

Document: NorthMet Project Plant Site Water Quality Model Quality Assurance Project Plan (ver. 1) April
Reviewers: Blaha, Clark, Engstrom, Kunz, Liljegren, Olson, Schwanz

| Comment # | Location (Page, Figure, or Table) | Number | Comment/Concern |
|-----------|-----------------------------------|--------|--|
| 1 | 1 | 1 | Given EPA's focus on evaluating the Mine Site Model QAPP using criteria and checklist in their specific guidance for QAPPs for Modeling, it seems that this should include a formal citation and reference to this EPA's report (i.e., "Guidance for Quality Assurance Project Plans for Modeling, EPA QA/G -5M"). |
| 2 | 5 | 3 | Suggest setting some acceptance criteria for model calibration to baseline surface water quality. This process is called "model corroboration" in Barr's QAPP, and "performance of acceptance criteria" in EPA QAPP modeling guidance (EPA QA/G-5M). An option would be a statistical test (e.g., t-test) to determine whether measured and model values for solutes in surface water probably represent the same populations. This would add symmetry to the report by applying metrics to model calibration that are analogous to those applied to mass balance (e.g., the technical review checklists require +/- 1% mass balance error for Modflow and <5% for hydraulic continuity in XP-SWMM models. |
| 3 | 4 | 4 | Under the explanation of "test cases," add the term "bench mark." Benchmarking generally means comparing rates in a model to an independent source (another model, analytical solution, or measurement), and is used widely in model guidance. This concept is covered in Barr's QAPP, and adding this term would identify this for reviewers. |

| Reviewer | Carlson Response | PolyMet/Barr Response | Carlson response |
|------------|--|--|--|
| H. Kempton | | Reference to this document is provided in Section 1.0. | Comment resolved |
| H. Kempton | Barr is developing document to disclose the calibration statistics and describe goodness of fit justifications, QAPP could reference this document | Comment refers to the compariston of model predicitions to observations within the Embarrass River and tributaries and is not directly realted to model calibration. Text added to acknowledge that this comparison will be quantitative but doesn't specify whether a specific statistical test will be used. | Barr to reference in QAPP the forthcoming document that discloses corroboration statistics |
| H. Kempton | | Text has been added to acknowledge that this is also known as benchmarking. | Comment resolved |