Spectral data set factor analysis and end-member recovery: Application to analysis of Martian atmospheric particulates

Joshua L. Bandfield and Philip R. Christensen
Department of Geology, Arizona State University, Tempe

Michael D. Smith
NASA Goddard Space Flight Center, Greenbelt, Maryland

Abstract. A method is described that uses target transformation factor analysis techniques to determine the number of independently variable components and recover the spectral end-members present in a set of mixed spectra. These techniques were tested on two sets of synthetic spectral mixtures and several subsets of Thermal Emission Spectrometer (TES) data. In both synthetic mixture sets, the correct numbers of components were determined, and the original spectral end-members were accurately recovered. An initial application of this method was used with several subsets of TES apparent emissivity spectra that contain only minor surface spectral components. This method has demonstrated that the spectra can be modeled using linear combinations of three spectral end-members: atmospheric dust (with atmospheric gas absorptions that vary in unison with the dust shape), water ice cloud, and blackbody. The atmospheric dust and water ice cloud spectral shapes were recovered from several orbits with a wide variety of atmospheric dust opacities during the southern spring and summer seasons. The atmospheric dust spectral shape is nearly constant except for the relative contribution of atmospheric gasses that vary in unison with the dust shape. The water ice spectral shape is also constant with the exception of a small shift in the position of the \( \sim 800 \text{ cm}^{-1} \) absorption that was observed during a period of high opacity.

1. Introduction
1.1. Technique

Determining the composition of a substance through its spectral characteristics is a common goal in a variety of laboratory and remote sensing applications. It is often the case that a single measured spectrum contains a mixture of several substances or compositions, which can greatly complicate the interpretation and analysis of the spectrum and lead to false or ambiguous results. However, accurate interpretation of a spectrum that contains several components allows for the detection of substances that could not otherwise be distinguished at the spatial resolution of the measurement. With multiple measurements of different portions of a mixed surface or of mixtures with different concentrations, the number and spectral characteristics of the individual components can be determined from the variations in the spectra of the mixtures.

These techniques are applied in the following method.
1. A set of measured spectra are selected. Any variability within this set is assumed to be described by varying linear combinations of spectral end-members.
2. The eigenvalues and eigenvectors of the covariance matrix of the measured spectra are calculated. The number of eigenvectors required to reconstruct the original measured spectra to within the noise level determines the number of linearly independent spectral end-members in the set of measured spectra.
3. The spectral end-members are identified by least squares fitting trial spectra with the first \( n \) eigenvectors, where \( n \) is the number of independent components in the set of measured spectra determined from step 2. These, in general, may not correctly identify the spectral end-members present. However, if the trial spectrum may be closely matched by a linear combination of the significant eigenvectors (the best fit spectrum), then the trial spectrum is a plausible spectral end-member. Where the trial spectrum and the best fit spectrum differ, the least squares fit more closely describes the actual end-member present in the system. A number of trials are attempted until \( n \) linearly independent end-members are identified.
4. The selected set of end-members is then fit in a linear least squares fashion to the measured spectra. Physically meaningful concentrations of the end-members as well as RMS errors between measured and modeled spectra provide a check on the set of end-members, as they are not automatically identified and may be subject to bias in their selection.

1.2. Application

We demonstrate here that this method may be applied to data returned from the Thermal Emission Spectrometer (TES) [Christensen et al., 1992] aboard the Mars Global Surveyor Spacecraft for identification and characterization of Martian atmospheric particulate dust and water ice. As Martian atmospheric gases, atmospheric particulates, and surfaces all display absorption features in the thermal infrared portion of the spectrum (5–50 \( \mu m \)) [Hanel et al., 1972; Conrath et al., 1973; Martin and Richardson, 1993; Christensen, 1998; Christensen et al., 1998; Smith et al., this issue (a), (b)], there is a great deal of overlapping compositional information that can be difficult to interpret in the individual spectra. Correct isolation of the precise spectral shape of atmospheric particulate end-
2. Method

2.1. Introduction

In the present application, R-mode factor analysis derives the eigenvectors and eigenvalues from a set of mixed spectra to determine the number of components that influence the system. The value of this analysis is the reduction of the dimensionality and the number of variables of the data (Figure 1). For example, if there are only three components varying within a system, the system can be reconstructed with linear combinations of the first three eigenvectors. As a result, the dimensionality of the data set is reduced from the number of spectral bands to three. However, linear combinations of these orthogonal eigenvectors can reproduce all of the spectra within the original data set. The eigenvectors may have no physical meaning and do not represent the spectra of the true components that are present in the mixture.

The target transformation technique attempts to find linear combinations (transformations or oblique rotations) of the set of the R-mode eigenvectors that produce a set of real spectral components that can closely reproduce the original spectra. There are several methods of oblique rotation of the eigenvectors such as quartimin and oblimax [Rummel, 1970] in addition to target transformation. Target transformation is the only one of these techniques that is used to identify the original spectral end-members, whereas the other oblique rotation methods are used to optimize various parameters such as the loadings along factor axes. Using a trial spectrum (also known as the “test” or “target” spectrum), the transformation rotates the matrix composed of the significant eigenvectors to produce a spectrum that best fits the trial spectrum [Malinowski, 1991] (Figure 2). The trial spectrum may be taken from a spectral library or be simply a guess. In other words, a linear combination of the eigenvectors is produced which attempts to match, or fit, the trial spectrum. If the trial spectrum fits within the dimensionality of the data, the trial spectrum is a component of the system and a possible spectral end-member and will be closely reconstructed using a combination of the derived eigenvectors. If the trial spectrum is not a possible component, the eigenvectors will not be able to match the trial spectrum. Thus the target transformation provides a tool for independently identifying spectral end-member components within a system.

A new end-member recovery technique described here takes advantage of the limited spectral shapes that can be constructed from the eigenvectors to isolate the spectral end-members present within a system, even if the data set contains only spectral mixtures. This technique uses the rotation defined by the target transformation to predict spectral end-members. In addition, end-member recovery does not require that the spectral end-member be present in the set of trial
spectral bands) of the measured values, $R$

In this equation, $D$ is the column matrix of end-member spectra, and $C$ is the column matrix of end-member concentration values. The goal of the analysis is to find $R$ and $C$ using the measured values $D$ and a set of end-member reference spectra. There is no unique $R$ and $C$, however, because even though it is possible to have more equations than unknowns in the system, the system will remain underdetermined [Liu and Koenig, 1987] as shown by (2):

$$\text{rank } [D] = \min \{\text{rank } [R], \text{rank } [C]\}.$$  

In other words, the dimensionality of $D$ is less than that of $R$ and $C$. Sasaki et al. [1984], Liu and Koenig [1987], and Tompkins et al. [1997] have successfully solved for $R$ and $C$ directly. These workers used methods for solving large systems of nonlinear equations and have restrictions on component sums equal to unity and nonnegative $R$ values. With correct estimation of the number of end-members and good initial guesses, these methods have converged on accurate solutions.

2.2. Factor Analysis

Factor analysis essentially decomposes the $D$ matrix into abstract $R$ and $C$ matrices (the bar over the matrix symbol indicates an abstract matrix). These matrices are composed of the orthogonal eigenvectors derived from the $D$ matrix. The eigenvectors composing the rows of $R$ are used for determining the number of components using the factor analysis, target transformation, and end-member recovery methods described below. A detailed discussion of the derivation of eigenvalues and eigenvectors as it is applied here is provided by Malinowski [1991].

Factor analysis reduces the dimensionality of the data set to the number of significant components and describes this space within the full dimensionality of the data through the derived eigenvectors. The eigenvectors describe a system as efficiently as possible by deriving orthogonal axes that maximize the variability in measured values along the derived axes [Lay, 1994]. Each successive eigenvector is orthogonal to the previous eigenvectors and accounts for the maximum variation in the data set not accounted for by the previous eigenvectors. The variance along each eigenvector within the system is described by the corresponding eigenvalues. The information contained within the eigenvalues and eigenvectors can be used to estimate the number of end-member components that influence the data.

If there are independent components varying within a system, the set of spectra composed of mixtures will be linear combinations of the spectral end-members if spectral mixing is linear. The maximum dimensionality that can be spanned by the spectra is the number of spectral end-members. If the number of spectral bands exceeds the number of spectral end-members, the measured spectra will lie in a space contained within the higher dimensionality of the data set with the number of dimensions equal to the number of spectral end-members. The derived eigenvalues and eigenvectors account for any systematic variation as efficiently as possible. As a result, the first $n$ eigenvectors, where $n$ is the number of end-member spectra, will span the space that is also spanned by the data to within the noise. All the data and the spectral end-members may be reconstructed by linear combinations of these first $n$ eigenvectors. Any deviation from this space is due to noise present in the data, and the subsequently derived eigenvalues will be smaller than the significant eigenvalues that are associated with detectable spectral variation [Malinowski, 1991].

There are numerous methods that use eigenvalues and eigenvectors to determine the number of components ($n$) that influence a system. Unfortunately, these methods require a good understanding of the application and error. In a remote sensing situation, where noise is highly variable and dependent on instrument conditions, wavelength, and temperature or lighting conditions of the surface, numerically characterizing noise using a single number is difficult. This makes an automated method of determining the number of end-members based on noise unreliable. Analysis of eigenvalues and eigenvectors can determine $n$ reliably from two methods that were used in this study. (1) the eigenvalues become small and relatively constant with increasing $n$, as they only account for random error [Malinowski, 1991]. (2) The eigenvectors may be plotted as spectra. Eigenvectors composing $R$ columns 1

![Figure 2](image.png)

**Figure 2.** The target transformation finds the orthogonal projection of a trial spectrum onto the plane defined by the eigenvectors in Figure 1 (ev1 and ev2). This orthogonal projection is the least squares fit of the best fit spectrum to the trial spectrum and is a possible end-member.
through $n$ contain spectral information. Columns $n+1$ and higher contain only random noise that is present in the system [Smith et al., 1985]. Systematic errors and other features, such as sloping in the spectra due to miscalibration may also be represented in the eigenvectors and eigenvalues.

Several methods exist with differences in accuracy and speed for deriving eigenvectors and eigenvalues from large matrices, including the iterative power method, the Jacobi method, singular value decomposition (SVD), and nonlinear iterative partial least squares (NIPALS) [Malinowski, 1991]. The power method as described by Press [1992] was used for this study. This method is relatively simple and has good accuracy and speed for deriving the first few eigenvalues and eigenvectors of a matrix [Lay, 1994].

In the analysis presented here, the mean is removed before the analysis is performed. This essentially reduces the rank of the data set by one. As a result, the true number of components in the system is the number of components determined by factor analysis plus one. Rozett and Peterson [1975] have evaluated the difference between performing factor analysis about the mean or origin on chemical data sets and found no significant difference in the interpretation based on the two methods. For this analysis, the mean is treated as a significant eigenvector when performing the target transformation described below. There was no significant difference in interpretation or results of this method performed about the mean or origin.

### 2.3. Target Transformation

To obtain physically significant columns of $R$ (i.e., end-member spectra), it is necessary to find a transformation vector $\mathbf{t}_n$ such that (summarized from Malinowski [1991])

$$\mathbf{x}_n = \mathbf{R}\mathbf{t}_n,$$

where $\mathbf{x}_n$ is the best fit spectrum and $\mathbf{R}$ is the abstract eigenvector matrix composed of the significant eigenvectors. In practice, the best fit spectrum $\mathbf{x}_n$ is constructed from each end-member trial spectrum by performing a least squares fit of the eigenvectors to the trial spectrum $\mathbf{x}_n$. If the trial and best fit spectra are sufficiently similar, then $\mathbf{x}_n$ is a valid column of $R$. This comparison may be quantified by finding the RMS error between the real and best fit spectra. However, RMS error in spectral curve fitting is affected by not only the overall fit of the two spectra being compared, but also the random noise level present in either spectrum. Minor components reconstructed from the eigenvectors are generally noisier than the major ones and will have a higher RMS error between the trial and best fit spectra. As a result, a poorly fitting, but low noise target transformation result may have a lower RMS error than a matching target transformation that is a minor component with a higher noise level. In practice, valid trials are best determined by individual inspection.

### 2.4. End-Member Recovery

Inspection of the best fit spectra from the target transformation shows that even for many failed transformations, in which the best fit and trial spectra are not similar, the best fit spectrum is an accurate spectral end-member. For successful target transformation trials, the best fit spectrum is a refined and more accurate end-member than the trial spectrum. It is possible to take advantage of this property in the analysis of a data set by selecting the set of true spectral end-members for the system from a group of best fit spectra that result from the target transformation. These derived end-members may be analyzed further for their properties and used to deconvolve the mixed spectra in order to obtain end-member concentrations.

The ability to obtain a spectral end-member using a trial spectrum that is not necessarily similar to the best fit spectrum is a new and somewhat surprising result. Because the target transformation is simply fitting the trial spectrum with a linear combination of eigenvectors in a least squares manner, there is no uniqueness of one combination over another. Spectral shapes of substances are broadly similar to each other, however, and the trial spectra produce realistic spectral shapes from the eigenvectors even if the specific spectral end-member is not present in the set of trial spectra. Because end-members show a greater spectral contrast than mixtures (i.e., features are more pronounced), the target transformation method is practical for obtaining the end-members present in a system. The target transformation matches as well as possible the relatively high spectral contrast present in the trial spectrum by fitting a combination of the eigenvectors to the trial. As a result, best fit spectra are often spectral end-members. Spectral end-members are generally recovered with trial spectra that contain absorptions near the same wavelengths as the spectral end-members, but these absorptions are not necessarily the same shape.

Many minerals demonstrate shifts in shape and position of spectral features based on compositional differences due to solid solution or ionic substitution [Farmer, 1974; Salisbury et al., 1991]. In addition, varying grain sizes may change spectral features even if composition remains constant [Hapke, 1981; Salisbury et al., 1987; Moersch and Christensen, 1995]. Spectral libraries will always contain a finite number of reference samples, and the spectral end-members within a set of spectra may be of a different particle size or composition than what is contained in the spectral library. Consequently, there will always be at least minor differences between a trial spectrum and the true spectral end-member within a system. Constructing and analyzing the end-members derived from the eigenvectors recovers minor deviations present between trial and end-member spectra that can aid in analysis of the individual end-members and the system.

The ability to obtain end-members through the best fit spectra does necessitate the use of trial spectra, but the end-members present in this suite need only be broadly similar in shape to those in the data. Small variations in features are recovered in the best fit spectra. If the spectral end-member is not present in the trial set, it may still be recovered in the best fit of a trial spectrum that is only grossly similar in shape to the end-member. Selection of end-member spectra from the set of best fit spectra, when the spectral end-member is not present in the reference set, is not automated. Selection is done by plotting best fit spectra, identifying major end-members, and eliminating spectra that are either mixtures of end-members or physically unrealistic. This selection obviously introduces both user bias and experience into any result.

The selected set of end-members may be checked to ensure their accuracy by deconvolution of the original mixed spectra with the set of best fit end-members. The deconvolution provides a least squares fit of the original mixed spectra using a linear combination of the best fit end-members. The weighting of each end-member represents the component fraction of the measurement. Deviation of component sums from 100%, in-
individual end-member component percentages <0 and >100%, and high RMS errors indicate that the set of selected spectral end-members is not correct. The most common indication that an incorrect spectral end-member has been selected is the presence of the incorrect end-member in significantly negative percentages as derived from the deconvolution. In addition, all spectral end-members should be linearly independent. To check for linear independence, it is established that no end-member can be modeled as a linear combination of the other end-members selected.

3. Results 1: Synthetic Mixture Sets
3.1. Introduction
The combined factor analysis, target transformation, and end-member recovery technique was applied to two synthetic spectral mixture sets to demonstrate the ability to both identify and predict the number of components and spectral end-members using this technique. The first set is an idealized two-end-member case that consists of noiseless mixtures of two simple synthetic spectral shapes. The second mixture set contains linear combinations of three mineral spectra: quartz, microcline, and albite with random noise added (Figure 3).

Mineral spectra used for the synthetic mineral mixture set and for the set of trial spectra were selected from the Arizona State University spectral mineral library [Christensen et al., this issue (a)]. All mineral spectra used were obtained in the laboratory using the thermal emission method described by Ruff et al. [1997] in the 2000–400 cm⁻¹ wavelength range at 2 cm⁻¹ spectral sampling (4 cm⁻¹ spectral resolution). All samples were reduced to 400–1600 cm⁻¹ wavenumber range and resampled to 10 cm⁻¹ sampling.

The synthetic spectral shape mixture set consists of 10 noiseless spectral mixtures of two components weighted from 10 to 90%. The mineral mixture data set consists of 200 linear mathematical mixtures of the three spectral end-members (Figure
4) weighted in varying proportions. The microcline is never included in a proportion above 20%, and the quartz and albite concentrations are never above 66%. Spectral end-members are not represented within the mixed data sets in order to demonstrate that the end-members can be recovered even if the data only contains mixtures and some components are only represented in relatively minor amounts. Random noise with an average deviation of $\sim 0.005$ emissivity (SNR = $\sim 200$) was added to the mixed mineral spectra (Figure 4).

3.2. Simple Mixture Set

Eigenvectors were derived from the first synthetic mixture set and are shown in Figure 5. The first eigenvector demonstrates spectral-like features and has a significant value relative to the succeeding eigenvectors which do not deviate from zero. The first eigenvalue is also much larger than succeeding eigenvalues due to the high variance associated with the significant eigenvector. Subsequent eigenvalues are zero, indicating that no information is contained beyond the first two dimensions in the original data set. In this idealized, noise-free case, any magnitude associated with the insignificant eigenvalues is due to numerical roundoff error. The results of the eigenvector and eigenvalue analysis methods agree with each other and correctly indicate that there are two end-member components. As the analysis was performed on data with the mean removed, the number of components indicated by the factor analysis is 2, the number of components in the mean-removed data set plus one.

The set of trial spectra (in which the spectral end-members are present) and the first eigenvector and the mean (which is used as a significant eigenvector) are used for the target transformation. Several of these results are shown in Figure 6, including a successful transformation, an unsuccessful transformation that successfully isolates the correct end-member, and an unsuccessful transformation that does not predict the correct end-member. The successful target transformation (Figure 6a) confirmed the presence of the trial spectra in the factor space as indicated by the similarity between the trial and best fit spectra. The unsuccessful target transformations (Figure 6b) that predict a correct end-member indicate how the end-member recovery technique uses an unsuccessful target transformation to return an accurate end-member from a trial spectrum that is somewhat similar in band shape and position. This is not always the case, as shown by the unsuccessful target transformation (Figure 6c) that returns a best fit spectrum that is a mixture of the two spectral end-members.

Distinguishing between the two unsuccessful transformations is important. In this case, it is easy to recognize that the best fit spectrum shown in Figure 6c is a mixture of the two end-members and the best fit spectra shown in Figure 6b are true end-members. The two spectral features in the first unsuccessful target transformation example are clearly varying independently of each other in the mixed spectra. The best fit spectrum shown in Figure 6c can also be rejected as a possible end-member as it can be modeled as a linear combination of the other two end-members. There will also be negative concentration values if the best fit spectrum from Figure 6c is used to deconvolve the original spectra.

Figure 4. Mineral synthetic mixture set and an example of the random noise level (centered about 1) added to each spectrum.

Figure 5. Eigenvalues (listed) and eigenvectors (plotted) for the idealized, noiseless synthetic spectral mixture set. All information is contained in the mean (solid line) and the first eigenvector.
3.3. Mineral Mixture Set

The correct number of end-members in the mineral mixture set was identified by analysis of the magnitude of the derived eigenvalues (Figure 7). Each successive eigenvalue is smaller than the previous and the component information is contained in both the magnitude of the eigenvalues and the rate of decrease between successive eigenvalues. In the mineral mixture set the first two eigenvalues are relatively large when compared to the succeeding eigenvalues. Eigenvalues 3–10 are decreasing at a relatively constant and much lower rate than the first two, which are decreasing by more than several times their magnitude between successive eigenvalues. In addition, inspection of the eigenvectors (Figure 7) shows that only the eigenvectors corresponding to the significant eigenvalues contain spectral information and the remaining eigenvectors only contain noise. Each eigenvector does not correspond to any particular end-member, but the low spectral contrast of the second eigenvector indicates that either two of the components are similar and/or a minor component is present in the system. This is the case with the minor abundance of microcline in the mixtures and its spectral similarity to the albite.

The target transformation was performed on the mineral mixture set using the first two eigenvectors and the mean. The best fit mineral spectra were matched to each spectral end-member present in the mineral mixture data set (Figure 8). The results of the analyses of the mineral mixture sets demonstrate the ability of the target transformation to confirm end-members present in the system. The target transformation closely reproduces the correct trial spectra with the best fit spectra because they are contained within the factor space defined by the significant eigenvectors. When a trial spectrum is not contained within the factor space, it is not reproduced well by the best fit spectrum.

Close inspection shows that not only are the best fit spectra similar to the trial, but where differences occur in band shape, the best fit spectra are recovering minor spectral features that are present in the original spectral end-members used to create the mixtures. Figure 8 shows the trial, best fit, and true quartz that is present in the mixtures. Points A–D in Figure 8 show where the best fit spectrum is accurately recovering minor spectral features not present in the trial spectrum. At all of these points the best fit spectrum more closely matches the true end-member rather than the trial spectrum. If one or more spectral features in the trial spectrum match the spectral end-member present in the system, the best fit spectrum will more closely match the spectral end-member than the trial spectrum.

Recovery of these minor features can provide detailed information about the spectral end-members that might not otherwise be gleaned from the data set. Use and interpretation of the best fit spectra to select and optimize spectral end-members, rather than for a blind numerical comparison to trial spectra, is a new and significant benefit of the recovery method. As with the first mixture set, the actual spectral end-members are not necessary to be able to isolate them from the data set (Figure 9). Trial spectra containing spectral features roughly matching those of the spectral end-members can often be used to define a transformation that can accurately isolate a spectral end-member. The target transformation is obviously unsuccessful in confirming the presence of variscite and serpentine in Figure 9; however, the method is successful in recovering the accurate microcline and albite spectral shapes.

The best fit spectra generally contain some noise because of

Figure 6. Examples of target transformation results for the synthetic mixture set. (a) A successful target transformation which confirms the presence of each spectral end-member. The best fit spectrum is offset from the original spectrum for clarity. (b) An unsuccessful target transformation that recovers accurate end-members. The trial spectra are not well matched; however, the best fit spectra closely match the original spectral end-members. (c) An unsuccessful target transformation that returns a best fit spectrum that is not a spectral end-member.

the random error added to the data sets. This is especially true with components which are only included in minor abundances. The correct numbers of components were selected in both mineral mixture sets and much of the random noise separated from spectral information. However, there is always some remaining imbedded error that can be minimized with a larger data set, reduced random noise, larger end-member
composition percentages, and increased spectral contrast [Malinowski, 1991]. However, despite the noise, the original spectral end-member is still accurately recovered, as has been demonstrated.

4. Results 2: Application to Mars

4.1. TES Data Calibration and Processing

Data were acquired from the TES instrument during the initial aerobraking phase of the Mars Global Surveyor mission. The TES instrument is a Fourier transform Michelson interferometer that covers the wavelength range from 1700 to 200 cm\(^{-1}\) (6 to 50 \(\mu\)m) at 10 or 5 cm\(^{-1}\) sampling [Christensen et al., 1992]. The instrument also contains bore sighted thermal (5–100 \(\mu\)m) and visible/near-infrared (0.3–3.5 \(\mu\)m) bolometers. The focal planes in each wavelength consist of three cross-track and two along-track detectors with an instantaneous field of view of \(\sim 8.5\) mrad. The TES instrument uses a pointing mirror that allows for limited targeting capability, limb observations, image motion compensation (IMC), and periodic calibration by observing space and an internal reference surface. The final 2 hour circular mapping orbit of MGS will provide a continuous strip three pixels wide with a spatial sampling of \(\sim 3\) km from a mean altitude of 379 km.

The data used for this study were obtained during the initial aerobraking and science phasing orbit phases of the mission,
where the geometry of observations, including emission angle and pixel size, is quite variable. The seasons covered were between $L_s$ 225 and 307, corresponding to southern spring and summer. Spectra used for this study were limited to emission angles less than 45° and spatial resolutions of 10–20 km per pixel. These spectra were obtained near the periapsis portion of the orbit where the distance of the spacecraft to the measured surface was less than 1500 km. Because of the high and variable velocity of the spacecraft at the periapsis of the elliptical orbit, IMC could not be implemented, and the field of view is changing during the measurement. The ground temperatures of the observations used for this study ranged from 240 to 300 K as local time of observations varied from 1600 to 1000 with increasing orbit number. Sample to sample random noise of the TES instrument is rated at $1.2 \times 10^{-2}$ W sr $^{-1}$ cm $^{-2}$ [Christensen et al., 1992], and the variety of temperatures observed result in a range of signal to noise values of 200–650 at 1000 cm $^{-1}$.

A linear response function for each of the six detectors is derived periodically from observations of space and an internal blackbody reference surface of known emissivity and temperature and is used to convert the raw spectra to calibrated radiance. Three scans each of space and the reference surface are taken and averaged for each detector. The linear response of the instrument allows for a relatively simple conversion from raw to calibrated radiance. All spectra used for this study were converted to apparent emissivity by dividing out a Planck curve of the highest brightness temperature within a band of 50 cm $^{-1}$ width in each spectrum. In all cases, the highest brightness temperature was near 1300 cm $^{-1}$.

The data in this study were selected to emphasize atmospheric components. This was done by focusing on periods of relatively high dust opacity [Smith et al., this issue (a)] as well as using data collected over surfaces with high emissivity in the wavelength regions studied. These surfaces were selected based on previous infrared thermal mapper (IRTM) results [Christensen, 1998]. While surface components are either not present or subdued in these locations and wavelength regions,

![Figure 9](image_url)
it has been demonstrated that surface components are consistently present at other locations [Christensen et al., this issue (b); (c); Smith et al., this issue (b)].

4.2. Arsia Mons Orbit 55

A subset of 510 spectra from orbit 55 (Table 1) was chosen for the application of the factor analysis and end-member recovery. This subset passes nearly directly over Arsia Mons, and a visual inspection of the data set shows a significant amount of variability in both the atmospheric water ice and dust absorption features (Figure 10). While dust opacities are low (0.1–0.3 at 1050 cm\(^{-1}\)), they comprise the range observed for most Martian conditions [Martin and Richardson, 1993; Christensen, 1998; Smith et al., this issue (a)] with the exception of dust storms, where 1050 cm\(^{-1}\) dust opacities may exceed 1. The water ice abundance is also quite variable and span conditions observed in most of the current daytime TES data set (J. C. Pearl et al., Mars water ice clouds: Observations by the Thermal Emission Spectrometer (TES) during the first Martian year, submitted to Journal of Geophysical Research, 1999) (hereinafter referred to as Pearl et al., submitted manuscript, 1999).

Eigenvectors and eigenvalues were derived over wavelength ranges 234–529 and 793–1301 cm\(^{-1}\). This range includes the fundamental absorptions in both water ice and dust. It also avoids both the 667 cm\(^{-1}\) CO\(_2\) fundamental, where atmospheric opacities are \(\gg 1\), and the high wavenumber range of the spectrum, where the noise is higher and absorptions are dominated by atmospheric water vapor and CO\(_2\).

The eigenvalues and eigenvectors (Figure 11) clearly indicate the presence of four components. The first two eigenvalues account for \(\sim 95\%\) of the variance in the mean removed data set and are \(\sim 2\) orders of magnitude larger than the small, relatively constant eigenvalues 4–10. The third eigenvalue is much smaller than the first two but is still 3–4 times larger than the succeeding eigenvalues, which are decreasing at a much lower and more constant rate. The first two eigenvectors also contain significant spectral information, while eigenvectors 4–10 are noisy and featureless. Again, while the third eigenvector is relatively featureless compared to the first two, it does contain a significant offset from zero in the low-wavenumber region. This offset may account for slight effects in the spectra such as slopes. Slopes may be introduced when the peak emissivity is less than 1 and the derived surface temperature is slightly less than the true kinetic temperature or when a range of surface temperatures are present within an observed footprint. The eigenvector and eigenvalue analyses indicate that all spectral absorption features in the data set can be reconstructed using linear combinations of three major components. A fourth minor component can account for slight sloping effects in the data.

Using the mean, the first three eigenvectors, and a set of trial spectra, the target transformation was used to recover three

**Table 1. Several Parameters for the Data Subsets Used in the Analysis**

<table>
<thead>
<tr>
<th>Orbit</th>
<th>ICK</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Emission Angle, (\deg)</th>
<th>Sun Inclination, (\deg)</th>
<th>Target Distance, km</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>922</td>
<td>1323</td>
<td>341</td>
<td>349</td>
<td>5</td>
<td>0–45</td>
</tr>
<tr>
<td>55</td>
<td>966</td>
<td>1050</td>
<td>122</td>
<td>123</td>
<td>0</td>
<td>0–45</td>
</tr>
<tr>
<td>61</td>
<td>758</td>
<td>802</td>
<td>275</td>
<td>276</td>
<td>13</td>
<td>3–5</td>
</tr>
<tr>
<td>107</td>
<td>900</td>
<td>934</td>
<td>273</td>
<td>274</td>
<td>17</td>
<td>2–4</td>
</tr>
<tr>
<td>225</td>
<td>1220</td>
<td>1450</td>
<td>149</td>
<td>153</td>
<td>15</td>
<td>0–21</td>
</tr>
</tbody>
</table>

ICK, Incremental Counter Keeper.

Figure 10. Orbit 55 data subset used for the analysis presented here. Though not all spectra are shown, the plotted spectra cover the variability of the entire set. Strong absorptions at \(\sim 400\) and \(\sim 1100\) cm\(^{-1}\) are due to atmospheric dust, and the variable absorption at \(\sim 850\) cm\(^{-1}\) is primarily due to ice.
spectral end-members present in the data set: dust, water ice, and blackbody (Figure 12). A feldspar shape was used to recover the dust spectral shape because it contains absorptions in the same general wavelength region as the dust, though the shapes of the absorptions are significantly different. It is obvious that the recovered dust end-member is not an emissivity spectrum of pure feldspar and its composition must be interpreted by other means. However, the feldspar spectrum did provide a close enough approximation to recover the dust end-member shape. This was applied in the same way as with the synthetic mixture sets described in the previous section. While a spectral library is necessary for this method, the actual spectral end-member does not need to be present in the library in order to be recovered and identified.

The recovered dust compares favorably to spectra that are free of the 850 cm\(^{-1}\) water ice cloud absorption in this and other data subsets [Smith et al., this issue (a)]. The dust spectral shape also contains minor CO\(_2\) and water vapor absorptions. This indicates that the CO\(_2\) and water vapor are varying in unison with the dust and are not separate spectral end-members.

Figure 11. Eigenvalues (listed) and eigenvectors (plotted) for the orbit 55 Arsia Mons subset. Eigenvectors 1 and 2 contain obvious spectral information. Eigenvector 3 also contains a slight slope that deviates significantly from zero. There is no significant spectral information in the succeeding eigenvectors. The mean is shown in solid.

Figure 12. Target transformation and end-member recovery results for the orbit 55 Arsia Mons subset. Trial (dashed) spectra were used to recover the best fit (solid) spectra which were used as the spectral end-members. Blackbody (top) and water ice (bottom) are well matched but also contain small deviations from the idealized blackbody and water ice trial spectra. The dust shape (which includes minor water vapor and CO\(_2\) bands) was not matched well by any single mineral; however, a trial plagioclase spectrum was used to recover the dust end-member because they both contain absorptions in the same wavelength regions. Each trial/best fit pair is offset for clarity.
A synthetic water ice spectrum was used to recover the Mars water ice cloud shape, which closely matches both the 825 and 250 cm$^{-1}$ water ice absorptions [Curran et al., 1973]. Finally, a featureless blackbody shape was recovered using a synthetic blackbody shape. No other spectral shape could be recovered that was linearly independent from these three spectral shapes.

The original spectral set was deconvolved using the three recovered spectral end-members to obtain the concentrations of the dust and water ice end-members. The deconvolution provides a least squares fit to the original spectra using linear combinations of the spectral end-members. All spectra were fit with a RMS error of better than 0.010 and an average RMS error of 0.006. As the end-members were normalized to a low and high emissivity of 0 and 1, respectively, the derived concentrations represent the amount of absorption due to the spectral end-member at the wavelength of its greatest absorption.

The distribution of dust and water ice is shown in Figure 13. The dust concentration shows a strong anticorrelation with topography, as the lowest concentration corresponds to the summit of the volcano. The dust concentration increases along the flanks of Arsia Mons with increasing distance from the summit. The strong dependence on atmospheric path length provides an inverse topographic profile of Arsia Mons. Water ice cloud abundance is greatest in two closely spaced regions near the summit and south flank of the volcano. In addition, there is a small amount of water ice concentration derived for the entire region.

### 4.3. Dust Shape Variability

Dust and water ice shapes were recovered from several orbits that covered a range of atmospheric conditions from high dust and low water ice opacities (orbit 50, $L_s = 225$, local time 1600) to low dust opacities and ubiquitous water ice hazes (orbit 225, $L_s = 307$, local time 1000) (Table 1). The dust spectral shapes recovered from these orbits are very similar except in spectral regions with minor atmospheric CO$_2$ and H$_2$O gas absorptions, such as those near 960, 1060, and 1260 cm$^{-1}$ (Figure 14). To further accentuate this, the shapes of the dust recovered from orbits 50 and 225 were normalized outside of the CO$_2$ and H$_2$O gas absorption regions, and the orbit 50 dust shape was subtracted from the orbit 225 dust shape. The difference compares quite closely to a synthetic spectrum of CO$_2$ and H$_2$O gas [see also Maguire, 1977]. The variation in the observed dust spectral shape is small for the daytime, warm surface conditions used in this study.

### 4.4. Water Ice Shape

Figure 15 shows water ice spectral shapes recovered for orbits 50 and 225. While water ice is not nearly as abundant in either of these data subsets as in the orbit 55 data set discussed above, its contribution is still well above the noise, and its spectral shape is well constrained. Although the spectral shapes are a bit noisier than the dust spectral shapes because of the relatively lower abundance of water ice, there appears to be a real but small difference in the shapes between 800 and 940 cm$^{-1}$. The shape from orbit 225 (which is much more typical of ice shapes from orbit 55 and other orbits) has a definite absorption centered around 825 cm$^{-1}$, while the shape from orbit 50 decreases all the way to the last data point at 794 cm$^{-1}$.

### 5. Discussion

#### 5.1. Linearity of Mixing

The spectral character of the original TES spectra from the Arsia Mons data subset is well modeled. Any residual error can be attributed to either noise in the original spectrum or a small slope present in the residual (Figure 16). This slope can be attributed to the fourth minor component present in the system. There are several possible sources of slopes in the emissivity spectra which are often associated with temperature effects: (1) If the maximum emissivity is less than unity, the surface temperature is underestimated, and a slope effect is introduced because the wrong Planck function is used to derive emissivity. This is the case with mixtures of ice and dust clouds because their peak transmittances are at different wavelengths. (2) Variable atmospheric temperatures and atmosphere/surface temperature contrast can introduce a slope component because of varying amounts of atmospheric emission as a func-
tion of temperature and wavelength. (3) A nonisothermal surface can produce possible slopes, as it can no longer be modeled using a single Planck function.

The analysis of the Arsia Mons data subset has demonstrated that the TES spectra can be closely reconstructed using linear combinations of three end-member spectral shapes. While spectral mixing of atmospheric components can be a highly nonlinear phenomenon, the relatively low atmospheric opacities typical for Mars and the daytime, warm surface temperature conditions used in this analysis allow for an accurate linear mixing model to be used. These conditions are consistent with a large portion of the daytime data to be returned from the final 1400 LT mapping orbit of MGS.

5.2. Component Spectral Shapes

The recovered spectral shapes from the orbits examined indicate that the only variability of the dust spectral shape is the relative amount of atmospheric CO$_2$ and water vapor varying in unison with the dust in each orbit subset. This relation is easily understood in terms of the highly varying opacity of dust compared to the much more constant opacity of CO$_2$ and

---

**Figure 14.** Synthetic Martian atmospheric gas spectrum (top) and the difference (middle) between the recovered dust end-members from orbits 50 and 225. The recovered dust end-members (bottom) are from orbits 50, 61, 107, and 225 and have been normalized to the same depth outside regions of CO$_2$ hot bands and water vapor. Orbits 50 and 225 represent the two extreme cases where the gas bands are least and most prominent, respectively.

**Figure 15.** Recovered water ice spectral end-member for orbits 50 (solid) and 225 (dashed). The orbit 50 ice end-member has an absorption at $\sim$800 cm$^{-1}$ that is shifted to $\sim$850 cm$^{-1}$ in the orbit 225 water ice spectral end-member.
water vapor in these observations. Where there is more dust, the relative contribution of CO$_2$ and water vapor is less.

Toon et al. [1977] investigated the sensitivity of several parameters on several Mariner 9 infrared interferometer spectrometer (IRIS) brightness temperature spectra. They determined that the spectral shape is most sensitive to the particle size distribution between 1 and 10 µm and the composition of the dust. Therefore the consistent spectral shape of the dust that we observe in a wide variety of TES data indicates that the composition and particle size distribution averaged through the atmospheric column is nearly constant during the southern spring and summer periods, when major dust storm activity occurred [Christensen et al., 1998; Smith et al., this issue (a)].

There are differences in the position and shape of the absorption present in the water ice spectral shape (Figure 15). A larger particle size (changing the radius from 2 to 4 µm, for example) could shift the position of the absorption peak near 800 cm$^{-1}$ as observed [Curran et al., 1973]. However, the very close similarity of the spectral shapes between 230 and 530 cm$^{-1}$ argues against particle size variations. Another possibility is that the ice spectral shape from orbit 50 is influenced somewhat by nonlinearities in the dust spectral shape as the dust opacity in this data subset exceeds unity. In any case, the variation in the ice spectral shape is not large, with the possible exception of a small shift in position of the absorption peak near 800 cm$^{-1}$ for the extreme case of orbit 50.

5.3. Water Ice Cloud Detection

While the atmospheric dust is a ubiquitous component in the TES spectra, water ice clouds are not always present with sufficient abundance to be visually detected. Thin hazes of water ice clouds slightly depress the 800 cm$^{-1}$ emissivity but do not necessarily produce the obvious absorption feature observed in the orbit 55 data (Figure 10). The target transformation and end-member recovery techniques provide powerful tools for detecting water ice through the deconvolution of the TES spectra into their component end-members. This deconvolution also provides a first-order estimate of relative abundance that can be used to map water ice clouds in space and time (Pearl et al., submitted manuscript, 1999).

6. Conclusions

The work presented here has demonstrated the following.

1. The combined techniques of R-mode factor analysis, target transformation, and end-member recovery can accurately determine the number of independently variable components and recover the spectral end-members present in a set of mixed spectra.

2. Where differences occur between the actual spectral end-member in the mixed spectra and a spectral library end-member, the recovered spectrum is more accurate than the end-member present in the spectral library. If the spectral end-member is not present in the spectral library used as the set of trial spectra, an accurate end-member may still be recovered using a trial spectrum with broadly similar absorption band positions.

3. The initial application of this combined technique toward several subsets of TES emissivity spectra with limited surface components has shown that these spectra are composed of three independently variable spectral end-members: atmospheric dust, water ice clouds, and blackbody. Linear combinations of these three shapes can reconstruct the spectral features in the original data with a high degree of accuracy over a wide variety of atmospheric and topographical conditions.

4. Spectral shapes of both atmospheric dust and water ice aerosols have been recovered for a variety of orbits. The spectral shape of the dust is nearly constant for all the orbits examined (not including the difference in relative contribution of CO$_2$ and H$_2$O gasses that are included in the “dust” shape). This indicates that, during southern spring and summer, the composition and particle size distribution of the dust averaged through an atmospheric column did not change significantly, despite an order of magnitude variation in dust opacity [Smith et al., this issue (a)]. The spectral shape of water ice shows some variability of the shape and position of the absorption

Figure 16. Original TES spectrum (solid) from the orbit 55 Arsia Mons subset and the least squares fit (long dash) using the recovered spectral end-members shown in Figure 10. Noise and a slight slope account for most of the residual error (centered about 1).
near 800 cm$^{-1}$ during a period of high dust opacity, but is otherwise relatively constant.

Acknowledgments. We wish to thank Barney Conrath, John Pearl, Tim Titus, and two anonymous referees for helpful reviews of this manuscript. Michael Abrams and Raymond Arvidson also provided reviews of a previous version of this work which also significantly improved this manuscript. In addition, Greg Mehall, Kelly Bender, and Noel Gorelick provided much appreciated operations, processing, and software support for the TES instrument and data.

References


Christensen, P. R., J. L. Bandfield, M. D. Smith, and V. E. Hamilton, Identification of a basaltic component on the Martian surface from Thermal Emission Spectrometer data, J. Geophys. Res., this issue (c).


Smith, M. D., J. C. Pearl, B. J. Conrath, and P. R. Christensen, Mars Global Surveyor Thermal Emission Spectrometer (TES) observations of dust opacity during aerobraking and science phasing, J. Geophys. Res., this issue (a).

Smith, M. D., J. L. Bandfield, and P. R. Christensen, Separation of atmospheric and surface spectral features in Mars Global Surveyor Thermal Emission Spectrometer (TES) spectra, J. Geophys. Res., this issue (b).


J. L. Bandfield and P. R. Christensen, Department of Geology, Arizona State University, Tempe, AZ 85287-1404. (joshband@imap3.asu.edu)

M. D. Smith, NASA Goddard Space Flight Center, Code 690.2, Greenbelt, MD 20771.

(Received May 26, 1999; revised August 2, 1999; accepted August 5, 1999.)