Mine Water and the Environment (2020) 39:630–646
https://doi.org/10.1007/s10230-019-00634-0

TECHNICAL COMMUNICATION

A Quantitative Framework for Comparing Observed and Simulated Pit-Lake Geochemistry

Connor P. Newman[1,2]

Received: 24 October 2018 / Accepted: 13 September 2019 / Published online: 19 September 2019

Abstract

Pit lakes are commonly studied using advanced geochemical models. However, quantitative post-audits of these studies to assess the degree of agreement with observations are rare. This study used a quantitative evaluator (error variance parameter; Δ) to compare predicted and observed geochemistry of nine pit lakes across three continents and five ore-deposit types. The compiled modeling studies applied a variety of methods to incorporate geochemical (sulfide oxidation and mineral precipitation) and hydrologic processes (groundwater flow and hydrodynamics). Several common ions (Ca, Mg, SO_4) and pH were shown to compare well with observations. In contrast, comparisons of model simulations and observations for trace metals and metalloids such as Fe and As were less congruent. The discrepancies for these and other species is potentially linked to inappropriate application of equilibrium thermodynamics and temporal model discretization. Model complexity was also shown to be related to the reproducibility of observed geochemistry. Quantitative analysis using Δ in a wider range of studies may increase understanding of prevalent geochemical processes invoked in pit-lake modeling.

Keywords Model evaluation · Geochemical modeling · Quantitative error analysis · Mineral equilibrium · Thermodynamics

Introduction

Pit lakes are becoming more common as open-pit mining is more widely used and more pits are developed below the water table. These mining operations are often overseen by a state and/or federal regulatory body. Of primary importance to these regulatory bodies is the degree to which future pit-lake water quality can be reliably predicted. In addition, pit-lake modeling is also commonly applied as a purely scientific endeavor aimed at improving our understanding of fundamental hydrologic, limnologic, and biogeochemical processes (Castendyk et al. 2015a, b, and references therein).

A common question asked by regulatory agencies is the degree of certainty associated with a specific pit-lake model. This question has substantial bearing on the future of a mining project and the protection of environmental resources in the area of a proposed mine. Unfortunately, answering this question is difficult as many models of future pit lakes do not have any realistic comparison point, since dewatered pits often take many tens of years to hydrologically recover, and predictions can require centuries to reach geochemical pseudo-equilibrium. Because of this, the degree to which pit-lake geochemistry can be predicted is, in many cases, unknown. Also, although numerous pit lakes have been geochemically investigated (e.g. Sánchez-España et al. 2008; Shevenell et al. 1999; Tucci and Gammons 2015), the results of many such studies are not publicly available. It is also desirable to evaluate predictive applicability over time in a given lake, as opposed to one snapshot in time.

The purpose of this study was to compile a subset of predictive geochemical models from the literature and quantitatively evaluate the similarities and differences between the predicted and observed water quality in each pit lake. One of the most important considerations in evaluating model predictions is who sets the evaluation criteria, and what these criteria are (Nordstrom 2012). A logical starting point for pit-lake modeling evaluation criteria is how well pit-lake water quality has been predicted in the past. Although Eary and Schafer (2009) approached this subject, no quantitative framework was described. In this analysis, a quantitative framework is suggested and applied to the question: can

✉ Connor P. Newman
connor.newman1988@gmail.com

[1] Nevada Division of Environmental Protection, 901 S. Stewart St., Suite 4001, Carson City, NV 89701, USA

[2] Present Address: US Geological Survey, Denver Federal Center, Bldg. 53, MS-415, Denver, CO 80225, USA

pit-lake geochemistry be predicted with any useful consistency given a reasonable conceptual model and appropriate modeling procedures? The method applies the error variance parameter (Δ), a quantitative evaluator for coherently corroborating the predictive quality of geochemical models of pit lakes. This paper also presents a summary of Δ values for various major (e.g. Ca, SO_4) and minor (e.g. Al, As, Fe) constituents for a variety of models from a subset of studies in the literature.

Compilation of Pit-Lake Models

A wide range of pit-lake predictive models were used for this analysis, including models prepared for regulatory permitting purposes as well as fundamental scientific research. A variety of different ore deposit types and commodities were incorporated. Most of these modeling endeavors involved a temporal prediction component, allowing evaluation of predictions over time in a given pit lake. A summary of the models is provided in Table 1, and a short description of each follows. The different modeling approaches and assumptions are also discussed, referencing to how each study incorporates (or does not incorporate) important hydrologic, geochemical, limnologic, and ecological processes.

Most of the compiled studies are in arid regions (Table 1). While this may not accurately reflect the distribution of pit lakes worldwide, the compilation was limited to published and publicly available studies that included a predictive aspect of future geochemistry [i.e. studies that included detailed geochemical analysis but no predictive element such as Tucci and Gammons (2015) were excluded from the compilation]. The distribution of pit lakes assessed herein simply reflects the availability of published literature available on this topic. The compilation is limited by the available literature, but this does not necessarily limit the applicability of the conclusions described herein because the overall chemistry of the pit lakes spans a wide range, as discussed below. Also, although most studies were conducted in areas of high potential evapotranspiration (PET), there is considerable variability in the estimated annual precipitation (Table 1). Therefore, this compilation includes environments in which catchment hydrology may play an important role in pit-lake chemogenesis.

In terms of their chemical composition, few of the studies contain enough details to allow full metal or major-ion chemistry to be evaluated (i.e. by Ficklin or Piper diagrams). However, pH was reported in each study, and comparing the range of pH values indicates the general similarity of pit-lake chemistry included to well-known pit lakes such as those of the Iberian Pyrite Belt or the Berkeley pit lake. Of the nine pit lakes in the study, four have been acidic at some period in their history (Copper Flat, Getchell, Bärwalde, and Kepwari; Table 1). This distribution of acidic vs. circumneutral pit lakes seems to reasonably reflect the literature, given that there are regions dominated by both acidic (Sánchez-España et al. 2008) and circumneutral lakes (Shevenell et al. 1999; Ramteke et al. 2016).

Five pit lakes are included from the state of Nevada, USA, which is one of the most active regions in the US for open-pit metal mining (Kramer et al. 2018). These pit lakes are predominantly formed in mines extracting Carlin-type gold deposits, with the exception of the Dexter mine, which is hosted in an epithermal gold deposit (Castor et al. 2003). Pit lakes throughout much of Nevada display generally circumneutral pH and moderate contaminant concentrations (with some notable exceptions), primarily due to interaction of acid-forming minerals with carbonaceous rocks (Shevenell et al. 1999). Of the five pit lakes in Nevada, three currently exist: the Getchell mine has been dewatered and the Post/Betze pit lake does not yet exist. The Getchell mine previously hosted two pit lakes; the North (Tempel et al. 2000) and the Main (Davis et al. 2006). The composition of modeling and observed waters for Post/Betze used herein are from batch testing (Eary and Schafer 2009; Schafer et al. 2006), and therefore offered an opportunity to compare a simple geochemical model and laboratory testing methods. For the modeling of the existing Cove pit lake (Hanna et al. 2016; Itasca 2016), various models were created using different assumptions of influent groundwater composition, and with ferrihydrite suppressed. The purpose of the different simulations was to evaluate the impact of predicted mineral precipitation on final chemistry. For this analysis, the base-case model allowing ferrihydrite precipitation was used.

The Harvard pit lake is an alkaline lake in California's historical Mother Lode gold district. The lake has elevated As concentrations and the composition indicates partial control by the dissolution of efflorescent sulfate salts present on the pit highwalls (Savage et al. 2009). The Copper Flat pit lake (located in New Mexico, USA) also hosts prevalent evaporative salts and displays a wide composition range, from slightly acidic with elevated metal concentrations to circumneutral with relatively low metal concentrations (SRK 2017). Several variations of modeling are summarized in SRK (2017), each using different assumptions and procedures. The results of the "unreclaimed pit lake model" were used for this analysis because this set of model assumptions is most applicable to the previously observed water quality. Both the Harvard and Copper Flat pit lakes have been present since the late 1980s/early 1990s.

The Kepwari pit lake is located in western Australia, southeast of Perth. This pit lake is unique in that several iterations of modeling have been completed with periods of observations between (Oldham et al. 2009; Salmon et al. 2017). Therefore, this pit lake offers the rare opportunity to evaluate how multiple modeling endeavors compare to

Table 1 Summary of models included in compilation, including description of climatic attributes and modeling methods

Pit lake(s)	Deposit type	Location	Mean annual precipitation (mm)^{1}	Mean annual evapotranspiration (mm)^{1}	Observed pH range	Reference(s)	Time-step length	Modeling code(s) used	Thermodynamic database used	Kinetic constraints approach	Phases included in mineral equilibrium	Adsorption processes	Sulfide oxidation quantification approach^{2}	Groundwater flow approach^{3}	Biogeochemical processes	Hydrodynamic modeling approach	Catchment processes
Getchell North and Main	Carlin-type gold	Nevada, USA	368^{a}	1160^{a}	3.0–7.9	Tempel et al. (2000)	Seasonal (summer and winter)	EQ3/6	data0.cmp	Psuedo-kinetic (relative dissolution rates)	Quartz, muscovite, kaolinite, anorthite, microcline, annite, calcite, dolomite, pyrite, realgar, orpiment, arsenopyrite, scorodite	Not included	Reaction path	Analytical	None described	None described	Highwall runoff
						Davis et al. (2006)	Only steady-state presented	PHREEQC	minteq.dat	None described	$Fe(OH)_3$, calcite, $O_{2(g)}$, $CO_{2(g)}$	$Fe(OH)_3$ based on Dzombak and Morel (1990)	Modified Davis-Ritchie	Numerical (MODFLOW)	None described	CE-QUAL-W2	Highwall runoff
Post/Betze	Carlin-type gold	Nevada, USA	280^{a}	1210^{a}	8.9^{4}	Eary and Schafer (2009); Schafer et al. (2006)	Only steady-state presented	PHREEQC	Not stated	None described	Fe(OH)3, calcite, barite, willemite, hypothetical solid solution species (Ni-silicate, Ni-carbonate), substitution of Zn and Ni into calcite, $O_{2(g)}$, $CO_{2(g)}$	$Fe(OH)_3$ based on Dzombak and Morel (1990), with modifications	Davis-Ritchie	Numerical (MODFLOW)	None described	None described	Highwall runoff
Cove	Carlin-type gold	Nevada, USA	235^{a}	1360^{a}	7.5^{4}	Hanna et al. (2016); Itasca (2016)	Annual	PHREEQC	minteqv4.dat	None described	Gibbsite, barite, calcite, dolomite, gypsum, otavite, malachite, fluorite, $Fe(OH)_3$, rhodochrosite, cerrusite, strontianite, smithsonite, $O_{2(g)}$, $CO_{2(g)}$	$Fe(OH)_3$ based on Dzombak and Morel (1990)	Davis-Ritchie	Numerical (MODFLOW)	None described	None described	None described
Twin Creeks	Carlin-type gold	Nevada, USA	258^{a}	1250^{a}	7.4–8.2	Kempton et al. (1997)	Variable (1 year initially, increasing at an unspecified rate through model progress)	MINTEQA2	minteq.dat	None described	Gibbsite, calcite, avicennite, malachite, barite, otavite, cerrusite, anglesite, $NiCO_3$, $Cd(OH)_2$, $Cr(OH)_3$, $O_{2(g)}$, $CO_{2(g)}$	$Fe(OH)_3$ based on Dzombak and Morel (1990)	Davis-Ritchie	Not described	None described	CE-QUAL-W2	Highwall runoff

Table 1 (continued)

Pit lake(s)	Deposit type	Location	Mean annual precipitation (mm)[1]	Mean annual evapotranspiration (mm)[1]	Observed pH range	Reference(s)	Time-step length	Modeling code(s) used	Thermodynamic database used	Kinetic constraints approach	Phases included in mineral equilibrium	Adsorption processes	Sulfide oxidation quantification approach[2]	Groundwater flow approach[3]	Biogeochemical processes	Hydrodynamic modeling approach	Catchment processes
Dexter	Epithermal gold	Nevada, USA	334[a]	1200[a]	7.1–9.0	Newman (2014)	Monthly	EQ3/6; Visual MINTEQ	data0; thermo.vdb and type6.vdb	Psuedo-kinetic (relative dissolution rates)	Calcite, goethite, managanite, biotite, orpiment, $Fe(OH)_3$	Goethite based on Weng et al. (2001); $Fe(OH)_3$ based on Dzombak and Morel (1990); manganite based on Tonkin et al. (2004)	Reaction path	Analytical	None described	None described	None described
Harvard	Orogenic gold	California, USA	742[a]	1470[a]	7.6–8.4	Savage et al. (2009)	Seasonal	PHREEQC	Not stated	None described	Calcite, magnesite, siderite, pyrite, arsenian pyrite, gypsum, epsomite, $Fe(OH)_3$, $O_{2(g)}$, $CO_{2(g)}$	Not included	Reaction path	Analytical	None described	None described	Highwall runoff
Copper Flat	Porphyry copper-molybdenum	New Mexico, USA	317[a]	1730[a]	4.4–8.2	SRK (2017)	Variable (0.5–25 years)	PHREEQC	minteqv4.dat	None described	Alunite, albite, anhydrite, azurite, barite, bronchiatite, brucite, calcite, chrysotile, epsomite, $Fe(OH)_3$, fluorite, gypsum, gummite, kaolinite, mirabilite, otavite, rutherfordine, schoepite, sepiolite, SiO_2, $Fe(OH)_3$, Ag_2Se, Cr_2O_3, U_3O_8, UO_3, $UO_2(OH)_2$, $CaMoO_4$, $CaSeO_3{:}2H_2O$, $CdMoO_4$, $CuMoO_4$, Cu_2Se, $Mg_3(PO_4)_2$, $MnSeO_3$, $NiMoO_4$, $Ni(OH)_2$, $Ni_2(AsO_4)_2{:}8H_2O$, $PbMoO_4$, SbO_2, $ZnMoO_4$	$Fe(OH)_3$ based on Dzombak and Morel (1990)	Geometric approach to reactive surface areas (based on assumed fracture density and oxygen penetration)	Numerical (MODFLOW)	None described	None described	Highwall runoff

Table 1 (continued)

Pit lake(s)	Deposit type	Location	Mean annual precipitation (mm)[1]	Mean annual evapotranspiration (mm)[1]	Observed pH range	Reference(s)	Time-step length	Modeling code(s) used	Thermodynamic database used	Kinetic constraints approach	Phases included in mineral equilibrium	Adsorption processes	Sulfide oxidation quantification approach[2]	Groundwater flow approach[3]	Biogeochemical processes	Hydrodynamic modeling approach	Catchment processes
Kepwari	Sub-bituminous Coal	Western Australia, Australia	840[b]	1550[b]	3.8–4.5	Oldham et al. (2009)	Not stated	Custom algorithm (parameterized similar to PHREEQC)	WATEQ4F	Psuedo-kinetic approach applied, but no details described	Gibbsite, $Fe(OH)_3$, Fe-sulfide (not specifically named), $O_{2(g)}$, $CO_{2(g)}$	Sorption included, but no details described	None described	Analytical	Ecological constraints included based on algal growth and organic geochemical species	CAEDYM	None described
						Salmon et al. (2017)	Hourly	Custom code (parameterized similar to PHREEQC)	Custom (based on WATEQ4F)	Mechanistic (based on parameterization of reactions)	Gibbsite, $Fe(OH)_3$, $O_{2(g)}$, $CO_{2(g)}$	$Fe(OH)_3$, sorption constants based on various references	Mechanistic (based on parameterization of sulfide oxidation reactions)	Analytical	Ecological constraints included based on parameterization	DYRESM and CAEDYM	Highwall runoff, river inflow, river outflow
Bärwalde	Lignite Coal	Saxony, Germany	545[c]	386[d]	2.9–3.5	Werner et al. (2006)	Not stated	PHREEQC (as implemented in MODGLUE)	Not stated	None described	Metal hydroxides (not specifically named), $O_{2(g)}$, $CO_{2(g)}$	Ion exchange included, but no details described	None described	Numerical (PCGEOFIM, as implemented in MODGLUE)	Carbon and nutrient availability included	CE-QUAL-W2 (as implemented in MODGLUE)	None described
						Werner (2009)	Not stated	PHREEQC (as implemented in MODGLUE)	Not stated	Kinetics mentioned as included, but no description supplied	Phases mentioned as included, but no list supplied	Not stated	None described	Numerical (PCGEOFIM, as implemented in MODGLUE)	Carbon and nutrient availability included	CE-QUAL-W2 (as implemented in MODGLUE)	None described

Table 1 (continued)

[1] Estimates of climatic attributes were from the literature or based on online estimator tools. Sources of climatic parameters; a = *Climate Engine* (Huntington et al. 2017), b = Salmon et al. (2008), c = Meteoblue (2019), d = ESRI (2019). The estimates provided by Meteoblue (2019) and ESRI (2019) are not as accurate as those from *Climate Engine*, and should be interpreted accordingly

[2] A reaction path approach to sulfide oxidation modeling applies mineral dissolution rates relative to one another and incorporates a step-wise mineral dissolution path, for additional details see Tempel et al. (2000). The Davis–Ritchie approach is described in Davis and Ritchie (1986; 1987) and Davis et al. (1986); the modified Davis–Ritchie approach is described in Fennemore et al. (1998)

[3] Analytical methods for simulating pit-lake inflow include those discussed in Vandersluis et al. (1995) and Marinelli and Niccoli (2000), numerical methods include finite-element and finite-difference groundwater models

[4] Only a single pH measurement was available for compilation, so no range is presented

observations. For this quantitative analysis, predictions for the period 2007–2011 were extracted from Oldham et al. (2009) and compared to monitoring data for the same time period from Salmon et al. (2017). Modeling results of Salmon et al. (2017) were also compared to observations for the time period 2007–2011.

Lake Bärwalde is located in the Lausitz mining district of eastern Germany. Surrounded by many other pit lakes hosted in previous coal mines, the pit lake is hydrologically complex as it is influenced by the nearby river as well as localized groundwater outflow (Werner et al. 2006). The lake has had an acidic pH throughout its entire observed history, although acidity has been decreasing in more recent modeling and observations (Werner 2009).

The models for all pit lakes considered in this analysis display a wide range in complexity and mathematical methodology. All studies used a method to simulate geochemical equilibrium (Table 1), with the most commonly used modeling code being PHREEQC (Parkhurst and Appelo 2013). Other codes used included EQ3/6 (Wolery and Jarek 2003) and MINTEQA2 (Allison et al. 1991). A number of different thermodynamic databases were applied. It is important to note that these databases may have different data for given species (Nordstrom and Archer 2003), potentially causing discrepancies between studies. In addition to thermodynamic constraints, several studies indicated that kinetic constraints were applied. These constraints generally only applied to dissolution reactions (e.g. Newman 2014; Tempel et al. 2000), and other studies that mentioned kinetics generally described their application in vague terms. The only study that rigorously incorporated kinetic constraints on mineral dissolution and precipitation was Salmon et al. (2017).

The approach used in each study to simulate sulfide oxidation (commonly assumed to be the major source of solute-loading to mine-impacted waters; e.g. Kempton et al. 2009) varied substantially. Several studies used the Davis–Ritchie "shrinking core" approach (Davis and Ritchie 1986, 1987; Davis et al. 1986) wherein sulfide oxidation in the field is simulated through calibration to a geochemical proxy (commonly kinetic testing results). Other studies used reaction path or mechanistic approaches wherein sulfide (and other) minerals were allowed to dissolve based on their laboratory dissolution rates. An alternative to these mechanistic approaches is to assign a specific fracture density and depth to the weathered highwall, as was done by SRK (2017). While this method allows for a "reactive mass" to be calculated, it does not address the geochemical reactions occurring in the pit, as suggested by Kempton et al. (2009). This approach is also not site-specific (unless measurements have been made of the weathered depth of the highwall), and therefore introduces additional uncertainty. Finally, several studies instituted no methodology to simulate sulfide oxidation. These studies may therefore have neglected one of the most important processes governing the generation of solutes in mine-impacted waters (Kempton et al. 2009; Nordstrom 2011).

The area in which the compiled studies showed the most variability was in the application of mineral and gas equilibria (Table 1). Several studies included large lists of minerals that were allowed to precipitate from the pit lake(s) on predicted saturation, while other studies included only one or two minerals and gases. For example, Salmon et al. (2017) included only gibbsite and ferrihydrite ($Fe[OH]_3$) as mineral phases in their predictions. This contrasts with Tempel et al. (2000), Hanna et al. (2016), and SRK (2017) who respectively included 13, 15, and 40 phases in their modeling. It is important to note that a number of efflorescent salt phases have been observed near the Copper Flat pit lake, so many of the phases included in the SRK (2017) model may be quite reasonable. Although numerous phases included in the various studies in Table 1 have been hypothesized to be equilibrium constraints in pit lakes (Eary 1999; Eary and Castendyk 2013), direct evidence as to their abundance in most of these specific pit lakes is lacking. Nordstrom (2009) specifically notes that equilibrium assumptions may not always be appropriate in mine-impacted waters. Most of the studies applied gas equilibria for atmospheric O_2 and CO_2. Related to mineral equilibrium, many of the studies included sorption processes as a possible control on aqueous concentrations, typically by applying the diffuse double-layer model (DLM; Dzombak and Morel 1990). Finally, with respect to reactions and reaction rates, only a subset of the models included any quantification of biogeochemical reactions and their impact on water quality. Lack of inclusion of biogeochemical processes may impact model results because the reaction-rate changing effects of biota (Nordstrom 2011) would not be accounted for. Variability in the application of mineral equilibrium, sorption processes, and biogeochemical reactions may lead to substantial differences between the compiled models.

Nearly all of the investigations evaluated limnologic aspects in some manner, although not strictly quantitatively. Only those studies that quantitatively addressed hydrodynamics are noted with the applicable modeling code in Table 1. Others simulated the epilimnion and hypolimnion separately and then mixed the composition of each layer as appropriate (Newman 2014; Savage et al. 2009; Tempel et al. 2000). Hanna et al. (2016) and SRK (2017) qualitatively evaluated limnology of the existing pit lakes and concluded that the respective lakes likely experience complete turnover annually. These studies therefore assumed that the respective pit lakes could be simulated as a single layer lake.

Approaches for predicting the water balance of pit lakes include analytical methods (e.g. Marinelli and Niccoli 2000; Vandersluis et al. 1995) or numerical modeling codes (e.g. MODFLOW, FEFLOW). Analytical methods are typically more easily implemented but may not produce sufficiently detailed or accurate results (depending on the goals of the study). Contrastingly, numerical modeling approaches may be more accurate, but are data intensive. The models compiled for this study are nearly equally split between numerical and analytical groundwater flow modeling approaches. In addition to subsurface hydrogeologic processes, many of the studies included runoff from highwall surfaces in their modeling. Effects of solute transport from rainstorms could be important in some instances, e.g. the Copper Flat pit lake where secondary mineral salts are abundant (SRK 2017). Aside from highwall runoff, only Salmon et al. (2017) implemented other catchment-hydrology specific boundary conditions. This was necessary because the Kepwari pit lake experiences both inflow from and outflow to a river.

The temporal resolution of the various predictions (hydrologic, limnologic, geochemical) for each model also varied substantially (Table 1). This temporal frequency, referred to as the time-step length, ranged from hourly (Salmon et al. 2017) to 25 years (SRK 2017). Several studies, however, did not specifically describe temporal resolution during discussion of model construction. As will be discussed further in this paper, temporal resolution may have important implications for model performance.

A different subset of geochemical parameters was included in predictions for each model. Because of this, not all parameters have the same number of possible quantitative evaluations. All values for both simulated and observed geochemistry used in this study were extracted from the references in Table 1 or the International Network for Acid Prevention (INAP) pit lakes database (Johnson and Castendyk 2012). In some instances, multiple publications were available for a particular pit lake (e.g. Lake Bärwalde; Werner 2009; Werner et al. 2006). In these instances, the values used in quantitative analysis presented herein were extracted from the most complete and easily transferrable record (except in the case of the Kepwari pit lake, where both studies were used, as explained above).

This compilation is not meant to be exhaustive of all ore deposit types and mining regions around the world. Instead, it is meant to include a subsample of pit-lake studies prepared by academia and consulting scientists from five regions of the world and in five different mining environments. Although it would be useful to include examples from some of the world's most well-known and geochemically extreme pit lakes (e.g. Berkeley, Iberian Pyrite Belt, etc.), to the author's knowledge no publicly available sources of both temporal predictions and observations exist for these lakes.

In total, the studies considered in this review included predictions of 26 geochemical species (pH, alkalinity, Ag, Al, As, Ba, B, Cd, Ca, Cl, Cr, Cu, F, Fe, K, Pb, Mg, Mn, Na, Ni, Sb, Se, Si, SO_4, Tl, and Zn) totaling 642 possible comparisons between observed and predicted concentrations. Most of these species had only one or two potential comparisons (i.e. no temporal evaluation was possible). However, nine species were included with enough regularity to facilitate a statistical evaluation of their corroboration (pH, alkalinity, Al, As, Ca, Fe, Mg, Mn, and SO_4). These constituents are also some of those that are commonly important for geochemical modeling studies.

Quantitative Framework for Geochemical Model Assessment

Although a variety of consulting and academic studies have produced predictive geochemical models of existing pit lakes, currently there is no calibration measure in geochemical modeling that is used as ubiquitously as those employed in the evaluation of groundwater flow models, such as root mean squared error (RMSE) or mean absolute error (MAE; Anderson et al. 2015). While the mathematical formulation of these calibration measures does not preclude their application to the results of geochemical modeling, their practicality is limited primarily by the number of constituents included in geochemical models and because each of the previously mentioned calibration measures has a unique ideal usage. Salmon et al. (2017) apply RMSE and normalized MAE (NMAE) to their modeling results, but report them only for Al, Fe, and pH. Hanna et al. (2016) provide a qualitative error analysis, but no quantitation is included. Likewise, Kuipers et al. (2006) summarize a post-audit of models related to mining impacts (some of which included pit lakes), but, again, no quantitation was performed. More general applications of quantitative analyses are described by Bennett et al. (2013) and Alewell and Manderscheid (1998). While potentially useful, none of the approaches suggested by those authors have been widely applied to pit-lake modeling.

Aside from a lack of existing quantitative framework, the validation of geochemical models is inherently difficult, as discussed by Nordstrom (2012). Expression of a model's validation or accuracy (geochemical, hydrodynamic, etc.) is a misstatement of the applicability of the model. All models contain uncertainties that are both quantifiable and unquantifiable (i.e. conceptual uncertainties; Bredehoeft 2003, 2005). Stating that a model is accurate implies that the processes implemented by the mathematical code were consistent with observations and mechanistic processes, that no important processes were excluded, and that all of the processes simulated in the model actually occur in reality, and vice versa. It is often very difficult to determine a priori the applicability of these conceptual models. As described by Nordstrom (2012), better terminology would be to state that the models are either corroborated or uncorroborated for a given set of assumptions and observations.

Although there are issues involved in evaluating the accuracy of pit-lake geochemical modeling (Nordstrom 2012), a more quantitative evaluation of model predictions would possibly enhance understanding of geochemical processes occurring in pit lakes. The error variance parameter, first suggested here, to evaluate geochemical model predictions is Δ, described by Eq. (1):

$$\Delta = 10^{\left[\frac{(C_{Mod} - C_{Obs})}{C_{Obs}}\right]} \tag{1}$$

where $[C]_{Mod}$ represents the modeled dissolved concentration of constituent C and $[C]_{Obs}$ represents the observed dissolved concentration of constituent C. The quantity Δ is therefore 1 for an exact match between modeled and observed concentrations, less than 1 for simulated concentrations that are less than observed, and greater than 1 for simulated concentrations that are greater than observed. Because the units of quantities in Δ cancel, either mmol L^{-1} or mg L^{-1} are acceptable for usage, and Δ itself is unitless. The use of Δ to evaluate geochemical modeling results may allow increased quantitative comparison of simulations and observations throughout the literature if applied in more studies. One of the primary goals of this study was to establish a baseline of Δ values from the published literature, so that future investigators can improve on these baseline values.

Results and Discussion

Major Ions

Most of the dissolved solutes in observed pit-lake chemistry are the major ions Ca, Mg, HCO_3, and SO_4, which are likely predominantly derived from dissolution of minerals in the highwalls of the open pits (Castendyk et al. 2015a; Savage et al. 2009; Shevenell et al. 1999). Of these constituents, Ca and SO_4 are commonly important from a geochemical process standpoint as these are the components of gypsum ($CaSO_4 \cdot 2H_2O$), one of the most common secondary minerals associated with mine sites (Nordstrom 2011) and an often-applied solubility control in pit-lake models (Eary 1999).

Figure 1 is a plot of observed vs. predicted SO_4 (Fig. 1a) and Ca (Fig. 1b) concentrations for all sites in the compilation with applicable data. Most of the studies included in the analysis were able to predict both SO_4 and Ca concentrations reasonably well, with several notable exceptions. Predictions for the Copper Flat pit lake (SRK 2017) illustrate a propensity for under-prediction of SO_4 concentrations (Fig. 1a), potentially indicating that not all sulfide-oxidation reactions occurring in the pit lake over time were implemented in the model. Of all the compiled models, Copper Flat was the only to use a simple geometric approach to simulate sulfide-linked solute release (Table 1). Had a more site-specific method been used, it is possible that SO_4 concentrations would have shown better corroboration. Elevated SO_4 concentrations, up to ≈ 7000 mg L^{-1}, could also be related to the dissolution of acid-storing efflorescent sulfate salts near the pit lake during rain events [as noted by SRK 2017 to be present near the pit lake]. Although the modeling approach incorporated highwall runoff, the method used may not have had the correct temporal resolution to simulate rapid geochemical responses to dissolution of efflorescent salts. A similar pattern of under-prediction of SO_4 is observed in the simulations of the Getchell Main pit lake (Davis et al. 2006), although to a lesser extent.

When comparing Ca concentrations, all models showed good reproducibility, aside from Tempel et al. (2000) for the Getchell North pit lake, where Ca was substantially under-predicted. (Itasca also displays minor under-prediction in the Cove pit lake and is discussed separately.) The likely cause of this discrepancy in Tempel et al. (2000) is the exclusion of gypsum from the dissolved minerals in the modeling procedure (Table 1). Observed water quality indicates that gypsum was close to saturation in the North pit lake at Getchell. The contrasting degrees of corroboration of SO_4 and Ca in the results of Tempel et al. (2000) illustrates that although both constituents are key components in one of the most common secondary minerals associated with mine sites (gypsum), corroboration of these species is not necessarily completely linked. In this instance, Tempel et al. (2000) included a likely mineralogic source for both species (pyrite, arsenopyrite, realgar, orpiment for SO_4 and calcite for Ca), but the sequential reaction steps used did not produce reasonable reconstructions for both constituents. Had gypsum been included, corroboration of Ca concentrations may have been achieved. The potential influence of gypsum solubility on the error apparent in SO_4 concentrations in Fig. 1a for SRK (2017) cannot be evaluated without further information, as Ca concentrations were not presented in a temporal format conducive to analysis in this study. SRK (2017) did note the presence of a number of efflorescent sulfate salts near the pit lake and these may have had an important impact on pit-lake water quality in terms of both Ca and SO_4 (e.g. Bowell and Parshley 2005).

Calculation of Δ for SO_4, Ca, and pH over time for the Cove pit lake (Hanna et al. 2016; Itasca 2016) and Copper Flat pit lake (SRK 2017) was used to evaluate how model corroboration varies temporally. As shown in Fig. 2, modeling of the Cove pit lake initially overestimated the SO_4 (Fig. 2a) and Ca (Fig. 2b) concentrations, followed by consistently under-predicting the Ca concentration for the remainder of the modeled time period (although to a lesser quantitative degree). These trends likely have different causes. Modeling of the Cove pit lake used the Davis–Ritchie approach to simulate sulfide oxidation (Table 1; Davis and Ritchie 1986, 1987; Davis et al. 1986).

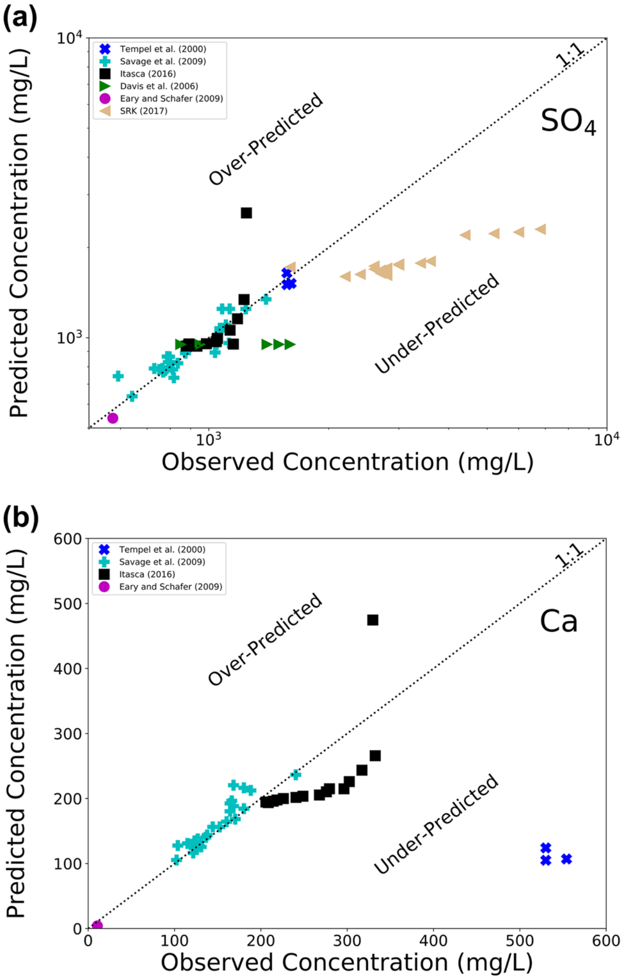

Fig. 1 Comparison of predicted and observed concentrations of SO_4 (**a**) and Ca (**b**)

The initial over-prediction could indicate that the sulfide-oxidation modeling over-estimated the mass of solutes (specifically SO_4) liberated by sulfide oxidation, which were then "flushed" in the simulation to the pit lake during initial infilling. However, this is somewhat speculative as the modeling overall was highly parameterized and thus sensitive to a variety of inputs. The Δ values less than 1 throughout the remainder of the simulation for Ca may be linked to several processes: (1) the simulated water balance of the model, (2) simulated groundwater solute concentrations, or (3) the inclusion of calcite as an equilibrium phase. The latter option is likely not the major cause of the discrepancy because sampling of precipitates from the water column indicated the presence of calcite, and calcite was predicted to precipitate by the geochemical model (Itasca 2016). Using slightly greater influent Ca concentrations in groundwater (which adds mass to the pit lake) resulted in slightly better matches with observations, but did not explain the entirety of the discrepancy (results not shown here). The model may

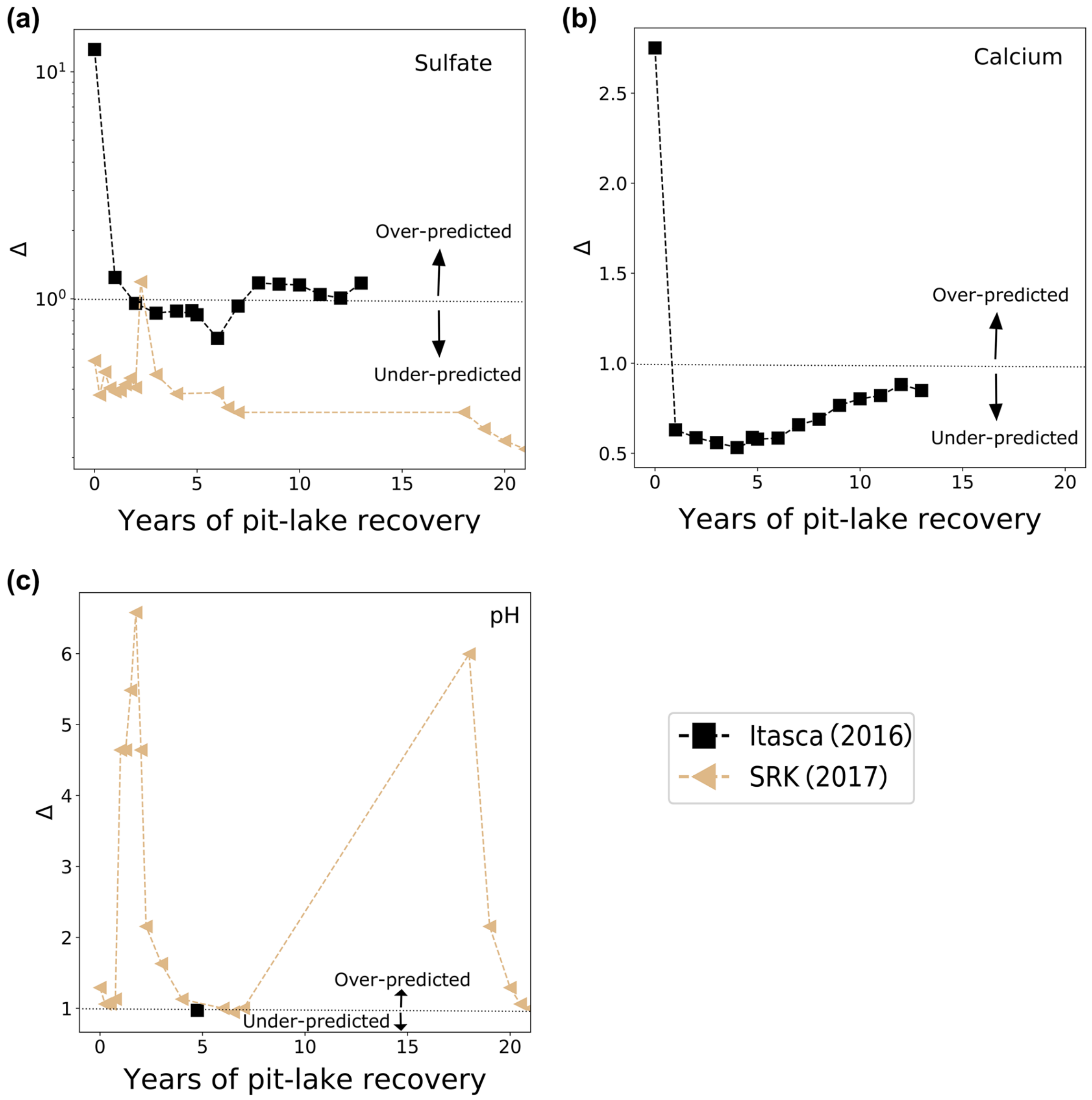

Fig. 2 Calculated Δ values through time for SO_4 (**a**), Ca (**b**), and pH (**c**) in the Cove pit lake (Itasca 2016) and Copper Flat pit lake (SRK 2017). Note logarithmic scale for panel (**a**). The dotted line denotes $\Delta = 1$

therefore have used an inaccurate water balance (calculated in this instance using MODFLOW) and diluted Ca more than actually occurs. The discrepancies are likely not linked to gypsum solubility, as the mineral is undersaturated in all simulations and observations in the pit lake (based on equilibrium calculations).

Evaluation of corroboration of SO_4 concentrations over time in the Copper Flat pit lake (SRK 2017) display a different behavior than the Cove pit lake, in that SO_4 concentrations in Copper Flat are under-predicted throughout nearly the entire prediction interval. A similar pattern is reflected in pH (Fig. 2c), with pH being consistently over-predicted (which represents an under-prediction of acidity). Prediction of these geochemical components is intimately related because of the relationship of acidity (and thus pH) with sulfide minerals (and thus SO_4). Concurrent under-prediction of acidity and SO_4 indicates that some acid-producing reaction that is occurring in the pit lake has been excluded by the model.

Also evident in the evaluation of Δ through time for the Copper Flat pit lake is that corroboration varies widely, with modeled pH corroborating well with observations in some, but not all, time periods. The variations in corroboration may be related to the temporal discretization of the model. As summarized in Table 1, modeling used variable time-step lengths from 0.5 to 25 years. These may not have incorporated fine enough temporal detail in the water balance and mineral equilibrium reactions to be able to sufficiently approximate the pit-lake system.

Simulated and observed pH are illustrated in Fig. 3, which includes more studies than the comparison for Ca and SO_4 because nearly every pit-lake model includes pH as one of the parameters of interest. As evident from Fig. 3, many pit-lake studies were able to replicate pH reasonably well across a full range of pH values, ranging from acidic to circumneutral. Notably, however, pH was substantially over-predicted in several studies but was not substantially under-predicted in any study. Three linear, roughly horizontal clusters of comparisons (Fig. 3) corresponding to Getchell Main pit lake (Davis et al. 2006), Copper Flat pit lake (SRK 2017), and Twin Creeks pit lake (Kempton et al. 1997), indicate that pH was generally over-predicted (i.e. the hydrogen acidity of the lake was under-predicted), although the difference was relatively minor for Kempton et al. (1997). For Davis et al. (2006) and SRK (2017) the linear nature of these features is likely due to the temporal resolution of the simulations. Davis et al. (2006) only present the long-term pseudo-steady state predicted chemistry of the pit lake at the Getchell mine; therefore, the temporal variation of constituent concentrations is not accounted for. The original modeling presented by Davis et al. (2006) was summarized in consulting reports, but these reports were not available for this analysis. It is possible that these consulting reports contained simulations with a finer temporal resolution that would result in closer corroboration. For SRK (2017), the model uses a variable timestep approach, wherein the composition was simulated at finer resolution (0.5–3 years) during early filling, and coarser resolution (25 years) during later filling. It is clear that pH corroborates well in both SRK (2017) and Davis et al. (2006) at some time intervals, but not at others. This illustrates that the time-step length is important in geochemical modeling of pit lakes. Lack of corroboration for pH in SRK (2017) could also be caused by sulfide oxidation processes or efflorescent sulfate salt dissolution not being captured, as discussed above with reference to SO_4 and Ca.

It is interesting to note the relatively close approximation of Oldham et al. (2009) and Salmon et al. (2017), as these authors did not calibrate to observed conditions but instead conducted highly parameterized modeling that included hydrodynamics, ecology, and geochemistry. Both Oldham et al. (2009) and Salmon et al. (2017) tested various input scenarios for mineral equilibrium (either allowing gibbsite/ferrihydrite precipitation or not), which resulted in a variety of models. All of these are shown in Fig. 3, which creates the patterns visible in the apparent linear features for Oldham et al. (2009) and Salmon et al. (2017). Also noteworthy when discussing the correspondence of predictions for the Kepwari pit lake is that both modeling endeavors used very simple mineralogic constraints (Table 1). It is apparent that, for pH, these simple constraints allowed for close corroboration.

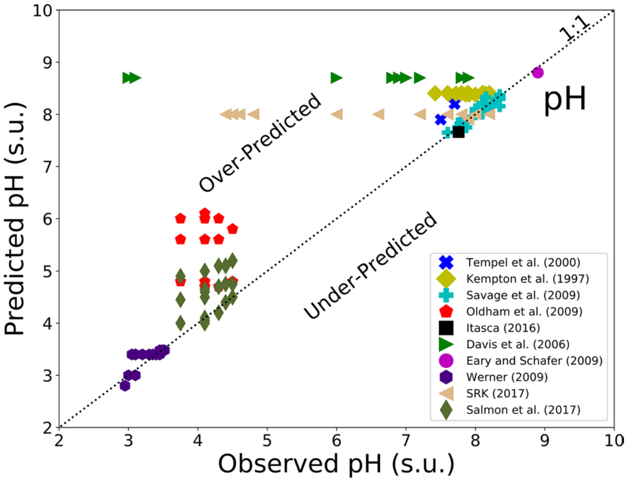

Fig. 3 Comparison of predicted and observed pH

Trace Metals and Metalloids

Of primary interest in many pit-lake modeling studies is the behavior of trace elements such as Al, As, and Fe. For Al and As, these elements may be toxic to humans or other life forms. While Fe is less toxic, it is often of primary concern because of the geochemical controls Fe has on many other trace elements through co-precipitation and sorption. Figure 4 illustrates the degree of corroboration for Al, As, and Fe for all studies with applicable simulations and observations.

Although Al was included in only a subset of models, it appears to have been simulated with reasonable reproducibility in all cases but Tempel et al. (2000), where Al was underestimated by several orders of magnitude. The likely cause of this is the slow dissolution rates of the aluminosilicate minerals included in the simulations (which were the only source of Al; Table 1). Simulating another Al source mineral may have resulted in better corroboration.

The comparison of As concentrations (Fig. 4b) includes more studies (likely due to the toxicity of As and focused study of this constituent in mining environments). This comparison indicates that As concentrations are generally well corroborated, although some spread exists. The same linear feature is evident in the predictions of Davis et al. (2006) as was previously discussed in reference to pH and SO_4. The degree of corroboration of As predictions from a number of studies may be somewhat surprising when observing the poorer corroboration for Fe (Fig. 4c), given the high-degree of correlation assumed by many authors for these constituents due to As-sorption on Fe-oxyhydroxides (e.g. Davis et al. 2006; Kempton et al. 1997; Tempel et al. 2000). Additional discussion is provided below on this topic.

Contrary to the corroboration of As predictions, Fe is not well corroborated except in studies by Newman (2014), Salmon et al. (2017), and in one event by Tempel et al. (2000). Concentrations of Fe were consistently under-predicted in the other studies. One likely cause for the discrepancy is the equilibrium manner in which mineral precipitation is often treated. As noted by Eary (1999), observed Fe concentrations are commonly greater than expected when compared to ferrihydrite solubility in circumneutral pit lakes, although ferrihydrite appears to be a good approximation for the upper limit of Fe solubility in acidic pit lakes. All of the studies included in Fig. 4c included Fe-oxyhydroxide in equilibrium calculations. Newman (2014) was able to calibrate the model of the Dexter pit lake to observed conditions for Fe by allowing mineral precipitation, then re-dissolving a portion of the precipitated mineral mass to reproduce observations (except in two events where corroboration remained poor, as illustrated in Fig. 4c). Salmon et al. (2017) allowed ferrihydrite to precipitate in one model, and it appears that the coupled nature of the modeling allowed for reasonable masses of the mineral to precipitate, as opposed to removing nearly all Fe mass from solution as in the other studies. The under-prediction of Fe by the other studies indicates that these models may have overestimated the amount of Fe precipitated from solution.

It is important to note that the comparison for some trace elements likely suffers due to the colloidal nature of many precipitates in the field. Nordstrom (2011) describes evidence from equilibrium calculations and electropotential measurements that indicates a substantial proportion of precipitated Fe-hydroxides may pass through field-filtration units. It is also possible that fine colloidal precipitates of Al and Si could pass through filters in much the same manner. If this were the case, trace-element mass that had precipitated would be assigned as dissolved in the observations, and any computations comparing dissolved concentrations predicted by geochemical models would be biased. The degree to which this process could impact the comparisons herein is not known, because the requisite data was not typically provided in each study. For a more robust comparison to be made, the sampling methods of each study would need to be fully explained including if filters were used in the field and their pore size (if applicable). Additionally, full water-quality analyses would need to be provided such that saturation indices could be calculated for each water sample. These saturation indices would then need to be compared to corresponding model outputs to assess discrepancies. Very few of the studies in this compilation included the requisite data to allow a

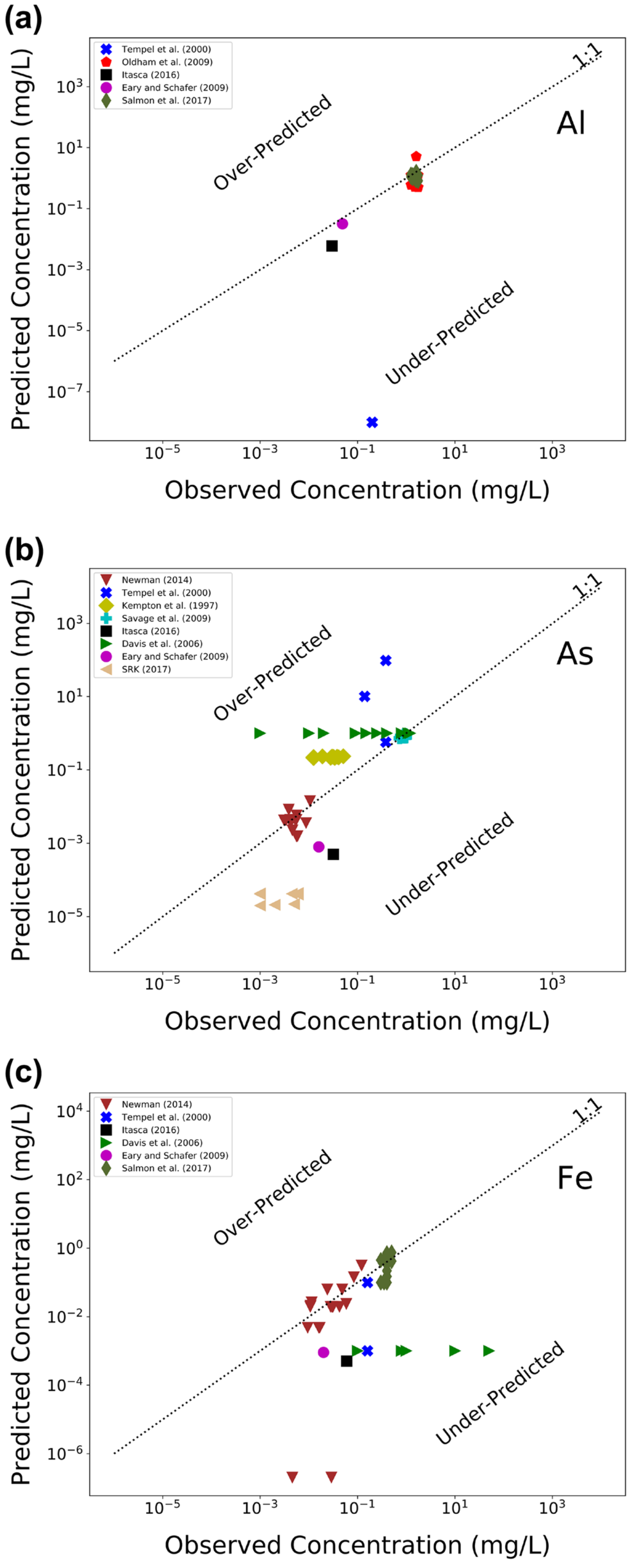

Fig. 4 Comparison of predicted and observed Al (**a**), As (**b**), and Fe (**c**)

comparison to take place; therefore, the impact of filtration (or lack thereof) on the comparisons is not known.

An excellent example of the temporal variation of Δ values for Al and As is illustrated in Fig. 5. Comparison of Al predictions and observations for various different model setups in Fig. 5a for Salmon et al. (2017) show that the degree of corroboration for metals with hypothesized equilibrium controls are sensitive to the mineral assemblages included in modeling. The Al predictions with no mineral equilibrium and ferrihydrite-only equilibrium did not differ from one another and showed $\Delta < 1$ (under-prediction). When gibbsite was included as an equilibrium phase, however, predictions corroborated much better. In this instance, the gibbsite was simulated as dissolving. When completing mineral equilibrium calculations such as this, it is important to note that thermodynamic data may vary between databases, and that the specific phase included could impact solubility (i.e. amorphous $Al(OH)_3$ vs. microcrystalline gibbsite; Nordstrom 2011). The comparison for As through time for the Harvard pit lake (Fig. 5b; Savage et al. 2009) indicates the degree to which limnology plays a role in controlling concentrations. Epilimnion As concentrations corroborated well, while As concentrations in the hypolimnion were more commonly under-predicted ($\Delta < 1$). These results indicate that an As source may have been under-represented in the hypolimnion.

Summary of Predictive Corroboration

As described previously, nine constituents included enough potential comparisons to allow statistical analysis of their corroboration (Fig. 6). For this analysis, all Δ values greater than 10,000 (i.e. > 4 orders of magnitude of over-prediction) were removed to facilitate comparison. This filtering resulted in 29 of the 642 total Δ values not being displayed in Fig. 6, most of which were As (24 of the 29 filtered values).

Figure 6 shows that pH, As, SO_4, Ca, Mg, and alkalinity display median Δ values close to 1, meaning that these constituents generally corroborated well. It is clear that nearly a quarter of the As comparisons were filtered out for being grossly over-predicted (24 of 101 total As comparisons), and thus this analysis suffers for As. The distribution of Δ values around the median (represented by the 25th and 75th percentiles at the edges of the boxes) illustrates the propensity for these constituents to be under- or over-predicted. Alkalinity, Ca, and Mg show a nearly equal distribution, indicating predictions did not have a tendency to systematically under- or over-represent observations. Contrastingly, pH has a slightly upper-skewed distribution, indicating that pH was more over-predicted than under-predicted, as noted above and displayed in Fig. 3. This result suggests that on the whole, acidic pit lakes may be more difficult to predict a priori. The opposite was true for As and SO_4, with slightly downward-skewed distributions.

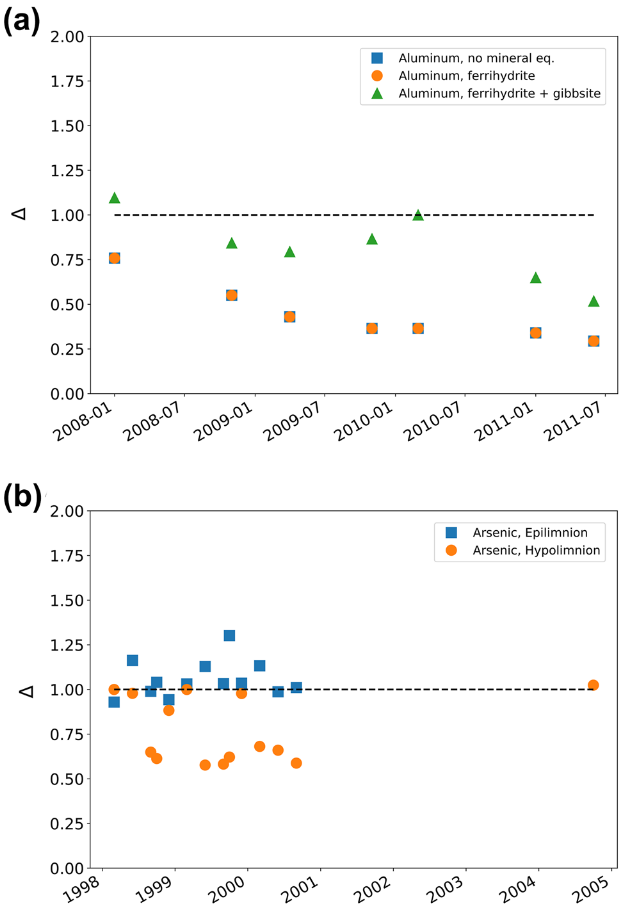

Fig. 5 Calculated Δ values through time for Al (**a**) in the Kepwari pit lake (Salmon et al. 2017) and As (**b**) in the Harvard pit lake (Savage et al. 2009). The dotted horizontal line denotes $\Delta = 1$

In contrast to the previous group of species that were well-corroborated, analysis of Δ values in Fig. 6 for Al, Fe, and Mn indicate that all had median Δ values less than 1, representing under-prediction. This is somewhat surprising considering the visual comparison shown on Fig. 4a for Al. This seeming contradiction illustrates the benefit of the quantitative comparison using Δ as opposed to a simple qualitative comparison of observed and simulated values (especially given that visual comparisons may be difficult due to the logarithmic scale required for most constituents). As discussed previously, Fe was commonly underestimated by the modeling considered in this review, likely due to overly simple equilibrium constraints being applied or due to filtration effects. The reason Mn was under-predicted may be similar to that for Fe, as some studies included Mn-oxyhydroxides in modeling (Eary and Schafer 2009; Newman 2014). However, Mn shows a wider spread in Δ values, and poor corroboration for this constituent may be due to other

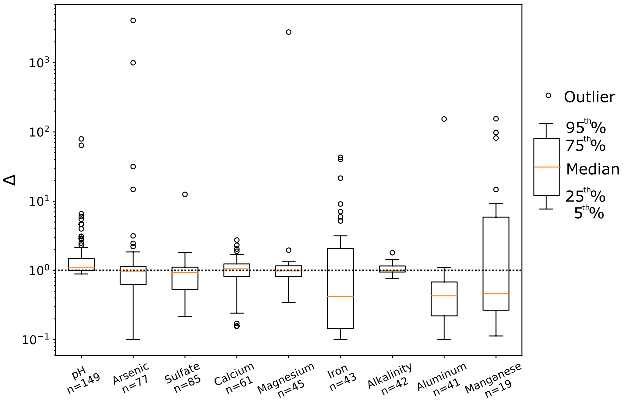

Fig. 6 Boxplot of calculated Δ values for all geochemical species with greater than 10 possible comparisons (number of applicable comparisons shown below parameter name). Δ values greater than 10,000 were removed from the analysis. The dotted horizontal line denotes Δ = 1

aspects of the modeling, such as water balance or sorption. Both Fe and Mn may have been better corroborated had investigators included benthic fluxes in models, as this process is known to impact concentrations of these metals in variable redox environments (e.g. Pakhomova et al. 2007), but benthic fluxes were consistently not included. Inclusion of this process may require highly advanced modeling codes, however, and may be difficult to benchmark without observations from existing pit lakes, which is generally lacking in the literature. Additionally, Al, As, Fe, and Mn all show the lowest range of 5th percentile Δ values, indicating that a statistically relevant portion of the predictions of these species greatly under-predicted observed concentrations.

Aside from the comparison of specific parameters shown in Fig. 6, several other commonalities of well-corroborated models may be observed. First, models that used finer temporal resolution (time steps) tended to result in better corroboration. For example, temporal predictions of Al and As (Fig. 5) by Salmon et al. (2017) and Savage et al. (2009) show better corroboration than predictions of SO_4 and Ca (Fig. 2) by Itasca (2016) and SRK (2017), despite the fact that Al and As were statistically more difficult to predict closely (Fig. 6). One of the differences driving the variable degrees of corroboration may be the temporal frequency of the model. Savage et al. (2009) and Salmon et al. (2017) each used sub-annual time steps related to the physical attributes of the site, whereas Itasca (2016) used annual time steps and SRK (2017) used a variable time-step ranging from 0.5 to 25 years. Temporal resolution has previously been described as likely impacting the predictive abilities of the Copper Flat (SRK 2017) and Getchell (Davis et al. 2006) pit lake studies, which did not include time steps at a frequency able to capture sub-annual geochemical changes in the pit lake.

As discussed above, various mineral equilibrium assumptions may also result in poor corroboration. Again comparing Itasca (2016) to Salmon et al. (2017), the former study applied 15 phases (13 minerals, 2 gases), although not all were predicted to precipitate, whereas the latter applied only four phases (2 gases, 2 minerals). Although Itasca (2016) showed generally reasonable corroboration for Al, As, and Fe, these elements were not as closely corroborated as they were in the results of Salmon et al. (2017; see Fig. 4 for comparisons). It is also interesting to note that although several studies noted that kinetic constraints were applied, no details were generally included as to the effect these kinetic constraints had on results. One instance in which kinetics were specifically discussed was Salmon et al. (2017), wherein the authors suggested that future research could incorporate kinetic constraints for the conversion of ferrihydrite to goethite.

One aspect which has not yet been discussed herein is the intended purpose of each of the modeling studies, which generally fall into two categories: academically driven (Newman 2014; Oldham et al. 2009; Salmon et al. 2017; Savage et al. 2009; Tempel et al. 2000; Werner 2009; Werner et al. 2006) or prepared for regulatory purposes (Davis et al. 2006; Eary and Schafer 2009; Hanna et al. 2016; Itasca 2016; Kempton et al. 1997; Schafer et al. 2006; SRK 2017). Although there are no strict differences, academically driven studies tended to be more complex in their numerical construction (i.e. incorporating biogeochemical or limnologic processes) and use finer temporal resolution. Consulting studies prepared for regulatory purposes contrastingly tended to use coarser temporal resolution (with simulations running for longer durations) and commonly included a greater number of equilibrium phases. One can assume that the reason for these differences is the needs of the end user. Regulatory bodies are commonly concerned with long-duration predictions of water quality (i.e. on the order of hundreds of years) while academic studies may be more focused on identifying the relevant biogeochemical processes in a mechanistic manner. Both approaches may lend useful conclusions as to reasonable methods for simulating pit-lake geochemistry.

Conclusions

A quantitative framework, the error variance parameter (Δ), has been used in this paper to evaluate the results of predictive geochemical modeling for pit lakes. Although simple, Δ allows for more robust analysis of model discrepancies, which may lead to a further understanding of the causative factors for these discrepancies. The application of Δ as opposed to the RMSE or NMSE has the advantage of retaining both the directionality (i.e. under- or over-prediction) as well as the magnitude of the difference. The most important aspect of this research, however, is not the suggestion of a simple error variance parameter, but observing the wide variability in how geochemical models for pit lakes are conceptualized and constructed, which processes may be important to include or exclude, and how the incorporation of specific processes leads to predictions that more closely approximate observations. Overall, this compilation and quantitative evaluation of pit-lake studies across three continents and five styles of mineralization indicates that the water quality of future pit lakes may be predicted with a relatively high degree of corroboration. It is important to note, however, that this corroboration does not indicate true accuracy, as it is conceivable that a different conceptual model and input parameters could result in equally corroborated predictions. The analysis also indicates that models that tended to have a high degree of corroboration were either highly parameterized and complex (e.g. Oldham et al. 2009; Salmon et al. 2017) or had data available for calibration (e.g. Hanna et al. 2016; Newman 2014). One model included a probabilistic approach to simulations (Kempton et al. 1997), and this model showed good corroboration. Kempton et al. (1997) also had no data with which to calibrate, indicating that indeed pit-lake geochemistry can be simulated a priori with a reasonable degree of agreement with future observations.

It is important to note that except for some predicted events in Davis et al. (2006) and SRK (2017) (Fig. 3), this study did not assess one of the most consequential failings of a pit-lake predictive model; when that model fails to predict the genesis of an acidic and contaminant-laden pit lake and instead predicts a circumneutral and low-contaminant water body. These instances can result in severe consequences for regulatory bodies and mining companies alike. An example of such a scenario is the Lone Tree pit lake in central Nevada (Croall et al. 2017), where some of the predictive modeling conducted prior to pit-lake filling failed to predict the acidic (pH ≈ 3) water chemistry. The Lone Tree pit lake model evolved over time, however, resulting in different conclusions as to the predicted geochemistry. The failings of models such as the Lone Tree models are worth a detailed study in their own right.

Although the pit lakes compiled herein represent only a small fraction of existing pit lakes in the world today, they are believed to be a representative sample based on their geographic locations, ore-deposit types, climatic attributes, and range of observed pH [which encompasses the range of pH expected in mine pit lakes worldwide, as summarized by Eary and Castendyk 2013]. For example, although the compilation included only one pit lake from the German lignite region, the distribution of pH conditions from all pit lakes in this compilation is similar to that observed for the lignite region as a whole (Schultze et al. 2010). Therefore, it is suggested that although these conclusions rely on site-specific information, pit-lake modelers in nearly all environments may gain insight on processes that should be included in modeling, and those that may not be critical to incorporate.

The results of this study can be summarized as a list of suggestions for those conducting pit-lake studies:

1. Equilibrium mineralogy controls may be critical to include (Eary and Castendyk 2013), but it is important to consider which thermodynamic database is used as some databases contain critical inconsistencies for key species such as As (Helz and Tossell 2008; Nordstrom and Archer 2003). Also the specific mineral included must be considered for the overall applicability to the geochemical conditions (pH, redox, etc.) in the system of interest (Eary 1999; Nordstrom 2009). Finally, it is possible that instituting kinetic constraints on mineral precipitation and dissolution would result in more favorable corroboration in some instances.
2. Although sorption of metals and metalloids may be thermodynamically predicted (Nordstrom 2011 and references therein), some models including sorption compare particularly poorly with observations. Also, co-precipitation may be an equally important process (e.g. Schafer et al. 2006). More field-based study of sorption and co-precipitation processes is required in pit lakes to advance modeling capabilities.
3. Implementation of this quantitative evaluation framework may lead to greater understanding of geochemical processes operating in pit lakes, given that the quantitative methodology allows for specific failings of the models to be recognized and addressed.
4. Model complexity should be critically assessed during the initial stages of a project. Several instances in this analysis indicated that highly complex and parameterized models produced good corroboration. However, those sites generally had abundant data available for inclusion (such as climatic attributes for hydrodynamic modeling). At more remote mine sites, these same data may not be available. It is important for model complexity to be balanced with data availability and worth. Initial modeling may be able to be completed with a less complex model, which can then drive future data collection to allow more complex processes to be simulated as warranted.

Acknowledgements Early portions of this work were completed when the author was in graduate school at the University of Nevada, Reno, where he benefitted from the expertise of Drs. Gina Tempel, Ron Breitmeyer, and Lisa Stillings. The help of Drs. Kaye Savage and Ursula Salmon in supplying modeling results are appreciated. Useful discussions were also had with Houston Kempton and Dr. Glenn Miller. Tom Gray, Dr. Alexandra Lutz, and two anonymous reviewers provided helpful comments on the first draft of this manuscript.

References

Alewell C, Manderscheid B (1998) Use of objective criteria for the assessment of biogeochemical ecosystem models. Ecol Model 107:213–224. https://doi.org/10.1016/S0304-3800(97)00218-4

Allison JD, Brown DS, Novo-Gradac KJ (1991) MINTEQA2/PRODEFA2, A geochemical assessment model for environmental systems. Version 3.0 User's Manual, US Environmental Protection Agency (EPA/600/3-91/021)

Anderson MP, Woessner WW, Hunt RJ (2015) Applied groundwater modeling: simulation of flow and advective transport. Academic Press, Cambridge

Bennett ND, Croke BFW, Guariso G, Guillaume JHA, Hamilton SH, Jakeman AJ, Marsili-Libelli S, Newham LTH, Norton JP, Perrin C, Pierce SA, Robson B, Seppelt R, Voinov AA, Fath BD, Andreassian V (2013) Characterising performance of environmental models. Environ Modell Softw 40:1–20. https://doi.org/10.1016/j.envsoft.2012.09.011

Bowell RJ, Parshley JV (2005) Control of pit lake water chemistry by secondary minerals, Summer Camp Pit, Nevada. Chem Geol 215:373–385

Bredehoeft J (2003) From models to performance assessment: the conceptualization problem. Groundwater 41(5):571–577. https://doi.org/10.1111/j.1745-6584.2003.tb02395.x

Bredehoeft J (2005) The conceptualization problem—surprise. Hydrogeol J 13(1):37–46. https://doi.org/10.1007/s10040-004-0430-5

Castendyk DN, Eary LE, Balistrieri LS (2015a) Modeling and management of pit lake water chemistry 1: theory. Appl Geochem 57:267–288. https://doi.org/10.1016/j.apgeochem.2014.09.004

Castendyk DN, Balistrieri LS, Gammons C, Tucci D (2015b) Modeling and management of pit lake water chemistry 2: case studies. Appl Geochem 57:289–307. https://doi.org/10.1016/j.apgeochem.2014.09.003

Castor SB, Boden DR, Henry CD, Cline JS, Hofstra AH, McIntosh WC, Tosdal RM, Wooden JP (2003) The tuscarora Au–Ag district: eocene volcanic-hosted epithermal deposits in the carlin gold region, Nevada. Econ Geol 98:339–366

Croall J, White P, Coleman B, Schmidt J, Skidmore S (2017) Novel approach to acid pit lake treatment and management. In: Wolkersdorfer C, Sartz L, Sillanpää M, Häkkinen A (eds) Proc. Conf of the International Mine Water Association, Lappeenranta, pp 495–501

Davis GB, Ritchie AIM (1986) A model of oxidation in pyritic mine wastes: part 1, equations and approximate solution. Appl Math Model 10(5):314–322. https://doi.org/10.1016/0307-904X(86)90090-9

Davis GB, Ritchie AIM (1987) A model of oxidation in pyritic mine wastes: part 3, import of particle-size distribution. Appl Math Model 11(6):417–422

Davis GB, Doherty G, Ritchie AIM (1986) A model of oxidation in pyritic mine wastes: part 2, comparison of numerical and approximate solutions. Appl Math Model 10(5):323–329. https://doi.org/10.1016/0307-904X(86)90091-0

Davis A, Bellehumeur T, Hunter P, Hanna B, Fennemore GG, Moomaw C, Schoen S (2006) The nexus between groundwater modeling, pit lake chemogenesis and ecological risk from arsenic in the Getchell Main Pit, Nevada, USA. Chem Geol 228:175–196. https://doi.org/10.1016/j.chemgeo.2005.11.023

Dzombak D, Morel F (1990) Surface complexation modeling: hydrous ferric oxide. Wiley, New York

Eary LE (1999) Geochemical and equilibrium trends in mine pit lakes. Appl Geochem 14(8):963–987. https://doi.org/10.1016/S0883-2927(99)00049-9

Eary LE, Castendyk DN (2013) Hardrock metal mine pit lakes: occurrence and geochemical characteristics. In: Geller W, Schultze M, Kleinmann B, Wolkersdorfer C (eds) Ch 3.2.2, Acidic Pit Lakes: the legacy of coal and metal surface mines. Springer, Berlin

Eary LE, Schafer WM (2009) Approaches for evaluating the predictive reliability of pit lake numerical models. In: Castendyk DN, Eary LE (eds) Mine pit lakes: characteristics, predictive modeling, and sustainability. Society for Mining, Metallurgy and Exploration, Littleton, pp 167–178

ESRI (2019) Evapotranspiration web viewer. https://www.arcgis.com/apps/OnePane/main/index.html?appid=b1a0c03f04994a36b93271b0c39e6c0f. Accessed 11 Feb 2019

Fennemore GG, Neller WC, Davis A (1998) Modeling pyrite oxidation in arid environments. Environ Sci Technol 32(18):2680–2687

Hanna BT, Newman CP, Sterrett RJ (2016) A calibrated geochemical model of leaching and attenuation reactions in a mine pit lake. In: Proceedings of the 20th international conference on tailings and mine waste, Keystone, CO, pp 711–722

Helz GR, Tossell JA (2008) Thermodynamic model for arsenic speciation in sulfidic waters: a novel use of ab initio computations. Geochim Cosmochim Ac 72:4457–4468

Huntington J, Hegewisch K, Daudert B, Morton C, Abatzoglou J, McEvoy D, Erickson T (2017) Climate engine: cloud computing of climate and remote sensing data for advanced natural resource monitoring and process understanding. B Am Meteorol Soc. https://doi.org/10.1175/BAMS-D-15-00324.1

Itasca (2016) Cove Pit-Lake 2016 water-quality prediction update, McCoy Cove Mine, WPCP#NEV0088009. Report to Premier Gold Mines Limited, Itasca Denver (unpublished)

Johnson EP, Castendyk DN (2012) The INAP pit lakes database: a novel tool for the evaluation of predicted pit lake water quality, In: Price WA, Hogan C, Tremblay G (eds) Proceedings of 9th international conference on acid rock drainage, Ottawa, Canada

Kempton JH, Locke W, Atkins D, Nicholson AD, Bennett M, Bliss L, Maley P (1997) Probabilistic prediction of water quality in the Twin Creeks mine pit lake, Golconda, Nevada, USA. In: Proceedings of 4th international conf on acid rock drainage (ICARD), Vancouver, Canada, pp 891–904

Kempton H, Conrad C, Parnow C, Philips S (2009) Linking sulfide mineral oxidation and metal release in models of acid rock drainage from mine waste. In: Proceedings of 8th ICARD, Skellefteå, Sweden

Kramer DA, Papp JF, LaTurno NM, Gambogi J (2018) Mine and mineral processing plant locations—supplemental information for USGS map I-2654. https://minerals.usgs.gov/minerals/pubs/mapdata/#datadocs. Accessed 09 Jan 2018

Kuipers JR, Maest AS, MacHardy KA, Lawson G (2006) Comparison of predicted and actual water quality at hardrock mines: the reliability of predictions in environmental impact statements. http://www.aktrekking.com/pebble/news/ComparisonsReportFinal.pdf

Marinelli F, Niccoli WL (2000) Simple analytical equations for estimating ground water inflow to a mine pit. Groundwater 38(2):311–314

Meteoblue (2019) https://www.meteoblue.com/en/weather/forecast/week/forst-%28lausitz%29_germany_6550656. Accessed 02 Nov 2019

Newman CP (2014) Geochemistry of arsenic, manganese, and iron in the Dexter pit lake, Tuscarora, Nevada. MSc Thesis, University of Nevada, Reno

Nordstrom DK (2009) Pitfalls and limitations of mineral equilibrium assumptions for geochemical modeling of water–rock interactions at mine sites. In: Proceedings of 8th ICARD, Skellefteå, Sweden

Nordstrom DK (2011) Hydrogeochemical processes governing the origin, transport and fate of major and trace elements from mine wastes and mineralized rock to surface waters. Appl Geochem 26(11):1777–1791. https://doi.org/10.1016/j.apgeochem.2011.06.002

Nordstrom DK (2012) Models, validation, and applied geochemistry: issues in science, communication, and philosophy. Appl Geochem 27(10):1899–1919. https://doi.org/10.1016/j.apgeochem.2012.07.007

Nordstrom DK, Archer DG (2003) Arsenic thermodynamic data and environmental geochemistry. In: Welch KG, Stollenwerk AH (eds) Arsenic in groundwater: geochemistry and occurrence. Kluwer Academic, The Netherlands

Oldham CE, Salmon SU, Hipsey MR, Ivey GN (2009) Modeling pit lake water quality: coupling of lake stratification dynamics, lake ecology, aqueous geochemistry, and sediment diagenesis. In: Castendyk DN, Eary LE (eds) Mine Pit Lakes: characteristics, predictive modeling, and sustainability. Society for Mining, Metallurgy and Exploration, Littleton, pp 127–136

Pakhomova SV, Hall POJ, Kononets MY, Rozanov AG, Tengberg A, Vershinin AV (2007) Fluxes of iron and manganese across the sediment–water interface under various redox conditions. Mar Geochem 107:319–331. https://doi.org/10.1016/j.marchem.2007.06.001

Parkhurst DL, Appelo CAJ (2013) Description of input and examples for PHREEQC version 3—a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. USGS Techniques and Methods, book 6, chap. A43

Ramteke S, Verma V, Chakradhari S, Sahu PK, Sahu BL, Rajhans KW, Yadav A, Patel KS (2016) Pit lake water quality of central India. J Geogr Inf Syst 8:28–39. https://doi.org/10.4236/jgis.2016.81003

Salmon SU, Oldham CE, Ivey GN (2008) Assessing internal and external controls on lake water quality: limitations on organic carbon-driven alkalinity generation in acidic pit lakes. Water Resour Res 44:W10414. https://doi.org/10.1029/2007WR005959

Salmon S, Hipsey M, Wake G, Ivey G, Oldham C (2017) Quantifying lake water quality evolution: coupled geochemistry, hydrodynamics, and aquatic ecology in an acidic pit lake. Environ Sci Technol 51(17):9864–9875. https://doi.org/10.1021/acs.est.7b01432

Sánchez-España J, López-Pamo E, Pastor ES, Ercilla MD (2008) The acidic mine pit lake of the Iberian pyrite belt: an approach to their physical limnology and hydrogeochemistry. Appl Geochem 23:1260–1287. https://doi.org/10.1016/j.apgeochem.2007.12.036

Savage KS, Ashley RP, Bird DK (2009) Geochemical evolution of a high arsenic, alkaline pit-lake in the Mother Lode gold district, California. Econ Geol 104(8):1171–1211. https://doi.org/10.2113/gsecongeo.104.8.1171

Schafer WM Logsdon M, Zhan G, Espell R (2006) Post-Betze pit lake water quality prediction, Nevada. In: Barnhisel RI (ed) Proceedings of 7th ICARD, St. Louis

Schultze M, Pokrandt KH, Hille W (2010) Pit lakes of the central German lignite mining district: creation, morphometry, and water quality aspects. Limnologica 40:148–155. https://doi.org/10.1016/j.limno.2009.11.006

Shevenell L, Connors KA, Henry CD (1999) Controls on pit lake water quality at sixteen open-pit mines in Nevada. Appl Geochem 14:669–687. https://doi.org/10.1016/S0883-2927(98)00091-2

SRK (2017) Predictive geochemical modeling of pit lake water quality at the Copper Flat Project, New Mexico. Prepared for THEMAC Resources Group, SRK Consulting (USA) (unpublished report)

Tempel RN, Shevenell LA, Lechler P, Price J (2000) Geochemical modeling approach to predicting arsenic concentrations in a mine pit lake. Appl Geochem 15:475–492. https://doi.org/10.1016/S0883-2927(99)00057-8

Tonkin JW, Balistrieri LS, Murray JW (2004) Modeling sorption of divalent metal cations on hydrous manganese oxide using the diffuse double layer model. Appl Geochem 19:29–53. https://doi.org/10.1016/S0883-2927(03)00115-X

Tucci NJ, Gammons CH (2015) Influence of copper recovery on the water quality of the acidic Berkeley pit lake, Montana, USA. Environ Sci Technol 49:4081–4088. https://doi.org/10.1021/es504916n

Vandersluis GD, Straskraba V, Effner SA (1995) Hydrogeological and geochemical aspects of lakes forming in abandoned open pit mines. In: Hotckkiss WR, Downey JS, Gutentag ED, Moore JE (eds) Proc. Water Resources at Risk, American Institute of Hydrology, Denver, pp 162–177

Weng L, Temminghoff EJ, Van Riemsdijk WH (2001) Contribution of individual sorbents to the control of heavy metal activity in sandy soil. Environ Sci Technol 35:4436–4443

Werner F (2009) Integration of prediction models. In: Castendyk DN, Eary LE (eds) Mine Pit Lakes: characteristics, predictive modeling, and sustainability. Society for Mining, Metallurgy and Exploration, Littleton, pp 159–165

Werner F, Müller M, Graupner B (2006) Predicted and observed water quality data from the coal mine pit lake Bärwalde, Laustz, Germany. In: Barnhisel RI (ed) Proceedings of 7th ICARD, St. Louis, pp 2333–2343

Wolery T, Jarek R (2003) Software Users Manual: EQ3/6, Version 8