European Journal of Soil Science

European Journal of Soil Science, November 2014, **65**, 897–906 doi: 10.1111/ejss.12160

Mapping contaminated soils: using remotely-sensed hyperspectral data to predict pH

C. C. H. Ong & T. J. Cudahy
Earth Science and Resource Engineering, CSIRO, 26 Dick Perry Ave, Perth, Western Australia 6151, Australia

Summary

This study assessed the feasibility of remote mapping and, thus, monitoring of soils contaminated by acid mine drainage. We report on the use of laboratory and airborne spectroscopy to determine pH. Reflectance spectra were obtained for rock and soil samples collected at our test site, the abandoned Brukunga Pyrite Mine in South Australia, using a laboratory-based Analytical Spectral Devices Inc. (ASD) spectroradiometer. A partial least squares (PLS) regression was used to develop a predictive equation for pH based on the reflectance spectra. The validation results indicated that it is possible to generate satisfactory predictions of pH from spectral data, as demonstrated by the ratio of performance to deviation (RPD) of 1.53, a relatively small root mean square error of prediction (RMSE) of 0.91 and an R^2 value of 0.58. Evaluation of the predictive equation indicated that it depends on diagnostic spectral features related to secondary iron minerals resulting from acid mine drainage (AMD). The presence of these minerals was validated independently using X-ray diffraction (XRD). The predictive equation was applied to airborne hyperspectral data, using a form of remote sensing that simultaneously acquires spatially co-registered images in many spectrally contiguous bands (> 50). Hyperspectral data acquired by the HyMap sensor between 1998 and 2001 were used to produce multi-temporal pH maps. Despite the inaccuracy of the global positioning system (GPS), locations of the validation samples and geographical inaccuracy of the airborne imagery, validation of the maps indicated that pH can be generated reliably from airborne hyperspectral data, as indicated by the relatively small RMSE of 0.57 and the R^2 value of 0.72. These maps demonstrate the potential to provide environmental practitioners, including soil scientists, with a spatially comprehensive view of pH related to AMD conditions. This has implications for the application of remotely sensed hyperspectral data for monitoring soil pH related to AMD conditions, especially in the near future when such data will be available from satellite sensors such EnMap (Stuffler *et al.*, 2009). Some examples of their use include a better understanding of the progress of restoration efforts and or to pinpoint areas where future efforts should be concentrated and to evaluate the extent of downstream impacts, including contamination of soil. More generically, soil pH data are routinely required by soil scientists as part of the suite of data for understanding soil characteristics and the ability to provide spatially comprehensive soil pH data, as demonstrated by this study, would be valuable.

Introduction

The oxidation of sulphidic mine wastes, and the consequential release of acid mine drainage (AMD), is one of the main strategic environmental issues facing the mining industry worldwide. The total cost of managing these contaminated soils from current mines over a 15-year period in Australia alone was estimated in 1997 to be US $670 million (Harries, 1997).

Quantification of the spatial extent of AMD-affected areas, their associated contamination levels and risk potential is important, especially in cases of accidental spills, which may cause large-scale contamination of soils. This was the case in the Aznacollar accident (Grimalt *et al.*, 1999). Hyperspectral imaging is a form of remote sensing defined as the simultaneous acquisition of spatially co-registered images in many spectrally contiguous bands (> 50), measured in calibrated radiance units from a remotely operated platform (Schaepman *et al.*, 2006; Goetz, 2009). Several studies have reported the value of hyperspectral imaging for characterizing AMD conditions and the resultant effects on soils (López-Pamo *et al.*, 1999; Swayze *et al.*, 2000; Mars & Crowley, 2003). These include producing maps of minerals using airborne hyperspectral data for a single date, and some studies based on laboratory spectral measurements (Chang *et al.*, 2001; Kemper & Sommer, 2002; Viscarra Rossel *et al.*, 2006). López-Pamo *et al.* (1999) demonstrated the value of the technology for quantifying the spatial extent of an

Correspondence: C. C. H. Ong. E-mail: cindy.ong@csiro.au
Received 12 May 2014; revised version accepted 12 May 2014

AMD spill and its effects. There are, however, no published results on the use of hyperspectral data for multi-temporal monitoring of AMD-related contaminated soils. Furthermore, for such data to be useful for environmental practitioners, including soil scientists, they must be translated into environmental measurements. For monitoring AMD, pH is important, but there is a lack of published data that relate mineralogical mapping to pH.

Previous studies found that the secondary minerals generated as a result of AMD varied with pH (Bigham *et al.*, 1996a). Hence, mineralogy may be used as a proxy to map pH, and this study investigated the relationship between spectra and pH, with the aim of producing multi-temporal maps of pH with remotely-sensed hyperspectral data. It is hypothesized that these maps could be potentially useful for soil scientists involved in AMD for quantifying the spatial extents of the contaminated soils, assessing the severity and potential risks at the baseline stage and monitoring the performance of remediation efforts. Generically, soil pH is an important property that is required routinely in soil studies. Therefore, the ability to provide spatially comprehensive soil pH maps using hyperspectral imagery would be of value.

Background

Acid mine drainage and its impacts on the Australian mining industry

Acid mine drainage is commonly defined as seepage with a pH < 5, originating from a tailings pile, waste rock pile or exposed sulphide-rich rock (Shaw *et al.*, 1998). Other terms used for this condition are acid drainage (AD) and acid rock drainage (ARD). The oxidation process may be induced by a variety of geochemical agents at pH values below 4.5, but is usually associated with Fe-oxidizing bacteria (*Thiobacilli*), the most common of which is *Thiobacillus Ferrooxidans* (Bigham *et al.*, 1996a).

Acidic waters formed by sulphide oxidation can precipitate a large suite of relatively soluble and insoluble iron (Fe)-bearing secondary minerals whose speciation is controlled by pH, degree of oxidation, moisture content and solution composition (Swayze *et al.*, 2000). With a prolonged drying period a suite of efflorescent salts is produced, which when dissolved, rapidly decreases the pH of acid waters and can result in short-term catastrophic effects on metal loadings in aquatic ecosystems (Hammarstrom *et al.*, 2005).

The oxidation of sulphidic mine wastes and the consequent release of AMD is one of the main strategic environmental issues facing the mining industry in Australia (and world-wide) (Harries, 1997). Results of a survey conducted from 317 operational and historic mine sites in Australia suggest that approximately 54 sites are managing significant amounts of potentially acid-generating wastes. More than 10% of the waste is potentially acid generating and is equivalent to > 10 million tonnes (Mt) of potentially acid-generating waste. A further 62 additional sites in Australia are managing some potentially acid-generating waste equivalent to less than 10% of the total waste (< 10 Mt) (Harries, 1997).

The additional cost of managing potentially acid-generating waste at operating mine sites for the Australian mining industry was last estimated at about US $45 million per year (Harries, 1997). This includes the costs of cover installation, selective placement of wastes, additional waste characterization and water treatment.

Studies conducted by Bigham *et al.* (1996a,1996b) found that the secondary minerals generated as a result of acid mine drainage vary with pH. Precipitates formed at pH 6.5 or higher are predominately ferrihydrite and or goethite. Schwertmannite was a major component of the precipitates at pH 2.6–4.5, with minor amounts of associated jarosite and goethite. Furthermore, jarosite increases in abundance with decreasing pH below pH 3 and goethite appears between pH 3.3 and 3.6. These secondary minerals are Fe-rich and many of them are hydroxyl- and or water-bearing, making it possible to identify them on the basis of their diagnostic spectral reflectance signatures (Swayze *et al.*, 1996). Because of the possibility of identifying these minerals with spectral methods and because they are related to pH, the feasibility of producing predictions of pH from laboratory-measured spectra was tested. The results were then applied to airborne hyperspectral data to demonstrate the ability to provide multi-temporal maps of pH using remotely-sensed hyperspectral data.

Imaging spectrometry/hyperspectral imaging

Hyperspectral sensors that are currently available commercially measure reflected light in the 400–2500 nm (visible to short-wave infra-red (VNIR to SWIR)) range. Swaths of discrete pixels of spectra are recorded along a flight path over the landscape, allowing images with pixel sizes of $3 \times 3\,m^2$ (using the HyMap sensor) to be obtained. Recent advances in the technology have led to the development of airborne hyperspectral sensors that are of comparable quality to laboratory spectrometers. These airborne instruments are radiometrically calibrated to National Institute of Standards and Technology (NIST) (2014) certifications and, therefore, the data recorded are quantitative, repeatable and traceable. Consequently, airborne hyperspectral sensors may be used in a quantitative manner similar to laboratory spectrometers for the measurement of materials such as paint (Newman, 1979) and grain (Finney *et al.*, 1987; Shenk *et al.*, 1992) for quality analysis.

Test site

The test site for this research was the abandoned Brukunga Pyrite Mine in the Mount Lofty Ranges, 4 km north of Nairne and 40 km east-south-east of Adelaide, the capital of South Australia (see bottom right of Figure 1). The AMD pollution from the Brukunga Pyrite Mine originates from leaching of sulphide minerals contained in the overburden of waste-rock dumps and tailings dumps that were discarded during mining and milling. The location of the waste-rock dumps and tailings dump are marked on the photograph in Figure 1. The Brukunga mine site is classified, together with Rum Jungle, Mt Lyell, Captains Flat and Mt Morgan, as a major source of AMD in Australia (Taylor & Cox, 2003).

The climate at Brukunga is essentially Mediterranean, with cool wet winters (average temperatures between 5 and 14°C,

Figure 1 Location of Brukunga Mine and key site features (image reproduced from Cox *et al.* (2006)).

average rainfall of 103 mm) and warm-hot dry summers (average temperatures between 11 and 26°C, average rainfall of 27 mm). The average rainfall is 766 mm per annum and pan evaporation is 1100–1400 mm per annum.

The immediate area of the mine site was described as containing a significant amount of outcropping to partial outcropping of exposed bedrock with only a thin (< 1 m) veneer of soil cover (Nandy *et al.*, 2011). Additionally, three major soil types were identified and include: (i) dark brown-red coloured, iron-rich, muscovite-bearing, *in-situ* soils formed directly above the pyritic schists, (ii) pale-light brown coloured *in-situ* soils formed directly over the quartz-mica-rich weathered sediments and (iii) colluvial deposits, which have filled the incised valleys, which begin from the break in the steep slopes throughout the area (Nandy *et al.*, 2011). The area surrounding the mine has been classified according to the Australian Soil Classification System (Isbell, 2002) as predominantly Chromosols with Tenosols to the west and Sodosols to the south (Johnston *et al.*, 2003).

Geologically, the host rock of the pyrite ore is the Talisker Calc-siltstone of the Cambrian Lower Kanmantoo Group. This consists of a metamorphosed phyllite with calc-phyllite interbeds (Gravestock & Gatehouse, 1995). The Nairne Pyrite Member from which the pyrite was mined is in the lower part of the Talisker Calc-siltstone formation. The Nairne Pyrite Member also hosts several small-scale base-metal deposits.

The iron-sulphide mineralization occurred as three steeply-east-dipping, conformable lenses, separated by waste beds. The mine benches are shown in Figure 1. Each of the 15–30-m thick ore zones consisted of iron sulphide-bearing muscovite schists and gneisses with minor lenses of calc-silicate and quartz plagioclase metasediments. The waste zones consisted of quartz plagioclase granofels and minor calc-silicate granofels, muscovite schists and gneisses. The sulphide lenses outcropped as ferruginous gossans, with weathering to a depth of 18 m (LaGanza, 1959). The mineralization was pyrite and pyrrhotite, with minor sphalerite, chalcopyrite, galena and arsenopyrite, which are sources of arsenic, cadmium, copper, zinc, antimony, lead, nickel, tin, barium, cobalt, manganese and sulphate (Burtt & Gum, 2000a,b). Pyrite and pyrrhotite, when exposed to air and water, oxidize and break down to generate AMD. The acid conditions enable the various heavy metals, which are potentially harmful to the environment and the wellbeing of inhabitants of the area, to dissolve. The resultant seepage enters nearby waterways, contaminating the flow as it passes through the mine site.

Mining was undertaken by quarrying into the eastern flank of a hillside that forms the western side of a valley traversed by Dawesley Creek, a tributary of the Bremer River. The current layout of the mine site is shown in Figure 1. The open pit is oriented N to S and is approximately 1800 m long and 150 m wide. It consists of an exposed foot-wall and main bench with two deep slots excavated to access deeper ore (left-hand side of Figure 1). Waste rock was placed in two large dumps on the western side of Dawesley Creek and a smaller dump on the eastern side. Material containing more than 5% sulphide was regarded as ore and processed on-site. Tailings were deposited in a valley-fill storage facility on the eastern side of Dawesley Creek (Taylor & Cox, 2003).

Materials and methods

Airborne imaging spectrometer

Airborne hyperspectral data were acquired at Brukunga at 3 m resolution on 3 April 1998, and, 5 m resolution on 9 April 1999, 21 March 2000 and 14 March 2001 using the Hyperspectral Mapper (HyMap; Integrated Spectronics Pty. Ltd., Sydney, Australia) instrument (Cocks *et al.*, 1998). The instrument measures 126 bands in the 450–2500 nm wavelength region, with bandwidths between 15 and 20 nm. This allows contiguous spectral coverage except in the atmospheric water vapour bands. Bands close to the latter regions, specifically around 1400 and 1900 nm, were not used. The total number of remaining bands used was 112.

Processing of the airborne data included: (i) atmospheric correction to surface reflectances using the MODTRAN® radiative transfer algorithm (Berk *et al.*, 1989); (ii) compensation for residual atmospheric artifacts unaccounted for by the atmospheric correction and instrument artifacts using gains generated from ground spectral measurements obtained at invariant spectral calibration targets that were also captured on the imagery; (iii) application of pixel-mean normalization (PMN) (Green & Craig, 1985) to remove any wavelength-independent multiplicative effects in the airborne data that may be significant because of the topographic relief of the test site and bi-directional reflectance difference factor (BRDF) effects (the assumption of multiplicative and wavelength independence is warranted as the surfaces to be mapped for acid drainage are mostly denuded of vegetation and also remaining small

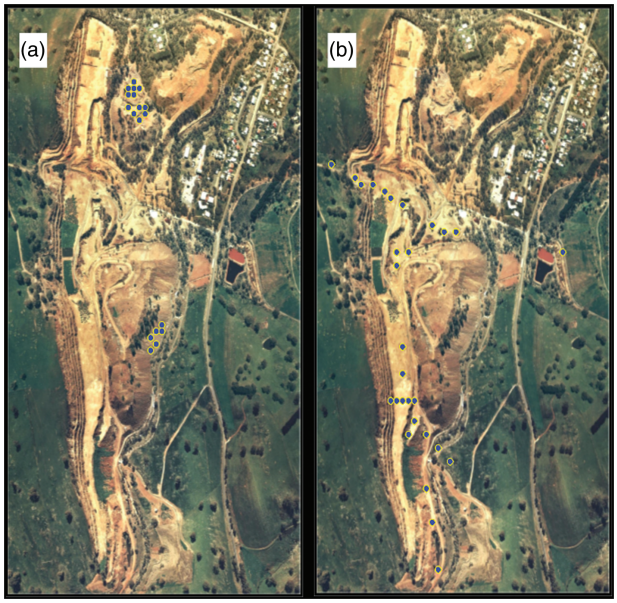

Figure 2 (a) Location of samples collected during the 1999 field campaign. (b) Location of samples collected during the 2001 field campaign.

amounts of vegetation were masked and removed from the imagery as described below); (iv) geo-rectification of the airborne imagery using the geometric lookup table (GLT) provided by the service provider who acquired the data (the positioning error is usually less than three pixels); and (v) masking of areas containing vegetation by setting a threshold based on the normalized difference vegetation index (NDVI) (Tucker, 1979). This is a well-known index used in the remote sensing community to characterize vegetation. Green vegetation usually ranges between NDVI values of 0.3 and −0.7 and so a threshold was set where all areas with NDVI values larger than 0.3 were masked.

Sample collection

A total of 59 samples were collected during two field campaigns in 1999 and 2001. The first field campaign was concurrent with an airborne acquisition in 1999. Samples were collected from 43 locations across the mine site and were selected on the basis of mineralogical maps produced from airborne hyperspectral data acquired in 1998. Samples, usually weighing less than 0.5 kg, were collected from the top 5 cm of the surface.

The second acquisition of 16 samples was concurrent with the airborne acquisition in 2001. Mineral maps and preliminary pH maps generated using the 1998, 1999 and 2000 airborne data guided the sample collection. The sizes of the samples were similar to those collected in 1999. The locations of the 1999 samples are plotted on Figure 2(a) and the 2001 samples are plotted on Figure 2(b).

Spectral laboratory measurements

The samples were characterized in the laboratory with an ASD FieldSpec® Pro FR spectroradiometer (ASD Inc.; Boulder, CO, USA) that records spectral data in the 350–2500 nm wavelength range using the bare fibre, giving a field of view (FOV) of 25°. The fibre was secured to a laboratory tripod so that the end was 10 cm vertically above the sample. The sampling area for this configuration is a circle of diameter 4.43 cm. A 1000-watt quartz-halogen, photoflood lamp was used to illuminate the samples. The lamp was placed at an angle of approximately 60° from vertical.

Spectral measurements were collected from both the rock and soil samples and the pulp sample (sample after grinding, drying and chemical analysis). The samples were placed in 11 16.5-cm-sized plastic containers for measurement. Care was taken to ensure that the pulp samples had even flat surfaces before spectral measurements were recorded. Multiple spectral measurements were acquired for each sample. The samples were rotated after each spectral measurement. A $20 \times 20\,cm^2$ 99% LabSphere Spectralon® (Labsphere, Inc.; North Sutton, NH, USA) white reference panel was recorded between samples and used to calibrate the sample measurements to produce reflectance.

Pixel-mean normalization was applied to the spectral data to account for illumination brightness effects related to differential lighting in the laboratory. Note that the same correction was also applied to the airborne hyperspectral data to ensure cross-calibration between the laboratory and airborne spectra.

X-ray diffraction analysis

As the prediction of pH is based on the spectral mineralogy, we used X-ray diffraction (XRD) analysis to validate independently the minerals on which the predictions depend. Only samples from the 1999 campaign were analysed by XRD and measurements were performed with a X'Pert instrument (Philips Electronics N.V., Almelo, the Netherlands). Qualitative identification of the relevant AMD minerals and their abundance was performed by identification of the presence or absence of the three major peaks (Table 1) from the XRD pattern. The XRD traces were matched to library traces of secondary AMD minerals expected to be present in the sample. The XPlot software (CSIRO, 2005) was used to perform this analysis. Furthermore, to facilitate more accurate comparisons the traces were calibrated before the comparisons. Forty of the 41 samples analysed contained quartz and, hence, these 40 XRD patterns were calibrated internally to The International Centre for Diffraction Data (ICDD: http://www.icdd.com/) with quartz as the standard.

There were some challenges with the identification of some of the minerals by XRD. For example, it was difficult to distinguish between jarosite and natrojarosite, especially in the presence of bassanite as the peaks of the three minerals overlap. Also, some of the AMD secondary minerals, such as schwertmannite and ferrihydrite, were in the class of 'poorly crystalline' or 'X-ray amorphous' minerals, with the XRD patterns expected to show broad peaks. Identification of these minerals was therefore difficult. Note also that the samples were dried at 100°C and, because the drying process dehydrated the gypsum to bassanite, only trace amounts of gypsum were detected in the samples even though some samples were collected at a gypsum pile.

Table 1 Minerals related to acid mine drainage and their three major peak locations in the X-ray diffraction (XRD) traces. Note that there are overlaps in peaks in XRD measurements for some minerals. For example, goethite, haematite, pyrite, gypsum and bassanite have overlapping peaks at around 2.70–2.80 and goethite, epsomite and halotrichite have overlapping peaks at around 4.20 to −4.30

	Peak location (d-spacing/nm)		
Mineral	Peak 1	Peak 2	Peak 3
Jarosite	3.080	3.110	5.080
Goethite	4.183	2.450	2.693
Schwertmannite	2.550	3.390	4.860
Ferrihydrite	2.500	2.210	1.960
Pyrite	2.700	1.633	2.421
Haematite	2.690	1.690	2.510
Gypsum	2.871	4.280	2.684
Bassanite	3.006	2.807	6.000
Natrojarosite	3.006	3.122	5.060
Muscovite	9.980	2.554	2.556
Melanterite	4.900	3.776	4.560
Rozenite	4.470	5.460	3.400
Epsomite	4.216	4.200	5.980
Copiapite	18.400	9.230	5.570
Halotrichite	4.810	4.300	3.500

The minerals identified by XRD were jarosite, goethite, haematite, melanterite, rozonite, muscovite, gypsum and bassanite. In the case of goethite and haematite, their identification was combined as there were overlaps in the XRD peak locations for these minerals and other minerals such as pyrite and epsomite. Melanterite and rozonite were combined as they have similar reflectance spectral features. Gypsum and bassanite also have similar spectral features and overlapping XRD peaks.

In summary, the samples consisted mainly of jarosite (38 out of 41) and muscovite (37 out of 41) samples, followed by goethite or haematite (25 out of 41 samples) and gypsum or bassanite (27 out of 41). There were few samples (4 of 39) of melanterite or rozenite.

Chemical analysis

Chemical analysis was performed by a commercial laboratory to determine the pH. Preparation of the samples included drying at a temperature of approximately 100°C and then milling in a LM1 pulverizer to a particle size of a nominal 90% passing through a 106-µm sieve. The pH was measured on the pulverized samples using a 1:5 ratio of sample to deionized water and analysed with a pH meter with an accuracy of 0.01 pH units.

Partial least squares analysis

The analysis of the laboratory spectral data was performed using partial least squares (PLS) (Geladi & Kowalski, 1986; Haaland & Thomas, 1988; Wold *et al.*, 2001), a multivariate data analysis technique designed to relate multiple response variables, to multiple explanatory variables. The PLS method has been adopted more recently in spectral laboratory studies to investigate the relationships between physicochemical properties and spectral data (McGovern *et al.*, 1999; Dury *et al.*, 2001; Kemper & Sommer, 2002; Polder *et al.*, 2004). We chose PLS as the analysis tool as it has the advantage of being able to analyse many variables efficiently that may be strongly correlated, as is the case for spectral bands. Partial least squares is also useful where the data contain substantial random noise, when the matrix of predictors (X) has more variables than observations (Y) (as in the case of the spectral bands and observations of pH) and when there is multicollinearity among X values (as for spectral bands). A caveat for using this technique is the underlying assumption that a linear relationship exists between X and Y, and by using PLS analysis we assume that the spectral data are linearly correlated to pH.

An important factor to consider when using PLS analysis is that the prediction will be constrained by the calibration or training dataset. Therefore, it is important that the calibration dataset is comprehensive enough to explain fully the condition that is being measured.

The PLS technique aims to identify the linear combination of X variables that best models the Y dependent variables. The linear combination of bands, called the 'final regression coefficient' (FRC) (sometimes called PLS regression coefficients or b-coefficients), provides further diagnostic information on the key wavelength regions for prediction (Haaland & Thomas, 1988). We used the information from the FRC to identify the underlying dominant spectral signatures to predict pH. That is, the FRC relates to the key spectral features, which can then be related to the different AMD minerals. The FRC values themselves cannot be interpreted easily, but the relative size within each model can indicate the following: the larger the absolute FRC values, either positive or negative, the more important the associated spectral regions. The FRC values are the factors or weights that are applied to each band of spectral data to enable predictions of the Y variable. In our case, we applied the FRC to the airborne hyperspectral data to generate maps of pH.

The PLS analysis undertaken to investigate the relationship between the laboratory spectral data and pH measurements used 55 samples out of a total of 59. Four of the samples were outliers, as they were not representative samples for the calibration suite of AMD minerals. Specifically, the outliers comprised a road surface, two precipitates and one gypsum sample from a waste pile at a nearby farm. The PLS used was implemented with a one-at-a-time cross-validation (Wold *et al.*, 2001). The diagnostic statistics used to evaluate the accuracy of the predictions were the root mean square error (RMSE) and the ratio of performance to deviation (RPD) (Williams, 1987). The equations used were:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{N} \frac{\left(\hat{y}_i - y_i\right) 2}{N}}, \tag{1}$$

and

$$\mathrm{RPD} = \frac{S_y}{\mathrm{RMSE}}, \tag{2}$$

where ŷ is the predicted value, y is the observed value, N is the number of data and S_y is the standard deviation of the observed data.

The RMSE is commonly used to describe the prediction capability of a model and the RPD has been adopted by many researchers in the soil community for evaluating the performance of the model. The interpretation of the RPD values varies (Nduwamungu *et al.*, 2009). We used the commonly adopted guidelines proposed by Chang *et al.* (2001). Specifically, RPD > 2 predicts well, 1.4 < RPD < 2 predicts moderately well and RPD < 1.4 indicates poor prediction.

Validation of airborne data

The validation of the pH maps was performed by using the classical technique of comparing the airborne mapping results with *in-situ* measurements collected concurrently with the airborne acquisitions. Twelve of 16 ground-truth samples were collected concurrently with the airborne data acquisition in 2001 and these were used for the validation. Four of the samples were considered to be outliers, as they were not representative of AMD samples. Most of the samples consisted of a single sample and two consisted of short transects where five samples were collected. The samples were sent to a commercial chemical laboratory for pH analysis. The laboratory used and the analysis undertaken were the same as those used for the calibration dataset collected in 1999. Geographical positioning system (GPS) locations for the ground-truth sites were acquired with a hand-held GPS unit. The ground-truth sites were located on the airborne data by using the recorded GPS position and visual location on the maps while in the field. An area of at least 5 × 5 pixels on the 2001 airborne data corresponding to the GPS position recorded for the validation sites was examined to extract the pH value mapped by the airborne hyperspectral data. These pH values were compared against the pH obtained from the laboratory analysis to provide validation of the pH map.

Results and discussion

Figure 3 shows the results from the PLS cross-validation analysis. It plots the measured values of pH against corresponding predicted values as determined from the PLS analysis. The results show a linear trend with an R^2 value of 0.58, indicating moderate agreement between measured and predicted values. The RPD of the model was 1.536, indicating that a fair or satisfactory prediction may be obtained. Further, the RMSE was also relatively small at 0.908.

Figure 4 shows the final regression coefficients (FRCs) plotted in red. To assist with the interpretation of these results, the FRC values were plotted together with some typical secondary minerals related to AMD (plotted in black, blue, green and cyan). Bands with large absolute FRC values are bands that can be interpreted as key spectral features that determine the relationship. In this analysis, the bands with a large contribution are 437, 582, 910, 1917 and 2260 nm (black broken vertical lines show their location). The relationship of these large FRC values to spectral attributes of the AMD-related secondary minerals can be attributed to the ferrous iron absorption band at 582 nm, which is a feature related to

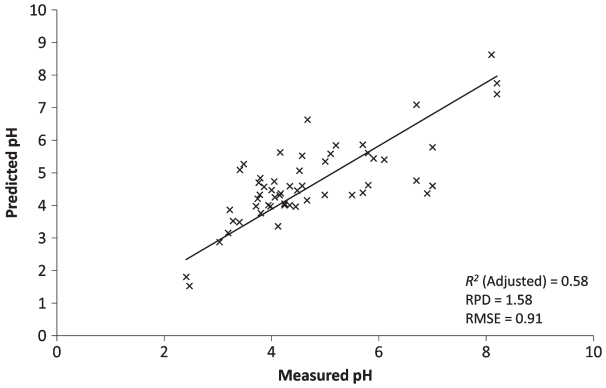

Figure 3 Scatter plot of chemically determined pH measurements against spectral-predicted pH measurements for the calibration dataset. The predicted values were obtained by applying the partial least squares (PLS) results to the spectral calibration dataset (laboratory Analytical Spectral Devices measurements).

charge transfer absorption of ferrous iron (Curtis, 1985). The FRC values also illustrate the importance of the 437 nm jarosite feature (Crowley *et al.*, 2003), which had the second largest coefficient. It is interesting to compare the contribution of this feature with that of the 2260 nm feature from O-H and Fe-O-H (Clark, 1999), which, although large, was much smaller than the contribution to the FRC from ferrous iron. Equally as important as the 437 nm feature is the large FRC value at 910 nm, which was related to a ferrous iron crystal field absorption feature (Curtis, 1985). Taken together, these results imply that the predictions of pH were related primarily to the iron oxide mineralogy. Independent validation of the mineralogy by XRD confirmed the presence of jarosite, goethite, haematite and gypsum or bassanite as the main minerals, which also further confirms the validity of the PLS spectral equation.

An important consideration is the large coefficients in the visible (VIS) region, especially below 550 nm, particularly when trying to apply these results to the airborne hyperspectral imagery. This region is complicated by aerosols, which are not usually properly accounted for in atmospheric correction of remote sensing data. In addition, the 1917 nm feature is in the water absorption region and this feature would not be detectable in the airborne data because of overwhelming absorption of the atmospheric water in this wavelength region.

Figure 5 shows a multi-temporal series of pH maps generated from the remotely-sensed airborne hyperspectral HyMap data with the application of the predictive equation (FRC) from the PLS analysis. Areas mapped with pH between 0 and 7 have a colour coding that shows the most acid areas mapped in 'hot' colours (red indicates pH of 0–2) and neutral areas in 'cool' colours (blue indicates pH 6–7). Areas outside these pH ranges have a black and white background image to provide a perspective of the area.

Figure 6 shows the result of the validation of the pH map for 2001, where measured pH from laboratory analysis of the ground-truth samples was plotted against the pH mapped using the airborne hyperspectral data. This graph shows a general linear trend and

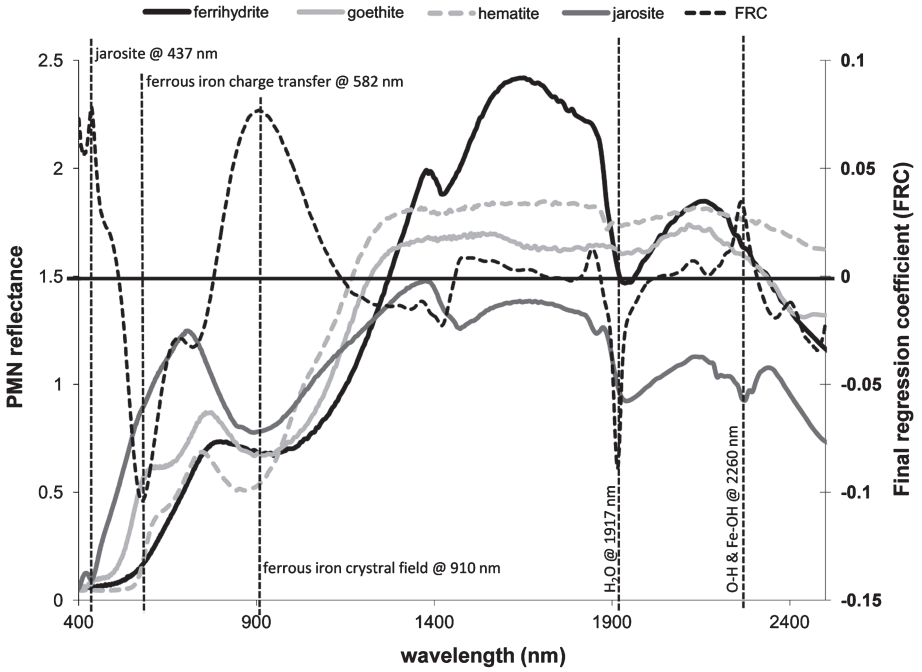

Figure 4 Final regression coefficient (FRC) results from partial least squares analysis of pH (red) and some typical acid mine drainage (AMD)-related secondary minerals from the the USGS library (Clark, 1999) (black, blue, green and cyan). The vertical black broken lines illustrate large coefficients and the physiochemical explanation for the features.

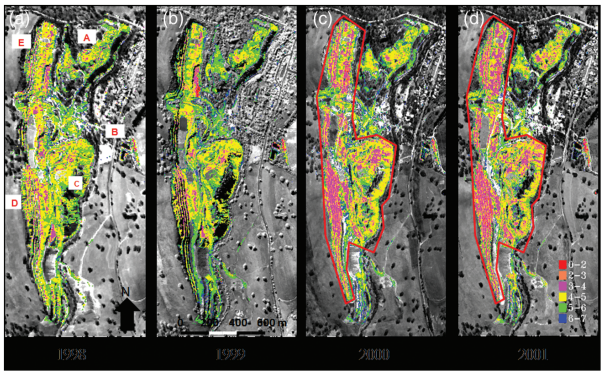

Figure 5 (a–d) pH maps generated from the application of a predictive algorithm produced from laboratory chemical and spectral measurements to multi-temporal airborne hyperspectral data collected over the Brukunga Mine using the HyMap sensor. (a) Indicates location of top waste rock dump, (b) the location of the acid treatment pond, (c) the middle waste rock dump, and (d) the south-western mine benches.

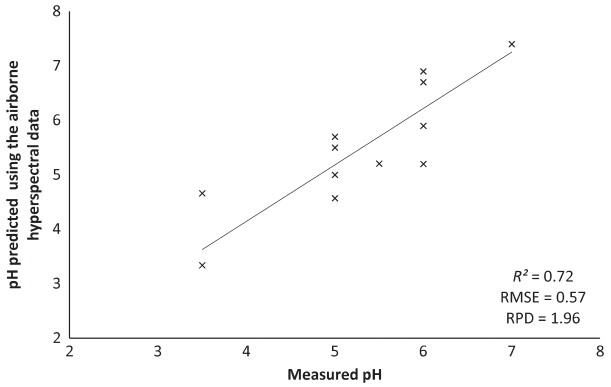

Figure 6 pH measured using chemical analysis plotted against pH values predicted from the airborne hyperspectral data.

R^2 of 0.72, RPD of 1.96 and a relatively small RMSE of 0.57. The agreement between the predicted pH values with the airborne imagery and those obtained from the ground-truth samples indicates that pH can be predicted reliably with airborne hyperspectral data.

When the pH maps were compared between years, the total areas with pH between 0 and 7 appeared to be smaller for data collected in 1998 and 1999. Similarly, smaller areas of jarosite were present in the earlier imagery, which is related to the low angle of the sun when the data were acquired, resulting in shadows and small values in the pixels representing the steep mine and tailing faces and benches. These data emphasize the importance of planning flights to optimize the quality of the data. There also appears to be consistency between the years in the mine area where the ranges of pH mapped were between 3 and 5. This agrees closely with pH data from water collected downstream from the mine for the same years (Randall & Cox, 2003).

As well as the steep mine faces where there are uncertainties with the mapping, there appears to be larger areas at pH 2–3 mapped in the 2000 and 2001 data, especially around the waste rock dumps and some of the benches of the mine, as shown inside the red polygon on Figure 5. These may be related to greater water flow in those 2 years than in 1998 and 1999, as reported by Randall & Cox (2003). Interestingly, there was a corresponding increase in sulphate, aluminium and iron loads for 2000 and 2001 (Randall & Cox, 2003).

These maps (Figure 5a–d) have the potential to provide environmental practitioners and soil scientists with a spatially-

comprehensive view of the whole mine, which may lead to improved understanding of the progress of restoration efforts, to identify areas on which to concentrate future remediation and to understand better the effects of downstream soil contamination. The following is an example of how these maps can be used to define areas where restoration or remediation has occurred and to monitor progress, as well as to define areas that need further investigation. The areas marked A, B and C had been identified as the 'problem' areas and, remedial efforts are performed here. At these locations, the pH maps show that the acid holding pond (marked B) was stable (no change) over the 4 years and the top right waste rock (marked A) shows only minor changes. In 1998 and 1999 this area was mapped as small patches at pH 3–4 surrounded by pH 4–5, which was surrounded by a larger area of pH 5–6. In 2000 and 2001 the patches of pH 3–4 were larger, with a corresponding decrease in the sizes of the pH 4–5 and 5–6 areas. The middle waste rock (marked C) showed changes mainly in the north-eastern corner of the waste rock, where in 1998 and 1999 the pH values were mainly in the 4–6 range and there were small patches of pH 3–4. In 2000 and 2001 areas at pH 3–4 were larger.

The largest changes were observed, however, in areas not previously identified by the environmental practitioners. Two such areas are located across the south-western mine benches (marked D) and across the north-western mine floors (marked E). Although the steep faces of mine benches at D were not mapped in 1998 and 1999 because of the poor signal (this area was in shadow) there was a general trend showing an increase in acidity in 2000 and 2001. This area was mapped at mainly pH 3–4 with small patches at pH 2–3 and 4–6 in 2000 and 2001, but was mapped at mainly pH 4–6 with small patches at pH 3–4 in 1998 and 1999. The variation observed at locations A, C, D and E could be a result of the greater precipitation in 2000 and 2001 compared with 1998 and 1999. Investigation of rainfall data and the water data (Randall & Cox, 2003) collected for those years confirms this. The larger amount of rainfall may have flushed out the system or reactivated or accelerated further sulphide oxidation and, hence, generated more acidic conditions.

Conclusions

This research examined the relationship between VNIR and SWIR spectral measurements and changes in pH caused by AMD at the Brukunga Pyrite Mine in South Australia. Laboratory spectral analysis of field samples from the mine compared with chemically determined pH measurements indicates that there is a significant linear relationship between pH and spectral data. The results of the PLS cross-validation analysis showed a general linear trend (at $R^2 = 0.58$ and a relatively small RMSE of 0.91) between pH measured using laboratory analysis and pH predicted using spectral data. The RPD of 1.53 indicates that the model generated from the PLS analysis may be able to provide sufficiently accurate predictions of pH from spectral data. The analysis of the FRCs produced from the PLS analysis also confirmed that diagnostic spectral features specific to secondary iron minerals were mainly responsible for the relationship, indicating that the PLS equation is appropriate. The strong reliance on the VNIR (350–1200 nm) region for the predictive equation, however, has significant implications when transferring this equation to airborne data. The VNIR region can be compromised severely by atmospheric and other acquisition conditions, such as time of day and orientation of the flight paths, which can affect the severity of bi-directional reflectance difference factor (BRDF) effects and the signal in terrains at high elevations. The reliance on the VNIR region emphasizes the need for proper pre-processing of the data for such issues as solar flux, atmospheric scattering (and atmospheric transmission) and good flight planning. Secondly, the confounding effects of vegetation on the spectra in the VNIR must be addressed. A simple solution for vegetation contamination is to mask all vegetation-affected pixels and produce pH maps for only non-vegetated areas. This strategy was used for this study. For most situations, this is a logical solution as most acid-affected areas contain only minimal amounts of green vegetation.

The pH maps produced from the airborne hyperspectral data acquired from 1998 to 2001 demonstrated the potential to provide environmental practitioners, including soil scientists, with a spatially-comprehensive view of contaminated areas that may help to improve understanding of the progress of rehabilitation efforts or pinpoint areas in which to concentrate further efforts, and to understand the spatial extent of the contamination and the resulting impacts, particularly regarding soils, on surrounding areas.

The validation approach adopted in this study was the classical technique of comparing an independent set of ground-truth data with the airborne maps. This technique is known to be limited, however, because of the mismatch in scale between airborne and ground measurements (McConnell & Weidman, 2009). Furthermore, there are technological challenges in locating the exact pixels. These challenges include the accuracy of the GPS (pre-2000 GPS measurements had an accuracy of > 20 m) and accuracy of the geo-rectification of the image. Despite the limitations, the results of the validation indicated that pH maps can be generated reliably from airborne hyperspectral data as indicated by the R^2 of 0.72, RPD of 1.96 and relatively small RMSE of 0.57. As a reconnaissance tool these maps could be used to pinpoint areas for further examination in the field, and this level of accuracy may be sufficient.

Acknowledgements

The support of the Department of Manufacturing, Innovation, Trade, Resources and Energy (DMITRE) in this study is gratefully acknowledged. We also acknowledge the assistance and contribution of our CSIRO colleagues.

References

Berk, A., Bernstein, L.S. & Robertson, D.C. 1989. *MODTRAN: A Moderate Resolution Model for LOWTRAN7*. Air Force Geophysical Laboratory, Hanscom AFB, MA.

Bigham, J.M., Schwertmann, U. & Pfab, G. 1996a. Influence of pH on mineral speciation in a bioreactor simulating acid mine drainage. *Applied Geochemistry*, **11**, 845–849.

Bigham, J.M., Schwertmann, U., Traina, S.J., Winland, R.L. & Wolf, M. 1996b. Schwertmannite and the chemical modelling of iron in acid sulfate waters. *Geochemica et Cosmochimica Acta*, **12**, 2111–2121.

Burtt, A.C. & Gum, J.C. 2000a. *The Dawesley-Bremer Catchment: Phase 2 Sampling Program for Potential Environmental Contaminants.* Report 2000/00019, PIRSA, Adelaide.

Burtt, A.C. & Gum, J.C. 2000b. *Soil and Stream Sediment Sampling of the Dawesley-Bremer Catchment for Potential Environmental Contaminants.* Report 2000/00002. PIRSA, Adelaide.

Chang, C.W., Laird, D.A., Mausbach, M.J. & Hurburgh, C.R. 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal*, **65**, 480–490.

Clark, R.N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. In: *Manual of Remote Sensing* (ed A. Rencz), pp. 3–58. John Wiley & Sons, Inc., New York.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I. & Shields, T. 1998. The hymap airborne hyperspectral sensor: the system, calibration and performance. In: *1st EARSEL Workshop on Imaging Spectroscopy*, pp. 37–42. EARSEL, Zurich.

Crowley, J.K., Williams, D.E., Hammarstrom, J.M., Piatak, N., Chou, I.-M. & Mars, J.C. 2003. Spectral reflectance properties (0.4–2.5 μm) of secondary Fe-oxide, Fe-hydroxide, and Fe-sulphate-hydrate minerals associated with sulphide-bearing mine wastes. *Geochemistry: Exploration, Environment, Analysis*, **2**, 219–228.

CSIRO 2005. *Xplot for Windows XRD Manipulation Program*. CSIRO Land & Water, Glen Osborne.

Curtis, B. 1985. *Evaluation of the physical properties of naturally occurring (III) oxyhydroxides on rock surfaces in arid and semi-arid regions using visible and near infrared spectroscopy*. PhD thesis, University of Washington, Seattle.

Dury, S., Turner, B., Foley, B. & Wallis, I. 2001. Use of high spectral resolution remote sensing to determine leaf palatability of eucalypt trees for folivorous marsupials. *International Journal of Applied Earth Observation & Geoinformation*, **3**, 328–336.

Geladi, P. & Kowalski, B.R. 1986. Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, **185**, 1–17.

Goetz, A.F.H. 2009. Three decades of hyperspectral remote sensing of the Earth: a personal view. *Remote Sensing of Environment*, **113**, S5–S16.

Gravestock, D.I. & Gatehouse, C.G. 1995. Stansbury basin in the geology of South Australia – the phanerozoic. In: *The Geology of South Australia – The Phanerozoic* (eds J.F. Drexel & W.V. Preiss), p. 347. Department of Mines and Energy, South Australia, Geological Survey of South Australia, Adelaide.

Green, A. & Craig, M. 1985. Analysis of aircraft spectrometer data with logarithmic residuals. In: *AVIRIS Workshop* (ed R.O. Green), pp. 111–119. NASA JPL, Pasadena, CA.

Grimalt, J.O., Ferrer, M. & MacPherson, E. 1999. The mine tailing accident in Aznalcollar. *Science of the Total Environment*, **242**, 3–11.

Haaland, D.M. & Thomas, E.V. 1988. Partial least squares. *Analytical Chemistry*, **60**, 1193–1200.

Hammarstrom, J.M., Seal, R.R. II, Meier, A.L. & Kornfeld, J.M. 2005. Secondary sulfate minerals associated with acid drainage in the eastern US: recycling of metals and acidity in surficial environments. *Chemical Geology*, **215**, 407–431.

Harries, J. 1997. *Acid Mine Drainage in Australia: Its Extent and Potential Future Liability*. Supervising Scientist, Kingston.

Isbell, R. 2002. *The Australian Soil Classification*. CSIRO Publishing, Collingwood.

Johnston, R., Barry, S., Bleys, E., Bui, E.N., Moran, C., Simon, D. *et al.* 2003. ASRIS: the database. *Soil Research*, **41**, 1021–1036.

Kemper, T. & Sommer, S. 2002. Estimate of heavy metal contamination in coils after a mining accident using reflectance spectroscopy. *Environmental Science Technology*, **36**, 2742–2747.

LaGanza, R. 1959. Pyrite investigations at Nairne, SA. *Economic Geology*, **54**, 895–902.

López-Pamo, E., Barenttino, D., Antón-Pacheco, C., Ortiz, G., Arránz, J.C., Gumiel, J.C. *et al.* 1999. The extent of the Aznalcóllar pyritic sludge spill and its effects on soils. *Science of the Total Environment*, **242**, 57–88.

Mars, J.C. & Crowley, J.K. 2003. Mapping mine wastes and analyzing areas affected by selenium-rich water runoff in southeast Idaho using AVIRIS imagery and digital elevation data. *Remote Sensing of Environment*, **84**, 422–436.

McConnell, M. & Weidman, S. 2009. *Uncertainty Management in Remote Sensing of Climate Data: Summary of a Workshop*. National Academies Press, Washington, DC.

McGovern, A.C., Ernill, R., Kara, B.V., Kell, D.B. & Goodacre, R. 1999. Rapid analysis of the expression of heterologous proteins in Escherichia coli using pyrolysis mass spectrometry and Fourier transform infrared spectroscopy with chemometrics: application to a2-interferon production. *Journal of Biotechnology*, **72**, 157–167.

Nandy, S., McLeary, M., Lightbody, P. & Carrigan, M. 2011. Conceptual hydrogeological model for brukunga mine site. In: *2011 Workshop on Australian Mine Rehabilitation* (eds O. Nichols & N. Vikuckis), pp. 181–195. JKTech Pty Ltd, Adelaide.

Nduwamungu, C., Ziadi, N., Parent, L.É., Tremblay, G.F. & Thuriès, L. 2009. Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: a review. *Canadian Journal of Soil Science*, **89**, 531–541.

NIST 2014. *Calibrations*. National Institute of Standards and Technology (NIST), Boulder, CO.

Polder, G., van der Heijden, G.W.A.M., van der Voet, H. & Young, I.T. 2004. Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectrometry. *Postharvest Biology & Technology*, **34**, 117–129.

Randall, J.E. & Cox, R.C. 2003. *2002 Water Quality Monitoring at Brukunga*. PIRSA, Adelaide.

Schaepman, M.E., Green, R.O., Ungar, S., Boardman, J., Plaza, A.J., Gao, B.C. *et al.* 2006. The future of imaging spectroscopy – prospective technologies and applications. In: *2006 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2005–2008. IEEE, Denver, CO.

Shaw, S.C., Groat, L.A., Jambor, J.L., Blowes, D.W., Hanton-Fong, C.J. & Stuparyk, R.A. 1998. Mineralogical study of base metal tailings with various sulfide contents, oxidized in laboratory columns and field lysimeters. *Environmental Geology*, **33**, 209–217.

Stuffler, T., Foerster, K., Hofer, S., Leipold, M., Sang, B., Kaufmann, H. *et al.* 2009. Hyperspectral imaging-an advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme). *Acta Astronautica*, **65**, 1107–1112.

Swayze, G.A., Clark, R.N., Pearson, R.M. & Livo, K.E. 1996. Mapping acid-generating minerals at the California Gulch Superfund site in Leadville, Colorado using imaging spectroscopy. In: *6th Annual JPL Airborne Earth Science Workshop* (ed R.O. Green), pp. 231–234. NASA JPL, Pasadena, CA.

Swayze, G.A., Smith, K.S., Clark, R.N., Sutley, S.J., Pearson, R.M., Vance, J.S. *et al.* 2000. Using imaging spectroscopy to map acidic mine waste. *Environmental Science & Technology*, **34**, 47–54.

Taylor, G.F. & Cox, R.C. 2003. The Brukunga Pyrite mine – a field laboratory for acid rock drainage studies. In: *6th ICARD*, pp. 93–106. ICARD, Cairns.

Tucker, C. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, **8**, 127–150.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. & Skjemstad, J.O. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, **131**, 59–75.

Williams, P.C. 1987. Variables affecting near-infrared reflectance spectroscopic analysis. In: *Near-infrared Technology in the Agricultural and Food Industries*, pp. 143–167. American Association of Cereal Chemists Inc., Saint Paul, MN.

Wold, S., Sjöström, M. & Eriksson, L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics & Intelligent Laboratory Systems*, **58**, 109–130.