Final Peer Review Summary Report

External Peer Review of Wobus et al. 2012: *Potential Hydrologic and Water Quality Alteration from Large-scale Mining of the Pebble Deposit in Bristol Bay, Alaska*

November 2, 2012

Peer Reviewers:

Michael N. Gooseff, Ph.D.
Andrew M. Ireson, Ph.D.
Thomas Meixner, Ph.D.
John D. Stednick, Ph.D.

Contract No. EP-C-12-045
Task Order 005

Prepared for:
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment
1200 Pennsylvania Avenue, NW (8623-P)
Washington, DC 20460

Prepared by:
Versar, Inc.
6850 Versar Center
Springfield, VA 22151

TABLE OF CONTENTS

I. INTRODUCTION

In May 2012, the U.S. Environmental Protection Agency (EPA) released a draft report entitled *An Assessment of Potential Mining Impacts on Salmon Ecosystems of Bristol Bay, Alaska*. The purpose of this report was to put forth a prospective risk assessment of large-scale mining in the Bristol Bay watershed, focusing on a specific case study for a hypothetical but realistic mine scenario at the Pebble deposit. Specifically, the assessment examines how future large-scale mining may affect water quality, habitat, and salmon fisheries in the Bristol Bay watershed. During preparation of this draft assessment, EPA identified the following report developed by non-EPA scientists that contained information relevant to this topic, but was not included because it had not been peer-reviewed: *Potential Hydrologic and Water Quality Alteration from Large-scale Mining of the Pebble Deposit in Bristol Bay, Alaska: Results from an Integrated Hydrologic Model of a Preliminary Mine Design* (Wobus et al. 2012).

The purpose of this letter peer review is to determine if the information contained in the Wobus et al. 2012 report is of sufficient scientific quality and credibility to be incorporated into EPA's revised Bristol Bay report.

PEER REVIEWERS

Michael N. Gooseff, Ph.D.
Pennsylvania State University

Andrew M. Ireson, Ph.D.
University of Saskatchewan

Thomas Meixner, Ph.D.
University of Arizona

John D. Stednick, Ph.D.
Colorado State University

II. CHARGE TO REVIEWERS

The report to be reviewed is entitled *Potential Hydrologic and Water Quality Alteration from Large-scale Mining of the Pebble Deposit in Bristol Bay, Alaska: Results from an Integrated Hydrologic Model of a Preliminary Mine Design* (Wobus et al. 2012). This report evaluates the reliability of pre-mining water quality predictions at U.S. hardrock mining operations, and analyzes the most common causes of water quality impacts and prediction failures. Please provide detailed explanations for responses to the charge questions below.

Charge Questions:

1. Are the conclusions of the report well-supported by the evidence provided? Why or why not?

2. What are the strengths and weaknesses of the Wobus et al. 2012 report, in terms of:

   a. Modeling approaches used to estimate streamflow changes (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

   b. Modeling approaches used to estimate copper leaching (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

   c. Application of modeling results to evaluate potential effects of streamflow changes and copper leaching?

3. Are there important limitations or uncertainties associated with applying results from the Wobus et al. 2012 report to the hypothetical mine scenario evaluated in the EPA draft report? If so, what are they?

III. GENERAL IMPRESSIONS

Michael N. Gooseff

This report outlines the findings of a modeling study to determine the impacts of mine development (i.e., pit development, establishment of tailings piles, processing, etc.) and associated impacts on stream water quality and quantity in the area of the proposed Pebble mine. In my opinion, the model choice and the objective to represent the coupled nature of the real system with the numerical model are both laudable. The data collected from publicly available sources including meteorology records, soil distribution, digital elevation models, etc., all appear to be accurate. Assuming that the model output was appropriately interpreted, I am confident that the results are accurately reported. The presentation is generally clear (I have a few further comments on assumptions made in the modeling that are detailed below), and based on the assumptions that went into the model (i.e., boundary conditions, state variables, etc.), the conclusions are sound.

Andrew M. Ireson

My comments focus on the credibility of the modeling and the conclusions based on outcomes from the model. I am unable to comment on some of the chemistry, in particular I would exempt most of the text on pp. 17-21 from my review. The modeling approach is defensible in principle. The problem is that it is extremely ambitious, in particular with respect to the groundwater component, and that insufficient evidence is presented that the model captures the system behavior adequately. The model will never be perfect, since this is a complex system. This means, unfortunately, any assessment of the credibility of the model will be subjective. By my judgment, as the report is currently presented, there is too much ambiguity in the groundwater simulations to be able to have a high degree of confidence in the predictions. I think the model outcomes and conclusions are plausible, but with a fairly low degree of confidence in the detailed predictions. If a revision of the report were considered, it would be relatively straightforward to include the information I believe is missing, outlined in Question 2a below, and a much more rigorous and informed assessment of this modeling study would be possible.

Thomas Meixner

The report by Stratus Consulting to the Nature Conservancy on the hydrologic and water quality impacts of a proposed mine in the watershed of Bristol Bay, Alaska is generally well written and reasonably supports their conclusion that there are potential adverse water quality impacts from mine operations. The report takes a simple and straightforward approach to modeling the coupled surface and groundwater system. They do so by using generally available or locally available data sets to build a MIKE-SHE model of the system. A key challenge for this model or any model of this system is the general lack of data, which makes any modeling effort subject to significant uncertainty. The authors make a number of assumptions that make their results subject to scrutiny. First, they assume copper (Cu) is transported conservatively in the subsurface. This is highly unlikely since Cu can be subject to a large number of chemical reactions in the subsurface. A counter point to this criticism would be that acid mine drainage can have sustained effects in altering the geochemistry of subsurface systems. Second, an assumption that makes their results have a lower impact on the environment than they otherwise would is the assumption that the mine will only operate for 25 years. These along with other assumptions the authors make result in a study that supports its conclusions, particularly the two regarding the need for extreme efforts to reduce leachate leakage and the need for further study.

John D. Stednick

As important as the environmental assessment of the proposed Pebble Project, the writing and tone of the report suggests less than an objective approach. In my estimation, there are three major issues in the report that cause this concern: 1) the use of a stable copper release rate is an oversimplification of the chemical dynamics that occur in the tailings/waste rock. The physical environment in Bristol Bay will significantly affect the chemical weathering processes and rates; 2) since the humidity cell test did not follow ASTM procedures, the leachate concentrations were increased by 1 standard deviation. No justification for this inflation was provided; and 3) The report inseparably links copper concentrations to streamflow rates by stating that copper was treated as a conservative constituent, which is incorrect given the variability of water quality in the study area and hence speciation of copper.

The report represents a significant work effort in data compilation, model calibration, and validation. The MIKE-SHE model is data intensive and model results of hydrologic responses are good. The model has been criticized as being over-parameterized, but has been used in a number of environmental assessments. In general, the model assumptions need more clarification or defense of the positions chosen.

IV. RESPONSE TO CHARGE QUESTIONS

Question 1. Are the conclusions of the report well-supported by the evidence provided? Why or why not?

Michael N. Gooseff

The conclusions, based on the interpretation of the model results, appear to be sound. Given the decisions that went into developing the simulated structure of the natural system – for example, the choice of thickness and properties of the subsurface, influence of vegetation, etc. – the conclusions of the report are well supported by the model output. My interpretation of the results is that the strong local impacts of mining operations at the head of the watersheds are expected to have a significant impact on streamflow generation in streams closest to direct mining activities, and have less and less influence further downstream, as streamflow generated from other, unimpacted parts of the landscape emerge in streams. Given the decisions for how the subsurface is represented in the numerical model, my opinion is that this is an expected outcome of the model. I am, however, a bit concerned about the assumptions that went into the model design, and in particular, the heterogeneity of the subsurface near the proposed pit location. I have detailed these concerns below (Question 2).

Andrew M. Ireson

The conclusions are weakly supported by the evidence provided.

Why: The conclusion that the model produces a "very good qualitative fit" to observations and "does predict the general degree and direction of potential impacts" is weak, for reasons I explain below (Question 2a). The conclusions about mine impacts are dependent on the model and, therefore, those too are not strongly supported.

Thomas Meixner

The conclusions are well supported by the evidence provided. They show that if leachate management fails for even a brief period then Cu concentrations would be high enough to adversely affect salmon.

John D. Stednick

The conclusions as currently written, read more like a summary of the findings rather than specifically identifying the modeled results. Quantitative model results are not presented and some of the comments read like editorial opinions rather than reporting scientific results. The groundwater-surface water interaction was recognized, but model efforts were not adequately described. Comments like 'a very good qualitative fit' and 'does predict the general degree and direction of potential impacts' (both on page 39) are value judgments rather than conclusions.

Other Comments:

'Reduced streamflow and water quality as measured by copper concentrations.' Water quality constituents are not measures of streamflow. 'Copper concentrations as measured may be acknowledged but persistence is not identified.' Copper concentrations were modeled with streamflow, thus, a streamflow change would result in a copper concentration change and exceedance could be quantified.

The Knight-Piesold reference in the conclusions is online, but the materials available were surprisingly void of detail and water quality data. Better support is needed.

Question 2. What are the strengths and weaknesses of the Stratus report, in terms of:
a. Modeling approaches used to estimate streamflow changes (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

b. Modeling approaches used to estimate copper leaching (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

c. Application of modeling results to evaluate potential effects of streamflow changes and copper leaching?

Michael N. Gooseff

a. Modeling approaches used to estimate streamflow changes (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

The strength of the modeling approach used is that the fully-coupled nature of the real system is at least mostly represented in the MIKE and MIKE-SHE codes. The coupling of atmosphere to surface to subsurface, and the processes by which each are coupled is important for the spatially and temporally distributed evaluations of streamflow generation in this case. I do not have any significant reservations about the modeling code that was chosen, or about the choices for discretization, but I have some concerns about the model calibration and validation and the representation of the state variables in the model. With respect to the model calibration and validation, the strength of the approach used is that it incorporated comparison of model simulations to observed streamflow and groundwater observations (Figures 6 and 7 in the report). One particular weakness, however, is that there is substantial mis-match in streamflow magnitude and timing between the model and the observations, particularly during the fall and winter periods (i.e., September through January). I would not be surprised if this was due in part to problems with the observations at these times (ice formation is inherently problematic for interpreting pressure sensor data). The general behavior of the annual hydrographs is indeed captured, and this suggests that the model may be representing the culmination of processes that produce streamflow well. However, there remains the potential of getting the right answer for the wrong reasons; that is, behavioral agreement between simulations and observations could be achieved by many potential combinations of moving water through/across this landscape. For example, the report states that a wide range of bedrock hydraulic conductivity values reproduced measured streamflows equally well. This would indicate that the model is producing most if not all of the streamflow from the overlying unconsolidated layers. This is not stated directly in the report. Further studies of natural (i.e., geochemical and isotopic) tracers may help better constrain this source water partitioning. The same issue of model design may apply to heterogeneity of the unconsolidated layer. The most dramatic finding stated here is the strong potential influence of mining activities on the quantity of streamflow generated at the headwater locations, largely due to consistent dewatering activities within the open pits. Given the potential to reduce streamflows by 60% during high flow events, the model must be very sensitive to the changes in the model, imposed by the mine’s presence. My most significant concern is that this most sensitive portion of the model domain be accurately represented. It is not clear to me that the spatial representation of the unconsolidated surface layers and bedrock layers are accurate in the model. I recognize that the modelers are working without perfect information across the watershed. The information that may significantly inform this gap may very well be privately held or even proprietary. However, this location is of particular concern as it is proposed to be the most heavily impacted, and apparently also the most important as a headwater stream source. Finally, it goes without saying that the future meteorological, and therefore hydrological conditions, will not be the same as in the observed record. Therefore, the comparison of model output to modified conditions is reasonable, but should not be considered a prognostication for the future.

b. Modeling approaches used to estimate copper leaching (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

I cannot comment on this extensively as the geochemistry of copper leaching is not my expertise. However, I note that the simulated copper concentrations vary inversely with streamflow in Figure 13. At site UT100D, this makes sense because streamflow at that site is largely sourced from land that would be impacted by the mine. However, the lack of any potential interactions of the dissolved copper in the stream as it travels from UT100D to UT100B suggests this is perhaps a worst-case result for this site. That is, I would expect dissolved copper to be retained by hyporheic exchange, possible biological uptake, or complexed with other materials (particularly DOM) while being transported downstream.

c. Application of modeling results to evaluate potential effects of streamflow changes and copper leaching?

My following comments will focus on the streamflow changes only. Ultimately, the modeling scenario is one potential scenario choice for the forcing functions and expected changes to the system. While I have confidence that the model is accurately manifesting the simulated changes to the system and how it operates, I am more concerned that the simulated changes to the system state are accurate, and that the system is well-parameterized in the subsurface near the proposed mine site. It is critical to get this characterization correct in the baseline model for proper comparison in the future condition.

Andrew M. Ireson

a. Modeling approaches used to estimate streamflow changes (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

The general flow modeling methodology is sound, but I do have some serious criticisms of the presentation of results and hence credibility of the model. MIKE-SHE is an appropriate platform. The flow model is built deductively on the basis of an understanding of the surface water and groundwater processes in the catchment. Sensitive parameters (specifically hydraulic conductivities) were calibrated to improve simulations of streamflow at various gages, but “extensive tuning of the model parameters,” was not performed (p. 39); hence this is not a rigorous calibration. The model will not be perfect since this is a complex system, so the authors’ approach of not trying to make “specific numeric predictions,” but rather to look at potential effects with a model (p. 3) is sound. The credibility of the model, therefore, rests on its ability to reproduce the observed behavior of the system. Setting up credible groundwater models is data intensive and time consuming, in particular in the calibration stage. I believe that there are sufficient data (see Schlumberger, 2011, referenced in the report) to do this for the Pebble mine area. However, in the Wobus et al. report, despite the fact that in the discussion various conclusions are based on outcomes from the groundwater model (see section 4.2.2), very little evidence (i.e., figures showing) that the groundwater model is a plausible representation of the system is presented. A critical look at Figure 7 (showing only annual average groundwater levels, with biases of up to 30 m) suggests that the groundwater model performance is probably poor. On the other hand, Figure 8 is very nice and supports the model credibility. These are the only two figures in the report presenting results from the groundwater model. At a minimum, to consider this model credible, I would want to see at least one figure showing the spatial pattern of simulated water levels across the study area, at least one figure showing the simulated and observed seasonal patterns of water levels at key locations, and at least two figures showing the impact of the mine on the groundwater levels (the depth and extent of the cone of depression that is extensively discussed and is a critical outcome) locally (in the direct vicinity of the mine) and over the entire model domain. All of this information must be available and was probably looked at by the authors. The fact that it wasn't included in the report is a serious shortcoming. This is important because the surface water systems, and salmon habitats that depend on them, are groundwater fed. Hence changes in the groundwater patterns will impact both flows and water quality, as copper leachate may be transmitted through groundwater pathways into the streams.

b. Modeling approaches used to estimate copper leaching (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

I am not qualified to discuss this in depth, but I do have two criticisms of this component of the report: 1) the (extremely simple) equation on page 17 is a mess. Symbols should be used and units should be clearly distinguished from the symbols. "L" appears to be used both as a symbol representing a volume and as a unit, presumably liters, so that the sentence below the equation makes no sense. This must be corrected in the report, however, I don't believe that what was done is actually wrong. 2) On page 19, paragraph 3, it is stated that one standard deviation was added to the concentrations of the waste rock leachate, because it was felt that the method used (HCT) produced an underestimate. There is no justification provided for the choice of adding one standard deviation, and this could be seen as an attempt to bias the outcome of the study. I suspect if the "underestimated" leachate concentration was used, it would not change the outcome of this study. In any case it is necessary to justify rigorously the value chosen, not arbitrarily add one standard deviation.

c. Application of modeling results to evaluate potential effects of streamflow changes and copper leaching?

In terms of predicting the impact of the mine, the report considers four scenarios – leachate discharge mitigation, no mitigation, mitigation failure for one month, and mitigation with failure for 6 months. The suitability of these scenarios is beyond the scope of this review, and should be examined carefully. There appears to be no discussion of the first scenario. The results from this scenario should be included. For the latter three scenarios, we can say without any modeling that, based on the assumptions being made, there will be some impact on flow and water quality. A model is needed to quantify the magnitude and duration of impacts on flow and in-stream copper concentrations. I find the general approach here, described in section 3.2.3, to be appropriate and well justified in the text. The major problem comes back again to the credibility of the baseline model (see my comments in Question 2a above). Secondary problems that are less critical are: 1) no results showing copper migration through the subsurface are presented (it should be easy to plot a plume for each scenario), and 2) despite listing a number of uncertainties in section 4.2.3, no attempt to quantify the impacts of these uncertainties was made. For the latter point it would be quite straightforward to use an ensemble approach to explore a range of possible leachate concentrations, and possible, though less straightforward, to look at a range of possible hydraulic conductivity values.

Thomas Meixner

a. Modeling approaches used to estimate streamflow changes (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

The model is explained in some detail and references are provided for more information. The appendices are also useful. It would be nice to have more performance statistics on how well the model performed. Calibration approach seems reasonable for a low data environment and certainly provides a sufficient calibration/validation spatial and time series split.

b. Modeling approaches used to estimate copper leaching (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

The hydrology of the model is robust, but the assumption of conservative Cu movement in subsurface and surface system is flawed.

c. Application of modeling results to evaluate potential effects of streamflow changes and copper leaching?

On stream flow changes, the modeling results are as robust as can be expected in a data sparse environment. On the other hand, the assumptions regarding Cu transport are flawed.

John D. Stednick

a. Modeling approaches used to estimate streamflow changes (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

The MIKE-SHE model is a parameter intensive model and once validated can be used to assess the effects on streamflow by modification of input parameters. The model is spatially explicit. The time-step is 3 hours from the NARR precipitation time-step, a higher resolution than most models for wildland settings. Although the authors recognize parsimony, it would still be useful to look at parameter sensitivity and scaling differences. How does a 3-hour precipitation rate on a 30km grid compare to soil surveys from 1:100,000 STATSGO on a 29m DEM?

The hydrologic characterization of the study area needs elaboration. The report states that the area has two different types of streamflow, snowmelt events in the spring and rainfall events in the fall, yet hydrograph shapes would suggest that it is not that simple. Even for the larger watershed areas, snowmelt peak streamflow rates occur in a matter of hours and have a comparable recession to rain dominant events. Precipitation events as rain can be modeled easily; the snowmelt events require a more complex approach to modeling. How was the degree-day model determined? What is the frequency of rain-on-snow events or rain-on-frozen ground? The model ‘underestimates the peak streamflows in the fall and thus underestimates the potential reductions in streamflow from mining activities.’ Why was this not better addressed? The NARR data are presented as annual and monthly data in the appendix, and a cumulative plot of daily (?) precipitation in Figure A.6. Since the 3-hour precipitation data were used in the model, why the explanation of these other time steps?

‘A number of gaining and losing reaches, and in some areas, streamflow is sustained by groundwater transfer from losing stream segments in other watersheds (Knight Piesold, 2011).’ The importance of the hyporheic flows in maintaining open waters is important. From other works, we know that the hyporheic flow networks are significant. The current report needs additional text to explain and clarify modeling efforts made to identify the groundwater upwelling. How was a 250m wide channel decided upon? Since groundwater is important, why were annual average elevations used? A specific effort was made to model groundwater movement from SK100C to tributary UT119A, based on Knight Piesold (2011). Are there other tributaries that are influenced by trans-basin groundwater movements? I could not find this information in the cited work.

b. Modeling approaches used to estimate copper leaching (e.g., in terms of general methodology, time steps, model calibration and validation, input data)?

The authors have done a good job characterizing the existing water quality and documenting the variability in pH, TOC, hardness, and show temporal and spatial variability in dissolved copper concentrations in surface waters (Figure 3), and the change in water quality standards with water hardness. Thus, it was surprising to read that copper was treated as a conservative constituent (page 38). The oversimplification of the dynamics of copper is bothersome, especially coming from a consulting group known for its toxicology work.

The range in water quality data (Table 1) show seasonal dynamics that are ignored in the copper solubility discussion. The assumption that copper concentrations in the leachate will reflect changes in surface water quality is doubtful. Since the humidity cell test was modified from ASTM standards (never justified) and they used a larger sample volume with larger particles (page 17), the predicted leachate values were increased by one standard deviation (page 19). No defense for this decision.

c. Application of modeling results to evaluate potential effects of streamflow changes and copper leaching?

Streamflow reductions were modeled under the mining scenario. Flow reduction was caused by precipitation interception and groundwater depression by the mining activities. Groundwater levels were recognized as having ‘strong variability’ in elevation, but annual average points were used.

'Without ongoing active management, these changes in streamflow and water quality would likely result in adverse effects on aquatic biota, ranging from reductions in habitat quantity and quality to acute or chronic toxicity to aquatic organisms.' This statement is from the Conclusion section (page 39). Presumably, this is a reference to salmonid habitat, but habitat was not included in the modeling effort. None of these observations are defended in the report and suggest a lack of objectivity. This lack of objectivity tempers the study results and leaves me questioning other results.

Question 3. Are there important limitations or uncertainties associated with applying results from the Wobus et al. 2012 report to the hypothetical mine scenario evaluated in the EPA draft report? If so, what are they?

Michael N. Gooseff

As I outlined above, my chief concern related to the most significant conclusions of the report is that the structure of the subsurface is adequately represented near the proposed pit areas. The change in hydrologic function of this part of the landscape is clearly reflected in the 'impact' simulations. My sense is that the model is limited by the data available to discretize and characterize the state variables within the domain. Hence, these results are useful to indicate potential change, but there are many other potential scenarios that could be implemented with respect to mine operation that might just as likely and directly influence the hydrology of headwater streams in the basin.

Andrew M. Ireson

Yes, the credibility of the model is questionable, as I have described in Question 2a.

Thomas Meixner

The modeling, as always in such data sparse environments, is highly uncertain. The authors make assumptions that try to minimize this uncertainty. In particular, their assumptions about conservative Cu transport are reason for concern. Their hydrologic representation appears to be very robust however.

John D. Stednick

No. The proposed mine scenario is from the PLP plan prepared by Ghaffari et al. and is a reasonable approach.