A unified approach to infrared aerosol remote sensing and type specification

L. Clarisse1, P.-F. Coheur1, F. Prata2, J. Hadji-Lazaro3, D. Hurtmans1, and C. Clerbaux3,1

1Spectroscopie de l’Atmosphère, Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles, Brussels, Belgium
2Climate and Atmosphere Department, Norwegian Institute for Air Research (NILU) P.O. Box 100, Kjeller, 2027, Norway
3UPMC Univ. Paris 6; Université Versailles St.-Quentin, CNRS/INSU, LATMOS-IPSL, Paris, France

Correspondence to: L. Clarisse (lclariss@ulb.ac.be)

Received: 29 August 2012 – Published in Atmos. Chem. Phys. Discuss.: 11 October 2012
Revised: 18 January 2013 – Accepted: 12 February 2013 – Published: 25 February 2013

Abstract. Atmospheric aerosols impact air quality and global climate. Space based measurements are the best way to observe their spatial and temporal distributions, and can also be used to gain better understanding of their chemical, physical and optical properties. Aerosol composition is the key parameter affecting the refractive index, which determines how much radiation is scattered and absorbed. Composition of aerosols is unfortunately not measured by state of the art satellite remote sounders. Here we use high resolution infrared measurements for aerosol type differentiation, exploiting, in that part of spectrum, the dependency of their refractive index on wavelength. We review existing detection methods and present a unified detection method based on linear discrimination analysis. We demonstrate this method on measurements of the Infrared Atmospheric Sounding Interferometer (IASI) and five different aerosol types, namely volcanic ash, windblown sand, sulfuric acid droplets, ammonium sulfate and smoke particles. We compare these with traditional MODIS AOD measurements. The detection of the last three types is unprecedented in the infrared in nadir mode, but is very promising, especially for sulfuric acid droplets which are detected in the lower troposphere and up to 6 months after injection in the upper troposphere/lower stratosphere.

1 Introduction

Atmospheric aerosols consist of primary (sea spray, crustal material, smoke, and organic matter) and secondary (sulfates, nitrates, ammonia, volatile organic compounds) components (Kondratyev et al., 2005). Their presence reduces air quality, affecting human health, visibility and life in the whole biosphere (Pöschl, 2005). Aerosols also have an impact on the radiation budget of the Earth, causing a net cooling effect on the climate (Forster et al., 2007). The magnitude of the cooling is highly uncertain, but essential for a better understanding of ongoing climate change (Hansen et al., 2011). The reasons for the uncertainty is the large variability in aerosol: on a temporal, spatial and vertical scale and of composition, size, shape, chemical, physical and optical properties (Kaufman et al., 2002; Li et al., 2009). Conversely, the greenhouse budget is determined by slow varying long lived gases with definite optical properties. Another element is the great number of different ways in which aerosols can alter the radiative budget of the Earth (Ramanathan et al., 2007). These include direct interactions with solar and terrestrial radiation (Yu et al., 2006) and indirect effects through a multitude of different interaction mechanisms with(in) clouds (Haywood and Boucher, 2000; Li et al., 2011). These effects can only partially be measured in a direct way and therefore have to be assessed through careful modeling.

High temporally and spatially resolved space measurements of aerosols can greatly help in constraining such modeling efforts. Ideally, space measurements should not only include full information on the distribution and physical characteristics of aerosols, but also on their composition. Chemical composition is recognized as a key variable determining both the direct and indirect effects of aerosols on climate.
Knowledge of the composition allows to distinguish natural from anthropogenic aerosols, hygroscopic from hydrophobic aerosols and absorbing from scattering aerosols (Mishchenko et al., 2004; Kim et al., 2007a; Mishchenko et al., 2007).

Despite its importance, relatively little attention has been given to the sounding of aerosol composition. The most commonly derived quantity from space-based aerosol sounders is the so-called aerosol optical depth (AOD) which is a (vertically integrated) measure of how much radiation is absorbed and scattered. Fortunately, the scientific focus is broadening, and state-of-the-art aerosol sounders exploit increasingly many measuring dimensions, such as more spectral bins, multi-angle observations and polarization measurements; all of which can be used to infer information on aerosol composition. Multichannel measurements can be used to differentiate fine vs. coarse mode aerosol which in turn can be used as a proxy to differentiate anthropogenic vs. natural aerosols (Kaufman et al., 2005; Yu et al., 2006; Remer et al., 2005; Jones and Christopher, 2011). Estimates of the single scattering albedo (SSA), which is the ratio of the scattering versus total extinction efficiency, also yield clues. Sulfate and seasalt aerosol for example mainly scatter and have a SSA close to one, while black carbon aerosols mainly absorb and have a small SSA (Takemura et al., 2002; Kim et al., 2007a; Torres et al., 2007). Multi-angle observations allow to differentiate non-spherical (mainly dust and cirrus) from spherical particles (e.g. Kahn et al. (2009)). The reader is referred to Omar et al. (2009) and Kim et al. (2007a) and references therein for recent efforts of aerosol type specification using shortwave measurements.

POLDER (Tanré et al., 2011) is currently the most advanced aerosol sounder in orbit, measuring polarization and intensity in different spectral channels and in a multi-angle geometry. Recently, it was demonstrated (Dubovik et al., 2011) that this instrument can be used to derive aerosol (chemical) composition via retrieval of refractive indices. Building on the success of POLDER, the APS instrument was proposed (Mishchenko et al., 2007). It was specifically designed to address scientific questions regarding the impact of aerosols on the global climate. One of the key objectives was exactly the capability of measuring aerosol composition through determination of the refractive index (Mishchenko et al., 2004; Cairns and Mishchenko, 2011).

Remarkably, like most dedicated aerosol sounders (Kokhanovsky and de Leeuw, 2009), the APS design does not include any spectral channels in the thermal infrared. However using thermal infrared radiation for aerosol sounding has a number of appealing advantages such as (i) the possibility of measuring in absence of solar light, at night and in the winter at high latitudes (ii) less problems with retrieval over bright surfaces (iii) enhanced sensitivity to coarse mode aerosols and (iv) large sensitivity to aerosol composition (Clarisse et al., 2010a). This sensitivity to aerosol composition is caused by strong variations of the aerosol refractive index both as a function of composition and a function of wavelength in the thermal infrared spectral region.

In this paper we present a unified method for aerosol specification using only thermal infrared radiation and demonstrate it on the detection of five different aerosol types, some of which have never been identified directly from space in nadir mode. Others are observed with unprecedented sensitivity. Our results underline the large potential of aerosol sounding in the thermal infrared, and its complementariness to existing shortwave sounding methods.

In the next section we review the state of the art of aerosol detection methods from infrared sounders, which leads up to a unified detection method based on classical discrimination analysis. In Sect. 3 we apply these techniques on measurements of the Infrared Atmospheric Sounding Interferometer (IASI) (Clerbaux et al., 2009; Hilton et al., 2011). IASI is one of the most versatile infrared sounders currently in orbit, designed for both operational and scientific research goals, it has excellent temporal, spatial and spectral coverage and instrumental characteristics. The infrared thermal window is captured from 645 cm$^{-1}$ to 2760 cm$^{-1}$ at an apodized spectral resolution of 0.5 cm$^{-1}$ and low instrumental noise (< 0.3 K almost everywhere). We discuss and present results for volcanic ash, windblown sand, sulfuric acid droplets, ammonium sulfate and smoke particles. Where possible, we compare our results to standard (absorption) AOD measurements from OMI and MODIS. In Sect. 4 we present our conclusions.

## 2 Detection methods

### 2.1 Feature detection

Feature detection methods use elementary (arithmetic) operations on spectral bands or channels to produce detection flags. Such methods have been and are still widely applied on measurements from broadband sounders for the detection of aerosols. The simplest of these is a single threshold on a given band for the detection of thick high altitude clouds. At the other extreme there is, e.g. Ackerman et al. (1998) who published a sophisticated cloud test consisting of 11 different tests on single and pairs of different MODIS bands.

The most commonly used method for aerosol detection is the brightness temperature difference (BTD) between two different spectral bands. Inoue (1985) showed that the difference of the AVHRR bands at 11 and 12 µm is a good indicator for the presence of cirrus clouds. Analogously, Prata (1989) showed a reverse absorption effect between 10 and 11 µm for volcanic ash clouds. These papers also featured the first bispectral diagrams between the BTDs and a reference band. Such graphs can be used to infer quantitative information of two other independent variables such as effective particle radius and optical depth of mass (see Wu, 1987; Parol et al., 1991; Rose et al., 1995 for ice – Wen and
Rose, 1994; Yu et al., 2002; Prata and Prata, 2012 for ash). Similar techniques have been applied for detection of wind-blown sand (Ackerman, 1989) and extended to 3 (Ackerman, 1997; Ellrod et al., 2003; Strabala et al., 1994) or more bands (Pavoloni et al., 2006).

Detection techniques based on BTDs have also been applied on high spectral resolution measurements, e.g. for the detection of ice (Kahn et al., 2003), ash (Carn et al., 2005) and sand (DeSouza-Machado et al., 2006) but also for weak absorbing trace gases, such as sulfur dioxide (SO$_2$) (Clarisse et al., 2008) and ammonia (NH$_3$) (Clarisse et al., 2009). The use of BTDs on hyperspectral measurements has the advantage of being able to largely avoid contamination with (other) trace gases. However, by only using a handful of spectral channels they do not fully exploit all the information content captured in such measurements. A full multichannel extension of BTDs for the detection of ash was presented in Gangale et al. (2010) where first and second order polynomials were fitted to high resolution brightness temperature spectra. Detection thresholds could then be introduced on fitted parameters and goodness of fit (see also Newman et al., 2012).

2.2 Spectral fitting

Feature detection methods have a physical basis, and can therefore often be used to derive quantitative information. Because of their simplicity they can be employed in operational applications or when large amounts of data need to be processed with limited computational power. In contrast, the most sophisticated methods rely on spectral fitting, where the observed spectrum $y$ is matched to a calculated spectrum generated by a forward radiative transfer model $F_x$ and varying physical parameters $x$ (Rodgers, 2000). The physical parameters should include all unknown parameters which influence the forward model in the spectral range of interest.

Spectral fitting is typically ill-conditioned so that the problem needs to be constrained with prior information of the physical parameters $x_a$. Mathematically, the problem is then often formulated as a minimization of the weighted least-squares cost function

$$J = (y - F_x)^T S_e^{-1} (y - F_x) + (x - x_a)^T S_a^{-1} (x - x_a).$$

(1)

Here the weights are the covariance matrices of the instrumental noise ($S_e$) and of the prior information ($S_a$). A common assumption is that of moderate non-linearity of the forward model, in which case the solution $\hat{x}$ can be achieved by Newtonian iteration of

$$x_{i+1} = x_a + (S_a^{-1} + K_i^T S_e^{-1} K_i)^{-1} K_i^T S_e^{-1} [y - F_{x_i} + K_i (x_i - x_a)],$$

(2)

with covariance

$$\hat{S}^{-1} = K_i^T S_e^{-1} K_i + S_a^{-1}.$$  

(3)

Here $K$ is the Jacobian built up of derivatives $K_{ij} = \partial F_j (x_i)/\partial x_j$. Under the assumption that both the instrumental noise and prior information can be accurately described as Gaussian probability density functions (pdf), this solution can be shown to be the maximum a posteriori solution. Since this solution is the one that maximizes $P(x|y)$, it can be interpreted as the Gaussian probability density function with mean $\hat{x}$ and covariance $\hat{S}$ and the method is therefore often referred to as “optimal estimation”.

Note that there are many other (iterative) methods for spectral fitting (Rodgers, 2000), and that such methods are widely employed for the quantitative retrieval of major trace gases (see e.g. Hurtmans et al., 2012). They can also be used to detect weak absorbers, and they are the ultimate way of confirming the presence of their spectral signature in the observed spectrum (see Coheur et al., 2009 and Clarisse et al., 2011 who reported a series of rare trace gas observations from IASI).

A more qualitative way of spectral fitting was introduced in Walker et al. (2011) (see also in this context Rodgers, 2000, p. 70–71 and von Clarmann et al., 2001), who proposed a non-iterative pseudo retrieval of a single physical variable or target species $x$ using

$$\hat{x} = x_0 + (K^T S^{-1} K)^{-1} K^T S^{-1} (y - F_{x_0}),$$

(4)

where the covariance $S$ includes the instrumental noise and the covariance of all physical parameters except $x$. The Jacobian $K$ is the derivative of the target species with respect to a fixed average atmosphere. When all constants are omitted, and the column matrix $K$ is written as a vector $\mathbf{k}$, a quantity

$$R_1 = k^T S^{-1} y = \sum_i \lambda_i k^T v_i v_i^T y$$

(5)

is obtained, which can serve as a (non-normalized) measure for $x$. Here $S = \sum_i \lambda_i v_i v_i^T$ is the eigenvalue decomposition of $S$.

This approach is powerful when $x$ is a parameter which does not affect the observed spectrum in a typical atmosphere, as in that case it is straightforward to generate a covariance matrix $S$ from an ensemble of observed spectra. For instance, SO$_2$ is rarely observed with IASI, and the covariance matrix $S$ can be generated from randomly observed spectra subject to a simple BTD test on SO$_2$. Not having to retrieve or to model all the other important physical parameters is of great advantage, especially as this method can be applied to large spectral ranges, exploiting the full information content from high resolution sounders. Note that this approach is not equivalent to a full retrieval since (1) it assumes linearity, (2) a fixed Jacobian is used, and (3) it assumes a Gaussian distribution of the probability density function describing the instrumental noise and all other physical parameters. This last condition is valid to some extent, but will fail for outlying features (e.g. rarely observed events such as volcanic eruptions, peculiar surface emissivity effects), causing
false detections. With these caveats in mind, the development of the pseudo-retrieval method is an important breakthrough in remote sensing and is extremely powerful, in particular for the detection of spectral features spanning a large spectral range.

Spectral fitting approaches have also been applied to the detection and quantification of atmospheric aerosols. The problem is non-trivial and is made more difficult by the intricacies of aerosol radiative transfer which ideally includes the effects of multiple scattering. Another added complexity is the fact that the optical properties of aerosol are never known exactly. These depend on size, shape and composition (e.g. the spectral properties of windblown dust depend largely on their mineral content (Sokolik and Toon, 1999)) and can undergo transformations (such as due to particle coagulation) during transport in the atmosphere. This becomes especially important for high spectral resolution measurements. The infinity of different possible aerosol optical properties contrasts with trace gas spectroscopy for which only a limited number of molecular parameters need to be known. In addition to these difficulties, because of spectral interference, trace gases and aerosols should be retrieved simultaneously.

A complete simultaneous fit of trace gas concentrations, aerosol optical depths and effective radii for high resolution IASI spectra was presented in Clarisse et al. (2010a) for scenes with a different dominant aerosol type. It was shown there, that at least 5 different aerosol compositions can be differentiated spectroscopically (ice, biomass burning, sand, volcanic ash and volcanic sulfate). Another recent example of a full fit approach (on selected window channels, and without trace gases) is presented in DeSouza-Machado et al. (2010) for dust retrievals on AIRS spectra.

2.3 Distance approaches

An alternative to a full iterative fit is to use a large amount of precalculated spectra as a lookup table (LUT). Varying parameters can include the target molecule/aerosol loading but also surface temperature, interfering trace gases, aerosol radius, aerosol height and viewing angle. Observed spectra $y$ can then be matched to one of the precalculated spectra $y_i$ using, e.g. the Euclidean distance, possibly weighted by the instrumental noise covariance matrix $S_e$:

$$J' = (y - y_i)^T S_e^{-1} (y - y_i).$$

This distance also appears as the first term in Eq. (1) and is called the Mahalonobis distance (Rencher, 2002). Advantages of LUT approaches are that the obtained solution is guaranteed to be physically meaningful, and that a global minimum can be found of the difference observed-calculated spectra (this is not guaranteed using iterative spectral fit approaches). The LUT approach is especially appealing for the retrieval of aerosols for which the radiative transfer calculations are slow. Examples include retrievals of (cirrus) clouds (Li et al., 2005; Yue and Liou, 2009), dust (Peyridieu et al., 2010 and references therein) and volcanic ash and ice (Gangale et al., 2010; Clarisse et al., 2008; Corradini et al., 2010) from AIRS and IASI. LUT approaches have also been employed for the retrieval of rare trace gases, as in Prata and Bernardo (2007) who proposed a detection algorithm for SO$_2$ by looking for a good correlation between precalculated absorbance spectra of SO$_2$ and absorbance spectra obtained by dividing observed spectra in a given scene. Their method also used a LUT approach for determining SO$_2$ abundances.

A purely qualitative detection method based on distance was proposed in Clarisse et al. (2010b) for the detection of volcanic ash. Rather than using calculated spectra, a set of real observed spectra was used as a LUT. Observed spectra were then matched against these using the linear Pearson correlation as distance measure. Methods based on auto correlation have also been proposed, both for trace gases (Beer and Norton, 1987) and for clouds (Serio et al., 2000; Masiello et al., 2002).

2.4 Methods based on singular value decomposition and principal component analysis

The number of spectral channels in high spectral resolution instruments such as AIRS and IASI far exceeds the number of independent pieces of information contained in them (in a typical observation of a terrestrial atmosphere). Making a principal component analysis (PCA) (Jolliffe, 2002) of an ensemble of spectra is a way of reducing the dimensionality by extracting the principal components of spectral variation and disregarding those that carry no information (Huang and Antonelli, 2001; Antonelli et al., 2004; Klüser et al., 2011). This allows to remove instrumental noise as was illustrated in Atkinson et al. (2010) with a better detection of NH$_3$ by the application of a BTD filter on reconstructed spectra.

PCA can also be applied in a different way for the detection of trace gases or aerosols. Principal components should be calculated from a large number of random training spectra to accommodate for all observed variability. However, as discussed in Atkinson et al. (2010), very rare events (volcanic eruptions, large fires) will typically be reconstructed poorly as their weight is too low for their spectral features to be represented in the principal components. This opens up the possibility of using principal components as a detection tool by explicitly avoiding the presence of the target species in the spectra of the training set. This was for instance done for cloud detection in MIPAS observations (Hurley et al., 2009).

In particular, suppose we have a set of clear spectra $\{s\}$ of length $n$ (so an ensemble of spectra with no detectable spectral signature due to the presence of the physical variable or target species $X$). The principal components are the eigenvectors corresponding to the $m < n$ (to be chosen cleverly) largest eigenvalues of their covariance matrix $S$. Equivalently, they can be obtained from a singular value
The eigenvalue decomposition of the data matrix consisting of all the clear spectra. The eigenvalue decomposition of $S$ can be written as

$$S = \sum_{i=1}^{\text{rank} \ S} \lambda_i v_i v_i^T \approx \sum_{i=1}^{\text{rank} \ S} \lambda_i v_i v_i^T$$

The set $\{v_i\}$ form a complete basis of the observation space, so that the identity matrix can be written as $I = \sum_{i=1}^{\text{rank} \ S} v_i v_i^T$.

By projecting an arbitrary spectrum onto

$$S' = \sum_{i=m+1}^{n} v_i v_i^T = I - \sum_{i=1}^{m} v_i v_i^T,$$

the clear component of the spectrum is disregarded. Doing this on a large number of polluted spectra (these can be simulated or observed but should exhibit a signature due to $x$), we obtain a set of spectra for which again the principal components can be calculated. Projection of an arbitrary observed spectrum $y$ onto the vector $I$ corresponding to the first principal “polluted” component, then gives a quantitative indication of the presence of the pollutant. This can be written as

$$R_2 = I^T y = I^T S' y = \sum_{i=m+1}^{n} I^T v_i v_i^T y.$$

Here the notation was chosen to make the relation with Eq. (5) apparent. In Eq. (9) the pollutant vector $I$ could be replaced by the Jacobian $k$ from Eq. (5) since it is being projected onto the polluted space. Both the pseudo retrieval method and the PCA method project the spectrum on the eigenbasis formed by the covariance matrix of an ensemble of clear spectra. They differ in the applied weights: in Eq. (5) the weights are $1/\lambda_i$ and thus inversely proportional to the eigenvalues, while in Eq. (9) the weights are $0$ (for the largest eigenvalues) and $1$ (for the others). So while both methods are clearly related, they will lead to different results. The advantage of the PCA method is that the Jacobian does not need to be known (it can be estimated from the set of polluted observed spectra); while the advantage of the pseudo retrieval method is that it exploits better the full space and does not depend on ad-hoc choices such as the number of principal components.

### 2.5 Use of geophysical information

The detection methods we have discussed until now work on a single spectrum basis. A great deal of information can be extracted by considering ancillary information, in particular observations adjacent in time or space. An example of a detection algorithm which relies heavily on time context is the infrared difference dust index (Legrand et al., 2001; Vergé-Dépré et al., 2006). Here for each location, observations of the past 15 days are taken into account to determine the unpolluted background. The ash correlation method (Clarisse et al., 2010b) takes advantage of spatial context by applying a weaker detection threshold in the neighborhood of certain detections. Another example of an algorithm which explicitly uses spatial context was presented in Watkin et al. (2003), which looks for typical volcanic cloud shapes (e.g., a downwind plume) across infrared images from Meteosat. Similarly, isolated false detections can be avoided by using a median or despeckle filter (Pavolonis and Siegaff, 2010). Pergola et al. (2004) use a climatology of a region to determine a background of natural variability which may be compared to the signal containing the anomaly (e.g., volcanic ash). The method, termed the Robust AVHRR Technique (RAT), uses statistical measures and is self-adaptive with dynamic thresholds requiring no a priori assumptions. With the increasing use of high temporal and spatial remote sensors, the use of context in detection algorithms should play an increasingly important role.

### 3 A general approach

In this section we use basic results of classical discriminant analyses to better understand and generalize some of the detection methods outlined in the previous section. For a complete account of this type of supervised classification we refer to Ripley (1996) and McLachlan (2004). From now on, we assume we deal with the problem of detection of aerosol from high resolution infrared measurements.

#### 3.1 First pillar: discriminant analysis

Suppose we have a large training set of observations (here spectra) $y$ which can be subdivided in classes or groups based on certain criteria of the observed scene (here aerosol composition). Discrimination analysis is then concerned with the allocation of arbitrary observations to the different classes. For each class $c$ we can calculate a mean spectrum $\mu_c$ and covariance matrix $S_c$. If the associated probability density function $p_c$ is Gaussian and the training set is large enough, then these parameters describe the different classes completely. We denote the prior probability of a randomly chosen observation $y$ to be in class $c$ as $p_c$. Using Bayes’ formula it is straightforward to write down the probability $p(c|y)$ that $y$ belongs to class $c$:

$$p(c|y) = \frac{p(c)p(y|c)}{p(y)} = \frac{\pi_c p_c(y)}{\sum_c \pi_c p_c(y)}$$

with

$$M_c(y) = [(y - \mu_c)^T S_c^{-1} (y - \mu_c)]^{1/2}$$

the Mahalanobis distance with respect to class $c$. Note that the summations with index $d$ in these equations are summations over all the different classes. Based on this formula we
can assign an observation to the class $c$ with the largest probability $p(c|y)$, leading to the so-called Bayesian discrimination. Note that this allocation rule is quadratic in $y$ and is therefore also referred to as quadratic discriminant analysis.

This quadratic rule can be simplified under certain assumptions. The prior probabilities are usually not known or hard to determine and are often taken equal $\pi_c = \pi_d$. When we also assume that the covariance matrices of the different classes are equal $S_c = S_d = S$, then the allocation rule reduces to assigning the observation to the class with smallest Mahalanobis distance. When comparing two different classes, the term $y^T S^{-1} y$ cancels out and the allocation rule becomes linear in $y$. In particular, for two groups $c$ and $p$ we have a linear allocation rule of the form

$$R_3 = m^T S^{-1} y = (\mu_c - \mu_p)^T S^{-1} y < \text{constant} \quad (13)$$

While it is easy to write down the analytical expression for this constant, in practice it is usually tuned manually on observations with known class assignments. This two-class discrimination rule also appears as a special case of the so-called Fischer discriminant analysis (or canonical variate analysis). This equation is the basis of our unified method. We will not use any multiclass discrimination techniques here, but will address the problem indirectly by combining two-class discrimination rules. In particular, we design two-class discrimination rules for each type of aerosol, where in each case the rule attempts to differentiate spectra containing a signature of a specific type of aerosol (polluted spectra $p$) with those that do not (clear spectra $c$).

Equation (13) will only be a good discriminator if the covariance matrix of clear spectra $S_c$ and polluted spectra $S_p$ are equal to each other. This is fortunately the case for low aerosol loadings (so those that are hardest to detect) as can be seen as follows. $S_c$ should be constructed from a representative set of clear spectra not containing a specific type of aerosol. So this covariance matrix will contain the variability and correlation due to all atmospheric parameters affecting the spectrum, except the specific aerosol type. Now for low aerosol loadings we have $S_p \approx S_c$ as the presence of a thin aerosol layer will affect the spectrum but will not affect the natural variability and correlations due to all other atmospheric parameters. In other words, the pdf of moderately polluted and clear spectra differ only in their mean and not in their covariance. For increasing aerosol loadings the covariance matrix $S_p$ will have a component due to aerosol covariance, but no problems are expected here, since the normality assumption is not critical for the detection of their large signature.

The similarities between Eqs. (5) and (13) are no coincidence as they are both derived from maximum likelihood under the assumption of normality. The covariance matrices appearing in both equations can be constructed from clear spectra (in the language of Walker et al. (2011), the ensemble approach). The vector $m = \mu_c - \mu_p$ can be interpreted as a Jacobian with respect to a changing aerosol loading. This makes the link with Eq. (5) complete. Note that Eqs. (5) and (13) complement each other. Equation (5) can be used when no large number of polluted training spectra are available, but where the Jacobian is known from laboratory measurements (e.g. for sulfuric acid aerosols). Equation (13) can be used if no representative Jacobians are known, but where numerous training spectra are available (e.g. for volcanic ash), and thus shares this advantage with PCA-type approaches.

To more easily interpret the quantities $R_1$ or $R_3$, it is convenient to scale and normalize them. We therefore use

$$R_N(y) = \frac{k^T S^{-1} (y - \mu_c)}{\sqrt{k^T S^{-1} k}}, \quad (14)$$

with $k$ the Jacobian or when not available the difference $\mu_p - \mu_c$. For clear spectra, this quantity has a mean of zero and a standard deviation of one. This normalization is done implicitly in (Walker et al., 2011) with the use of $Z$-numbers. An allocation criterion is found by comparing $R_N$ to a predefined threshold, which following our normalization can be expressed in standard deviations. So for example 68.3% of unpolluted spectra will have an $R_N$ within $\pm 1$ and 99.7% of unpolluted spectra will have a value of $R_N$ within $\pm 3$. In practice (as also explained below), as we are never really dealing with a normal distribution the detection threshold is choosing manually by looking at a large number of assumed clear and polluted observations. Also note that the expected standard deviation (which we can interpret as an error) of $R_N$ is exactly 1.

In Walker et al. (2011, 2012), the quantity in Eq. (4) is called an apparent column (see the discussion leading up to Eq. (4)), and the same can be said of $R_N$. Even though it is scaled and normalized, $R_N$ is linear to the column amount when the magnitude of the aerosol extinction in the spectrum is directly proportional to the column loading. There are three conditions necessary for this apply. (1) fixed atmospheric conditions. Values obtained in different atmospheric conditions cannot be compared directly, as the magnitude of the aerosol extinction then also depends on atmospheric temperature, plume altitude, thermal contrast and interference with other molecules. For low altitude plumes the thermal contrast is probably the most important parameter, while for higher altitude plumes the dependence on plume altitude is largest. (2) constant aerosol properties. Aerosol properties like size can affect the absorption/scattering efficiency greatly, and therefore again the signal strength. (3) non-saturated regime. As aerosol loadings within a given plume increase, the plume will gradually turn opaque. A saturation effect occurs first in spectral regions with large extinction, where extinction has reached its maximum magnitude (which will again depend on other atmospheric parameters). Once saturated, a further increase in aerosol column will not result in a further increase of $R_N$. Indeed for a completely opaque aerosol plume it is
even possible to have a decrease in the aerosol extinction signature because of saturation throughout the thermal infrared region, in which case the aerosol plume will appear as a black or gray body. For the vast majority of aerosol plumes observed however, this saturation regime rarely occurs, and the magnitude of the aerosol signature is proportional to the magnitude of aerosol extinction and therefore to $R_N$. In view of these remarks, we conclude by saying that the quantity $R_N$ is broadly speaking proportional to the aerosol extinction and scaled and normalized in such a way that its value expresses the distance in standard deviations from the unpolluted observations used in the training set.

3.2 Second pillar: Mahalanobis distance

Despite their name, outlying events occur frequently in spectral observations. The pdf of the ensemble of spectra is not Gaussian, and large pollution events or peculiar type of surfaces will in general not be described well in the Gaussian mean and covariance matrix. Such events need to be dealt with when applying the above techniques. The problem can be easily understood with an analogy. Suppose we have a linear discrimination routine to differentiate apples and pears based on a few features (size, weight, color, best matching shape, ...). What will the algorithm do when we present it with an outlier, such as a banana? Since the observation space is cut in two by the linear discrimination, it will be classified as either an apple or pear based on the Mahalanobis distance; even if that distance is very far from either mean. To avoid such obvious misclassifications, we suggest to combine the linear discrimination approach – which serves as a relative distance measure, with a test based on the absolute distance of an observation to the class mean. In the analogy, it makes common sense to classify a fruit as an apple only if it looks at least a little bit like an apple. As a distance measure we use the Mahalanobis distance with respect to the mean polluted spectra.

$$A(y) = (y - \mu_p)^T S^{-1} (y - \mu_p),$$

(15)

This is akin to the ash correlation algorithm (Clarisse et al., 2010b), but of course here we want this distance not to serve as the primary criterion for detection, but only to remove some of the false detections. Note also that as a substitute of $\mu_p$ we can use $k + \mu_c$. As for the relative distance $R_N$ it is convenient to normalize the absolute distance. We use

$$A_N(y) = \frac{(y - \mu_p)^T S^{-1} (y - \mu_p)}{N},$$

(16)

with $N$ so that for unpolluted spectra the mean of $A_N$ equals one. As this quantity is used as a secondary criterion (to exclude anomalous observations), the threshold value for which we reject observations $A_N$ does not need to be very small. A value of 0.5 or 1 is usually enough. However, when the avoidance of false detections is essential, much smaller values can be used.

3.3 Summary of the algorithm

In this section we summarize the proposed unified algorithm for aerosol detection, and the different required (computational) steps.

1. Collection of a large number of varied but unpolluted or clear observations/spectra (with respect to the target aerosol species).

2. From this collection, the mean unpolluted spectrum $\mu_c$ and covariance matrix $S = S_c$ are calculated.

3. Spectral information on the target aerosol species is gathered. Here two distinct approaches are possible

(a) Collection of a large number of polluted observations/spectra. In this case we can directly calculate $\mu_p$. A Jacobian $k$ can be obtained as $k = \mu_p - \mu_c$.

(b) Construction of a Jacobian $k$ via radiative transfer simulations and knowledge of the spectrally resolved refractive index, with assumptions on shape, size distribution and typical plume altitude. A polluted mean vector can be obtained as $\mu_p = k + \mu_c$.

4. With the above constant vectors and matrices we can calculate the two distance measures $R_N$ and $A_N$ on arbitrary spectra. These can be evaluated on known clear and polluted observations to find suitable thresholds for detection. The thresholds themselves are obtained by finding a suitable compromise between detection sensitivity and avoidance of false identifications.

5. In view of the interpretation of $R_N$ it can be meaningful not to make a strict binary classification decision, but to use $R_N$ as a pseudo-quantitative property. In this case it can also be averaged in time or spatial to increase the signal to noise ratio. The absolute distance criterion can still be applied to reject anomalous observations. In the examples below we will sometimes follow this approach.

While this general algorithm is followed consistently in this paper, there are some intricacies and practical problems to be dealt with, which depend on the aerosol type and what we are trying to achieve. The next section deals with the most important ones, others are mentioned in the detailed discussion of examples in Section 4.

3.4 Practical considerations

A practical point we need to address is the case where the class of aerosol we wish to detect is insufficiently homogeneous and Gaussian. This is pertinent for mineral aerosol. As noted before, the spectral signature of ash or sand depends largely on the specific mineralogical composition. So while
we might want to classify aerosol based on origin (e.g. volcanoes or deserts), their spectral composition might not allow the different possible aerosols signatures to be treated as one aerosol type within the framework of linear discrimination. For volcanic ash, a global class mean spectrum \( \mu_p \) will contain signatures due to rhyolite but also due to basaltic ash. Such signatures can cancel each other out and the mean spectrum might not correspond to something observable or representative.

The problem of inhomogeneity also appeared in Clarisse et al. (2010b) for the selection of an ash reference spectrum. As a solution, not one, but a large number of reference spectra were selected to accommodate for the different spectral signatures. Here we propose a solution along the same lines: instead of trying to detect one type of ash or sand, we divide each category in a number of homogeneous subgroups and devise detection tests for each of them. But rather than composing these subgroups in an ad-hoc fashion, we use a more systematic approach based on spectral clustering. We have opted here for the \( k \)-means algorithm (Ripley, 1996). It is a method to cluster data in \( k \) groups in such a way that each data point is assigned to the class with smallest distance to its class mean. To start, the algorithm assigns classes randomly, and class means and class assignment are then updated iteratively until convergence. For the distance metric, we use the Mahalanobis distance (with a covariance matrix calculated from clear spectra). We have applied this spectral clustering both for the detection of ash and sand, with \( k = 10 \) as detailed in the next section.

A last practical point is the construction of the clear covariance matrices. For some of the rarer type of aerosol (e.g. volcanic ash), this is straightforward. For others, a prior detection method is required to filter out polluted spectra. Here we used mainly BTD type methods. Also, rather than using a single covariance matrix, we can increase sensitivity by making use of several covariances for different time/space intervals. For example volcanic ash detection over the South Pacific ocean should be easier than detection over the Saharan desert. The typical covariance over these different regions is clearly different and this can be exploited by the use of a different covariance matrix. In next section, depending on the specific aerosol type, we will make use of different covariance matrices for observations over ocean or land, for observations contained in different latitude–longitude grid boxes or for different periods of the year.

4 Applications

In this section we present examples of the different types of aerosol that we were able to distinguish with IASI. Note that for the analysis we only use a subset of IASI channels, namely 100 window channels between 750 cm\(^{-1}\) and 1250 cm\(^{-1}\) in brightness temperature space. Channel selection is a constraint based on a priori knowledge of the window channels. While the above methods can clearly utilize the full spectrum, using only a subset of channels is resource-friendly and allows for efficient storage and reprocessing of large time periods of IASI data. Using the full spectrum should lead to slightly better results, although we expect that these 100 channels capture the majority of the aerosol information content. In the examples below, the goal is not to present an ultimate or finished detection product, but rather to demonstrate the huge and largely unexplored possibilities of infrared instruments for aerosol type specification. Where applicable, Jacobians were calculated for a standard atmosphere (NOAA-NASA-USAF, 1976), with infrared spectra simulated using an advanced radiative transfer forward model (Clarisse et al., 2010a) with optical properties of aerosol calculated with Mie theory.

4.1 Sulfuric acid droplets

Oxidation of SO\(_2\) produces sulfuric acid, which leads to the formation of sulfate aerosol in particular sulfuric acid–water solution drops (Steele and Hamill, 1981; Turco et al., 1982; Hamill et al., 1982). Here, our main focus is on upper tropospheric and stratospheric sulfuric acid droplets which are formed within hours to weeks after injection of SO\(_2\) gas and have an atmospheric lifetime of months to years. A direct effect of sulfuric acid aerosols is their interaction with radiation, i.e. scattering of shortwave radiation and absorption of longwave radiation (Stenchikov et al., 1998). In general, large stratospheric injections of sulfuric acid aerosols lead to both a local atmospheric warming and a global cooling of the climate (Robock, 2000). Very large injections can have a dramatic effect on air quality and human health (Schmidt et al., 2011). Sulfuric acid aerosols also have a detrimental effect on aircrafts, potentially causing damage to the windscreen, turbine, engine and airframe (Carn et al., 2009).

In the satellite era there have been several large volcanic stratospheric SO\(_2\) injections, most notably due to the eruptions of El Chichón in 1982 and Mount Pinatubo and Cerro Hudson in 1991. A couple of instruments were in orbit long enough to measure the associated increase and decrease in aerosol loading. Examples include the limb viewing SAGE II instrument at visible and near infrared wavelengths (Thomason et al., 1997; Bauman et al., 2003) and CLAES and ISAMS at infrared wavelengths (Rogers et al., 1998; Lambert et al., 1997). While shortwave instruments can measure an increase in aerosols, infrared instruments are sensitive to particle composition and in ISAMS and AT-MOS infrared measurements performed after the Pinatubo eruption, it was possible to unambiguously identify the specific infrared spectral signature of sulfuric acid–water drops (Grainer et al., 1993; Echle et al., 1998; Eldering et al., 2001, 2004; Steele et al., 2006). Nadir observations of sulfuric acid aerosols include the optical depth measurements made by the NOAA/AVHRR sounder (Stowe et al., 1992; Long and Stowe, 1994) and infrared HIRS observations.
Fig. 1. View of the Northern Hemisphere on five different days (from top to bottom): 12 June (onset of the volcanic eruption of Sarychev), 19 June, 26 June, 12 July and 12 September 2009. On the left, the $\text{SO}_2$ BTD is shown and on the right $R_N(\text{H}_2\text{SO}_4)$. 
(Ackerman and Strahala, 1994). More recently, sulfuric acid observations from the moderate (1–2 Tg SO$_2$) eruptions of Kasatochi and Sarychev were reported using the CALIPSO lidar, OSIRIS limb sounder and ACE occultation measurements (Kravitz et al., 2011; Doeringer et al., 2012; O’Neill et al., 2012; Haywood et al., 2010; Bourassa et al., 2010; Vernier et al., 2011). Example IASI spectra of sulfuric acid aerosol observations were shown in Karagulian et al. (2010), Clarisse et al. (2010a) and Haywood et al. (2010).

In the current detection framework, a clear covariance matrix can be built from time periods which are assumed to be relatively clear of enhanced concentrations of sulfuric acid aerosol. In particular, we used a complete year of IASI observations: April and May 2009 and 2010 for the other months (because of the Eyjafjallajökull eruptions in April and May 2010). Taking such a large time period and so many days minimizes the influence of smaller and short lasting volcanic eruptions. To increase sensitivity, not a global but a local covariance matrix was built, representative for each 10 by 10 degree grid cell. For the calculation of the Jacobian, a choice has to be made for a representative size distribution, temperature and sulfate concentration (Steele and Hamill, 1981; Grainer et al., 1993). A concentration of 75% H$_2$SO$_4$ was chosen, with droplets following a logarithmic size distribution with an effective radius of 1 µm. Reflective indices were chosen, with droplets following a logarithmic size distribution – Eq. (16). Volcanic ash and other rare events can also cause interferences, but such spectra are easily excluded using the absolute distance – Eq. (16). We conservatively filtered out observations with an absolute distance exceeding one.

The results for 3 yr of IASI observations in the Northern Hemisphere are shown in Figs. 1 and 2. Figure 1 shows SO$_2$ BTD (negative values indicative for SO$_2$) and $R_N$(H$_2$SO$_4$) following the eruption of Sarychev on 12 June 2009, after 0, 7, 14, 30 and 90 days. From the very onset, H$_2$SO$_4$ is detected and follows nicely the SO$_2$ distribution. Already after two weeks, the H$_2$SO$_4$ detection exceeds the SO$_2$ detection. After
30 days sulfuric acid aerosol is detected everywhere north of 40° N and it continues to be detectable for many more months after the eruption, with average $R_N(H_2SO_4)$ values over 0.25 until December 2009.

This is also evident from Fig. 2 which shows the Northern Hemisphere minimum daily value SO$_2$ DBT and average value of $R_N(H_2SO_4)$. The different peaks in the SO$_2$ DBT plot can all be attributed to volcanic eruptions, but note that the magnitude of the peaks are not related to the magnitude of the eruptions. Only two clear broad peaks can be seen in the $R_N(H_2SO_4)$ plot, and these can be attributed to the Kasatochi and Sarychev eruptions which had a clear long term impact on sulfuric acid concentrations in the Northern Hemisphere. Although the total SO$_2$ injection from Kasatochi ($>1.5$ Tg) was larger than of Sarychev ($\approx 1$ Tg) (Krotkov et al., 2010; Clarisse et al., 2012), it seems that the lifetime of the sulfuric acid aerosol from Sarychev was the largest. This is consistent with observations from the OSIRIS limb sounder which found a larger maximum mean aerosol optical depth from Sarychev (Haywood et al., 2010;
Bourassa et al., 2012). A similar observation can be made from sulfuric acid aerosol retrievals from ACE (Doeringer, 2011; Doeringer et al., 2012). In the absence of large injections, as was the case in 2010, there is some fluctuation in the mean value of $R_N/(H_2SO_4)$ and looking at the inter-annual variability on Fig. 2 there is also a recurring seasonal minimum between February and June, possibly related to ozone seasonality. Although we have not tried this here, it is likely that this seasonality would disappear when making use of dedicated seasonal covariance matrices.

The previous examples focus on the larger eruptions of 2008 and 2009, but $H_2SO_4$ can also be detected in smaller eruptions or sustained degassing. Examples of the latter are presented in the left panel of Fig. 3 which shows monthly averaged maps of $R_N/(H_2SO_4)$ for three volcanoes. The two first panels on the left show a large plume over Central Africa originating from the Nyamuragira volcano. It is Africa’s most active volcano and one of the largest recent natural sources of $SO_2$ emissions (Carn and Bluth, 2003; Bluth and Carn, 2008). On 6 November 2011 the Nyamuragira volcano violently erupted and continued to do so for 5 months. Injection altitudes were between 4–8 km (Theys et al., 2012). Throughout November, $SO_2$ can be detected in the vicinity of the volcano and drifting westwards. The November monthly mean shows $SO_2$ mainly over the D.R. of Congo (see also the monthly mean from OMI shown in inset).

$H_2SO_4$ detection extends thousands of kilometers, well over the South Atlantic with a maximum detection at some distance away from the emitting source (1000–2000 km). The $SO_2$ and $H_2SO_4$ plots are consistent, as sulfate aerosol with a much longer lifetime is produced from atmospheric $SO_2$ which is hence only detected in the vicinity of the volcano. The transport patterns are consistent with a lifetime of $SO_2$ of the order of one day (Bluth and Carn, 2008), wind velocities of the order of 500 km per day ($\sim 6$ ms$^{-1}$) and the fact that sulfuric acid aerosols that have undergone some growth or coagulation are easier to detect. The third figure in the left panel of Fig. 3 shows the August 2008 $H_2SO_4$ plume from Kilauea (Hawaii). The plume has a maximum close to the volcano and another 1500–2500 km downwind. The detection is remarkable as the injection height of the relatively mild eruptions at Kilauea is low. In fact, almost no $SO_2$ is measured using the BTD-based IASI $SO_2$ retrieval (Clarisse et al., 2012) which uses spectral channels with a penetration depth between 3 and 6 km. Similarly, an $H_2SO_4$ signal was picked up near the Galápagos Islands in April 2009 for which no coincident $SO_2$ was measured with IASI. The plume depicted in the lower left corner of Fig. 3, can be attributed to activity of the Fernandina volcano.

It is instructive to compare these results with MODIS AOD measurements. Sulfate aerosols are efficient scatterers of solar radiation, and we therefore expect these plumes to show up in the monthly AOD averages shown in the right panel of Fig. 3. While the two quantities ($R_N/(H_2SO_4)$ and AOD) are clearly different, in the non-saturated regime, they are both proportional to the aerosol loading. The best match is obtained for the Kilauea plume. Small differences in the location of the plume maximum, direction and extend can be attributed to instrumental dependency on aerosol radius (larger particles are easier to detect in the thermal infrared than the visible) and altitude (higher altitude particles are easier to detect in the infrared because of larger thermal contrast). But overall we note that there is an excellent agreement between the two sounders. Note that MODIS AOD monthly averages for 2008 around Kilauea were analyzed in detail in Beirle et al. (2012) in the context of sulfate formation and depletion. The other sulfate aerosol plumes detected by IASI also appear to be present in MODIS AOD monthly averages, but they are hard to discern because of the presence of other aerosol particles: smoke and dust for Central Africa and background (mostly anthropogenic) aerosols in the Pacific ocean (see again Fig. 3). These comparative examples demonstrate the sensitivity and accuracy of the current algorithm. At the same time they illustrate the great advantage of being able to single out aerosol sulfate from all other types of aerosol as small loadings can easily be missed in AOD measurements due to the presence of background aerosols (as is the case here for e.g. the Fernandina volcano).

4.2 Windblown sand

Sand, dust and other soil derived aerosol is emitted directly rather than formed in the atmosphere. As one of the most abundant types of aerosol it has a multitude of climatological and environmental impacts (Gassó et al., 2010). Sand interacts with a wide range of radiation and is readily sensed by most aerosol sensors, see e.g. Carboni et al. (2012) for a comparison of AOD of Saharan dust from eight different instruments. Retrievals over land, and especially deserts is challenging in the thermal infrared because of variable surface emissivity and in the solar spectral range because of high reflectivity. Thermal infrared instruments are very sensitive to sand aerosols, and they have the advantage of being able to distinguish them from other aerosol types. They can also be used to retrieve effective radius and height (Pierangelo et al., 2004, 2005; Peyridieu et al., 2010; DeSouza-Machado et al., 2010).

As explained above, to deal with the large variety of different sand signatures, we have constructed 10 polluted class mean spectra $\mu_p$ using the $k$-means algorithm. Twelve sand/dust storm scenes were selected (including transport) from the Sahara, Gobi, Arabian and Simpson deserts both during night and day. Effort was made to confirm the presence of sand aerosols using either visible (MODIS) or lidar (CALIPSO) imagery. From these different scenes, sand polluted spectra were selected based on DBT tests and fed to the $k$-means algorithm. The 10 resulting class mean spectra are shown in Fig. 4. The observed large variability in the spectral signatures supports the choice to work with different classes, although the number of classes could probably...
be reduced. For most class means, the characteristic V-shape between 800 and 1200 cm\(^{-1}\) is obvious. This shape is due to the presence of Si-O fundamentals in minerals such as quartz, illite and montmorillonite (Hunt, 1982; Clark, 1999; Sokolik and Toon, 1999). Other tentative assignments of calcite and kaolinite have also been indicated.

The problem of constructing a clear covariance matrix is not trivial, and here we have opted for a recursive approach as a better sand detection allows us to construct a better covariance matrix. First, to better accommodate for surface emissivity effects in the detection of windblown sand we treat land and ocean separately. In the first iteration, a sand detection test was designed based on a single relative distance \(R_N\). Here, the clear covariance matrices (one for land and one for ocean) were constructed from all spectra of one day (so also contaminated spectra) and the polluted class mean from the mean of the twelve dust scenes described above. Detection in this way is noisy, but allows to filter out the most obvious dust scenes. In the second iteration, this basic test was used to design better clear covariance matrices from 24 days of spectra. These matrices were used to replace the covariance matrices in the first test to obtain a better detection test. This second detection test was then used to construct the final clear covariance matrices.

Using these, and the 10 different class means, 10 different sand detection tests were designed, each using the value of \(R_N\) as primary detection criterion and \(A_N\) to remove false alerts. The corresponding detection thresholds were manually tuned. The result for a day in March and one in April are shown in Fig. 5. The colors refer to the individual tests, but in cases of detection by multiple tests, only the color corresponding to the first successful detection is indicated. The March plot (top panel) shows large synoptical scale transport patterns westwards from the Saharan desert and eastwards from the Taklamakan and Gobi desert. Looking at North Africa and the land/ocean transitions it is clear that detection over land is almost as good as over ocean for these large plumes. The April plot (bottom panel) depicts a very large plume of Saharan dust being transported towards the UK and the rest of Europe. The plume originated from the North African coast on 5 April and was transported over the Atlantic and Spain reaching the UK, Norway and Sweden on 8–10 April and reaching back to Saudi-Arabia on the 13th. Note that most of the sand is detected with tests 2, 4, 7 and 8. This has to do with parameter tuning but also with the fact that some signatures only manifest themselves in special thermal conditions. An example is test 1 which is observed very close to the west coast of Africa in the bottom panel and seems to occur when there is almost no or negative thermal contrast between the dust plume and the underlying ocean.

Figure 6 shows MODIS AOD and OMI absorption AOD (AAOD) for the same two days as Fig. 5. Comparing the figures side by side, a good agreement is found. The AOD maps exhibit the same transport patterns found with IASI: westwards/northwards from the Saharan desert and eastwards from the Taklamakan and Gobi desert. Enhanced AOD values over parts of India and South China are probably due to aerosol from anthropogenic origin. Note that the MODIS product shown here (dark target) does not provide retrievals over bright surfaces such as the Saharan desert. The plume which IASI detects on 23 March over the Northern part of the Sahara is hence not present in the MODIS AOD.
Fig. 5. Sand aerosols detected over the Northern Hemisphere on 23 March 2010 (top) and 8 April 2011 (bottom) based on strict thresholds for $R_N$ and $A_N$ for each of the ten tests. The colorcode corresponds to the test which passed (see Fig. 4). In case multiple tests flagged an observation, only the lowest test number is indicated.

this reason in the infrared spectral range is just like we have also shown the OMI AAOD where it does show up. Overall note that the AOD maps are rather patchy in comparison with the IASI detections using the current algorithm. This conclusion becomes even stronger when realizing that the detection problem we are tackling with IASI attempts to assign a binary dust/no dust detection on a per pixel basis. This way of presenting (Fig. 5) is not flattering as it exposes even single pixel false detections, while making the problem of detecting plume edges much harder. In contrast, anomalies in smoothened AOD images (a continuous quantity) are averaged out in the contour plot and/or obscured by how the brain perceives continuous data.

Resulting monthly averages of detected sand are shown in Fig. 7 for the months January and June 2011. The figure was obtained by gridding the daily detection maps and calculating the percentage of detected sand spectra in each grid cell. The plots were saturated at 20%. A qualitative comparison with AIRS optical depths (Peyridieu et al., 2010) over the Atlantic ocean and Arabian sea shows an excellent correspondence. Detection over land is more difficult than over ocean; especially for lower loadings (e.g. in January). The false detection rate is in general low (smaller than 1% globally, larger in the tropics) except in specific locations marked by peculiar surface emissivity features (e.g. over Lake Eyre, Australia). Almost no interference with other aerosol types is found, except with ash, which will be dealt with in next section.

4.3 Volcanic ash

Volcanic ash generally refers to small (less than 2 mm) pulverised particles of pieces of volcanic glass and rock. Their atmospheric lifetime range from minutes to months depending on injection height and particle size (Rose and Durant, 2009). The main motivation for routine monitoring of fine volcanic ash from space is mitigation of the safety threat it poses to aviation. Volcanic ash causes erosion and abrasion to the aircraft (e.g. to the windscreen) and readily melts inside the engine, potentially leading to engine failure (see Casadevall (1994); Miller and Casadevall (2000); OFCM (2004) and references therein). Currently a global network of nine volcanic ash advisory centers (VAACs) reports ash observations and forecasts. These rely on transport and dispersion models fed by ground-based data (lidar/radar), satellite observations and reports from pilots. Until April 2010 the emphasis of the volcanic ash problem was on the safety aspect. However, during and after the April–May eruption of Eyjafjallajökull eruption in 2010, a lot of attention was given to the financial aspect. Indeed the cost of rerouting or canceling flights can be very large, e.g. for the period of 15 to 21 April 2010, airlines lost at least US$ 1.7 billion in revenue (Mazzocchi et al., 2010). The eruptions of Eyjafjallajökull in 2010 and Grímsvötn reinforced the need for, and importance of accurate high spatial and temporal ash measurements and modeling forecasts (Stohl et al., 2011; Zehner, 2012).

The infrared spectral range is just like for windblown dust, the preferred way of differentially observing volcanic ash. For a review of different techniques and instruments see...
Prata (2009). Traditionally broadband instruments have been used for the detection of volcanic ash and their general small footprint and high revisit time remain a decisive advantage (Pavolonis and Sieglaflf, 2010; Prata and Prata, 2012). Only a few papers deal with ash measurements from high spectral resolution instruments, and these report better sensitivity to thin plumes and better differentiation of ash with other atmospheric parameters (clouds, water vapor, sand and surface emissivity features) (Gangale et al., 2010; Clarisse et al., 2010b; Klüser et al., 2012).

In the current framework, as volcanic ash is only sporadically observed, it is easy to generate the covariance matrices. For simplicity, we have used the same ones as for the detection of sulfuric acid aerosols. The same distinction was also made for spectra with and without a detectable SO$_2$ signature. So, to avoid cross contamination, a reduced spectral signature is used in case of positive SO$_2$ detection. There is no unique spectral signature of ash because of variable composition and size distribution. For the calculation of the polluted class mean(s) we have adopted therefore the same approach as for sand, selecting spectra of ten different ash scenes from eruptions of Eyjafjallajökull, Grímsvötn, Puyehue-Cordón Caulle, Soufrière Hills, Shiveluch and Sarychev. These spectra were then clustered using the $k$-means algorithm for which class means are shown in Fig. 4. The algorithm separates the different eruptive plumes remarkably well and the different classes clearly exhibit a difference in composition but also in magnitude. Most striking is the difference around 860 cm$^{-1}$ (830 to 900 cm$^{-1}$) going from concave for felsic ashes (e.g. from Puyehue-Cordón Caulle) to convex for mafic ashes (Grímsvötn). The concavity of felsic ashes was noted in Gangale et al. (2010) in spectra of the rhyolitic ash of Chaitén and is only partially reproduced in the widely used refractive measurements of Pollack et al. (1973) which has the concave peak closer to 800 cm$^{-1}$. See in this context also Newman et al. (2012) for

Fig. 6. MODIS (Terra) aerosol optical depth measurements at 550 nm and OMI absorption aerosol optical depth measurements at 500 nm on 23 March 2010 (top) and 8 April 2011 (bottom). Data downloaded from the Giovanni online data system (http://disc.sci.gsfc.nasa.gov/giovanni/overview).
Fig. 7. Ash and sand detection in January (top panel) and June (bottom panel) 2011 in percentages per grid cell. In case both sand and ash were detected in the same grid cell all observations were assigned to one category based on the relative number count: when the number of sand detections was lower than twice the number of ash detections it was marked as ash, otherwise as sand.

example plots of high resolution spectral fits of rhyolitic and basaltic volcanic ash.

In the same way as was done for sand, detection threshold were manually tuned. While the use of different covariance matrices in different latitude/longitude gridboxes already guarantees few false detections due to sand and surface emissivity features, an improvement can be made to reduce these even further while increasing the sensitivity. First the detection test for sand was expanded with a so-called weak ash detection test. The corresponding detection thresholds were chosen to be quite relaxed, and the resulting detected aerosols were labeled as “containing minerals”. This set of observations was then subjected to a strong ash detection test with stricter detection thresholds. Observations which passed the strong ash test were labeled ash. In addition, all observations within 150 km of these observations and which passed the weak ash detection test were also labeled ash. All other mineral containing aerosols were labeled as sand. While the above method is ad-hoc and can obviously be improved, it was found to have a larger sensitivity and less false detections than the ash-correlation method (Clarisse et al., 2010b). We have applied the ash detection algorithm on 4 years of IASI data, and compared with available literature data. This analysis confirmed the robustness of the current algorithm, as no major ash-rich eruption was missed, while excellent agreement was obtained with published data for the eruptions of Eyjafjallajökull and Chaitén.

Monthly averages of ash detection are indicated together with sand detection on Fig. 7. Ash in January can be attributed to eruptions of Kirishima (Japan), Tengger Caldera...
Fig. 8. Average MODIS (Terra) aerosol optical depth measurements at 550 nm for June 2011. This product is not sensitive to aerosol composition and therefore shows a variety of different aerosols and their transport patterns which can be attributed to natural aerosol (sand mainly between 0 and 30° N and ash below −30° N), and anthropogenic (mainly urban/industrial pollution in India and Asia and biomass burning in Central Africa). This figure can be compared with Figure 7 for the minerals and with Figure 10 for ammonium sulfate aerosol. Data downloaded from the Giovanni online data system (http://disc.sci.gsfc.nasa.gov/giovanni/overview).

(Indonesia) and Kizimen/Karymsky (Kamchatka, Russia) and in June to the eruptions of Puyehue-Cordón Caulle (Chile). The eruption of Kirishima illustrates nicely how ash and sand can be differentiated even over areas with significant background sand transport. Ash from the eruptions of Puyehue-Cordón Caulle was very long-lived. On the daily plots (not shown) ash from the June 4 and subsequent eruptions of Puyehue-Cordón Caulle eruption (Klüber et al., 2012; Theys et al., 2012) could be seen to circulate the Southern Hemisphere four times between −40° and −60° latitude until about 20 July. As can be seen from the daily plots, the original ash layer was enriched after 10 days by later eruptions. Most aerosol was correctly identified as ash until the end of June, with increasing amounts of thin filaments labeled as sand/minerals after that.

To conclude the discussion of mineral detection, we compare our monthly June average Fig. 7 (bottom panel) to the MODIS AOD average for the same time period as shown in Fig. 8. The sand/dust transport patterns and their extent correspond generally really well, especially transport of Saharan sand westwards and of the Arabian desert over the Middle East region. Note that our monthly average was constructed using binary data, which was designed to yield almost no false detections. This also means that data near the edges of plumes is lost. The MODIS AOD monthly average on the other hand, benefits a great deal from averaging. Observations that would otherwise never been attributed to dust (because of having an only marginally augmented AOD) do contribute to the monthly average, and help in establishing the harder to detect plume edges. The fact that both maps agree is not-trivial and is a clear demonstration of how well the detection of sand works with the unified method. Also note that the (limited) transport from the Taklamakan and Gobi desert shows up really well in the IASI monthly average. It is harder to discern in the MODIS AOD average, partly because of problems with bright surfaces, partly because of interference with anthropogenic aerosol (which is seen over the whole of Asia and the North Pacific ocean). Ash from Puyehue-Cordón Caulle is detected with MODIS across the Southern Hemisphere, but limited to above −45° because of the Austral Winter. This eruption is a good example of the necessity of being able to measure aerosols globally and during all periods of the year.

4.4 Ammonium sulfate

Sulfuric acid droplets can exist in pure form in remote areas or high up in the troposphere and stratosphere, but in polluted areas ammonium sulfate readily forms as ammonia condenses on sulfuric acid droplets. Ammonia sulfate aerosols are abundant and may be found in either solid or aqueous form, and can be pure, mixed with other secondary aerosol or coated on primary aerosol (Wang et al., 2008; Yu et al., 2012; Zhang et al., 2007). They represent a major part
of anthropogenic aerosol, affecting air quality and impacting significantly the radiative budget (Adams et al., 1999; Martin et al., 2004). To our knowledge remote sensing measurements of ammonium sulfates have not been reported until now.

It is however well known that ammonium sulfate has a strong absorption band near 1115 cm$^{-1}$ (see Fig. 9), which is favorably located in the atmospheric window. Refractive indices of ammonium sulfate have been measured since the 1970’s both in crystalline (Volz, 1973; Toon et al., 1976; Earle et al., 2006; Segal-Rosenheimer et al., 2009) and aqueous form (Remsberg, 1973; Downing et al., 1977; Boer et al., 2007). For the calculation of the Jacobian, we use the recent publicly available dataset from Earle et al. (2006) assuming a logarithmic size distribution with an effective radius of 0.2 µm. Because we expect ammonium sulfate aerosols close to the surface, for this type of aerosol we only used cloud filtered spectra. Covariance matrices were constructed separately for land and ocean on global data of 24 days covering all months. As we have a priori no easy way of filtering out spectra with an ammonium sulfate signature, the idea of using a global covariance matrix is to reduce the weight of such spectra.

Ammonium sulfate was unambiguously identified over East China for May–July 2010, see Fig. 10; and confirmed by visual inspection of observed spectra over this region. An example of such a spectrum is given in Fig. 9. It exhibits the ammonium sulfate absorption band around 1115 cm$^{-1}$; but also a clear slope from 750 to 1000 cm$^{-1}$. This slope appears together with the ammonium sulfate absorption band in refractive indices of water soluble aerosol, reported in Volz (1972a,b) and Shettle and Fenn (1979). The time and location of the ammonium sulfate plume coincides very well with increased AOD in MODIS, LIDAR and AERONET measurements (Kim et al., 2007b). This is also apparent in our MODIS AOD June average (Fig. 8). The peak in June/July is explained by the availability of ammonia and favorable stable meteorological conditions in terms of build up and growth of aerosols (Kim et al., 2006, 2007b). Globally, ammonium sulfate was not confidently detected anywhere else over background noise and interferences of dust and smoke.

4.5 Smoke

Smoke from biomass burning consists of 50–70 % of organic and black carbon, representing less than 5 % of burned carbon. In addition secondary particles are also formed within smoke plumes. The majority of smoke particles is in the accumulation mode (< 1 µm diameter) and efficiently absorb and scatter solar radiation. A non-negligible fraction (∼10 % in mass) of coarse mode particles consisting of ash, soil, carbons and partially combusted biomass with sizes is in the coarse mode (Reid et al., 2005a,b). This suggest the possibility of interaction with infrared radiation. One of the few laboratory measurements of smoke particles in the infrared (Sutherland and Khanna, 1991) reports significant extinction between 1000 and 1800 cm$^{-1}$, see Fig. 11. Also IASI observations of smoke particles have been reported (Clarisse et al., 2010a). Since the launch of IASI, the most clear observations of smoke particles were probably made in plumes from the Australian bush fires of 2009, where a global drop in the baseline of ∼10 K (and ∼20 K between 1050 and 1250 cm$^{-1}$) was observed (Clarisse et al., 2011). Although in this case, arguably, part of the observed signature was also due to the sum of all enhanced trace gas absorptions (e.g. from large molecules as PAN and acetic acid). Other observations of smoke with IASI were reported in Köhler et al. (2011), who point out the importance of thermal contrast and the fact that downwelling infrared radiance is much more affected by the presence of smoke than upwelling radiance.

For the covariance matrices we have used the same ones as for the detection of ammonium sulfate, with again land/ocean separation and cloud filtering. A Jacobian was generated from a simulation using the refractive indices of smoke from a mixed weed sample reported in Sutherland and Khanna (1991). While the detection is noisy on an
individual spectrum basis, it picks up (aged) smoke plumes over the South Atlantic and Indian ocean especially in September and October as shown in Fig. 12 for October 2010. It is well known that fires in Southern Africa and South America peak in these months. Also the transport pattern from West to East matches well with earlier observations (Edwards et al., 2006; Giglio et al., 2006; Roy et al., 2008). Comparison with the MODIS (Fig. 13) monthly AOD average for October shows a good similarity in the transport pattern, but also notable differences over land where infrared detection seems to be more difficult.

5 Conclusions

In this paper we have presented a unified aerosol detection scheme based on a primary relative and secondary absolute distance criterion. The primary criterion uses classical linear discrimination analysis, and generalizes previously reported detection scheme using optimal estimation and PCA’s. The secondary criterion was used here to filter out false detections which becomes necessary when the underlying probability distributions depart from normality. We have demonstrated the methods on IASI observations using five different types of aerosol namely volcanic ash, windblown sand, sulfuric acid droplets, ammonium sulfate and smoke particles. For each of these we have introduced small adaptations in the exact implementation, mainly in the way the Jacobian and covariance matrices were generated. It is clear that there are lot of different variations, which will work equally well. The presented results should therefore be seen as proof of concept rather than finished products.

It would indeed be a natural step to use the presented first generation products as a basis for a second generation product. As we demonstrated for sand detection, first generation products can be used to build better covariance matrices. Other improvements could include making use of larger spectral range (in particular the window between 2000–2200 cm\(^{-1}\)) and to make more use of space and time differentiated covariance matrices. Improvements outside the current detection framework include making use of some of the more advanced techniques of class discrimination. Indeed, we have only implemented here what is essentially the most basic form of class discrimination, while the state of the art is vast and highly advanced (McLachlan, 2004). Also note that nothing in the method is IASI specific, so that our techniques

www.atmos-chem-phys.net/13/2195/2013/

can easily be applied to measurements of other high spectral resolution infrared sounders (TES, AIRS, GOSAT, ...).

Notably absent in the current scheme are clouds and sea spray. As noted before, ice crystals found in e.g. cirrus clouds found have a pronounced spectral signature. The presented method can therefore be applied for their detection without any complications. We have not attempted the detection of regular clouds, as these do not have a pronounced spectral signature in the infrared, and we therefore do not expect the current scheme to be easily adaptable for their detection. Sea spray on the other hand does have a marked, but slowly varying spectral signature (Irshad et al., 2009). We have attempted detection of sea spray but no convincing results emerged, likely due to a combination of (1) the use of too small spectral range and (2) constructed covariance matrices as sea spray is omnipresent, (3) low thermal contrast over oceans and (4) no obvious test cases of enhanced loadings.

For the aerosol types for which the method does work, the results exceed our expectations. Aerosol composition is an essential climate variable, but state of the art aerosol sounders offer only limited information. As we have demonstrated here, the thermal infrared has a (largely unexplored) potential of measuring aerosol composition directly, which can contribute to our understanding of climatic issues. The presented algorithm allows to detect volcanic sulfate aerosol with unprecedented sensitivity, both from large and from sustained volcanic eruptions down to the mid/lower troposphere. This advance could lead to better constraints on the large uncertainties of sulfate aerosol climate forcing (Schmidt et al., 2012). It was already well known that infrared sounders are very good at measuring mineral aerosol. But the current detection scheme seems to be particular robust, especially in differentiating ash from sand. These mineral detections would not require too much work to be implemented in an operational setting, with applications to e.g. aviation hazard

Fig. 12. October 2010 mean value of $R_N$ for smoke over the South Atlantic and Indian Ocean.

Fig. 13. MODIS (Terra) aerosol optical depth measurements at 550 nm for October 2010 and for the same area as Fig. 12. Data downloaded from the Giovanni online data system (http://disc.sci.gsfc.nasa.gov/giovanni/overview).
mitigation. The demonstration of smoke and ammonia sulfate detection was unforeseen, especially since these aerosols consist mainly of small particles (< 1 μm) and are located in the lower troposphere. Also note that it is possible for multiple aerosol types to be detected in the same observation, provided that their respective spectral signatures cover at least partially different spectral regions. A good example of this is the Central Africa region, where we were able to detect sulfate aerosol (Fig. 3), windblown sand (Fig. 7) and smoke (Fig. 12).

Acknowledgements. IASI has been developed and built under the responsibility of the Centre National d’Etudes Spatiales (CNES, France). It is flown onboard the Metop satellites as part of the EUMETSAT Polar System. The IASI L1 data are received through the EUMETCast near real time data distribution service. We also acknowledge the MODIS and OMI mission scientists and associated NASA personnel for the production of the data used in this research effort. MODIS and OMI optical depth data was downloaded from the Giovanni online data system, developed and maintained by the NASA GES DISC. L. Clarisse and P.-F. Coheur are respectively Postdoctoral Researcher (Chargé de Recherches) and Research Associate (Chercheur Qualifié) with F.R.S.-FNRS. C. Clerbaux is grateful to CNES for scientific collaboration and financial support. The research in Belgium was funded by the F.R.S.-FNRS (M.I.S. nF.4511.08), the Belgian State Federal Office for Scientific, Technical and Cultural Affairs and the European Space Agency (ESA-Prodex arrangements and the Support to Aviation Control Service (SACS+) project). We also acknowledge financial support from the “Actions de Recherche Concertées” (Communauté Française de Belgique).

Edited by: E. Highwood

References


L. Clarisse et al.: Aerosol type specification in the infrared

2217


Klúser, L., Martynenko, D., and Holzer-Popp, T.: Thermal infrared remote sensing of mineral dust over land and ocean: a spectral
L. Clarisse et al.: Aerosol type specification in the infrared

2219


