Memo – Technical Comments for Work Products Supplied for TAC Meeting #3

United States Government and State of Montana
December 13, 2013

Hydrology

1. Page 11 – *The available observed data is generally not sufficient to support the calibration and validation of multiple parameters in a complex model, so a simple approach was preferred.*

   Please provide information supporting this statement (i.e., are there no climate gages, soils, topography, and landcover in the watershed) which are the primary data needed for such a model. Why were such data available for the Line Creek and Dry Creek UBC model?

2. Page 11 – *An empirical model (rather than a mechanistic or first principles model) was selected, partly, because the fundamental hydrologic processes occurring in watersheds with large waste rock spoil piles are not well understood at this time. Once more is known about waste rock hydrology, a more mechanistic approach could be used to refine the information described in this report, if such a change can provide added value.*

   Please describe what processes are unknown relative to typical understandings of surface water and groundwater hydrology and what future work is planned to fill these data gaps.

3. Page 12 – *In hydrology, flows for target sites are commonly estimated by transposing local flow data from another monitoring location (i.e., from an "analogue" watershed).*

   It would be useful to provide a citation for this approach. It seems analogous to the SCS unit hydrograph procedures, but those were developed for single events.

4. Page 20 – *Dry Creek Analogue Watershed*

   Why not apply this approach (U.B.C. watershed model) to the entire Elk River Valley. This undoubtedly would be a more robust framework to evaluate many other watershed processes such as net percolation through spoil piles, changes in hydrology from topographic/aspect changes induced from mining, etc. A linked hydrologic and water quality model that represents management actions of interest (i.e., cover management, reclamation, etc.) seems appropriate.

5. Page 37 – *Cataract Creek Analogue Watershed*

   For this and other watersheds, please describe the methods used to make the flow measurement and associate quality assurance and control procedures with respect to those

measurements. Also were they done via wading with a current meter, calibrated control structures such as a weir or flume (and what type), acoustic Doppler meter, etc.?

6. Page 44 – *Adjustment for Reclamation*

   Please provide a table listing the number of reclaimed areas being simulated in the model.

7. Page 48 – *Each analysis used four curve fitting methods, 3 parameter Log Normal (3P), Extreme Value (EV), Log Pearson III (LP3), and Weibull (Gumbel III), and the best fitting curve was chosen for each month.*

   Please comment on the validity of using a different frequency distribution for each month. Does it make sense that there is this type of variability each month? If the results of all frequency models were similar, it is recommended that a single procedure be used.

Watershed Modeling

8. Page 5 – *"An empirical approach was also selected to reflect the current level of understanding of the geochemical processes and conditions that occur within the waste rock spoils and other mine features, which are not sufficiently detailed to support the development of a mechanistic, first principles model. Although a mechanistic model may be developed in the future, it is not required to support the development of the EVWQP, because the adopted empirical approach can provide the necessary information."*

    We suggest that this statement is too sweeping and dismissive of mechanistic models. In fact, such models are often used to simulate management scenarios that extend beyond the training data sets of empirical models (see Chapra 1993; Reckhow 1994; Reichart and Omlin, 1997; Martin and McCutcheon 1999). Likewise models are often used in an exploratory analysis to gain knowledge of how environmental systems work and also to predict the behavior of those systems (Omlin and Reichart, 1999). Hence a process-based model for the EVWQP may in fact be appropriate, particularly in the case where hydrologic and geochemical conditions change appreciably between the model calibration-validation and that of the prediction period (i.e., referring to cover management scenarios which alter net percolation, geochemistry, and associated pollutant generation rates).

9. Page 6-7 – *Conceptual Model (Section 2.2.).*

    Based on the conceptual model described in Section 2.2, it appears that some components of a process-based approach should be considered. For example, a simple water balance model could be developed to estimate net percolation through the rock spoil piles using the UBC watershed model (Quick and Pipes, 1977, or equivalent). This would describe snow accumulation, melt, and evapotranspiration behavior thereby proving unique percolation information for each of the spoil piles. Thereafter, physical properties of the spoils (or a simple lag approach) could be used to inform residence time distribution calculations (i.e., exit age) and whether kinetic and thermodynamic approaches are most appropriate. At least then some processes of interest (which may differ with various management scenarios) are appropriately characterized.

    Note: The approach described above has already been undertaken in the United States (Tetra Tech, 2012). In this case, a watershed model (LSPC which is similar to the UBC model) was linked with the chemical transport code from MINTEQA2, equilibrium code from MINEL, and kinetics from PHREEQC to dynamically simulate ionic composition of receiving waters and time variable loadings in mining settings (called LSPC-MDAS; Mining Data Analysis System). While the use of this exact model is not necessary, a more representative approach (even if empirical) should be considered for the various pollutant generation mechanisms which would greatly extend the reliability and predictive capability of the model for the EVWQP.

10. Page 16 – *"Sinks, such as mine-related active water treatment facilities, were not considered in the simulation of historical conditions, because there are no mine-related active water treatment facilities in operation in the Elk Valley. In addition, no biological, physical and chemical decay of the parameters in the surface water were included in the Model. Adsorption, partitioning and absorption of substances were also not modelled. These assumptions are anticipated to result in a conservative estimation of potential loads in water."*

Please provide calculations supporting why these sinks (i.e., biological or chemical) are not important during transport. A simple order of magnitude analysis should suffice. On a similar regard, for completeness, a sink term should be included in Eq. 1 since mass per time loading (i.e., $c_i q_i$) is being determined indirectly. Lastly, it should be noted that since sinks are being ignored in calibration of each modeling node, the actual release rate (from spoil piles; Table 3-2) could be underestimated if assimilation of pollutants (biological uptake and incorporation into tissue, biogenic volatilization, etc.) either between the toe of the spoil and the calibration point, or between subsequent monitoring points downstream, are important.

11. Page 17 – *"Conceptualization of sources and release mechanisms combined with interpretation of the data led to division of the water quality parameters of interest into two groups based on whether release would be expressed dominantly as loading-based or concentration-based. Nitrate, selenium and sulphate were identified as being appropriately characterized as mainly loading-based, as a result of their high mobility under the prevailing weathering conditions in the Elk Valley and their narrow range of release rates when normalized to rock volume (SRK 2013a)."*

Since nitrate, selenium, and sulphate are loading-based (i.e., linked strongly to net percolation through the spoil piles), why isn't this process accounted for in the modeling (even if it is done simplistically)? A simple coefficient related to the infiltration volume may improve model results rather than relying on the annual release rate, an associated calibration factor, and flow adjustment. It is recommended such a factor be included in Eq. 2.

12. Page 17 – *Equation 2.*

This equation has too few degrees of freedom and relies too heavily on the calibration factor (F_c) as the only variable to alter loading rates (L) from waste rock volume. At a minimum, two calibration factors should be included, one describing the effect of net percolation and another reflecting *insitu* processes with the spoil pile. The former could readily be developed using a combination of site data including the ratio of current year to average annual precipitation, drainage/spoil aspect, cover type, etc. If such data are not available, precipitation lapse rates (adjusted for elevation) can be calculated from existing gages and applied to a specific spoil pile; aspect can be determined via GIS using the most

recent Shuttle Radar Topography data (or other up-to-date topographic information); and vegetation type can be determined from LANDSAT TM classifications.

13. Page 18 – *Table 3-3*.

These percentages (i.e., α) are a good start in estimating seasonal distributions of loadings (on average). However, in agreement with our prior comment, the addition of another factor to reflect inter-annual variability of net percolation would greatly enhance the loading analysis and would likely provide better inter-annual model predictions.

14. Page 18 – *Eq. 3*.

We note this as a dimensionless attempt to remedy the deficiency highlighted previously (in the last two comments) which is valid only when all of the spoil piles are uncovered (i.e., $L \propto R_{avg}*Q_i/Q_{avg}$). However, as cover options are considered (thereby reducing infiltration), this relationship will undoubtedly change as the volume of infiltrating water is reduced (i.e., the dependency of L on $R_{avg}*Q_i/Q_{avg}$ becomes non-linear). Case in point, if there were no percolation, there would be no loadings from the spoils (hence no relationship would exist). As currently configured, this assumption will overestimate L for the cover management scenarios, perhaps to the detriment of considering the cost-benefit of each action. More thought needs to be given to how Eq. 3 will be used to accommodate future loading scenarios, or if it is even valid.

15. Page 20 – *Figure 3-6*.

The linear relationship between the dimensionless quantities Q_i/Q_{avg} and Se_i/Se_{ave} is interesting and seems to have good explanatory power. How would this relationship hold if the amount of water infiltrating the spoil pile is changed or if geochemical conditions change? How does the existing model represent this dependency? Would some other assumption need to be made? Finally, assuming the solubility limit can somehow address this concern, does equilibrium in fact exist (for which K_{sp} is based), and to what volume of water would this K_{sp} be applied to since only the loadings, not the volume of water for each spoil pile were determined?

16. Page 21 – "*Initial concentrations were compared to the geochemical solubility limit (Table 3-2), and loadings were reduced, if required, so that predicted concentrations of selenium and sulphate in waters draining from waste rock structures did not exceed the solubility limit.*"

Can you please provide the calculated concentrations in the water draining from the waste rock (range, mean, median) in the Table 3-2?

17. Page 21 – "*The Ferguson and Leask (1988) method was developed using data from coal mines in the Elk Valley, and it describes the mass of nitrate released from waste rock as a function of the ratio of slurry to ANFO used in a given year. The method assumes that all*

explosives residue is washed off the waste rock within one year. Following this method, active spoils release nitrate but non-active spoils do not."

Thank you for acknowledging that this is not a particularly good assumption, and for incorporating a time-dependent release rate for waste rock spoils older than 1 year.

18. Page 21 – *Eq. 13*.

How was the flow (*Q*) through the waste rock determined?

19. Page 32 – "*Where there was scatter in the measured data compared to the modelled results, a level of conservatism was maintained in the Model (i.e., on average, having the simulated concentrations exceed those measured in the field). A level of conservatism in the Model was maintained in an effort to avoid under-predicting future concentrations when the Model is applied to the evaluation of potential mitigation or management scenarios. Conservatism in model simulations will be considered when using results to make management decisions."*

While conservatism is good from the perspective of having a model that doesn't underpredict constituents of interest, an equally valid point is having a model that best represents the observed data. As such, it is suggested that the model be refined to minimize simulation error (using an objective function such as the sum of squared residuals or other quantity), which will reduce uncertainty in model predictions and better informative for water quality managers. Uncertainty can then be addressed through other measures after-the-fact (i.e., by doing an uncertainty analysis).

20. Page 45 – *Table 4.5*.

Selenium calibration factors in this table range an order of magnitude (0.11-2.48). Is it not concerning to see such a wide range of factors? Can any other calibration coefficients be added to the model (that have some physical meaning) to reduce this range? Has this been considered?

21. Page 45 – "*In addition, though simulated and measured data were compared and differences minimized by adjusting the geochemical release rate, no curve fitting was done to try to replicate measured data."*

Please revise this text. In our opinion, adjusting a single calibration factor in an arbitrary way to match observed data is in fact curve-fitting.

22. Page 56 – "*The selenium calibration factors assigned to each watershed in the Elk Valley were typically within 30% of the corresponding sulphate calibration factors (Figure 4-23)."*

Please expound on what this means with respect to underlying model. It seems like given the predictable behavior of both selenium and sulphate within the individual tributaries, it

would of great benefit to try to further understand why there are such pronounced differences between each of the tributaries [i.e., can this information be used to better describe the wide discrepancy in calibration factors (F_c)].

23. Page 60 – *Table 4-9.*

    Was any attempt made to compare nitrate calibration factors with that of selenium or sulfate? If this was done, do they correlate well? What does that tell us about the underlying mechanism of loading if there is correlation?

24. Page 70 – *The Model was validated by initially updating the inputs to include information from 2011 and 2012 and then running the Model for these two years without making any changes to the configuration of the Model or the calibration factors discussed in Section 4.*

    Please characterize hydroclimatic conditions of 2011 and 2012 with respect to the model calibration period.

25. Page 91 – *The results of the model validation indicated that the Model generally performed well for selenium, sulphate and nitrate with respect to simulating in-stream conditions in the Elk River mainstem, in Michel Creek and in Line Creek.*

    Please provide some context for the descriptor "well" in this statement from the literature.

26. Page 92 – *Model Re-calibration*

    This section would benefit by having a matrix for each watershed describing what changes were made after the model calibration-validation. This will help better understand what the final approach was (i.e., flow relationship, re-adjustment of calibration factor, change in monthly release distributions, etc.) and the differences between all of the watersheds.

27. Page 94 – *Kilmarnock, Swift and Greenhills Creeks*

    Please describe the physical justification for modifying the monthly release distributions for selenium, sulphate, and nitrate.

28. Page 106 – *Table 6-4.*

    It is surprising that by including 2 additional years of data for the validation in re-calibration (i.e., as opposed to the 7-8 years used in calibration), the selenium F_c changed so greatly in some watersheds (40+ percent in some instances). Is there an explanation for this difference?

29. Page 130 – *6.3 Summary.*

A more robust description of model sensitivity and uncertainty would be helpful in understanding the behavior of the model. While the underlying equations are linear (hence sensitivity analysis is useless), input uncertainty may not be, therefore some description of uncertainty should be included. A simple Monte Carlo analysis in GoldSim would be useful. In addition, when describing uncertainty, the modeling report should include a brief description of how the 0.10 probability high/low flow will also be used in this regard (as described in other documents). Finally, some description of how the model will be adaptively managed would be of great benefit. For example, will a model post-audit be done and on what frequency will that be considered? How will the model be adaptively changed in the future and how will those changes be distributed to stakeholders? How will the results of the R&D program be integrated into the modeling framework? How often will literature reviews be undertaken to investigate new mechanisms to add to the model?

30. Appendix B Page 1 – *Figure B-1.*

Please describe why an older version of the dimensionless regression relationship (SRK 2011 vs. SRK 2013) was used in the model ultimately?

References

Chapra, S.C. 2003. Engineering Water Quality Models and TMDLs. J. Water Res. Plan-ASCE. 129 (4): 247-256.

Martin, J.L. and S.C. McCutcheon, 1999. Hydrodynamics and Transport for Water Quality Modeling. CRC Press, Inc., Boca Raton, Florida.

Omlin, M. and P. Reichart. 1999. A comparison of techniques for the estimation of model prediction uncertainty. Ecol. Model. 115 (1): 45-59.

Quick, M.C. and A. Pipes. 1977. U.B.C. Watershed Model/ / Le modèle du bassin versant U.C.B. Hydrolog. Sci. J. 22 (1): 153-161.

Reckhow, K.H. 1994. Water Quality Simulation Modeling and Uncertainty Analysis for Risk Assessment and Decision Making. Ecol. Model. 72 (1-2): 1-20.

Riechert, P. and M. Omlin. 1997. On the usefulness of overparameterized ecological models. Ecol. Model. 95 (2-3): 289-299.

Tetra Tech. 2012. Updated MDAS Model Capabilities Within the LSPC Modeling Framework. Technical Memorandum. Fairfax, VA 22030.

Geochemistry

31. Page 10 – *Step 2. Concentration Calculation (and subsequent calculations).*

    How is Q through the waste rock pile determined (i.e., net percolation)? If it from gaged flow downstream of the waste pile, is any attempt made to discriminate between groundwater Q that may originate from somewhere else and net percolation through the tailings? Have any studies been conducted to compare these values (i.e., estimated net percolation against observed flow downstream of the tailings annually)? If so, how do these estimates compare?

32. Page 18 – *Step A – Calculate Daily Loads*

    Please provide further detail about how the daily concentration data (C_j) were developed so that $L_{j,day}$ could be determined. What was the underlying relationship and its associated explanatory power? How many data points were used to develop this relationship?

33. Page 20 – *Figure 5*

    The similarity between loading for all constituents on a monthly basis are remarkable. Does this not suggest that flushing of spoil piles (and associated infiltrated water) are perhaps more important than geochemical processes (i.e., that nitrate from blasting residuals and selenium from weathering show an approximate 1:1 loading relationship)?

34. Page 22 – *Figure 6*

    While there is no definitive relationship, it does seem like there is an inverse correlation between age of spoil pile and calculated loading rate. Please characterize whether a significant correlation exists (i.e,. using Spearman's rho or some other approach) and what the associated significance of the correlation is. Are these results consistent with that reported in the literature? In addition, please characterize what this means for long-term loadings estimates with respect to the modeling (and your associated treatment options)? Finally, if "age of pile" is in fact an important consideration, should not this be included as a separate factor in the model (which would be easy since the model is linear)?

35. Page 25 – *Sulphate and Selenium Release*

    While both selenium and sulphate should be correlated due to stoichiometric considerations with respect to pyrite oxidation, why the apparent (or perhaps coincidental) correlation with nitrate?

36. Page 27 – *Figure 9.*

This flow relationship is presumably for current conditions under which there are no covered spoil piles. Does this flow relationship in fact hold true if covers were placed in each of the watersheds?

37. Page 28 – *Selenium Solubility Limit*

How does the K_{sp} for selenium compare with the daily calculated concentrations from the spoil piles? Is equilibrium a valid assumption if compared to the estimated residence time of water within the spoil piles? Has residence time ever been determined via tracers (something like rhodamine or more complex methods like stable oxygen isotopes)?
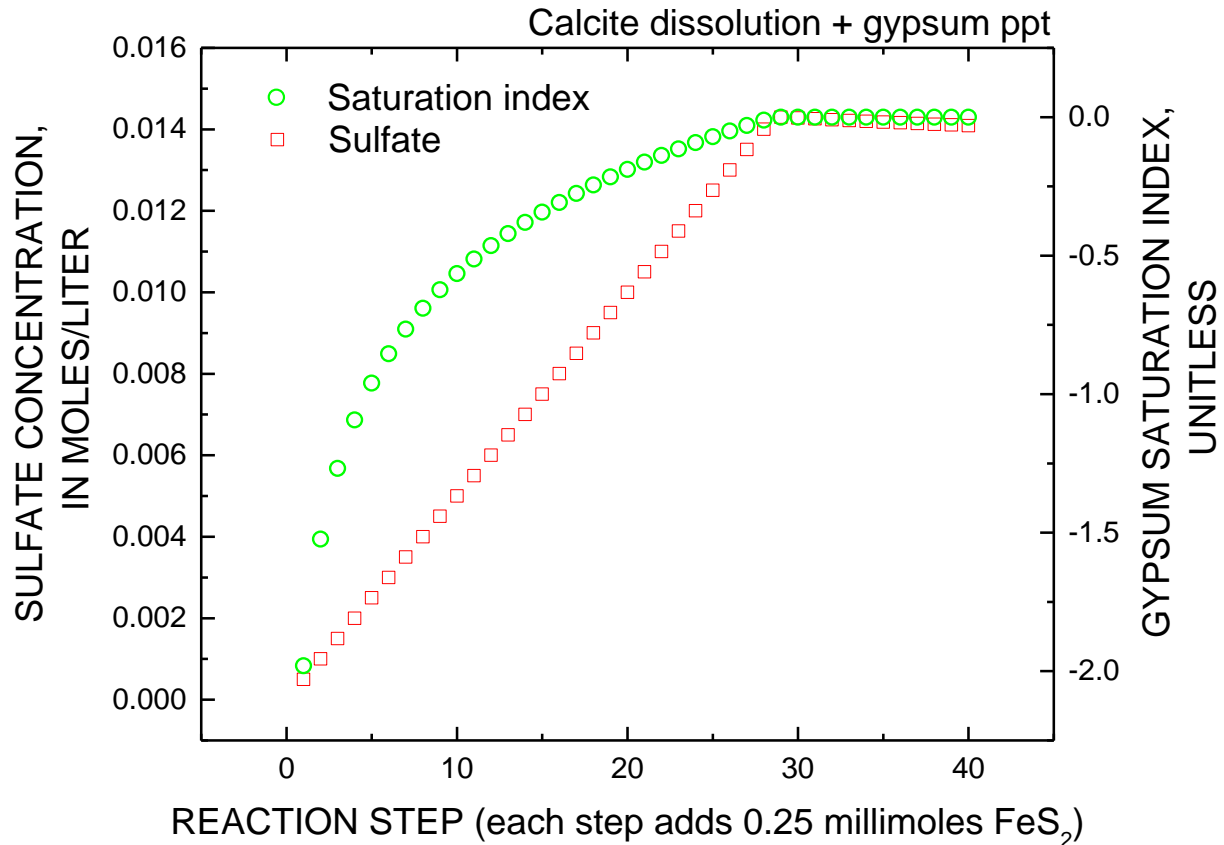
38. Page 32 – *Figure 11*

This figure (and associated discussion) seem to be in conflict with the release rates and age of pile shown in Figure 6. Am I missing something missing here? How can the release rate (Figure 6) show some decline with age but the concentration (Figure 11) stay the same? Are there water management activities buried somewhere in this data?

39. As noted in various sections of the Water Quality Modeling report, a static solubility limit is used for both gypsum and selenium (i.e. Table 3-2, p. 18) without adequate supporting information or the recognition that this solubility limit is dependent on a combination of other dissolution/precipitation processes and mineralogy that may change as the waste rock piles age. A series of PHREEQC reaction path simulations that considered different combinations of plausible waste rock/water reactions were modeled and resulted in a wide range of sulfate concentrations to attain gypsum saturation in the infiltrating water. These simulations are not conclusive and the concentrations of the modeled constituents should only be used for comparative purposes to illustrate how minor changes in the waste rock pile can have large relative impacts on corresponding sulfate concentrations. The model descriptions and results are summarized below.

Simulation 1: Water infiltrating the waste rock pile is equilibrated with a soil zone pCO_2 of $10^{-2.5}$ atm. This infiltrating water is then reacted with pyrite (dissolves 20 millimoles in 80 reaction increments). During each reaction step, the solution is equilibrated with calcite ($CaCO_3$) and amorphous iron oxyhydroxide ($Fe(OH)_3$). In this simulation, aragonite is the carbonate phase assumed to buffer the pyrite oxidation in the waste pile and the iron released from pyrite oxidation forms the oxidized iron. The saturation index of gypsum is calculated after each reaction step, as is the total sulfate concentration in solution. Results from simulation #1 are shown in the graph below and this reaction model indicates that a sulfate concentration of 21.8 mmoles/liter (2,100 mg/L) can be obtained at anhydrite saturation (saturation index = 0). This compares well with the sulfate saturation concentration of 2,400 mg/L noted on Table 3-2 of the Water Quality Modeling report.

Calcite dissolution + anhydrite ppt
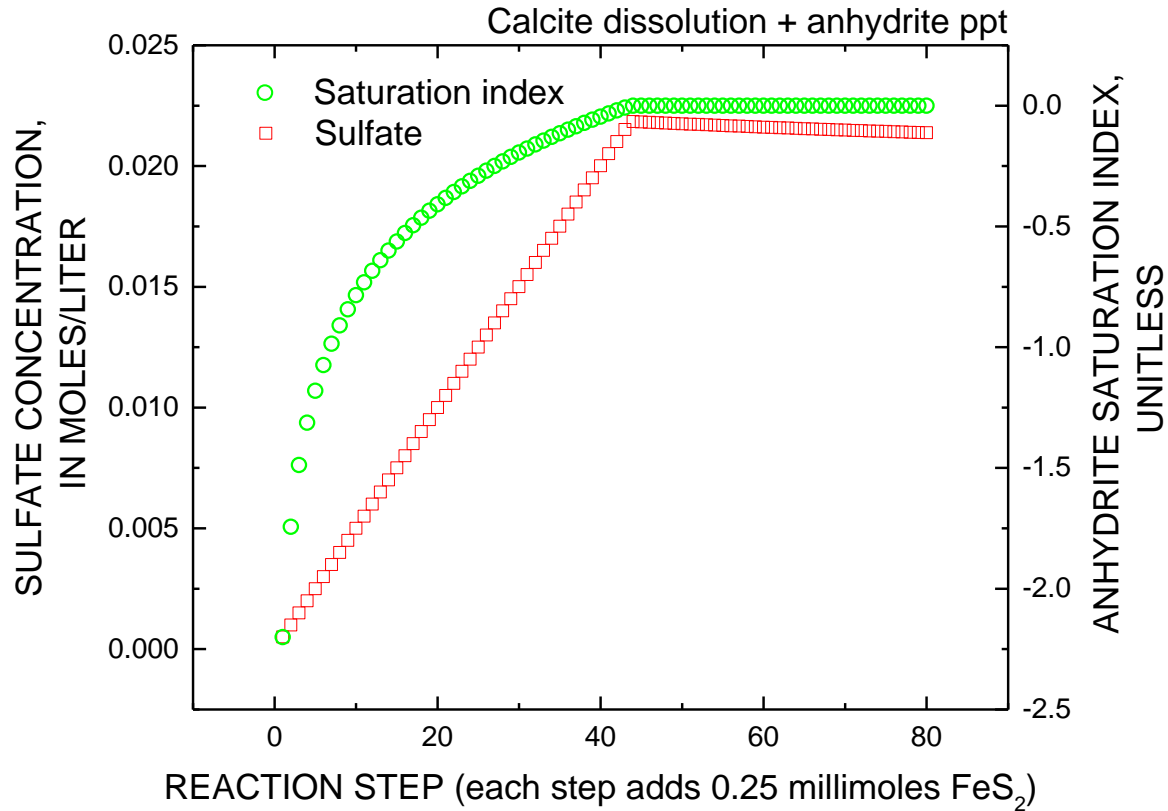
Simulation 2: The same initial reactions that were used in simulation 1 were applied in simulation 2. In simulation 2, gypsum, instead of anhydrite was allowed to reach saturation with the infiltrating water. Results from simulation #2 are shown in the graph below. This reaction model indicates that a much lower sulfate concentration is obtained at gypsum saturation (14.3 mmoles/liter; 1,370 mg/L) relative to the previous simulation using anhydrite saturation (21.8 mmoles/liter; 2100 mg/L), using a different equilibrium mineral phase to control sulfate concentration during reaction progress.

Calcite dissolution + gypsum ppt

(Graph: SULFATE CONCENTRATION, IN MOLES/LITER vs REACTION STEP (each step adds 0.25 millimoles FeS$_2$), with GYPSUM SATURATION INDEX, UNITLESS on right axis. Legend: ○ Saturation index, □ Sulfate)

Simulation 3: This simulation considers the likely possibility that the carbonate buffering material is a Ca-Mg carbonate mineral phase (50:50 mole ratio), instead of a pure Ca-carbonate phase such as aragonite or calcite. The other model boundary conditions are the same as simulation 1, with anhydrite selected as the equilibrium mineral phase to control sulfate concentration in the reaction solution. Results from simulation #3 are shown in the graph below. This reaction model indicates that the carbonate mineral phase composition can potentially have a large impact on the sulfate concentration. Since there is 50 percent less calcium available in solution from each reaction step, calcium instead of sulfate becomes the limiting reactant for anhydrite formation after mineral equilibrium is reached. As shown in the graph of reaction progress, there is a decrease in the rate of sulfate increase after anhydrite saturation is reached; however, sulfate continues to increase after solution saturation. The reaction path model was terminated after the oxidation of 40 mmoles of pyrite, resulting in a final sulfate concentration of 55.6 mmoles/liter (5,340 mg/L). Under this reaction scenario, the sulfate concentration in solution will likely be limited by either the amount of pyrite dissolution or a higher solubility sulfate salt.

Ca-Mg carbonate dissolution + anhydrite ppt
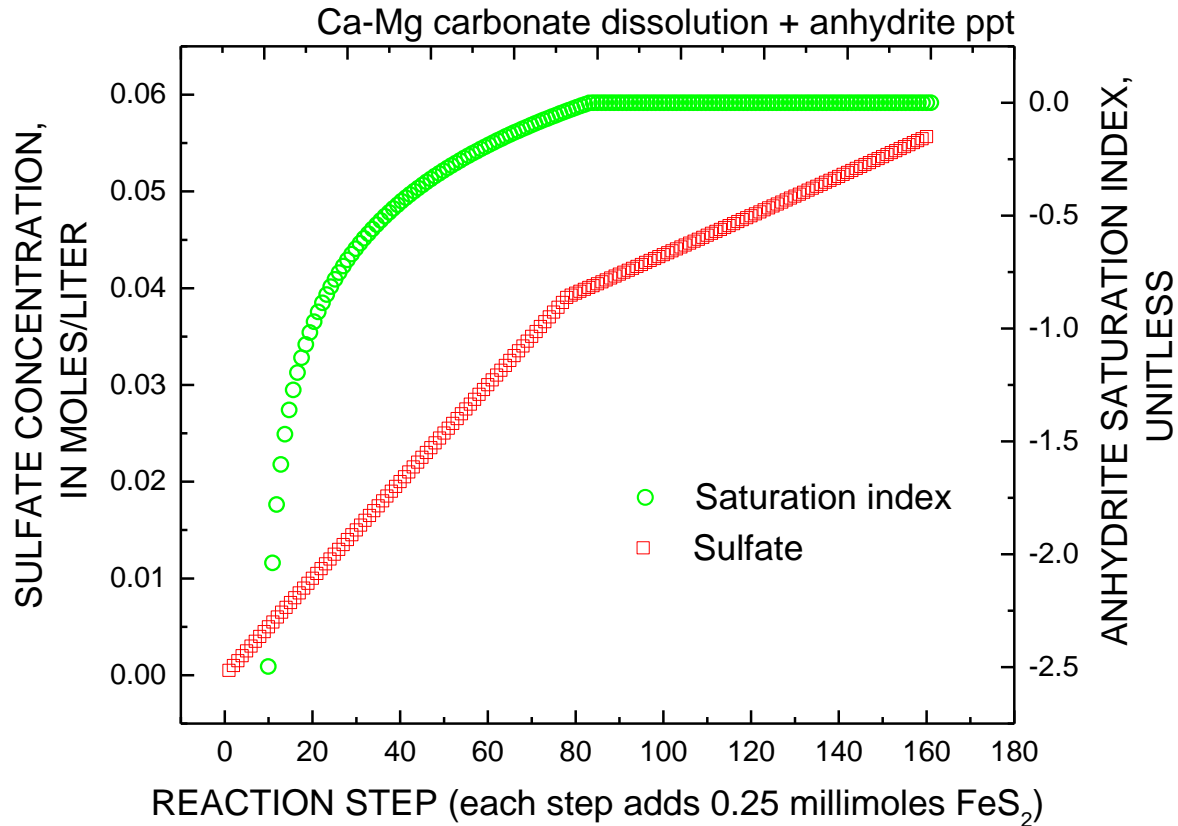
While these PHREEQC models are for illustrative purposes only, they support our concern regarding the need for more quantitative geochemical information to support the development of a process-oriented approach during future iterations of the basin wide water-quality model. As noted previously, one of the main concerns to U.S. state, federal, and tribal agencies is the long-term leaching potential of waste rock materials in the Elk and Fording River valleys. Increasing this concern is the lack of a quantitative understanding of the current and long-term geochemical reactions in combination with waste rock hydrology that may continue to release contaminants such as selenium and sulfate for time periods beyond the operating life of the mines and associated water treatment facilities to Lake Koocanusa.

40. As noted in Water Quality Modelling Report and verified in the PHREEQC modeling discussed earlier, gypsum saturation and subsequent precipitation (and non-verified substitution of SeO_4^{2-} for SO_4^{2-} in the mineral structure) driven by sulfide oxidation is a potential reaction path during initial weathering of the waste rock piles. If and when the sulfide mineral phase(s) and (or) oxygen become a limiting reactant in the waste rock piles, this reaction path will no longer generate sulfate and corresponding gypsum/anhydrite precipitation. Under this evolving set of geochemical conditions, the infiltrating water in the waste rock piles will now be undersaturated with respect to typical sulfate mineral phases and will dissolve the previously precipitated sulfate mineral phases, likely resulting in a long-term source of sulfate and co-precipitated selenium originally precipitated during sulfide weathering. As an example, PHREEQC was used to simulate contact of recharge water with previously precipitated anhydrite. Based on these modeling results, re-solution processes could mobilize up to 23.4 mmoles of sulfate per liter of water (2,220 mg/L) under equilibrium conditions. Furthermore, any co-precipitated selenium in the anhydrite mineral structure could be re-released. Assuming a molar ratio of 0.00015 of selenate-to-sulfate in the anhydrite mineral structure, dissolution could result in a re-release of up to 280 μg/L of dissolved selenium.

While sulfate mineral phase saturation/precipitation processes may be an initial contaminant removal mechanism in the waste rock piles, additional quantitative information on the re-release potential of these contaminants are needed. Potential approaches for obtaining this information include: (1) long-term monitoring of $\delta^{34}S$ and $\delta^{18}O$ in the dissolved sulfate within existing waste rock piles of various age ranges and weathering regimes; (2) solid phase analysis of secondary sulfate minerals obtained from buffered sulfide weathering zones in selected waste rock piles; and (3) laboratory column/weathering cell tests simulating buffered sulfide oxidation and sulfate precipitation that then transitions into anhydrite/gypsum dissolution.