There are no options in the user interface for setting up the control file to select or determine if the placement of the inflows within these boundaries is uniform or density-based. However, based on review of the snapshot output files (e.g., the file received for the calibration run output named "Final Calibration_Restart.snp"), it was clear that the flows output by the model in the "Discharge Boundary Condition" section of the snapshot file that correspond to the inflows are uniformly distributed in the vertical direction.

GEMSS) have the capability to compute a withdrawal zone from which only certain layers contribute to the outflow, based on each structure, each outlet flowrate, and in-reservoir water density. This methodology has been developed to replicate how water is physically discharged from a reservoir. The model can then take that information and output the resulting water quality in the downstream river.

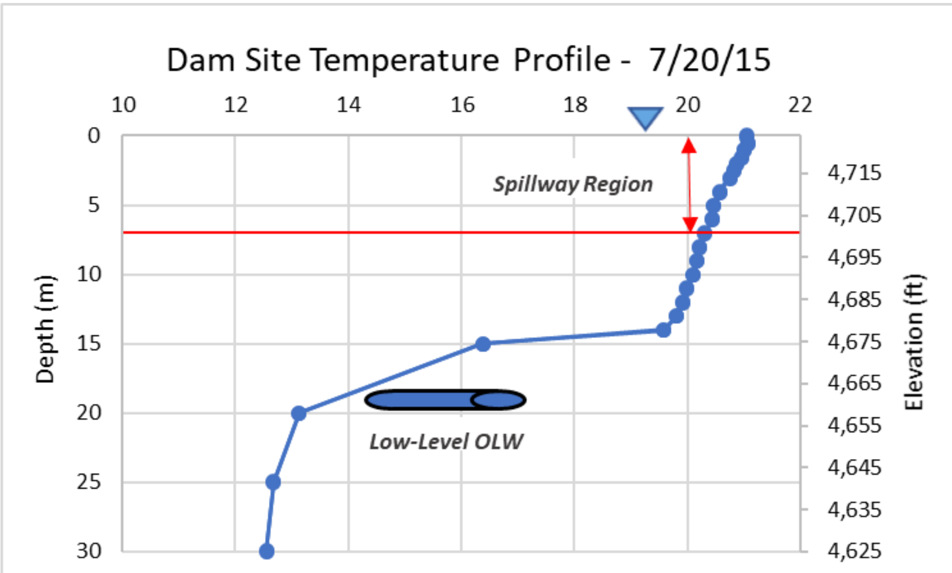

Figure 6: Temperature Profile Showing Stratification and Elevations of Releases (OLW = Outletworks)

11. Mischaracterization of Reservoir Release

Although the GEMSS model software includes the ability to characterize different structures and compute withdrawal zones, the modelers chose to release water uniformly in the vertical direction within each column - from the top layer of the reservoir to the bottom layer[4]. Thus, there is no differentiation between the outletworks and the spillway and outlet operations that control downstream water quality are completely ignored. Again, this is a serious flaw. Also, the assumptions made are not described in the Report and were only determined based on review of model files.

Reason 2: Model Performance Was Not Appropriately Evaluated – Erroneous Conclusions Reached

After the model was completed, ERM compared the results to certain targets to show that the model was calibrated, validated, and adequate for use to prediction of future conditions with the project. There are several instances where misleading information is provided. The reservoir model cannot be

[4] For the outflow, the control file specifies that the vertical limits are the bottom of the reservoir at the location of the outflow and the water surface. There is an option to choose either density placement or area-based placement of the outflow within these vertical boundaries. The area-based option was chosen in the control file, as provided. Review of the snapshot output also reveals that the area-based option is equivalent to the vertical uniform distribution of flow for each column of cells where the outflow takes place. Because the outflow takes place in two columns of cells located at the dam of the reservoir, and one column is deeper than the other one, the net vertical distribution of flow is not completely uniform. It is uniform from the water surface to the bottom of the shallowest column and between the bottom of the shallowest column and the bottom of the deepest column. However, overall, there is more flow coming from the upper layers than from the bottom layers, when in reality, more water is likely to flow out from deeper sections due to the low-level outlet location.

considered to be calibrated or adequate for simulating water quality in-reservoir or downstream in the Wind River.

Evaluation and Reporting of Wind River Simulation Results

The focus of the work conducted by ERM is to ensure protection of the downstream Class I segment of the Wind River. Analyses were conducted to determine produced water flows that would meet antidegradation requirement at that location. Thus, a very critical part of the analysis involves the quantification of model results for release water quality.

12. Evaluation and Reporting of Wind River Results are Wrong and Misleading

Through review of the model files, our team determined that the graphs displaying calibration and validation results for the Wind River below the reservoir are misleading and severely flawed. An example graph for temperature is shown in Figure 7. The top of the graph is labeled as “Outflow” and the caption says “Wind River Below Boysen Reservoir.” The data (green markers) are reportedly Aethon’s temperature measurements in the Wind River below the dam. The reader is led to believe that the blue line represents the temperature of the water released from the reservoir (via the low-level outlet and/or the spillway) and delivered to the river.

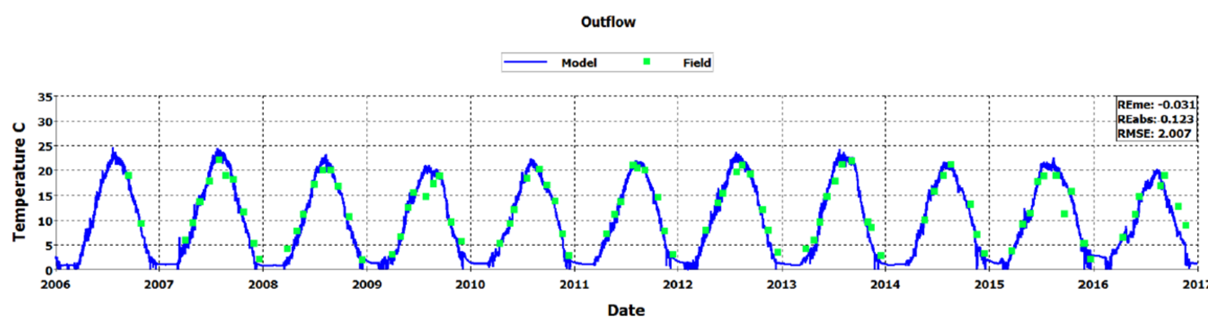

Figure J 3 Wind River Below Boysen Reservoir Temperature Calibration at Surface Level

Figure 7: Temperature Calibration Figure for Outflow (from Report)

According to the model files, the blue line actually represents the simulated temperature at the top model layer (~top 2 feet) of the most downstream location (at the dam). This is wrong and misleading and it is unclear why this was done. Note that ERM added “at Surface Level” at the end of the caption (Figure 7) and perhaps thinking this makes it not misleading, even though there is a low-level outlet used for power production?

This is a very serious problem since the water quality at the top two feet of the reservoir is being represented as Wind River water quality and there are observed (but not simulated) vertical variations in the reservoir. In reality, the water quality at the top of the reservoir is often different from the bottom of the reservoir[5] (see Figure 6). Since water is removed predominantly through the low-level

[5] See Figure 6 as an example for temperature. Many other constituents (e.g. iron, manganese, arsenic) often show significant differences in top versus bottom concentrations in a reservoir, especially during stratification.

outlet, Wind River water quality would generally reflect the water flowing through that outlet or a combination of lower level releases and spills[6].

An example of temperature variations in a stratified reservoir is provided in Figure 8. The location of the outlet is an important factor in determining the water quality of the river downstream. ERM mistakenly compared the simulated reservoir surface temperature (top 2 feet) to the samples in the Wind River.

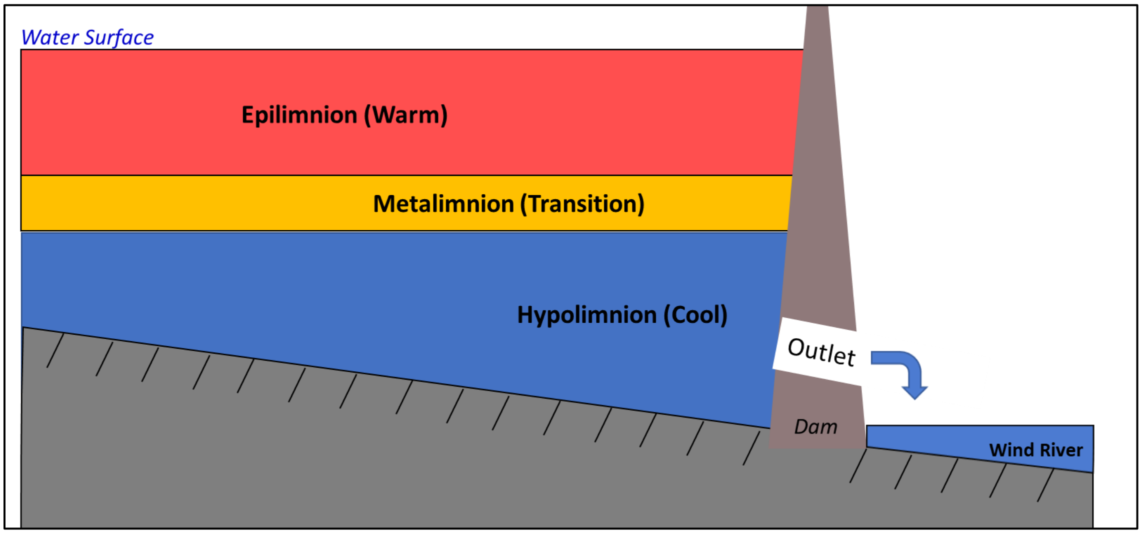

Figure 8. Temperature Differences in a Stratified Reservoir

If the modeler had differentiated between the outlets and simulated withdrawal zones, the release water quality output file would have reflected these dynamics. This was not done by ERM and incorrect and very misleading comparisons were made.

Choice of Observed Dataset Used for Comparisons

During calibration/validation, comparisons are made between observations and simulation results. Thus, the observed dataset used is important when evaluating model performance.

13. <u>ERM Removed Numerous Observed Data Points from Analysis without Justification</u>

There are several cases where measured data were removed from the analysis without justification. A few examples are highlighted below:

Removal of In-Reservoir Data

Table 5-2 of the Report includes a list of all available data for calibration and validation in Boysen Reservoir (Figure 9). A footnote at the bottom indicates that more than 300 data points were excluded after "thorough QA/QC." There is no discussion to justify the exclusion of all data associated with 15 constituents in the reservoir. The only data that were kept and considered were profile data (conductivity, pH, and temperature).

[6] Unless an outage or maintenance resulted in flow restrictions through the low-level outlet.

As a result, for the reservoir, there is absolutely no calibration or any ground-truthing of numerous constituents, including TDS, chloride, sulfate, and numerous metals.

Table 5-2 Available Data for Calibration and Validation within the Reservoir

Parameter	Total # of Stations	Total # of Data Points	Date Range	Vertical Profiles Available
Aluminum (µg/L)	0	0	N/A	No
Arsenic (µg/L)	1	4*	1999 only	No
Barium (µg/L)	0	0	N/A	No
Boron (µg/L)	0	0	N/A	No
Calcium (mg/L)	1	32*	1999-2001	No
Chloride (mg/L)	1	59*	1999-2006	No
Chromium (µg/L)	0	0	N/A	No
Conductivity (uS/cm)	10	1226	2002-2016	Yes
Copper (µg/L)	1	4*	1999 only	No
Fluoride (mg/L)	1	32*	1999-2000	No
Iron (µg/L)	1	4*	1999 only	No
Magnesium (mg/L)	1	32*	1999-2000	No
Manganese (µg/L)	1	4*	1999 only	No
Mercury (µg/L)	1	4*	1999 only	No
Nickel (µg/L)	1	4*	1999 only	No
Oil and Grease (mg/L)	0	0	N/A	No
pH	11	1404	1999-2016	Yes
Total Petroleum Hydrocarbons mg/L)	0	0	N/A	No
Total Suspended Solids (mg/L)	1	33*	1999-2001	No
Radium-226 (pCi/L)	0	0	N/A	No
Sodium (mg/L)	1	32*	1999-2000	No
Sulfates (mg/L)	1	32*	1999-2001	No
Sulfides (µg/L)	0	0	N/A	No
Total Dissolved Solids (mg/L)	1	61*	1999-2006	No
Temperature (°C)	10	1343	2002-2016	Yes
Total Hardness (mg/L)	1	33*	1999-2001	No

*These data points were excluded from the analysis after a thorough QA/QC process was performed

Figure 9: ERM Table Indicating that Over 300 Data Points were Excluded

Removal of Winter Data

Although water-quality impacts in the spring through fall period are very important, the winter period is critical. Due to low tributary flows in the winter, any produced water added will result in the highest % effluent in Badwater Creek (and highest changes in salinity, etc.), as it enters the reservoir. ERM chose to exclude winter data, with no valid justification. ERM states:

> *"temperature data overlapping with model predicted periods of non-zero ice thickness were excluded from the calibration and validation comparisons to field data. This is because grab sample measurements recorded during predicted periods of ice cover are highly uncertain. The uncertainty arises because these samples could have been taken from localized areas that may not have ice or may have been collected from below the ice cover. These field measurements did not contain such information and were deemed unsuitable for comparison to model results."*

The reasons given for exclusion do not make sense and this is unconventional. In fact, several studies focus on accurate modeling under ice-cover conditions and/or simulated conditions over a number of years and include data collected during ice cover (e.g., Brodzeller and McGinley, 2016; LimnoTech, 2016; Hydros Consulting, 2017). It is suspect that ERM chose to remove data from a critical period for this project.

The percent of produced water in the inflow from Badwater Creek into the reservoir (using flows from ERM input files) is displayed in Figure 10. Results from the calibration model run are shown along with the three compliance analysis cases considered by ERM. Large increases are seen in July and August and maximum levels are reached in December and January. The highest percentages occur in the winter months and reach values of over 90% produced water under Case 03 (the case considered in the Statement of Basis). These periods are when "maximum concentrations entering Boysen Bay and the reservoir" occur, as noted by ERM. We acknowledge that a portion of the water is to be treated, but also note that concentrations of several constituents are not reduced via treatment (examples include arsenic, chromium, nickel, magnesium, manganese, copper, sulfide, and mercury – Table 6-4 in the Report).

Winter conditions are critical for this analysis and ERM's exclusion of winter data is unwarranted and wrong.

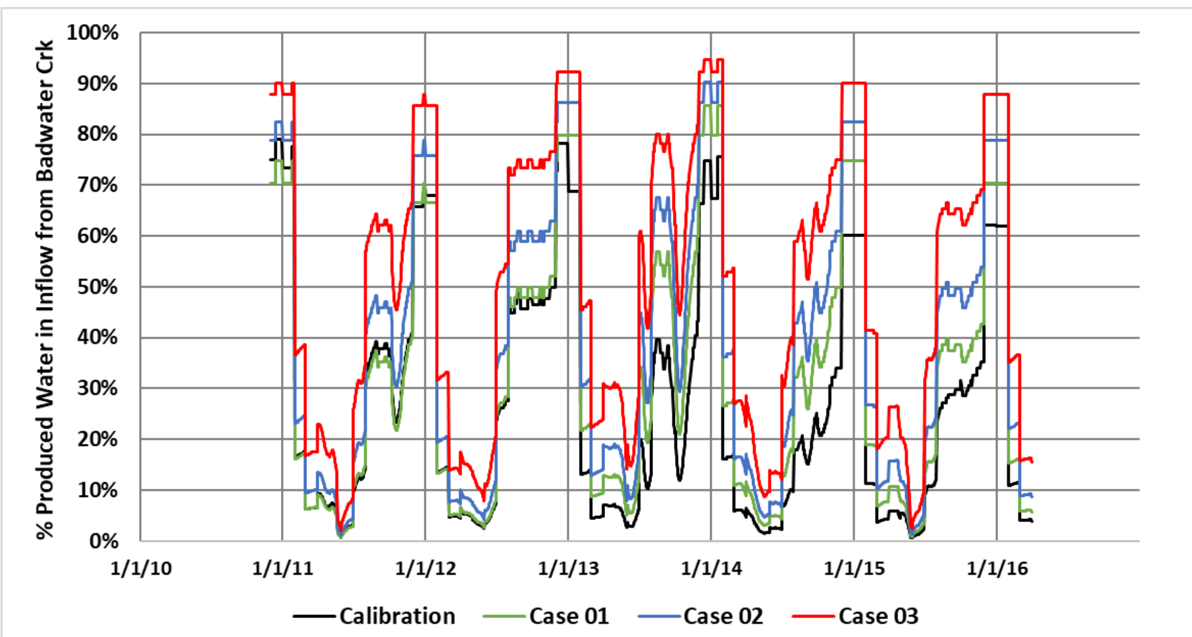

Figure 10: Proportion of Produced Water in the Inflow from Badwater Creek into the Reservoir

Reservoir Model Calibration Targets

Calibration targets are used to evaluate model performance and to determine if the model can be used for desired purposes. This is an important aspect of model development.

14. <u>ERM Used Overly-Lenient Calibration Targets</u>

In Section 5.4 of the Report, ERM describes the calibration targets used to evaluate the reservoir water-quality model.

> *"EPA-based metrics for evaluating watershed model performance (Donigian 2000) were used to evaluate the GEMSS model performance for important water quality parameters to the study."*

Although they are evaluating a <u>reservoir</u> model, ERM chose to use targets that were developed for <u>watershed</u> modeling, specifically HSPF. It is easier to more accurately simulate reservoir dynamics than watershed dynamics, due to the smaller spatial scale and greater homogeneity of the physical environment represented. Thus, calibration targets used for reservoir modeling are more stringent and should have been used for this effort. For example, developers of the well-used CE-QUAL-W2 model (from which GEMSS is reportedly based on) note that temperature simulations (important for simulating flow patterns accurately) should have an average mean absolute error within 1 °C. This means that the simulated model value is, on average, within 1 °C of the measured temperature. This target is met by numerous model applications of CE-QUAL-W2 (Cole and Wells, 2016 lists 70 applications in Table 4).

Given that the Boysen Reservoir is developed in 3 dimensions (versus a using 2-dimension assumption for CE-QUAL-W2 applications), one could expect the targets for Boysen Reservoir could be more stringent than the ones used in W2. Note that the commonly-used temperature target for reservoirs is not met by the Boysen Reservoir application (at least at the dam). This indicates that the model is not performing well enough to be called calibrated or adequate for making predictions.

In addition, ERM represents "% differences" in a manner that is highly unusual, dividing the mean of the RMSE by the average model prediction. It is unclear why this metric was created and used for this effort. In addition, ERM does not present the % differences computed. Only the final categories are presented (fair, poor, etc.) for a particular constituent. Thus, the actual % differences computed are not disclosed anywhere in the text, which results in lack of transparency.

Display of Results

Modeling results need to be complete and transparent. This is not the case for the Report reviewed.

15. <u>Information Was Concealed by Limiting Bottom Elevations Displayed on Profile Graphs</u>

In-reservoir observed and simulated results are shown in the Report in Appendices J and K for temperature, TDS, and pH with depth. All of the graphs provide data and results for elevations above 4,680 feet. This elevation is not at the bottom of the reservoir (at least near the dam) and cutting off the elevations in the figures leads the reader to assume that the reservoir is typically well-mixed summer and does not stratify or have much vertical variation. In addition, the model results show something similar. An example is shown in ERM's Figure K 39 (Figure 11) for July 30, 2002 near the dam, where the reservoir appears to be well-mixed with good model predictions (and hot from top to bottom).

However, the full profile of observed data indicates stratified conditions (Figure 12). In addition, Figure 12 shows that ERM failed to display about 90% of the observed profile. The bottom of ERM's model grid at the calibration location is also indicated, showing that the model grid is not deep enough near the dam. Nor is it deep enough to reach the lower-level outlet. This also highlights significant issues with the development of the model grid (which isn't deep enough to reach the lower-level outlet).

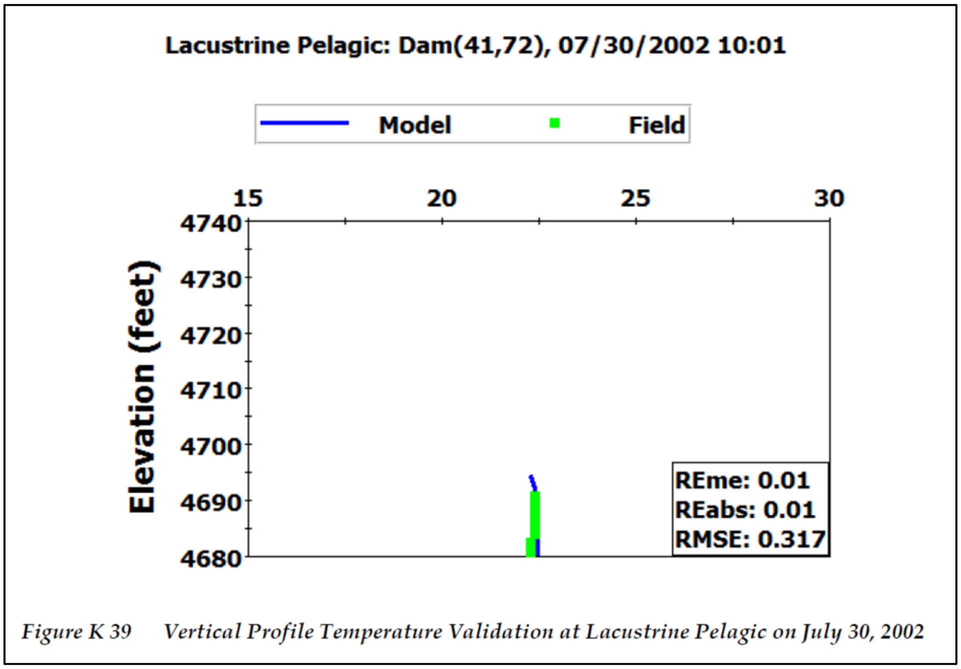

Figure 11. ERM's Figure Showing Observed and Simulated Temperature Profiles

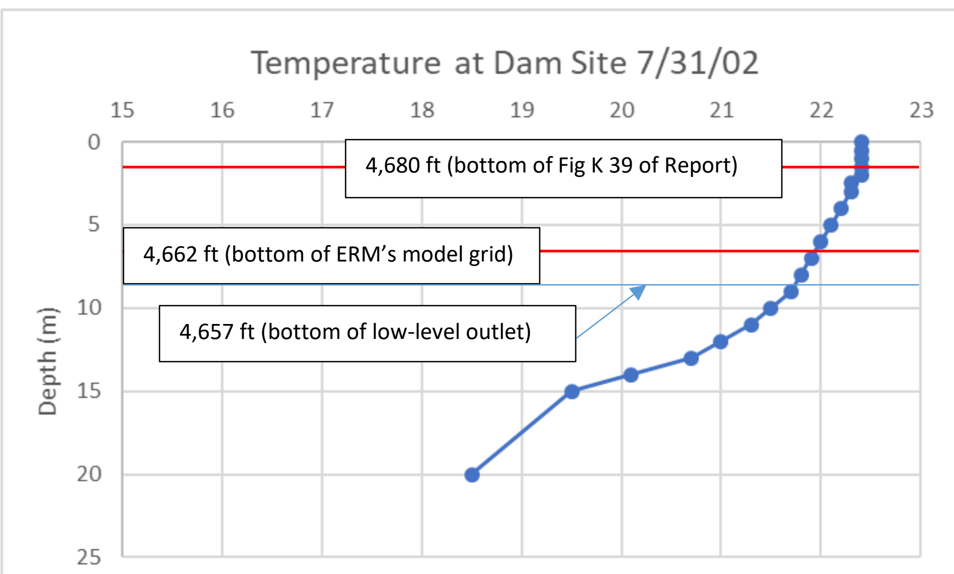

Figure 12. Full Profile at the Calibration Site Closest to the Dam

16. Information Was Omitted by Failing to Include All Profile Dates

In-reservoir observed and simulated results are shown in the Report in Appendices J and K for temperature, TDS, and pH. Several profiles were omitted, including temperature profiles near the dam for 2014-2016. It is not clear why this is the case.

Actual Model Performance

17. Simulation Results are Poor

Capturing observed flow patterns and hydrodynamics with the model is important. This is necessary to be able to use the model to predict conditions with increased flows at higher concentrations at Badwater Creek. Fortunately, a few temperature and specific conductivity profiles are available. Both of these constituents are good indicators of flow and thermal patterns and hydrodynamics.

Even though a number of adjustments were made during model development and calibration, the model results are very poor. Temperature profiles near the dam (Figure 13) show that the model is not capturing observed stratification in the summer and shows very little variation top to bottom. Reservoir temperature calibration is an initial and very important step in modeling. Recall that water in Boysen Reservoir is released to the Wind River via a low-level outlet (elevation 4,657 ft) and an upper spillway, at times. This makes it even more critical to be able to capture the vertical variations. As described earlier, using commonly-accepted calibration targets, the ERM model is not adequate for use. Also note that the bottom of the model grid at this calibration location is so high that water in the bottom 35 feet of the reservoir is ignored. Thus, water quality in this region (near the lower level OLW) is not even simulated.

Instances where modeled outflow temperatures to the Wind River correspond closely with observed temperatures downstream are strong indications of poor reservoir model performance. This is because the observed temperatures were compared to the temperatures at the top 2 feet of the reservoir near the dam, as described previously. The observed downstream temperatures should be the result of outflows that depend on release location (low level outlet, spillway), amount released at that location, and vertical density distribution. Thus, even when the reported model results seem to be acceptable, they are not generated as a result of a physically realistic simulation. This renders modeled predictions at the Wind River Class I segment unreliable.

Specific conductivity profiles are shown in Figure 14. Again, the vertical variations are not captured and sometimes the magnitudes are overestimated by 100’s of uS/cm. Note that most of these are in mid-summer, when the % produced water increases (see Figure 10).

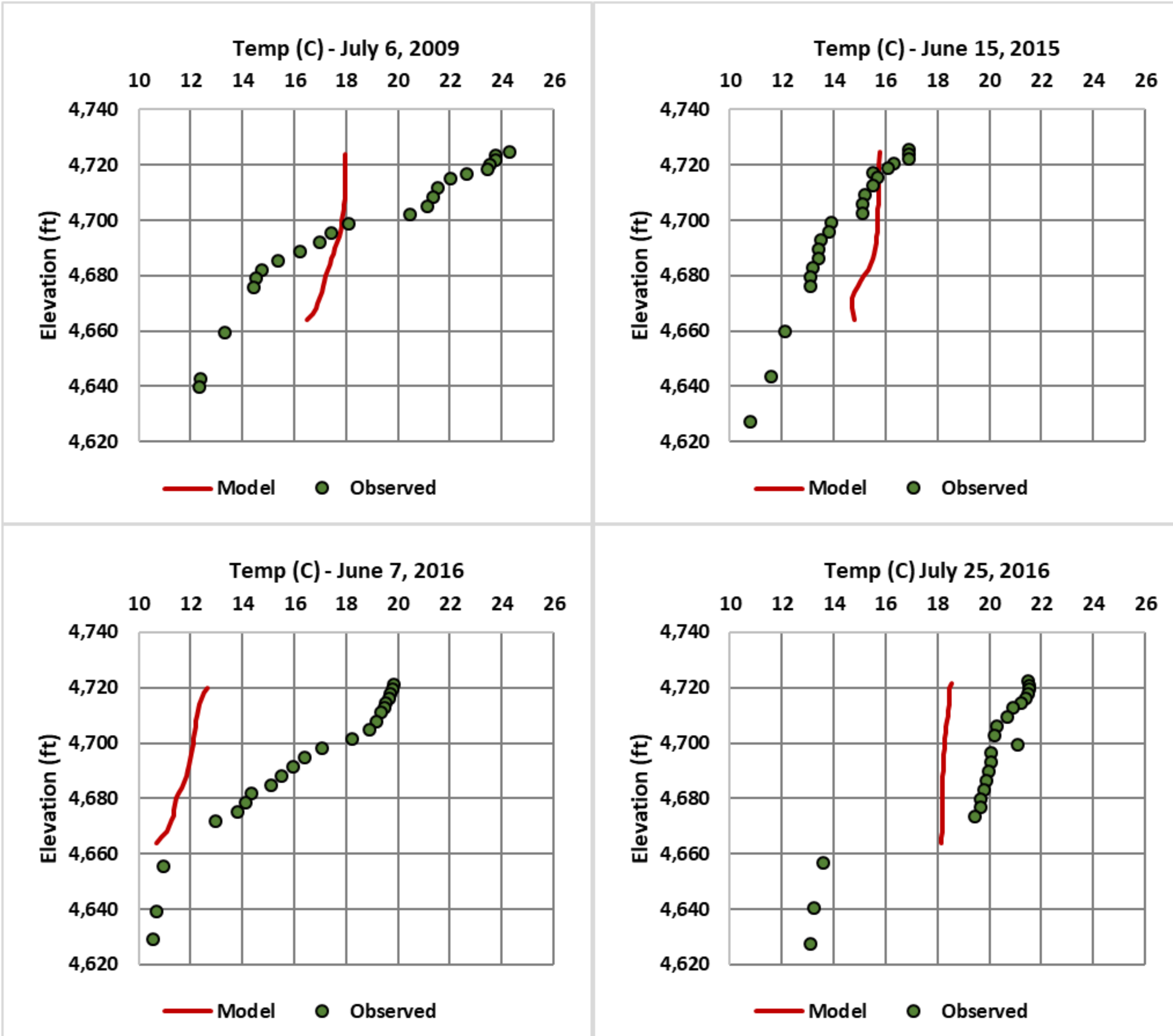

Figure 13. Example Temperature Profiles Displaying All Observations and Model Results from Top to Bottom of the Reservoir. Data from Lacustrine Pelagic: Dam Site

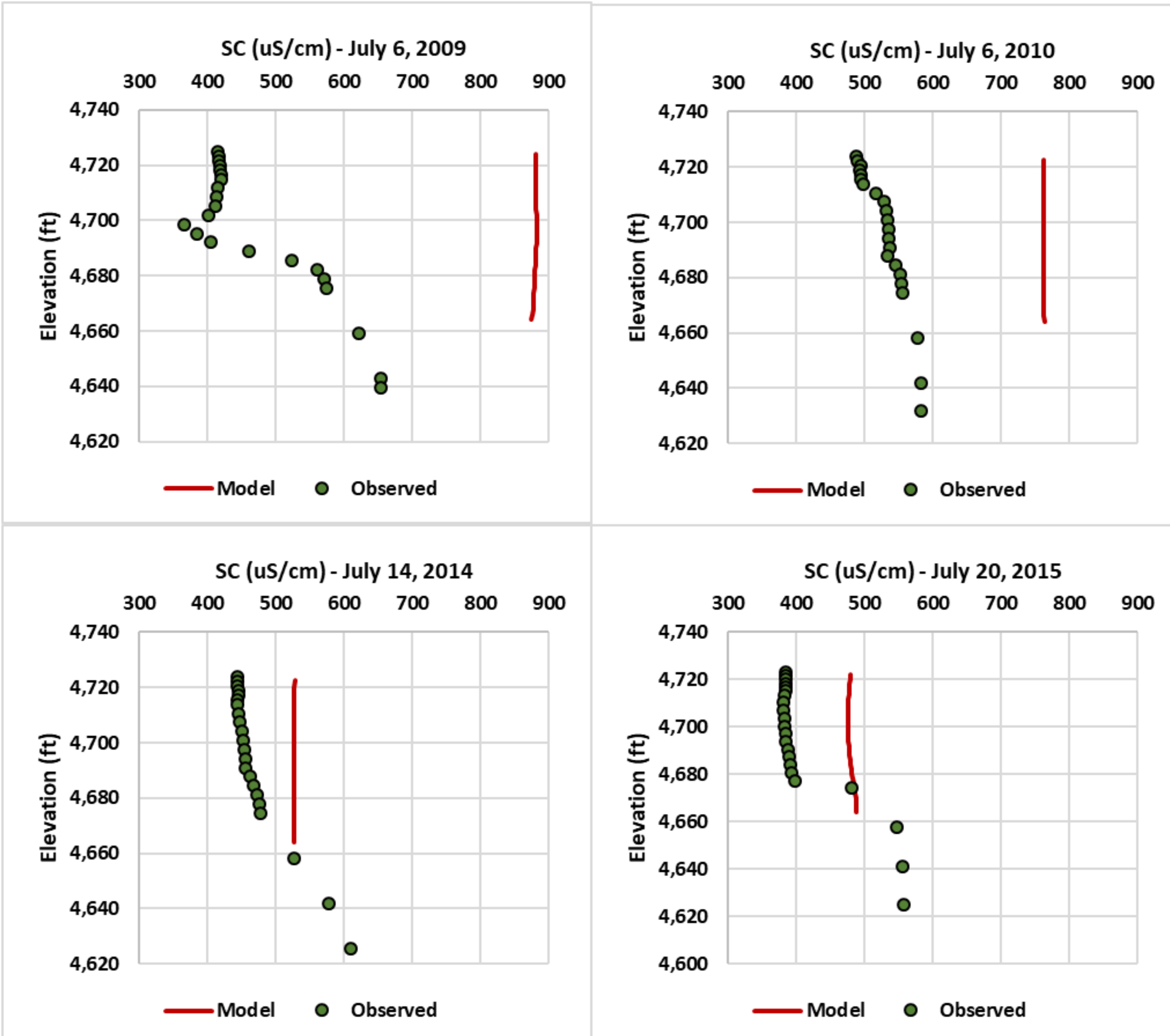

Figure 14. Example Specific Conductivity Profiles Displaying All Observations and Model Results from Top to Bottom. Data from Lacustrine Pelagic: Dam Site

MAJOR POINT: "COMPLIANCE ANALYSIS" METHODS AND FINDINGS ARE INCORRECT

Even if the reservoir model was developed adequately, the methodology used by ERM to evaluate compliance is severely flawed and is biased. Comments below are considered in 3 areas:

1. Data Used to Define Baseline Conditions in Class 1 Section;
2. How ERM Shows Compliance; and
3. Boysen Reservoir Antidegradation.

Data Used to Define Baseline Conditions in Class 1 Section

Baseline conditions in the Wind River are very important because these conditions are the basis for protection.

18. <u>ERM Failed to Use USGS Data for Defining Baseline in Class I Segment</u>

Only Encana/Aethon-collected data were considered when defining baseline conditions for the Class I segment of the Wind River. Yet, there are hundreds of approved water-quality measurements from the USGS below the reservoir for the period ERM defined as baseline (December 2010 – March 2016). Approved USGS data are considered to be of very high quality. In some instances, there are more data from the USGS in this period (Encana/Aethon did not report data for 8 months of the baseline period; Figure 15). In addition, there is much more variability in much of the data collected by Encana/Aethon than the USGS (see Figure 15 as an example). This variability would serve to increase a standard deviation.

In addition, the Encana / Aethon baseline data provided to Hydros Consulting did not appear to be raw data. The forms of the constituents were not noted (dissolved or total). The dates were often the 1st of the month and appear to be reported as a monthly value. The values could be averages or single points, and this is not clear. Aethon switched labs (going from "Lab 1" to "Lab 2") in November 2013. This resulted in an increase in detection limit for 12 of 14 metals, most of which were already below detection limits.

Baseline conditions should be defined using USGS data which is of higher quality and more complete, in most cases.

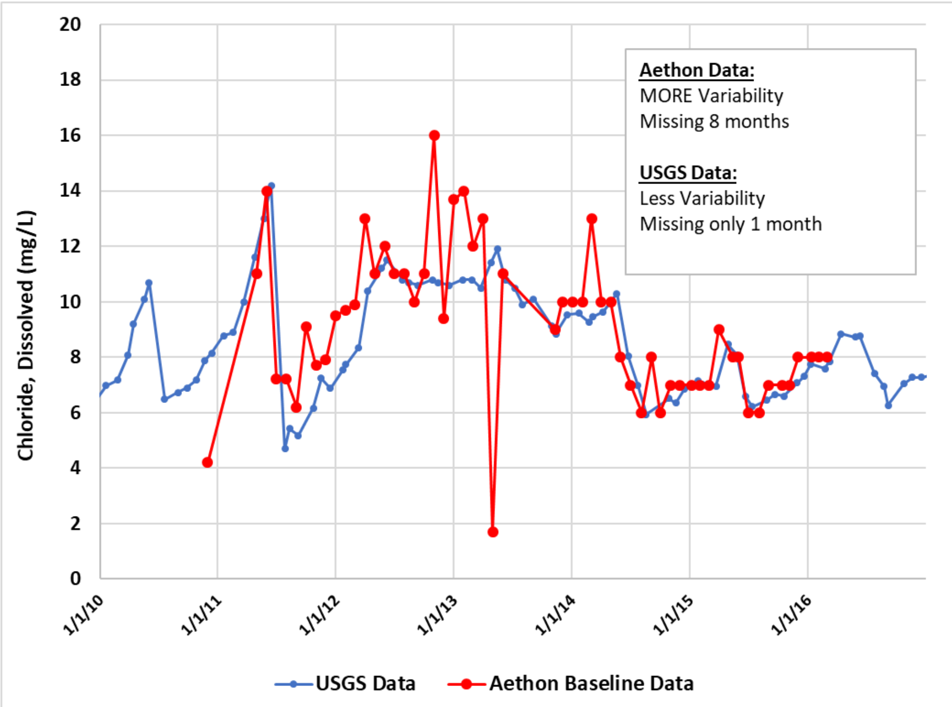

Figure 15: Chloride Observations in the Wind River below Boysen Reservoir (2010-2016)

How ERM Shows Compliance in Class I Segment

19. Used Monthly Averages, Obscuring Results

ERM chose to complete the compliance analysis on an average monthly basis. So, all Januarys are averaged together, Februarys are averaged together etc. This method lumps the data, reduces observed variability, and also serves to hide important differences that occur year-to-year, especially since time-series of the results are not displayed. This point is illustrated for chloride in Figure 16.

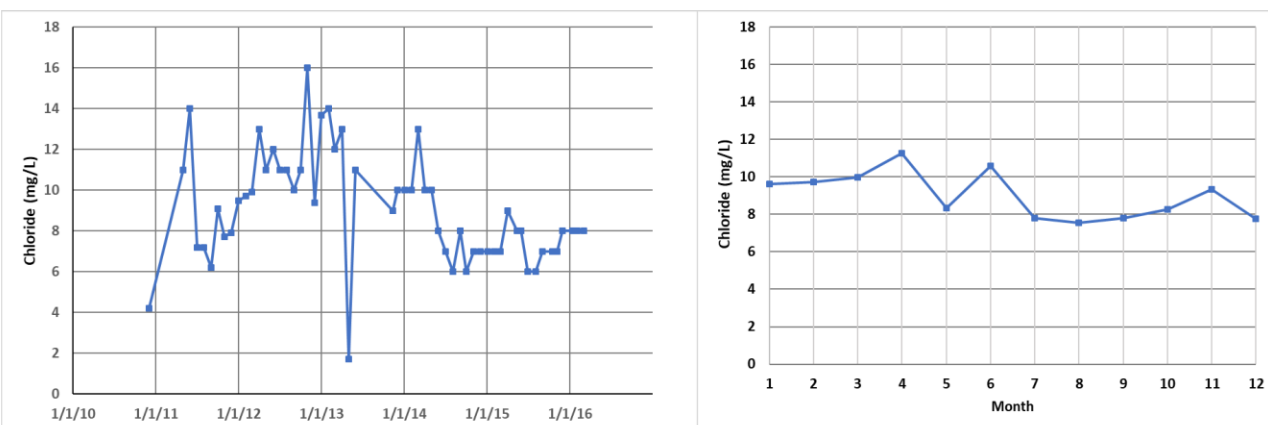

Figure 16: Chloride Measurements, Wind River Below Boysen Reservoir, Aethon Data. Individual Measurements (Left); Lumped Average Monthly Values (Right); Illustration of Reduction in Variability.

ERM did display chloride results in an attempt to justify the model "spin-up" period and to only focus on model results from December 2010 – March 2016 (Figure 17). Note that the project results in <u>lower</u> chloride concentrations in the "outflow" (top 2 feet of the reservoir) in 2010-2011.

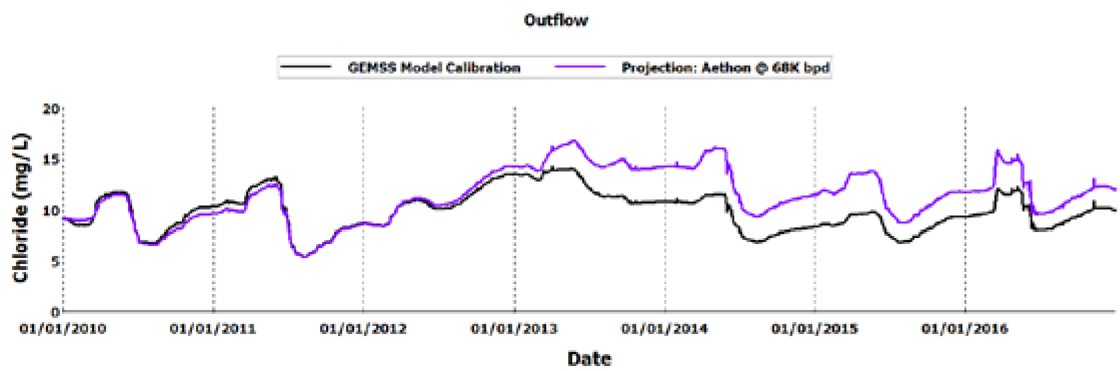

Figure 6-4 Projected Chloride Concentrations Compared to the Calibrated Model to Demonstrate the Model Spin-Up Period

Figure 17. Model Output for "Outflow" (Top 2 ft of the Reservoir); Model Calibration and Case 01

The results are clearly not due to "spin-up" if one considers the flow inputs into the model (Figure 18).

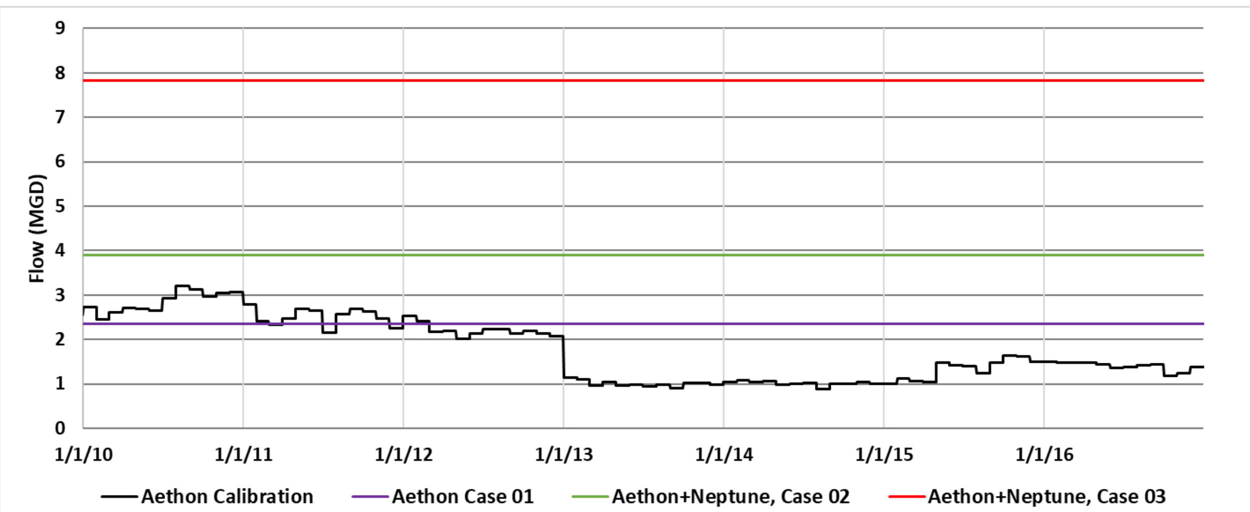

Figure 18. Aethon Produced Water Flowrate Assumptions

Because the projection flows (68,000 bpd; Case 01) are lower than what actually occurred in 2010 and 2011, the model shows an improvement (lower concentrations) with the project. Starting in 2013, when the projection flows are much higher than the actual, the model shows some significant increases in concentrations (Figure 17).

This is an additional illustration as to why the method of lumping into monthly averages is inappropriate. In this case, the conclusion reached depends on the period analyzed. If one only considered the period 2010 – 2012, the results would show an improvement with the project. If, on the other hand, one only considered the period 2013-2016, the results would show a greater impact than reported in the Report.

The analysis should be presented on a daily basis so that periods of larger impact are transparent. For example, from Figure O 1 (Figure 19), the reader could assume that April concentrations may only increase by up to 2.7 mg/L chloride, while the time-series data (Figure 17) show increases of up to 4.6 mg/L at times.

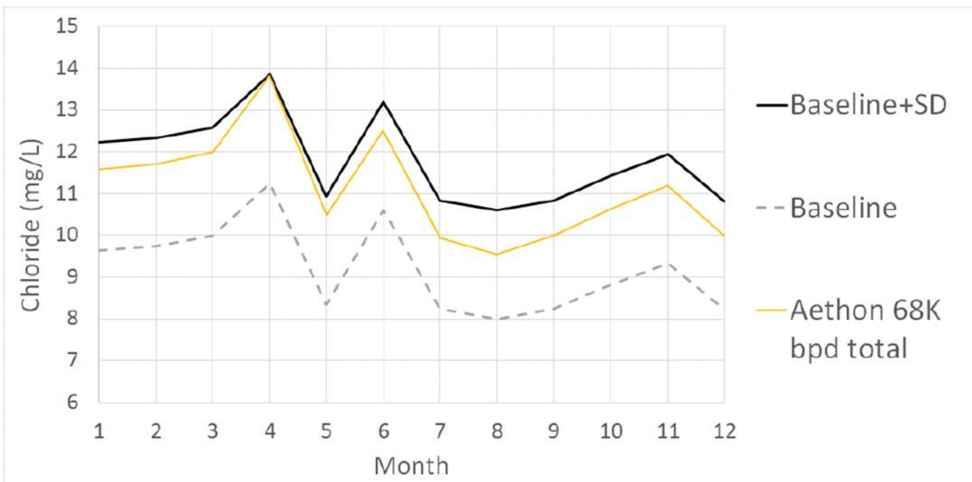

Figure O 1 Chloride Resulting in Wind River Below Boysen Dam from Aethon's Projected Discharge of 68,000 bpd Total (Modeling Case 01)

Figure 19. ERM Figure Showing Impact of the Project for Case 01

20. <u>Used Inflated Standard Deviation</u>

As described above, ERM chose to conduct the analysis on a monthly basis. If this is done to quantify the baseline and if the analysis is to be based on a standard deviation, the estimate of the baseline + 1 standard deviation (SD) must be performed using the SD of the lumped monthly data. ERM chose to use the SD of the original data points. This is incorrect and results in allowing a greater load to the reservoir.

For the example above (Figure 16), the SD for the un-lumped data (left) is 2.7 mg/L, while the SD for the lumped data (right) is 1.2 mg/L. This makes a considerable difference in the antidegradation analysis, since the larger SD allows for larger decreases in water quality (Figure 20).

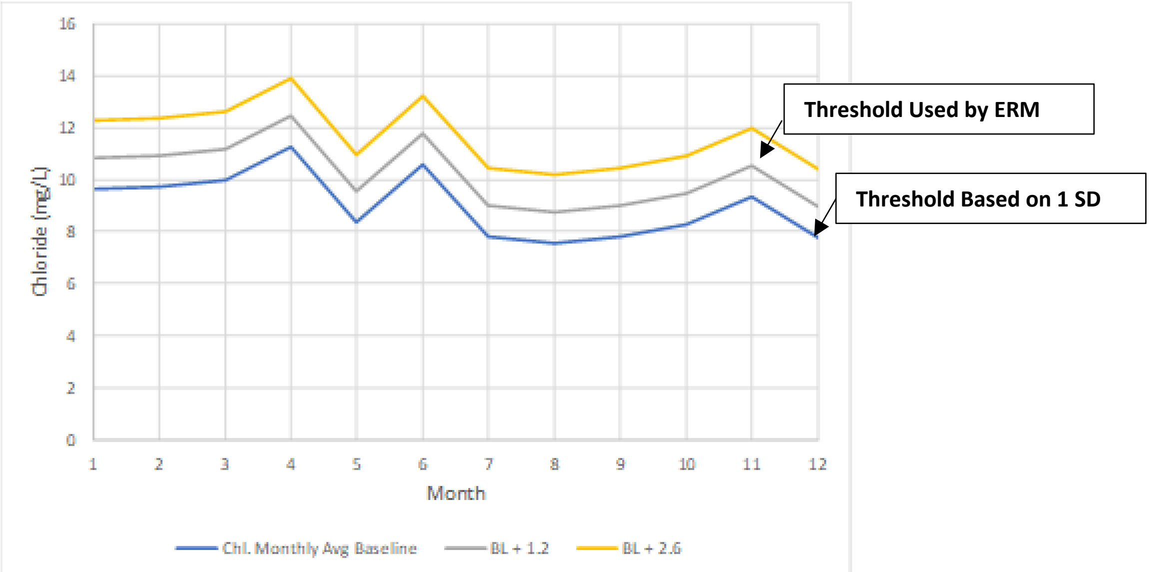

Figure 20. Impacts of Different Standard Deviations

21. Favorable Assumptions Made for Category III Constituents

ERM created an analysis category to include constituents that are present in Aethon's discharge above required detection limits, yet are often below detection limits (more than 50% of the time) in the Wind River below the reservoir. This was called Category III and includes total chromium, dissolved copper, dissolved nickel, dissolved aluminum, and dissolved mercury. ERM chose to evaluate these constituents for compliance by:

- Taking the model results simulated at the top 2 feet of the reservoir near the dam. (which is not representative of the outflow)
- Lumping the results together on a monthly basis and averaging (thus removing observed variability and eliminating the need to display a time series of results and changes each year)
- Comparing the results to the detection limit.

ERM noted that the detection limits varied over time (since they changed labs in 2013, most often resulting in an increased detection limit for some reason), so the decision was made to use the maximum detection limit. Issues associated with this decision include:

- It makes it easier to show compliance; and
- It is inconsistent with use of ½ the DL in the rest of the analyses.

As an example, more detail is provided here for dissolved nickel. For the Aethon sampled Wind River data (which was exclusively used to determine baseline versus using USGS data), the detection limit was 0.5 ug/L from December 2010 through June 2013. Then a different lab was used for November 2013 – March 2016[7] and reported a 5 ug/L detection limit. It is not clear why the detection limit would

[7] No data were collected for the 4-month period between July 2013 – October 2013.

increase, in light of the fact that of the 27 samples collected before the lab change, 13 were below the detection limit[8].

Using the approach developed by ERM, the results are displayed as Figure O 28 in Appendix O (see Figure 21). Using the threshold of 5 ug/L (based on the 2nd lab's detection limit), it appears that the project will not result in degradation in the Class I section. However, if the lab change had not occurred (or if the minimum DL was chosen), then the conclusion would be that degradation would occur.

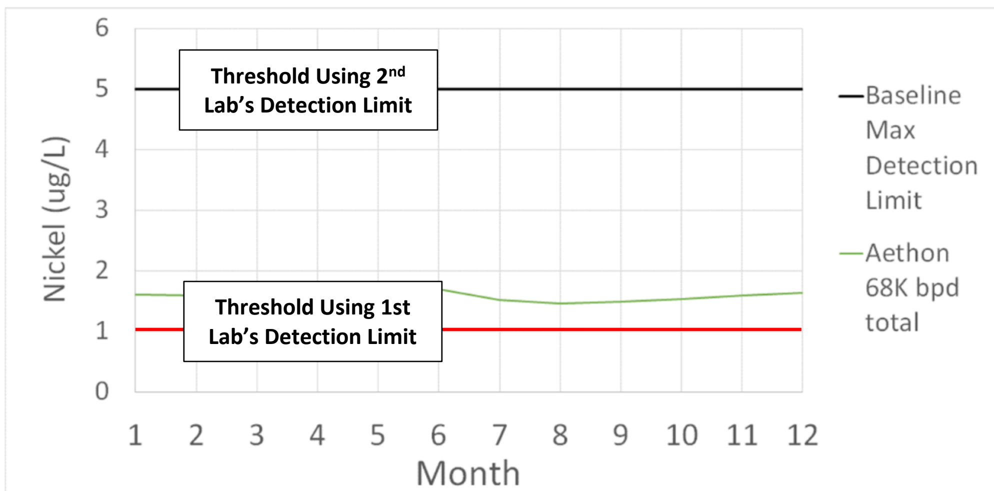

Figure O 28 Nickel Resulting in Wind River Below Boysen Dam from Aethon's Projected Discharge of 68,000 bpd Total (Modeling Case 01)

Figure 21. ERM Results for the Compliance Analysis for Dissolved Nickel (Red Line and Two Text Boxes Added)

Overall, this method and its implementation are flawed.

22. Created Alternative Threshold for Aluminum

Extending the discussion for Comment 22 for an in-depth look at how dissolved aluminum was evaluated (another Category III constituent), it appears that an alternative tactic was used. For this constituent, the 1st lab's detection limit was 4 ug/L and the 2nd lab's detection limit was actually lowered to 3 ug/L. The observed data are shown in Figure 22, along with USGS data for comparison (which were not considered to quantify baseline conditions).

[8] Although it is interesting that the 50% threshold for Category III constituents was being approached.

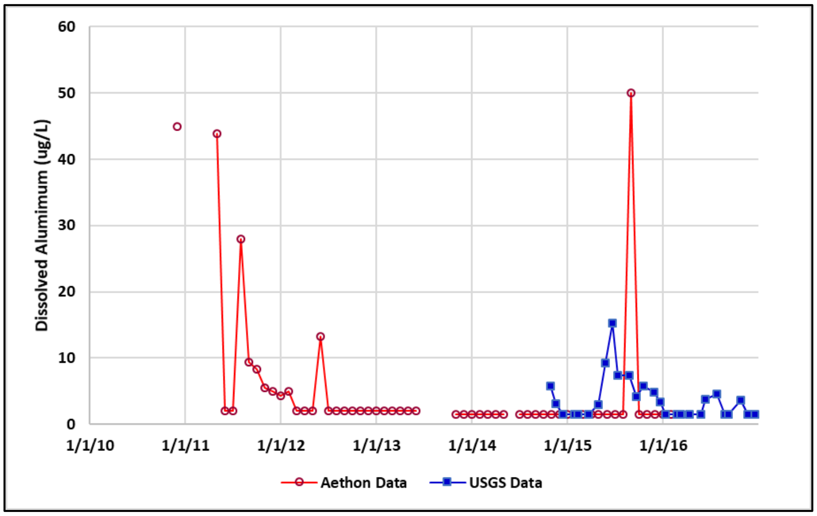

Figure 22. Dissolved Aluminum Data for Wind River below Boysen Reservoir

After lumping the model results and applying the adjustment factor (described above), the monthly model results are in the range of ~22-33 ug/L (Figure 23). This would show a problem if one compares these values to the 3-4 ug/L detection limits. ERM chose to set an alternative threshold of 50 ug/L and using that threshold, the project would not result in degradation.

The source of the 50 ug/L threshold appears to be the required detection limit for dissolved aluminum at the <u>end of the pipe</u> (WDEQ, 2019a). This 50 ug/L detection limit <u>does not apply</u> to the Class I segment of the Wind River (since it is not effluent) and use of this value by ERM for compliance is wrong.

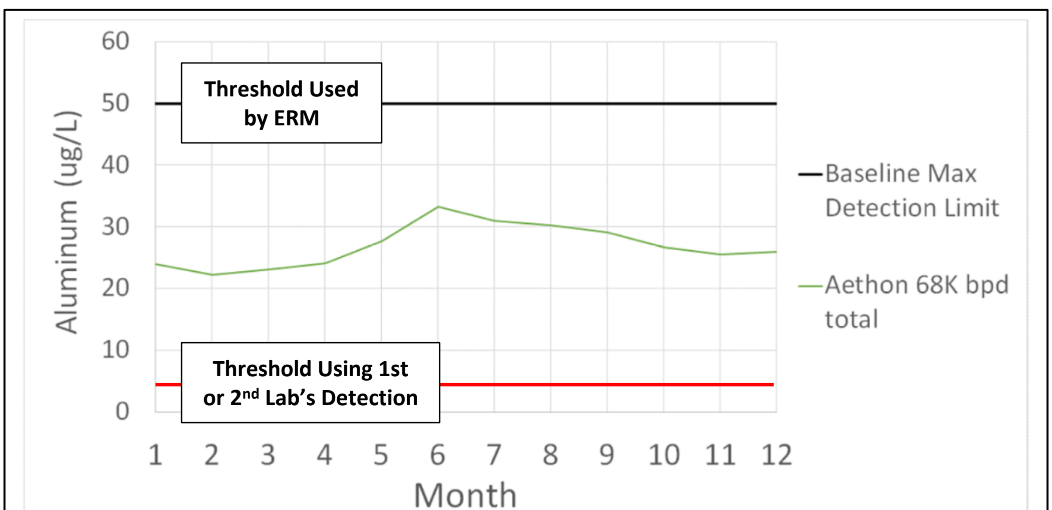

Figure O 31 Aluminum Resulting in Wind River Below Boysen Dam from Aethon's Projected Discharge of 68,000 bpd Total (Modeling Case 01)

Figure 23. ERM Results for the Compliance Analysis for Dissolved Aluminum (Red Line and Two Text Boxes Added)

Boysen Reservoir Antidegradation

23. <u>ERM Did Not Conduct an Antidegradation Analysis for Boysen Reservoir</u>

For Boysen Reservoir, which is classified as a High Quality Water (Class 2AB), “a lowering of water quality may be allowed if it is determined that the amount of degradation is insignificant” (WDEQ, 2013). The determination of significance of the degradation is to be determined using either of the following tests:

- The increased loading is less than 10% of the existing total load for critical constituents; or
- The increased loading will consume, after mixing, less than 20% of the assimilative capacity for critical constituents.

The only time loading to the reservoir is described in the Report is in Chapter 8, which focuses exclusively on chloride. ERM assumes that Aethon can discharge 23.8 tons/day of chloride to the reservoir, based on a flow of 68,000 bpd (Case 01 – no treatment) and a concentration of 2,000 mg/L end-of-pipe limit. ERM states *“This resulting load is the total allowable chloride load that can be discharged by Aethon’s operations while complying with the Antidegradation criteria.”* This is <u>not</u> based on Boysen Reservoir antidegradation, but on the Wind River below. ERM did not consider the 10% load increase criterion for Boysen Reservoir.

The only time project impacts to the water in Boysen Reservoir were considered in the Report is in Chapter 7, the Mixing Zone Study. ERM describes a mixing zone and claims that *“Chronic water quality criteria outside the mixing zone within the reservoir is (sic) met in all three flow conditions.”* Thus, ERM considered it to be acceptable to consume all of the assimilative capacity in the reservoir for this project. ERM did not consider the 20% limit for assimilative capacity.

Thus, ERM failed to conduct an antidegradation analysis for Boysen Reservoir.

SUMMARY

ERM developed a mechanistic hydrodynamic water-quality model of Boysen Reservoir to support permitting and to determine conditions for Aethon’s project that would “protect downstream surface water quality in Badwater Creek, Boysen Reservoir and the downstream Class 1 segment of the Wind River Below Boysen Reservoir, as well as require Aethon to uphold Wyoming’s antidegradation policies.”

There are very serious issues related to the development, evaluation, and use of the Boysen Reservoir Model. Our review of the reservoir model documentation and reservoir model files revealed critical concerns. Highlights include:

<u>The Model was not Developed Properly and Does not Account for Factors Important for this Project</u>

- Density changes anticipated in the future for water flowing into Badwater Bay, (important for flow patterns) were completely ignored.
- Releases to the Wind River (low-level outlet vs. spills) were not differentiated.
- Releases to the Wind River were not density based.
- Wind speeds were severely and unrealistically reduced without discussion.
- Reservoir evaporation was not considered.

- Several water balance and water quality input assumptions and adjustments were made without justification.

Model Performance was Not Evaluated Appropriately and is Misleadingly Communicated

- ERM misleadingly claims that the reservoir model is calibrated and adequately simulates Wind River (Class I segment) water quality. This is done by comparing water-quality measurements in the river to water quality simulated in the top two feet of the reservoir. This is disturbing, wrong, and was done even though the reservoir stratifies and has a low-level outlet.
- There are numerous instances of excluding meaningful data during the calibration/validation process (including all non-profile reservoir data and all data during periods of highest percent produced water).
- Information was misleadingly concealed from the reader by only displaying the top portion of profile results and observations.
- The model is not calibrated and the results are poor.

"Compliance Analysis" Methods and Findings are Flawed and Incorrect

- Baseline conditions for the Class I segment excluded valid USGS data.
- Methods used to show compliance for the Class I segment:
  - Used monthly averages, leading to the conclusion of reduced impacts
  - Used inflated and incorrect values for standard deviation
  - Relied on favorable assumptions for Category III constituents
- An antidegradation analysis for Boysen Reservoir was not conducted.

Based on how the model was developed and the results, the reservoir model cannot be used for projections or decision making. In addition, even if the model adequately simulated water quality, the methods used to determine compliance are inadequate, sometimes wrong, and several assumptions were made to show favorable results.

According to the WDEQ (2019b), "Model was designed to ensure compliance with WQS applicable to Boysen and to maintain existing quality in the Wind River below Boysen." Unfortunately, this is not a true statement.

REFERENCES

Brodzeller, J. and P. McGinley. 2016. A CE-QUAL-W2 Model for Dissolved Oxygen in the Big Eau Pleine Reservoir, Wisconsin to Understand and Manage Winter Anoxia. Center for Watershed Science and Education at the University of Wisconsin-Stevens Point. June 2016.

Cole, T.M., and S.A. Wells. 2016. CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydro-dynamic and Water Quality Model, Version 4.0. Department of Civil and Environmental Engineering. Portland State University. Portland, OR. June 2016.

ERM (Environmental Resources Management). 2018. Water Quality Compliance Analysis for the Long Range Development Plan at Moneta Divide, Wyoming. A Hydrologic, Hydrodynamic and Water Quality Study of the Boysen Reservoir Watershed. Prepared for Aethon Energy. April 23, 2018.

Hydros Consulting. 2017. Three Lakes Water-Quality Model Documentation. Submitted to Grand County, Northern Water, and the U.S. Bureau of Reclamation. November 27, 2017.

LimnoTech. 2016. CE-QUAL-W2 Lake Response Modeling of Petenwell and Castel Rock Reservoirs for the Wisconsin River TMDL. Prepared for U.S. EPA Region 5. June 2016.

NOAA (National Oceanic and Atmospheric Administration). 1982. Technical Report NWS 33: Evaporation Atlas for the Contiguous 48 United States.

WDEQ (Wyoming Department of Environmental Quality). 2013. Wyoming Surface Water Quality Standards. Implementation Policies for Antidegradation, Mixing Zones and Dilution Allowances, Turbidity, Use Attainability Analysis. September 24, 2013.

WDEQ (Wyoming Department of Environmental Quality). 2019a. Wyoming Pollutant Discharge Elimination System Application for Permit to Surface Discharge Produced from Oil and Gas Production Unit Discharges. DEQ_WQD SHORT FORM C.pdf. January 2019.

WDEQ (Wyoming Department of Environmental Quality). 2019. Moneta-Public-Meeting-Presentation_2019-0521.pdf. May 21, 2019.

USDOI (U.S. Department of Interior). 1981. Project Data. Water and Power Resources Service.