Geological Survey of Canada

CURRENT RESEARCH
2006-C4

A comparison of automatic and supervised methods for extracting lithological end members from hyperspectral data: application to lithological mapping in southern Baffin Island, Nunavut

J.R. Harris, P. Ponomarev, S. Shang, P. Budkewitsch, and D. Rogge

2006

CURRENT RESEARCH

Natural Resources Canada Ressources naturelles Canada

Canada

ISSN 1701-4387
Catalogue No. M44-2006/C4E-PDF
ISBN 0-662-43901-5

A copy of this publication is also available for reference by depository libraries across Canada through access to the Depository Services Program's Web site at http://dsp-psd.pwgsc.gc.ca

A free digital download of this publication is available from GeoPub: http://geopub.nrcan.gc.ca/index_e.php

Toll-free (Canada and U.S.A.): 1-888-252-4301

Critical reviewer
David Viljoen

Authors' addresses

J.R. Harris (harris@nrcan.gc.ca)
Geological Survey of Canada
615 Booth Street
Ottawa, Ontario K1A 0E9

S. Shang (Shang@agri.gc.ca)
Agriculture and Agri-food Canada
K.W. Neatby Building
Room 4135, 960 Carling Avenue
Ottawa, Ontario K1A 0C6

D. Rogge
(drogge@ualberta.ca)
Department of Earth and Atmospheric Sciences
University of Alberta
1-26 Earth Sciences Building
Edmonton, Alberta T6G 2E3

P. Ponomarev
(peter.ponomarev@waypointinfo.com)
Waypoint Information
223 Colonnade Road South, Suite 114
Ottawa, Ontario K2E 7K3

P. Budkewitsch
(budkewit@nrcan.gc.ca)
Canada Centre for Remote Sensing
588 Booth Street, Room 3-355
Ottawa, Ontario K1A 0Y7

Publication approved by GSC Central Canada

Original manuscript submitted: 2006-06-03
Final version approved for publication: 2006-06-15

Correction date:

A comparison of automatic and supervised methods for extracting lithological end members from hyperspectral data: application to lithological mapping in southern Baffin Island, Nunavut

J.R. Harris, P. Ponomarev, S. Shang, P. Budkewitsch, and D. Rogge

Harris, J.R., Ponomarev, P., Shang, S., Budkewitsch, P, and Rogge, D., 2006: A comparison of automatic and supervised methods for extracting lithological end members from hyperspectral data: application to lithological mapping in southern Baffin Island, Nunavut; Geological Survey of Canada, Current Research 2006-C4, 19 p.

Abstract: This paper compares two automatic methods (unsupervised classification and an end-member extraction algorithm) and a supervised method (ground collection of training areas followed by matched filter classification) for producing a spectral map over a small area in eastern Baffin Island in the Canadian Arctic. The spectral maps and automatically generated spectra used to produce the spectral maps are compared to the spectral map and spectra produced using the supervised approach.

Both the supervised and unsupervised methods discussed in this paper provide spectral maps that provide similar lithological information when compared to the mapped geology. The supervised approach provides the results that are the most comparable to the mapped geology and this is not unexpected as this approach benefits from fieldwork and supporting geochemical and thin-section analysis.

Résumé : Dans le présent article, nous comparons deux méthodes automatisées (classification non dirigée et algorithme d'extraction de terme extrême) et une méthode de classification dirigée (collecte au sol de zones-échantillons suivie d'une classification par filtrage adapté) pour la production d'une carte spectraloïde d'un petit secteur de l'est de l'île de Baffin, dans l'Arctique canadien. Les cartes spectraloïdes et les spectres automatiquement générés servant à la production de ces cartes sont comparés à la carte spectraloïde et aux spectres produits en utilisant la méthode de classification dirigée.

Les méthodes de classification dirigée et non dirigée examinées dans cet article fournissent toutes deux des cartes spectraloïdes qui présentent de l'information lithologique similaire lorsque comparées à la carte géologique. La méthode de classification dirigée fournit les résultats les plus comparables à la cartographie géologique, ce qui n'est pas inattendu puisqu'elle s'appuie sur des travaux de terrain et sur des analyses géochimiques et de lames minces.

INTRODUCTION

Hyperspectral remote sensors offer significant advantages over multispectral optical remote sensors (i.e. Landsat) for geological mapping. The increased spectral resolution that hyperspectral remote sensors offer allows for identification of specific minerals and various rocks groups (i.e. groups of minerals) as opposed to simple discrimination. Reflectance properties of minerals have been well studied (Hunt and Salisbury, 1970, 1971; Hunt, 1977; Goetz et al., 1982; Clarke et al., 1990), although fewer studies have been conducted on the reflectance spectra of rocks (Hunt et al., 1973a, b, 1974). Significant research in hot desert areas of the world have shown that a wide range of minerals, particularly clay, hydroxyl-group minerals, and carbonate can be uniquely identified and mapped through spectral analysis (*see* van der Meer and de Jong, 2001, and references therein). Less research has been conducted in cold-desert environments typical of Canada's north; however, recent work using airborne hyperspectral (PROBE) data in the Canadian Arctic indicates that dolostone can be uniquely separated from limestone (Budkewitsch et al., 2000) and that metasedimentary, metatonalitic, and metagabbroic rocks can be discriminated (Harris et al., 2001, 2005).

Many methods can be envisioned with which to produce a spectral map. A spectral map is based on differences in spectral reflectance in a hyperspectral or multispectral, remotely sensed data set. In many cases the spectral map may correlate with lithogical and/or surficial units on the ground; however, field verification is generally required to geological calibrate a spectral map from a hyperspectral data set. This paper compares two automatic methods (unsupervised classification and an end-member extraction algorithm) and a supervised method (ground collection of training areas followed by matched filter classification) for producing spectral maps. The spectral maps and spectra generated automatically are compared to those produced using the supervised approach, as well as the existing lithological map. The supervised approach involved field verification, collection of field samples, and subsequent lab analysis using an Analytical Spectral Device (ASD) ground spectrometer, whereas the automatic approach did not.

STUDY AREA

The study area comprises a small portion (~ 5 km by 7 km) of southeastern Baffin Island along the Meta Incognita Peninsula located in northern Canada (Fig. 1). The area has been chosen as it represents a typical Arctic geological and biophysical environment with relatively good bedrock exposure and little vegetation cover (lichens excluded), comprising a variety of rock types ranging from gneissic and/or granitoid basement rocks to rift-related volcanic and sedimentary rocks, and has recently been mapped (St-Onge et al., 1998a, b; 2001). The recent bedrock mapping and associated field observations provide a basis for comparison between spectral maps derived from the hyperspectral data and the geological maps that have been produced by traditional methods (i.e. field traverses, aeromagnetic signatures, and analysis of airphotos).

The bedrock geology of this area (Fig. 2) was mapped by St-Onge et al. (1998a, b, 20012001) using traverses spaced 1 km to 3 km apart. The area comprises of Paleoproterozoic metasedimentary rocks, monzogranite, metatonalite and rare ultramafic rocks. Polyphase deformation and metamorphism during the Trans-Hudson Orogeny affected the Paleoproterozoic rocks, resulting in the current complex fold interference geometry (St-Onge et al., 1998a, b). The surficial environment includes rolling topography (~ 350–600 m) with appreciable areas of exposed and weathered bedrock. Lichen is prevalent and covers from 0% to 80% of exposed surfaces depending on rock type. Lichen and stunted vegetation (e.g. grasses, moss, shrub willows) grows selectively on specific rock types including metagabbro, metadiorite, and amphibolite.

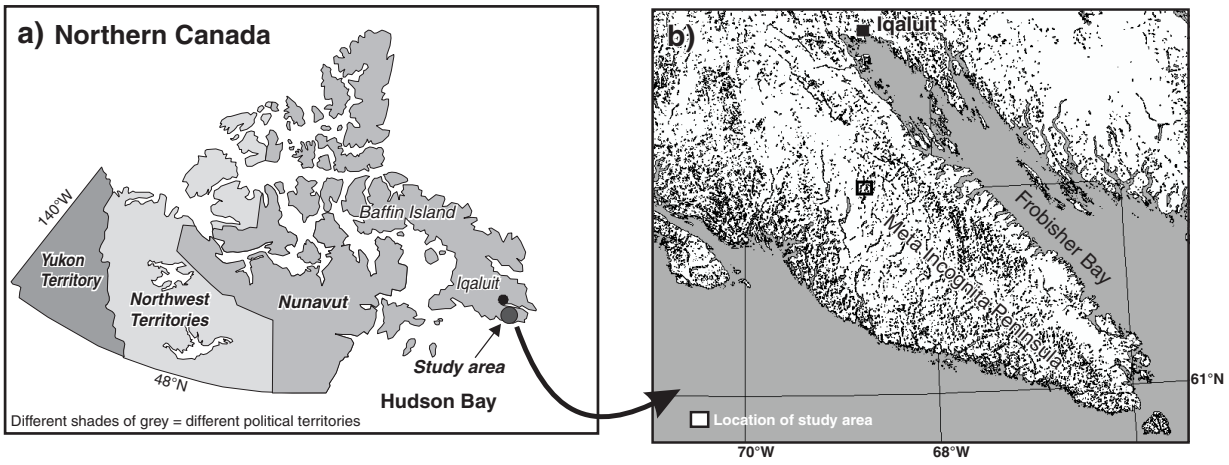


Figure 1 a) b). Study area.

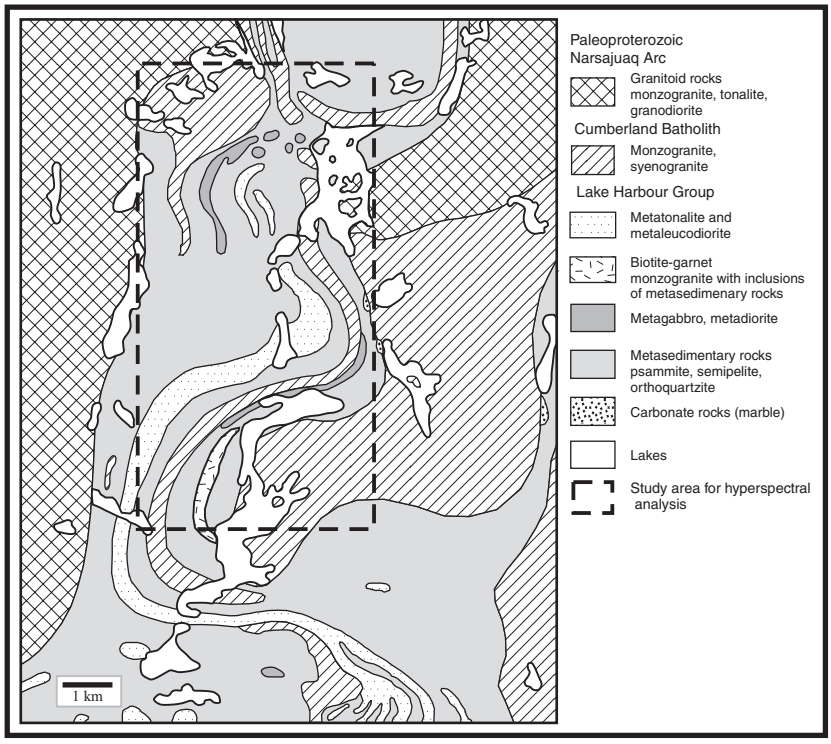


Figure 2. Geology of study area (*after* St-Onge et al., 1998a, b).

DATA

Noranda Inc. collected hyperspectral data along several flight lines over the study area (Fig. 1), using the airborne PROBE sensor in mid-July 2000. Swath width for each line was approximately 7 km and length was 50 km. Flight line orientation was north-south to minimize bi-directional reflectance artifacts. Solar elevation angles varied from approximately 45º to 37.5º during the period of acquisition. Spatial resolution, which is partially defined by the flight altitude, was approximately 5 m for this data set. The data consisted of 128 bands with a bandwidth of 15 nm ranging from 0.4 µm to 2.5 µm over the visible-near-infrared (VNIR) and short-wave infrared (SWIR) portions of the electromagnetic spectrum. The analysis presented here is focused on a small portion of one flight line (~14 km by 5 km, Fig. 2) that comprises a representative subset of the geology of southern Baffin Island.

METHOD

The data were processed in a number of steps summarized in Figure 3.

Data preprocessing

The data were converted from digital numbers to radiance values using calibration coefficients measured in the field (vicarious calibration) and then atmospherically corrected

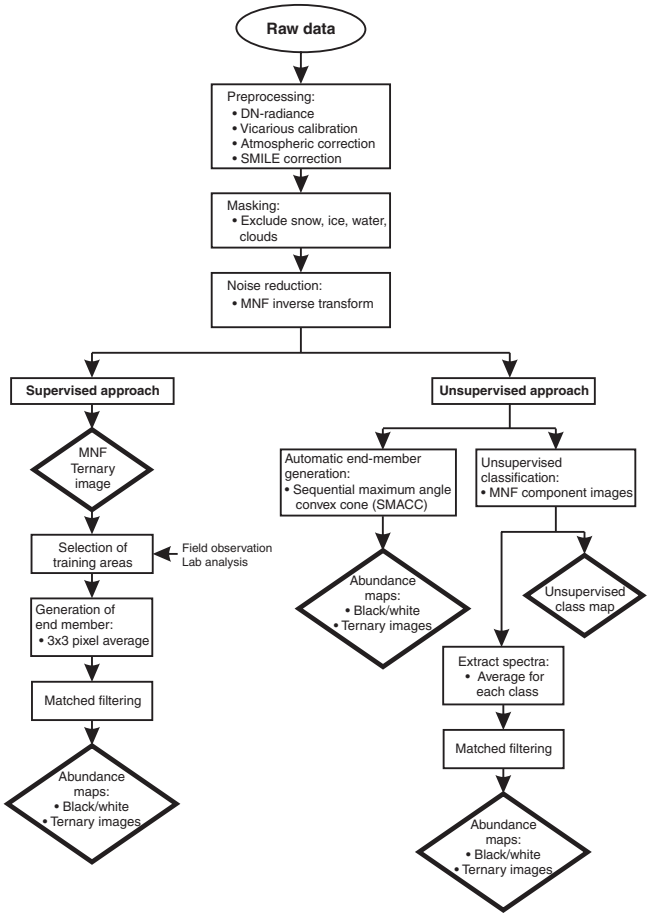


Figure 3. Data processing flow chart. Diamonds represent spectral maps useful for mapping different lithological and compositional units.

using a MODTRAN 4 algorithm (Berk et al., 1989). A flat field correction (SMILE; Neville et al. (2003; R.A. Neville, unpub. technical note, Canada Centre for Remote Sensing, 2004) was then applied to the data to correct for spectral line curvature.

Vicarious calibration was used to correct the radiometric coefficients supplied with the data (Secker et al., 2001). Ground spectra measured in situ were used to provide spectra for vicarious calibration. Atmospheric correction was based on a look-up-table (LUT) approach (Staenz and Williams, 1997), which considers the wavelength, pixel position, atmospheric water vapour, aerosol optical depth, and terrain elevation for the generation of LUT data and then a MODTRAN 4 algorithm is used to correct the data. The SMILE or 'slit-curvature' is a method to correct for spectral line curvature caused by the spectrograph's dispersing component (prism or grating) (Neville et al., 2003; R.A. Neville, unpub. technical note, Canada Centre for Remote Sensing, 2004).

The atmospherically corrected data were not geometrically corrected to preserve the spectral integrity of each pixel; however, the resulting maps and classifications were geometrically corrected using a second-order polynomial ('rubber-sheeting') to warp the data to the existing geological map for comparison purposes.

Data enhancement

Areas of snow, ice, water, and vegetation were masked out following the procedures set forth by Harris et al. (1998, 2005). Eliminating these areas allowed analysis to focus on areas of exposed or partially weathered bedrock; however, Harris et al. (2005) found that vegetation could be used as a proxy for mapping metagabbroic rocks within the study area and carbonate outside the immediate study area. This was due to the favourable chemistry of these rocks (high levels of MgO and or CaO) that nourished vegetation growth. Thus, in some cases, the exclusion of vegetation in northern regions may lead to a loss of potential lithological information. This must be evaluated on a case-by-case basis.

The data were noise reduced using the method proposed by Harris et al. (in press). This simple technique involves transforming the data over the masked areas (bedrock and weathered surface) using a minimum noise fraction (MNF) algorithm (Green et al., 1988), selecting the component images that account for more than 3% of the variance (Harris et al., 2005), and then applying an inverse MNF transform using only the selected component images to produce a noise-reduced data set comprising the same number of channels as the original data set. This is an effective technique since much of the noise in the MNF transformed data set resides in the higher component images that account for only a small variation of the information content within the data set. Ten component images were used in the inverse transform.

Once corrected and noise reduced, the hyperspectral data set was transformed again using an MNF algorithm. The MNF was employed as an enhancement technique as ternary images of the components (spectral maps) were found to provide good colour separation of the various lithological units (Harris et al., 2005), enabling the selection of training areas (end members) in concert with supporting field observations as well as a valuable spectral map for photogeological interpretation of various lithological units. In fact, Harris et al. (2005) found that ternary hyperspectral images produced using a MNF transform provided more lithological information than the existing geology map in the same area as this study.

Supervised approach

An MNF ternary image (Fig. 4) in conjunction with analysis of the existing geology map (Fig. 2) and field observations were used to select representative training areas for each lithological unit as well as check areas used for evaluating the various classifications produced from matched filtering (MF) (Fig. 4). Training areas (end members) were visually identified on the basis of colour homogeneity on the MNF colour ternary image as well as by inspecting N-dimensional scatterplots of the MNF components. Training areas (shown on Fig. 4) comprising a single pixel were collected for each lithological unit, derived from the MNF image; however, the end-member spectra used for classification were generated from an average of nine pixels in total surrounding the single pixel defining the training area. It was found that for mapping lithological units comprising a number of minerals an average spectrum based on nine, 5 m pixels, representing an area of 225 m^2, produced more representative spectra than spectra based on a single pixel. Furthermore, the averaging process helped to produce spectra that were less noisy. The training areas were selected in areas where topographic effects were likely to be minimal by inspecting the shaded DEM (not shown) of the study area. The bright and dark areas on the DEM (slopes facing the sun azimuth direction at the time of acquisition and the adjacent, poorly illuminated back slopes) were avoided when possible. Check areas (Fig. 4) were chosen based on the colour separation seen on the ternary image (Fig. 4), inspection of the mapped geology (Fig. 2), and field observations. The check areas were used as 'ground truth' with which to evaluate the various lithological maps produced by the supervised and automatic methods using crosstabulation (confusion matrix) analysis.

Matched filtering (MF) was applied to the masked and noise-reduced data (Fig. 3) and used to find the abundances of four end members (monzogranite, psammite, quartzite, and metatonalite) using a partial unmixing. Unmixing is a technique whereby the relative abundances of materials present in a multi- or hyperspectral data set are determined based on the material's spectral characteristics. The reflectance for each pixel is assumed to be a linear combination of each material (termed end member) present within each pixel. This technique maximizes

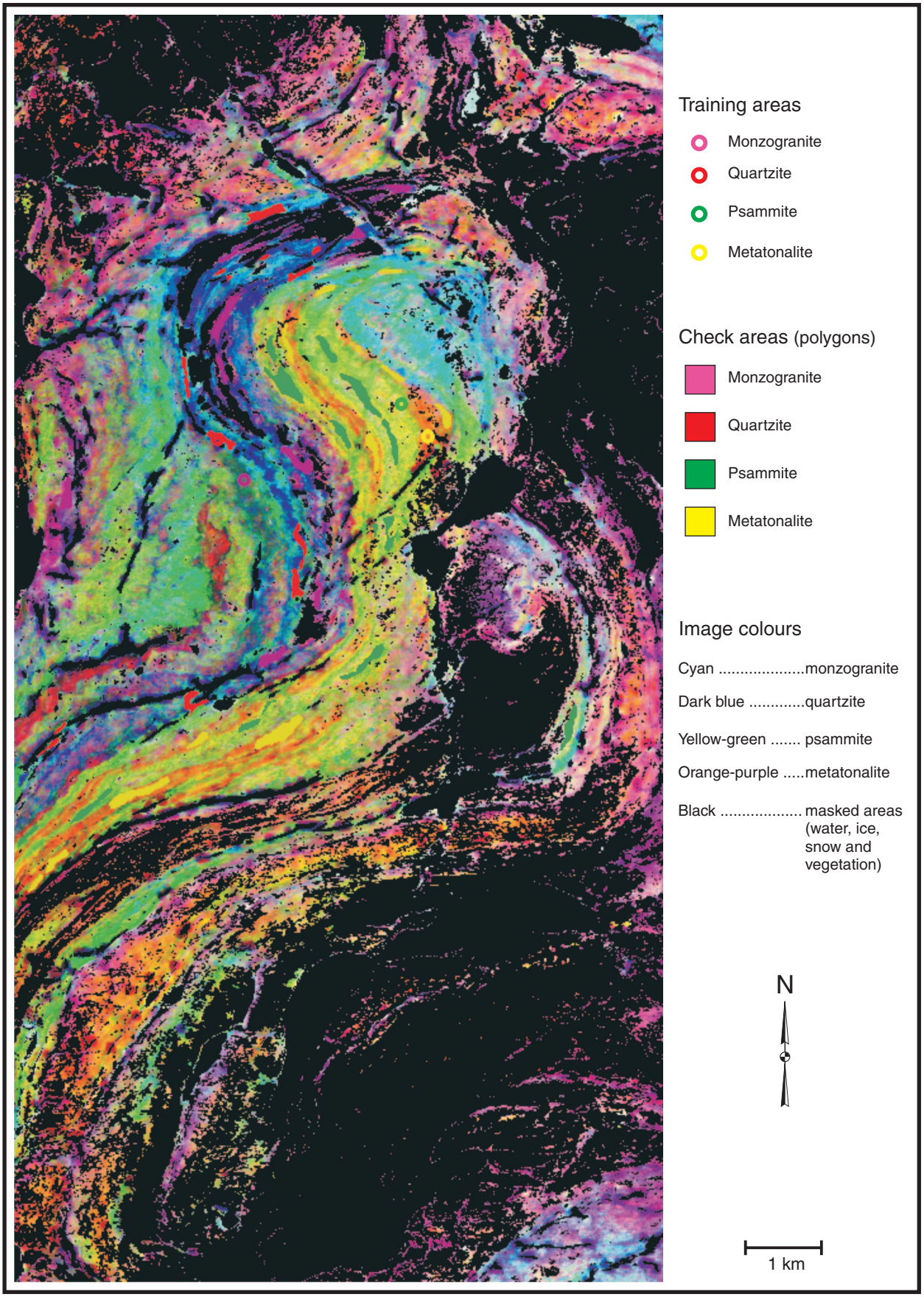


Figure 4. Minimum noise fraction ternary image (spectral map) showing different colours that represent different lithological and compositional units as well as training and check areas used for classification experiments.

the response of the known end member and suppresses the response of the composite unknown background, thus 'matching' the known signature. Unlike linear unmixing, matched filtering does not require knowledge of all the end members within the scene. Thus in areas of highly mixed rocks (typical of this study area) where identification of all the end members is difficult, matched filtering may be a better choice for classification. Matched filtering was selected to classify the data sets as Harris et al. (2005) found that better results were achieved using matched filtering on this data set as opposed to Spectral Angle Mapper (SAM), another popular algorithm for classifying hyperspectral data.

The results of the matched filtering appear as a series of gray-scale images (fraction or abundance maps), one for each selected end member. These fraction maps have values that range from 0 to 1 where 0 represents a nonmatch to the end-member (training) spectrum and 1 is a perfect match. The fraction maps can be combined in ternary images that are useful for displaying and interpreting lithological units. Thresholds using standard deviations, quantiles, or probability plots can then be identified to create binary maps from the fraction maps to show areas with relatively good matches to the end-member spectra or combined with other fraction maps as ternary images producing a lithological and/or compositional map. The binary maps can also be combined into class maps that can be used to compare to the check sites to evaluate the performance of the classification

Automatic approaches

Many methods exist for automatically deriving end members (spectra) from hyperspectral data including ORASIS (Bowles et al., 1998), N-FINDR (Winter, 1999) and Iterative Error Analysis (IEA) (Neville et al., 1999). In this study the authors have used the sequential maximum angle convex cone (SMACC) method as this offers advantages for hyperspectral data sets that have high correlation between channels that can impair unmixing by standard techniques (Gruninger et al., 2004).

Sequential Maximum Angle Convex Cone (SMACC) method

The SMACC algorithm, which is part of the ENVI™ image processing software, is based on a convex cone model for representing vector data where the end members are extracted directly from the hyperspectral data set. Extreme points in N-dimensional space are identified, forming a convex cone that comprises the first end member. The next end member is derived from applying a constrained oblique projecting to the existing cone and the cone is then increased to include the new end member. This process is repeated until a projection derives an existing end member or until the desired number of end members have been found. A set of abundance images (equalling the number of end members found) are generated which show the reflectance of each pixel contributed by each resulting end member. In simple terms SMACC finds the brightest pixel in the data set, then finds the pixel most different from the brightest and then the pixel most different from the first two. This is an iterative process that continues until a pixel already included in a previous group has been found or until the specified number of end members has been found. More detail on this algorithm including the mathematical derivation can be found in Gruninger et al. (2004).

The SMACC was applied to the masked and noise reduced data set (*see* Fig. 3). Initially the number of end members to be found was set at 30; however, the actual number found that appeared to be related to continuous lithological units, through preliminary visual analysis of the spectra and associated abundance images, was reduced to fourteen. The spectra of these derived end members and associated abundance images were then visually assessed in more detail and compared to the geology map (Fig. 2) and MNF ternary image (Fig. 4) and only those that appeared to be spatially continuous and related to lithological patterns were selected for further analysis. The total number of lithological end members after this screening process was six. These spectra and abundance images were evaluated by comparing them to the end members derived from supervised analysis (*see* above).

Unsupervised classification

A K-means clustering algorithm was applied to the masked and noise-reduced data set. This commonly used and popular algorithm initially uses a user-specified number of means (classes) evenly distributed in N-dimensional space and then iteratively assigns pixels to clusters reclassified after each iteration. More details on this algorithm can be found in Tou and Gonzales (1974).

K-means clustering was applied to the first 10 MNF component images that were calculated from the masked and noise-reduced data (Fig. 3). A total of six classes were derived for the classification and the average spectra for each of these classes were calculated and used as end members. It was felt that six classes would be sufficient to capture the signature (spectra) for the four known rock types as well as any possible extraneous signatures due to other factors (i.e. noise). These spectra were then used to classify the masked and noise-reduced data using MF, similar to the supervised process discussed above. These spectra and the associated unsupervised classes and abundance maps generated from MF were evaluated by comparing them to the end members and abundance images derived from the supervised analysis.

RESULTS

Supervised approach

The spectra for the quartz-bearing rocks (quartzite, monzogranite; Fig. 5a for spectra and Fig. 6a, b, c, d for type exposures and hand samples of the various rock types) have the highest intensity and this is expected as felsic rocks generally have higher albedos and thus greater overall reflectance. The difference between the two spectra other than the quartzite being more reflective occurs in the short-wave-infrared (SWIR) portion where a moderate absorption band typifies the quartzite spectra around 2.32 µm, indicating the presence of another mineral other than pure quartz or an effect due to lichen. A reflectance peak or absorption trough is absent for the monzogranite in this portion of the spectra whereas the metatonalite and psammite have reflectance peaks.

The monzogranite is characterized by the second highest reflectance (Fig. 5a, 6c, d) next to the quartzite, but the spectrum is relatively featureless with no noticeable peaks or troughs in the VNIR or SWIR.

The metatonalite (Fig. 5a, 6e, f) is typified by lower reflectance than the psammite throughout the VNIR and SWIR. The psammite (Fig. 5a, 6g, h, i) has a strong iron signature in the visible portion of the electromagnetic spectrum

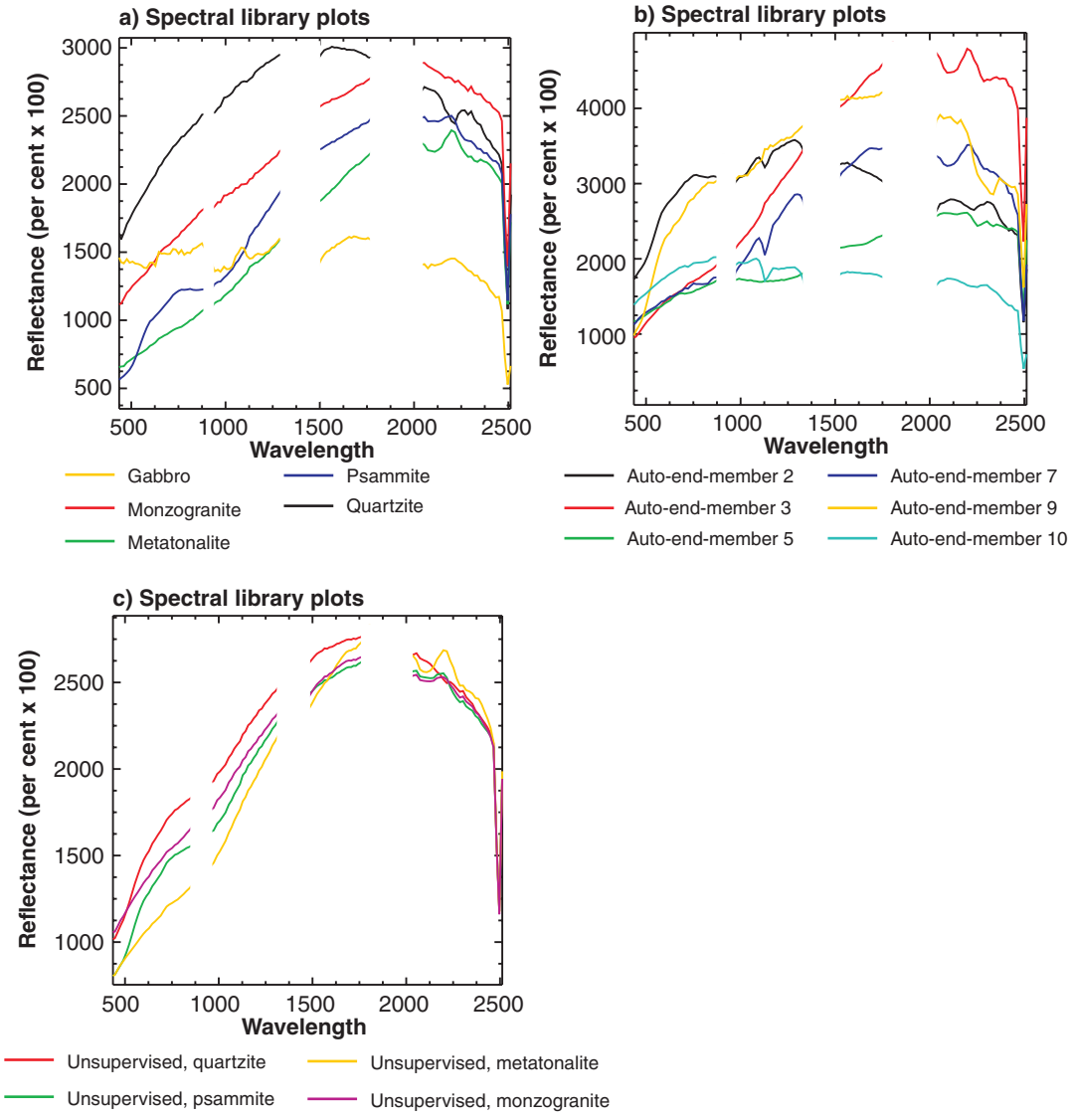


Figure 5. Spectra extracted from the airborne PROBE hyperspectral data set for various rock types, **a)** spectra extracted from training areas for each major lithological unit (training areas shown on Fig. 4), **b)** spectra extracted automatically using the Sequential Maximum Angle Convex Cone (SMACC) algorithm from the PROBE data, **c)** spectra extracted from an average of the classes derived from an unsupervised classification (K-means) of the PROBE data.

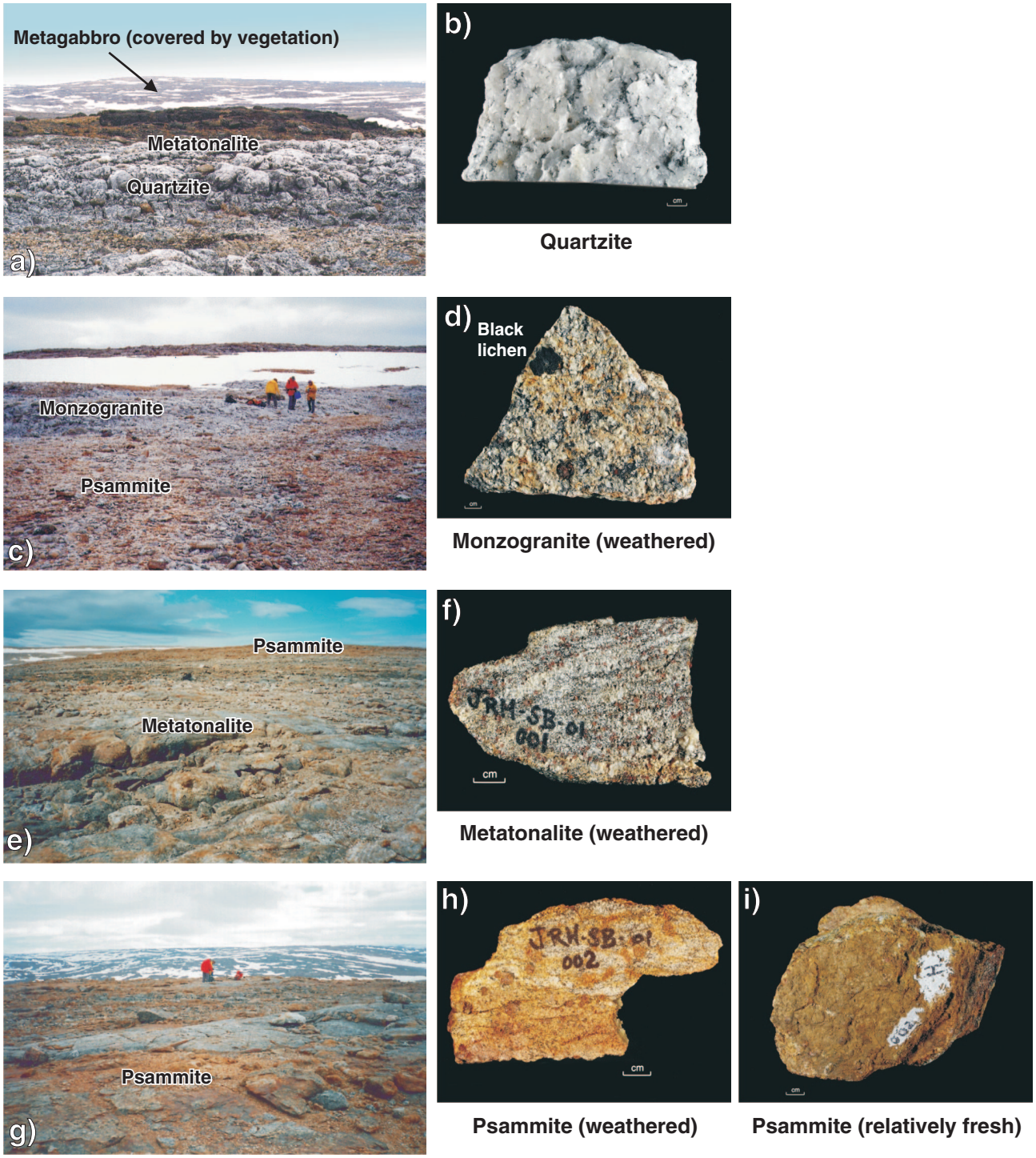


Figure 6. Ground photographs and hand samples of the various lithological units: **a)** quartzite and metagabbro, **b)** quartzite hand sample (fresh surface), **c)** contact between psammite and monzogranite, **d)** monzogranite hand sample (weathered surface), **e)** metatonalite and psammite, **f)** metatonalite hand sample (weathered surface), **g)** psammite, psammite hand sample, **h)** fresh and **i)** weathered surface.

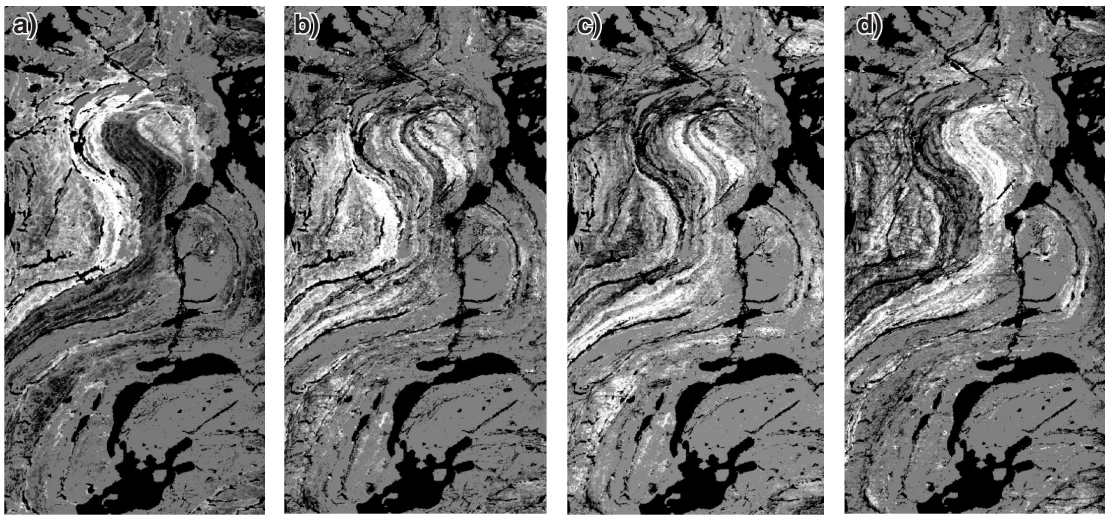
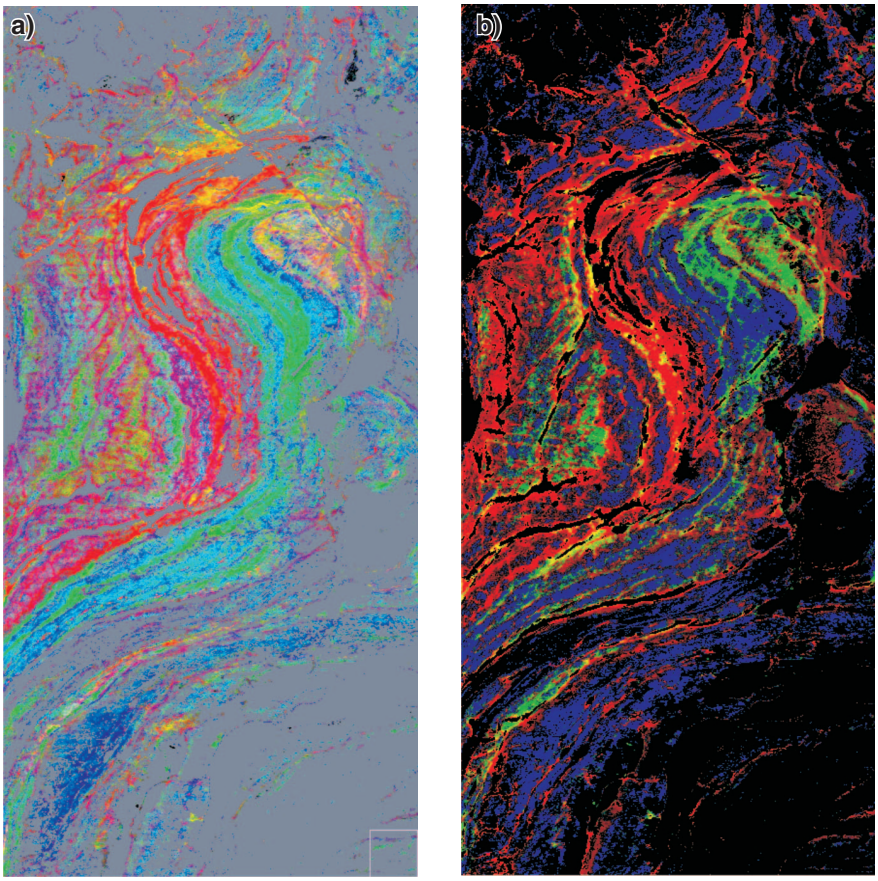


Figure 7. Abundance maps for each major lithological unit derived from training area spectra (Fig. 5a) and matched filtering (MF) classification based on noise-reduced hyperspectral data set, **a)** quartzite abundance map, **b)** monzogranite abundance map, **c)** metatonalite abundance map, and **d)** psammite abundance map.



Red = quartzite, green = psammite, blue/cyan = metatonalite

Red = auto-end-member 2 (quartzite), green = auto-end-member 9 (psammite), blue = auto-end-member 3 (metatonalite)

Figure 8. Ternary images based on abundance maps produced from supervised and automatic approaches for identifying end members, **a)** based on the supervised approach, **b)** based on automatic approach (Sequential Maximum Angle Convex Cone, SMACC).

typified by a steep slope from the blue to infrared wavelengths with a peak around 0.834 µm. Harris et al. (in press) provide a more detailed interpretation of the spectra.

The individual rock types are quite distinct on the MNF ternary image (Fig. 4); the psammite units are green-yellow, the monzogranite units are cyan, the metatonalite units are orange-red, and quartzite units are blue. The MF fraction maps produced (Fig. 7), which in fact represent mixtures of the various rock types, in some cases do show reasonable separation of each lithological unit when compared to the MNF ternary image (Fig. 4) and the existing geology map (Fig. 2). The brighter tones in Figure 7 indicate areas that have a better spectral match to the end member (training spectra), the locations of which are shown on Figure 4, and the actual spectra on Figure 5a.

The quartzite abundance map (Fig. 7a) shows a sinuous (folded) zone of resistant quartzite that has been included in one large metasedimentary unit on the geology map (Fig. 2). In addition the dark, sinuous band in the central portion of the image closely correlates with the metasedimentary unit comprising psammite. The psammite unit is a dark tone on the quartzite abundance map (Fig. 7a), indicating a low match and thus a large difference in spectral response (*see* Fig. 5a). The monzogranite abundance map (Fig. 7b) roughly correlates with the mapped monzogranite (Fig. 2) but appears to be more extensive and continuous than mapped. The metatonalite abundance maps (Fig. 7c) appears to comprise two separate units differing from the one larger mapped unit (Fig. 2). The psammite abundance map (Fig. 7d) comprises one moderately thick sinuous unit and a thinner unit that have been lumped into the mapped metasediment unit on the regional geology map (Fig. 2).

A ternary image (Fig. 8a), comprising the abundance maps shown in Figure 7 provides a first-order lithological map in which the quartzite, psammite, and metatonalite units are clearly separated displayed in red, green, blue and/or cyan colours, respectively.

Each of the abundance maps shown in Figure 7 were converted to a classified map using a threshold of three standard deviations above the mean value for each map. A threshold of three standard deviations was chosen to create the class maps as this includes the areas of the best match (i.e. > 95% match) to each of the spectra (Fig. 5a) as well minimizing possible areas of overlap (confusion) between end members. These were then combined into one image (Fig. 9a) and compared to

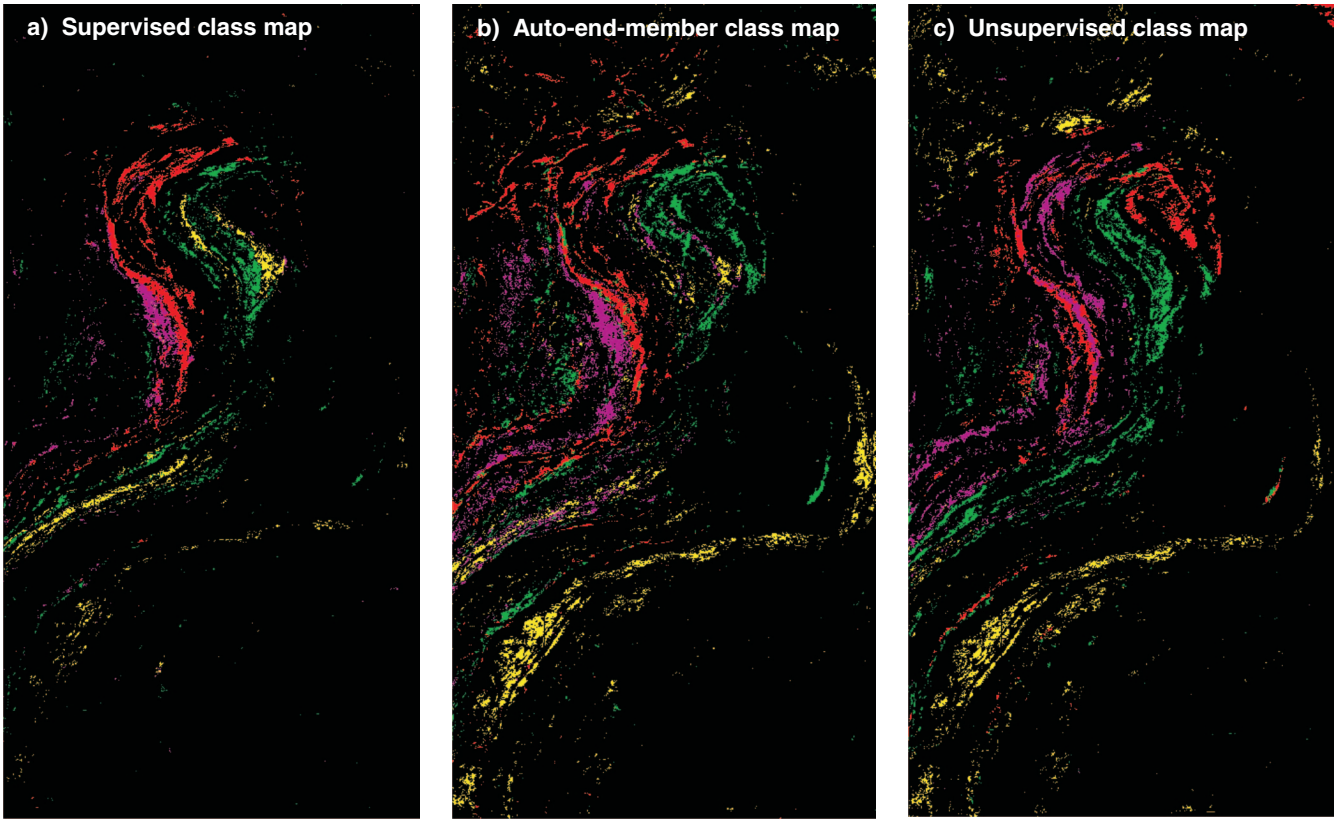


Figure 9. Class maps derived from thresholding abundance maps **a)** based on the supervised approach where abundance maps (Fig. 7) were created using spectra shown on Figure 5a (training areas for spectra shown on Fig. 4) and matched filtering, **b)** based on the automatic approach (Sequential Maximum Angle Convex Cone, SMACC) to define end members (spectra shown on Fig. 5b) and abundance maps (*see* Fig. 10), **c)** based on the automatic approach (unsupervised classification) to define end members (spectra shown on Fig. 5c) and abundance maps (*see* Fig. 13).

check areas (Fig. 4) identified in the field for each lithology using crosstabulation (confusion matrix) analysis. An 88.7% accuracy was achieved when comparing the class map to the check areas and greater than 90% when comparing the class map to the training areas (Table 1). The quartzite and monzogranite were the most confused as they have similar spectra confirmed by transformed divergence (TD) analysis (Table 2) that shows a smaller TD value (1.31) that is indicative of only moderate statistical separation between spectra. Both units are characterized by high reflectance due to the high abundance of quartz (Fig. 5a). The kappa coefficients which are generally high (> 0.4) (Table 1) also indicate good agreement between the class map and check training areas. Quartzite shows good spectral separation between psammite and metatonalite (TD > 1.9, Table 2) and moderate separation from monzogranite (TD of 1.31, Table 2). The psammite and metatonalite shows moderate separation (TD of 1.72) with psammite characterized by higher reflectance in both the VNIR and SWIR as well as strong iron reflectance in the visible (V) range (Fig. 5a).

Table 1. Results of comparing class maps (Fig. 9) for each end member extraction method to check areas shown on Figure 4.

	Accuracy (%)	Kappa coefficient
1. Supervised		
Training areas	92.7	0.90
Check areas	88.7	0.83
2. Auto-end-members		
Training areas	69.4	0.63
Check areas	49.9	0.29
3. Unsupervised		
Training areas	16.6	0.06
Check areas	88.1	0.83

Automatic approach

Sequential Maximum Angle Convex Cone method

The spectra of the six automatically derived end members (Fig. 5b) are quite distinct and provide four abundance images (Fig. 10) that can be directly related to mappable lithological units. Using the ENVI™ Spectral Analyst™ function the best match between the automatically derived spectra (auto-end-members, Fig. 5b) and the supervised spectra (Fig. 5a) can be obtained (Table 3, Fig. 11). The output from Spectral Analyst is a ranked score between 0 and 1 (based on spectral feature fitting); the highest score (*see* Table 3) indicates the closest match and higher confidence in spectral similarity. Auto-end-member 2 (Fig. 5b) matches well with the quartzite end member (Fig. 11a) and to a lesser

Table 2. Matrix of Transformed Divergence (TD) values for each end member derived from training areas shown on Figure 4. (TD >1.9 = good separability, <1 is poor; colours based on MNF ternary image in Fig. 4).

	Quartzite	Psammite	Metatonalite	Monzogranite
Quartzite		**1.9**	**1.95**	1.31
Psammite			1.72	**1.99**
Metatonalite				**1.99**
Monzogranite				

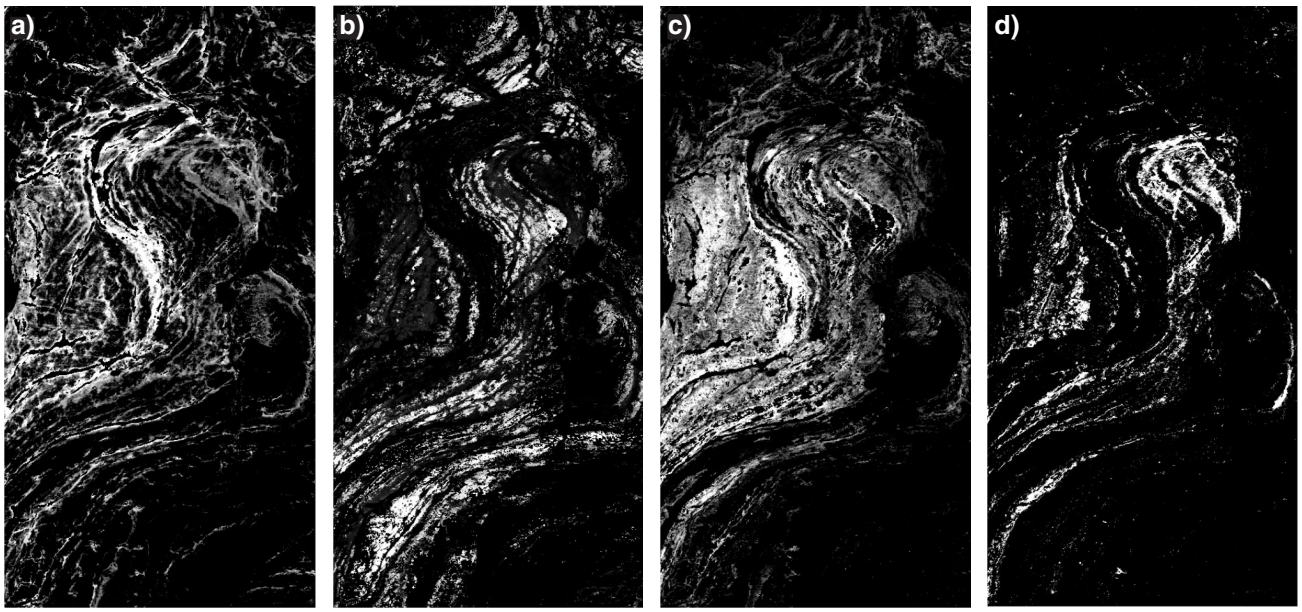


Figure 10. Abundance maps derived from the automatic method (Sequential Maximum Angle Convex Cone, SMACC) to derived end members: **a)** auto-end-member 2, mostly quartzite, **b)** auto-end-member 3, mostly metatonalite, **c)** auto-end-member 5, mostly monzogranite, **d)** auto-end-member 9, mostly psammite.

extent the monzogranite end member. The associated abundance map for the auto-end-member 2 (Fig. 10a) has the strongest correlation (*see* last column on Table 3) with the quartzite abundance map (Fig. 7a). Auto-end-member 3 (Fig. 5b) has the strongest match with the metatonalite spectra (Fig. 11b) and the abundance map (Fig. 10b) is moderately correlated (0.4, Table 3) with the metatonalite abundance map derived from the supervised approach (Fig. 7c). Auto-end-member 5 (Fig. 5b) best matches the monzogranite spectrum (Fig. 11c) and the abundance map (Fig. 10c) has the highest correlation (0.51, Table 3) with the monzogranite abundance map (Fig. 7b) whereas auto-end-member 9 (Fig. 5b) best matches the psammite

Table 3. Best spectral matches between end members derived automatically using SMACC and end members derived from training areas shown on Figure 4.

Auto-end-member	Best spectral match (SFF)	Second best spectral match (SFF)	Lithological type	Highest correlation (with MF-abundance maps)
2	Quartz (0.61)	Monzogranite (0.51)	Quartz-monzogranite	Quartz = 0.68
3	Metatonalite (0.79)	Psammite (0.67)	Metatonalite	Metatonalite = 0.4
5	Monzogranite (0.47)	Metatonalite (0.45)	Monzogranite-metatonalite	Monzogranite = 0.51
9	Psammite (0.59)	Metatonalite (0.57)	Psammite	Psammite = 0.3

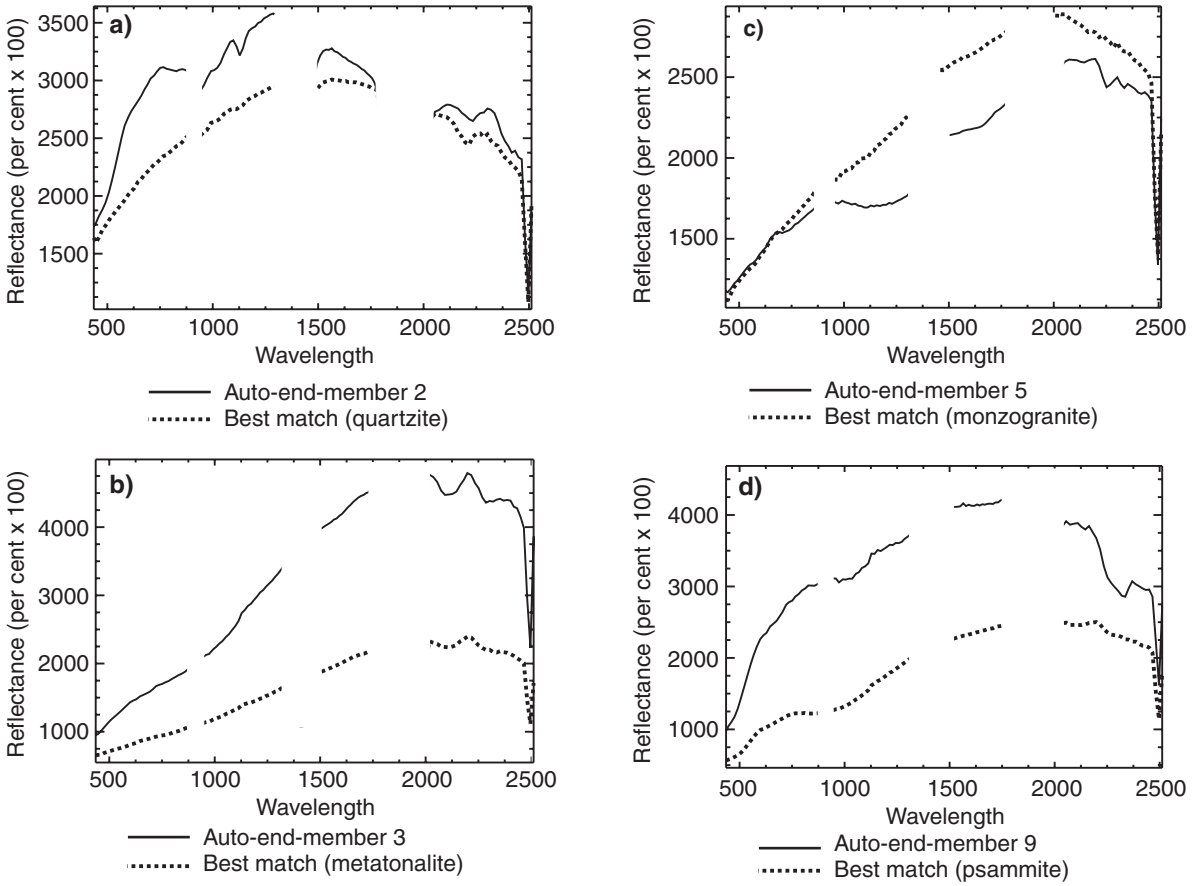


Figure 11 (left). Spectra derived from automatic generation of end members (Sequential Maximum Angle Convex Cone, SMACC), **a)** auto-end-member 2, mostly quartzite and best match from spectra (Fig. 5a) derived from training sites (Fig. 4), which in this case is quartzite, **b)** auto-end-member 3, mostly metatonalite and best match from spectra (Fig. 5a) derived from training sites (Fig. 4), which in this case is metatonalite, **c)** auto-end-member 5, mostly monzogranite and best match from spectra (Fig. 5a) derived from training sites (Fig. 4), which in this case is monzogranite, **d)** auto-end-member 9, mostly psammite and best match from spectra (Fig. 5a) derived from training sites (Fig. 4), which in this case is psammite.

spectra (Fig. 11d) and the abundance map (Fig. 10d) shows the highest correlation (0.3, Table 3) with the psammite abundance map (Fig. 7d).

A ternary map of the abundance images (Fig. 8b) provides reasonable separation of the quartzite and metasedimentary rocks, although not nearly as clear as the ternary map derived from the MF abundance maps based on the training areas (Fig. 8a).

A class map (Fig. 9b) was derived from each of the abundance maps for each auto-end-member using the same approach used to produce class maps of the MF fraction maps, discussed above. The class map was compared to the check areas using a crosstabulation. Reasonable accuracies were achieved between the class map and the check areas (49.9%) and supervised training areas (69.4%) (Table 1). Again confusion was highest between auto-end-member 2 (reflecting mostly quartzite) and auto-end-member 5 (mostly monzogranite). Some confusion between auto-end-member 9 (psammite) and quartzite and auto-end-members 3 (metatonalite) and 9 (psammite) was also found.

Unsupervised method

An unsupervised classified map of the first 10 MNF component images (Fig. 12) produces a reasonable lithological map when compared to the MNF ternary image (Fig. 4) and

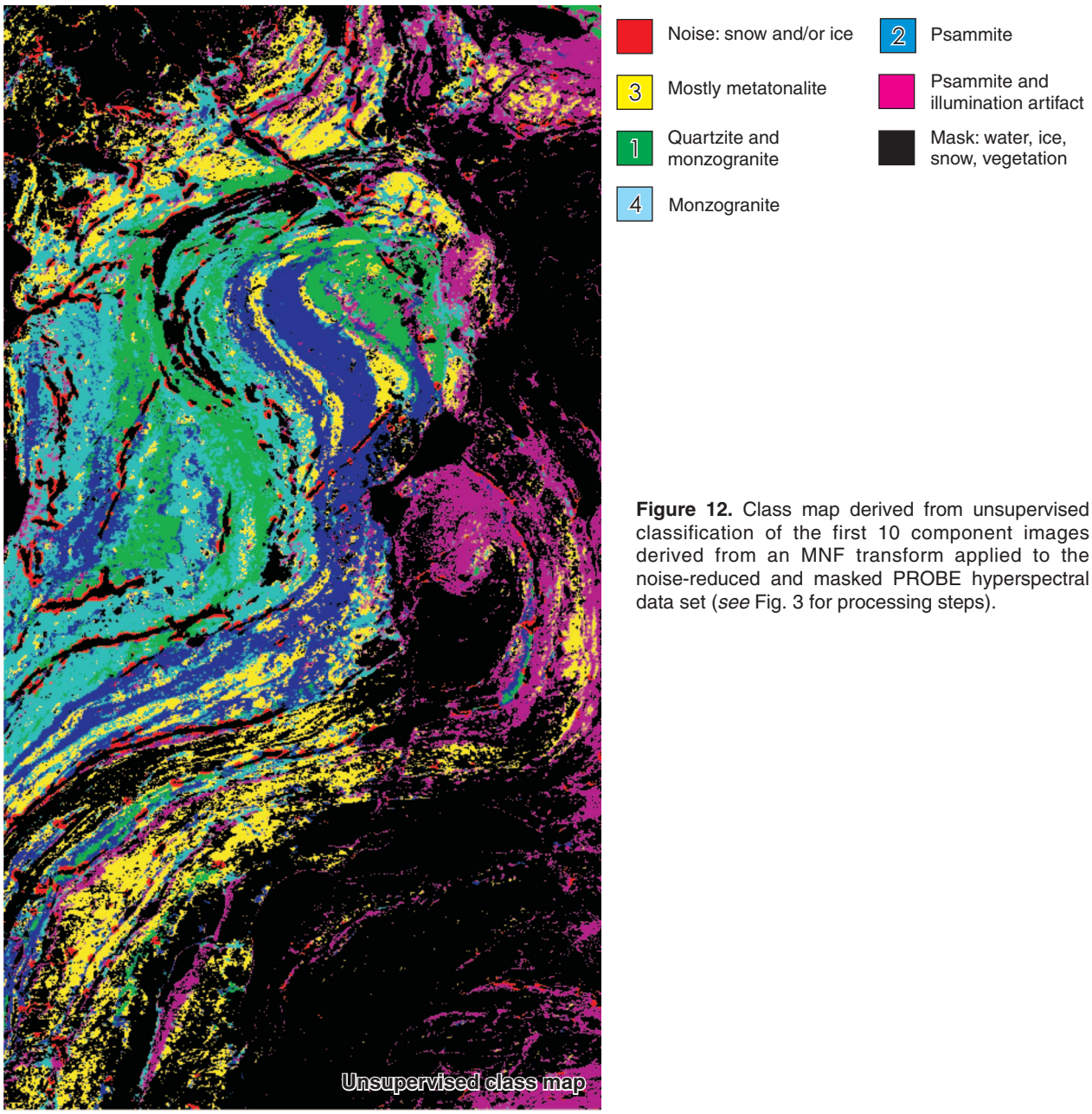


Figure 12. Class map derived from unsupervised classification of the first 10 component images derived from an MNF transform applied to the noise-reduced and masked PROBE hyperspectral data set (*see* Fig. 3 for processing steps).

the mapped geology (Fig. 2); however, the classification has highlighted an illumination problem in the raw data reflected by the magenta colour on the right side of the image. This is in part a reflection of lithological differences, but also of unequal illumination across the swath. Methods to correct for this unequal illumination effect are presently under investigation. Comparison to the MNF ternary image (Fig. 4) and geological map (Fig. 2) as well as field observations indicates that the yellow class represents mostly metatonalite, the blue class psammite, the green and cyan, quartz-bearing rocks (quartzite and monzogranite). The magenta class represents psammite and/or sandstone, but also is a manifestation of unequal illumination across the swath as mentioned above.

The average spectra for each unsupervised class (Fig. 5c) were used to classify the noise-reduced and masked data using MF and Figure 13 shows the resulting abundance maps.

The spectrum for unsupervised class 1 (Fig. 5c for spectrum and Fig. 12 for class) best matches the quartzite spectrum (Table 4, Fig. 14a) and the associated abundance map (Fig. 13a) has the highest correlation with the quartzite abundance map (Fig. 7a). The spectrum from unsupervised class 2 (Fig. 5c, 12) best matches the quartz-bearing rocks (quartzite and monzogranite; Table 4; Fig. 13b); however, the highest correlation of the abundance map (Fig. 13b) is with the psammite abundance map (Fig. 7d). Indeed, visual comparison of these abundance maps indicates that unsupervised class 2 is more representative of psammite, which is the third best spectral match (Table 4) according the 'Spectral Feature Fitting' algorithm. This stresses the need for independent evaluations of the spectra through a comparison of abundance maps (as above), as spectral matching algorithms are by no means foolproof. This is further illustrated by unsupervised classes 3 and 4, the spectra of which best match quartzite according to

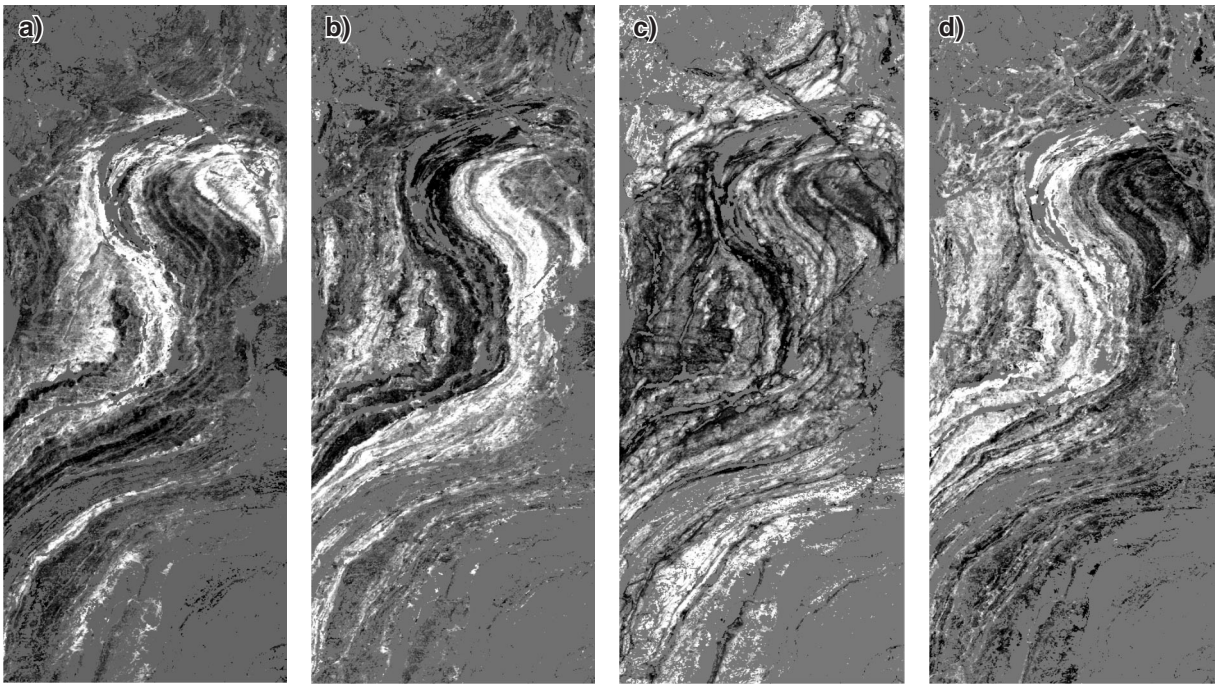


Figure 13. Abundance maps producing using average spectra (Fig. 5c) for each class derived from the unsupervised classification (Fig. 12), using matched filtering (MF), **a)** quartzite and/or monzogranite, **b)** psammite, **c)** metatonalite, **d)** monzogranite and/or quartzite.

Table 4. Best spectral matches between end members derived automatically using unsupervised classification and end members derived from training areas shown on Figure 4.

Unsupervised end-member class	Best spectral match	Second best spectral match	Lithological type
2, green	Quartz (0.864)	Monzogranite (0.777)	Quartz or monzogranite
3, blue	Quartz (0.8)	Monzogranite (.74) / psammite (0.69)	Monzogranite or psammite
4, yellow	Quartz (0.754)	Metatonalite (0.74)	Metatonalite or quartzite
5, cyan	Quartz (0.863)	Monzogranite (0.775)	Monzogranite or quartzite

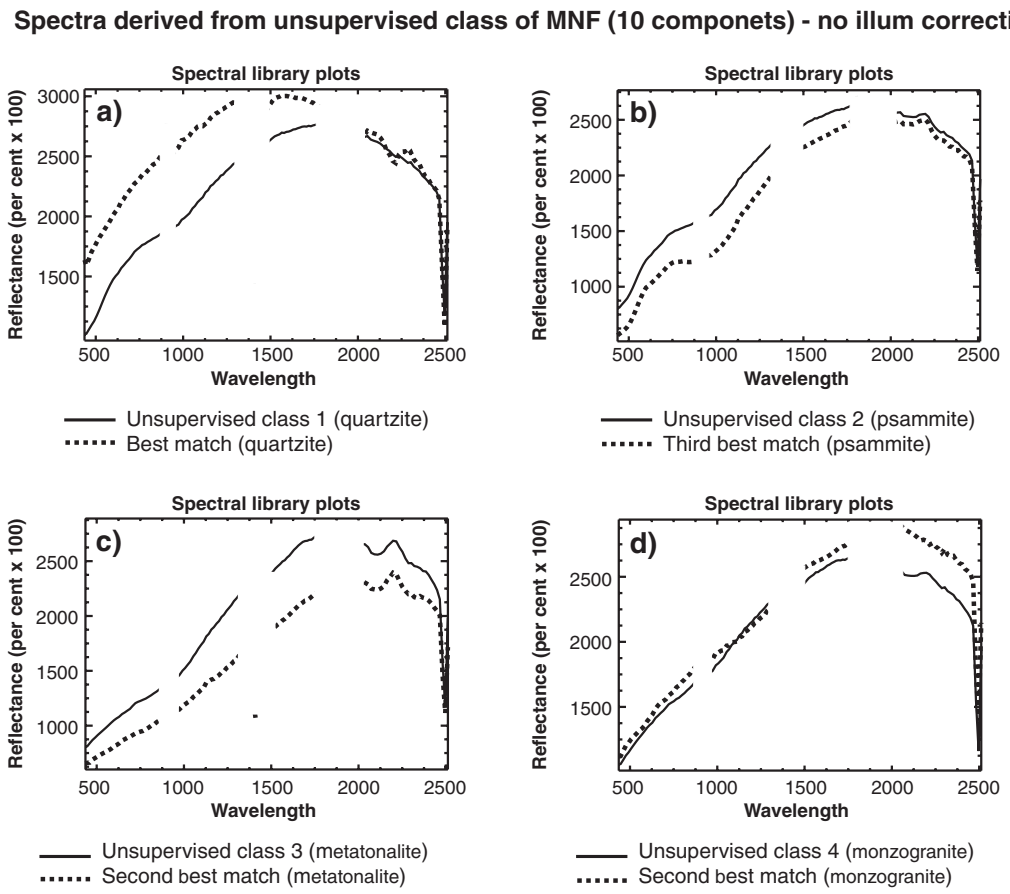


Figure 14. Spectra derived from automatic generation of end members (unsupervised classification, K-means) represents average spectra for each unsupervised class (*see* Fig. 12), **a)** class 1, mostly quartzite and best match from spectra derived from training sites (Fig. 4) which in this case is quartzite, **b)** class 2, mostly psammite and third best match from spectra derived from training sites (Fig. 4), which in this case is psammite, **c)** class 3, mostly metatonalite and second best match from spectra derived from training sites (Fig. 4), which in this case is metatonalite, **d)** class 4, mostly monzogranite and second best match from spectra derived from training sites (Fig. 4), which in this case is monzo-granite.

the spectral feature-fitting algorithm (Table 4); however, a better visual match is obtained by the spectra that have the second best match, these being metatonalite and monzogranite for unsupervised spectra 3 and 4, respectively (Fig. 14c, d; Table 4). Indeed the abundance maps for these spectra (Fig. 13c, d) show the highest correlations with the metatonalite and monzogranite abundance maps (Fig. 7c, b).

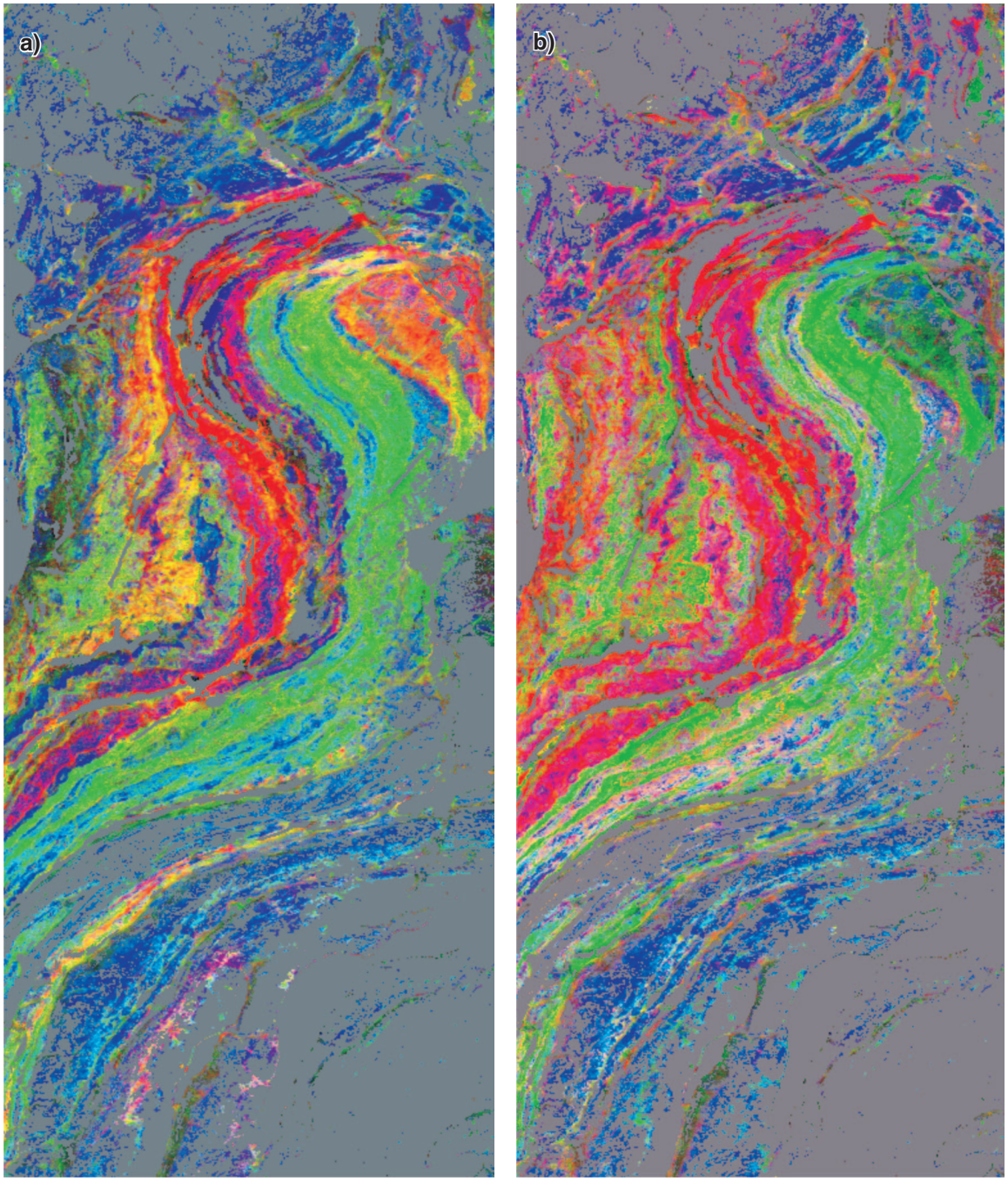
In general, the averaged spectra for each unsupervised map class (Fig. 5c) appears more similar to the supervised spectra (Fig. 5a) than the automatically derived spectra (Fig. 5b).

Comparison of the class map (Fig. 9c) derived by thresholding each abundance map in Figure 13 (*see* above) indicates a good match between the ground verified check areas and predicted lithology derived from the unsupervised classification (Table 1). In fact the overall accuracy is equal to the accuracy derived from the supervised approach when comparing the class maps to the check areas. The greatest confusion is between class 1 representing mostly quartzite and class 4 (mostly monzogranite).

A reasonable lithological map can be created by combining the MF abundance maps (Fig. 13) derived from the unsupervised classification into ternary images (Fig. 15). Red on these maps reflect primarily quartzite; magenta and yellow, monzogranite; green, psammite; and cyan-blue, metatonalite.

DISCUSSION

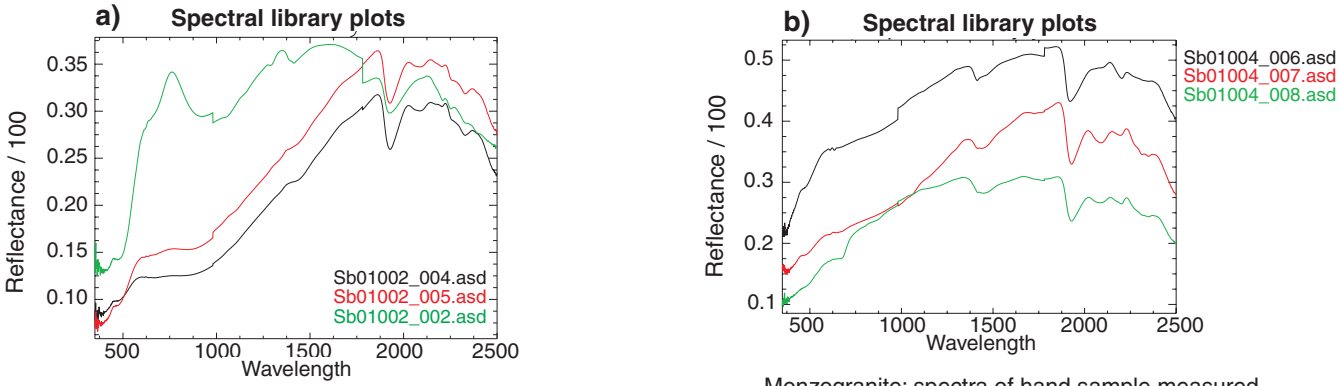
All methods discussed in this paper (Fig. 3) provide reasonable spectral maps that appear to represent different lithological units when compared to the mapped geology (Fig. 2) and the MNF ternary image (Fig. 4) that we have used as 'ground truth' in this paper as the colours in this image have been geologically calibrated in the field (Harris et al., 2005).



RGB class 2 (quartzite and/or monzogranite)
RGB class 3 (psammite)
RGB class 4 (metatonalite)

RGB class 5 (quartzite and/or monzogranite)
RGB class 3 (psammite)
RGB class 4 (metatonalite)

Figure 15. Ternary images (spectral maps) based on abundance maps derived from average spectra from unsupervised classification, **a)** using abundance maps derived from average spectra from unsupervised classes 2, 3, and 4, **b)** using abundance maps derived from average spectra from unsupervised classes 5, 3, and 4; RGB = red-green-blue.



Psammite: spectra of hand sample measured using an ASD spectrometer (internal light source)
002: fresh surface, about 30% garnet
004: weathered rusty surface, about 50% lichen
005: weathered surface, 10–30% lichen

Monzogranite: spectra of hand sample measured using an ASD spectrometer (internal light source)
006: weatherd surface, < 10% lichen
007: weathered surface, about 30% lichen
008: weathered surface, about 90% lichen

Figure 16. Spectra of hand samples using internal light source from the ASD spectrometer contact probe **a)** psammite **b)** monzogranite.

In fact these maps provide apparently more lithological information than on the geology map (Fig. 2). The various lithological maps produced in this study (especially Fig. 4, 8a, 9, 12, 15) reveal more compositional variation than is present on the regional geological map (Fig. 2); however, it must be kept in mind that the geological map is based on regional scale transects that are spaced on average 3 km apart and filled in from interpretations of airphotos in contrast to information derived from 5 m pixels on the airborne hyperspectral data. In addition compositional units have been grouped together on the geological map (i.e. quartzite is included in the mapped metasedimentary unit that comprises semipelite, psammite, orthoquartzite, and monzogranite). Therefore, this result might be expected; however, the hyperspectral data offers compositional information providing a distinct signature for psammite, quartzite, and monzogranite.

There is no doubt that lichen imparts a similarity to the spectra throughout the VNIR and SWIR range by decreasing reflectance and in part suppressing diagnostic reflectance peaks and troughs (*see* Fig. 16). In both the spectra for psammite (Fig. 16a) and monzogranite (Fig. 16b) derived from field samples measured using an ASD spectrometer in the lab (internal light source), reflection decreases with increased lichen cover and spectral features (troughs and peaks) are suppressed especially in the psammite spectra; however, in this study, enough variability in the spectra as a function of different mineralogy and associated weathering styles (particularly colour) was present to successfully discriminate one major lithological group (metatonalite) and three compositional units (psammite, quartzite, and monzogranite) that comprise a mapped metasedimentary lithological unit as well as vegetation which, in this area only, is a good proxy for metagabbroic rocks (*see* Harris et al., 2005).

The supervised approach arguably provides the 'best' results (Fig. 7, 8a) and this is not unexpected as this approach benefits from fieldwork and supporting geochemical and thin-section analysis (Harris et al., 2005); however, supporting fieldwork may not always be possible especially when providing a spectral map to support the initial stages of fieldwork (i.e. a first year of a mapping project). In this case a supervised approach, which involved the inspection of existing (albeit often old and out-of-date) geological maps in concert with analysis of remotely sensed data (e.g. Landsat, airphotos) to define potential end members (training areas), may provide the best alternative for producing a spectral map which gives some indication of lithological variability in a given study area.

The lithological maps produced from particularly the unsupervised approach (as opposed to the SMACC method) (Fig. 12, 13) provide reasonable renditions of the mapped lithology and as discussed above offer more detail on specific lithological packages than what has been mapped. The unsupervised classification approach may be favoured as the 'best' automatic method as it gave better results when comparing the derived class map (Fig. 9c) to the check areas; however, much more analysis on other hyperspectral data sets in varying geological environments is required to unequivocally verify this result.

Each method has its advantages and disadvantages. The supervised approach although not requiring field analysis, does benefit from it by allowing for the fine-tuning of representative training areas (end members), the selection of check areas and for the geological calibration of the resulting maps. The automatic approaches are quick and are easily implemented although the associated uncertainty in such maps may be high. They do, however, produce maps which provide a first 'geological look' at a given area, assist in focusing

fieldwork, and/or provide stand-alone maps for areas where no fieldwork is possible. Both approaches afford the opportunity to collect representative spectra of mixed rocks that can be stored in a spectral library and applied to adjacent areas or areas with similar geology and surficial cover.

CONCLUSIONS

Hyperspectral remote sensing holds promise for assisting in producing spectral maps that can be directly related to different rock types. These spectral maps are useful for focusing field mapping in areas with complex geology and also as a data source for producing stand-alone maps that can serve as first-order lithological maps in remote areas that cannot be mapped in detail.

Two methods (supervised and automatic) have been presented to produce a spectral map. For all methods the implementation of masking and noise-reduction techniques is recommended to improve classification results. The supervised method produced, arguably the 'best' results, but requires field input. Spectral maps that are almost as accurate as the supervised map, when compared to ground-truth information, are produced using both automatic methods (auto-end-member selection algorithm and unsupervised classification). The spectra and maps produced using unsupervised classification are in general more similar to the spectra and maps produced using the supervised approach. The maps produced using either method provide valuable geological information in frontier areas such as Canada's north.

ACKNOWLEDGMENTS

This research was funded by the Earth Science Sector's Remote Predictive Mapping Project (RPM) part of the Northern Resource Development program. David Viljoen provided constructive criticism that has resulted in a clearer presentation of this work.

REFERENCES

Berk, A., Bernstein, L.S., and Robertson, D.C.
1989: MODTRAN: A moderate resolution model for LOWTRAN7, Final Report, GL-TR-122, Air Force Geophysics Laboratory, Hanscom Air Force Base, Maryland, 42 p.

Bowles, J., Daniel, M., Grossman, J., Antoniades, J., Baumback, M., and Palmadesso, P.
1998: Comparison of output from ORASIS and pixel purity calculations; *in* Proceedings SPIE Imaging Spectroscopy IV, v. 3428, p. 148–156.

Budkewitsch, P., Staenz, K., Neville, R.A., Rencz, A.N., and Sangster, D.
2000: Spectral signatures of carbonate rocks surrounding the Nanisivik MVT Zn-Pb mine and implications of hyperspectral imaging for exploration in Arctic environments; *in* Ore Deposit Workshop: New Ideas for a New Millenium, Technical Volume, east Kootenay Chamber of Mines, Cranbrook, British Columbia, May 5–6, 2000, p. 1–11.

Clarke, R.N., Gallagher, A., and Swayze, G.A.
1990: Material absorption and band depth mapping of imaging spectrometer data using a complete band shape least-squares fit with library spectra; *in* Proceedings of the Second Airborne Visible/infrared Imaging Spectrometer (AVRIS) Workshop, (ed.) R.O. Green, National Aeronautics and Space Administration–Jet Propulsion Laboratory Publication 90-54, p. 176–186.

Goetz, A.F.H., Rowan, L.C., and Kingston, M.J.
1982: Mineral identification from orbit: initial results from the shuttle multispectral infrared radiometer; Science, v. 218, p. 1020–1031.

Green, A.A., Berman, M., Switzer, P., and Craig, M.D.
1988: A transformation for ordering multispectral data in terms of image quality with implications for noise removal; *in* Institute of Electrical and Electronic Engineers, Inc. Transactions on Geoscience and Remote Sensing, v. 26, no. 1, p. 65–74.

Gruninger, J., Ratkowski, A.J., and Hoke, M.L.
2004: The sequential maximum angle convex cone (SMACC) endmember model; *in* Proceedings, International Society for Optical Engineering, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X; v. 5425-1, Orlando, Florida, p. 1–14.

Harris, J.R., Eddy, B., Rencz, A.N., De Kemp, E., Budkewitsch, P., and Peshko, M.
2001: Remote sensing as a geological mapping tool in the Arctic: preliminary results from Baffin Island, Nunavut; Geological Survey of Canada, Current Research 2001-E12, p. 1–9.

Harris, J.R., Ponomarev, P., Shang, S., and Rogge, D.
in press: Noise reduction and best band selection techniques for improving classification results using hyperspectral data: application to lithological mapping in Canada's Arctic; Canadian Journal of Remote Sensing.

Harris, J.R., Rencz, A.N., Ballantyne, B., and Sheridon, C.
1998: Mapping altered rocks using Landsat TM and lithogeochemical data: Sulphurets-Brucejack Lake district, British Columbia, Canada; Photogrammetric Engineering and Remote Sensing, v. 64, no. 4, p. 309–322.

Harris, J.R., Rogge, D., Hitchcock, R., Ijewliw, O., and Wright, D.
2005: Mapping lithology in Canada's Arctic: application of hyperspectral data using the minimum noise Fraction (MNF) transform and matched filtering; Canadian Journal of Earth Sciences, v. 42, no. 12, p. 2173–2193.

Hunt, G.R.
1977: Spectral signatures of particulate minerals in the visible and near-infrared; Geophysics, v. 42, p. 501–513.

Hunt, G.R. and Salisbury, J.W.
1970: Visible and near-infrared spectra of mineral and rocks, I. Silicate minerals; Modern Geology, v. 1, p. 283–300.

1971: Visible and near-infrared spectra of minerals and rocks, II. Carbonates; Modern Geology, v. 2, p. 22–30.

Hunt, G.R., Salisbury, J.W., and Lenhoff, C.J.
1973a: Visible and near-infrared spectra of minerals and rocks, VI. Additional silicates; Modern Geology, v. 4, p. 85–106.
1973b: Visible and near-infrared spectra of minerals and rocks, VI, Acid igneous rocks; Modern Geology, v. 4, p. 217–224.
1974: Visible and near-infrared spectra of minerals and rocks, IX. Basic and ultra-basic igneous rocks; Modern Geology, v. 4, p. 15–22.

Neville, R.A., Staenz, K., Szerdi, T., Lefebvre, J., and Hauff, P.
1999: Automatic end member extraction from hyperspectral data for mineral explorational; *in* Proceedings of the 4th International Airborne Remote Sensing Conference and Exhibition–21st Canadian Symposium on Remote Sensing, v. II, p. 891–897.

Neville, R.A., Sun, L., and Staenz, K.
2003: Detection of spectral line curvature in imaging spectrometer data; *in* International Society for Optical Engineering Proceedings: Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery IX , (ed.) S.S. Shen, and P.E. Lewis, v. 5093, Bellingham, Washington State, 5093, p. 144–154.

Secker, J., Staenz, K., Gauthier, R.P., and Budkewitsch, P.
2001: Calibration of hyperspectral sensors in operational environments; Remote Sensing of Environment, v. 76, issue 1, p. 81–92.

St-Onge, M.R., Scott, D.J., and Wodicka, N.
1998a: Geology, Hidden Bay, District of Franklin, Northwest Territories; Geological Survey of Canada, Open File 3536, scale 1:100 000.
1998b: Geology, McKellar Bay, District of Franklin, Northwest Territories; Geological Survey of Canada, Open File 3537, scale 1:100 000.
2001: Geology, Meta Incognita Peninsula, Baffin Island, Nunavut; Geological Survey of Canada, Map 2009A, scale 1:100 000.

Staenz, K. and Williams, D.J.
1997: Retrieval of surface reflectance from hyperspectral data using a look-up table approach; Canadian Journal of Remote Sensing, v. 23, p. 354–368.

Tou, J.T. and Gonzales, R.C.
1974: Pattern Recognition Principals; Addison-Wesley Publishing Co., Inc., London, United Kingdom, 377 p.

Van Der Meer, F. and De Jong, S.M.
2001: Imaging Spectrometry: Basic Principles and Prospective Applications; Kluwer Academic Publishers, London, United Kingdom, 403 p.

Winter, M.E.
1999: N-FINDR: An algorithm for fast autonomous spectral end member determination in hyperspectral data; *in* Proceedings International Society for Optical Engineering Imaging Spectroscopy, v. 3753, p. 266–275.

Geological Survey of Canada Project Y03