Atmos. Chem. Phys., 13, 3133–3147, 2013
www.atmos-chem-phys.net/13/3133/2013/
doi:10.5194/acp-13-3133-2013

Atmospheric Chemistry and Physics
Open Access

Quantitative comparison of the variability in observed and simulated shortwave reflectance

Y. L. Roberts[1,2,3], P. Pilewskie[1,2], B. C. Kindel[2], D. R. Feldman[4], and W. D. Collins[4,5]

[1]Department of Atmospheric and Oceanic Science, University of Colorado – Boulder, Boulder, CO, USA
[2]Laboratory of Atmospheric and Space Science, Boulder, CO, USA
[3]NASA Langley Research Center, Hampton, VA, USA
[4]Lawrence Berkeley National Lab, Berkeley, CA, USA
[5]Department of Earth and Planetary Science, University of California – Berkeley, Berkeley, CA, USA

Correspondence to: Y. L. Roberts (yolanda.l.roberts@nasa.gov)

Received: 19 August 2012 – Published in Atmos. Chem. Phys. Discuss.: 26 October 2012
Revised: 14 February 2013 – Accepted: 22 February 2013 – Published: 15 March 2013

Abstract. The Climate Absolute Radiance and Refractivity Observatory (CLARREO) is a climate observation system that has been designed to monitor the Earth's climate with unprecedented absolute radiometric accuracy and SI traceability. Climate Observation System Simulation Experiments (OSSEs) have been generated to simulate CLARREO hyperspectral shortwave imager measurements to help define the measurement characteristics needed for CLARREO to achieve its objectives. To evaluate how well the OSSE-simulated reflectance spectra reproduce the Earth's climate variability at the beginning of the 21st century, we compared the variability of the OSSE reflectance spectra to that of the reflectance spectra measured by the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY). Principal component analysis (PCA) is a multivariate decomposition technique used to represent and study the variability of hyperspectral radiation measurements. Using PCA, between 99.7 % and 99.9 % of the total variance the OSSE and SCIAMACHY data sets can be explained by subspaces defined by six principal components (PCs). To quantify how much information is shared between the simulated and observed data sets, we spectrally decomposed the intersection of the two data set subspaces. The results from four cases in 2004 showed that the two data sets share eight (January and October) and seven (April and July) dimensions, which correspond to about 99.9 % of the total SCIAMACHY variance for each month. The spectral nature of these shared spaces, understood by examining the transformed eigenvectors calculated from the subspace intersections, exhibit similar physical characteristics to the original PCs calculated from each data set, such as water vapor absorption, vegetation reflectance, and cloud reflectance.

1 Introduction

Reflected solar radiation from Earth contains information about several variables relevant to changes in Earth's climate, including cloud properties, aerosols, land surface albedo, and sea ice (National Research Council, 2007; Loeb et al., 2007; Roberts et al., 2011; Wielicki et al., 2012). Changes in these and other atmospheric and surface variables impact the spectral, spatial, and temporal variability of reflected solar radiation through spectrally dependent scattering and absorption processes. Monitoring solar reflectance from space to study climate requires highly accurate, hyperspectral measurements (Wielicki et al., 2012). In this context hyperspectral refers to spectrally contiguous, overlapping spectral radiation measurements (Goetz et al., 1985; Goetz, 2009); solar (shortwave) radiation includes wavelengths ranging from the near ultraviolet to the near infrared, 300–2500 nm, accounting for approximately 95 % of the solar radiation incident at the top of the atmosphere. Since the 1970s, the information in shortwave hyperspectral measurements has facilitated the identification of individual surface materials and the application of sophisticated atmospheric correction techniques to obtain surface spectral reflectance (Goetz, 2009). The information about Earth's surface and

atmospheric properties in space-based hyperspectral shortwave measurements can also be used in climate change detection and attribution studies. This information in spectrally resolved shortwave radiation can be used to understand the variability of the climate system using spectral decomposition techniques such as principal component analysis (PCA) (Rabbette and Pilewskie, 2001; 2002; Grenfell and Perovich, 2008; Roberts et al., 2011).

Roberts et al. (2011) quantified the spectral variability of Earth-reflected hyperspectral solar radiance to study the information contained in direct satellite measurements for climate change detection and attribution. Although complete separation of all atmospheric and surface variables represented in a reflectance spectrum is challenging even using information-rich hyperspectral measurements, occasionally it is possible to spectrally identify the physical variance drivers such as clouds, sea ice, and vegetation using spectral decomposition techniques (Rabbette and Pilewskie, 2001, 2002; Huang and Yung, 2005; Roberts et al., 2011). For example, Roberts et al. (2011) applied PCA to Arctic Ocean radiance spectra, separating contributions to the data variance from clouds and sea ice. These results demonstrated that hyperspectral reflected radiation contains physical information about the Earth's climate system that can be extracted with multivariate spectral decomposition techniques.

Highly accurate climate observation systems are being designed that will include spectrally resolved measurements in the visible and near infrared. Such systems include the Climate Absolute Radiance and Refractivity Observatory (CLARREO) (National Research Council, 2007; Wielicki et al., 2012) and the Traceable Radiometry Underpinning Terrestrial- and Helio- Studies (TRUTHS) (Fox et al., 2003, 2011). The shortwave instruments proposed by both of these projects will provide high spectral resolution measurements with unprecedented absolute radiometric accuracy and SI traceability.

Feldman et al. (2011a) designed a climate Observation System Simulation Experiment (OSSE) as a CLARREO shortwave instrument emulator used to derive measurement and mission requirements. For the OSSE, Feldman et al. (2011a) used global climate model output with a radiative transfer model to simulate CLARREO shortwave instrument reflectance measurements. By comparing OSSE output from forced and unforced scenarios, changes in variables such as clouds, aerosols, sea ice, and snow cover were evident in zonally averaged spectra, implying that spectrally resolved reflectance may be capable of detecting changes in key climate variables by the middle and end of the 21st century (Feldman et al., 2011a). Using the OSSE, Feldman et al. (2011b) also found that spectrally resolved reflectance improves time-to-detection over broadband shortwave measurements. The results from the climate OSSE studies further support the need for a highly accurate climate observation system that includes hyperspectral shortwave reflectance measurements (Feldman et al., 2011a,b; Wielicki et al., 2012).

The climate OSSE is a powerful tool; however, we need to evaluate how realistic the variability of these simulated spectral reflectance spectra is relative to observations of spectral reflectance. The ability of the OSSE to reproduce present-day climate variability is necessary to use OSSE simulations to make confident statements about climate change detection and attribution. There is still the possibility, however, that even if the OSSE is able to meet this necessary condition of reproducing present day climate variability, its twenty-first century climate change predictions may not be realistic depending on how well the underlying climate model simulates future changes in climate. The spectral variability of simulated and observed hyperspectral reflectance can be compared both qualitatively (Feldman et al., 2011b), and quantitatively, by the methods presented here.

In this study, we evaluate how well simulated shortwave hyperspectral reflectance reproduces the variability in satellite-measured reflectance. To address this question we will explore the utility of the variability of shortwave reflectance to serve as an appropriate measure of the similarity between two data sets. Roberts et al. (2011) showed that it is possible to extract physical variables from directly measured radiance using principal component analysis rather than inverse modeling techniques or any other model-based analysis. Therefore, we compare the variability of measured and simulated reflectance using PCA and other multivariate analysis techniques to quantify their similarity.

The next section is an overview of the observed and simulated reflectance spectra used in this study. Section 3 details the multivariate techniques used in the comparisons. In Sect. 4, we present an example that exhibits the quantitative comparison techniques with data, and Sect. 5 provides a summary of the study, conclusions, and a discussion of future work.

2 Data

2.1 Observed reflectance – SCIAMACHY measurements

This study uses hyperspectral reflectance measured by the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) (Bovensmann et al., 1999). SCIAMACHY flew on the European Space Agency's ENVISAT (Environmental Satellite), a sun-synchronous satellite in near polar orbit that operated from March 2002 to April 2012. In May 2012, the European Space Agency declared the official end of the ENVISAT mission after a spacecraft failure in April 2012. SCIAMACHY was designed to study the effect of natural and anthropogenic sources on global atmospheric composition (Gottwald et al., 2011c). Additional objectives included understanding the global

distribution, chemistry, and physics of trace gases, aerosols, and clouds in the troposphere, stratosphere, and mesosphere (Bovensmann et al., 1999). SCIAMACHY measured across eight channels covering the spectral ranges 214–1773 nm, 1934–2044 nm, and 2259–2386 nm with spectral resolutions ranging between 0.22 nm and 1.48 nm (Gottwald et al., 2011a). Ice deposited on channels seven and eight (spanning 1934–2386 nm), interfering with the optical throughput (Gottwald et al., 2011b). For the present study, analysis is restricted to the wavelength range 300 nm to 1750 nm. Nadir pixel size is dependent on the integration time and swath width, causing footprint sizes to vary between 26 km (along track) by 30 km (across track) and 32 km (along track) by 930 km (across track). For nadir sampling, SCIAMACHY has a scanning angular width of ±32° across track, which corresponds to a maximum nadir swath width of 960 km (Gottwald et al., 2011a). The measurement characteristics of SCIAMACHY make it the best candidate to compare space-based observations of shortwave reflectance with climate OSSE-simulated reflectance spectra.

2.1.1 SCIAMACHY reflectance spectra

Figure 1 presents three examples of SCIAMACHY reflectance spectra for different scene types: a thick cloud (black), cloud-free green vegetation (blue), and cloud-free ocean (red). Spectral reflectance, R_λ, is defined by,

$$R_\lambda = \frac{\pi I_\lambda^{\uparrow}}{\cos(\theta) F_\lambda^{\downarrow}} \tag{1}$$

where $I_\lambda^{\uparrow}$ is reflected spectral radiance, $F_\lambda^{\downarrow}$ is the incident spectral solar irradiance at the top of the atmosphere, and θ is the solar zenith angle. Similarities among the three spectra in Fig. 1 include absorption features such as the oxygen-A band centered around 762 nm and water absorption bands centered at 940 nm, 1140 nm, and 1350 nm. There are also differences among these reflectance spectra that are characteristic of the different scenes. Throughout the visible and much of the near infrared outside water absorption bands, the reflectance values for the thick cloud spectrum are generally higher than the two surface reflectance spectra shown. For a cloud-free ocean spectrum measured at nadir, the reflectance values are low throughout the spectral range except between 300 nm to 400 nm where atmospheric molecular and aerosol scattering increase reflectance. The spectrum measured over green vegetation has low reflectance in the visible with a local maximum known as the "green peak", centered around 550 nm. The vegetation "red edge" is the increase in reflectance between 700 and 750 nm. These examples of reflectance serve as a frame of reference when examining the spectral shapes of principal components, such as those shown in Fig. 4, which often resemble reflectance spectra, albeit often with a linear and nonlinear mixture of source signatures.

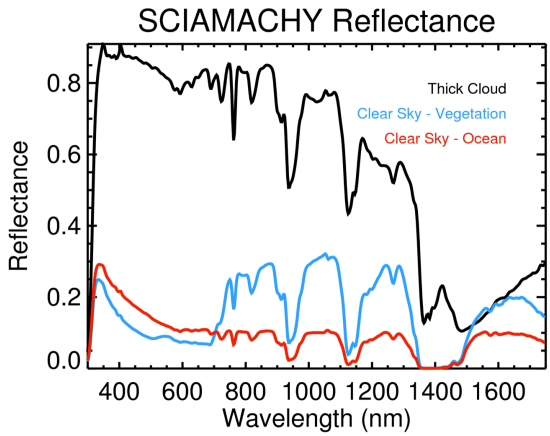

Fig. 1. SCIAMACHY-measured reflectance spectra for three scene types: a thick cloud (black), cloud-free vegetation (blue), and cloud-free ocean (red).

2.2 Simulated reflectance – Observation System Simulation Experiments

Feldman et al. (2011a) constructed OSSEs using input from the Community Climate System Model version 3.0 (CCSM) (Collins et al., 2006) Global Climate Model and using MODTRAN 5.3 (Berk et al., 2006) to simulate CLARREO shortwave spectral reflectance measurements during the twenty-first century. Monthly averaged fields from two IPCC AR4 emission scenarios were used to produce the OSSE results. The all-sky (cloud-inclusive scenes) reflectance spectra used in this study were simulated using the unforced constant CO_2 emission scenario model results, in which well-mixed radiatively active atmospheric greenhouse gases and aerosols were held constant at levels observed in the year 2000 throughout the model run (Meehl et al., 2005, 2007). Results from the forced A2 emission scenario (IPCC, 2007), in which concentrations of well-mixed greenhouse gases were steadily increased to the year 2050 then reduced to 1900 levels by the year 2100, were also used to simulate reflectance using the OSSE (Feldman et al., 2011a). Changes in the climate system over the course of the century include a tripling of CO_2 relative to pre-industrial levels, surface warming, and decreases in snow and ice cover. In the present study we have used the unforced scenario results because we are comparing individual months at the beginning of the 21st century, when differences between the two scenarios are minimal.

3 Quantitative multivariate methods

3.1 Principal component analysis

Principal component analysis (PCA) is a spectral decomposition technique used to quantify the variance distribution in a multivariate data set. The principal components (PCs)

are linear combinations of the original variables, in this case the spectral reflectance values. If the original variables are correlated (as is generally the case with reflectance spectra), PCA can significantly reduce the number of variables needed to explain the majority of the variance in a data set. Roberts et al. (2011) presented a literature review of how PCA has been used in various atmospheric science applications, specifically to understand the variability of spectral radiation in the longwave (longer than 4 µm) and the shortwave (350 nm–2500 nm).

3.1.1 PCA input

Shortwave reflected radiation provides a number of options for what form of the data to use in PCA, each of which has its own advantages. Roberts et al. (2011) used standardized spectral radiance rather than unstandardized radiance because of the large differences in spectral variance. Wavelength bands with the most variability in spectral radiance (largest spectral variance) dominate the spectral shapes of the PCs. By standardizing the data, principal components typically represent physical variables more prominently than PCs calculated from unstandardized data (Preisendorfer, 1988). Standardized radiances are calculated by spectrally mean-centering the radiances and normalizing them by the spectral standard deviation. The standardized PCs can be used when comparing two data sets, but not to quantify how much information is shared between them. Unstandardized data must be used in quantitative comparisons because the standard deviation contains pertinent information about each data set. Haskins et al. (1999) compared both standardized and unstandardized infrared radiance PCs, taking advantage of both the qualities of the standardized and unstandardized principal components. When the information in the standard deviation is removed, it becomes difficult to identify genuine differences and similarities between two data sets, the main purpose of this study.

Using reflectance rather than radiance is one way to avoid using standardized radiance while retaining the information in the standard deviation. Normalization by incident solar irradiance removes this known and dominant spectral shape from the PCs, leaving the influences due to atmosphere and surface properties. In the case of SCIAMACHY, the calculation of reflectance may also remove instrument anomalies because the downwelling TOA solar irradiance and Earth-reflected radiance are measured using the same sensors, in part diminishing the impact of spectrally dependent noise on the PCA results.

3.1.2 PCA descrption

The first step in principal component analysis is calculating the covariance matrix, **C**, from the spectrally mean-centered reflectance spectra. Using a spectral decomposition technique, the eigenvalues (Ω) and eigenvectors (**E**) of the covariance matrix are determined such that they satisfy the characteristic equation, $\mathbf{CE} = \Omega\mathbf{E}$, where **E** is a $K \times K$ matrix, and Ω is a $K \times K$ diagonal matrix composed of the eigenvalues (ω). K is the number of variables, or spectral bands, contained in each spectrum. Each eigenvalue, ω_k, is the variance explained by each eigenvector. In this study we define the principal components (PCs) to be the eigenvectors scaled by the square root of their corresponding eigenvalues:

$$\mathrm{PC}_k = \sqrt{\omega_k}\boldsymbol{E}_k \tag{2}$$

The spectral shapes of the principal components may provide insight into which physical variables are explained by each PC dimension. Projection of the mean-centered data onto the eigenvectors are the PC scores, the weighted averages of the input data with the eigenvectors as the weights. Depending on the spatial or temporal distribution of the data, the scores can be used to evaluate how the principal components vary in space or time. The physical significance of PCs is often difficult to determine because the PCs are linear combinations of the original variables. The original variables may have nonlinear dependencies on the physical variance drivers; however, some of the dominant physical variables can occasionally be identified. For a more detailed and mathematically rigorous description of PCA, consult Roberts et al. (2011) and Jolliffe (2002).

3.1.3 Boundary between data signal and noise

There are several methods that can be applied to estimate the number of dimensions that define the boundary between signal and noise in a data set (Jolliffe, 2002). There is no clear, quantitative boundary between signal and noise. To maximize the variance explained by each principal component, information from both the signal and noise are included in each eigenmode. Unless the noise variance exceeds the signal variance, the signal in the data typically dominates the variance explained by the first few eigenmodes. The best estimate of the boundary between signal and noise is the dimension at which noise begins to consistently dominate the variance, which may be difficult to determine at times because the noise and signal variances are not always strictly anticorrelated with increasing dimension number. This can make it difficult to assess if a particular dimension is dominated by signal or by noise. There are methods that can be used to more clearly separate the signal and noise in a data set, such as the Minimum Noise Fraction Transform (Green et al., 1988), which is discussed briefly in Sect. 5.

Cattell (1966) suggested using a plot of the eigenvalues on a linear scale to determine the location of this boundary. It is identified graphically by visually locating the initial change in slope in the eigenvalue spectrum. Craddock and Flood (1969) presented a similar technique, but instead used a logarithmic eigenvalue scale, a method that has been justified with the PCA of simulated data with known variance structures (Farmer, 1971). In studying the eigenvalue

spectrum calculated from PCA of solar radiation, we have found that the logarithmic plot of the eigenvalues is one of the best tools to identify how many PCs explain signal in the data. Kaiser (1960) suggests that all principal components associated with eigenvalues larger than the average eigenvalue explain the signal. A more liberal criterion can be used in which some fraction (typically 0.7) of the average eigenvalue is used as the cut-off, in an attempt to account for possible sampling variations (Jolliffe, 1972). In the Broken Stick Method (Jolliffe, 2002), $\omega_k > \frac{1}{K}\sum_{j=k}^{K}\frac{1}{j}\omega_{\text{avg}}$ is the criterion for determining the location of the boundary. Although the PCs are mathematically independent, it is still possible for the sampling distribution of each PC to be related to either of its neighboring PCs. One test of this separation is called the *North et al. Rule of Thumb* (North et al., 1982). This rule uses eigenvalue confidence intervals to determine if neighboring PCs are statistically separate from each other. If the neighboring eigenvalues fall outside the confidence interval of a given PC, then that PC is statistically different those on either side. The 95 % confidence intervals for the eigenmodes can be calculated using $\triangle\omega = z(0.975) \times \sqrt{\frac{2}{N}\hat{\omega^2}}$, where N is the number of observations (number of spectra in this study). This is another method that can be used to determine the approximate boundary between signal and noise; the boundary would be located where the eigenmodes are no longer statistically separated.

The subjectivity of locating the signal-noise boundary is recognized by all the studies discussed here, but these suggested techniques often help to provide guidance in making this decision. The information provided by these selection criteria help us to make sense of the subspace comparison techniques applied to the two data sets below. An approximation of the boundary between signal and noise in these two data sets puts the results from the significance test described in Sect. 3.2.4 into context as we will discuss in Sect. 4.

3.2 Quantitative comparison description

3.2.1 Comparing spectral variability in the infrared

Comparing the variability between two data sets of radiance data using second-moment statistics (statistics derived from the squared values of the data, such as variance) as an objective test of climate models is a technique that has been used by several others (Goody et al., 1998). For example, second-moment statistics have been used to evaluate climate model variability in temperature (Polyak, 1996) and longwave emission spectra (Haskins et al., 1997, 1999; Huang et al., 2002). All of these studies state that the model needs to exhibit correct second-moment statistics with respect to sufficiently accurate measurements to be considered rigorously validated. Haskins et al. (1997) compared the variance contribution and the spectral shapes of principal components to evaluate how well simulated infrared spectra reproduced the variability observed in Infrared Interferometer Spectrometer (IRIS) measurements. In a subsequent study, Haskins et al. (1999) inverted IRIS radiance principal components to derive cloud fraction, relative humidity, and temperature. Those principal component inversions quantified the constraints imposed upon climate models by infrared radiance measurements confirming that clouds are a dominant driver of the climate system and explain the largest fraction of the variance in the measured data (Haskins et al., 1999). Haskins et al. (1999) also concluded that if a model is unable to reproduce the observed cloud variability represented in the most dominant principal component, it is unlikely that it would simulate realistic changes in climate. Huang et al. (2002) combined principal component analysis with statistical regression techniques to quantify how well a GCM represented cloud variability relative to IRIS measurements and found that the model underestimated cloud variations by 2 to 6 times compared to measurements. The latter two studies converted the PCs of observed radiance to physical quantities using inversions and regression techniques to quantitatively evaluate the performance of climate models. In the present study, we only use the information provided by the principal components that explain nearly all the variance in the data set to evaluate simulated reflectance.

3.2.2 Quantitative comparison using subspace intersection

There are several methods that can be used to compare the variability of two data sets. Similar to variability analysis in the infrared introduced above, one method is to compare the spectral shapes of the components. This comparison can be helpful in that it is a preliminary, qualitative, representation of the relationship between the two data sets. Here, we examine the information shared by two data sets by applying a method similar to that used by Goetz et al. (1998) to develop a novel atmospheric correction lookup table method to retrieve AVIRIS surface reflectance. This method compared subspaces of the measured AVIRIS radiance spectra with that simulated by MODTRAN under a variety of atmospheric conditions. The spectral decomposition of the intersection between these subspaces determined how many dimensions the two data sets shared. The intersection was used as a transformation between the two data sets providing the means to relate the simulated atmospheric conditions with those observed in the AVIRIS spectra. The Goetz et al. (1998) primary objective was to develop a computationally efficient method of atmospheric correction for surface reflectance studies. In the present study, a similar mathematical framework is applied to determine how much of the total variance is shared between two data sets as a quantitative measure of their similarity.

Quantitative methods similar to those presented in this paper have been used in other areas of atmospheric science and other scientific fields to evaluate multivariate data (e.g. Krzanowski, 1979; Crone and Crosby, 1995). For example,

Crone and Crosby (1995) used the spectral decomposition of the intersection between subspaces of independent satellite measurements to determine their similarity. Determining the significance of the difference between two subspaces is instrumental for principal component regression because this decision gives guidance for determining if a subspace defined by one set of principal components is appropriate to explain the variability of another (Crone and Crosby, 1995; Jolliffe, 2002). The ultimate goal of the present study is to evaluate one data set based on its relationship to another, so we employ similar spectral decomposition analysis techniques.

3.2.3 Mathematical details of intersection decomposition

The intersection comparison method described here is largely derived from a technique described by Krzanowski (1979) for comparing groups of principal components. First, we calculate the principal components as described in Sect. 3.1.2. The following process is repeated p times, where $1 \leq k \leq p$, and p is some number less than the total number of PCs. Using the eigenvectors calculated from PCA, we calculate the intersection (**I**) between the two data sets using:

$$\mathbf{I} = \mathbf{E}_A \mathbf{E}_B^{\mathrm{T}} \mathbf{E}_B \mathbf{E}_A^{\mathrm{T}} \tag{3}$$

The intersection will be a $k \times k$ square matrix. The eigenvector matrices ($\mathbf{E}_A$ and $\mathbf{E}_B$) used to calculate the intersection are composed only of the k eigenvectors used to define the subspace. Singular value decomposition determines the eigenvalues (Γ) and eigenvectors (**Y**) of **I**. Because **I** is a symmetrical matrix, the two sets of eigenvectors calculated in this decomposition are equivalent:

$$\mathbf{I} = \mathbf{Y} \Gamma \mathbf{Y}^{\mathrm{T}} \tag{4}$$

The eigenvalues on the diagonal of the $k \times k$ diagonal eigenvalue matrix (Γ) can also be represented as a vector, Γ, k elements long.

The spectral decomposition provides information from which we can understand the amount of shared variance between the two subspaces. The eigenvector matrix, **Y**, is used to determine the transformed eigenvector matrices for each data set in the shared intersecting space:

$$\mathbf{A} = \mathbf{E}_A^{\mathrm{T}} \mathbf{Y} \tag{5a}$$

$$\mathbf{B} = \mathbf{E}_B^{\mathrm{T}} \mathbf{E}_B \mathbf{A} \tag{5b}$$

Each of the k vectors in **A** and **B** are mutually orthogonal and are used to understand the spectral nature of the overlap between the two data sets.

The eigenvalues in $\boldsymbol{\gamma}$ provide a measure of similarity between each pair of subspaces. If the sum of all the eigenvalues (also the trace of the intersection matrix in Eq. 4) is equivalent to the number of dimensions, k, included in the analysis, the two subspaces are equivalent. If the sum of all k eigenvalues is zero, then the two data sets are completely orthogonal and do not share any information.

3.2.4 Subspace similarity significance

We adopt the Crone and Crosby (1995) method for determining if two subspaces are significantly close at the 95 % confidence level using their distance. That is, if the distance between the two subspaces is significantly small, they are similar. The result from this significance test provides an upper limit for the number of dimensions that two data sets share and can be used as a guideline. This significance test determines how much of the total variance of each data set is shared between the two data sets. Determining if two subspaces are equivalent is not the same as concluding that the individual PCs are the same, nor is it equivalent to concluding that the covariance matrices calculated in the PCA process are equal to each other; rather, this analysis helps to determine to what degree the subspaces spanned by the k PCs are similar.

To address this question, we use a metric called the subspace distance, which is defined using the intersection eigenvalues:

$$D(\widehat{\mathrm{Obs}}, \widehat{\mathrm{Sim}})_k = \sqrt{k - \sum_1^k \boldsymbol{\gamma}_k} \tag{6}$$

where k is the number of PCs used to define the subspace and $\widehat{\mathrm{Sim}}$ and $\widehat{\mathrm{Obs}}$ represent the original simulated and observed reflectance data sets, respectively. The distance defined in Eq. (6) is the sample distance calculated from the spectral decomposition of the intersection between the two data sets. We use this sample distance to test the null hypothesis that the population distance between the two subspaces is zero, $D(\mathrm{Obs}, \mathrm{Sim})_k = 0$, against the alternative hypothesis that $D(\mathrm{Obs}, \mathrm{Sim})_k > 0$.

The distance metric is used in the triangle inequality to construct a confidence interval that tests the null hypothesis (Crone and Crosby, 1995). The form of the triangle inequality we use is:

$$D(\widehat{\mathrm{Obs}}, \widehat{\mathrm{Sim}})_k \leq D(\widehat{\mathrm{Obs}}, \mathrm{Obs})_k + D(\mathrm{Obs}, \mathrm{Sim})_k + D(\widehat{\mathrm{Sim}}, \mathrm{Sim})_k \tag{7}$$

and is rearranged to give:

$$D(\mathrm{Obs}, \mathrm{Sim})_k \geq D(\widehat{\mathrm{Obs}}, \widehat{\mathrm{Sim}})_k - D(\widehat{\mathrm{Obs}}, \mathrm{Obs})_k - D(\widehat{\mathrm{Sim}}, \mathrm{Sim})_k \tag{8}$$

where Obs and Sim are bootstrap-generated observed and simulated reflectance data sets. This equation allows us to estimate a one-sided confidence interval for the true parameter distance between the observed and simulated reflectance. To estimate the distributions of $D(\widehat{\mathrm{Obs}}, \mathrm{Obs})_k$ and

$D(\widehat{\text{Sim}}, \text{Sim})_k$, we generate new reflectance data sets with the same number of spectra, N, as the original observed and simulated data sets by using *bootstrap with replacement*. The bootstrapped data sets are formed by randomly selecting reflectance spectra in the observed and simulated sets of spectra until the newly formed data sets are the same size as the originals. This is done *with replacement*, meaning that it is possible for each spectrum to be chosen more than once. We perform this procedure using a random number generator.

We calculate the principal components for the bootstrap-generated data sets and find the intersections between the bootstrap-generated and original observed and simulated data sets. Then we calculate the distances between the generated and original data sets for each number of k components used to define the subspaces. We repeat this process 500 times to estimate distributions of $D(\widehat{\text{Obs}}, \text{Obs})_k$ and $D(\widehat{\text{Sim}}, \text{Sim})_k$, using those distributions to find the distances in the 97.5 percentile to use in Eq. (8) for the estimation of $D(\text{Obs}, \text{Sim})_k$. 500 repetitions was the value used by Crone and Crobsy (1995), but we also investigated the impact different numbers of repetitions had on the results, finding that at least 500 repetitions create continuous distance distributions. Creating these distributions with 1000 repetitions resulted in equivalent estimations of $D(\text{Obs}, \text{Sim})_k$ compared to using 500 repetitions. If $D(\text{Obs}, \text{Sim})_k > 0$, then the null hypothesis is rejected; otherwise, we fail to reject the null hypothesis. We can also think of these population distances as 95 % confidence intervals. If $D(\text{Obs}, \text{Sim})_k > 0$, then the confidence interval does not include zero; otherwise, the confidence interval does include zero, and it is possible for the population distance to be zero.

4 Hyperspectral reflectance variability

4.1 SCIAMACHY and OSSE data processing

Because we are quantitatively evaluating the similarity between the SCIAMACHY and OSSE reflectance spectra, it is important that the spectral, spatial, and temporal resolution and sampling of the two data sets are comparable. We created identically sized reflectance data sets by spectrally, spatially, and temporally resampling both of them and by including only averaged spectra located in grid boxes with data in both the SCIAMACHY and OSSE resampled data sets. Both sets of reflectance spectra were resampled to 10 nm full-width at half maximum spectral resolution and 3 nm sampling resolution. The OSSE spectra are produced using monthly averaged data on a 1.25° grid (Sect. 2.2). To ensure that SCIAMACHY pixels from at least every three days (the approximate time over which SCIAMACHY obtains near global coverage) throughout each month were represented in the monthly average of each grid box, we expanded the grid to 5.625°, four times the size of the original OSSE grid. We temporally aligned the data by calculating monthly averages of the SCIAMACHY reflectance by linearly averaging the SCIAMACHY pixels falling into each 5.625° grid box within each month.

It is a challenge to entirely eliminate the sampling differences between the two data sets. The OSSE spectra were generated using gridded input data from monthly averaged GCM output. The SCIAMACHY reflectance spectra, on the other hand, were instantaneous measurements from a satellite in sun-synchronous, near polar orbit. Even with the data resampling, inherent differences between satellite-measured and model-generated reflectance may remain because of the inherent nonlinearity in the equation of radiative transfer. The objective of the steps presented above is to mitigate the impact on the quantitative comparison due to sampling differences.

To understand the effect of computing comparable spatial grids and monthly averages, we performed PCA on all SCIAMACHY reflectance spectra measured in January, April, July, and October 2004 for each month separately. The eigenvalue spectra from these all-inclusive PCA results (Fig. 2: gray) show that the variance of the SCIAMACHY dominant modes is higher than when the spatial and temporal averages are computed. The shapes of the eigenvalue spectra in black, calculated from the resampled SCIAMACHY data are much closer to the shape of the red OSSE eigenvalue spectra, implying that the distribution of information is also more comparable after resampling. Despite the sampling differences between the SCIAMACHY and OSSE data sets, the spectral, temporal, and spatial resampling performed here aligns the distribution of the variability within the data sets, lending confidence to the appropriateness of the applied resampling.

4.2 Spectral reflectance variability

To illustrate the quantitative methods presented in Sect. 3.2.3 this study will focus on the four months for which we had daily SCIAMACHY data in 2004, January, April, July, and October as an initial evaluation of the OSSE performance at the beginning of the twenty-first century. Before employing the quantitative comparison tools described above, we first calculate the principal components from the unstandardized OSSE and SCIAMACHY reflectance spectra. The eigenvalues (i.e. the variance of each PC dimension) for each of the four cases are shown in Fig. 2. The shapes of the eigenvalue spectra show that the general distributions of the variance for both data sets are similar, at least for, approximately, the first 15 or 16 dimensions. The cumulative variance contribution in Fig. 3 shows some differences in variance in the first few PCs, but for both data sets and all four months, six PC dimensions explain between 99.7 % and 99.9 % of the total data variance.

In addition to studying the distribution of variance for the two data sets, we also examine the spectral shapes of the first several components that dominate the data variability.

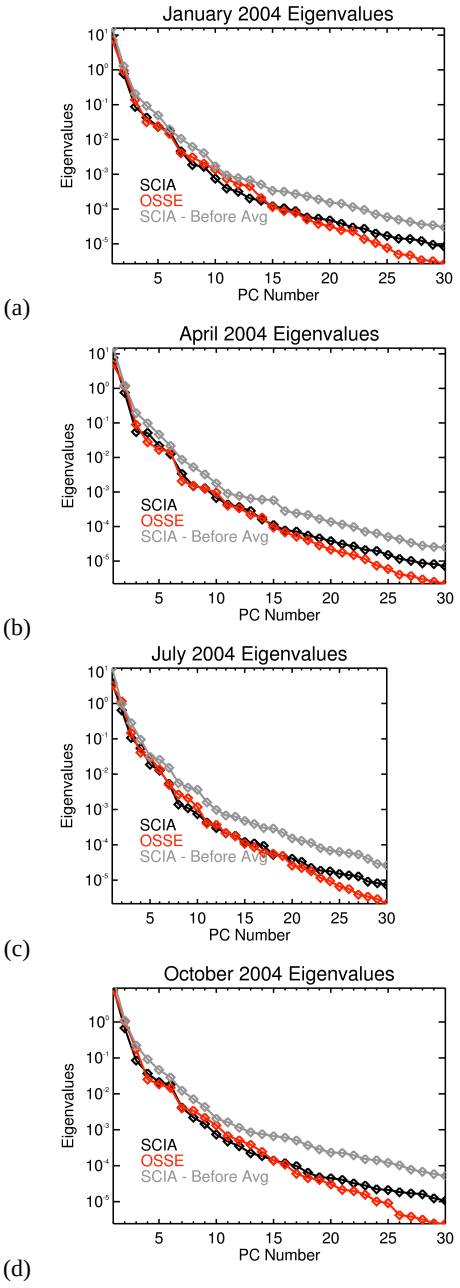

Fig. 2. The first 30 eigenvalues for the January **(a)**, April **(b)**, July **(c)**, and October **(d)** 2004 SCIAMACHY and OSSE Reflectance PCs. The difference in shape between each of the grey lines and the black lines shows how well the SCIAMACHY data resampling performed prior to PCA aligned the SCIAMACHY distribution of information with that of the OSSE.

Figure 4 compares the first nine October 2004 SCIAMACHY and OSSE PCs. Generally among the four cases the spectral shapes of the SCIAMACHY and OSSE components are very similar. The spectral shapes of the first two components are nearly identical for the four cases and together explain 94.7–97.5 % (OSSE) and 95.7–98.3 % (SCIA) of the data variance.

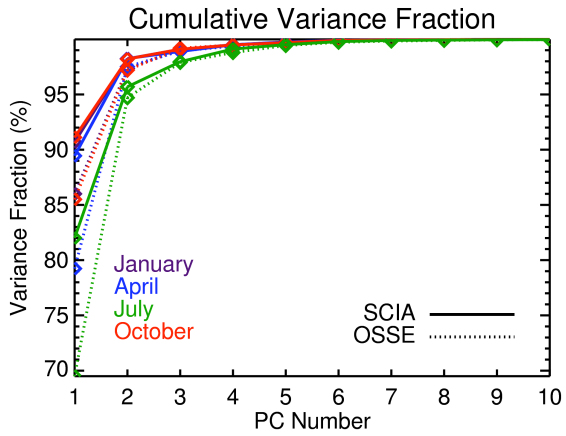

Fig. 3. The cumulative variance fraction for each of the four cases for the first ten PC dimensions. For all four months, both SCIAMACHY (solid) and OSSE (dashed), six PC dimensions explain between 99.7 % and 99.9 % of the total data variance in both data sets.

In addition to there being similarities between the PCs from the two data sets, there are spectral features that are indicative of physical variables. Water absorption bands are evident in at least the first four PCs for both data sets. The first PC resembles a cloud reflectance spectrum, and PC4 resembles a green vegetation reflectance spectrum (Fig. 1). It is likely that the other PCs explain physical variables but they cannot be uniquely identified. An illustration of this point is presented in Fig. 5, which shows the October 2004 OSSE and SCIAMACHY scores for PC4 and PC5. The spectral shape of PC4 is indicative of vegetation. This is confirmed in the spatial distribution of the scores by the relatively high scores over regions that are green in October such as the Amazon, the Southeastern US, sub-Saharan Africa, and Southern Asia. Moreover, negative scores are seen over areas typically devoid of green vegetation such as the oceans, polar regions, and semi-arid regions. Although similar spatial patterns are also observed in the PC5 scores, evidence that PC5 is partly explained by vegetation is not apparent by the spectral shape of PC5. This point also helps to support the importance of comparing entire subspaces when evaluating the data set similarity rather than solely relying on one-to-one PC comparisons.

There are some cases, however, in which individual comparisons of the PCs can reveal important differences between data sets. For example, although the first October 2004 SCIAMACHY PC contains some aspects of a cloud reflectance spectrum (Fig. 4), its spectral shape also contains characteristics of a frozen surface, such as ice clouds or ice or snow at the surface. The local maximum that occurs between 1400 and 1450 nm occurs because of its position between the water vapor absorption band centered at 1350 nm and the ice absorption band centered at approximately 1500 nm. Although it appears that the first OSSE PC

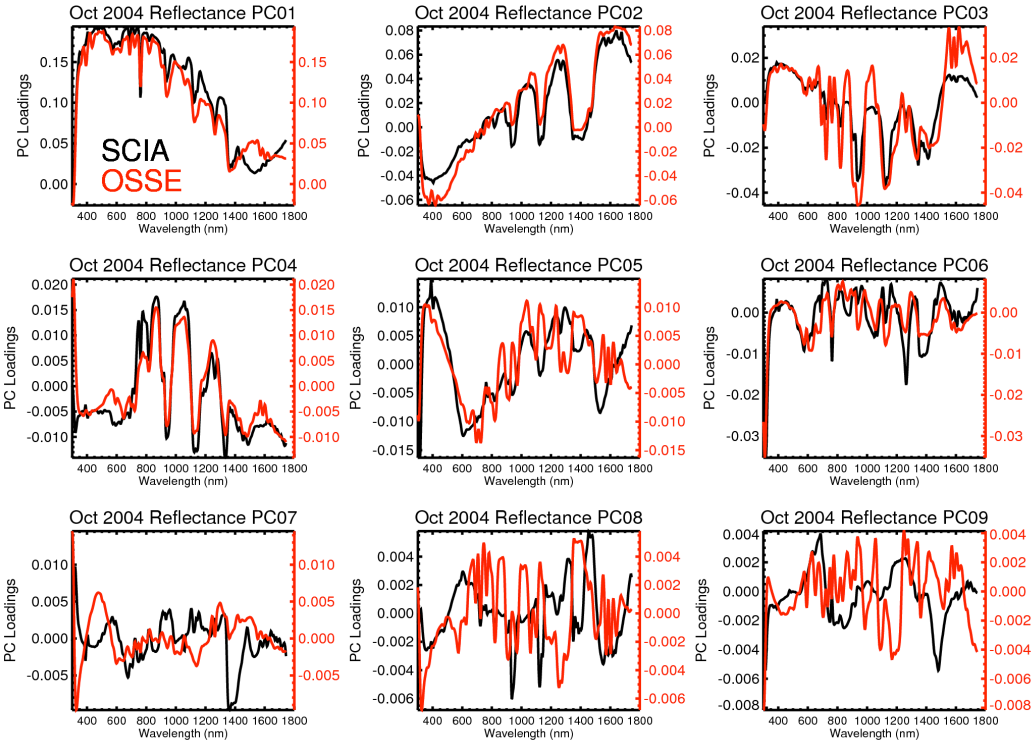

Fig. 4. The first nine October 2004 SCIAMACHY (black) and OSSE (red) principal components show a close comparison between the two data sets. Physical variables are identifiable within some PC spectral shapes including clouds (PC1), vegetation (PC4), and water absorption (PC1–4).

does not have the same ice spectral feature at 1400 nm as the first SCIAMACHY PC, this feature is in the OSSE PC, but it is broadened so that the peak occurs at a longer wavelength. We also see this difference in the January and April 2004 PC1 cases. It is likely that we do not see this ice feature and difference between the data sets in the July PC1 because of the reduction in Arctic snow and ice in July and the Antarctic night that occurs during this time. The way in which this feature is manifested in the PC may be representative of how snow BRDF values from MODIS are used as input for MODTRAN within the OSSE. The BRDF under snowy conditions was determined from snow-covered and snow-free MODIS surface reflectance and was created by linearly interpolating over the MODIS channels to obtain an estimate of the spectral BRDF function for each grid box. This estimate was input into MODTRAN. The necessary linear interpolation over the coarse band coverage in the near infrared may be the cause of the broadened ice feature around 1400, which is visible in the PC1 comparison in Fig. 4.

4.3 Quantitative subspaces comparison

Initial evaluation of the comparison of the eigenvalues and PC spectral shapes suggests that the variance distribution between these data sets is similar. To quantify how much of the variance is shared between the observed and simulated reflectance, we begin by using the selection criteria described in Sect. 3.1.3 to estimate the number of dimensions that define the boundaries between signal and noise. For example, using the October 2004 logarithmic eigenvalue plot suggested by Craddock and Flood (1969) (Fig. 2d) it appears that six dimensions may be sufficient to represent the signal explained by the variability in the data set, which is also how many dimensions the fractional Kaiser method (Jolliffe, 1972) (Sect. 3.1.3) suggested. The dip between the sixth and seventh eigenvalues and the change in slope before and after these eigenvalues likely indicates that the first six dimensions explain most of the variance in the data sets. This is also supported by the increasing amount of noise in the PC shapes after PC6 in Fig. 4. The Broken Stick Method (Jolliffe, 2002) suggested 14 dimensions, but the Broken Stick Method typically suggests the largest number of dimensions among the PC selection criteria described in Sect. 3.1.3. Using these criteria for the other three months as well, we have estimated that seven dimensions explain the reflectance signal for January and April, and eight dimensions explain the signal for July. Because the Broken Stick Method suggested that between 14 and 16 dimensions were above the noise level, we calculated intersections using the first twenty eigenvectors. Even though we estimated that fewer dimensions than define the boundary between signal and noise, we calculated 20

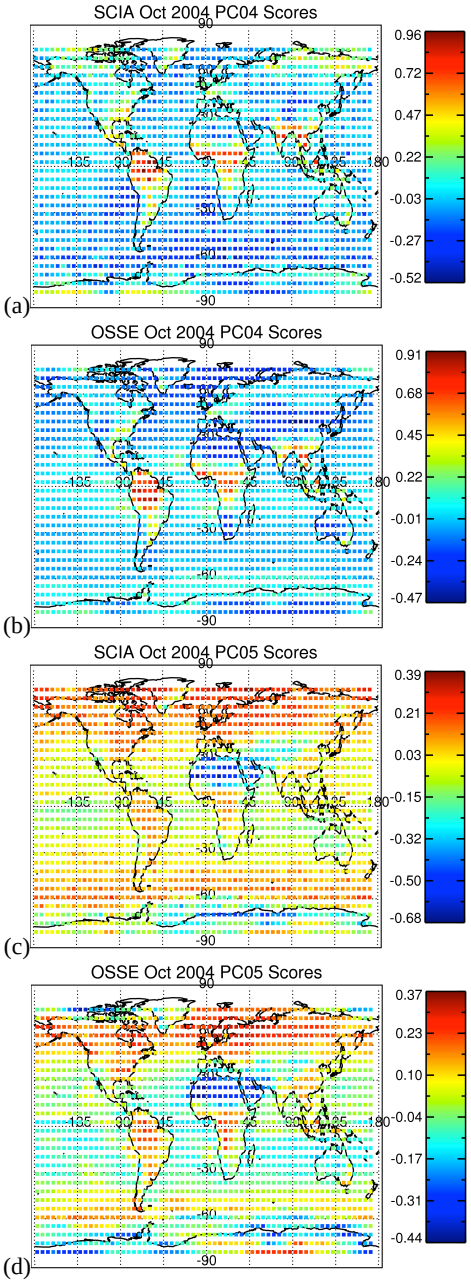

Fig. 5. The spatial distribution of the October 2004 PC4 SCIAMACHY **(a)** and OSSE **(b)** scores and the PC5 SCIAMACHY **(c)** and OSSE **(d)** scores. Both PC4 and PC5 scores show evidence of vegetation, implying that this physical signal is distributed between at least these two components. Data set similarities seen in the overlap of PC shapes in Fig. 4 are reinforced by the similarities seen in the spatial distribution of the scores.

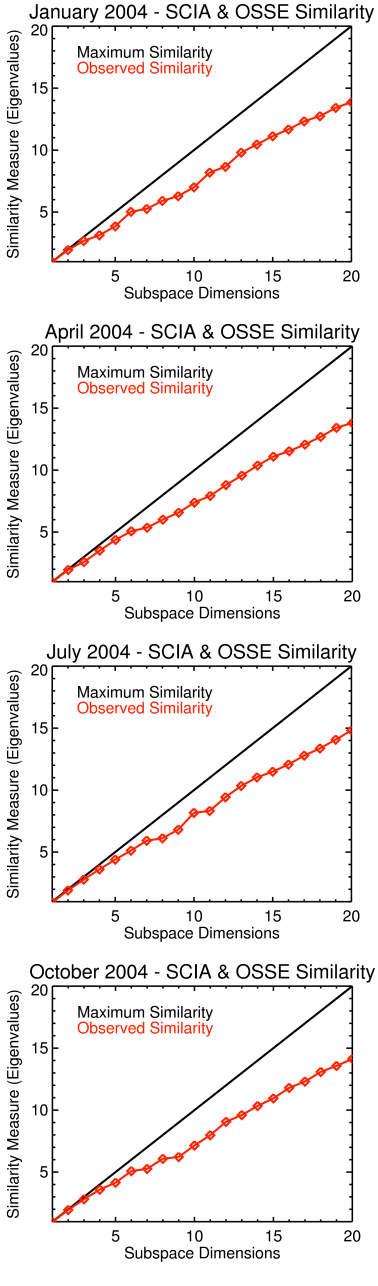

Fig. 6. The comparison of the eigenvalues from the spectral decomposition of the intersection to the maximum possible eigenvalue for each number of subspace dimensions. As more dimensions are used to define the subspaces, the similarity identified by the eigenvalues decreases.

different intersections with between one and twenty eigenvectors to find the number of dimensions at which the two data sets are different at the 95 % confidence level.

For each month, the twenty intersections were computed using the subspaces spanned by $1 \leq k \leq 20$ eigenvectors of the SCIAMACHY and OSSE data, and the spectral decompositions of each of these intersections were performed. Recall that the eigenvalues from the intersection decomposition are measures of similarity between the subspaces. The eigenvalues for each of these subspaces are shown in Fig. 6 for

each month, with the maximum possible similarity, k, shown in black. As k, the number of dimensions used to define each subspace, increases between one and twenty dimensions, the observed similarity between the two data sets decreases. This is illustrated by the increasing difference between the red lines and the black lines in Fig. 6. To quantify the largest number of dimensions that the two data sets share, we first calculate the distance between each set of subspaces. The calculated subspace distances, the maximum possible distances ($\sqrt{k}$) and the ratio between the calculated distance and the maximum distance are shown in Fig. 7. The subspace distances also confirm the result shown by the eigenvalues in Fig. 6, most clearly demonstrated by the relative distances, which generally increase with the number of dimensions used to define each subspace.

The subspace distances shown in Fig. 7 are the observed distances on the left side of Eq. (8). Continuing with the process, the triangle inequality is used to estimate the population distances, shown in Fig. 8. This statistical significance test shows how many k-dimensional subspaces are the same at a 95 % confidence level, and vertical lines indicate the largest k-dimensional subspace for which this is true for each case in Fig. 8. We also note that the selection criteria results using the logarithmic eigenvalue plot shown in Fig. 2 give similar values to those determined by the statistical significance test. The statistical significance test found that the two data sets agree over seven dimensions in April and July and eight dimensions in January and October. This alignment demonstrates that the two data sets are generally similar at the signal to noise boundary, discussed at the beginning of this section and estimated to be located at seven (January), seven (April), eight (July), and six (October) dimensions using the techniques described in Sect. 3.1.3. Using the cumulative variance explained (Fig. 3) by the number of dimensions indicated by the vertical lines in Fig. 8 we can determine how much OSSE and SCIAMACHY variance is explained in the k-dimensional space in which they are similar at the 95 % confidence level. The results in Figs. 3 and 8 show that for the number of dimensions over which the two data sets agree in January, April, July, and October, approximately 99.9 % of the SCIAMACHY and OSSE data variance is explained.

It is also informative to inspect the spectral shapes of each pair of transformed eigenvectors (Fig. 9). Using the October 2004 results from the statistical significance test, we show the transformed vectors of the eight-dimensional shared space between the SCIAMACHY and OSSE reflectance data. The first three eigenvectors exhibit several spectral characteristics that are also present in the original PCs in Fig. 4. The fourth transformed eigenvector in part resembles the original PC4, but the others contain only segments of recognizable spectral features, if any. Because some of the transformed vectors resemble the original PCs, this means that overlapping information between the two data sets is very similar to that of the original dominant modes of observed variability. By applying the intersection decomposition and the statistical significance technique, we have presented an objective method with which to quantitatively compare two multivariate subspaces, a technique which has several other applications, as described below.

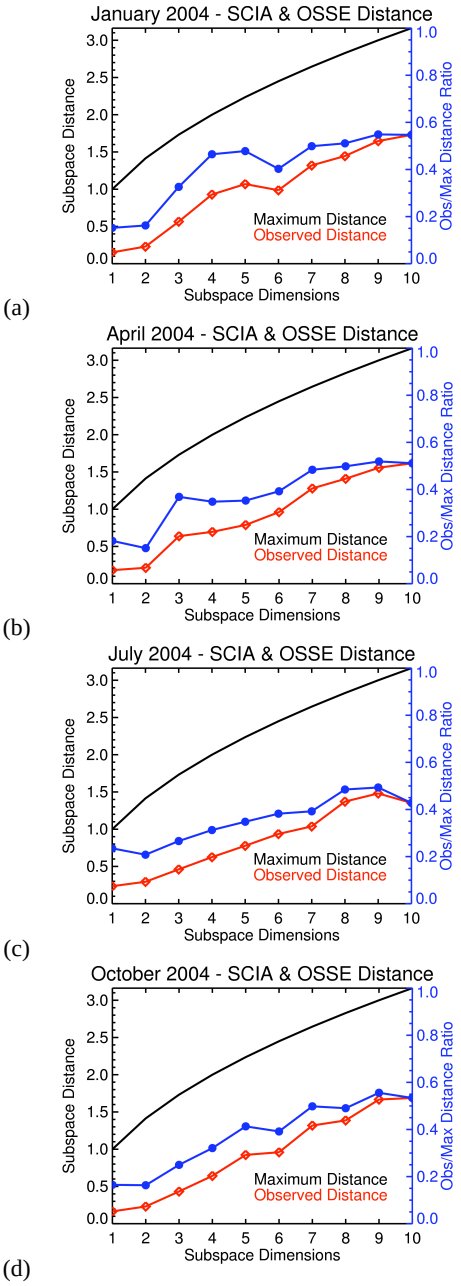

Fig. 7. The observed subspace distances between the SCIAMACHY and OSSE reflectance subspaces for ten subspaces (red) compared to the maximum possible distance between the two subspaces (black). The blue line shows the ratio between the observed subspace distance and the maximum possible distance for each number of subspace dimensions, that is, the ratio of the values on the red line to the values on the black line.

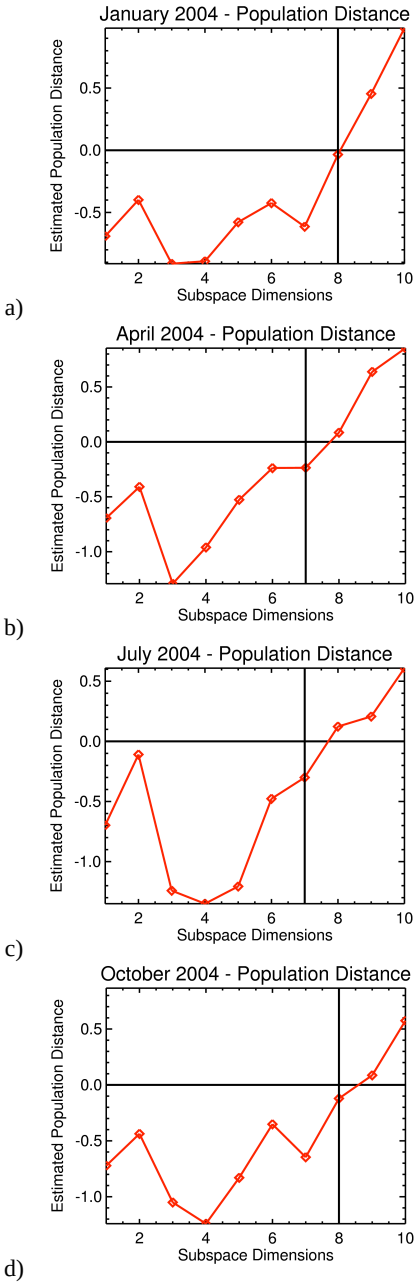

Fig. 8. The population distances between the OSSE-simulated and SCIAMACHY-observed reflectance spectra for January **(a)**, April **(b)**, July **(c)**, and October **(d)** 2004. The distances less than zero correspond to subspaces of k dimensions that are the same at the 95 % confidence level. The vertical lines indicate the maximum number of dimensions the two data sets share at the 95 % confidence level for each case.

5 Conclusions and future work

In this study, we used SCIAMACHY-measured hyperspectral solar reflectance to evaluate how well OSSE-simulated hyperspectral reflectance captures variability in the Earth's climate system. We presented two primary ways in which the information between two data sets can be compared. First, we qualitatively compared the most dominant principal components that explained the majority of the variance in both data sets and found that the two data sets appear to share similar variance distributions. We also found that linear interpolation of surface reflectance in the OSSE manifests as a difference in the first principal component for the January, April, and October cases. Second, we quantitatively compared the spectral variability of the two data sets using their principal components. This analysis showed that the OSSE and SCIAMACHY reflectance spectra share a large fraction of their spectral variability and that this variability shares spectral characteristics with the original PC transformation of the measured data set. From these results we conclude that at the beginning of the century, the OSSE appears to give a realistic representation of the Earth's variability relative to SCIAMACHY-measured reflectance. These findings provide a necessary, initial condition that helps us to understand the predictive potential of the OSSE for understanding how Earth's variability may change during the 21st century.

In Sect. 4.1, we discuss the differences between the SCIAMACHY and OSSE data sets despite our attempts to align the spectral, spatial, and temporal sampling. Our main objective in this study was to compare the variability of the two data sets, which we did using the intersection of the principal components calculated from the covariance matrices. It is possible that the sampling differences between the two sets of reflectance spectra were manifested as differences in the pairs of principal components, but despite those sampling differences, the spectral shapes of the principal components were qualitatively similar. The test that we proposed in this study to quantitatively compare the variability between two data sets does not rely on the similarity between each pair of spectra, nor does it evaluate the equivalency of the covariance matrices or the resulting principal components. Rather, this test evaluates the similarity of the subspaces that are spanned by some number of principal components. If the two subspaces are found to be statistically similar, we interpret this to mean that the variability of those subspaces is similar as well.

There are several other research questions that the quantitative comparison method applied in this study could be used to address. For the OSSE simulations used in this study, CCSM3 output was used, but other climate model results could also be used for the OSSE simulations. The comparison method presented here can provide rigorous objective testing of these different climate models to determine which model best reproduces Earth's present-day climate variability and is likely better to study future changes in Earth's climate.

Feldman et al. (2011a) used the OSSE to compare the shortwave reflectance signal observed between two different emission scenarios simulated using CCSM3: the constant CO_2 and the A2 emission scenarios. The quantitative method described in this study can be used to understand how changes in different climate forcing scenarios are manifested

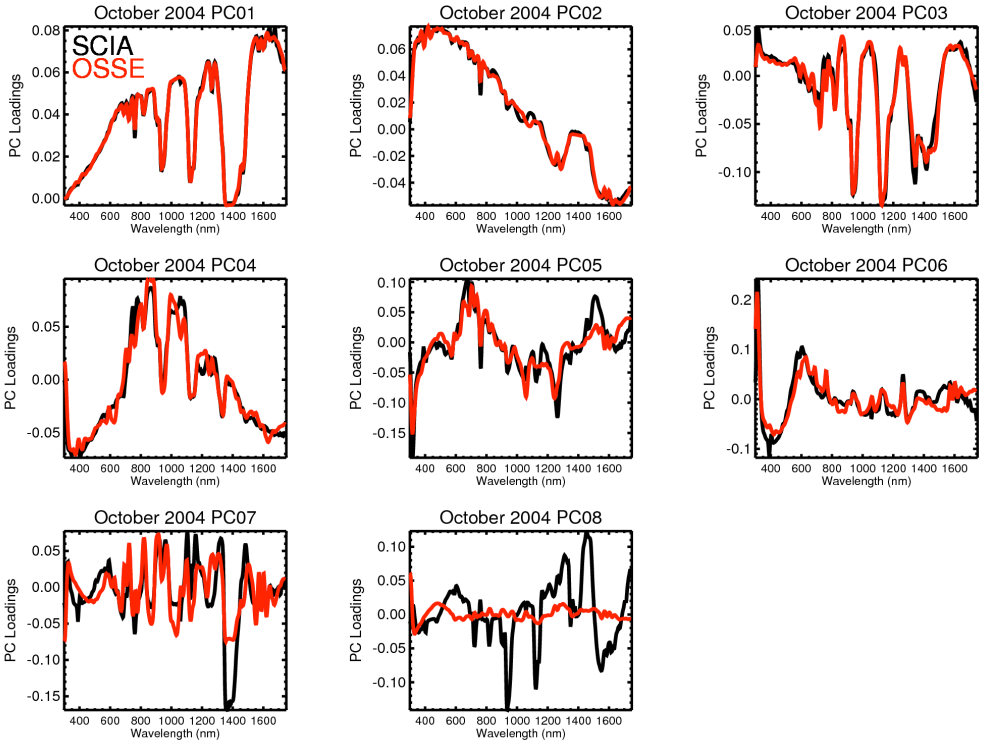

Fig. 9. The spectral shapes of the transformed eigenvectors for October 2004 in the shared intersection space. The first six or seven dimensions align well, and the first four dimensions share spectral characteristics with the first four original October 2004 PCs, shown in Fig. 4.

in the variability of hyperspectral reflectance. These results can be studied during the first decade of the 21st century, for comparison to SCIAMACHY reflectance, and during the entire 21st century, to attempt to understand how changes in climate contribute to changes in reflected shortwave spectral variability on a centennial time scale. This may provide insight into which variables contribute to changes in the measured reflectance over different time scales. In a subsequent paper, we will evaluate how well the OSSE reproduces the temporal variability of the Earth's climate system over the decade for which we have SCIAMACHY measurements.

In addition to the ideas presented above, there are other ways to improve and expand upon the analysis presented in this study. We focused on the similarities between the observed and simulated data, but it may also be useful to investigate the nature of the differences between the two data sets. One approach to address the differences in the variability between data sets would be to conduct a radiative transfer simulation study in which specific variables that may be the cause of variability differences were modified. For example, regarding the difference discussed in Sect. 4.2, we could rerun the OSSE using snow BRDFs with a higher spectral resolution than the current MODIS input to evaluate if such a change accounts for the difference observed between the first pair of SCIAMACHY and OSSE eigenvectors.

The Minimum Noise Fraction (MNF) transform (Green et al., 1988) is a method that can be used in conjunction with the comparison method described in this study. The MNF is a two-part PCA that whitens or decorrelates the noise in the data set, so if a well-defined noise characterization is available from noise-equivalent dark spectra during an instrument's lifetime, this transformation can be applied to the radiances before the quantitative comparison method is used. One of the benefits of the MNF transform is that it typically provides a clearer boundary between the signal and noise levels using the eigenvalue spectrum.

Another improvement to this work involves the method used to spatially and temporally resample the sun-synchronous satellite-measured reflectance. Although the methods aligned the spectral, temporal, and spatial sampling of the two data sets, it would be beneficial to establish more appropriate methods for gridding sun-synchronous satellite data to minimize the potential sampling differences in comparisons such as these. As climate observation systems are deployed, we will be able to apply the techniques described here to further improve the development of climate OSSEs as future instruments are designed. The results presented in this paper provide a foundation for how these quantitative comparisons between two hyperspectral data sets can be made. These results also provide the community with a measure of

how well the OSSEs are able to reproduce the variability of the Earth's climate system.

Acknowledgements. The authors would like to thank the European Space Agency (ESA) for their use of SCIAMACHY data ©ESA: 2009. The authors would also like to thank Xianglei Huang and another anonymous reviewer for their comments and suggestions, which improved the quality of this manuscript.

Edited by: Q. Fu

References

Berk, A., Anderson, G. P., Acharya P. K., Bernstein L. S., Muratov L., Lee J., Fox M., Adler-Golden S. M., Chetwynd J. H., Hoke M. L., Lockwood R. B., Gardner J. A., Cooley T. W., Borel C. C., Lewis P. E., and Shettle E. P.: MODTRAN5: 2006 Update, Proc. SPIE, 6233, 62331F, 2006.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: mission objectives and measurement modes, J. Atmos. Sci., 56, 127–150, 1999.

Cattell, R.: The scree test for the number of factors, Multivar. Behav. Res., 1, 245–276, 1966.

Collins, W., Bitz, C., Blackmon, M., Bonan, G., Bretherton, C., Carton, J., Chang, P., Doney, S., Hack, J., Henderson, T., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. J., and Smith, R. D.: The community climate system model version 3 (CCSM3), J. Climate, 19, 2122–2143, 2006.

Craddock, J. and Flood, C.: Eigenvectors for representing the 500 mb geopotential surface over the Northern Hemisphere, Q. J. Roy. Meteor. Soc., 95, 576–593, 1969.

Crone, L. and Crosby, D.: Statistical applications of a metric on subspaces to satellite meteorology, Technometrics, 37, 324–328, 1995.

Farmer, S.: An investigation into the results of principal component analysis of data derived from random numbers, Statistician, 20, 63–72, 1971.

Feldman, D. R., Algieri, C. A., Ong, J. R., and Collins, W. D.: CLARREO shortwave observing system simulation experiments of the twenty-first century: simulator design and implementation, J. Geophys. Res., 116, D10107, doi:10.1029/2010JD015350, 2011a.

Feldman, D., Algieri, C., Collins, W., Roberts, Y., and Pilewskie, P.: Simulation studies for the detection of changes in broadband albedo and shortwave nadir reflectance spectra under a climate change scenario, J. Geophys. Res., 116, D24103, doi:10.1029/2011JD016407, 2011b.

Fox, N., Aiken, J., Barnett, J., Briottet, X., Carvell, R., Frohlich, C., Groom, S., Hagolle, O., Haigh, J., Kieffer, H., Lean, J., Pollock, D. B., Quinn, T. J., Sandford, M. C. W., Schaepman, M. E., Shine, K. P., Schmutz, W. K., Teillet, P. M., Thome, K. J., Verstraete, M. M., and Zalewski, E. F.: Traceable radiometry underpinning terrestrial-and helio-studies (TRUTHS), Adv. Space Res., 32, 2253–2261, 2003.

Fox, N., Kaiser-Weiss, A., Schmutz, W., Thome, K., Young, D., Wielicki, B., Winkler, R., and Woolliams, E.: Accurate radiometry from space: an essential tool for climate studies, Philos. T. Roy. Soc. A, 369, 4028–4063, 2011.

Goetz, A.: Three decades of hyperspectral remote sensing of the Earth: a personal view, Remote Sens. Environ., 113, 5–16, doi:10.1016/j.rse.2007.12.014, 2009.

Goetz, A., Vane, G., Solomon, J., and Rock, B.: Imaging spectrometry for Earth remote sensing, Science, 228, 1147–1153, 1985.

Goody, R., Anderson, J., and North, G.: Testing climate models: an approach, B. Am. Meteorol. Soc., 79, 2541–2549, 1998.

Gottwald, M., Hoogeveen, R., Chlebek, C., Bovensmann, H., Carpay, J., Lichtenberg, G., Krieg, E., Lützow-Wentzky, P., and Watts, T.: SCIAMACHY – Exploring the Earth's Changing Atmosphere, Springer, New York, chap. 3, 29–46, 2011a.

Gottwald, M., Hoogeveen, R., Chlebek, C., Bovensmann, H., Carpay, J., Lichtenberg, G., Krieg, E., Lützow-Wentzky, P., and Watts, T.: SCIAMACHY – Exploring the Earth's Changing Atmosphere, Springer, New York, chap. 6, 77–97, 2011b.

Gottwald, M., Hoogeveen, R., Chlebek, C., Bovensmann, H., Carpay, J., Lichtenberg, G., Krieg, E., Lützow-Wentzky, P., and Watts, T.: SCIAMACHY – Exploring the Earth's Changing Atmosphere, Springer, New York, chap. 10, 175–216, 2011c.

Green, A., Berman, M., Switzer, P., and Craig, M.: A transformation for ordering multispectral data in terms of image quality with implications for noise removal, IEEE T. Geosci. Remote, 26, 65–74, 1988.

Grenfell, T. C. and Perovich, D. K.: Incident spectral irradiance in the Arctic Basin during the summer and fall, J. Geophys. Res., 113, D12117, doi:10.1029/2007JD009418, 2008.

Haskins, R. D., Goody, R. M., and Chen, L.: A statistical method for testing a general circulation model with spectrally resolved satellite data, J. Geophys. Res., 102, 16563–16581, 1997.

Haskins, R. D., Goody, R. M., and Chen, L.: Radiance covariance and climate models, J. Climate, 12, 1409–1422, 1999.

Huang, X. and Yung, Y. L.: Spatial and spectral variability of the outgoing thermal IR spectra from AIRS: a case study of July 2003, J. Geophys. Res., 110, D12102, doi:10.1029/2004JD005530, 2005.

Huang, X., Farrara, J., Leroy, S. S., Yung, Y. L., and Goody, R. M.: Cloud variability as revealed in outgoing infrared spectra: comparing model to observation with spectral EOF analysis, Geophys. Res. Lett., 29, 1270, doi:10.1029/2001GL014176, 2002.

IPCC – Climate Change2007: Synthesis Report, Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Pachauri, R. K. and Reisinger, A., IPCC: Geneva, Switzerland, 36–41, 2007.

Jolliffe, I. T.: Discarding variables in a principal component analysis. I: Artificial data, Appl. Stat., 21, 160–173, 1972.

Jolliffe, I. T.: Principal Component Analysis, Springer-Verlag, New York, 2002.

Kaiser, H.: The application of electronic computers to factor analysis, educational and psychological measurement, Educ. Psychol. Meas., 20, 141–151, 1960.

Krzanowski, W.: Between-groups comparison of principal components, J. Am. Stat. Assoc., 74, 703–707, 1979.

Loeb, N., Wielicki, B., Su, W., Loukachine, K., Sun, W., Wong, T., Priestley, K., Matthews, G., Miller, W., and Davies, R.: Multi-instrument comparison of top-of-atmosphere reflected solar radiation, J. Climate, 20, 575–591, 2007.

Meehl, G., Covey, C., McAvaney, B., Latif, M., and Stouffer, R.: Overview of the Coupled Model Intercomparison Project (CMIP), B. Am. Meteorol. Soc., 86, 89–93, 2005.

Meehl, G., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J., Stouffer, R., and Taylor, K.: The WCRP CMIP3 multi-model dataset: a new era in climate change research, B. Am. Meteorol. Soc., 88, 1383–1394, 2007.

National Research Council: Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, National Academy Press, Washington DC, 2007.

North, G. R., Bell, T. L., Cahalan, R. F., and Moeng, F. J: Sampling errors in the estimation of empirical orthogonal functions. Monthly Weather Review, 110(7), 699-706, 1982

Polyak, I.: Observed versus simulated second-moment climate statistics in GCM verification problems, J. Atmos. Sci., 53, 677–694, 1996.

Preisendorfer, R. W. and Mobley, C. D.: Principal Component Analysis in Meteorology and Oceanography, Elsevier, New York, 1988.

Rabbette, M. and Pilewskie, P.: Multivariate analysis of solar spectral irradiance measurements, J. Geophys. Res., 106, 9685–9696, doi:10.1029/2000JD900582, 2001.

Rabbette, M. and Pilewskie, P.: Principal component analysis of Arctic solar irradiance spectra, J. Geophys. Res., 107, 8049, doi:10.1029/2000JC000566, 2002.

Roberts, Y., Pilewskie, P., and Kindel, B.: Evaluating the observed variability in hyperspectral Earth-reflected solar radiance, J. Geophys. Res., 116, D24119, doi:10.1029/2011JD016448, 2011.

Wielicki, B. A., Young, D. F., Mlynczak, M. G., Thome, K. J., Leroy, S., Corliss, J., Anderson, J. G., Ao, C., Bantges, R., Best, F., Bowman, K., Brindley, H., Collins, W., Dykema, J. A., Doelling, D. R., Feldman, D. R., Fox, N., Huang, X., Holz, R., Huang, Y., Jin, Z., Jennings, D., Johnson, D. G., Jucks, K., Kato, S., Knuteson, R., Kopp, G., Kratz, D. P., Liu, X., Lukashin, C., Mannucci, A. J., Phojanamongkolkij, N., Pilewskie, P., Platnick, S., Ramaswamy, V., Revercomb, H., Rice, J., Roberts, Y., Roithmayr, C. M., Rose, F., Sandford, S., Shirley, E. L., Smith, W. L. S., Soden, B., Speth, P. W., Strow, L., Sun, W., Taylor, P., Tobin, D., and Xiong, X.: Climate Absolute Radiance and Refractivity Observatory (CLARREO): achieving climate change absolute accuracy in orbit, B. Am. Meteorol. Soc, in press, 2013.