Comment 1:

Quality Assurance Project Plan (QAPP) Document

ERM to address. I think we should determine what else we could do on QA/QC. Given that they devoted a full page to this topic, I think we need to agree to additional work (easy for me to say, it would be ERM doing it).

Comment 2:

The potential for low or high flow from the Peter Mitchell Pit (PMP) was included in the screening sensitivity analysis that was performed for the Mine Site model ("Initial Sensitivity Analysis Results of the NorthMet Mine Site GoldSim Model (Version 2)", March 28, 2013). This input parameter will also be included in the detailed sensitivity analysis to be completed following direction of the Co-lead Agencies.

Depending on the chemical constituent, flow from the PMP may either increase or dilute the concentration in the Partridge River relative to background conditions; therefore reductions in the modeled flow may cause the concentration to either increase or decrease. Note that the Mine Site model was calibrated assuming a PMP flow of 1 cfs, the screening sensitivity analysis was performed with PMP flows ranging from 0.5 to 10 cfs.

Despite the observed sensitivity of the estimated concentrations in the Partridge River to the amount of flow from the PMP, the estimation of potential impacts from the NorthMet project is not sensitive to the amount of flow from the PMP. A model run with PMP discharge at zero can be included as part of the forthcoming more detailed sensitivity analysis.

Comment 3:

The commenter is correct that Attachment C does not provide the necessary detail on the empirical information that is being used to address potential indirect wetland impacts associated with pit dewatering. The detail that is being requested can be found in Section 5.2.1.2.2 of the Wetland Data Package, Version 7.
Comment 4:

ERM to address. Response may be something along the lines of the representativeness of the samples was considered when the Co-lead Agencies agreed that the new data did not need to be incorporated into the model.

Comment 5:

ERM to address. Correct errors in first paragraph. Reference decision by Co-lead agencies to not consider correlations. Other than hardness and TDS, correlations not important. Hardness discussed in a comment below. TDS addressed as part of a position paper that EPA probably hasn’t seen (I don’t think we included in our data package).

Comment 6:

It is not clear why the decision was made to use the sample means in lieu of the complete sample. More commonly, upper confidence limits are calculated and used to represent the uncertainty in the mean of sampling data. It means that upper confidence limits should have been calculated and used similarly to the regulatory standard values listed for ground water and surface water. The guidelines for calculating upper confidence limits are well established and easy to follow. They may be a more appropriate approach for this project. Essentially, using the sample means reduces the natural variability of background concentrations and may eliminate those potential “worst-case” scenario conditions that, in combination with the proposed actions, would lead to environmental impacts, unnecessary future monitoring, additional documentation and cost for the proposed plant.

Stochastic elements are supposed to be independent of all other elements. It is clear that the chemical elements are dependent on each other and should not be imputed as stochastic elements into the GoldSim program. These are points within the GoldSim to bring for the actual sample analysis.

As discussed in the Water Modeling Data Package Volume 1 – Mine Site, Section 5.2.3.8.1, the uncertainty in the mean background groundwater concentration is used in the models “because
groundwater sources cover a wide area and groundwater travel times are relatively slow. Thus, loading from groundwater can be expected to reflect the integration of the spatial and temporal variability observed in the baseline data.” The uncertainty in the mean background groundwater quality is used to model the average quality of baseflow in the Partridge and Embarrass Rivers, the average quality of recharge to the surficial aquifers, and the average quality of groundwater inflow to the mine pits.

In order to address the concern expressed by the reviewers, the Mine Site GoldSim model was re-run using the distribution fit to the full population of background groundwater quality data, rather than the uncertainty in the mean. This approach is akin to assuming that any single groundwater quality sample could represent the constant quality of baseflow and recharge everywhere in the Partridge River watershed over the 200-year model period. Clearly this assumption is contradicted by the weight of the groundwater and surface water quality data; nevertheless this represents the probability of a “worst-case” scenario as described by the reviewers based on the observed occurrences of high groundwater concentrations in the Mine Site data set.

For example, the surficial aquifer groundwater quality data and distributions for copper at the Mine Site are shown in the figure below. The diamonds represent the individual samples (132 in total), the dashed line represents the uncertainty in the mean concentration (the distribution in the base model), and the red line represents the lognormal distribution fit to the entire data set (used for this analysis). For reference, the maximum observed copper concentration in the Partridge River is 6.3 μg/L, well below the upper end of the observed groundwater concentrations.

![Copper](image)

The results of this extreme model run can be compared to the results presented in the Water Modeling Data Package Volume 1 – Mine Site (Version 12). Several summary figures from that document are
reproduced and attached here for key surficial flow paths and Partridge River evaluation locations, along with new versions reflecting the model run with the full range of sampled background groundwater quality. It is clear from these figures that although the range of estimated groundwater and surface water concentrations may change, in no case does the model predict a 90th percentile concentration that is over the standard in the extreme model run but below the standard in the base model run. This observation should address the concern that additional impacts could be observed if the model were run in an unrealistic worst-case scenario.

Regarding independence and correlation between modeled constituent concentrations, see the response to the previous comment.

Comment 7:

Use of the GoldSim Model Geochemical Data Files Within the Model

Modeling Hardness

The first use of the geochemical data in the GoldSim program is the calculation of hardness. This is done using the hardness equation:

\[ \text{Hardness} = 2.5[\text{Ca}] + 4.1[\text{Mg}] \]

Values for calcium and magnesium are independently chosen at random, independent of each other, by Goldsim. This means that the combination of these two geochemical elements may never be seen or interpreted from the actual samples and the hardness may not be within the range of the actual hardness. [It should be noted that in the actual samples for the surface water (Partridge River watershed), there is a strong correlation between calcium and magnesium, but in the GoldSim average surface water sample, there is no correlation.]

Hardness is then used further in the model to calculate the release rates of various chemical elements, which are then added to the original values that the GoldSim program had independently chosen at random for those geochemical elements.

Some clarification is needed on the first statement of the comment regarding “the first use...”. Barr does not agree with the statement, potentially because Barr may not fully understand the statement. The calculation of hardness as shown in the equation above is present in the model, but it is unclear why this equation is called out as “the first use of the geochemical data”.

Barr would also like to correct the misconception (as we understand it) in the final paragraph of the comment where it is stated that “Hardness is then used further in the model to calculate the release rates of various chemical elements...”. There is not a single release rate defined in the model, from waste rock, overburden, ore, or tailings that is defined as a function of hardness (see the Waste Characterization Data Package, Version 10, March 2013, Sections 7, 8, 9 and 10; also see the input tables describing release rate methods for the waste rock, overburden, ore, and tailings included as attachment B in both volumes of the Water Modeling Data Package). Some clarification is needed to understand the origin of this comment and correct any misunderstandings between parties.

Finally, the reviewer is correct that values for calcium and magnesium (release rates, background groundwater quality, surface runoff, etc.) are chosen at random and independently and that modeled
hardness theoretically may not be in the range of the actual hardness. Barr compared the modeled hardness for the No Action models in the Partridge River and in the Embarrass River at PM-12 to the measured values to see how the model is performing at these locations without considering possible correlations between calcium and magnesium. The plots below show the comparison. This first plot shows the comparison of modeled versus measured hardness in the Partridge River. The model does a good job of matching measured hardness in the Partridge River.

This second plot shows the comparison of modeled versus measured hardness at PM-12 on the Embarrass River. The model does a good job of matching measured hardness in the Embarrass River.

It is true that concentrations of calcium and manganese are correlated in the surface waters. However, not considering this correlation in the GoldSim model does not result in unrealistic estimates of hardness.
Comment 8:

Use of Average Samples
The use of average samples in determining when there is a violation of standards is questionable. By the time the average sample show a violation, half of the actual samples would have been in violation and there is no way to determine when the first actual violation occurred.

Need clarification.

Comment 9:

“Seeding” of GoldSim Actual Samples
The reseeding option in GoldSim was turned off for the program. This prevents GoldSim from choosing a different random number to use in creating the average samples that are used in the program. The end result is that GoldSim uses the same set of originally random average samples every time the program is run. This results in the same set of average samples being used for the calibration of the model and for the regulatory submittal of the model. This could be interpreted as tweaking the model to give a desired result for a specific set of data.

As a definition, a random seed (or seed state, or just seed) is a number (or vector) used to initialize a pseudorandom number generator. Random seeds are often generated from the state of the computer system such as the time. If the same random seed is deliberately used and shared, it becomes a “secret key” to the pseudorandom number generator, so two or more systems using matching pseudorandom number algorithms and matching seeds can generate matching sequences of non-repeating numbers. This is the case of the models which were delivered for review. This was done so that the reviewers would be able to see the same values and results that were generated by Barr to facilitate discussion, comments, results comparison, and review.
The calibrations that were performed for the modeling (background surface runoff for example) were not performed within the same models that were submitted for regulatory review. Therefore, the seed of the models submitted does not affect the calibration that was performed of the models. Submitting a model with the reseeding option turned off should in no way be interpreted as “tweaking the model”, because the calibrations and model runs were not performed in the same models and the choice to force GoldSim to repeat the sampling sequence was intentional to facilitate review and discussion.

Additionally, Barr provided on March 27, 2013 a second player file with the reseeding option turned on so that reviewers of the model could see the impact of this model switch. Along with the second player file, Barr generated a technical memo (Sufficiency of the Number of Model Realizations) dated March 18, 2013 that was distributed to the greater group on March 22, 2013. In this memo, Barr attempted to disclose the minor impact that a different random seed had on the complete model results.

Comment 10:

General Model-Related Comments
On March 25, 2013, we provided to the co-lead agencies “EPA Questions/Comments on GoldSim Model Review.” We have had subsequent conversations with the co-lead agencies and others. Many of these initial comments have been resolved, which we appreciate. We list below those comments that are unresolved, with a footnote reference to the modeling issues noted in the March 25th package. As a result of these challenges, we have prepared the March 25th package to the agency with comments in red indicating which comments have been resolved and which still need to be resolved.

For the mine site, flowpath widths in the GoldSim model range from 240 to 1,735 meters. Cell lengths range from 26.2 to 29.8 meters. At the upper end, these dimensions result in cells that are 12 acres in size. The following assumptions in the GoldSim model appear to be invalid for such large cells:

- well mixed water and contaminant balance
- homogeneous, low resistance flow path
- singular values for the equation and solution
- uniform geochemical reactions.

Our concern relates mainly to modeling flow via a homogeneous (rather than variable) medium. Unless these assumptions remain a constant (homogenized) block in the model, adjusting the model to vary in both spatially and temporally is essential to more accurately capture the variability.

**ERM** to address. Can the response talk about how an assumption of homogeneity is appropriate given the objectives of the model and is how contaminant transport is commonly done? We are giving some thought as to how a sensitivity analysis addressing this could possibly be done.

Comment 11:
In response to the first part of the comment, related to the correlation between precipitation and recharge, Barr disagrees with the statement that “years with high precipitation will have high recharge and vice versa”. Barr does not believe there is strong evidence of a correlation at this location, and it was decided to not include correlations unless clearly required by theory or empirical evidence. Additionally, annual variations in recharge within the prescribed distribution are not expected to affect the assessment of impacts. Annual variations will be smoothed out or integrated given the long travel times of decades to centuries in the groundwater flow paths.

In response to the second part of the comment, related to the distribution used for recharge, some literature references would support the use of a triangular distribution, and may even suggest it over a log-normal distribution. Hammonds et al state that “where data are limited but uncertainty is relatively low (less than a factor of 10), a range may be used to specify a uniform distribution. If there is knowledge about a most likely value or midpoint, in addition to a range, a triangular distribution may be assigned.” (Hammonds et al, 1994). And later, “the assumption of normal, lognormal, or empirical distributions is usually dependent on the availability of relevant data” (Hammonds et al, 1994).

However, Barr could perform a sensitivity analysis related to this comment to see if groundwater quality and surface water estimates are significantly impacted by making the assumptions mentioned in the comment. Barr could assume a log-normal distribution for recharge which has a similar cumulative density function (CDF) to the currently assumed triangular distribution, allow recharge to be resampled annually, and correlate the recharge (as an extreme scenario) to the annual precipitation using a correlation coefficient of 1.

Comment from Tyler Kaspar, 1854 Treaty Authority

The conservatism of assuming the West Pit is well or poorly mixed depends on the dominant source of loading to the pit for a given constituent. Note that, because of the large volume of water, a well-mixed pit lake has a long “memory” of previous loading conditions and will reach a new steady-state...
equilibrium very slowly. The surface layer of a stratified pit lake, on the other hand, will be primarily affected by more recent conditions and will equilibrate relatively quickly.

Consider the following cases:

a) High loading from initial temporary sources but relatively low long-term loading from low-concentration surface inflows once the pit is flooded

b) Lower initial loading but significant long-term loading from high-concentration surface inflows once the pit is flooded

In case (a), modeling the pit as well mixed would result in a high concentration due to the lingering effects of the initial loading from temporary sources, with the concentration decreasing over time as the pit is slowly “flushed” by the cleaner surface inflows. Modeling the pit as permanently stratified would result in a lower concentration at the surface, because the original higher-concentration water would be trapped in the lower layers and the upper layer would be primarily affected by the low-concentration surface inflows. Therefore, in the loading scenario described in case (a), the assumption of a well-mixed pit is conservative with respect to the estimated concentration at the surface outflow.

In case (b), the opposite is true. Modeling the pit as well mixed would dilute the higher-concentration surface inflows with cleaner water from the depths of the pit. Modeling the pit as permanently stratified would result in a higher concentration at the surface, mostly affected by the high-concentration surface inflows. In the loading scenario of case (b), the assumption of a stratified pit is conservative with respect to the estimated concentration at the surface outflow.

Virtually all of the constituents considered in the Mine Site GoldSim model have the loading scenario described in case (a), with long-term inflows that are cleaner than the initial loading to the pit (the initial loading is primarily from wall rock flushing). For these constituents, therefore, the assumption of a well-mixed pit is conservative, as discussed in Section 6.2.3.3.3 of the Water Modeling Data Package Volume 1 – Mine Site. The only exception is barium, as noted in the text.

For mercury, however, the dominant source of loading is the long-term deposition on the lake’s surface rather than temporary sources. Mercury therefore fits the scenario described in case (b), and the assumption of a stratified pit is conservative, as discussed in Section 6.6.2.2 of the Water Modeling Data Package Volume 1 – Mine Site.

The intent of including these two differing assumptions with regard to the West Pit mixing is not to make a judgment about which state is likely to prevail in the actual pit lake. Rather, given the uncertainty over the likelihood and completeness of future pit mixing, the intent is to be conservative with respect to the estimated concentrations at the surface of the pit for the two different analyses.