A critical review of acid rock drainage prediction methods and practices

Anita Parbhakar-Fox, Bernd G. Lottermoser

School of Physical Sciences, University of Tasmania, Private Bag 79, Hobart, Tasmania 7001, Australia
Camborne School of Mines, University of Exeter, Penryn Campus, Penryn, Cornwall TR10 9FE, UK

ABSTRACT
Failure to accurately predict acid rock drainage (ARD) leads to long-term impacts on ecosystems and human health, in addition to substantial financial consequences and reputational damage to operators. Currently, a range of chemical static and kinetic tests are used to evaluate the acid producing nature of materials, from which risk assessments are prepared and waste classification schemes designed. However, these well-established tests and practices have inherent limitations, for example: (i) best-practice sampling is not pursued; (ii) risk assessments rely on limited static and kinetic test data, thus compromising the accuracy of resulting ARD block models; (iii) static tests are completed off-site and do not reflect actual field measurements; (iv) kinetic test data do not become available until later stages of mine development; (v) waste classification schemes generally categorise materials as only three types (i.e., PAF, NAF and UC) with other drainage forms (e.g., neutral metalliferous or saline) not considered; and (vi) conventional testing fails to consider that reactivity of waste is controlled by parameters other than chemistry (e.g., microbiology, type and occurrence of minerals, texture and hardness). Thus, accurate prediction is challenging because of the multifaceted processes leading to ARD. Hence, risk assessments need to consider mineralogical, textural and geometallurgical rock properties in addition to predictive geochemical test data. Instead, a new architecture of integrative, staged ARD testing should be pursued. Better ARD prediction must start with improving the definition of geoenvironmental models and waste units. Then, a range of low-cost and rapid tests for the screening of samples should be conducted on site prior to the performance of established tests and advanced analyses using state-of-the-art laboratories. Such an approach to ARD prediction would support more accurate and cost-effective waste management during operation, and ultimately less costly mine closure outcomes.

Crown Copyright © 2015 Published by Elsevier Ltd. All rights reserved.

Contents
1. Introduction .................................................. 108
2. Acid rock drainage formation .................................. 109
3. Geoenvironmental deposit modelling ......................... 112
4. Sampling strategies ............................................. 112
5. Short-term field tests ............................................ 113
  5.1. Drill core evaluation ........................................ 113
  5.2. Paste pH .................................................. 113
  5.3. USGS field leach test .................................... 113
  5.4. Wall washing ............................................. 113
  5.5. Total metals ............................................. 114
6. Static test methods .............................................. 114
  6.1. Carbon .................................................. 114
  6.2. Sulphur .................................................. 114
  6.3. Acid base accounting ................................... 115
  6.4. Net acid generation (NAG) tests ................. 115

* Corresponding author.
E-mail address: anitap1@utas.edu.au (A. Parbhakar-Fox).

http://dx.doi.org/10.1016/j.mineng.2015.03.015
0892-6875/Crown Copyright © 2015 Published by Elsevier Ltd. All rights reserved.
1. Introduction

Predicting acid rock drainage (ARD) is usually not an aspect that is strongly embedded into the development of mineral deposits as other aspects such as resource evaluation and testing for beneficiation, mineral processing and recovery attributes of ores and different ore types take priority. However, published evidence for the consequences of failing to predict and manage ARD for individual operations and for the mining industry as a whole is abundant, with estimated costs of US$ 100 billion for total worldwide liability associated with current and future remediation (Tremblay and Hogan, 2001). For large mines in settings favourable to the generation of ARD, unplanned closure costs have frequently been in the range of US$50–100 million, and sometimes beyond (Dowd, 2005).

Therefore, today’s mine regulators will only permit mining if robust waste management plans have been developed. However, this is not just a government requirement, with many financiers and stakeholders adopting guidelines to minimise environmental risk as published by the International Finance Corporation (IFC, 2007) and the Equator Principles III (2013). The typical requirement is for mine wastes to be appropriately characterised as part of an environmental impact assessment, with future characteristics of the materials also predicted (Azapagic, 2004; Price, 2009). Most significantly, the International Network for Acid Prevention (INAP) published the Global Acid Rock Drainage guide (GARD, 2014), a web-based Wikipedia-style handbook that covers pertinent topics relevant to ARD including prediction, rehabilitation and management. Whilst these up-to-date handbooks provide systematic information on how to undertake site-by-site ARD prediction, they do not greatly deviate from Morin and Hutt’s (1998) wheel approach to drainage chemistry prediction (Fig. 1). This comprises a variety of tests which are either field or laboratory based, and can be geochemical or mineralogical in nature.

Management and treatment of ARD affected sites can vary, but typically active additive approaches are preferred whereby neutraliser such as limestone (CaCO₃) and quick lime (CaO) are used to treat waste rock piles and tailings storage facilities (Kuyucak, 2001; Johnson and Hallburg, 2005; Akcil and Koldas, 2006; Carabello et al., 2009; Zhang, 2009; Simate and Ndlovu, 2014). Alternative strategies include those which focus on physical–chemical methods e.g., electrowinning (Vegliò et al., 2003; Gorgievski et al., 2009) and cation–anion exchange resins (Akcil and Koldas, 2006; Fu and Wang, 2011); and biological–chemical methods e.g., selective sequential precipitation of metals (Tabak et al., 2003; Luptakova et al., 2012) and packed bed bioreactors (Diz, 1997). Considering the magnitude and persistent occurrence of ARD liabilities, the long-term impacts on the environment, and the financial consequences to industry and society (Azapagic, 2004; Johnson and Hallburg, 2005), there is a growing need to provide accurate information of intrinsic rock characteristics likely to result in ARD. Such information is required early in the life-of-mine cycle because it impacts in particular on waste management throughout each phase. Early acid rock drainage (ARD) characterisation and risk assessment at the exploration, pre-feasibility and feasibility stages supports more effective management and valuation of ores and wastes during mineral processing, subsequent storage of waste and ultimately improved mine closure outcomes. Consequently, the existing ARD predictive tools and protocols need to be evaluated whether they provide robust, accurate and cost-effective characterisation in an industrial setting.

The objective of this paper is to critically review the methods and practices currently used for characterising mine and processing wastes for their ARD potential. The review does not present a complete catalogue of predictive tools nor does it document these
tools in detail. It briefly summarises the principles of ARD formation and aid an understanding of the predictive tests, reviews some of the strengths and weaknesses of the various predictive methods, and outlines opportunities for future improvement of ARD prediction practices.

2. Acid rock drainage formation

Acid rock drainage is produced by oxidation of sulphide minerals, particularly pyrite (FeS₂), and occurs via inorganic and biologically-mediated pathways (Evangelou and Zhang, 1995; Egiebor and Oni, 2007; Lottermoser, 2010). Solutions are characterised by low pH and high metal concentrations, and pose significant environmental problems because ARD waters can be harmful to humans and other life forms (Akcil and Koldas, 2006; Ma and Banfield, 2011). General oxidation reactions for pyrite are given in Eqs. (1)–(5). The initial step in the presence of atmospheric oxygen is described in Eq. (1) (Evangelou and Zhang, 1995). Oxygen directly reacts with pyrite and forms Fe³⁺, which is subsequently oxidised to Fe⁴⁺ (Eq. (2)). Ferric iron has a low solubility around pH 6. However, Morin and Hutt (2010) is likely an appropriate general approximation, encompassing the principles of ARD formation in the presence of microorganisms (Evangelou and Zhang, 1995; Schippers and Sand, 1999; Bond et al., 2000; Kuyucak, 2002; Baker and Banfield, 2003) particularly mosaicoacidiphilic chemolithotrophic species i.e., Acidithiobacillus ferrooxidans (oxidises Fe²⁺, S⁰ and metal sulphides as well as other reduced inorganic sulphur compounds), Acidithiobacillus thiooxidans (oxidises S⁰ and sulphides to sulphuric acid but not Fe²⁺) and Leptospirillum ferrooxidans (oxidises Fe²⁺ ions only). Singer and Stumm (1970) stated that microbial mediation accelerates the rate of pyrite oxidation by a factor >10⁶. However, Morin and Hutt (2010) regarded this as an overestimation, giving evidence from Morth et al. (1972) and Leathen et al. (1953a,b), who reported oxidation to proceed 10–50 times and 2–5 times faster, respectively. In reality, the estimate presented by Singer and Stumm (1970) is likely an appropriate general approximation, encompassing the cumulative contribution from all ARD microbes present in a given community.

Processes of pyrite oxidation are generally controlled by several other factors including pH (Evangelou and Zhang, 1995; Dold, 2010), oxidant type (O₂ or Fe⁴⁺; Moses and Herman, 1991; Moses et al., 1987; Hustwit et al., 1992), oxidant concentration (Lottermoser, 2010), morphology (Weber et al., 2004; Lottermoser, 2010), and trace element contents (Kwong, 1993, 1995; Jambor, 1994; Plumlee, 1999; Blanchard et al., 2007). Oxidation is further complicated by the electrochemical nature of pyrite which is a semiconductor, thus the exact chemical composition of the sample will dictate the specific oxidation pathway (Evangelou and Zhang, 1995; Rimmer and Vaughan, 2003; Egiebor and Oni, 2007; Savage et al., 2008; Chandra and Gerson, 2010).

Other sulphides have differing acid-forming potential and rates of reaction. This is dependent on the amount of Fe present, with Fe-sulphides generating the most acidity (Plumlee, 1999; Dold, 2010). Sulphides, which do not contain Fe in their crystal lattice (e.g., galena, Fe-poor sphalerite), do not have the capacity to directly generate large amounts of acid, but can be sources of potentially deleterious metals such as Cd, Pb and Zn (Dold, 2010; Lottermoser, 2010). General acid producing oxidation reactions for pyrrhotite (Fe₃₋₀.₅S), arsenopyrite (FeAsS) and chalcopyrite (CuFeS₂) for which oxygen is the oxidant, are given in Eqs. (6)–(8). Further details on oxidation pathways and the stability of reaction products are provided in Mok and Wai (1994), Nicholson and Scharer (1994), Corkhill and Vaughan (2009), Abbassi et al. (2009), Dold (2010) and Lottermoser (2010). Oxidation of sphalerite (ZnS) and galena (PbS) are shown in Eqs. (9) and (10), and in the presence of Fe⁴⁺ oxidation of MoS₂ (Me = divalent metal) will produce acidity (Eq. (11)) as documented in Dold (2010), Thurston et al. (2010) and Lottermoser (2010).

Fe₃₋₀.₅S + (2 – x)/2O₂ + xH₂O → (1 – x)Fe²⁺ + SO₄²⁻ + 2xH⁺ (6)
FeAsS + 7/2O₂ + 6H₂O → Fe(OH)₃ + SO₄²⁻ + H₂AsO₄⁻ + 3H⁺ (7)
2CuFeS₂ + 17/2O₂ + 5H₂O → 2Cu²⁺ + 2Fe(OH)₃ + 4SO₄²⁻ + 4H⁺ (8)
ZnS + 2O₂ → Zn²⁺ + SO₄²⁻ (9)
PbS + 2O₂ → Pb²⁺ + SO₄²⁻ (10)
2MeS + 4Fe⁴⁺ + 3O₂ + 2H₂O → 2Me²⁺ + 4Fe²⁺ + 2SO₄²⁻ + 4H⁺ (11)

Generally, sulphide reactivity has been reported in the order of: pyrrhotite > galena – sphalerite > pyrite – arsenopyrite > chalcopyrite (Keith and Vaughan, 2000; Moncur et al., 2009). However, this order is subject to change when other factors such as galvanic interactions (Kwong et al., 2003; Cruz et al., 2005) and biological interactions (Kuyucak, 2002; Egiebor and Oni, 2007) are considered. On oxidation product layers may develop which can control the rate of diffusion and thus the overall rate of oxidation (pyrite and pyrrhotite, Blowers and Jambor, 1990; Lottermoser, 2010; galena, Garcia et al., 1995; Diehl et al., 2006; sphalerite, Weisner et al., 2003; arsenopyrite, Harvey et al., 2006; Murcigo et al., 2011). Additionally, secondary efflorescent minerals may form, particularly in semi-arid and arid regions (Jambor et al., 2000; Harris et al., 2003; Nordstrom, 2009). These represent temporary stores of sulphate and metals and possibly hydrogen ions, which on dissolution will be released. A significant source of acidity in mine waste is also the precipitation of Fe⁴⁺ (Eq. (4)) and Al³⁺ hydroxides (Stumm and Morgan, 1996; Lottermoser, 2010). In addition, secondary Fe(III)-hydroxy-sulphates minerals (e.g., schwertmannite, Fe₆O₉(OH)₄SO₄ to Fe₅O₆(OH)₃(SO₄)₂) and jarosite, KFe₃(SO₄)₂(OH)₆ are commonly found in the oxidising environment of pyrite/pyrrhotite mine waste, or in the oxidised portion of weathering zones from ore deposits (i.e., gossans; Dold, 2010). Direct precipitation and dissolution of these phases is potentially acid forming with H⁺ ions liberated, in addition to transformation of these meta-stable phases to more stable Fe-hydroxides and oxides. Examples for jarosite are shown in Eqs. (12)–(14) (Welch et al., 2008; Dold, 2010).

3Fe²⁺ + K + 2SO₄²⁻ + 6H₂O → KFe₃(SO₄)₂(OH)₆ + 6H⁺ (12)
KFe₃(SO₄)₂(OH)₆ + 3H₂O → 3FeO(OH) + K⁺ + 3H⁺ + 2SO₄²⁻ (13)
jarosite → goethite : KFe₃(SO₄)₂(OH)₆ → 3FeO(OH) + K⁺ + 2SO₄²⁻ + 3H⁺ (14)

Acid formed by sulphide oxidation can be consumed through reaction with gangue minerals. Neutralisation is primarily offered by dissolution of carbonate minerals of which calcite is the most effective (Sherlock et al., 1995; White et al., 1999; Lapikko, 2002; Frostad et al., 2002; Lottermoser, 2010). Some neutralisation is
<table>
<thead>
<tr>
<th>Technique</th>
<th>Test Principle</th>
<th>Description</th>
<th>Advantages</th>
<th>Limitations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geo-environmental deposit modelling</td>
<td>Desk-study during early life of mine stages</td>
<td>Information on environmental characteristics of a deposit</td>
<td>* Highlights potential environmental problems that arise from the unique geological attributes specific to a deposit type&lt;br&gt; * Allows for early forecasting of potential liabilities and identification of appropriate rehabilitation options</td>
<td>* Too many assumptions made on ARD controls&lt;br&gt; * Bias of current models to North American sites&lt;br&gt; * Current published models are either too site specific, or too general</td>
<td>Kwong (1993)&lt;br&gt; Du Bray (1995)&lt;br&gt; Ficklin et al. (1992)&lt;br&gt; Plumlee et al. (1999)&lt;br&gt; Seal and Foley (2002)</td>
</tr>
<tr>
<td>Sampling strategies</td>
<td>Determination of appropriate sample numbers to use in ARD prediction testwork</td>
<td>General guidelines for number of samples based on stage of operations or deposit mass</td>
<td>* Early apportioning of funds to finance the predictive campaign permitted through early identification of the minimum sample number required</td>
<td>* No global standard procedure for sampling&lt;br&gt; * Sample selection may be non-systematic, can result in bias sampling</td>
<td>Downing (1999)&lt;br&gt; Price (2009)&lt;br&gt; Kentwell et al. (2012)</td>
</tr>
<tr>
<td>Short-term field tests</td>
<td>Drill core evaluation</td>
<td>Environmental logging of intact drill core by codes such as the acid rock drainage index</td>
<td>* Cost effective deposit-wide evaluation of ARD forming characteristics on intact samples</td>
<td>* Subjective&lt;br&gt; * Additional time added to geological logging</td>
<td>Parbhakar-Fox et al. (2011)&lt;br&gt; GARD Guide (2014)</td>
</tr>
<tr>
<td></td>
<td>Paste pH</td>
<td>Water added to powdered or crushed sample, pH measured after a period of time</td>
<td>* Rapid indication of leachable acidity, resulting data can be used as part of ARD domaining</td>
<td>* Not a predictive test&lt;br&gt; * No global protocol standard exists</td>
<td>Hammarstrom et al. (2003)&lt;br&gt; Smart et al. (2002)&lt;br&gt; ASTM D4972.01 (2007)&lt;br&gt; Noble et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>USGS Field leach test</td>
<td>Water added to sieved sample, solution analysed for dissolved metals</td>
<td>* Fast, simple, and cost-effective, can be performed in the field with no specialist equipment required</td>
<td>* Not a predictive test</td>
<td>Hageman (2007)</td>
</tr>
<tr>
<td></td>
<td>Wall washing</td>
<td>Measurement stations constructed on exposed rock surface, rinsed with deionised water, and leachate analysed&lt;br&gt; XRF, ICP or portable XRF (pXRF) analyses</td>
<td>* Results can support modelling of future mine water quality</td>
<td>* Complex experimental setup&lt;br&gt; * Infrequent sampling does not permit an understanding of net chemical changes</td>
<td>Morwijk (1995)&lt;br&gt; Price et al. (1997)&lt;br&gt; Price (2009)</td>
</tr>
<tr>
<td>Total metals</td>
<td>Quantitative assessment of metals and metalloids</td>
<td></td>
<td>* Quantitative data of total metals and metalloids</td>
<td></td>
<td>Price (2009)</td>
</tr>
<tr>
<td>Static tests</td>
<td>Carbon</td>
<td>Assessment of acid neutralising capacity (ANC)</td>
<td>* Relatively simple and low-cost procedure to determine C\textsubscript{organic}&lt;br&gt; * Efficient procedure to determine C\textsubscript{organic}&lt;br&gt; * Some carbonate minerals do not contribute to ANC&lt;br&gt; * Net carbon value (NCV) methodology is not widely used</td>
<td>* In-situ ARD prediction&lt;br&gt; * MPA based on S\textsubscript{total} can lead to overestimation if not acid forming minerals (e.g., barite, gypsum; non-Fe bearing sulphides) are present</td>
<td>Bucknam (1997)&lt;br&gt; Ross et al. (2013)&lt;br&gt; Gazley et al. (2014)&lt;br&gt; Weber et al. (2004)&lt;br&gt; Price (2009)&lt;br&gt; Lengke et al. (2010)&lt;br&gt; White et al. (1999)&lt;br&gt; Paytan et al. (2004)&lt;br&gt; Praharaj and Fortin (2004)&lt;br&gt; Burton et al. (2008)&lt;br&gt; Price (2009)</td>
</tr>
<tr>
<td></td>
<td>Sulphur</td>
<td>Assessment of maximum potential acidity (MPA)</td>
<td>* Simple to measure&lt;br&gt; * Occasionally, data collected by routine S\textsubscript{total} assay for other mine-site disciplines</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique</td>
<td>Test</td>
<td>Principle</td>
<td>Description</td>
<td>Advantages</td>
<td>Limitations</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Acid base accounting (ABA)</td>
<td></td>
<td>Caused by Net Acid Producing Potential (NAPP) or Neutralisation Potential Ratio (NPR)</td>
<td>MPA calculated based on $S_{\text{Total}}$ or $S_{\text{Sulphate}}$ or ANC determined by laboratory titration methods</td>
<td>Industry-wide practice, with commercial laboratories offering this testing, therefore results are easy to understand and interpret</td>
<td>MPA overestimated if based on $S_{\text{Total}}$ or ANC and MPA tests are not performed on exactly the same sample material, ABA procedure is not standardised, site-by-site comparisons are not possible, Organic matter interferes with measurements</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Acid Generation (NAG) test</td>
<td></td>
<td></td>
<td>Addition of H$_2$O$_2$</td>
<td>Industry-wide practice, NAGpH versus NAPP geochemical plots routinely used for waste classification</td>
<td>H$_2$O$_2$ quality must be ensured, Calcite and dolomite may react during NAG testing, Ca(OH)$_2$ and alkaline pH values result</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Interpretation of results is cumbersome and not standardised, Inappropriate in replicating a natural environment with regards to extracting fluid concentration and pH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Technically challenging, Not an appropriate screening tool, Role of other ARD-forming bacteria not considered</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acid buffering characteristic curve (ABC) test</td>
<td>Leachable metals</td>
<td>Measurement of acid neutralising capacity</td>
<td>HCl titration</td>
<td>Only measures the effective acid neutralising portion</td>
<td>Not an appropriate screening tool, Role of other ARD-forming bacteria not considered</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological acid production potential testing</td>
<td></td>
<td></td>
<td>Synthetic precipitation leach procedure (SPLP method 1312)</td>
<td>Screening-tool for leachable metals</td>
<td>No integration of mineralogical and microtextural analyses</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinetic tests</td>
<td>Humidity cell tests</td>
<td>Mimics weathering of waste rock, with leachate chemistry assessed</td>
<td>ASTM D5744 method</td>
<td>More realistic indication of leachable metals/metalloids than short-term SPLP tests, Attempt to mimic climatic cycle</td>
<td>Mineral precipitation does reflect actual mine site conditions, Day 7 rinse may remove all reaction products, Equilibrium conditions not reached due to short contact times, No integration of mineralogical and microtextural analyses, Unrealistic replication of climatic conditions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field test methods</td>
<td>Mineralogical characterisation</td>
<td>Monitor leachate quality of waste rock directly under natural conditions</td>
<td>Tubs, barrels or test piles of wastes left to weather under natural conditions</td>
<td>Exposure to site-relevant climatic and microbiological conditions</td>
<td>Not standardised, Larger scale piles expensive to construct and maintain</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optical microscopy</td>
<td></td>
<td>Identification of acid forming and neutralising minerals</td>
<td>Sulphide alteration index (SAI) and acid rock drainage index (ARDI) proposed</td>
<td>Examination of sulphide and carbonate textures</td>
<td>Not standardised, Technically challenging, Subjective, Not statistically valid, Expensive, Detection limits, Identification of trace phases impossible</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulk mineralogy</td>
<td></td>
<td>Quantification of acid forming and neutralising minerals</td>
<td>X-ray diffractometry (XRD)</td>
<td>Quantitative mineralogy, Data can be used for MPA/ANC calculations</td>
<td>Not statistically valid, Expensive, Detection limits, Identification of trace phases impossible</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
offered from silicate mineral dissolution, particularly olivine, wollastonite and serpentine phases (Jambor et al., 2002, 2007). However, the rate of dissolution is much slower than that of carbonates. Additionally, clay minerals have a neutralisation capacity, but the likely net-neutralising contribution is small compared to that of calcite (Lottermoser, 2010).

Considering the diversity of sulphide oxidation controls and ARD generation pathways, it is appreciable that predicting ARD is by no means a ‘one-test’ field. Rather, to undertake adequate ARD prediction, the geology, mineralogy and geochemistry of a deposit must be carefully considered in addition to the performance of the most appropriate tests on a suitable number of samples. Moreover, such testing is most commonly pursued by the mining industry and therefore, practical, accurate, rapid and cost-effective methods and best practices need to be followed. A summary of the most common tests and methods is presented in Table 1 and discussed in the following sections.

3. Geoenvironmental deposit modelling

Mineral deposits are natural concentrations of one or more mineral commodities that have formed within specified geological settings. Different genetic processes result in different mineral deposit types that share a set of geological attributes and contain a particular commodity (or commodities) that collectively distinguish them from other types. Consequently, classifications of ore deposits are based on genetic models or empirical features that tend to show the natural groupings of different deposit types. Ore deposit classifications are of particular use in ARD prediction if the deposit classification has been based on empirical features (e.g., minerals, host–rock association) rather than genetic models. The reasoning is that ore deposits of the same empirical type have the same ore and gangue minerals and accordingly, the same acid producing and acid buffering material. Hence, Kwong (1993) proposed a general ranking of ore deposits based on their potential to form ARD, whereas Plumlee (1999) approached this differently by examining the chemical and physical ARD effects of the most common host rock lithologies and alteration styles associated with hydrothermal deposits. Similar to the geoenvironmental modelling of sulphidic wastes, modelling of mine waters involves the classification of the deposit and the deduction of water quality problems using Ficklin diagrams that plot pH versus total metal contents (Ficklin et al., 1992; Plumlee et al., 1999). Ficklin plots are based on the chemical properties of drainage waters from various ore deposits, and different deposit types plot in distinct regions of the diagram. Clearly, such modelling only provides general insights in the overall likelihood and nature of acid generation, yet it assumes that factors influencing acid generation such as pyrite surface area, abundance of sulphides and carbonates, or waste dump characteristics are constant for the mine sites and ore deposits being compared. Thus, geoenvironmental modelling of an ore deposit is only an initial step in ranking deposits in terms of their potential to produce ARD.

Ore deposit classifications are periodically updated (most recently, Verplank et al., 2014). Similarly, ranking of ore deposits according to their ARD potential requires an urgent iteration since its inception in 1993 (Kwong, 1993). This revised ranking should be based on empirical mineral deposit classifications (rather than genetic ore groupings) because empirical deposit classes are descriptive and based on mineral abundances and host–rock associations.

4. Sampling strategies

Sampling for ARD prediction occurs at several stages of mine operations, from exploration to mine closure. Similar sampling
principles apply at each of these stages. Sample selection is the most critical aspect of an investigation aimed at determining the acid-forming potential of mine wastes using field and laboratory techniques. In the early stages of operation, sampling must be representative of all geological, lithological and alteration units related to the mine development plan, and must also be representative of the relative amounts and particle size of each type of material (Price, 2009).

However, adequate and accurate sampling is influenced by several factors such as: (i) budget available for environmental characterisation; and (ii) selection of appropriate sample mass and materials. Inadequate sampling or even inaccurate sample selection can contribute to excessive variance, difficulties in interpretation, and incorrect assessment (Downing, 1999).

Most recently, Lotter et al. (2014) discussed sampling issues in the context of flotation and process mineralogy; many of these points can be applied to ARD testwork sampling. For example, Lotter (2005) argue for the definition of a geometallurgical unit, which is defined as an ore type, or group of ore types, that possesses a unique set of textural and compositional properties leading to similar metallurgical performances. For ARD prediction, this concept must be adapted to define geoenvironmental units (previously referred to as mesotextural groups in Parbhakar-Fox et al., 2011). Such units must be defined based on similar textural, mineralogical and chemical properties from which a similar ARD forming and metal leaching potential can be predicted. Furthermore, identifying geoenvironmental units must not be restricted to ore, but must extend to waste. Through defining geoenvironmental units, representative samples can be chosen based on a sampling interval dictated by site-based definitions and abundances of local rock types.

Downing (1999) reasoned that whilst rigorous sampling can be costly, remediation and reclamation plans developed as a consequence of poor sampling and analysis can become even more expensive, hence the importance of a broad initial assessment. He argued that costs should not pre-determine the number of samples taken and analysed, but it should be dependent on the amount necessary to increase confidence in the data. Despite this argument, ultimately, ARD testing is limited by financial constraints; therefore, low-cost short-term field tests are needed that support deposit-wide ARD characterisation.

5. Short-term field tests

Low-cost, on-site field tests are an essential part of any ARD testwork program. The most commonly pursued tests and practices include drill core observations, field leach tests for pH and dissolved metals, and analyses for total metal concentrations using portable X-ray fluorescence instruments.

5.1. Drill core evaluation

Interpretation of the nature and relationships of sulphide and carbonate mineralogy on a core-scale is by far the most efficient way to gather a deposit-scale understanding of the potential for acid formation. Logging is generally undertaken for metallurgical, geotechnical and resource evaluation. The GARD Guide (2014) identified parameters that need to be observed for understanding the potential for ARD formation whilst logging. However, guidelines of how to evaluate these were not given. Parbhakar-Fox et al. (2011) addressed this through development of the Acid Rock Drainage Index (ARDI), a site-specific logging code designed to numerically score five key parameters that are known to control ARD formation, including: (i) sulphide and carbonate mineralogy; (ii) primary and secondary neutralisation mineralogy; (iii) alteration variations; (iv) modal mineralogy; and (v) grain boundaries as reaction sites for acid formation and neutralisation.

5.2. Paste pH

Paste pH tests represent basic testing procedures for leachable acidity. The methods involve adding a quantity of water to a powdered or crushed sample, with the pH measured after a period of time (<24 h; Price, 2009). The application of paste pH tests in predictive ARD testwork programs has been debated (e.g., Kania, 1998), because the test only measures the current acid forming characteristics of a material. Despite this, the test provides rapid information on ‘first flush’ characteristics, and has been used by various authors in their predictive ARD testwork (cf. Shaw, 2000; Weber et al., 2006; Hughes et al., 2007; Parbhakar-Fox et al., 2011). However, several pH test protocols are being pursued, with each method requiring a different solid: solution ratio, extraction solution, mass, grain size, equilibration time and procedure. In fact, Hammarstrom et al. (2003) and Noble et al. (2012) undertook comparative studies of various paste pH procedures and reported significant differences in pH measurements between methods. Noble et al. (2012) recommended that the methodology for pH tests applied to ARD risk assessment should (i) use 0.01 M CaCl₂ as the extraction solution; (ii) have a high solid to solution ratio; (iii) be conducted on a more than one particle size (e.g. pulverized <0.075 mm and coarse crushed material <20 mm); and (iv) include the analysis of certified reference materials (CRMs). Moreover, Noble et al. (2012) identified the American Society for Testing and Materials (ASTM) D4972.01 (2007) Method as the most appropriate to achieve reproducible paste pH data. By stabilising the electronic conductivity (EC) through using an electrolyte solution and not deionised water (i.e., CaCl₂) liquid junction effects associated with the pH anode are reduced, and data accurately collected.

5.3. USGS field leach test

The USGS field leach test (FLT; Hageman, 2007) uses deionised water to leach samples for its soluble metals, with an example of its application presented in Diehl et al. (2006). This test is ineffective as a predictive tool (i.e., similar to paste pH testing), because it only allows an evaluation of the leachable metal load which may eventuate during a ‘first-flush’ weather event. Recognising this, McLemore et al. (2014) proposed a mine waste decision tree whereby samples with a FLT pH ≤ 5 are assumed to have toxicity, and are therefore prioritised for laboratory based leach testing. Instead, the development of a better predictive field-leach test is needed, which uses a stronger chemical reagent than deionised water and ultimately resembles actual field conditions.

5.4. Wall washing

A methodology for determining the ARD characteristics of in-situ rock (e.g., pit walls, rock surfaces of adits, stopes and other underground workings) was developed by the Canadian Mine Environment Neutral Drainage (MEND) and the British Columbia ARD Task Force termed MINEWALL (Morwijk, 1995). First, MINEWALL stations are constructed using flexible bathtub edging, pure silicon sealant and plastic sheeting on an area of 1 m × 1 m of exposed rock surface (Price et al., 1997). This area is carefully rinsed with 11 of deionised water, with the leachate collected and geochemically analysed by typical ARD parameters (e.g., pH, sulphate, acidity, alkalinity and total metals; Price et al., 1997). High rinse frequencies are recommended to measure ARD and metal leaching changes over time (e.g., Bell Mine and Island Copper Mine, British Columbia; Morwijk, 1995). However, reports
of single flushing have also been documented (Equity Silver Mine, British Columbia; Morwijk, 1995). Limitations associated with this method include the technical difficulty in constructing the washing stations on fractured rock, as well as estimating the source area contribution of unsewn fractures. Such small fractures can be sealed, however, to test incompetent rock surfaces (i.e., weathered and hydrothermally altered) is challenging, and therefore, testing of all rock units may not be possible. MINEWALL 2.0 software was also developed to assist with mine water predictions for complex scenarios continuously through the operational and closure phases of a mine (Morwijk, 1995). To date, this methodology has only been performed at a small number of mines in Canada, the United States, Mexico, Indonesia and the Caribbean (Price, 2009).

5.5. Total metals

ARD testwork programmes should not only focus on determining acid forming characteristics of materials, but also assess the presence and abundance of environmentally significant metals and metalloids present within the rocks that may become mobile. Traditionally, a representative set of samples is chosen and analysed for a suite of metals/metalloids by off-site XRF (X-ray fluorescence) or ICP (inductively coupled plasma) methods. By contrast, field portable XRF (pXRF) has the potential to allow for chemical analyses of materials on-site prior to scheduling of costly analyses at an off-site laboratory (Ross et al., 2013; Gazley et al., 2014). Limitations of utilising pXRF technology exist with regards to which sample media to select for analysis (i.e., homogenised powdered sample vs. intact drill core). To reduce sampling bias, pXRF should be performed on powdered materials and appropriate CRMs during analyses (Hall et al., 2014; Piercey and Devine, 2014).

6. Static test methods

Once field tests are completed, static tests are applied to processed samples. Static tests are short-term (typically measured over a short time period i.e., hours or days) and relatively low-cost tools that quantify a sample’s acid forming or neutralising potential (White et al., 1999; Chotpantarat, 2011). The most widely used static test protocols fall into two types, acid base accounting (ABA) and net acid generation (NAG) tests. Additional static tests have recently been proposed such as the acid buffering characteristic curve (ABCC) method (Smart et al., 2002) and the carbonate bomb test (Hughes et al., 2007). Whilst the role of carbonate minerals in ARD neutralisation has been well established and is measured in static testing protocols, that of the silicate minerals has not been fully characterised, though it is understood that they contribute to neutralisation over a longer-time scale (e.g., Miller et al., 2010). Despite limitations associated with these tests documented in this section, the results of static tests are widely used to classify waste materials as potentially acid forming (PAF), non-acid forming (NAF) or uncertain (UC) waste types (e.g. Smart et al., 2002).

6.1. Carbon

Total and inorganic carbon can be measured to allow for the direct calculation of acid neutralising capacity. Typically, total carbon is measured using induction furnace methods (Crock et al., 1999; Price, 2009). Carbon is converted to CO₂ by heating a sample to approximately 1650 °C in a high-frequency electrical field under an O₂ atmosphere, and gases analysed by infrared spectrometry. The evolved CO₂ can be quantitatively measured by other standard volumetric or titrimetric methods. Organic carbon is reacted off by HCl digestion, with the dried residue analysed also by high-temperature furnace methods to determine the organic carbon fraction (i.e., plants material, graphite; Price, 2009). Alternatively, coulometric methods can be used to determine carbonate-carbon (Price, 2009). However, this is less commonly undertaken due to time and expense per sample. Total carbon values are subtracted from the organic values, to calculate inorganic carbon, which is taken to reflect the presence of carbonates. The carbonate neutralising capacity (CO₃-NP) is calculated by multiplying the percentage of inorganic carbon by 83.4, with resulting values expressed in kg CaCO₃/tonne (Price, 2009). Borden (2003) and Weber et al. (2004) presented examples in which total-carbon values were used in ARD calculations.

The net carbonate value (NCV) methodology also uses total carbon values to determine acid neutralising capacity (ANC), with such analyses routinely performed as part of acid base accounting by companies such as Newmont Mining Corporation. It can be calculated directly from semi-quantitative mineralogical information including percentage of carbonate and sulphide minerals, or from metallurgical assay analysis with total carbon and total sulphur values used (Bucknam, 1997). Mineralogical NCV is calculated based on using carbon dioxide (% CO₂) as the common unit for balancing the acid neutralising capacity or neutralising potential (ANC or AP) and maximum potential acidity or acid potential (MPA or AP) calculated from stoichiometric equivalents of one sulphide being oxidised to sulphuric acid and consuming one CO₂. Bucknam (1997) recommended that at least 100 samples (bench waste composites) are tested for NCV and 10–12 for mineralogical NCV confirmation tests. Representative samples of core should next be selected for each major geological waste-NCV classification (>5%) for further testing for quality assurance purposes. The classification criteria have since been modified by the Newmont Mining Corporation (NMC), and this NCV method has been standardised by the American Society for Testing and Materials Method E 1915-07 (ASTM, 2007). Application of this method is presented in Lengke et al. (2010).

6.2. Sulphur

The first insight into the likely magnitude of acid production of a sample is through measurement of total sulphur (S_{total}) with examples given in Downing and Giroux (1993), Borden (2003), Weber et al. (2005) and Hakkou et al. (2009). Sulphur contents are commonly determined by placing a sample in high-temperature furnace and analysing the developed gases by infrared spectrometry. However, the accuracy of basing calculations of acid potential on S_{total} is often debated (White et al., 1999; Parbhakar-Fox et al., 2011), as non-acid forming S-bearing species are included in the analyses (e.g. gypsum, barite). Consequently, methods such as chromium reducible sulphur (CRS) have been developed that target the removal of sulphates allowing for direct measurement of S_{Sulphide} (Prahraj and Fortin, 2004; Burton et al., 2008). Schumann et al. (2012) compared the accuracy of S_{total} versus CRS in the context of acid base accounting on coal mine and base metal mine wastes. They found that CRS was useful for the assessment of coal wastes with high organic sulphur, but the technique underestimated the acid potential of weathered wastes.

Alternative methods for sulphide speciation involve chemical pre-treatment (i.e., HCl) of samples to remove soluble sulphates prior to analysis. However, refractory barite is not removed by this method (Paytan et al., 2004). Typically, this is overcome by also performing an elemental assay to measure Ba and then calculate barite quantity. Alternatively, Tuttle et al. (2003) proposed a six-step sequential extraction procedure to separate the common forms of sulphur in rock samples. The procedure included various chemical treatments including washing with acetone, HCl and Cr²⁺. Whilst sulphide speciation is achieved, this sequential...
procedure represents a multi-faceted, laboratory-based and time-consuming process.

6.3. Acid base accounting

Acid base accounting requires the determination of MPA (or AP), and ANC (or NP). Calculation of MPA (or AP) is directly from total sulphur ($S_{\text{Total}}$) or sulphide–sulphur ($S_{\text{Sulphide}}$) values expressed in wt.%. The $S_{\text{Total}}$ or $S_{\text{Sulphide}}$ value is multiplied by the stoichiometric factor of 30.6 to give the MPA value in kg $\text{H}_2\text{SO}_4$/t (Weber et al., 2005; Stewart, 2005). Alternatively, a factor of 31.25 is used to give MPA in kg CaCO$_3$/t (White et al., 1999). However, the inaccuracy of using these factors for samples containing sulphides other than pyrite has long since been recognised. For example, Paktunc (1999) stated that if the sample being tested contains pyrrhotite in addition to pyrite, then overestimation of MPA/AP values may occur up to 1.5 times. This likely reflects that the ABA protocol was first developed for assessing ARD potential of coal mine wastes, in which pyrite is the dominant acid-forming sulphide mineral. Despite this, the procedure for calculating MPA remains unchanged.

The ANC (or NP) is a quantitative measurement of a solid phase sample’s capacity to neutralise aqueous acidity (Morin, 2009; Morin and Hutt, 2009). Typically, these tests cost approximately $30–$50 AUD. It is typically calculated by a titration method. The most widely used is the Sobek method (White et al., 1999; Bezaazoua et al., 2004; Jambor et al., 2006), which has subsequently been modified, and several additional methods exist (White et al., 1999; Price, 2009). However, conducting wet-chemical tests introduces potential for experimental error. For example, the strength of HCl for use in a Sobek test is determined by the fizz-reaction initially performed, which is ranked intuitively from ‘no fizz’ to ‘strong fizz’ (White et al., 1999). The interpretation of a fizz-reaction is subjective, potentially resulting in the wrong strength of reagents later used. As these tests are not repeatable, care must be ensured to apply appropriate QA/QC procedures. However, CRMs are not always identified and recommended for use in published methodologies (e.g., Smart et al., 2002; Price, 2009; Mills et al., 2011). Instead, the use of ABA CRMs such as KZK-1 and NBM-1 which were used by Paktunc (2001), Goodall (2008) and Parbhakar-Fox et al. (2011) must be ensured. There is further potential for conflicting results due to varying ANC protocol variables. These include particle size, back-titration end-point, temperature and digestion duration (White et al., 1999). Indeed, comparative studies between methods have reported different ANC values. For example, Capenema and Ciminelli (2003) identified the modified Sobek method as more accurate when characterising three sulphidic ore samples. Morin (2009) reported that Sobek ANC values when compared to modified methods showed variations (either higher or lower) 40–60% of the time when properly conducted. However, whilst there is much potential for discrepancy, Morin and Hutt (2009) stated there is no logic to arguing which ANC method is correct because first, ANC is an intrinsic material property and second, all methodologies utilise correction factors. Alternative methods to ANC calculations were presented in Bucknam (1997) and Hutt and Morin (2000), who based these directly on carbonate contents. Methods to determine silicate-ANC have been presented by Nesbitt and Jambor (1998), Jambor et al. (2002, 2007) and Miller et al. (2010), but these have yet to be applied and adopted by industry.

Despite the arguments presented by Morin and Hutt (2009), selection of an appropriate standard method for determining ANC is required by environmental legislation in a particular jurisdiction. For example, the U.S. EPA-600-compliant Sobek et al. (1978) NP Method is standard in the United States (Morin and Hutt, 2009). In Australia, the Sobek and modified Sobek methods are the most widely used in accordance with legislation specified by Commonwealth, state and local Governments (Comarmond, 1997). In Europe, the PRen 15875 Standard was recently established (cf. van der Sloot and van Zomeran, 2012), which follows the modified ABA method of Lawrence and Wang (1996).

After MPA and ANC have been calculated, the net acid producing potential (NAPP = MPA–ANC), net neutralising potential (NNP = ANC–MNA), or neutralising potential ratio (NPR = ANC/MPA) value is calculated and expressed either in terms of kg $\text{H}_2\text{SO}_4$/t (e.g., in Asia-Pacific), or kg CaCO$_3$/t (e.g., in North America). Using cut-off values for NPR and NAPP (or NNP), samples are then classified as potentially acid forming (PAF), non-acid forming (NAF) or uncertain (UC) waste types (e.g., Skousen et al., 2002; Fey, 2003; Bezaazoua et al., 2004). Despite limitations associated with ABA calculations, they remain the most widely used, as they are the quickest and most cost-effective way by which to quantify ARD forming potential.

6.4. Net acid generation (NAG) tests

NAG tests quantify acid forming potential by reacting the powdered material with hydrogen peroxide ($\text{H}_2\text{O}_2$), and back titrating with NaOH to determine the acid forming capacity (also expressed in kg $\text{H}_2\text{SO}_4$/t). A single-addition NAG test typically costs < $50 AUD. The first part of the procedure is determination of NAGPh. This value is measured after the powdered sample (typically 2.5 g) has reacted with 250 ml of $\text{H}_2\text{O}_2$, and cooled to room temperature, with the pH measured prior to the back titration (Smart et al., 2002). The NAGPh value is commonly plotted against the net acid producing potential (NAPP) value to allow for waste classification. A significant limitation of the commonly used single-addition NAG method is incomplete oxidation of sulphides in samples containing > 0.7–1 wt.% pyritic sulphide (Stewart, 2005). Therefore, a single-addition method is inappropriate for the majority of mineralised waste materials. To address this, three further NAG tests were developed, the multi-addition (mNAG), sequential (sNAG) and kinetic NAG (kNAG) tests (Smart et al., 2002). The soundness of the various NAG test procedures for assessing mine wastes (and other waste products) for their acid production can be challenged on a number of points. For example, carbonate minerals such as calcite and dolomite dissolve during NAG testing (cf. Becker et al., 2014) which in turn likely produce Ca(OH)$_2$, and therefore NAG testing of calcite-rich samples results in unrealistic alkaline NAGPh values and buffering of sulphuric acid produced during NAG testing. Also, quality control of the applied hydrogen peroxide is vital for accurate NAG measurements because inadequate storage may reduce the shelf life of the chemical and the presence of stabilisers leads to variable solution pH, with both aspects influencing NAG measurements.

6.5. Acid buffering characteristic curve (ABCC) test

The ABCC test is regarded as a higher cost alternative test to the ANC procedure. It intends to provide an indication of the ANC portion available for neutralisation (Miller and Jeffery, 1995; Smart et al., 2002; Weber, 2003; Tran et al., 2003; Stewart et al., 2006). The test involves the addition of 100 ml DI water to a sample (2.0 g; <75 µm) and then slow titration with HCl while continuously stirring and monitoring to pH 2.5 (Smart et al., 2002; Tran et al., 2003). Generally, it is recommended for use when assessing if a sulphidic sample with NAPP < 0 and NAG pH < 4.5 has enough readily available carbonate to render it non-acid forming (Smart et al., 2002). The major limitation of the ABCC test is the difficulty in interpreting the output results, as the graphic outputs are...
compared against reference curves (e.g., Stewart et al., 2006). If the shape of the output curve does not approximate that of the reference curves, then resolving carbonate mineralogy and understanding the reaction pathway becomes cumbersome.

6.6. Leachable metals

Chemical leach test methods for determining metal mobility exist, with the synthetic precipitation leach procedure (SPLP; EPA Method 1312) the most widely used to assess the risk of water contamination posed by leaching of mine wastes (Hageman et al., 2000). The validity of the SPLP test procedure for assessing mine wastes (and other waste products) for their metal leaching can also be challenged on a number of points. For example, there remains uncertainty whether the SPLP leachate concentrations represent the actual pore water concentrations in the waste, or whether they represent diluted concentrations as might be expected in the receiving aquifer or surface water. Also, the choice of nitric acid for the extraction fluids does not reflect ARD environments and its use may overestimate dissolution of minerals that are only soluble in nitric acid (e.g. lead arsenates). Moreover, the mildly acid pH conditions (pH 4.2, 5.0) of the SPLP test is rather high for an ARD environment and does not consider that leaching of some amphoteric metalloids and metals may occur in an alkaline environment. The test can also be challenged on the basis of the particle size reduction requirements (<9.5 mm). Size reduction of samples does not reflect the coarse grain size in which mining and mineral processing wastes are generally produced and managed. Finally, there is no consideration of redox potential and its influence on the leach behaviour of elements. Metal release from mining and mineral processing waste is (i) influenced by pH, redox potential, equilibration time and particle size, and (ii) cannot be determined by a single leach test such as the SPLP method. Clearly, the SPLP test can be applied as a screening tool for mining, mineral processing and metallurgical waste, however, its application as a waste classification standard appears to be limited.

6.7. Biological acid production potential testing

Whilst the role of biotic processes in accelerating ARD formation has long been established (cf. Singer and Stumm, 1970; Bond et al., 2000; Baker and Banfield, 2003), tests which directly attempt to measure this are the least often performed. The American Society for Testing and Materials (ASTM) E1915-11 Standard includes a testing procedure in its appendices (X1) for a biological acid producing potential (BAPP) test, which is based on the BC Research Confirmation testing method. Essentially, this test determines if the sulphide oxidising bacteria can generate enough H2SO4 from the sulphides in the sample to satisfy its acid demand. If they can, microbial action will continue on a self-sustaining basis, and ARD will result. The testing procedure recommends using the same powdered samples as used in routine ABA testing. It involves partial acid leaching of samples inoculated with A. ferrooxidans and subsequent metal analysis of the leachate. These data are compared against metals data resulting from simple leaches (e.g., paste pH test or USGS FLT). In practice, this methodology is rarely performed, possibly due to the length of time per analysis (i.e., several days). Furthermore, contributions from other acidophilic bacteria (e.g., A. thiooxidans; L. ferrooxidans) are not measured by this method. Alternative biological testing procedures have more recently been proposed including the biokinetic test (Hesketh et al., 2010). However, applications of this methodology have been in a research context only. Thus, the application of microbiological tests for commercial-scale ARD prediction test-work has yet to be established.

7. Kinetic test methods

The greatest criticism of static laboratory methods is the vast differences between testing conditions and the field. One of the greatest challenges in these methods is the successful upscaling of results to field scale. Currently, there is no consensus on which static method/test most accurately reflects field conditions. Questions have arisen regarding the length of laboratory test times and extrapolation to field weathering time. Consequently, a variety of simulated kinetic weathering tests have been developed and studied as drainage quality predictors (e.g., humidity cells, column leach tests). These tests incur higher costs per sample (frequently $1000s AUD).

The details of kinetic tests differ, but all methods attempt to mimic cyclic wetting/ drying and flushing of wastes. Kinetic tests are long-term (i.e. months to years) weathering tests conducted to aid prediction of drainage quality from mine wastes. Kinetic tests are generally performed after static testwork, with samples classified as acid forming prioritised to allow for measurement of the lag time to, and longevity of, acid formation. Samples classified as uncertain can also be tested to define their behaviour, allowing for a definite classification to be assigned. The most common kinetic tests are laboratory-based columns, humidity cells and field-based test pads (Sapsford et al., 2008). Lesser used procedures include Soxhlet, BCRC and shake flask/batch reactor tests, the latter two of which examine impacts of mine waste bacteria on acid formation (Sobek et al., 1982; Bradham and Caruccio, 1999; White et al., 1999; Price, 2009).

7.1. Humidity cell tests

The American Society for Testing and Materials (ASTM) D5744 and the Minnesota Department of Natural Resources (MN DNR) protocols are the most widely used humidity cell procedures (e.g., Lapakko et al., 2006; Sapsford et al., 2008; González-Sandoval et al., 2009). These tests are designed to accelerate weathering and model rates of oxidation in the field as well as the chemical release (Lapakko and Antonson, 2006). A standard humidity cell is operated on a weekly cycle that comprised of three days of dry air, three days of moist air and a rinse with distilled water on the seventh day. A limitation of humidity cell tests is that they are conducted using a single, weekly flushing of a constant nature. However, precipitation at an actual mine site may vary significantly during wet and dry seasons. Additionally, humidity cell tests use a sufficiently large volume of water with reaction products expected to be removed. Removal of all reaction products may not occur in the field due to much lower water to rock ratios, incomplete flushing, and the occurrence of chemical reactions along the flow-path. Furthermore, equilibrium conditions expected in waste rock piles may not be realised in the test due to short contact times between the test solid and solution. Most recently, Erguler et al. (2014) evaluated the physical effects of column dimension and particle size on leachate chemistry in humidity cell tests. They reported that the lag time to ARD formation increased with length, diameter and volume of columns for both fine grained and coarse grained particle sizes. Considering this, data used from such tests only allow for conclusions on likely lag times to acid formation, and on relative abundances and approximate magnitudes of environmentally significant elements and compounds in future mine waters.

7.2. Column leach tests

Column tests monitor leachate quality over time by cyclic (weekly or monthly) sampling. Data is gathered on the sulphide
reactivity, oxidation kinetics, metal solubility and the leaching behaviour of the test materials. The AMIRA P387A ARD Handbook (Smart et al., 2002) free draining column leach test methodology is the most commonly used in Asia-Pacific (e.g., Stewart, 2005; Miller et al., 2010; Munksgaard and Lottermoser, 2011; Parbhakar-Fox et al., 2013a). An advantage of this protocol is that it allows flexibility in the experimental design and therefore can be modified with respect to grain size, sample mass and frequency of leachate collection. A limitation of this testing protocol is the lack of integrated mineralogical and microtextural assessments for the duration of the experiment. Hence, Parbhakar-Fox et al. (2013a) proposed that samples of leached material should be taken during testing for bulk mineralogical and geochemical characterisation. By collecting these additional data, trends in leachate data can be related to mineralogical and geochemical changes (Parbhakar-Fox et al., 2013a). Such an integrated approach would yield greater insights into the weathering behaviour of geoenvironmental units.

7.3. Field test methods

Field based tests are small to large scale trials using tubs, barrels or test piles (e.g., Andrina et al., 2006; Smith et al., 2009). Large scale piles require installation of instrumentation including thermistors and lysimeters to monitor the in-pile geochemical conditions. Leachates collected from these piles are compared against laboratory based geochemical tests to check their accuracy. The benefit of field tests is that mine wastes are exposed to site-relevant climatic and microbiological conditions; therefore, field tests are the most accurate method for testing the long-term behaviour of mine wastes. However, a significant limitation apart from being costly and time-intensive is that they cannot be established early enough to aid with waste planning at the pre-feasibility or feasibility stages of mine operation. Consequently, the greatest value of field tests is in assessing potential capping and remediation methods of waste storage facilities upon mine closure (e.g., Andrina et al., 2006; Mauric and Lottermoser, 2011).

8. Mineralogical characterisation

Routine mineralogical characterisation relies on drill core logging, optical microscopy and X-ray diffractometry (XRD). This allows for characterisation of the acid forming and neutralising minerals, and for their bulk contents to be measured. Additional mineralogical tools include scanning electron microscopy (SEM; Chappell and Craw, 2002; Weisener and Weber, 2010; Parbhakar-Fox et al., 2013a; Parbhakar-Fox et al., 2014), electron microprobe analysis (EPMA; Weisner et al., 2003; Hudson-Edwards and Edwards, 2005), laser ablation inductively coupled plasma mass spectrometry (LA-ICPMS; Al et al., 2007; Ohlander et al., 2007; Parbhakar-Fox et al., 2013a; Parbhakar-Fox et al., 2014), micro-particle induced X-ray emission (micro-PIXE, Cabri and Campbell, 1999; Belcher et al., 2004; Jamieson et al., 2005), and micro-XRF (Adams et al., 1998; Katsuta et al., 2007). These relatively high-cost per analysis tools allow an understanding of mineral chemistry and mineral surfaces. Selection of the most appropriate tools for mineralogical characterisation is dependent on the objective of the study. For example, whether the bulk mineralogical composition is required for calculating the calcite:sulphide ratio (e.g., Paktunc, 1999), or whether the elemental composition of a specific mineral is required to understand its relative susceptibility to oxidation (e.g., Fe in sphalerite). In this context, the use of EPMA, LA-ICPMS, micro-PIXE and micro-XRF is restricted to sulphide minerals where compositional abnormalities affect ARD test work interpretation (e.g. Mills et al., 2011). Mineralogical tools commonly used in ARD prediction are summarised the following sections.

8.1. Optical microscopy

ARD focussed optical mineralogy studies have developed the sulphide alteration index (SAI) for tailings classification (Blowes and Jambor, 1990; Gunsinger et al., 2006; Moncur et al., 2009). However, the application of the SAI is limited as it only examines the degree of alteration but not any other physical parameters which might influence oxidation. Instead, microscopy studies should document and quantify all parameters that control ARD formation using appropriate indices (e.g., ARDI, Parbhakar-Fox et al., 2011).

8.2. Bulk mineralogy

Bulk mineralogy of samples can be assessed by XRD methods, yet poorly crystalline materials such as iron and aluminium phases commonly associated with ARD cannot be readily identified unless the Rietveld Method is used (Lapakko, 2002; Jambor, 2003; Raudsepp and Paní, 2003). Estimates of the modal mineralogical and sulphide alteration index can be gained using both XRD and whole-rock geochemical data (Paktunc, 2001; Hermann and Berry, 2002; Posch and Kurz, 2007; Weber et al., 2005; McLemore et al., 2009). However, the limitation of such calculated mineralogies is that inaccurate results will arise if bulk XRD analyses cannot resolve the nature and abundance of clay minerals. In this case, short-wave infrared spectrometry or special sample treatment prior to XRD analysis will allow accurate phase identification and quantification.

8.3. Automated mineralogy

Tools such as the mineral liberation analyser (MLA) and Quantitative Evaluation of Minerals by SCANning electron microscopy (QEMSCAN) are automated instruments that uniquely combine back scattered electron (BSE) image analysis, X-ray mineral identification and advanced imaging and pattern recognition analysis (Gu, 2003; Fandrich et al., 2007). Applications of MLA and QEMSCAN in applied mineralogy and metallurgical processing are well established (e.g., Bruckard et al., 2010; Chapman et al., 2011; Hunt et al., 2011; Rizmanoski, 2011). However, there are comparatively few published examples of automated mineralogy application in predictive ARD studies (e.g., Goodall, 2008; Aranda et al., 2009; Parbhakar-Fox et al., 2011; Parbhakar-Fox and Lottermoser, 2014). The primary applications of this tool have been to collect modal mineralogy data through the X-ray modal mineralogy (XMOD) point count method, and to characterise individual sulphide particles (e.g., tailings feed using SPL-lite data) in terms of their size, shape and mineral associations. However, automated mineralogical analyses are time-consuming and costly procedures that focus on minute sample masses, and hence a significant number of analyses need to be completed to resolve the need for ARD prediction in a large-scale industrial setting.

9. Waste classification

Samples are classified in terms of their acid forming potential following extensive geochemical, static, kinetic and mineralogical characterisation (as described in the previous sections). Firstly, classification can be based purely on sulphide and calcite contents (Paktunc, 1999). However, more commonly classifications are based on static geochemical results. NAPP, NAG, NAG pH, NPR and sulphur values are used to classify samples, using carefully selected cut-off values for the discrimination of acid forming and
acid neutralising materials. Presently, only the waste classification terms potentially acid forming (PAF), non-acid forming (NAF) and uncertain (UC) waste are favoured (cf. Smart et al., 2002). Generally, if a site contains PAF material, then kinetic test data must be acquired, and ARD management practices need to be established (Lottermoser, 2010). Lottermoser (2010) also emphasises that samples classified as NAF but with high sulphur content (>1%) must be evaluated to deduce if they host soluble secondary sulphate minerals such as gypsum. Furthermore, drainage from these samples may be neutral to alkaline but exceptionally saline, thereby exceeding water quality guidelines for sulphate. In addition, neutral to alkaline drainage waters may carry exceptionally high contents of metals such as zinc, molybdenum or cadmium and metalloids such as arsenic, antimony or selenium (Lottermoser, 2010). Clearly, current waste classifications based on static test data are not only limited by the quality of test data itself, but also by the fact that such simple classification schemes do not consider waste types that could lead to neutral metalliferous or saline drainage.

10. Future of ARD prediction

10.1. Existing approaches

Whilst tests and protocols for ARD prediction are well established, there are a series of limitations with individual tests and methods (Table 1). Moreover, the mining industry largely relies on chemical static and kinetic testing for ARD risk assessment. As stated, the most widely used predictive protocol is the wheel...
New and improved approaches for ARD prediction need to consider these missing parameters. For example, the Geochemistry-Mineralogy-Texture or GMT approach is a new ARD prediction and waste classification tool (Parbhakar-Fox et al., 2011). It considers that ARD reactions are controlled by and can be measured using integrated mineralogical, geochemical and textual analyses. The prediction methodology is pursued in stages. As one progresses, the number of samples decreases and the analytical sophistication increases. This new protocol predicts waste types with greater certainty, because of its reliance on multiple scientific tools. It creates more waste knowledge and does so at a lower cost than conventional approaches.

Accurate ARD prediction still represents a challenging concept because of the multifaceted mineralogical, geochemical and microbiological processes leading to ARD. ARD risk assessments need to consider predictive geochemical test data, but also mineralogical, textual and geometallurgical rock properties (Fig. 2; Parbhakar-Fox et al., 2013b). Moreover, a staged approach to ARD risk assessment and associated screening of different sample types allows a more cost-effective identification of ARD risks associated with specific rock types (Parbhakar-Fox et al., 2013b). Such a new architecture of integrative ARD testing should rely on quantitative measurements and integration of scientific tools and data, backed up by field instruments and state-of-the-art laboratories. The major advantage of staged approaches like the GMT approach is that through screening analyses (e.g., paste pH, ARDI, total sulphur), it is more financially viable to analyse best practice number of samples (cf. Price, 2009), for deposit-wide waste characterisation.

To further facilitate deposit-wide characterisation and to add value to already existing datasets, proxies for ARD tests and data must be identified. Geometallurgical tools show promise to aid in ARD prediction. Hypospectral mineralogy instruments (i.e., HyLoggerTM, Corescan) provide a relative estimate of the modal mineralogy. Its advantages over other mineral identification techniques (e.g., XRD and automated SEM instruments including MLA and QEMSCAN) are its rapidity (up to 1000 m of core per day), its low cost per sample, and its non-contact, non-destructive approach (Huntington et al., 2006). Its current limitation is its non-quantitative nature, however, with targeted supplementary XRD data; it can be transformed to provide semi-quantitative data. Hypospectral images combined with total sulphur values allow the recognition of ARD risk domains and the identification of carbonate minerals in drill core (Parbhakar-Fox et al., 2013b; Parbhakar-Fox and Lottermoser, 2014). Also, the application of Equotip on intact drill core materials can be used with total sulphur values to predict ARD risks (Parbhakar-Fox et al., 2013b; Parbhakar-Fox and Lottermoser, 2014).

10.2. A new architecture for ARD prediction

The future of ARD prediction lies in re-inventing the wheel of Morin and Hutt (1998), with one such option proposed in Fig. 3. First, geoenvironmental ore-deposit models should be consulted at the exploration stage to understand potential outcomes and anticipated environmental consequences of mining for the deposit in question. Through consultation of these models, forecasting of potential rehabilitation requirements is permitted (Lottermoser and Ashley, 2012). Next, the number of samples required for ARD testwork must be determined at the exploration stage so sufficient funds to support this can be apportioned. Sampling guidelines discussed in Section 4 should be followed, with consideration also given to the sampling curve shown in Downing (1999) which advises sample number based on deposit mass (Mt). Pre-screening testwork as per stage-one of the geochemistry-mineralogy-texture-geometallurgy (GMTG) approach (Fig. 2) should next be followed. Paste pH tests should be performed in batches using the same pulverised materials as those used for total sulphur measurements. This will provide basic pre-screening data to help identify samples most likely to require detailed geochemical ARD testwork. For the purpose of ARD domaining at early life-of-mine stages,
calculating maximum potential acidity (MPA) from $S_{\text{Total}}$ is an appropriately conservative risk approach, minimising mis-classification of PAF (potentially acid-forming) samples as NAF (non-acid forming) materials (Downing and Giroux, 1993; Borden, 2003). Disadvantageously, the turn-around time with off-site analyses is likely a number of days rather than hours as would be required to support best practice sampling. Therefore, to speed up the measurement of sulphur, industry should invest in benchtop element analysers that offer rapid and accurate data collection (<1 min/sample) from a powdered sample, with minimal data processing, making them ideal for use in an on-site field laboratory. In addition, if corresponding geometallurgical data (e.g., Equotip, hyperspectral mineralogy, MLA/QEMSCAN point counting/modal mineralogy) exist, such information can also be used in the domaining process (cf. Parbhakar-Fox and Lottermoser, 2014). Together with geoenvironmental modelling as well as $S_{\text{Total}}$ and paste pH values, geoenvironmental units should be identified and ARD risk-domains of the deposit should be established. Based on such screening, samples identified as low-risk (e.g., $S < 0.05 \, \text{wt.\%}$; $pH > 5.6$; Hutt and Morin, 1999; Lottermoser, 2010) and low metals content relative to the local baseline, should not be subjected to any further analyses. All other geoenvironmental units require a set of representative samples for further ARD testwork following an integrated GMTG approach (e.g., Stage two onwards).

At the screening stage, a simple microbiological test must be performed, but is currently not included in the GMTG approach. The BAPP test is considered a standard procedure (ASTM E1915-11), however, the development of simple and quicker tests is required, with the biokinetic test (Hesketh et al., 2010) potentially representing one such option. Systematic bulk-mineralogical testwork must be performed, and is now a possibility with the advent of benchtop XRD instruments. They allow for rapid collection of semi-quantitative data which can be used for mineralogical ABA calculations; however, caution is to be exercised with regards to data processing, as this requires technical expertise. Detailed mineralogical testwork using other analytical tools including SEM, LA-ICPMS and MLA should be considered to understand controls on ARD formation on a select number of representative samples (Stage three, GMTG approach, Fig. 2). This type of integrated mineralogical analyses should be adopted during kinetic testing (cf. Parbhakar-Fox et al., 2013a). Representative samples of each geoenvironmental unit should be tested because, even non-acid forming materials can pose an environmental risk i.e., neutral or saline mine drainage with high associated metalloid (e.g., As, Sb) contents. Such mineralogical and mineral-chemistry data must be effectively used when performing computer-based ARD predictive modelling which must be the next step, otherwise it collection is meaningless. Geochemical modelling using software packages such as Geochemists Workbench, PHREEQC and MINTEQ should be undertaken to allow for prediction of water–rock interactions, and an understanding of long-term mineralogical transformations (Maest and Kuipers, 2005; Martin et al., 2005; Lengke et al., 2010). The use of machine learning (e.g., Bietri et al., 2014), fuzzy logic and data mining (e.g., Aroba et al., 2007) computed methods for predicting ARD chemistry should also be explored.

Based on these data, waste classification should be performed following the criteria given in the GMT/GMTG approaches (Parbhakar-Fox et al., 2011, 2013b). However, additional criteria must be set to include an evaluation of microbiological testwork, once an appropriate rapid methodology has been established as standard practice. Ultimately, a mine site must develop an ARD block model and waste management strategy to the satisfaction of regulators to allow for mining to proceed to the next stages. Whilst a relatively large number of samples should have been selected and analysed as part of the pre-screening stage, their ARD properties should be entered into the mine-site geological database, whereby geostatistical techniques and transfer functions should be applied (e.g., principal component analysis, GARD Guide, 2014; kriging; Rossi, 2006; Kentwell et al., 2012). Only through application of these processes will the ARD testwork data be appropriately used, resulting in the development of robust ARD models and waste management strategies.

11. Conclusions

Life-of-mine planning requires a solid understanding of future ARD risks. Presently, the mining industry predominately relies upon static geochemical test methods established in the 1970s to predict ARD. Limitations associated with these tests include the potential for over- or under-estimation of acid forming potential through the use of over-aggressive tests (e.g., Sobek test, sequential NAG), and calculation of NAPP using total-sulphur as opposed to sulphide–sulphur. Furthermore, a lack of consideration is given to other ARD forming minerals (e.g., jarosite) and indeed, long-term silicate neutralisers (e.g., chlorite). Fundamental controls on ARD formation such as mineralogy, texture and microbiological processes are also not appropriately characterised in established predictive protocols such as the wheel of Morin and Hutt (1998). Despite these short-comings, the tests, and these protocols remain widely used.

Best practice for ARD prediction should be performed at the earliest stage of mineral resource development and, ARD predictive tests and methodologies should rely on integrated and staged field and laboratory measurements using mineralogical, geochemical, textural and geometallurgical tools. In future, microbiological and bioaccessibility tests should be incorporated into the evolving ARD prediction protocol. Furthermore, the effective use of computing tools such as data mining and machine learning must be established. Research efforts should also focus on the evaluation and validation of downhole drill core scanning technologies, with the use of sensors (e.g., laser induced breakdown spectroscopy; prompt gamma neutron activation analysis) for sulphide identification and element quantification. Such data should be combined with hyperspectral data to create a fully integrated element and mineral map, from which ARD simulations can be run.

While the research community needs to establish new practical state-of-the-art ARD prediction tools and approaches, industry has to accept and use these, if we are to achieve more cost-effective mine closure and reduce environmental liabilities in the long term. Such progress also requires the application of predictive ARD tools at the beginning of the life-of-mine cycle. A more integrated and staged approach to early ARD prediction supports more effective management and valuation during operation, and ultimately less costly mine closure outcomes. Finally, better ARD prediction practices will only be achieved through education of current and future mining practitioners.

Acknowledgements

The authors would like to acknowledge the support of CRC ORE (Project 20090048-10-P1B), established and supported by the Australian Government’s Cooperative Research Centres Programme. Dr. Taryn Noble is thanked for providing access to paste pH data. Two anonymous reviewers are thanked for their comments on an earlier manuscript draft.

References


Glossary

ABA: acid base accounting
ABCC: acid-buffering characteristic curve
ANC: acid neutralising capacity
AP: acid potential
ARD: acid rock drainage
ARDI: acid rock drainage index
BAPP: biological acid producing potential
BSE: back scattered electron
CRM: certified reference material
CRS: chromium reducible sulphur
EPMA: electron probe microanalysis
GMT: geochemistry-mineralogy-texture approach
LA-ICPMS: laser ablation inductively coupled plasma mass spectrometry
MIA: mineral liberation analyser
MPA: maximum potential acidity
NAF: non-acid forming
NAG: net acid generation
NAPP: net acid producing potential
NCV: net carbonate value
NNP: net neutralising potential
NP: neutralising potential
NPR: neutralising potential ratio
PAF: potentially acid forming
PIXE: particle induced X-ray emission
pXRF: field portable X-ray fluorescence
QEMSCAN: Quantitative Evaluation of Minerals by SCANning electron microscopy
SAI: sulphide alteration index
SPLP: synthetic precipitation leach procedure
UC: uncertain
XMOD: X-ray modal analysis
XRD: X-ray diffractometry
XRF: X-ray fluorescence