Best Practices Checklist for Modelling Mine Waters

Jerry Vandenberg[1], Kristin Salzsauler[1], Scott Donald [2]

[1]Golder Associates, Vancouver, BC, Canada
Jerry_Vandenberg@golder.com, Kristin_Salzsauler@golder.com
[2]Golder Associates, Cambridge, ON, Canada
Scott_Donald@golder.com

Abstract

Modelling the flow, concentrations and load of mine waters often entails the use of different tools and software to represent different aspects of the mine and the receiving environment. However, there are steps common to all modelling applications that can improve accuracy, transparency and interpretation by the modeller. These steps will also help the client who may need to make management decisions based on the model results, or regulators or stakeholders who will ultimately rely on the model results to make permitting or policy decisions.

This paper provides a 15-point modelling checklist, along with examples of what to do and what to avoid, drawn from case studies. The case studies include modelling completed for oil sands, diamond and copper mining projects in Western and Northern Canada. The models used in the case studies include mass balance, hydrogeological and geochemical modelling of mine waters; dispersion modelling of effluent; hydrodynamic modelling of pit lakes; and watershed modelling in the receiving environment.

The best practice modelling checklist is as follows: 1. Define objectives for the model; 2. Develop a conceptual model; 3. Establish output metrics; 4. Establish screening criteria; 5. Select the model framework; 6. Gather input data; 7. Implement quality assurance procedures; 8. Set up model physical domain, boundary conditions, and parameterization; 9. Calibrate stepwise: physical, chemical and biological; 10. Simulate scenarios; 11. Compare output to criteria; 12. Iteratively modify mitigation, models and screening criteria, as appropriate; 13. Conduct an uncertainty and/or sensitivity analysis; 14. Identify and simulate “black swan” events; and 15. Conduct a post-audit.

Key words: Modeling, calibration, validation, hydrodynamic, dispersion, mine waters

Introduction

Many types of models are applied to mine waters. The basic mine water model is a water balance, which is often constructed as a “spreadsheet-like” accounting system. Components of the water balance may be modelled in more detail using numerical models, such as a hydrogeological model or a hydrodynamic model. A model’s utility and range of application can be improved by coupling it with a larger water quality model. Generally this requires a geochemical model to derive the composition of contact waters from waste rock, tailings, ore, and roadways; waters originating from the mill; and other site-specific inputs. For complex moving and dynamic flows, a hydraulic, hydrodynamic, or watershed model may be required. Most mines require a hydrogeological model to understand pit dewatering, refilling and exfiltration. A whole-system model that represents the mine throughout the life of mine and post closure will actually require most of the aforementioned models in a system of linked or coupled models. All of these types of models share a common set of fundamental steps, along with their unique set of requirements, limitations, uncertainties, data needs and quality assurance practices. It is the aim of this paper to present a basic set of “Best Practices” that are applicable to mine water models, based on a number of case studies drawn from projects that included multiple sets of coupled models.

Recognizing that there is no “recipe book” approach to modelling that will fit all applications, the list of steps presented in this paper is put forward for consideration when developing a model and serves as a checklist for reviewing models by others.

1. Define objectives for the model

Definition of the model objective(s) is the most important step in mine water model development, as it determines many aspects of subsequent steps. The objective should be clearly defined and agreed to by the entire modelling team and the client. There may be multiple objectives for different scientific disciplines within a project, but there must be one defined objective for a given model. A few examples of a mine water model objective are to: test a hypothesis; support a permit application; evaluate mitigation; conduct an environmental assessment; forecast future conditions; guide research by identifying knowledge gaps.

The objective should be clear, actionable and achievable. All subsequent modelling tasks should be aimed at meeting this objective. For example, the choice of inputs to a model will change depending on the objectives. For permitting, the emphasis will be on constructing a defensible model, and inputs which are known to be biased to produce conservative results are common and acceptable. Conversely, for hypothesis testing or design work, accurate or likely inputs are usually preferred over conservatively high inputs.

In systems of linked or coupled models (e.g., Vandenberg et al. 2015), there may be multiple modellers simulating different aspects within the overall model domain, each with individual objectives. While it is important for the entire team to be working toward a shared overarching objective, it is also important that modellers understand each other's individual objectives as those may affect the data required, as well as how the data being transferred from one model to another are generated.

2. Develop a conceptual model

A conceptual model is a description of the system of study, preferably accompanied by a schematic or diagram. Conceptual models can be the traditional source-pathway-receptor diagram as used in risk assessment, or more commonly in mine water applications, they are schematics of mine facilities with simplified water balances that illustrate storage, inflows and outflows. More complex, mechanistic models may warrant detailed conceptual models. Generally the complexity of the conceptual model informs the complexity of the numerical model; it follows that if a process is important enough to illustrate in the conceptual model, it is likely relevant to include it in the numerical model.

The conceptual model serves two main purposes. First, it helps the modelling team brainstorm and agree on which processes are most important to the system of study. Equally important, it communicates these important processes to clients, reviewers, regulators and stakeholders who will have a more limited view of the modelling process that was used to generate the results they will rely on to make decisions.

3. Establish output metrics

Before commencing the processes of numerical model development, it is important to consider the type of output that is required to achieve the model's objectives. The most basic information will be flow and concentration; virtually all models will be capable of producing this information. However, the granularity of output required can vary greatly depending on the system to be modelled. For example, knowledge of high-flow and low-flow events are often the most critical to results even though they occur infrequently. Similarly, high concentrations may occur with limited spatial and temporal extents, but may be critical to the overall system (e.g., a seepage to spawning habitat during low flow). Knowing which types of model output will be required and the temporal periods of interest will help select the right model platform, and/or help plan for level of effort in developing post-processors that convert or format the model output.

4. Establish screening criteria

In most mine water applications, model results will need to be compared to some threshold or criteria. Establishing the screening criteria in advance will help interpret results as they are being generated. It is not unusual to generate several iterations of results; a pre-defined set of criteria make it easier to judge whether mitigation or model refinements are necessary.

Ideally, the frequency of model output should be aligned with the given screening criteria. For example, many water quality guidelines are derived from (and are therefore applicable to) a given exposure duration. If the model cannot meaningfully output as frequently as the specified criteria due to an inherent limitation such as model time step, the model or criteria will need to be adjusted. Conversely, if the model outputs too frequently, the results can be post-processed for comparison to screening criteria.

Will you be concerned with relative changes or absolute values? When dealing with low water concentrations, the relative change can be very large but not ecologically significant. In contrast, a relative change may be required to determine significance when the results are near the margins of a threshold. In some models, a relative change may be all that can be confidently relied upon, because there is a high uncertainty in the absolute values. Provided a transparent reason is stated, either of these types of output may be justified.

It is sometimes useful to have two or more levels of criteria, based on return period of flows, biological response, generic guideline, or regulatory limit. The frequency of compliance for the different levels of criteria may differ, which is another capability to consider when selecting a model.

Most chemical models can output all water quality parameters that are simulated, but the units can vary. If this is going to be the case, plan for pre-processing and additional quality control of results. Where models are linked or coupled, the output of one model becomes the input to another model. In this case, it is critical to have matching input/output frequencies, spatial nodes, and units.

Regardless of the model output and screening criteria, it is up to the modeller to provide conclusions regarding the model results. For the model to have maximum credibility and persuasiveness, the justification for the screening criteria should be clearly laid out.

5. Select the model framework

There are many software packages available to develop a mine water model, and even multiple choices available for a given sub-discipline within the broader category of mine waters. A num-ber of criteria should be considered when selecting a framework, including: built-in functionality, dimensionality, availability of input data, ability to simulate sensitivity analysis or Monte Carlo simulations, scalability, linkages to other models, and computational efficiency.

Of these criteria, built-in functionality and dimensionality will normally be the first differentiators. If hydraulic, hydrologic, hydrogeolologic, hydrodynamic or geochemical functionality is required, the choice of model for those processes is limited. For analytical (or semi-analytical) water and mass balance models, generic software such as Excel, GoldSim or Matlab are commonly applied. These programs may also be used to integrate the results from several more complex models within the overall system. A good rule of thumb for selecting a model is that it must be capable of simulating the processes illustrated in the conceptual model.

Computational efficiency can be a major factor when simulating mine closure, because the time frames of interest extend for decades to centuries. A long-term simulation to understand the vertical mixing regime of a pit lake, for example, can take days to run, even if care is taken to make the grid efficient. Likewise, transient simulations of groundwater flows to a refilling pit can take substantial amounts of time. If this is the case, consider standalone simulations using the minimum required dimensions (which are site-specific) to reduce total run times. Similarly, if the input data required for a complex 3-D model are not available, consider using a more simple system until additional data can be acquired.

The choice of open-source versus proprietary software is sometimes made on the basis of license fees, but there are additional considerations. The fees associated with proprietary soft-ware often include technical support, which is important for new modellers who do not have a large support network within their institution. The drawback of proprietary models is that reviewers and stakeholders may not have access to the software, in which case they sometimes view it as “black box” software, which, rightly or not, can diminish the credibility of the overall project.

Lastly on this topic, modellers should be wary of becoming too invested in a given software package and automatically applying that tool to every mine water application. The applicability of a given software package to one mine site does not necessarily confer applicability to any other site.

6. Gather input data

The task of gathering input data can be one of the most time consuming and challenging; the effort required is often under-estimated. In mine water model applications, it is not unusual to have sub-sets of the input data maintained by the mining company, previous consultants, various government agencies, and combinations thereof. For example, the mining company may have a site-specific weather station that has detailed meteorological information – but from a limited time period that needs to be integrated into a long-term dataset from a nearby government station. Similarly, baseline water levels, flows and hydrogeological information may be available from neighbouring developments or with another consulting firm, and these may need to be adjusted based on long-term datasets from a regional monitoring station that is maintained by government. Geochemical data used for input to water quality models requires both computational organization, and an evaluation of the results of geochemical testing in the context of the conceptual model (e.g., the estimated metal leaching and acid generation potential of materials in laboratory conditions versus field conditions, the effect of project specific mitigation on the acid generation potential of a material, etc.). Each of the organizations supplying data will inevitably use different formats and conventions for storing data, which means the modeller may need considerable time to re-format and assemble datasets once compiled.

Once all data are obtained and compiled, there will often be gaps in time or space that need to be filled. For geochemical data, if site-specific information is not available for all water quality parameters, proxy data from a mine with similar geological characteristics may be applicable, such as that presented in Plumlee et al (1999).

Be aware that using proxy data of any type (i.e., geochemical, hydrologic, meteorologic, etc.) will increase the uncertainty of results. Recommendations for monitoring to reduce data gaps should accompany any model results that were derived using incomplete data sets, if the gaps could affect the conclusions drawn from the model.

Ideally, an independent dataset should be used for calibration and validation; that is to say, calibration and validation data should not be used as inputs to the model.

7. Implement quality assurance procedures

Quality assurance (QA) is an overarching process that begins before modelling and includes quality control (QC) of field and laboratory data. Ideally, the quality control applied to model inputs and outputs will be integrated into an overall QA program and the modeller will have input into the data collection process.

In practice, most mine water models are developed by modellers after the data have been collected by other groups and provided to the modeller. In this case, the modeller should examine the QC steps and results that have already been applied, such as screening of analytical data for outliers, ion balances, checking total versus dissolved metals, measured versus calculated TDS, etc. If these types of checks have not been documented, they should be completed prior to accepting data as valid.

After data have been validated, compiled and assembled into input files, all input data should be graphed in a time series. This applies to flow, chemistry, meteorology or any other time-varying input. If ranges of dates have been interpolated or substituted to fill gaps, the raw data should be graphed along with the synthetic timeseries to check for differences. This QC step has proven to be a reliable method for exposing anamalous data in model inputs. At a glance, one can usually spot the following types of errors in an input dataset: outliers; zeros or nulls in a dataset; negative values; changing detection limits; and nonsensical values, which are those that do not fit the data type, season, or variable.

While the visual method can usually identify extreme outliers, batch statistical tests can be used to identify statistical outliers. "Outlier" is a collective term that refers to either a contaminant or discordant observation (Beckman and Cook 1983). A discordant observation appears surprising or discrepant, but the origin of the result is unknown. A contaminant is an observation that may induce bias to a fitted probability distribution, and the reason for its characteristics can be explained. If a measurement is identified as a statistical outlier, the origin of the result can be investigated to validate the results of outlier detection tests performed on the dataset. Outliers of explainable origins (i.e., there was a reason for the anomalous result) are considered contaminants and are generally justified to be removed from the dataset without further consideration. Outliers of unknown origin are considered discordant observations and may be retained in the dataset for further investigation.

Given that statistical tests can identify, but usually not explain outliers, a conservative approach to removing outliers based on statistical tests is to apply multiple tests, and only remove those that are flagged by all tests. Three commonly used tests are: the three times standard deviation (3σ) rule; the Hampel identifier; and the combined Rosner – Dixon test, descriptions of which are presented elsewhere (Lauzon et al. 2011).

As a final step in validating input files, compare the ranges of values and units to the original dataset to confirm that any unit conversions along the way have been done correctly.

During the modelling process, a helpful QC check is to compare preliminary results against simple hand calculations or spreadsheets to confirm that output falls within the right range. This step will identify gross errors, which are common during initial runs. A peer or senior review of inputs and coefficients provides additional QC during the model simulation step.

Final model output can also be checked against simple calcuations as a QC step. This can be done by the modeller or the peer reviewer. If the output from one model becomes an input to another model (e.g., flows from a hydrologic model being input to a water quality model), those outputs should be critically reviewed by the receiving modeller.

8. Set up model: physical domain, boundary conditions, parameterization

The model setup process will vary greatly depending on the type of model and complexity of the system. Generally applicable steps and lessons learned related to model parameterization are provided below.

Whether using a single model or a series of linked or coupled models, the first step is usually to set up a physical domain, which may include a grid or series of compartments in one or more dimensions. At discrete points or along planes of the domain, boundary conditions such as flows, heads, solar radiation and winds are applied. When linking multiple models, it is important for both teams to agree on the dimensions and fluxes that will be output and input at these locations. In models that have coupled hydrogeology-solute transport or hydrodynamics-water quality, the physical model should generally be set up, run and verified before proceeding to chemical inputs. Even in simpler systems such as a water and load balance, the water balance should be fully operable before simulating any chemical loads. Likewise, chemical parameters that involve fate processes, and biological parameters, should be included after all conservative chemical species have been set up. In models with complex feedback interactions, such as an eutrophication model, the set up and calibration will be an iterative process, because the simulation of one species can influence another species that was previously thought to be fully set up.

When generating model input files, be aware that many variables in mine waters will covary seasonally and may also exhibit autocorrelation. Knowing the degree of autocorrelation to include in a given input can be challenging. If the variable has been measured frequently, the historical values will demonstrate this pattern; however, sufficient datasets are usually only available for existing mines with stringent monitoring conditions, such as those with active discharges to the receiving environment. If a comprehensive dataset is unavailable, proxy datasets can be examined; the dataset with the most similar hydraulic residence time and seasonality will give a good indication of autocorrelation (and perhaps, by extension, a template to follow for the sparser but site-specific dataset).

Depending on the amount of calibration data available, parameterization may be a simple step of inputting values from previous applications, literature, or professional judgement. If calibration data are not available, a sensitivity analysis can be completed on key inputs to evaluate the potential effect on results if another value were chosen for a given coefficient.

Lack of calibration data is generally a knowledge gap that the modeller can identify and help rectify by recommending monitoring that could be undertaken in the existing or future mine water system that is being modelled. Collection of monitoring data is a key component of the post-audit, discussion in Step 15.

9. Calibrate stepwise: physical, chemical, biological

The process of calibration will also vary depending on the specific model applied, the complexity of the system and the amount of data available. The term 'calibration' itself is not consistently applied, but in this context, it is considered to be

> *the process of adjusting model rates, coefficients and unknown or unmeasured variables such that the numerical model produces a reasonable hindcasted match compared to past observations.*

'Validation' is a related step, usually undertaken after calibration, and is defined here as

> *comparing model predictions to measurements that are collected after the model predictions have been made; or, extending a calibrated model with recent boundary condition data, and comparing predictions to field data, without adjusting the calibration.*

Similar to the model set up, calibration should proceed stepwise from the physical domain to the chemical and then to biological, which reflects the order of key drivers acting upon the system as well as an increasing level of complexity.

When applying any fate process to a mine water model, which will commonly be a loss of some mass with time, consider the actual mechanism that is being simulated. Often in calibrating a model, some rate that is observed in an existing facility will be applied to a future facility. If that is the case, critically examine the following: Is the mechanism known to the degree of confidence that justifies applying the calibration to future settings? Are conditions in the observed system the same as those in the future system? Are byproducts or feedback loops being generated that should be accounted for elsewhere in the model?

10. Simulate scenarios

Model simulation is generally straightforward. The main things to consider are that the simulation period, output nodes of interest and output format should all fit the model objectives. It is a given that models are imperfect representations of reality, and the degree of imperfection within a model will vary with time and space. What is important is that the model reflects a level of realism at the time and location of interest to the predictions, in a way that achieves the overall model objectives.

11. Compare output to criteria

After generating model results, the metrics and criteria derived or set in Step 3 can be applied to the results. Again, it is important to consider the model's objectives when comparing results to criteria. For a detailed design project, a high level of accuracy may be required, whereas for a regulatory application or risk assessment, a lower level of accuracy may be acceptable, provided conservatism is applied to uncertain inputs.

While models may produce vast arrays of results, not all results may be applicable to the selected criteria. For example, water quality guidelines are often derived using some averaging period, which can differ from an instantaneous result, or a model result that is averaged over a different time step. In some cases, it may be possible to line up the applicable time periods; if not, a description of the difference should be noted to those who rely on tables of data that compare model results against criteria.

Likewise, when comparing a statistical result to criteria, the probability chosen to represent model results should align with the derivation of the criteria or the design objectives. Using probabilistic or Monte Carlo simulations, described more under Step 13, is a common approach to assigning a probability to a model result.

If a deterministic simulation is completed using a single statistic to represent inputs (e.g., 95th percentiles), the probability of the output from that simulation is not possible to estimate (except to say that it will be much less likely). Therefore, the results should not be labelled with the same statistic (i.e., the likelihood of an outcome that assumes multiple, independent events, each with a likelihood of one in twenty, will be far lower than one in twenty).

12. Iteratively modify: mitigation, models & screening criteria, as appropriate

It is usual in mine water modelling to require several iterations with refinements to any of the previous steps. Often, conservatively high inputs in the first round of modelling will lead to results that exceed screening criteria, and the modeller may adjust the inputs to be more accurate and less conservative. Similarly, the screening criteria themselves may be too stringent and may need to be discarded in favour of site-specific thresholds.

One of the main reasons to employ mine water models is to identify and select mitigation measures. If an initial simulation identifies the need to add mitigation, and that measure is adopted, it will need to be incorporated into the model to assess potential efficacy. The mitigation may take many forms, such as a water treatment plant, barrier wall, soil or water cover, water diversion, or changes to the mine plan. Accordingly, all parts of the model process, including the conceptual model, are subject to modifications and may require iterations. In some modelling projects of dynamic mine environments, a model scenario may be deemed 'frozen', so that the entire team has a fixed basis that can no longer iterate. This can be an essential step in preparing large regulatory approvals where the preparation of a linked set of models takes extended time periods. The modeller needs to be aware that assumed changes in one part of the model may lead to discrepancies in a regulatory application with another part of the model that has not been frozen.

Ideally, a mine water model will be constructed with the intention of keeping the model as a living tool that will updated throughout the life of mine and into post-closure. Updates can entail frequently appending input data that are continually being collected, and occasional updates to the domain as the mine plan progresses, as outlined in Step 15.

13. Conduct an uncertainty and/or sensitivity analysis

There are many sources of uncertainty to a mine water model. Some can be understood and quantified using standard methods, while others cannot. The former type of uncertainty (known unknowns) is discussed in this section, and the latter (unknown unknowns) in the next section.

Some types of uncertainty can be identified, understood and quantified (at least approximately). For example, analytical variability, seasonality, anisotropy and heterogeneity in a given system can be quantified or estimated through intensive monitoring. If you are fortunate enough to know the approximate range of values for a given variable based on a thorough dataset (i.e., adequate length, spatial coverage and frequency of measurement over a range of seasonal and climactic conditions), along with the probability of each value, then the uncertainty due to this single variable may be estimated. If the uncertainty profile of multiple variables are known, they can be combined in a stochastic simulation. If all dominant variables are well characterized, they can be combined in a Monte Carlo simulation to quantify known unknowns.

In order for the resulting output profiles to provide realistic likelihoods associated with each result, the variables that co-vary must be identified, and the covarying nature of each input must be aligned. The alignment need not be perfect, but any discrepancies will add uncertainty to the results, and that uncertainty can be hard to quantify. Correlation plots provide a quick way to establish temporal patterns among inputs, such as flow and TSS concentrations, and if these are not already linked mechanistically in the model, they can be linked statistically when deriving inputs and generating model input files.

The results of a Monte Carlo simulation can be evaluated over time or in aggregate – again, whichever fits the model objectives. In either approach, there is a likelihood associated with any concentration, which provides a basis for estimating and communicating the overall model uncertainty. But the key limitation to consider when applying these methods is that they only account for uncertainty that has been identified and incorporated into model inputs.

A more straightforward and less computationally and data-demanding exercise that can be used to quantify uncertainty is a sensitivity analysis. This is done by changing one variable, which could be a coefficient or time series of input data, keeping all other model inputs fixed, and re-running a scenario. The difference in results indicate the effect or ‘sensitivity’ to that single input. This can provide valuable information when the modeller cannot obtain specific input data, has uncertainty about aspects of the calibration, or cannot defensibly choose one input over another to apply to future conditions. The sensitivity analysis can answer the question of *what if some other input turns out to be real?* The limitation of this approach is that it may oversimplify the problem because few, if any, variables are truly independent of all other variables.

14. Identify and simulate "black swan" events

To discuss the unknown unknowns, we borrow and adapt two metaphors of the black swan to modelling:

The first black swan metaphor by Sir Karl Popper deals with falsifiability. Prior to 1697, every swan that had ever been observed by Europeans was white, which inductively led to reasoning that "all swans are white". Were a biologist to go searching for more swans to prove this theory, they would indeed find that all empirical evidence supported that theory and they might be tempted to accept this as "proof". However, with the discovery of a black swan in Australia in 1697, a single observation falsified that theory. In this context, a black swan is a single observation that falsifies a previously held theory.

The second metaphor by Nassim Nicholas Taleb deals with predictability. In his book *The Black Swan* (Taleb 2007), he describes events that come as a surprise, have a major effect, and are often inappropriately rationalized after the fact with the benefit of hindsight. His metaphor is primarily concerned with events that have a major impact on civilization.

Extending these metaphors to a mine water model, a black swan can be thought of simply as a process, event, or input to the actual mine water system that, if encountered in the future, will invalidate the model results because it was not considered by the model. A black swan event is not reasonably foreseeable and its probability cannot be quantified with any accuracy. Therefore, the approach to deal with black swans is to attempt to identify possibilities and to have contingencies available where they are identified. The challenge is to distinguish between black swans from implausible or impossible scenarios.

An example of a black swan event in a mine water context is a landslide into a meromictic pit lake with submerged mine waste. The likelihood and consequence of lake overturn and release of contaminants to surface water and introduction of oxygen to the submerged waste are possible to estimate considering a plausible range of limnologic and climactic conditions. But if an earthquake or other geological event intervenes, all bets are off. The earthquake is a black swan that would invalidate all model results. In this case, the event needs to be evaluated and planned for using means besides the model. If the model can be modified appropriately, it may be able to be used to understand the consequence of such an event by assuming it happened.

The modeller need not arrive at contingencies for all such events (though that would be a value added service). The modeller's challenge, and their responsibility, is to identify and communicate to managers or stakeholders the types of events and processes that may render the model results invalid – the black swans.

15. Conduct a post-audit

While there is some debate over the definition and utility of model validation and verification (Konikow and Bredehoeft 1992; Nordstrom 2012), it is in the interest of all to understand whether model predictions turn out to be accurate. In the broader context of managing mine waters, regulators need to know whether to place confidence in models; stakeholders need to know whether the clean water, mitigation, or other resulting prediction came true; mine managers need to be able to adaptively manage based on differences in reality versus what was predicted; and the modeller needs to know which predictions were accurate and which were not in order to improve their approach with subsequent models.

A post-audit is a series of measurements, collected after the model predictions have been posted, and during or after the mine development proceeds, which are compared to the previously made model predictions. A post-audit provides a means of testing the model predictions against reality and verifying whether (a) the model was accurate;(b) the assumed mitigation was effective; and (c) that the correct decision was made (i.e., that the model objectives were ultimately met).

Conclusions

A checklist of best practices is provided above for developing and applying mine water models. The most important steps are to set clear objectives and to develop a conceptual model, as these steps will affect nearly all subsequent steps. The final steps of identifying uncertainty and comparing predictions to post-development observations are the most critical in terms of understanding the limitations and performance of the model and for communicating those to stakeholders.

References

Beckman, R. J. and R. D. Cook (1983). Outliers. Technometrics. 25:119-149.

Konikow, L.F. and Bredehoeft, J.D. (1992). Ground-water models cannot be validated. Advances in water resources, 15(1), pp.75-83.

Lauzon, N., Vandenberg J.A. and J.P. Bechtold J.P. (2011). Probabilistic Modelling Applied to the Mining Industry to Address Water Quality Uncer-tainty. MODSIM 2011. Modelling and Simulation Society of Australia and New Zealand, December 12-16. Perth, Australia

Nordstrom, D.K. (2012). Models, validation, and applied geochemistry: Issues in science, communication, and philosophy. Applied geochemistry, 27(10), 1899-1919.

Plumlee, G. S., Smith, K. S., Montour, M. R., Ficklin, W. H., and Mosier, E. L. (1999) Geologic controls on the composition of natural waters and mine waters draining diverse mineral deposits, in Plumlee, G. S., and Filipek, L. H., eds., The Environmental Geochemistry of Mineral Deposits Part B: Case Studies and Research Topics, Reviews in Econ. Geology, vol. 6B: Littleton, CO, Soc. Econ. Geologists, p. 373-409.

Taleb, N.N. (2007). The black swan: The impact of the highly improbable. Random House.

Vandenberg, J.A., M. Herrell, J.W. Faithful, A.M. Snow, J. Lacrampe, C. Bieber, S. Dayyani and V. Chisholm (2015). Multiple Modeling Approach for the Aquatic Effects Assessment of a Proposed Northern Diamond Mine Development. Mine Water and Environment, http://dx.doi.org/10.1007/s10230-015-0337-5, 1-9.