Noise Simulation and Correction in Synthetic Airborne TIR Data for Mineral Quantification

Christoph Hecker, Dean Riley, Mark van der Meijde, and Freek D. van der Meer

***Abstract*—Rock-forming minerals (such as feldspar and quartz) can be identified and quantified from thermal infrared (TIR) laboratory spectroscopy using spectral models. This paper uses synthetic airborne TIR spectra to test whether the hyperspectral Spatially Enhanced Broadband Array Spectrograph System (SEBASS) would theoretically be able to detect quartz and feldspar minerals and quantitatively predict mineral modes in felsic igneous rocks. Data from a previous laboratory study were used to simulate TIR spectra with band locations and noise levels of the SEBASS sensor. The quantitative partial least squares regression (PLSR) models from that study were applied to newly created synthetic SEBASS data, and results were compared with the predictions from the previous study. Predicted compositions based on SEBASS band positions are nearly identical ($\rho = 0.995$) to those based on laboratory resolution. Results are still reliable [prediction errors within 0.4% (absolute)] to the original laboratory PLSR predictions when adding up to 1% noise (about five times the SEBASS noise level) to the synthetic data. Prediction errors rapidly increase when noise levels beyond 1% are used. These results show that SEBASS' spectral resolution, spectral coverage, and signal-to-noise levels are sufficient to quantitatively predict quartz and feldspar amounts, and feldspar compositions with models based on PLSR. Spectral distortions, such as reduced spectral contrast, tilts, and vertical shifts, must be compensated for before these quantitative models are applied. A mean and standard deviation (MASD) normalization is proposed using a set of ground data for compensating systematic errors that are common to all image pixels.**

***Index Terms*—Airborne, hyperspectral, mean and standard deviation (MASD) normalization, noise simulation, partial least squares regression (PLSR), quantitative predictions, Spatially Enhanced Broadband Array Spectrograph System (SEBASS), synthetic data, thermal infrared (TIR).**

I. Introduction

QUANTITATIVE imaging spectroscopy for mineral identification in the visible to short-wave infrared region of the electromagnetic spectrum is widespread through literature (e.g., see [1]–[6]), whereas quantitative mineralogical predictions from airborne imaging data in the thermal infrared (TIR) wavelength range (8–12 μm) are still uncommon (e.g., see [7]–[11]). In the laboratory domain, TIR spectroscopy is routinely applied. TIR spectra can be successfully used to identify rock-forming minerals [10], [12] and to give quantitative information on minerals in a rock sample [13]–[16]. In a recent study, we have demonstrated a quantitative prediction model that linked TIR laboratory spectroscopy with mineral modes of alkali feldspar, plagioclase, and quartz (Afsp, Plg, and Qtz, respectively), as well as plagioclase composition (Plgcomp) [17]. In that research, we used partial least squares regression (PLSR) to link laboratory spectroscopy measurements [18] with the compositional information derived from traditional thin-section analysis (point counting) to predict quantitative mineral composition in samples, with error margins similar to the reference method itself. In addition, TIR spectroscopy offers the possibility to upscale the models to airborne TIR data, unlike point counting in thin-section analysis. Consequently, results from laboratory studies can be extended toward the airborne hyperspectral TIR imaging domain, which can highlight spatially continuous mineral patterns in the landscape rather than discrete sampling points.

This paper uses laboratory TIR data and noise characteristics (additive noise and spectral distortions) to simulate airborne image spectra of the hyperspectral Spatially Enhanced Broadband Array Spectrograph System (SEBASS) to test whether SEBASS would be able to detect quartz and feldspar minerals and quantitatively predict modes and compositions in igneous rocks. For this purpose, the laboratory spectra from our earlier study [17] were spectrally resampled and deteriorated to mimic airborne SEBASS spectra. Then, the same quantitative prediction models (PLSR) as in the earlier study were applied, and the results were compared with the original predictions. In the process, a new signal normalization method was developed so that laboratory-derived retrieval algorithms can be successfully applied to airborne hyperspectral TIR data.

II. Data

A. Laboratory Data

This paper builds on the data that had been previously employed in [17], where PLSR was used to determine mineral modes of alkali feldspar, plagioclase, and quartz, as well as plagioclase composition from the laboratory TIR spectra. Directional–hemispherical TIR reflectance spectra were recorded with a customized Bruker Vertex 70 laboratory Fourier transform infrared (FTIR) spectrometer at a 4-cm^{-1} resolution from 1330 to 625 cm^{-1} (7.5–16.0 μm) [18]. One hundred five

Manuscript received November 25, 2014; revised April 2, 2015 and July 13, 2015; accepted September 3, 2015. Date of publication November 2, 2015; date of current version February 24, 2016.

C. Hecker, M. van der Meijde, and F. D. van der Meer are with the Department of Earth Systems Analysis, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, 7514 AE Enschede, The Netherlands (e-mail: c.a.hecker@utwente.nl).

D. Riley was with The Aerospace Corporation, Chantilly, VA 14745 USA, and also with SpecTIR, LLC, Fairfax, VA 22033 USA. He is now with Booz Allen Hamilton Inc., McLean, VA 22102 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2482386

samples of igneous rocks from a porphyry copper deposit in Yerington, NV, USA, were used, and the modal mineralogic composition results were validated with traditional polarization microscopy on thin sections. The spectral models based on the laboratory spectra performed with estimated errors to be smaller or equal to traditional methods, with errors of $\pm 5.1\%$ (absolute) for alkali feldspar, $\pm 8.5\%$ (absolute) for plagioclase, and $\pm 6.9\%$ (absolute) for quartz. The regression model for plagioclase composition predicted with estimated errors of ± 7.8 mol% anorthite.

B. SEBASS Sensor Characteristics

This paper simulates spectra of SEBASS to use in the spectral modeling of mineralogic composition. Synthetic SEBASS responses were used instead of actual SEBASS data to avoid issues with validating the results with ground data acquired at a different scale. Actual SEBASS data were merely used to determine realistic SNR values and to conceptually demonstrate the data normalization method (see Section IV-C). SEBASS is a hyperspectral TIR imaging instrument, which was built and operated by The Aerospace Corporation, Chantilly, VA, USA, and was first described by the work in [19]. The sensor and its data are becoming increasingly accessible to the general public, and there is a need to develop the necessary tools to work with these data. SEBASS is a push-broom imaging system with two detector arrays covering the atmospheric windows between 3.0 and 5.2 μm (midwave infrared), as well as those between 7.8 and 13.5 μm [TIR or long-wave infrared (LWIR)]. The two focal plane arrays have dimensions of 128×128 pixels, resulting in two datacubes with 128 spectral bands, a narrow swath width of 128 pixels, and a field of view of 7.8°. Since the minerals of interest (silicates) show their most diagnostic features in the Reststrahlen band (around 10 μm), only the LWIR SEBASS data are treated here. The SNR of the LWIR sensor is quoted in the range of $\approx$3000:1 [20] and $\approx$2000:1 [21], and noise levels (NEΔT) are in the region of 0.05 K [19]. To achieve these high-SNR and low-noise levels, the Si–As detector is cooled with liquid He to temperatures of 4 K. These high SNR numbers are only applicable, however, to the raw radiance data set. Since the majority of the radiance signal is caused by temperature differences rather than variations in emissivity, the SNR levels after temperature–emissivity separation are significantly lower. Additionally, residual atmospheric features in the data set cause a further deterioration of the actual target signal. Our own tests on homogeneous and spectrally flat targets (e.g., water bodies) show that noise levels in the processed emissivity data sets of actual SEBASS scenes are around 0.002 or 0.2% absolute emissivity percentages (see Fig. 1) that correspond to an equivalent SNR of $\approx$500:1.

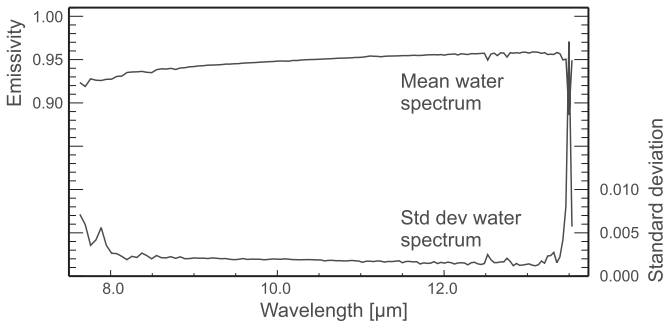

Fig. 1. SEBASS mean and standard deviation spectra of a homogeneous and spectrally flat target (the water body inside the flooded Yerington open cast mine). The statistics were calculated on a SEBASS *apparent emissivity* standard data product (Yerington Campaign on June 17, 2008, Flightline 20). The standard deviation, which is a proxy for the noise levels, is around 0.002 (or 0.2%) absolute emissivity values for most of the spectral range of the sensor.

III. Methods

A. Spectral Resampling

The high-resolution laboratory spectra were resampled to the SEBASS spectral resolution (hereon referred to as "SEBASS-resampled"). Since the SEBASS spectral response functions (the functions that define how each detector element reacts to incoming energy) are not public domain, they were approximated by a Gaussian function using a central wavelength and a full width at half maximum (FWHM) value. The central band positions were taken from the header of an actual SEBASS scene, whereas the extent of the bandwidth (i.e., FWHM) was approximated by the halfway points between adjacent bands. The spectrally resampled laboratory spectra were then recalculated from reflectance to emissivity using Kirchhoff's Law [22] in order to be visually more comparable with the spectra recorded by emissive SEBASS.

B. Spectral Distortions

Noise was added to the SEBASS-resampled spectra to simulate the reduced SNR in airborne data. Uniform random noise was produced with peak-to-peak values of ± 0.5, ± 1.0, ± 2.0, and $\pm 5.0\%$ (in absolute emissivity percentages) and then added to the SEBASS-resampled spectra. This process of adding noise was repeated (applying new random noise values to the original SEBASS-resampled spectra) to produce a second set of noise-added spectra. The first set of 105 spectra was used to build the PLSR model, whereas the second set of 105 spectra was used to validate the results.

Existing literature reports substantial shape differences between airborne SEBASS spectra and the TIR field or laboratory spectra of the same targets [8], [10]. Typical are reduced spectral contrast and drop-offs (spectral tilting) toward shorter wavelengths. These are caused by incomplete atmospheric correction, temperature–'emissivity separation algorithms that were developed for lower spectral resolution data, and sensor specific behavior. Part of the issues could be alleviated by using more sophisticated atmospheric correction algorithms (e.g., see [23]). In reality, these algorithms are rarely available to the general public; thus, we opted for the standard apparent emissivity product of The Aerospace Corporation.

The SEBASS-resampled spectra (without added noise) were subjected to a number of additive and multiplicative functions (see Table I) to assess the influence of the reported spectral distortions (i.e., reduced contrast and spectral tilt) on the quantitative prediction models. Each of the listed functions in Table I was chosen to simulate a particular spectral distortion

TABLE I
FUNCTIONS APPLIED TO SPECTRA TO SIMULATE ADDITIVE, MULTIPLICATIVE, AND TILT DISTORTIONS

Distortion	Function[a,b]	Comments
Additive	$y_{new} = y_{old} - 0.2$	Spectrum down-shifted by 0.20; spectral contrast kept same as in original
Multiplicative	$y_{new} = y_{old} \times 0.8$	Top edge of spectrum down-shifted by 0.20 and spectral contrast lowered to 80%
Additive and multiplicative	$y_{new} = (y_{old} \times 0.8) + 0.2$	Top edge of spectrum remains same as original; spectral contrast lowered to 80%
Additive tilting	$y_{new} = x - \left[1 - \left(\frac{LN(y_{old})}{3} + 0.13\right)\right]$	Shortest wavelengths down-tilted by ≈0.2; longest wavelengths same as original; spectral contrast kept same as original
Multiplicative tilting	$y_{new} = x\left[\frac{LN(y_{old})}{3} + 0.13\right]$	Shortest wavelengths down-tilted by ≈0.2; longest wavelengths same as original; spectral contrast lowered to ≈80% of original

[a] Emissivity spectra (y_{old}, y_{new}) in [0–1] value range
[b] Wavelength of x in [μm]

(and their combinations) that was reported in literature. Spectral contrasts of 80% of the original values and tilts to 80% of the absolute emissive values were assumed for this simulation because there is a lack of published quantitative information on the degree of distortion in SEBASS data. Each type of function, factor values, and constants were chosen to highlight a change in a particular scenario, i.e., a change in the absolute values or the spectral contrast, and/or introduce a spectral tilt. For details regarding the spectral distortions applied, see Table I.

TABLE II
SUMMARY OF LABORATORY PLSR MODELING RESULTS BEFORE DISTORTIONS (FROM THE WORK IN [17])

	Mineral			
	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

C. Partial Least Squares

PLSR is a quantitative multivariate statistical tool that can predict sample composition from a spectral signature. PLSR is commonly used with data sets that have strong correlations, noise, and more variables than samples. Therefore, it is particularly useful with spectroscopic data that contain strongly correlated spectral values at hundreds to thousands of wavelengths [24]. Although 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 [25], rock composition [17], soil composition and mechanics [26], [27], and land degradation [28]. PLSR is applied by first building a model using a training set of samples for which the spectral information (predictor variables) and the modal mineralogy (response variables) are known. In the second step, the resulting PLSR model is applied to new samples for which only spectra are available, and the modal mineralogy is modeled from the TIR spectra alone. For a more detailed introduction to PLSR theory and algorithms, see [24], [29], and [30].

In this paper, we used the identical data set and followed the same preprocessing procedures as we described in [17]. In summary, the work in [17] used TIR laboratory spectra as PLSR predictors, and PLSR models were built to predict alkali feldspar, plagioclase, and quartz modes, as well as the average plagioclase composition. The reference method for validating the composition was based on point counting petrographic thin-section analysis. The response variables (mineral modes) of Afsp, Qtz, and Plgcomp were transformed with an arcsine function ($x_{\text{new}} = \arcsin\sqrt{x}$), and Plg was transformed using a power transformation ($x_{\text{new}} = 10^x$). This improved the symmetry of their histograms, which is preferred in PLSR modeling [24]. After the laboratory spectra had been resampled and adjusted to mimic SEBASS spectra (see Section III-A and B), PLSR models were built on the distorted data sets and applied, as described in [17], with the following two exceptions: 1) the complexity of the PLSR model was not determined by the minimum in the root-mean-square error of prediction (RMSEP) plots, but the number of factors was fixed (see the first row in Table II) to allow for consistency and correlation between the different spectral treatments; and 2) the set of spectra with noise added was validated with a second set of noise-added spectra (same base spectra but different random noise added) to avoid overfitting in the model during the calibration stage. This procedure prevents the PLSR models from giving overly optimistic error estimates. The PLSR coefficients were applied to all spectra in the prediction stage, and the results were plotted against the values from thin-section analysis and the original laboratory PLSR results in [17]. Correlations between data sets were calculated using nonparametric Spearman's correlation since most variables show strongly asymmetric distributions, which make the use of a parametric Pearson's correlator inappropriate.

IV. RESULTS AND DISCUSSION

A. Changes in Spectral Shape

Due to the high spectral resolution of SEBASS, the SEBASS-resampled spectra look virtually identical to the original laboratory spectra (example of the quartz–oligoclase spectrum in Fig. 2, i.e., the first and second spectra from the top). The main difference between the two spectra is the reduced spectral

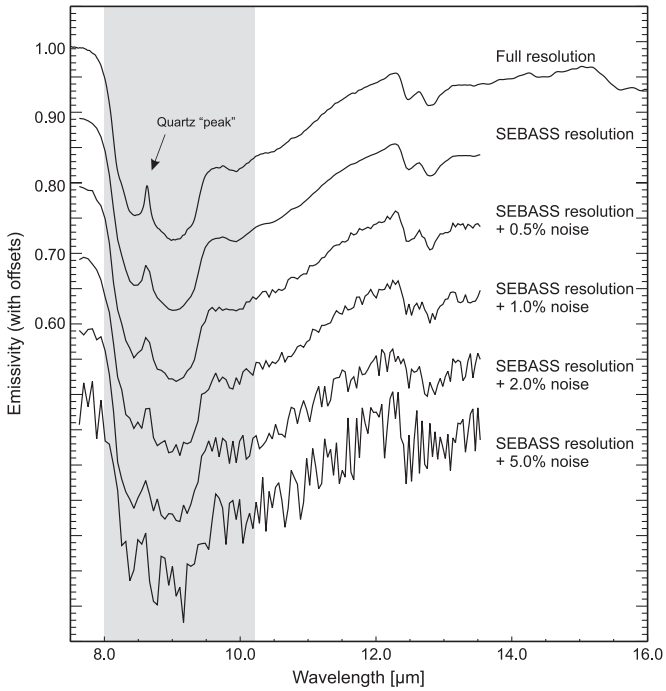

Fig. 2. (From top to bottom) Example of the TIR spectrum at full laboratory resolution, resampled to SEBASS band positions, and with progressively more noise added. Except for a full-resolution spectrum, all other spectra are downshifted by 0.10 emissivity units each for clarity. The spectral features in this sample (y653) are mainly due to quartz and plagioclase (for details, see the Auxiliary Dataset S1 in [17]). The zone in gray indicates the location of the main Reststrahlen features.

range of the airborne data, which covers the most important Reststrahlen features of the silicate minerals (see the gray zone in Fig. 2). Small differences can be also seen in areas with strong spectral changes (e.g., the central peak between the quartz doublets at 8.63 μm), whereas the SEBASS bandwidth is still able to resolve the feature; however, SEBASS' spectral resolution modifies the feature's shape.

As higher noise levels are progressively added, spectral features begin to deteriorate (see the bottom four spectra in Fig. 2). For up to 1% noise, both quartz doublets are still identifiable with confidence. At 2% noise levels, the quartz doublets merge to single features, but the (oligoclase) plateau at 9.6 μm is still discernible. At 5% noise levels, all individual spectral features have disappeared, except for a general Reststrahlen feature centered at about 9 μm.

The functions applied to the SEBASS-resampled spectra (described in Table I) simulate spectral distortions that have been noted in SEBASS data. The distortions modify the general shape of the spectra while leaving the spectral features intact (see Fig. 3). The main changes to the resampled spectra are: 1) the modification of the absolute emissivity values; 2) the reduction in the spectral contrast; and 3) the tilting of the spectrum toward shorter wavelengths.

B. PLSR Modeling Results

1) Additive Noise: The spectral resolution of SEBASS is sufficient to model the spectral features in the rock samples appropriately. The PLSR model predictions on the *SEBASS-resampled spectra* (with no noise added) are virtually identical to the predictions based on the original laboratory-resolution spectral data sets. As an example, the results of the alkali feldspar model are presented (see the leftmost plot in Fig. 4). Spearman's correlation between the SEBASS-resampled and original laboratory results is $\rho = 0.995$.

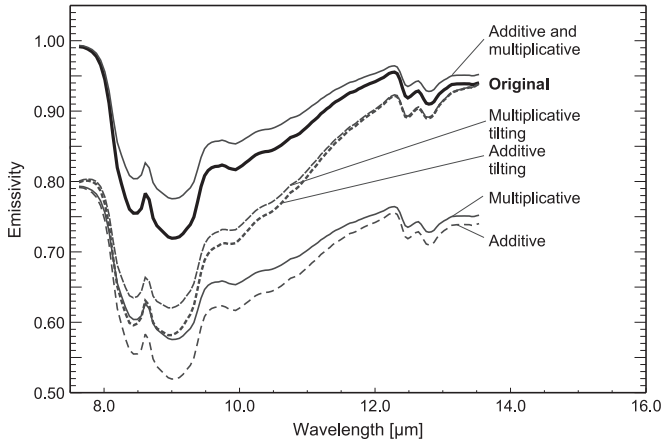

Fig. 3. Effects of the additive, multiplicative, and tilt distortions on the spectrum. Sample spectrum y653 is displayed as an example.

As expected for the *noise-added spectra*, the correlation drops with increasing noise levels: At 1% noise, the correlation is still almost 0.98. After that, the correlation rather quickly deteriorates, with $\rho \approx 0.91$ (2% noise) and $\rho \approx 0.76$ (5% noise, see the rightmost plot in Fig. 4).

To assess the influence of the added noise on the absolute prediction quality of the PLSR models (as compared with the thin-section reference method), RMSEP values were calculated for all mineral/noise level combinations. The results show that RMSEP values generally do not increase when the laboratory spectra were resampled to SEBASS wavelengths (see Fig. 5). Adding noise up to 1% slightly increases the prediction errors in the models (less than 0.4% absolute). At noise levels $\geq 2\%$, the RMSEP quickly and apparently increases in an exponential manner. The main exception is the quartz model as its prediction error levels stay almost constant, even at higher noise levels of 5%. This behavior can be attributed to quartz having strong spectral features that prevail in mixtures. As a result of this dominant spectral behavior in quartz, its PLSR model is very robust, which was built on only two factors that make the model less sensitive to noise.

Noise tests on actual SEBASS data showed that noise levels in the standard apparent emissivity product are around 0.2% of the absolute emissivity values (see Fig. 1). Therefore, the noise levels in the standard SEBASS apparent emissivity product are about five times lower than the 1% threshold, and the SNR is sufficiently high to model feldspar amounts and compositions with PLSR.

2) Spectral Distortions: In the case of the *spectrally distorted data*, PLSR prediction models were run on the distorted and original laboratory data, and the composition prediction results were compared in scatterplots (see Fig. 6) as follows.

- The PLSR quartz model predicts good relative and absolute values for all distorted spectra used in this paper

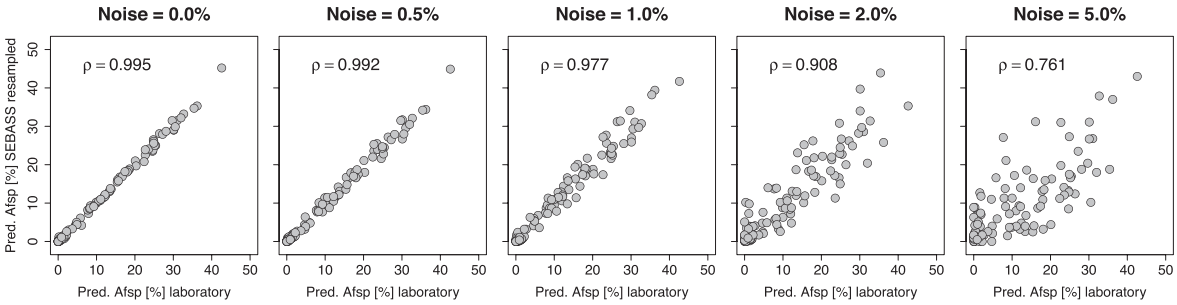

Fig. 4. Scatterplots comparing PLSR predictions (in areal percentages) for alkali feldspar based on the original laboratory spectra against those from SEBASS-resampled noise-added spectra. (From left to right) Noise levels: no noise, 0.5%, 1%, 2%, and 5%. ρ indicates Spearman's correlation factor of the two variables.

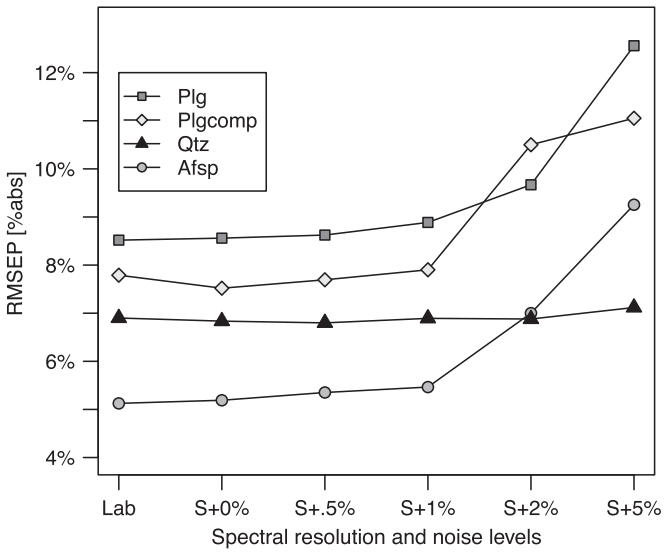

Fig. 5. PLSR model prediction errors (RMSEP (in % absolute), vertical axis) for laboratory spectra (from the work in [17]) and the SEBASS-resampled spectra (this paper) with noise levels from 0% to 5% (from $S + 0\%$ to $S + 5\%$, horizontal axis). The mineral abbreviations used are as follows: Afsp = alkali feldspar; Plg = plagioclase; Qtz = quartz; Plgcomp = plagioclase composition.

[see Fig. 6(a)]. The scatterplot shows values close to the diagonal 1:1 line, indicating good absolute values as compared with predictions based on undistorted spectra.

- In the case of the plagioclase model [see Fig. 6(b)], absolute prediction values have changed, but the relative patterns remained the same (samples plot in the correct relative order). These predictions based on shifted and tilted spectra are still giving useful relative plagioclase patterns, but the values cannot be used in absolute quantitative terms.
- In the plagioclase composition models [see Fig. 6(c)], only the *additive+multiplicative* model shows relative patterns, whereas the two tilting models show prediction patterns with a global maximum (at about 55%), after which very high values are erroneously plotting at lower prediction values again. This effect is caused by the arcsine square-root transformation that was used in [17] to make the data set symmetric. In the last stage of the PLSR modeling, the inverse transformation was used to return the results to the compositional information in the original units of areal percentages (see [17] for details). The squared sine function will cause data folding if the modified spectra cause (nonbacktransformed) prediction values outside the $[0-(\pi/2)]$ range. The data all range $[-(\pi/2)-0]$ for the additive and multiplicative plagioclase composition model [see Fig. 6(c)], and folding causes a total inversion of the data points.
- The alkali feldspar model shows an additional problem with data folding [see Fig. 6(d)]. In this model, a few of the undistorted spectra also cause negative prediction values, causing data folding as these data points cross the y-axis. Furthermore, the alkali feldspar model shows similar inversion of the data values as the plagioclase composition model does, but in this case, the tilted spectra are inverted.

The results aforementioned show that the quartz model behaves quite different to the feldspar models. The added noise up to levels of 5% has virtually no influence on the quartz prediction performance (see Fig. 5), but the tilt and spectral contrast changes are handled by the quartz model much more effectively than the feldspar models (see Fig. 6). We believe that this discrepancy can be attributed to the robustness of the quartz model to distortions; quartz has very distinct, broad, and deep TIR spectral features. The broad features prevail in the noise-added data (see Fig. 2). The distinct and deep features also cause the quartz PLSR model to only use two factors in the PLSR laboratory models (see Table II) (a setting that was also fixed for distorted spectra) that in turn makes the model more robust to spectral distortions and less prone to overfitting. Feldspars, on the other hand, have more and finer features that cannot be captured in simple two-factor models, which makes these models more susceptible to spectral distortions.

The experiment modeling the additive, multiplicative, and tilting distortions revealed that only the *additive+multiplicative* distorted spectrum (same maximum values but lower spectral contrast) provided relative and absolute results with all mineral models (Qtz, Plg, Plgcomp, and Afsp). Spectra with other distortions provided different absolute results, with even relative patterns often being incorrect. This illustrates that vertical shifts and tilts to the spectra were more detrimental for the PLSR results than a simple change in the "absorption depth" of the spectra (dark blue asterisks in Fig. 6).

Airborne SEBASS spectra do show, however, changes in the absorption depth and tilts. These preliminary results therefore

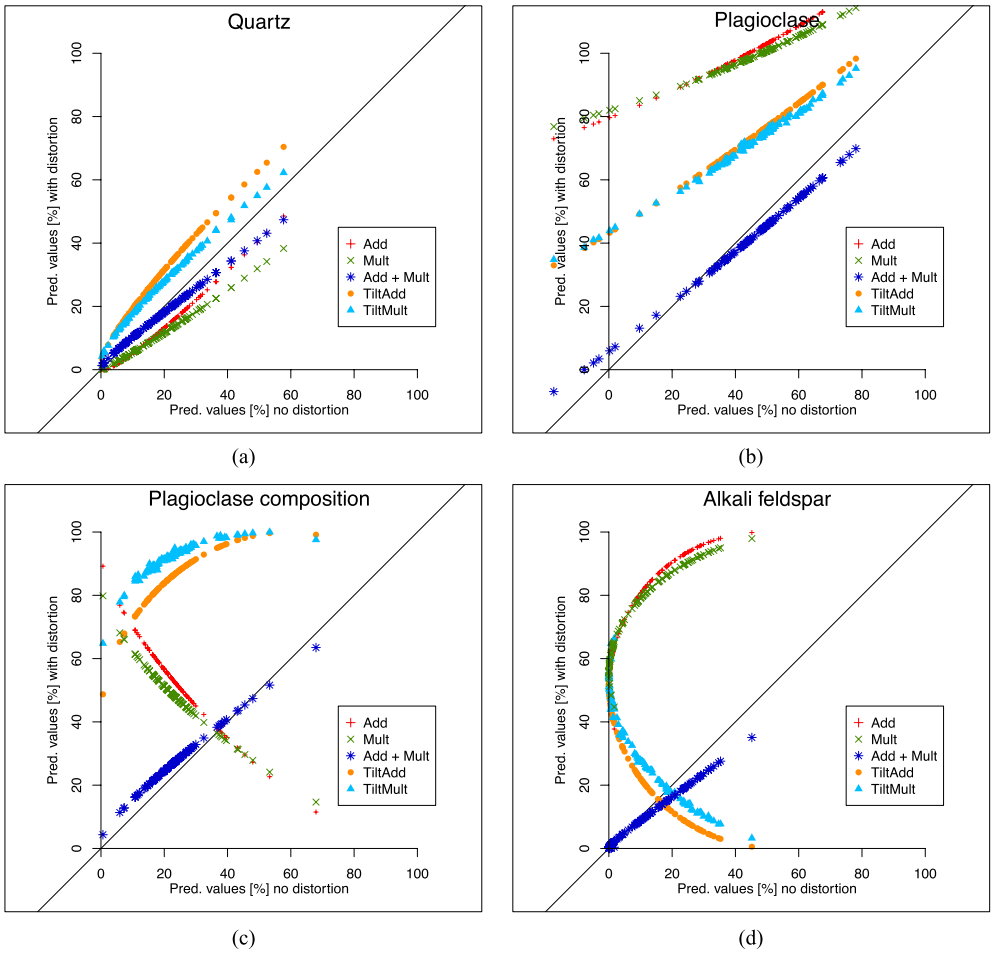

Fig. 6. (a)–(d) PLSR modeling results based on undistorted spectra (horizontal axis) against those from spectra with additive, multiplicative, and tilt distortions (vertical axis). The thin solid line represents the 1:1 line. Values are in areal percentages.

demonstrate the acute need to correct these distortions prior to applying PLSR models to SEBASS data in order to retrieve reliable and quantitative results. We will discuss a possible correction approach that adjusts the mean and standard deviation of airborne spectra to match the laboratory data, and this compensates for the additive and multiplicative distortions, as well as the spectral tilt, in the next section.

C. Compensating Signal Distortions

PLSR models for an area always require a number of samples with known composition and known spectral information. For reliable results, the samples should represent the spectral variation of the area under study. The knowledge from those ground-truth samples can be then used to predict values for the entire study area through a PLSR model.

Additionally, these ground samples could be used to correct for systematic distortions in the airborne hyperspectral data and to normalize the airborne and laboratory data to occupy the same range in values. A mean and standard deviation (MASD) normalization is proposed. For each airborne image band, the mean and standard deviation of the pixels are compared with the mean and standard deviation of the laboratory spectra of the samples from the study area. The airborne pixels at the same wavelength of the laboratory spectra are normalized using the following equation:

$$\epsilon_{\text{air\_new}}(\lambda) = [\epsilon_{\text{air}}(\lambda) - \mu_{\text{air}}(\lambda)] \frac{\sigma_{\text{lab}}(\lambda)}{\sigma_{\text{air}}(\lambda)} + \mu_{\text{lab}}(\lambda) \quad (1)$$

where $\epsilon_{\text{air\_new}}$ are the MASD normalized pixel values, ϵ_{air} are the pixel input emissivity values, μ_{air} and σ_{air} are the mean and standard deviation of the entire airborne image band, respectively, and μ_{lab} and σ_{lab} are the mean and standard deviation of the same band in the laboratory spectra, respectively. (λ) represents an image band at a given wavelength, and the normalization is then executed on a band-by-band basis. The main underlying assumption behind the MASD normalization is that the laboratory samples are a representative sample of the rocks in the airborne scene. The same assumption was already used in the PLSR model predictions, where the error estimates of the prediction models are only valid if the ground samples (for model training) are a representative sample of the rocks in the study area. To increase the validity of this assumption during the MASD normalization, a mask can be applied to the airborne data, where the normalization statistics are only extracted from the general region where the ground

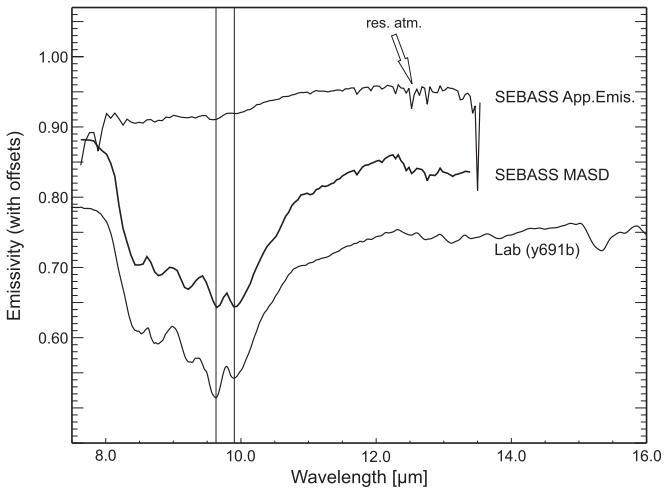

Fig. 7. Comparison between the standard SEBASS apparent emissivity product (Yerington Campaign on June 17, 2008, Flightline 23, with col = 15, row = 2022, no downwelling atmospheric compensation applied, and the temperature and spectral emissivity retrieved through the emissivity normalization method) and the same pixel after the MASD normalization. The arrow indicates residual atmospheric features in the data. *Lab(y691b)* is the laboratory spectrum of the quartz monzodiorite sample taken from the same location imaged by the pixel (see [17]). Vertical lines mark the features that are visible in all three spectra. Spectra are downshifted by 0.10 emissivity units each for clarity, except for the apparent emissivity spectrum.

samples were collected (e.g., the portion of the study area that was visited during the ground campaign). This normalization can be applied afterward to the entire extent of the airborne data set.

When we created synthetic SEBASS spectra with spectral distortions, each set of synthetic spectra contained the same systematic distortions. Therefore, applying the proposed MASD normalization to the synthetically distorted SEBASS signals always results in a perfect correction at all wavelengths, although the approach to create distortions (in the spectral domain) and that to correct for distortions (using information from other pixels in the spatial domain) are independent of each other. The distortions in the SEBASS data from an acquisition over Yerington, NV, USA, were used to run a preliminary test of the method to better understand how the MASD normalization will affect an actual airborne data set. Ground samples collected from the same area (described in [17]) were used for the MASD normalization. Fig. 7 shows an example of a single airborne SEBASS pixel, which shows a spectral tilt toward shorter wavelengths, reduced spectral contrast, and some artifacts between 7.5 and 8.5 μm. Although the spectral contrast is very low (mainly due to a lack of downwelling atmospheric radiance correction), some of the main Reststrahlen features (marked by vertical lines in Fig. 7) are already identifiable in this uncorrected spectrum. After the MASD normalization, the spectral tilt and artifacts near 8.0 μm have been corrected for, and the spectral contrast has been restored (see the bottom two spectra in Fig. 7). The general spectral shape of the Reststrahlen band and the relative depths of individual features match well with the laboratory spectrum.

This proposed MASD normalization is applicable to methods beyond PLSR, as any classification or unmixing algorithm that requires comparison with ground or laboratory spectra (rather than extracting the "endmembers" from the airborne data itself [31]) should benefit from accurate spectral shape and spectral contrast in the airborne data. The limiting factor of this method is the need for representative ground or laboratory spectra, which may not be available in all circumstances (e.g., planetary applications and early-stage exploration in terrestrial areas with poor accessibility). Spectral information for a number of ground sampling locations will be available for most terrestrial hyperspectral mineral mapping examples at the stage where the quantification of mineral composition is attempted, and results need to be validated.

V. Conclusion

Airborne TIR imaging spectroscopy can be used to quantitatively predict felsic rock compositions. In this paper, results from an existing PLSR laboratory study on feldspar and quartz quantification have been used with synthetic SEBASS spectra to test whether SEBASS could also theoretically map and quantify these minerals in airborne data. The results show that the spectral resolution and coverage of SEBASS are suitable for capturing spectral features in feldspars and quartz that fall within the LWIR atmospheric absorption window. Furthermore, the PLSR model results based on the SEBASS band positions were nearly identical (Spearman's correlation $\rho = 0.995$) to those based on laboratory resolution. Predictions became gradually less reliable when successively increasing the noise component in modeled SEBASS data. However, PLSR modeling results still show prediction error (RMSEP) values within 0.4% (absolute) to the original laboratory PLSR predictions, with up to 1% noise added. Noise levels in the standard SEBASS emissivity product were found to be around 0.2%, even without having compensated for downwelling atmospheric radiance. Therefore, the signal-to-noise levels in the standard SEBASS emissivity product are sufficiently high to model quartz and feldspar amounts, and feldspar compositions with PLSR. In order to achieve reliable mineral estimates prior to applying PLSR models, spectral contrast, tilts, and vertical shifts must be compensated for in the spectral data. For that purpose, a MASD normalization is proposed using the same PLSR ground data to compensate for systematic errors that are common to all image pixels. This MASD normalization successfully restored the spectral contrast of the airborne SEBASS data, after which laboratory-developed PLSR models could be applied, yielding quantitative results from airborne TIR data.

Acknowledgment

The authors would like to thank The Aerospace Corporation and SpecTIR LLC for providing the airborne data acquired during the Joint Airborne Collection using Hyperspectral Systems (JACHS) over Yerington, NV, USA. The authors would also like to thank the four anonymous reviewers for their comments that helped improve the clarity of the document.

References

[1] D. Roberts *et al.*, "Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267–279, 1998.

[2] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.
[3] J. Zhang, B. Rivard, and A. Sanchez-Azofeifa, "Derivative spectral unmixing of hyperspectral data applied to mixtures of lichen and rock," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1934–1940, Sep. 2004.
[4] F. J. A. van Ruitenbeek *et al.*, "Mapping white micas and their absorption wavelengths using hyperspectral band ratios," *Remote Sens. Environ.*, vol. 102, no. 3/4, pp. 211–222, Jun. 2006.
[5] N. M. Knox *et al.*, "Dry season mapping of savanna forage quality, using the hyperspectral Carnegie airborne observatory sensor," *Remote Sens. Environ.*, vol. 115, no. 6, pp. 1478–1488, Jun. 2011.
[6] F. D. van der Meer *et al.*, "Multi- and hyperspectral geologic remote sensing: A review," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 14, no. 1, pp. 112–128, Feb. 2012.
[7] M. S. Ramsey, P. R. Christensen, N. Lancaster, and D. A. Howard, "Identification of sand sources and transport pathways at the Kelso Dunes, California, using thermal infrared remote sensing," *Geol. Soc. Amer. Bull.*, vol. 111, no. 5, pp. 646–662, 1999.
[8] T. J. Cudahy, K. Okada, Y. Yamato, A. Cornelius, and R. Hewson, "Mapping the skarn–porphyry–epithermal alteration system at Yerington, Nevada, using VNIR–SWIR–TIR remotely sensed data," CSIRO Exploration and Mining, Dickson ACT, Australia, Tech. Rep. 1121R, 2001.
[9] R. G. Vaughan, W. M. Calvin, and J. V. Taranik, "SEBASS hyperspectral thermal infrared data: surface emissivity measurement and mineral mapping," *Remote Sens. Environ.*, vol. 85, no. 1, pp. 48–63, Apr. 2003.
[10] R. G. Vaughan, S. J. Hook, W. M. Calvin, and J. V. Taranik, "Surface mineral mapping at Steamboat Springs, Nevada, USA, with multi-wavelength thermal infrared images," *Remote Sens. Environ.*, vol. 99, no. 1/2, pp. 140–158, Nov. 2005.
[11] K. A. Reath and M. S. Ramsey, "Exploration of geothermal systems using hyperspectral thermal infrared remote sensing," *J. Volcanology Geothermal Res.*, vol. 265, pp. 27–38, Sep. 2013.
[12] J. W. Salisbury and D. M. D'Aria, "Emissivity of terrestrial materials in the 8–14 μm atmospheric window," *Remote Sensing of Environment*, vol. 42, no. 2, pp. 83–106, Nov. 1992.
[13] K. C. Feely and P. R. Christensen, "Quantitative compositional analysis using thermal emission spectroscopy: Application to igneous and metamorphic rocks," *J. Geophys. Res.-Planets*, vol. 104, no. E10, pp. 24 195–24 210, Oct. 1999.
[14] K. A. Milam, H. Y. McSween, and P. R. Christensen, "Plagioclase compositions derived from thermal emission spectra of compositionally complex mixtures: Implications for Martian feldspar mineralogy," *J. Geophys. Res.-Planets*, vol. 112, no. E10, Oct. 2007, Art. ID E10007.
[15] M. S. Ramsey and P. R. Christensen, "Mineral abundance determination: Quantitative deconvolution of thermal emission spectra," *J. Geophys. Res.-Solid Earth*, vol. 103, no. B1, pp. 577–596, 1998.
[16] S. W. Ruff, "Quantitative thermal infrared emission spectroscopy applied to granitoid petrology," Ph.D. dissertation, Arizona State University, Tempe, AZ, USA, 1998.
[17] C. Hecker, J. H. Dilles, M. van der Meijde, and F. D. van der Meer, "Thermal infrared spectroscopy and partial least squares regression to determine mineral modes of granitoid rocks," *Geochem. Geophys. Geosyst.*, vol. 13, no. 3, pp. 1–15, 2012.
[18] C. Hecker *et al.*, "Thermal infrared spectrometer for earth science remote sensing applications — instrument modifications and measurement procedures," *Sensors*, vol. 11, no. 11, pp. 10 981–10 999, 2011.
[19] J. A. Hackwell *et al.*, "LWIR / MWIR imaging hyperspectral sensor for airborne and ground-based remote sensing," in *Proc. SPIE*, 1996, vol. 2819, pp. 102–107.
[20] L. Kirkland *et al.*, "First use of an airborne thermal infrared hyperspectral scanner for compositional mapping," *Remote Sens. Environ.*, vol. 80, no. 3, pp. 447–459, Jun. 2002.
[21] D. N. Riley and C. Hecker, "Mineral mapping with airborne hyperspectral thermal infrared remote sensing at Cuprite, Nevada, USA," in *Thermal Infrared Remote Sensing: Sensors, Methods, Applications*, C. Kuenzer and S. Dech, Eds. Dordrecht, The Netherlands: Springer Science+Business Media, 2013, pp. 495–514.
[22] F. E. Nicodemus, "Directional reflectance and emissivity of an opaque surface," *Appl. Opt.*, vol. 4, no. 7, pp. 767–773, 1965.
[23] C. C. Borel and R. F. Tuttle, "Recent advances in temperature-emissivity separation algorithms," in *Proc. IEEE Aerosp. Conf.*, 2011, pp. 1–14.
[24] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemometrics Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, Oct. 2001.
[25] G. P. Asner and R. E. Martin, "Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3958–3970, Oct. 2008.
[26] A. Eisele *et al.*, "Advantages using the thermal infrared (TIR) to detect and quantify semi-arid soil properties," *Remote Sens. Environ.*, vol. 163, pp. 296–311, Jun. 2015.
[27] F. A. Yitagesu, F. van der Meer, H. van der Werff, and C. Hecker, "Spectral characteristics of clay minerals in the 2.5–14 μm wavelength region," *Appl. Clay Sci.*, vol. 53, no. 4, pp. 581–591, Oct. 2011.
[28] J. Farifteh, F. van der Meer, C. Atzberger, and E. J. M. Carranza, "Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN)," *Remote Sensing of Environment*, vol. 110, no. 1, pp. 59–78, 2007.
[29] K. H. Esbensen, *Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design*, 5th ed. Oslo, Norway: Camo, 2006.
[30] P. Geladi and B. R. Kowalski, "Partial least-squares regression: a tutorial," *Analytica Chimica Acta*, vol. 185, pp. 1–17, 1986.
[31] C. Hecker, M. van der Meijde, H. van der Werff, and F. van der Meer, "Assessing the influence of reference spectra on synthetic SAM classification results," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4162–4172, Dec. 2008.

Christoph Hecker was born in Basel, Switzerland, in 1974. He received the M.Sc. degree in earth sciences from the University of Basel, Basel, Switzerland, in 1999 and the Ph.D. degree in remote sensing of earth resources from the University of Twente, Enschede, Netherlands, in 2012.

From 2000 to 2001, he was with the Belize Audubon Society, Belize City, Belize. Since 2001, he has been with the Department of Earth Systems Analysis, Faculty of Geo-Information Science and Earth Observation, University of Twente. His research focuses on the mineralogy of hydrothermal and geothermal alteration systems to facilitate the exploration of mineral deposits and renewable energy sources. To study these systems, he applies laboratory, field, and imaging spectroscopy in the thermal-infrared wavelength band.

Dean Riley was born in Petersburg, VA, USA, in 1968. He received the B.Sc. degree in geology from Oregon State University, Corvallis, OR, USA, in 1991 and the M.Sc. and Ph.D. degrees in geological sciences from The Ohio State University, Columbus, OH, USA, in 1994 and 2004, respectively.

From 2000 to 2004, he worked with Ball Aerospace and Technology Corporation, Fairborn, OH, USA, as a Spectral Research Analyst, and from 2004 to 2006, he was with Raytheon UTD, Springfield, VA, USA, as a Senior Geologist. From 2006 to 2012, he was with The Aerospace Corporation, Chantilly, VA, USA, as a Senior Project Engineer, and from 2012 to 2013, he was with SpecTIR, LLC, Fairfax, VA, USA, as a Chief Geologist. Since 2013, he has been a Lead Scientist with Booz Allen Hamilton Inc., Herndon, VA, USA, and a Lieutenant Colonel in the Army Reserves. His research interests include geostatistics, image processing, and algorithm development in the geospatial sciences for geological remote sensing; the correlation of physical, mechanical, and chemical properties of rocks and materials with visible-to-long-wave-infrared spectroscopy using core-logging, airborne, and satellite hyperspectral imagers; target detection and identification; material quantification; multipolarization and multiwavelength synthetic aperture radar for geological mapping; bistatic radar image processing; and temporal processing of electrooptical, multispectral, and thermal-infrared imagery.

Dr. Riley was the recipient of the Balls Northern Operations Technical Excellence Award in 2002 and The Aerospace Corporations Technical Excellence Award in 2012. He was also the recipient of the Commission as Second Lieutenant in the United States Army Reserve, Fort Bragg, NC, USA, in 1990.

Mark van der Meijde received the M.Sc. degree in earth sciences from Utrecht University, Utrecht, The Netherlands, in 1998 and the Ph.D. degree in geophysics from the Swiss Federal Institute of Technology Zurich (ETH Zurich), Zurich, Switzerland, in 2003.

In the past, he worked with the Dutch Seismological Institute (KNMI) and the Dutch Geological Survey (NITG-TNO). In 2003, he started working with the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands, where he has been an Associate Professor in geophysics with the Department of Earth Systems Analysis since 2009. He is the author or coauthor of several publications in IEEE letters and transactions. His main research interests include making the link between remotely sensed surface and subsurface information. His fields of application focus on geophysics, 3-D geological modeling, integrating space and airborne (hyperspectral) remote sensing with seismology for geohazards and tectonics, and 3-D environmental modeling.

Dr. van der Meijde is a member of the American Geophysical Union, the European Geosciences Union, and the European Association of Geoscientists and Engineers.

Freek D. van der Meer was born in 1966. He received the M.Sc. degree in structural geology and tectonics from the Free University of Amsterdam, Amsterdam, The Netherlands, in 1989 and the Ph.D. degree in remote sensing from Wageningen Agricultural University, Wageningen, Netherlands, in 1995.

He started his career with Delft Geotechnics (now Geodelft), where he worked on geophysical processing of ground penetrating radar data. In 1989, he was appointed as a Lecturer in geology with the International Institute for Geo-Information Science and Earth Observation (currently the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, Netherlands), where he has been working in various positions and is currently the Vice Dean of the faculty and the Vice Chairman of the Department of Earth System Analysis. In 1999, he was appointed as a Professor with the Delft University of Technology, Delft, The Netherlands. In 2005, he was appointed as a Professor of geological remote sensing with the Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands. He is the author or coauthor of over 150 papers in international journals and more than 150 conference papers and reports. He is the Editor of the books *Spatial Statistics for Remote Sensing* (Kluwer, 2000), *Imaging Spectroscopy: Basic Principles and Prospective Applications* (Kluwer, 2001), and *Remote Sensing Image Analysis: Including the Spatial Domain* (Springer, 2004). His research is directed toward the use of hyperspectral remote sensing for geological applications with the specific aim of using geostatistical approaches to integrate airborne and field data into geologic process models. This has been directed toward exploration for oil and gas, mineral potential in hydrothermal systems, and the exploration of geothermal systems.

Dr. van der Meer was the Chairman of the Netherlands Society for Earth Observation and Geoinformatics and of the Geological Remote Sensing Special Interest Group of the European Association of Remote Sensing laboratories. He is a member of the Board of Earth and Life Sciences of the Royal Netherlands Academy of Sciences. He is the Editor-in-Chief of the *International Journal of Applied Earth Observation and Geoinformation*, the Associate Editor of *Terra Nova*, the Associate Editor of the *Brazilian Journal of Geology*, the Associate Editor of *Computers and Geosciences*, and the Series Editor of the Remote Sensing and Digital Image Processing Book Series of Springer.