Geochemistry Geophysics Geosystems G3

Published by AGU and the Geochemical Society

Article
Volume 13, Number 3
24 March 2012
Q03021, doi:10.1029/2011GC004004
ISSN: 1525-2027

Thermal infrared spectroscopy and partial least squares regression to determine mineral modes of granitoid rocks

Christoph Hecker
Department of Earth Systems Analysis, Faculty of Geo-Information Science and Earth Observation, University of Twente, PO Box 217, NL-7500 AE Enschede, Netherlands (hecker@itc.nl)

John H. Dilles
College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331–5506, USA

Mark van der Meijde and Freek D. van der Meer
Department of Earth Systems Analysis, Faculty of Geo-Information Science and Earth Observation, University of Twente, PO Box 217, NL-7500 AE Enschede, Netherlands

[1] In this paper, we present an approach to extracting mineralogic information from thermal infrared (TIR) spectra that is not based on an input library of pure mineral spectra nor tries to extract spectral end-members from the data. Instead, existing modal mineralogy for a number of samples are used to build a partial least squares regression (PLSR) model that links the mineralogy of the samples to their respective TIR spectral signatures. The resulting PLSR models can be applied to a larger group of samples for which the mineralogic composition can be estimated from the TIR spectra alone. Thermal infrared reflectance spectra were recorded from 1330–625 cm^{-1} (7.5 to 16.0 μm). The method is tested on igneous rocks from a porphyry copper deposit in Yerington, Nevada. As a reference, modal mineralogic composition was determined with traditional polarization microscopy on thin sections. Partial least squares regression models were developed to link the thermal infrared spectra to the thin section determined mineral modes of alkali feldspar, plagioclase and quartz, as well as the average plagioclase composition information. Results indicate that rock samples can be classified successfully in a quartz-alkali feldspar-plagioclase diagram based on thermal infrared spectroscopy and partial least squares regression modeling. Estimated errors for the mineralogic composition model results were found to be smaller or equal to traditional methods with errors of ±5.1% (absolute) for alkali feldspar, ±8.5% (absolute) for plagioclase and ±6.9% (absolute) for quartz. The regression model for plagioclase composition predicted with estimated errors of ±7.8 mol% anorthite.

Components: 9000 words, 5 figures, 1 table.

Keywords: partial least squares regression; quantitative mineralogy; remote sensing; spectroscopy; thermal infrared; thin section analysis.

Index Terms: 3934 Mineral Physics: Optical, infrared, and Raman spectroscopy.

Received 16 December 2011; **Revised** 8 February 2012; **Accepted** 8 February 2012; **Published** 24 March 2012.

Hecker, C., J. H. Dilles, M. van der Meijde, and F. D. van der Meer (2012), Thermal infrared spectroscopy and partial least squares regression to determine mineral modes of granitoid rocks, *Geochem. Geophys. Geosyst.*, *13*, Q03021, doi:10.1029/2011GC004004.

1. Introduction

[2] Optical mineralogy has been a successful and commonly used analytical method by petrologists for many decades. With the help of a petrographic microscope, thin sections of rock samples can be analyzed to reveal, among others, mineralogic composition, textures and overprinting relationships between minerals. Alternatively, infrared spectroscopy has been used in mineralogic studies [e.g., *Hunt and Turner*, 1953; *van Ruitenbeek et al.*, 2005]. Different analytical approaches measure the reflected, transmitted, absorbed or emitted portion of infrared radiation as a function of wavelength, which influence the information we can extract from the spectroscopic data [e.g., *Hecker et al.*, 2010; *Ramsey*, 2004]. Historically, reflectance measurements in the short-wave infrared region have received most of the attention. A number of important rock-forming silicate minerals (such as quartz and most feldspars), however, are indistinct in short-wave infrared spectra and show conspicuous spectral features only at longer wavelengths [*Salisbury et al.*, 1991]. Existing literature has shown that thermal infrared (TIR) spectra can be used successfully to identify rock-forming minerals in samples [e.g., *Salisbury and D'Aria*, 1992; *Vaughan et al.*, 2005] as well as to give quantitative information on minerals in a rock sample [e.g., *Feely and Christensen*, 1999; *Milam et al.*, 2007; *Ramsey and Christensen*, 1998; *Ruff*, 1998].

[3] Most existing studies [e.g., *Feely and Christensen*, 1999; *Gillespie*, 1992; *Milam et al.*, 2004; *Ramsey*, 2004; *Ramsey et al.*, 1999; *Thomson and Salisbury*, 1993] extract quantitative mineral abundance information from TIR spectra by means of a linear spectral mixture analysis (SMA). The fundamental assumption is that the measured rock spectrum is a linear combination of its components (minerals), weighted by each mineral's areal fraction. Typically, the fractions of each possible mineral are varied until the resulting rock spectrum shows the least root mean square error as compared to the measured rock spectrum. Earlier studies have shown that the assumption of linear mixtures is essentially true in the thermal infrared wavelength region for rocks or coarse particulate samples, where surface reflection dominates and volume scattering (energy passing through the grains) is minimal [e.g., *Feely and Christensen*, 1999; *Ramsey and Christensen*, 1998; *Thomson and Salisbury*, 1993]. To function properly and give reliable results, SMA algorithms require a database that contains spectra of all mineral components that are to be expected in the rocks to be analyzed. Missing end-members can either result in large root mean square errors or are partially (but incorrectly) compensated by linear combinations of other minerals [*Rogge et al.*, 2006]. Especially with complex mineral groups (like feldspars and amphiboles) it can be challenging to acquire pure sample material and representative spectra for all compositional and structural variations within the group. Furthermore, an over-complete spectral database (i.e., containing spectra that do not occur in the rock spectrum to be unmixed) can lead to an increased error of the abundance estimates similar to an incomplete database [*Rogge et al.*, 2006].

[4] In this paper, we present a different approach to extracting mineralogic information from TIR spectra that is not based on an input library of pure mineral spectra nor tries to extract spectral end-members from the [e.g., *Bandfield et al.*, 2000]. Instead, modal mineralogy information on a number of samples is used to build a quantitative multivariate partial least squares regression (PLSR) model that links the mineralogy of the samples to their respective TIR spectral signatures. The resulting PLSR model can be applied to another set of samples, for which only TIR spectra are available, and modal mineralogy can be estimated quickly and inexpensively from TIR spectra. The methodology has potential to a) partially replace expensive and time consuming laboratory analyses (such as thin section analysis or x-ray diffraction), b) conduct modal determination from TIR spectra directly in the field and c) be upscaled to airborne data for mapping of spatial mineralogic patterns on TIR imagery. Here we assess the methodology on a petrographically well-defined sample data set from Yerington, Nevada [*Dilles*, 1984; *Dilles and Einaudi*, 1992] and test how well thermal infrared reflectance spectroscopy and PLSR can estimate modal percentages of the main rock forming minerals in samples.

2. Methodology

2.1. Modal Mineralogy

[5] Modal mineralogy from optical microscopy was chosen as the reference method to compare the PLSR results to. This petrographic information (see Data Set S1 in the auxiliary material) stems from *Dilles* [1984] and *Dilles and Einaudi* [1992] and

was collected on igneous samples (ranging from gabbroic and andesitic to granitic) from a porphyry copper deposit in Yerington, Nevada.[1] Hydrothermal alteration in the area created distinct zones of potassic, sodic-calcic, quartz-sericite and tourmaline alteration. Areas close to neighboring sedimentary rocks show skarn alteration. Alkali feldspar, plagioclase and quartz (Afsp, Plg, Qtz) modes were determined in percent of the thin section area (areal%) with traditional polarization microscopy [*Dilles*, 1984]. These areal% data were not converted to typical volume% used for mineral modes *sensu stricto* since spectroscopy is sensitive to areal% as well. Error estimates for the modal mineralogy are similar to those typically quoted in literature, with accuracies of ±5–15% for major minerals (≥10% of area) and ±<5% for minor minerals (<10% of area) [*Feely and Christensen*, 1999; *Hamilton and Christensen*, 2000]. In altered samples, for each mineral phase the percentage of alteration as well as the alteration product was taken into account. This ensures that the tabular modal descriptions represent the current mineral composition of the samples (after alteration) and gives the best possible basis for comparison to TIR spectra.

[6] Plagioclase composition (Plgcomp) was determined using the a-normal method [*Deer et al.*, 1992] on several grains per thin section, and confirmed with electron microprobe on selected samples, with corresponding errors approximately ±1–2 mol% and ±1 mol%, respectively. In samples with zoned plagioclase, compositions were estimated on several spots from core to rim of individual grains and the results were averaged.

2.2. Thermal Infrared Spectroscopy

[7] Since coverslips ruled out spectroscopy directly on the thin sections, the corresponding offcuts from the thin section production were used in the TIR reflectance spectroscopy assuming that mineral modes were similar. The 105 sample blocks were prepared for the spectral measurements by grinding the measurement surface with 80 grit silicon carbide. This removed remnants of epoxy glue, sodium cobaltinitrite staining and saw markings from the surface to be measured. The blocks were washed with distilled water and left to air dry for a day.

[8] The thermal infrared reflectance spectra were recorded with the Bruker Vertex 70 FTIR spectrometer situated at the University of Twente, Faculty of ITC (UT-ITC). In order to measure spectra that are quantitatively comparable to remote sensing emission imagery [*Hecker et al.*, 2010; *Korb et al.*, 1999; *Salisbury et al.*, 1994], the standard laboratory instrument was customized to allow for directional-hemispherical reflectance measurements of large rock samples. A custom-built, diffuse gold-coated integrating sphere of 150 mm diameter was attached to an external port of the instrument. The beam path was modified such that it leaves the spectrometer through a connecting funnel and enters the equatorial plane of the sphere. A folding mirror placed slightly off the sphere's center deflects the near-collimated beam downward and through the 30 mm sampling port at the sphere's south pole onto the sample material. The incidence angle onto the sample is 10 degrees off-normal and all reflection directions are integrated by the sphere and measured by the Mercury-Cadmium-Tellurium detector situated near the north pole of the sphere. The folding mirror in the center of the sphere works as a baffle and prevents the first specular reflection from entering the detector. The entire spectrometer, funnel and integrating sphere assembly was purged with N_2 gas to suppress CO_2 and H_2O features in the resulting spectra. For more details on the UT-ITC instrument setup and data reduction to absolute reflectance spectra, see *Hecker et al.* [2011].

[9] Spectra were recorded from 1330–625 cm^{-1} (7.5 to 16.0 μm) at a resolution of 4 cm^{-1}. A Labsphere Infragold diffuse gold standard was used as a reference before each new sample. For the reference as well as sample measurement, 512 scans were co-added. Additionally, each sample was measured a total of eight times (without moving the sample) and averaged again to improve signal-to-noise ratio. Repeatability of the FTIR measurements is high for most of the spectral range with standard deviations of the eight repeat measurements per sample around 0.0015 reflectance units (roughly 0.6% of a typical reflectance value of 0.25). The standard deviation increases to 0.0080 toward the far end of the wavelength range where the limits of the detector sensitivity range are reached. Total measurement time for a reference and eight sample measurements was around 30 min (including purge delays of 2 min after placing the gold standard or sample under the integrating sphere).

[10] For this particular study a 20 mm port reducer with the same diffuse gold coating was mounted to the sampling port to prevent incident energy from leaking past the edges of the sample blocks. This

[1] Auxiliary material data sets are available at ftp://ftp.agu.org/apend/gc/2011gc004004. Other auxiliary material files are in the HTML. doi:10.1029/2011GC004004.

reduces the measurement spot from about 25 mm to 20 mm diameter on the sample surface (about a third of the entire offcut's surface area). The central part of each block was measured except for very heterogeneous samples, where the most representative part of the block was selected for measurement. For two samples, duplicate measurements were done at the beginning and the end of the measurement series (spanning several months) to test for possible drift in detector sensitivity as well as for sampling errors (i.e., slightly different measuring spot on the sample surface). The difference in the duplicates usually stays within 0.01 reflectance units, with an occasional value up to 0.02.

2.3. Partial Least Squares Regression

[11] Partial least squares regression (PLSR) is a quantitative multivariate statistical tool that allows for the analysis of data with strong correlations and with noise [*Wold et al.*, 2001]. In a first step, a PLSR model is built, using a training set of samples for which the spectral information as well as the modal mineralogy is known. In the second step, the resulting PLSR model can be applied to new samples for which only spectra are available, and the modal mineralogy is modeled from the TIR spectra alone. In this study four PLSR models were built to predict alkali feldspar, plagioclase and quartz modes, as well as the average plagioclase composition in the samples.

[12] Contrary to the more general multiple linear regression model, PLSR can also handle data sets with more variables than samples. Hence, it is especially useful for spectroscopic data sets that contain reflectance values at hundreds to thousands of wavelengths. While originally developed for the field of chemometrics, PLSR has been applied to a number of spectroscopic studies in diverse application fields such as vegetation studies [e.g., *Asner and Martin*, 2008], soil mechanics [*Yitagesu et al.*, 2009] and land degradation [*Farifteh et al.*, 2007]. *Goetz et al.* [2009] used PLSR to build a predictive mineral model from near and short-wave infrared spectra recorded on gangue material on a conveyor belt. [*Cudahy et al.*, 2001, 2009] applied PLSR for the first time to thermal infrared spectra of rock samples and compared them to X-ray diffraction d-spacing results.

[13] PLSR creates a regression model that uses a set of predictor variables **X** (in our case the TIR spectra) to predict the occurrence and concentration of a set of response variables **Y** (in our case the modal mineralogy of the samples from the thin section analysis). Similar to a principal component analysis (PCA), the high dimensional **X** matrix is reduced to a few factors or latent variables by a projection to an orthogonal system of small dimensionality. The main difference being that in a PCA, the variance in **X** is maximized while in a PLSR the covariance between **X** and **Y** is maximized [*Esbensen*, 2006]. This causes the first few factors to contain the spectral content that is most representative and predictive of the **Y** values while higher number factors contain spectral content that is either not related to the particular predicted **Y** or contains noise. In this paper we used the nonlinear iterative partial least squares (NIPALS) algorithm as implemented in The Unscrambler software [*Camo Software*, 2010]. For a more detailed introduction to PLSR theory and algorithms, the reader is referred to excellent existing literature [e.g., *Esbensen*, 2006; *Geladi and Kowalski*, 1986; *Wold et al.*, 2001].

2.3.1. Pre-processing

[14] A typical pre-processing chain for PLSR input data consists of data transformation, mean-centering and scaling [*Wold et al.*, 2001]. While PLSR does not have any particular prerequisites in terms of variable distribution [*Chin*, 1998], a symmetric histogram is preferred [*Wold et al.*, 2001]. For this particular data set, the response variables (mineral modes) were transformed with arcsine (modes of alkali feldspar and quartz as well as the composition of the plagioclase) or power transformation (plagioclase modes) to improve the symmetry of their histograms. To improve numerical stability and facilitate interpretation [*Wold et al.*, 2001], predictor and response variables were mean-centered by subtracting the average of the individual variable [*Geladi and Kowalski*, 1986]. Scaling of variables is important if they do not share the same value ranges. In those cases, variables with small value ranges will be under-represented in the final model. For this data set, no scaling was applied since all variables are already expressed in percentages and cover, therefore, the same value range. Not scaling the spectral variables has two further advantages: first, the resulting regression coefficients are directly interpretable in terms of spectral features of minerals and second, it prevents measurement noise in low reflective parts of the spectrum from influencing the model results.

2.3.2. Validation

[15] During the building of the PLSR model (i.e., projection of the TIR spectra into orthogonal factors), the main spectral information that is

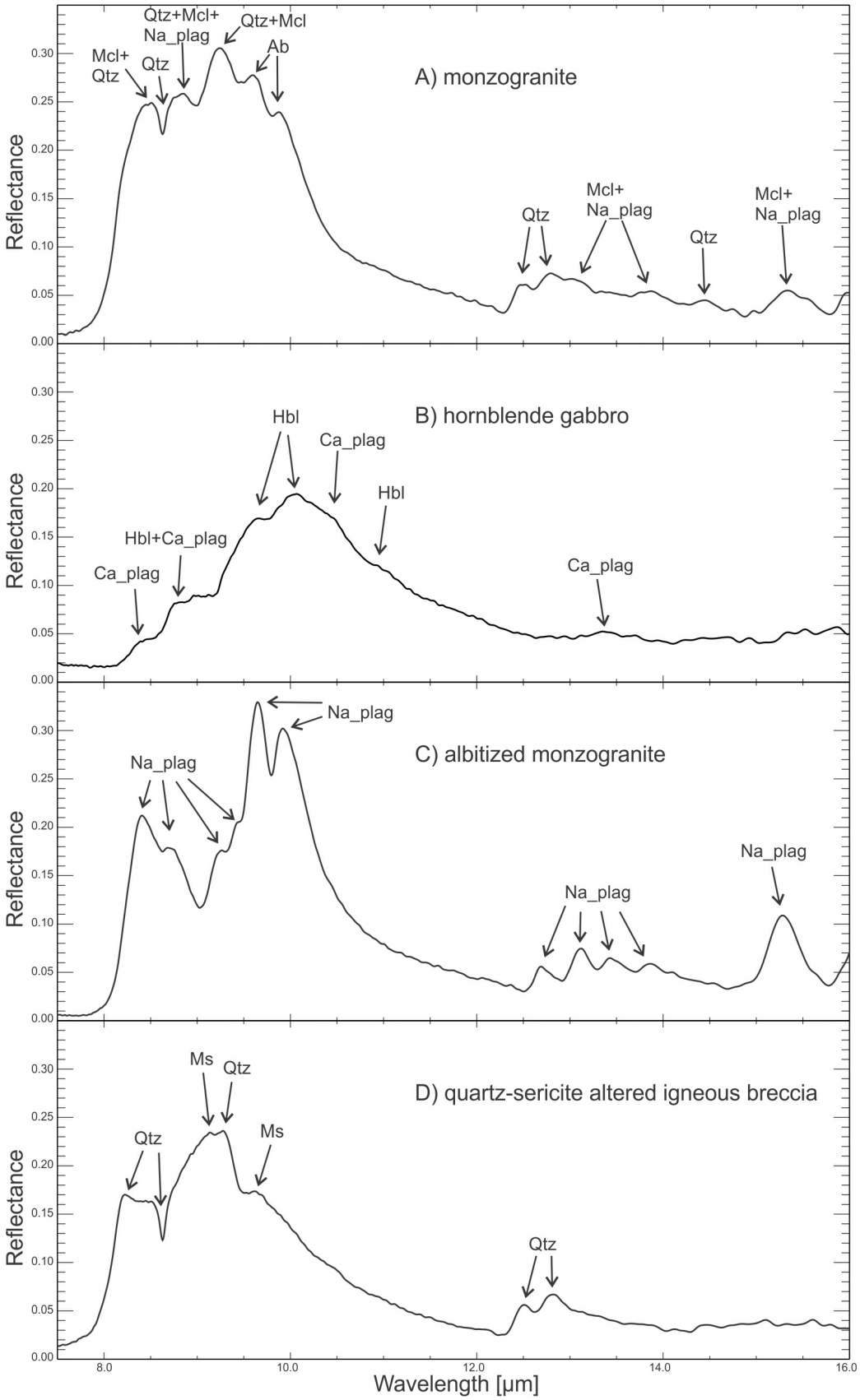
A) monzogranite
Mcl+ Qtz
Qtz
Qtz+Mcl+ Na_plag
Qtz+Mcl
Ab
Qtz
Mcl+ Na_plag
Qtz
Mcl+ Na_plag
B) hornblende gabbro
Ca_plag
Hbl+Ca_plag
Hbl
Ca_plag
Hbl
Ca_plag
C) albitized monzogranite
Na_plag
Na_plag
Na_plag
Na_plag
Na_plag
D) quartz-sericite altered igneous breccia
Qtz
Ms
Qtz
Ms
Qtz
Reflectance
Wavelength [μm]
0.00
0.05
0.10
0.15
0.20
0.25
0.30
8.0
10.0
12.0
14.0
16.0

Figure 1

explaining the mineral modes is compressed in the first few factors, while later factors contain mainly noise. The validation of the PLSR model has two main purposes: a) to separate the factors that contain useful information from those that are consisting mainly of noise and b) to supply a quantitative measure of the expected error in the model. This quantitative measure of the overall "goodness of fit" is represented by a root mean square error of prediction (RMSEP) that is defined as

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_{i,ref})^2}{n}}$$

where $\hat{y}_i$ is the modeled response variable, $y_{i,ref}$ the true measured response variable and n is the number of samples.

[16] RMSEP can either be calculated on an entirely separate validation data set (not included during the model building), or by cross validation, an iterative process that withholds each time a number of samples from of the model building set. The limited number of samples (105, plus two duplicates) and the level of uncertainty in the thin section data make the use of an independent validation data set (e.g., 1/3 of all samples not available for model building) undesirable. Instead, a segmented cross validation was used, where samples were split up into 10 segments. The samples were first ordered by the variable to be predicted and successively one-for-one distributed into the 10 segments. This technique ensures that each segment contains an equal number of samples with low, medium and high values of a particular response variable. The two duplicates were each assigned to the same group as their respective original measurement, to prevent the same sample from occurring in the model building as well as validation set.

[17] The PLSR model is built 10 times, each time a different segment is kept out of the model building set and used for validation. RMSEP values are calculated for each segment and averaged to form the overall RMSEP of the model.

2.3.3. Number of Factors

[18] The selection of the proper number of factors is imperative for a good regression model: too small and the model will not be able to account for the complexity of the mixed samples, too large and the model will over-fit the data set, as it begins to model the noise rather than the data itself. As a guideline, the number of factors that creates the best prediction (i.e., smallest RMSEP) was considered optimal (Figures S1, S3, S5, and S7 in Text S1). If the predicted error only decreased minimally when including an extra factor, we decided to not include the extra factor, as a model with less factors is more robust when applied to new data [*Esbensen*, 2006]. In the case of the plagioclase composition model, the RMSEP plot had a local minimum at five and a global minimum at eight factors. The model using eight factors was observably overfitting the calibration data: the loading weights started to become very unspecific and noisy, and the regression coefficients stopped showing spectral features characteristic of the mineral being modeled (for examples of loading weights and regression coefficients see the results section). In this particular case, the local minimum at five factors in the RMSEP plot was used as the cutoff.

2.3.4. Prediction

[19] As the final stage in the PLSR modeling, the previously built models are used to predict sample fractions for the modeled minerals from the TIR spectroscopy. For that purpose the PLSR model's regression coefficients are applied to thermal infrared reflectance spectra with unknown mineral modes, where the predicted variable

$$\hat{y} = B_0 + B_1X_1 + B_2X_2 + \ldots + B_nX_n$$

where $\hat{y}_i$ is the predicted response variable (e.g., quartz modes), B_1 to B_n the regression coefficients at all wavelengths, X_1 to X_n the reflectance values at all wavelengths and B_0 the regression coefficient bias. In a final step, the inverse of the earlier data transformations are applied to revert the prediction results to compositional information in the original units of areal percentages.

3. Results

3.1. TIR Spectra

[20] The igneous rocks studied here are typically intermediate in acidity and contain two feldspars, quartz, and sometimes amphiboles or pyroxenes

Figure 1. Thermal infrared spectra from a selection of samples as an illustration of their spectral content. The major spectral features are labeled with the mineral that causes it. Qtz = quartz, Mcl = microcline, Na_plag = sodic plagioclase, Hbl = hornblende, Ca_plag = calcic plagioclase, Ms = muscovite.

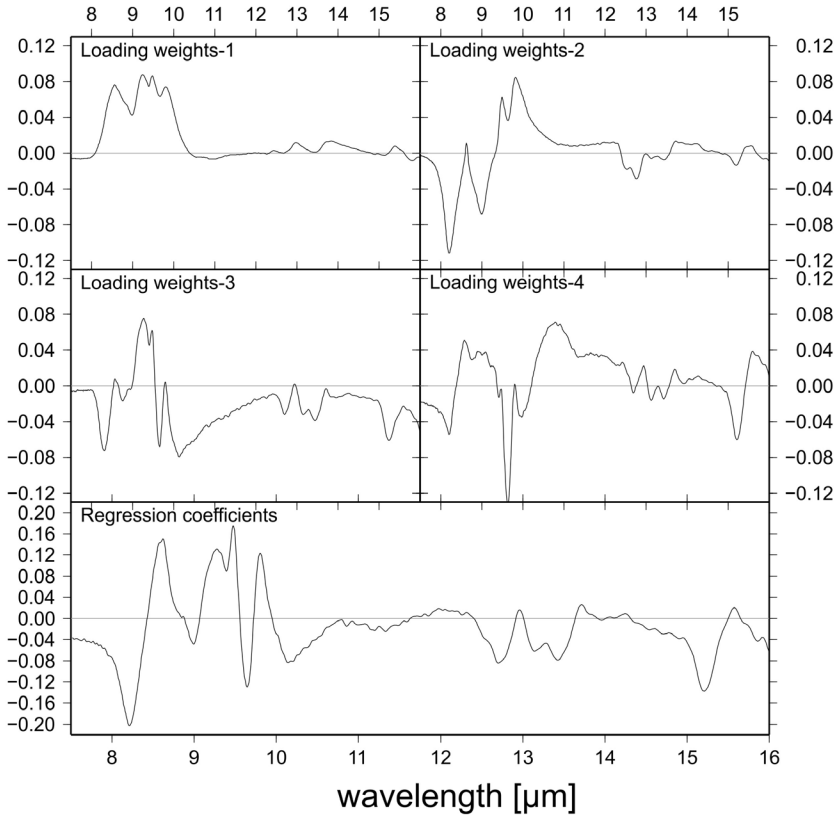

Figure 2. Loading weights (factors 1 to 4) and regression coefficients of the alkali feldspar model displayed as functions of wavelength. Wavelengths with regression coefficient values far from zero are influential for PLSR model.

that have features in the TIR rock spectra. Additionally, alteration minerals that occur frequently are sericite, secondary alkali and plagioclase feldspar, epidote and chlorite. Figure 1 shows examples of TIR rock spectra from this study with the most diagnostic features in the reststrahlen band (about 8–12 μm). The range of 12.5–14.0 μm contains additional features of feldspar (Figure 1c) and quartz (Figure 1d). Figure 1c is an example in which all TIR active minerals have been albitized by alteration and the TIR spectrum only shows features of Na-plagioclase. Most other samples, however, are strongly mixed spectra with overlapping features of several minerals in the reststrahlen band (e.g., Figures 1a and 1b) that can make mineral identification difficult and quantification impossible without multivariate modeling or mixing analysis.

3.2. PLSR Model Outputs

[21] The main outputs of a PLSR model are the RMSEP plot, the loading weights and the regression coefficient spectra (Figure 2 and Text S1). We will illustrate the use and the interpretation of each of these output data sets on the example of the alkali feldspar model. The plagioclase, quartz and plagioclase composition models are explained in detail in Text S1.

3.2.1. RMSEP Plot

[22] The RMSEP plot shows the predicted mean error (from the cross validation analysis) as a function of the number of factors used. For the alkali feldspar model, the predicted error increases again after a global minimum at six factors, as the model starts fitting noise in the data (Figure S1 in Text S1). Since the gain in RMSEP between a model of five and six factors is minimal, we chose to use the more robust five factor model.

3.2.2. Loading Weights

[23] Since our predictor variables are spectral data, the resulting loading weights and regression coefficients can be plotted and interpreted as spectra as well. The loading weights of a factor show which wavelengths are important in defining that particular factor. Through spectral interpretation of the loading weights, we can understand which minerals

and mineral combinations drive the particular PLSR model factor and, therefore, have a positive or negative link with the predicted variable.

[24] Under a negative link we understand that inverse features of a particular mineral reduce the modeled amount in case those features are present. This is best explained in the case of the simple quartz model (Text S1): If only factor 1 was used, samples with large albite content would have modeled quartz modes that are too high (due to overlaps in the quartz and albite reststrahlen features). The second factor is used to compensate for this effect and will reduce the quartz estimates in samples that are rich in albite, and increase the quartz estimates if samples are poor in albite. Hence, loading weights as well as regression coefficients (see below) can contain positive as well as inverse features of minerals to come up with an optimal modal estimate for the predicted mineral.

[25] For the alkali feldspar model, the loading weights of factor 1 (Figure 2) are strongly dominated by microcline, with reststrahlen features at 8.56, 9.24, 9.48 and 9.81 (all wavelengths are in units of μm). There are also features at 12.98, 13.37 and 15.35 that are caused by microcline. Loading weights of factor 2 show a negative link with a quartz spectrum and a positive link with a couple of microcline features at 9.49 and 9.82. Loading weights of factor 3 show a negative link with a few albite features at 9.78 and 15.25, and a weak positive link with microcline features at 8.57, 8.76, 9.28, 9.47, 9.66, 12.69, 13.45. The 8.31 feature could not be assigned to a mineral. The loading weights of factor 4 are dominated by albite. The features show a negative link with albite at 8.58, 9.64, 9.81, 9.95, as well as the four features between 12.5 and 14 and the feature at 15.21. Furthermore, a broad feature at 10.78 indicates a link with more Ca-rich plagioclase, such as labradorite or bytownite. The loading weights of factor 5 (not displayed here) show clear structure but the features could not be assigned to a particular mineral spectrum.

3.2.3. Regression Coefficients

[26] Regression coefficients summarize the relationship between the predictor variables and the response variable. While the loading weights show the influential wavelengths that define a particular factor, the regression coefficients show the influential wavelengths for the entire PLSR model. Wavelengths where the regression coefficients are close to zero have a negligible influence on the prediction of a particular response variable, while large positive and large negative values indicate wavelengths that have a strong positive and negative link with the response variable in question. Applying the regression coefficients to thermal infrared spectra with unknown mineral modes results in a predicted value of that particular mineral.

[27] The regression coefficient spectrum of the alkali feldspar model combines the information from the loading weights of factors 1 to 5. Most spectral features are linked positively to spectral features of microcline (8.62, 8.99, 9.28, 9.40, 9.47; all wavelengths in units of μm). The features at 8.21 and 8.62 are negatively linked to quartz, and features at 12.69, 13.12, 13.43 and 15.21 are negatively linked to albite. At the 10.78 wavelength, typical for Ca-rich plagioclase, the coefficients are nearly zero, thus showing that this wavelength region is not relevant in determining alkali feldspar modes. No evidence of other minerals was detected in the regression coefficients of the alkali feldspar model.

[28] In a similar fashion, the regression coefficients of the other three models were analyzed (see Text S1 for details). The plagioclase model is guided by a combination of different albite features, as well as a Ca- plagioclase feature. Quartz and microcline features are included to a lesser degree in predicting the plagioclase amount in the sample.

[29] The quartz model is a very simple model that contains only two factors. The regression coefficients also represent the simplicity of this model, which is driven mainly by the presence of quartz features and to a lesser degree by the absence of albite features.

[30] The plagioclase composition model is the most complex model. The local minima in the RMSEP plot show that the projection of the predictor spectra to latent variables is not yet ideally predicting the response variable, and that the data set possibly contains one or several outliers. Since all reasonable outliers (e.g., vague data description, uncertain mineral identification in thin section study) had been removed from the data set earlier, no unjustifiable outlier removal, to artificially improve model performance, was implemented at this stage. As a consequence, the loading weight spectrum of the plagioclase composition model contains an accumulation of many mineral features and possibly noise. From the identifiable features, it appears that the plagioclase spectrum is mainly driven by sodic and calcic plagioclase as well as quartz content.

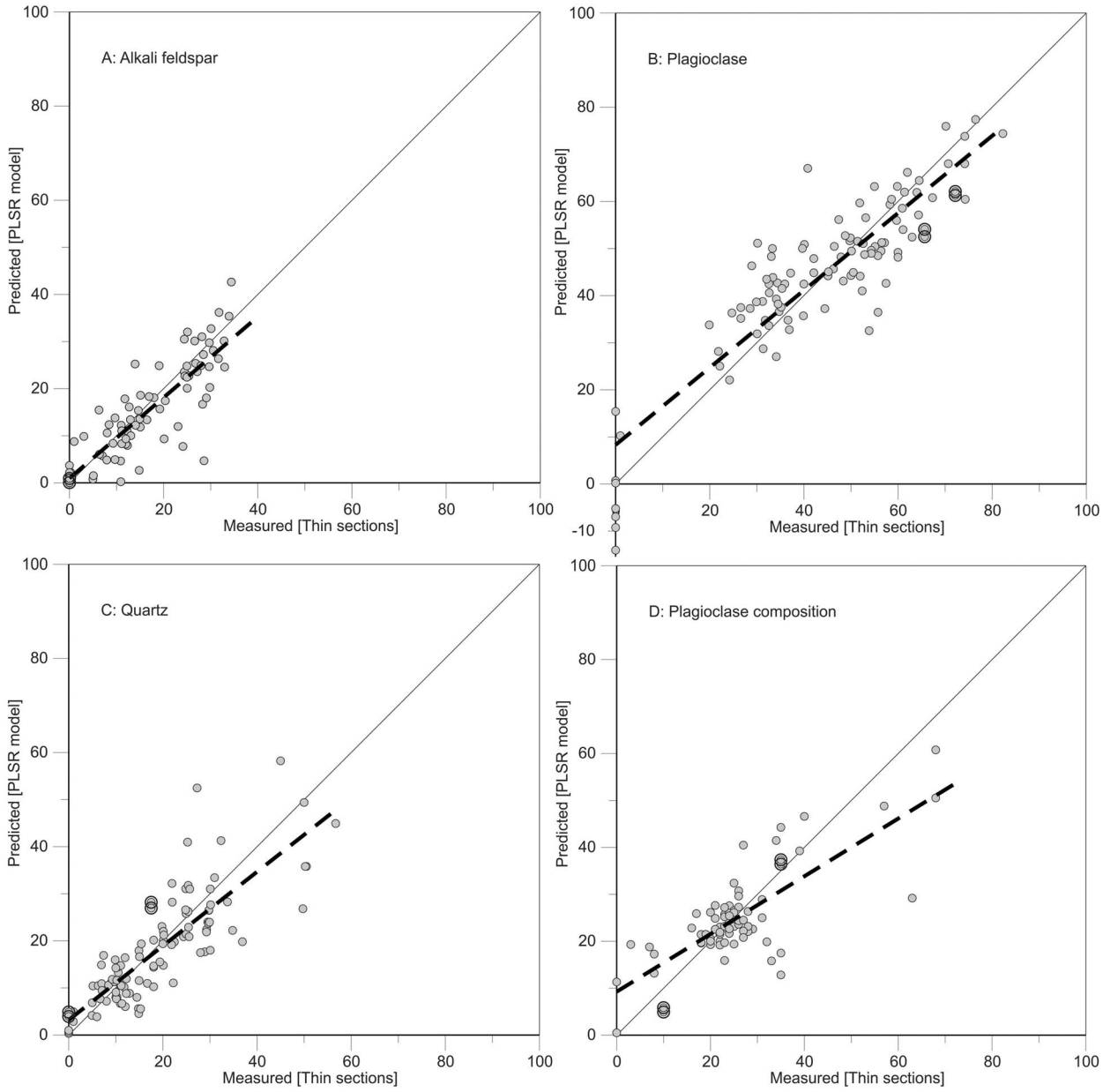

Figure 3. PLSR modeling results predicted (from spectroscopy) versus measured (from thin section) in areal percentages within the samples. Dashed line is regression line between measured and predicted values. Thin solid line represents the 1:1 line. Open circles indicate duplicate samples.

3.3. Measured Versus Predicted

[31] After applying the regression coefficients to the spectral measurements and inverting the preprocessing transformation, the TIR spectrum of each sample predicts alkali feldspar, plagioclase and quartz modes, as well as a plagioclase composition. The PLSR modeling prediction results show the modeled percentage of a mineral in the TIR spectrum of the sample compared to the measured composition from the thin section analysis (Figure 3). Performance parameters of the various models are summarized in Table 1. The alkali feldspar (Afsp) model (Figure 3a) works well ($R^2 = 0.81$). It does contain, however, a small group of samples that contain 20–30% Afsp in the thin section measurements but only 5–15% in the PLSR model results. They do not show any common lithology or alteration patterns that could explain this discrepancy and for this reason they were kept in the model as is. The RMSEP for the Afsp model is ±5.1% absolute, which is equivalent to the lower end of the error estimate of the reference method.

Table 1. Summary of PLSR Modeling Results per Mineral

	Mineral[a]			
	Afsp	Plg	Qtz	Plgcomp
Number of factors used	5	4	2	5
RMSEP [in % abs.]	5.1	8.5	6.9	7.8
R^2	0.81	0.80	0.70	0.59
Slope (of regr. line)	0.86	0.82	0.79	0.61

[a]Mineral abbreviations used are as follows: Afsp = alkali feldspar; Plg = plagioclase; Qtz = quartz; Plgcomp = plagioclase composition.

[32] The plagioclase model (Figure 3b) performs well (R^2 = 0.80). Samples with very low to no plagioclase show a larger spread in their model results than the rest of the samples. The RMSEP of the plagioclase prediction is ±8.5% absolute and can be attributed to the value range of the plagioclase percentages.

[33] The quartz model (Figure 3c) has an R^2 of 0.70. The prediction residuals show that the variance increases with increasing quartz content. This result can (at least partially) be attributed to the thin section reference method, which has a growing estimated error with increases in quartz content. The RMSEP for quartz is ±6.9% absolute for the entire model. However, samples with less than about 30% quartz (in our case most igneous samples without the influence of quartz-sericite alteration) model better with a RMSEP of ±5.8% absolute.

[34] The relationship between the measured versus predicted plagioclase compositions (Figure 3d) has an R^2 of 0.59, which is weaker than for the previous models. We suspect that the reference values from the thin section analysis overestimate the average anorthite content of plagioclase for samples with >An_{40} due to the common procedure to measure core, intermediate zone and rim. The intermediate zones (and rims) dominate the area, but with most samples showing normally zoned plagioclase crystals (i.e., with calcic cores), averaging the measurements overestimates the calcic component as compared to spectroscopic results. The RMSEP of the plagioclase composition prediction has an estimated error of about ±8 mol % anorthite. There are clusters of samples that plot well off the regression line: a cluster at measured An_{38} (that plots above the line) represents hornblende gabbroic rocks and other rocks with a distinct lack of Afsp. The cause for this is not understood and can also not be identified in the loading weights of the plagioclase model. A second cluster at a measured An_{33} (that plots below the regression line) represents samples that have been slightly altered to chlorite, epidote ± sericite. The main spectral feature of chlorite around 9.86 μm could easily interfere with the oligoclase-feature at the same wavelength. It has to be noted, however, that no quantitative, consistent link could be made between the residuals of the plagioclase composition prediction and the amount of chlorite (or epidote/sericite for that matter) in the sample.

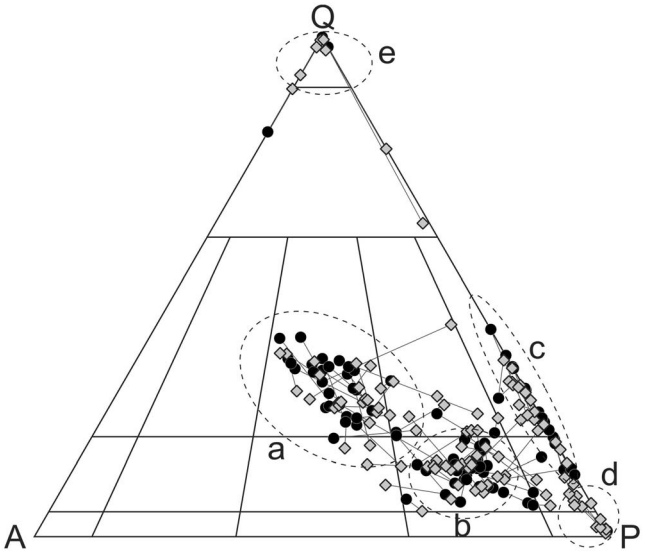

Figure 4. Quartz-alkali feldspar-plagioclase diagram with thin section measurements (black dots) connected by a line to the corresponding composition predicted from the spectroscopic data (gray diamonds). Clusters a to e are described in the text. Triangle subdivisions are according to *Le Maitre* [1989]. For field labels see Figure 5.

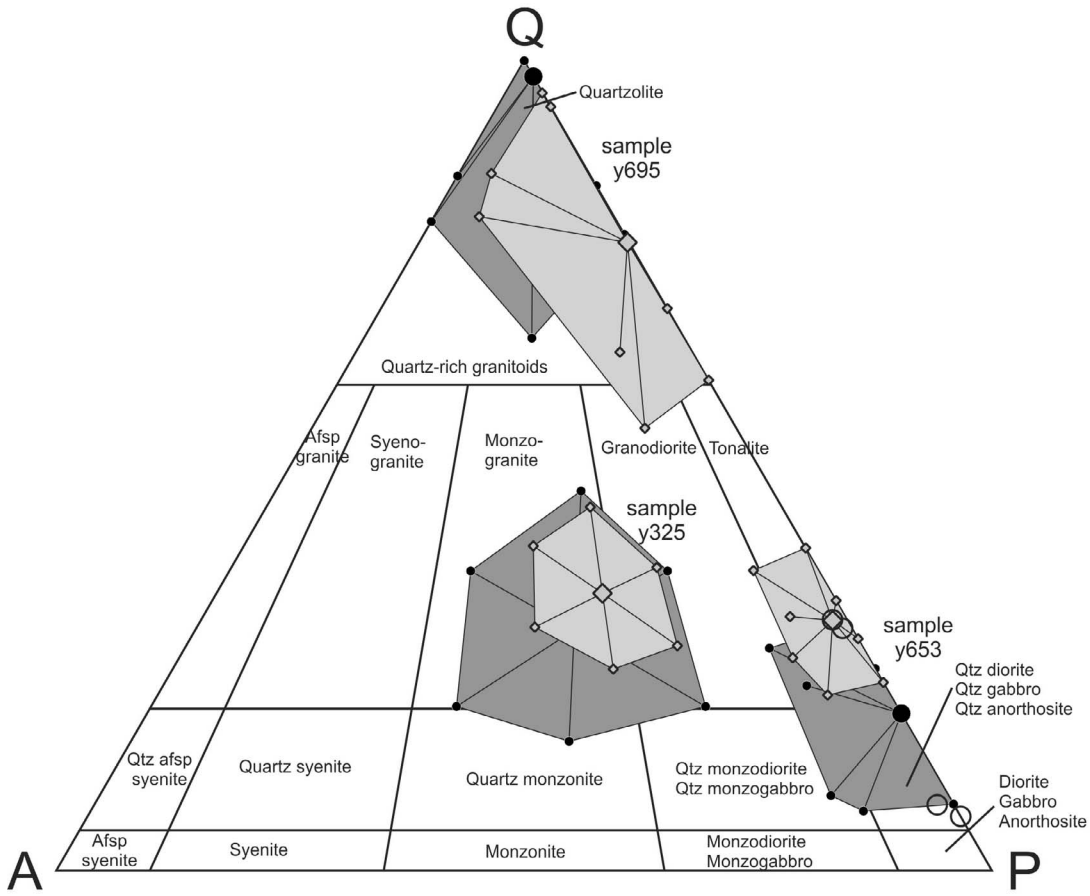

Figure 5. Spatial display of uncertainties on one random sample from each of the clusters a, c and e in the quartz-alkali feldspar-plagioclase diagram. The large symbols are the thin section measurements (black dots) and the spectroscopic predictions (gray diamonds). The smaller symbols and the shaded areas indicate the corresponding error regions. Open circles indicate locations of duplicate samples. Triangle subdivisions and field labels are according to *Le Maitre* [1989].

3.4. Ternary Plot

[35] The modal mineralogy from thin section measurements and PLSR modeling on spectroscopic data are represented in a QAP (quartz-alkali feldspar-plagioclase) plot for plutonic rocks (Figure 4) [*Le Maitre*, 1989]. The modal mineral percentages of the three minerals were normalized to 100% before plotting. Although not strictly applicable to altered igneous rocks, we plot both the altered and unaltered rocks on the QAP diagram to compare the results of both methods. Contrary to the recommendations of *Le Maitre* [1989], the extreme sodic end-member of the plagioclase group (An_0-An_5) was kept with the plagioclase count rather than summing them up with the alkali feldspar modes since the sodic plagioclases form part of a continuous albite-oligoclase mixture sequence. Five clusters of samples are discernible in the QAP diagram: a) the monzogranite group, b) the quartz monzodiorite group, c) the tonalite to quartz diorite group, d) the gabbros and e) a cluster in the top third of the QAP diagram containing samples rich in quartz and very poor in feldspars. Those of cluster e) have undergone quartz-sericite (±pyrite, ±tourmaline) alteration, where the feldspars are mainly replaced by fine-grained sericite. The same five clusters are also visible in the PLSR predicted compositions from the spectroscopic data (Figure 4). Only a few predictions (under 8%) do cross into another cluster. This demonstrates that we can use the combination of TIR spectroscopy and PLSR modeling to classify rock samples in a QAP ternary diagram.

3.5. Error Estimates

[36] In order to illustrate the order of magnitude of the QAP plot uncertainties, ternary error regions were plotted for the reference method as well as for the PLSR prediction (Figure 5). The error regions

show how their spatial extent changes with location in the QAP plot. One random sample each was selected from clusters a, c and e. For the reference method uncertainty, the average of the quoted error estimates (±10% for major minerals and ±5% for minor minerals) were added and subtracted from the original areal percentages. These error-modified points were normalized to 100%, added to a QAP diagram and the convex hull of the resulting points was used as the ternary error region. For the uncertainties in the spectroscopic predictions, the quartz, alkali feldspar and plagioclase model RMSEPs (Table 1) were added and subtracted from the respective original predicted values, normalized to 100% and the convex hull of the resulting points was used as the ternary error region.

[37] Sample y325 with roughly similar proportions of quartz, alkali feldspar and plagioclase shows a relatively small and isotropic prediction uncertainty field (Figure 5). Other samples (e.g., y695) have modal values along the edge of the QAP diagram or close to a corner. As negative values are not physically possible, the uncertainty fields are truncated at the edge of the triangle and become anisotropic in shape. Samples with Q + A + P totaling much less than 100% prior to normalization (e.g., y695 Q + A + P = 52%) show larger error fields as small errors of 6–8% result in large error regions after the normalization. Figure 5 shows that in all three cases the measured and predicted uncertainty fields overlap and are not significantly different from each other.

[38] To illustrate the reproducibility of the method, the duplicate samples are highlighted by open circles in Figures 3 and 5. In all cases the PLSR predictions based on the duplicate measurements plot very close to the originals and always within a 2% absolute margin. Since the duplicate measurements were taken several months later and on a different spot on the sample, the small deviations show that error propagation through sampling error and instrument noise and drift contributes less than 2% absolute to the overall uncertainty.

4. Discussion

[39] Table 1 and Figure 3 summarize some of the key parameters of the four PLSR models applied in this study and Text S1 provides some additional information. The loading weights and regression coefficients show that the models are driven by the most important mineral components in which we are interested. No spectral features of unexpected or accessory minerals were detected in the models, which implies that the bases for the models are robust. The quartz model is based on only two factors due to its distinct spectral features with no variations due to composition. Using only two factors also greatly reduces the risk of overfitting the data, which is important when the model is applied to another data set or study area of interest. The alkali feldspar and plagioclase models use five and four factors, respectively. Many of the factors show similar features in the loading weights but with different relative feature depths and slight shifts in wavelengths. These factors are required to account for changes in composition and mineralogy in the alkali feldspar and plagioclase group and are not a sign of redundancy or overfitting. The cross-validation results and the clear structure in all loading weights give confidence that these 4–5 factor models are also robust and applicable to other data sets. Since "plagioclase composition" does not have a spectral signature as such, and depends on the amount and compositions of the different plagioclase minerals in the rock, the plgcomp model is more complex and less straight forward than the other three. The RMSEP plot (Figure S7) shows local minima, which are a sign of remaining outliers in the data. While the loading weights of the plagioclase composition model are clearly driven by spectral features of various plagioclase compositions, the regression coefficients show less structure than in other models and display some noise.

[40] Figure 2 shows that the reference method and the predicted values from spectroscopy have discrepancies but the relative contributions of reference method and model toward these errors are irresolvable. Ideally, we would like to confirm the results with a reference method that has a smaller uncertainty and is less dependent on specialist interpretation than thin section analysis. X-ray diffraction (XRD) is generally a qualitative process that identifies minerals in the sample, but not their amounts. Quantitative XRD strives to give quantitative mineralogic information but uncertainties with thin section analysis is often cited around 10% [*Amaral et al.*, 2006] which does not represent a significant enough reduction in reference method uncertainty as compared to the ±5–15% for major minerals in thin section analysis. X-ray fluorescence (XRF) and scanning electron microprobe in turn give quantitative chemical information on a sample or on a small spot on a sample, respectively. Both require a model (based on order of mineral crystallization from melt) to get from oxide chemistry to normative mineralogy. Many of the samples

in the study area have undergone hydrothermal alteration by different fluids at different temperatures and intensities. They do not necessarily follow the normative models. Hence, the methods mentioned above would not significantly improve the uncertainties in the reference method.

[41] Three samples (sample numbers y538, y645 and y697b) of the monzogranite cluster (Figure 4, cluster a) show large differences between thin section and PLSR prediction. In two cases (y645 and y697b) the thin section observation included a fine-grained matrix, which may have resulted in misidentification of minerals. Two quartz-sericite (±pyrite, ±tourmaline) altered samples (y695 and y574) show large QAP differences with the thin section results (Figure 4, cluster e). Also in the case of these two samples, the discrepancies can partially be attributed to the fine-grained nature of the samples, which could hamper the identification of remaining feldspar relicts under the microscope. An additional source of increased uncertainty for these samples is inherent to the modeling method itself: A PLSR model developed on a sample set with many typical and only a few extreme compositions has difficulties in predicting the extremes of the value range. This problem is evident in low plagioclase (Figure 3b) and high quartz (Figure 3c) samples that show large spreads in the measured versus predicted plots. The problematic samples are in both cases the quartz-sericite altered rocks. Since most (non-altered) samples do not contain any substantial amounts of white mica, the PLS models do not include any sericite features in the loading weights and appear to have problems modeling those samples properly. To increase the prediction accuracy it can be considered to model pervasively altered rocks separately in the future.

[42] In order to put the modeled accuracies of this study into context, we compare it to those reported with the commonly applied spectral mixture analysis (or unmixing): *Feely and Christensen* [1999] reported typical unmixing uncertainties for major minerals in igneous rock samples of ±5–10% absolute. For those results all types of feldspars were considered together. The results for alkali feldspar and plagioclase alone were considerably less reliable and were not specified. *Wyatt et al.* [2001] determined plagioclase modes to within 5–10% absolute. In comparison, the results presented here perform well (Afsp ± 5.1%, Plg ± 8.5%, Qtz ± 6.9% all absolute). *Milam et al.* [2004, 2007] deconvolved TIR spectra of igneous rocks, coarse plagioclase sand, and polymineralic sand mixtures and modeled plagioclase compositions to within 6 mol% An. By comparison, plagioclase compositions modeled in this study had uncertainties of 7.8 mol% An.

[43] While results are comparable between PLSR and SMA, the two techniques have thoroughly different approaches and requirements. As mentioned in the introduction, the challenge of SMA is the a priori definition of a set of representative reference spectra for all expected mineralogic compositional and structural variations in the rocks to be analyzed. Its strength, on the other hand, is that if a set of reference spectra can be defined accurately, SMA can be applied to different study areas without any additional field information available (e.g., planetary sciences). In contrast to SMA, PLSR does not link sample spectra to a spectral end-member reference library. Instead, existing modal mineralogy for a number of samples are used to train the model to extract the required information automatically from the thermal infrared spectra. While not required for building the model, a spectral library can be very helpful when interpreting the internal structures of a PLSR model: by comparing loading weights and regression coefficients to mineral spectra, one can analyze what minerals and mineral combinations drive a particular PLSR model.

[44] PLSR is especially useful for many ongoing projects that do have existing information on composition for a number of samples on hand, which can be used to train the PLSR models and get modal estimates from TIR spectra on the rest of the samples. The most prospective sources of compositional information are mineral modes by thin section analysis or QXRD, since features in TIR spectroscopy are coupled to the mineralogic compounds in rocks. Alternatively, one could also attempt the use of whole rock or trace element geochemistry results as long as their concentrations are correlated with the mineral modes present in the rocks.

[45] The minimum number of samples required to build a reliable PLSR models based on spectroscopy is 20–50 well defined samples [*Esbensen*, 2006]. For mineralogic data sets with considerable uncertainties most authors prefer a slightly larger sample set of 50–100 samples. A prime example of where PLSR could be used on existing data sets would be a mining operation: data from different analytical methods for samples from diverse parts of the deposit exist and can be used to link TIR spectra to the existing data sets. While mineral modes are the TIR spectroscopy could be used at

the mine face, on drill cores or in sorting operations to determine mineralogy and possibly ore content. However, care has to be taken when PLSR models are transported from one study area to another. The models and the estimated prediction errors (RMSEP) remain valid as long as the new samples come from the same population, i.e., they have comparable mineralogic compositions. If that is not the case, prediction errors will increase and the model should be re-calibrated using samples from the new study area.

[46] Thermal infrared spectra give us information on the average mineralogic composition of the sample, while thin section analysis with optical methods or electron microprobe can reveal information on compositional zonation of grains or whether a mineral was stable in a given assemblage. To combine spatial information with spectral information, drill cores or flat samples can be imaged in the shortwave infrared by hyperspectral sample imagers, such as SisuRock (SisuRock Hyperspectral Core Imaging Station, Spectral Imaging Ltd, 2011, available at http://www.specim.fi/media/product-brochures/sisurock-ver1-11.pdf) for which a thermal infrared option is currently being developed. Furthermore, thermal infrared spectroscopy has the capacity to be upscaled to routine measurements in the laboratory or field, or to airborne imagery. While the current measurement setup of directional-hemispherical reflectance was chosen for future comparison to airborne data, routine laboratory and field applications would benefit from the larger energy throughput of a bi-directional setup, which would reduce the measurement time to well under a minute per sample. A similarly operating system (bi-directional) has been developed by CSIRO Australia for routine drill core measurements and has recently been extended to TIR wavelengths [*Huntington et al.*, 2010]. Future work will have to demonstrate whether PLSR derived models, as described in this paper, can also be applied to airborne TIR imagery.

5. Conclusions

[47] Thermal infrared spectroscopy in combination with partial least squares modeling can be used to estimate the modal mineralogy of igneous rock samples. In this study, TIR spectra were linked to thin section-determined alkali feldspar, plagioclase and quartz contents, as well as plagioclase composition. Rock samples were successfully classified in a QAP diagram based on TIR spectroscopy and PLSR modeling. PLSR-modeled alkali feldspar modes are within a ±5.1% estimated absolute error of those determined by thin sections. The plagioclase model performs well for all samples except those that contain very low or no plagioclase at all. Estimated errors for all samples are ±8.5% absolute. Modeled quartz modes stay within ±5.8% absolute for samples containing less than 30% quartz. If samples with high quartz contents (i.e., quartz-sericite altered) are included in the model, the estimated error increases slightly but remains below ±7% absolute. For the plagioclase group, a PLSR model was built that predicts plagioclase composition with discrepancies of ±7.8 mol% An.

Acknowledgments

[48] The authors would like to thank Keith Milam, Steven Ruff and Benoit Rivard for their suggestions that greatly improved the clarity of this document.

References

Asner, G. P., and R. E. Martin (2008), Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels, *Remote Sens. Environ.*, *112*(10), 3958–3970.

Amaral, P. M., J. Cruz Fernandes, and L. Guerra Rosa (2006), A comparison between X-ray diffraction and petrography techniques used to determine the mineralogical composition of granite and comparable hard rocks, *Mater. Sci. Forum*, *514–516*, 1628–1632.

Bandfield, J. L., P. R. Christensen, and M. D. Smith (2000), Spectral data set factor analysis and end-member recovery: Application to analysis of Martian atmospheric particulates, *J. Geophys. Res.*, *105*(E4), 9573–9587, doi:10.1029/1999JE001094.

Camo Software (2010), The Unscrambler X: All-in-one multivariate data analysis and design of experiments package, version 10.0.1, software, Oslo.

Chin, W. W. (1998), The partial least squares approach to structural equation modeling, in *Modern Methods for Business Research*, edited by G. A. Marcoulides, pp. 295–336, Lawrence Erlbaum Assoc., Mahwah, N. J.

Cudahy, T. J., J. Wilson, R. Hewson, P. Linton, P. Harris, M. Sears, K. Okada, and J. A. Hackwell (2001), Mapping variations in plagioclase felspar mineralogy using airborne hyperspectral TIR imaging data, paper presented at International Geoscience and Remote Sensing Symposium, Inst. of Electr. and Electron. Eng., Sydney, N. S. W., Australia.

Cudahy, T. J., et al. (2009), Drill core logging of plagioclase feldspar composition and other minerals associated with Archean Gold Mineralization at Kambalda, Western Australia, using a bi-directional thermal infrared reflectance system, in *Remote Sensing and Spectral Geology*, edited by R. Bedell, A. P. Crosta, and E. Grunsky, pp. 223–235, Soc. of Econ. Geol., Littleton, Colo.

Deer, W. A., R. A. Howie, and J. Zussman (1992), *An Introduction to the Rock-Forming Minerals*, 696 pp., Addison Wesley Longman, Harlow, Essex.

Dilles, J. H. (1984), The petrology and geochemistry of the Yerington batholith and the Ann-Mason porphyry copper

deposit, western Nevada, 389 pp., Stanford Univ., Stanford, Calif.

Dilles, J. H., and M. T. Einaudi (1992), Wall-rock alteration and hydrothermal flow paths about the Ann-Mason porphyry copper deposit, Nevada—A 6 km vertical reconstruction, *Econ. Geol.*, *87*(8), 1963–2001, doi:10.2113/gsecongeo.87.8.1963.

Esbensen, K. H. (2006), *Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design*, 5th ed., 598 pp., Camo, Oslo.

Farifteh, J., F. van der Meer, C. Atzberger, and E. J. M. Carranza (2007), Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN), *Remote Sens. Environ.*, *110*(1), 59–78, doi:10.1016/j.rse.2007.02.005.

Feely, K. C., and P. R. Christensen (1999), Quantitative compositional analysis using thermal emission spectroscopy: Application to igneous and metamorphic rocks, *J. Geophys. Res.*, *104*(E10), 24,195–24,210, doi:10.1029/1999JE001034.

Geladi, P., and B. R. Kowalski (1986), Partial least-squares regression: A tutorial, *Anal. Chim. Acta*, *185*, 1–17, doi:10.1016/0003-2670(86)80028-9.

Gillespie, A. R. (1992), Spectral mixture analysis of multispectral thermal infrared images, *Remote Sens. Environ.*, *42*(2), 137–145, doi:10.1016/0034-4257(92)90097-4.

Goetz, A. F. H., B. Curtiss, and D. A. Shiley (2009), Rapid gangue mineral concentration measurement over conveyors by NIR reflectance spectroscopy, *Miner. Eng.*, *22*(5), 490–499, doi:10.1016/j.mineng.2008.12.013.

Hamilton, V. E., and P. R. Christensen (2000), Determining the modal mineralogy of mafic and ultramafic igneous rocks using thermal emission spectroscopy, *J. Geophys. Res.*, *105*(E4), 9717–9733, doi:10.1029/1999JE001113.

Hecker, C., M. van der Meijde, and F. D. van der Meer (2010), Thermal infrared spectroscopy on feldspars—Successes, limitations and their implications for remote sensing, *Earth Sci. Rev.*, *103*(1–2), 60–70, doi:10.1016/j.earscirev.2010.07.005.

Hecker, C., S. J. Hook, M. van der Meijde, W. Bakker, H. van der Werff, H. Wilbrink, F. van Ruitenbeek, B. de Smeth, and F. van der Meer (2011), Thermal infrared spectrometer for earth science remote sensing applications—Instrument modifications and measurement procedures, *Sensors*, *11*(11), 10,981–10,999, doi:10.3390/s111110981.

Hunt, J. M., and D. S. Turner (1953), Determination of mineral constituents of rocks by infrared spectroscopy, *Anal. Chem.*, *25*(8), 1169–1174, doi:10.1021/ac60080a007.

Huntington, J., L. Whitbourn, P. Mason, M. Berman, and M. C. Schodlok (2010), HyLogging—Voluminous industrial-scale reflectance spectroscopy of the Earth's subsurface, paper presented at Art, Science and Applications of Reflectance Spectroscopy Symposium, ASD Inc., Boulder, Colo., 23–25 Feb.

Korb, A. R., J. W. Salisbury, and D. M. D'Aria (1999), Thermal-infrared remote sensing and Kirchhoff's law: 2. Field measurements, *J. Geophys. Res.*, *104*(B7), 15,339–15,350, doi:10.1029/97JB03537.

Le Maitre, R. W. (1989), *A Classification of Igneous Rocks and Glossary of Terms: Recommendations of the International Union of Geological Sciences, Subcommission on the Systematics of Igneous Rocks*, 193 pp., Blackwell, Oxford, U. K.

Milam, K. A., H. Y. McSween, V. E. Hamilton, J. M. Moersch, and P. R. Christensen (2004), Accuracy of plagioclase compositions from laboratory and Mars spacecraft thermal emission spectra, *J. Geophys. Res.*, *109*, E04001, doi:10.1029/2003JE002097.

Milam, K. A., H. Y. McSween, and P. R. Christensen (2007), Plagioclase compositions derived from thermal emission spectra of compositionally complex mixtures: Implications for Martian feldspar mineralogy, *J. Geophys. Res.*, *112*, E10005, doi:10.1029/2006JE002880.

Ramsey, M. S. (2004), Quantitative geological surface processes extracted from infrared spectroscopy and remote sensing, in *Infrared Spectroscopy in Geochemistry, Exploration Geochemistry and Remote Sensing*, edited by P. L. King, M. S. Ramsey, and G. A. Swayze, pp. 197–213, Mineral. Assoc. of Can., Quebec, Que., Canada.

Ramsey, M. S., and P. R. Christensen (1998), Mineral abundance determination: Quantitative deconvolution of thermal emission spectra, *J. Geophys. Res.*, *103*(B1), 577–596, doi:10.1029/97JB02784.

Ramsey, M. S., P. R. Christensen, N. Lancaster, and D. A. Howard (1999), Identification of sand sources and transport pathways at the Kelso Dunes, California, using thermal infrared remote sensing, *Geol. Soc. Am. Bull.*, *111*(5), 646–662, doi:10.1130/0016-7606(1999)111<0646:IOSSAT>2.3.CO;2.

Rogge, D. M., B. Rivard, J. Zhang, and J. Feng (2006), Iterative spectral unmixing for optimizing per-pixel endmember sets, *IEEE Trans. Geosci. Remote Sens.*, *44*(12), 3725–3736, doi:10.1109/TGRS.2006.881123.

Ruff, S. W. (1998), Quantitative thermal infrared emission spectroscopy applied to granitoid petrology, PhD thesis, 234 pp., Ariz. State Univ., Tempe.

Salisbury, J. W., and D. M. D'Aria (1992), Infrared (8–14 μm) remote sensing of soil particle size, *Remote Sens. Environ.*, *42*(2), 157–165, doi:10.1016/0034-4257(92)90099-6.

Salisbury, J. W., L. S. Walter, N. Vergo, and D. M. D'Aria (1991), *Infrared (2.1–25 μm) Spectra of Minerals*, Johns Hopkins Univ. Press, Baltimore, Md.

Salisbury, J. W., A. Wald, and D. M. Daria (1994), Thermal-infrared remote sensing and Kirchhoff's Law: 1. Laboratory measurements, *J. Geophys. Res.*, *99*(B6), 11,897–11,911, doi:10.1029/93JB03600.

Thomson, J. L., and J. W. Salisbury (1993), The midinfrared reflectance of mineral mixtures (7–14 μm), *Remote Sens. Environ.*, *45*(1), 1–13, doi:10.1016/0034-4257(93)90077-B.

van Ruitenbeek, F. J. A., T. Cudahy, M. Hale, and F. D. van der Meer (2005), Tracing fluid pathways in fossil hydrothermal systems with near-infrared spectroscopy, *Geology*, *33*(7), 597–600, doi:10.1130/G21375.1.

Vaughan, R. G., S. J. Hook, W. M. Calvin, and J. V. Taranik (2005), Surface mineral mapping at Steamboat Springs, Nevada, USA, with multi-wavelength thermal infrared images, *Remote Sens. Environ.*, *99*(1–2), 140–158, doi:10.1016/j.rse.2005.04.030.

Wold, S., M. Sjöström, and L. Eriksson (2001), PLS-regression: A basic tool of chemometrics, *Chemom. Intell. Lab. Syst.*, *58*(2), 109–130, doi:10.1016/S0169-7439(01)00155-1.

Wyatt, M. B., V. E. Hamilton, H. Y. McSween Jr., P. R. Christensen, and L. A. Taylor (2001), Analysis of terrestrial and Martian volcanic compositions using thermal emission spectroscopy: 1. Determination of mineralogy, chemistry, and classification strategies, *J. Geophys. Res.*, *106*(E7), 14,711–14,732, doi:10.1029/2000JE001356.

Yitagesu, F. A., F. van der Meer, H. van der Werff, and W. Zigterman (2009), Quantifying engineering parameters of expansive soils from their reflectance spectra, *Eng. Geol. Amsterdam*, *105*(3–4), 151–160, doi:10.1016/j.enggeo.2009.01.004.