

## A REVIEW OF DIMENSIONALITY REDUCTION TECHNIQUES FOR PROCESSING HYPER-SPECTRAL OPTICAL SIGNAL

Ana del Águila, Dmitry S. Efremenko, and Thomas Trautmann

*Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Methodik  
der Fernerkundung (IMF), 82234 Oberpfaffenhofen, Germany  
E-mail: dmitry.efremenko@dlr.de*

---

### ABSTRACT

Hyper-spectral sensors take measurements in the narrow contiguous bands across the electromagnetic spectrum. Usually, the goal is to detect a certain object or a component of the medium with unique spectral signatures. In particular, the hyper-spectral measurements are used in atmospheric remote sensing to detect trace gases. To improve the efficiency of hyper-spectral processing algorithms, data reduction methods are applied. This paper outlines the dimensionality reduction techniques in the context of hyper-spectral remote sensing of the atmosphere. The dimensionality reduction excludes redundant information from the data and currently is the integral part of high-performance radiation transfer models. In this survey, it is shown how the principal component analysis can be applied for spectral radiance modelling and retrieval of atmospheric constituents, thereby speeding up the data processing by orders of magnitude. The discussed techniques are generic and can be readily applied for solving atmospheric as well as material science problems.

**Keywords:** passive remote sensing, hyper-spectral data, principal component analysis, full-physics machine learning, trace gas retrieval

### 1. INTRODUCTION

Hyper-spectral sensors record the transmitted or reflected radiance in the narrow contiguous bands across the electromagnetic spectrum. The goal is to detect a certain object or a component of the

medium, which has a unique spectral signature, i.e. a fingerprint. Hyper-spectral imaging has emerged as one of the most powerful technologies in various fields including astronomy, mineralogy, agriculture, medicine and chemistry. For instance, hyper-spectral data (sometimes referred to as hypercube data or as an image cube) are used in astronomy and Earth remote sensing to create a spatially-resolved spectral image allowing more accurate recognition and classification of the objects in the instrument field of view. At the same time, significant data storage and computational power are required to process the hyper-spectral information and to retrieve a certain parameter of the scattering medium.

In this survey we are focused on hyper-spectral remote sensing of the atmosphere. The passive atmospheric composition sensors (ACS) detect and record the radiance reflected by the Earth atmosphere in the ultraviolet (UV), visible (VIS), and thermal infrared (IR) regions. The information about the atmosphere is then retrieved from the spectral data by using the so called atmospheric processors, i.e. codes which are specifically designed to invert ACS measurements [1]. Extracting the information about geophysical parameters (level-2 data) from spectral radiance distributions (level-1 data) turns out to be a major computational challenge and requires high performance computing (HPC) [2].

The recent developments in optics, sensor design and measurement techniques significantly improve the characteristics of hyper-spectral ACS, such as the spatial resolution and the signal-to-noise

**Table. Characteristics of Atmospheric Composition Sensors**

Instrument	GOME	GOME-2	TROPOMI
Platform	ERS-2	MetOp (A, B, C)	Sentinel 5 Precursor
Spatial resolution (km <sup>2</sup> )	320x40	80x40	7x3.5
Amount of level-1 data (TB per year)	0.8	4.2	240
Operational	1995–2011	2006–present	2017–present

ratio. Table shows a comparison between previous generation instruments like the Global Ozone Monitoring Experiment (GOME and GOME-2) and the newest, the TROPOspheric Monitoring Instrument (TROPOMI) [3] on board of the Copernicus Sentinel 5 Precursor (S5P) satellite. The spatial resolution of TROPOMI is two orders of magnitude higher, providing 21 million level-1B spectra per day, i.e. almost 8 milliards spectral points, while the signal-to-noise ratio in the UV/VIS channels reaches the values of about 1500. Fig. 1 shows the example of a retrieved map of tropospheric nitrogen dioxide (NO<sub>2</sub>) from S5P measurements (the data is freely available at <https://s5phub.copernicus.eu/dhus>). Observe that air pollution emitted by big cities and shipping lanes is clearly visible. With the high resolution data it is possible to detect air pollution over individual cities as well as to locate where pollutants are being emitted, and so, identifying pollution hotspots. Such high resolution satellite remote sensing observations are extremely useful for diagnosing the impact of atmospheric constituents on a global scale, in particular, allowing detection of small-scale sources, and increasing the fraction of

cloud-free observations. However, the high spatial resolution of the state-of-the-art ACS results in very challenging data volumes to be processed – 240 TB (terabyte) per year of level-1 data.

In fact, the amount of satellite data increases faster than the computational power [4]. The remote sensing data is recognized as Big Data [5] since it satisfies Doug Laney's 3V criterion: significant growth in the volume, velocity and variety. New efficient techniques have to be developed for next generation atmospheric processors to cope with these high efficiency requirements.

The radiative transfer modelling (RTM) is the key component and the major performance bottle-neck in the atmospheric processors. Furthermore, the hyper-spectral RTMs involve a hierarchy of nested computational loops [6] as shown in the pseudo-code in Fig. 2. Recent surveys such as those provided by V. Natraj [7] and D. Efremenko et al. [8, 9] showed that a significant performance enhancement can be achieved by optimizing the framework in which the radiative transfer solver is called rather than accelerating the RTM solver itself. In fact, the efficiency of monochromatic radiative transfer solvers hardly

can be further improved [10]. Several attempts have been made to optimize loops over ground pixels and geometry (see, e.g., [11, 12, 13, 14] and references therein). Currently the loop over wavelengths (which expresses the hyper-spectral processing) remains the most computationally demanding part.

The essential part of the Big Data analysis is the dimensionality reduc-

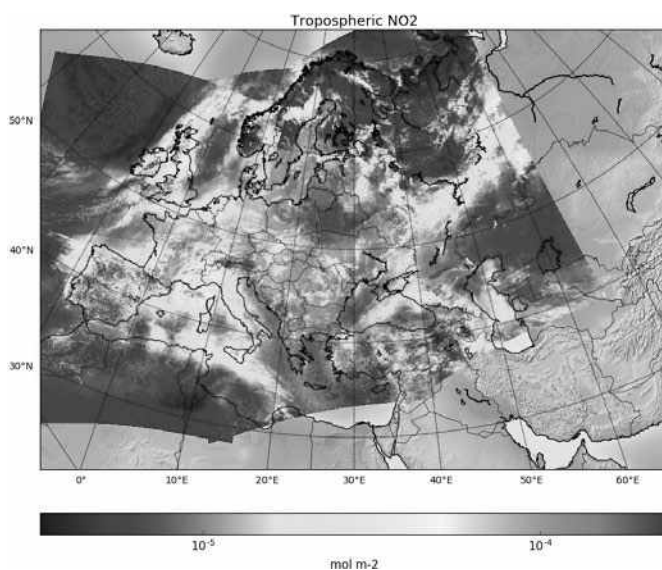


Fig. 1. Example of Sentinel-5P tropospheric nitrogen dioxide (NO<sub>2</sub>) measurements on 1.04.2019 (data is freely available at <https://s5phub.copernicus.eu/dhus>)

```

for each ground pixel:
  for each wavelength: #<--loop for hyper-spectral processing
    for each geometry:
      call Radiative_transfer_solver;

```

Fig. 2. Hierarchy of computational loops, in which the radiative transfer solver is called

tion procedures which are related in the context of RTMs to the loop over wavelengths. The principal component analysis (PCA) is one of the famous representatives of them. PCA was proposed in 1901 by K. Pearson [15] and today has become an integral part of hyper-spectral RTMs.

In this regard, many efforts have been made to develop hyper-spectral RTMs which explicitly take into account the interdependency and statistical relations between level-1 and level-2 data [16]. The motivation of this survey is to present basic concepts of dimensionality reduction for design of atmospheric processors in a systematic way and to put in one context recent developments in this field.

## 2. DIMENSIONALITY REDUCTION

### 2.1. Heritage from the $k$ -Distribution Technique

The techniques of dimensionality reduction in atmospheric science and astrophysics were preceded by methods which exploit a strong interdependency in the hyper-spectral data, although do not use PCA explicitly. Ambartsumian [17] noted that the transmission within a spectral interval does not depend on the line-by-line (LBL) variation of the absorption coefficient  $k$  with respect to wavelength  $\lambda$ , but rather on the distribution of absorption coefficient within the spectral interval. Such concept is called the  $k$ -distribution approach. It is based on the cumulative frequency distribution of  $k$ , namely,  $G(k)$ ; as a matter of fact, the inverse distribution  $k(G)$  is the  $k$ -distribution function. Since  $G(k)$  is a smooth function and  $k(\lambda)$  has a strong variation with respect to  $\lambda$ , it is beneficial from the numerical point of view to replace the integration in  $\lambda$ -space by that in  $G$ -space (less number of quadrature points for numerical integration is required). This method was extended by Goody et al. [18] to the cases of inhomogeneous atmosphere assuming that there is a correlation between  $k$ -distributions at different pressure levels, and that is the correlated- $k$  distribution method. Other techniques (e.g. exponential sum fitting [19], spectral map-

ping [20],  $k$ -binning approach [21], opacity sampling method [22], multi-dimensional  $k$ -distribution method [23], and fast  $k$ -distribution models [24, 25]) use similar ideas although based on a more elaborative mathematical basis. In [26], a modification of the  $k$ -distribution technique was considered, in which the integration is performed in the original  $\lambda$ -space. In [27], an empirical procedure based on uniform spectral grids was proposed for choosing the most representative spectral points in LBL computations. Together, the studies related to the  $k$ -distribution function outline that the hyper-spectral radiances can be modelled by much fewer number of monochromatic computations than required by the LBL-framework, thereby providing a basis to dimensionality reduction of the problem.

### 2.2. Basic Concept of Dimensionality Reduction

In simple words, dimensionality reduction means representing the initial data set with less number of parameters than it is initially represented. It can be considered as one of the lossy data compression paradigms [28]. Dimensionality reduction is crucial for stable and high-performance processing of spectral measurements. It excludes redundant information from the initial dataset, reduces the number of independent parameters and improves the efficiency of machine learning.

There is a distinction between linear and non-linear techniques for dimensionality reduction. A more detailed review can be found in [29, 30, 31, 32] and references therein. Linear and non-linear techniques have been inter-compared in [33]. Results of these numerical experiments reveal that non-linear techniques perform well on selected artificial tasks. However, they hardly outperform PCA on real-world tasks. Similar conclusions were reported in [34], where several methods for dimensionality reduction were inter-compared in the context of accelerating radiative transfer performance. Bearing in mind that no obviously superior method has emerged in the benchmarking studies (increasingly time-consuming and sophisticated dimensionality

reduction techniques lead to more accurate results, and vice versa) our analysis will be limited with the classical PCA.

### 2.3. Principal Component Analysis

Although the dimensionality reduction techniques are well-known and covered by many statistical libraries (e.g., scikit-learn [35] for Python), we make a short mathematical exposition to put the above considerations in a proper context. For clarity, we specify sizes of matrices using the notation  $\in \mathbb{R}^{\text{rows} \times \text{columns}}$ . Let  $\mathbf{y} = (y(\lambda_1), y(\lambda_2), \dots, y(\lambda_W))$ ,  $\mathbf{y} \in R^{1 \times W}$ , be a row-vector of atmospheric radiances at  $W$  wavelengths  $\{\lambda_w\}_{w=1, \dots, W}$ . A set of  $S$  spectra are assembled into a matrix  $\mathbf{Y} \in R^{S \times W}$  whose  $i$ -th row is  $\mathbf{y}_i$ . Then,  $\mathbf{y}_i$  can be represented in a new basis system as follows:

$$\mathbf{y}_i = \bar{\mathbf{y}} + \sum_{k=1}^W t_{ik} \mathbf{f}_k.$$

Here,  $\bar{\mathbf{y}} = \sum_{i=1}^S \mathbf{y}_i / S$ ,  $\bar{\mathbf{y}} \in R^{1 \times W}$  is the sample mean of the spectra (the average spectrum),  $t_{ik}$  is the  $k^{\text{th}}$  coordinate of the vector  $\mathbf{y}_i$  in the new basis system and  $\mathbf{f}_k = (f_k(\lambda_1), f_k(\lambda_2), \dots, f_k(\lambda_W)) \in R^{1 \times W}$  is the  $k^{\text{th}}$  basis vector. Noting that high-dimensional real data are often situated on or near a lower-dimensional manifold, the spectrum  $\mathbf{y}_i$  can be projected onto the  $K$ -dimensional subspace ( $K < W$ ) as follows:

$$\mathbf{y}_i \approx \bar{\mathbf{y}} + \sum_{k=1}^K t_{ik} \mathbf{f}_k, \quad (1)$$

or in matrix form for the initial dataset:

$$\mathbf{Y} \approx \bar{\mathbf{Y}} + \mathbf{T}\mathbf{F},$$

where  $\bar{\mathbf{Y}} = \{\bar{\mathbf{y}}, \dots, \bar{\mathbf{y}}\} \in R^{S \times W}$ ,

$\mathbf{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K\}^T \in R^{K \times W}$ ,  $\mathbf{T} \in R^{S \times K}$  is the matrix

whose entries are  $\{t_{ik}\}_{i=1, \dots, S}^{k=1, \dots, K}$ . Hereinafter the super-

script  $^T$  stands for ‘‘transpose’’. The transformation (1) can be done using dimensionality reduc-

tion techniques, such as PCA [15]. In the latter, basic vectors  $\mathbf{f}_k$  in (1) are referred to as ‘‘principal components’’ (PCs) or empirical orthogonal functions (EOFs) and are taken as  $K$  eigenvectors related to the  $K$  most significant eigenvalues of the

covariance matrix  $\text{cov}(\mathbf{Y}, \mathbf{Y}) \in R^{W \times W}$ . The coordinates  $t_{ik}$  in the new coordinates system and the corresponding matrix  $\mathbf{T}$  are called ‘‘principal component scores’’.

### 3. PRINCIPAL COMPONENT-BASED RADIATIVE TRANSFER MODEL FOR HYPER-SPECTRAL SIGNALS

The most conceptually simple approach uses the training data set of spectra in order to establish a set of EOFs by using PCA and then to restore hyper-spectral signal in  $W$  spectral points by using  $K$  monochromatic radiances. Naturally, we have  $K < W$ .

The theory of PCA briefly discussed in the previous section reveals a linear relationship between PC scores and monochromatic radiances:

$$\mathbf{y}(\lambda) = \bar{\mathbf{y}}(\lambda) + t_1 \mathbf{f}_1(\lambda) + \dots + t_K \mathbf{f}_K(\lambda). \quad (2)$$

Hence, for a given set of  $K$  EOFs and  $K$  spectral points it is possible to obtain a closed linear system of  $K$  equations:

$$\begin{cases} \mathbf{y}(\lambda_1) = \bar{\mathbf{y}}(\lambda) + t_1 \mathbf{f}_1(\lambda_1) + \dots + t_K \mathbf{f}_K(\lambda_1), \\ \mathbf{y}(\lambda_2) = \bar{\mathbf{y}}(\lambda) + t_1 \mathbf{f}_1(\lambda_2) + \dots + t_K \mathbf{f}_K(\lambda_2), \\ \dots \\ \mathbf{y}(\lambda_K) = \bar{\mathbf{y}}(\lambda) + t_1 \mathbf{f}_1(\lambda_K) + \dots + t_K \mathbf{f}_K(\lambda_K). \end{cases} \quad (3)$$

The key point here is that the radiance values in  $K$  spectral points are represented through the same EOFs. Then, by solving (3) we obtain PC scores  $t_1, \dots, t_K$ , and, by using (2), the full spectrum in  $W$  spectral points can be readily restored.

This approach requires a set of precomputed EOFs which is derived from a training data set of simulated or measured spectra. Fig. 3 shows the example of the PCA applied to the dataset of spectra computed in the Hartley-Huggins band used for ozone retrieval. The data set consists of  $10^5$  spectra. The following parameters are varied for the generation of reflectance spectra: the solar zenith angle, the viewing zenith angle, the relative azimuthally

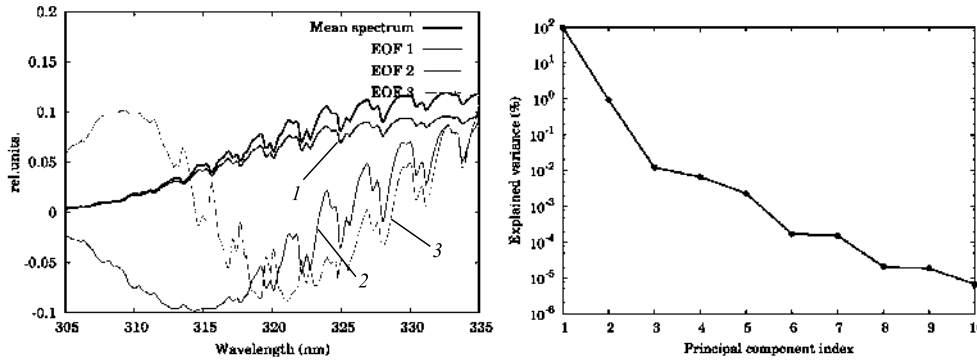


Fig. 3. (Left) Mean spectrum and first three empirical orthogonal functions computed in the Huggins band; (right) explained variance in percentage as a function of the principal component index

angle, the surface albedo, the ozone total column, the surface height, and the temperature. The right plot in Fig. 3 shows that almost 99.9 % of the variance in the data can be explained just with 5 principal components.

To obtain best efficiency, the parameters of this scheme, such as  $K$  and a set of chosen wavelengths, have to be tuned empirically. The number  $K$  depends on the desired level of variance to be captured by the principal components. Several semi-empirical rules have been proposed for the optimal number of principal components (e.g. the broken stick model [36]). However, there is no universal rule for the selection of  $K$ ; the choice is application-specific.

Regarding the choice of wavelengths, in [37] a method is proposed for selecting the location of monochromatic wavelengths by using a correlation function. In particular, this method involves the following steps:

- The correlation coefficients are computed for the radiance values and then converted to vector angles by an arccosine function;
- The spectral data is rearranged according to the magnitudes of the correlation coefficients;
- The monochromatic radiances are selected by choosing predictors with equal distances in the values of the correlation coefficients.

The schematic representation of the principal component-based hyper-spectral radiative transfer model is shown in Fig. 4. For the input data set containing optical parameters of the atmosphere for a set of wavelengths, the monochromatic radiative transfer solver is called. To obtain the most representative dataset, the smart sampling method [38] is recommended based on Halton sequences [39]. That produces a data set of spectra, which is divided into training data set and validation data set. By applying PCA and the correlation analysis to the training set, the system of EOFs is computed and a sub-

set of spectral points is chosen (spectral sampling), respectively. These two outputs are stored and used for computing PC scores for the validation data set. The spectra in the full wavelength range are restored using Eq. (2) and the error of this reconstruction can be estimated. If the error is larger than required, the number of generated spectra and the number of principal components) are increased. The main output of the training phase are empirical orthogonal functions and spectral sampling (marked with red in Fig. 4), which allow to process new data in the online phase (as shown in Fig. 5).

In [37], it was noted that the slit function convolution operator and the PCA are linear. Therefore the PC scores of the *convolved* spectra are linear functions of monochromatic radiances. Then at the training stage, the corresponding weighting factors of linear dependency are stored together with the system of EOFs and spectral sampling. Authors claim that for an infrared spectrum ranging from  $650\text{ cm}^{-1}$  to  $3000\text{ cm}^{-1}$  this representation for convolved spectra reduces the number of monochromatic computations from a few thousands to a few hundreds.

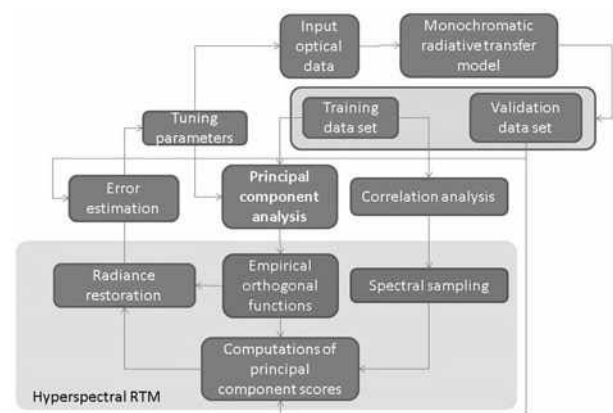


Fig. 4. Schematic representation of the PCA-based radiative transfer model with precomputed empirical orthogonal functions – offline phase



Fig. 5 The same as in Fig. 4, but online phase

The presented approach proved to be efficient and implemented in several packages (e.g. PCTRM [37], RTTOV [40] and others [41, 42]). The main drawback is that it requires time consuming computations of the training data set.

#### 4. PRINCIPAL COMPONENT ANALYSIS IN THE FRAMEWORK OF DIFFERENTIAL OPTICAL ABSORPTION SPECTROSCOPY

In this section we consider the application of the PCA in the framework of the differential optical absorption spectroscopy (DOAS) [43]. DOAS is widely used to retrieve the trace gas content. The main advantages of this technique are the simplicity and robustness as it is able to filter out the influence of factors, which are not taken into account properly in the RTM. The main idea of DOAS consists in the fact that the absorption caused by gases leads to a strong spectral signatures in the spectral radiances, while the influence of multiple scattering and other factors is usually smooth in the wavelength space. Following [44], the radiance reflected by the atmosphere consisting of  $N_g$  gas species with absorption coefficients  $\sigma_{abs,g}$  is represented using the weak absorption Beer-Lambert law as

$$\ln y(\lambda) = -\sum_{g=1}^{N_g} S_g \sigma_{abs,g}(\lambda) - P(\lambda) - RRS(\lambda), \quad (4)$$

where  $S_g$  is the number density of gas  $g$  along the optical path (also referred to as the slant column density),  $P(\lambda)$  is the polynomial term which represents the impact of Rayleigh and aerosol/cloud scattering as well as the surface reflectance, while RRS is a term representing the rotational Raman scattering [45, 46]. In the conventional DOAS,  $S_g$  is retrieved through least squares fitting that minimizes the residual between the measured (left part of Eq. (4)) and simulated (right part of Eq. (4)) radiance spectra. Then  $S_g$  is converted into the vertical column density ( $\Omega_g$ ) through the air mass factor (AMF). The latter is computed at a single wavelength as-

suming a prescribed vertical profile of gas  $g$  [47, 48].

In [49] the modification of the DOAS approach was proposed for retrieving the  $SO_2$  total column. The PCA is applied for the measured spectra in regions with no significant  $SO_2$ , e.g. the equatorial Pacific:

$$\ln y(\lambda) = \overline{\ln y}(\lambda) + \sum_{i=1}^K t_i \mathbf{f}_i(\lambda). \quad (5)$$

In this way, the EOFs capture the variability of the data caused by physical processes (i.e. Rayleigh and Raman scattering and ozone absorption). In addition, the features of the instrument (e.g. the instrumental degradation, the slit function and measurement artefacts) are implicitly accounted for by EOFs. That is the training phase. Then, for polluted regions with  $SO_2$ , representation (5) will produce a residual which is associated with  $SO_2$  content. Thus,

$$\ln y(\lambda) = \overline{\ln y}(\lambda) + \sum_{i=1}^K t_i \mathbf{f}_i(\lambda) + \Omega_{SO_2} \frac{\partial \ln y(\lambda)}{\partial \Omega_{SO_2}}, \quad (6)$$

where  $\Omega_{SO_2}$  is the  $SO_2$  vertical column density. The derivative in the last term can be estimated either by finite differences or by using linearized radiative transfer models [50, 51]. Then  $\Omega_{SO_2}$  can be readily retrieved from Eq. (6).

This method has been applied to the Ozone Monitoring Instrument (OMI) [52] data in the spectral range (310–340) nm. As the high order principal components represent the noise rather than a useful signal, the truncation over the principal components also acts as a filter. To reconstruct the spectral radiances, at least 20–30 principal components were required while in the presence of relatively strong  $SO_2$  signals that number could be reduced to 8. Authors claim that the noise in the data was decreased by factor of 2 thereby providing greater sensitivity to anthropogenic sources of  $SO_2$ .

So far, there are no reports of applying the similar approach to other trace gases. One reason for that is the difficulty of obtaining the system of EOFs for regions without a certain trace gas. The second reason is that, strictly speaking, representation (6) is approximate. For  $SO_2$  it works correctly and the residual is associated with the  $SO_2$  signal. For other trace gases representation (6) might be not valid and more elaborated approach is required.

## 5. DIMENSIONALITY REDUCTION IN RTMS WITH MACHINE LEARNING

### 5.1. General Consideration

Following [53], the inverse problem is solved by reducing it to an exercise in optimization. The main idea behind this method is to find the state vector that minimizes the residual between simulated data and measurements. A non-linear inverse problem is solved iteratively [54, 55]. Assuming an *a priori* state vector  $\mathbf{x}$ , a non-linear forward model is linearized about  $\mathbf{x}$ . Then, the linearized model can be easily inverted and a new estimation for the state vector can be found. This iterative approach is widely used for trace gases retrieval as well as for estimating aerosol and cloud properties [56]. However this inversion method is very time-consuming, due to repeated calls to complex radiative transfer forward models that simulate radiances and Jacobians (i.e. matrices of the first-order partial derivatives of spectral radiances with respect to  $\mathbf{x}$ ), and subsequent inversion of relatively large matrices. These considerations motivate the development of alternative inversion techniques for remote sensing real-time applications, which are based on machine learning and therefore sometimes referred to as full-physics inverse learning machines (FP-ILM) [57].

### 5.2. The Concept of Learning Machines for Atmospheric Retrievals

Machine learning algorithms do not consider the optimization problem explicitly. Rather, they *learn* from a given dataset and make predictions regarding parameters of interest. Conceptually, the machine learning algorithm consists of a training phase, wherein the inversion operator is obtained using synthetic data generated by the radiative transfer model, which expresses the “full-physics” component, and an operational phase, in which the inversion operator is applied to real measurements. Here the main advantage over the classical optimization approach is that the time-consuming training phase involving complex radiative transfer modeling is performed off-line; the inverse operator itself is robust and computationally simple.

Fig. 6 is a schematic representation of the possible implementation of the learning machine. During the training phase, a training dataset is computed using a full-physics forward model, which

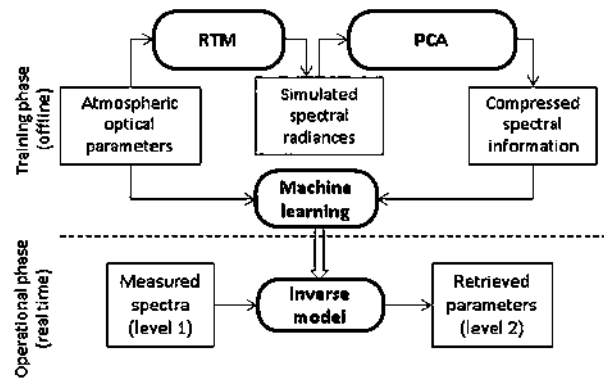


Fig. 6. Schematic representation of the machine learning retrieval algorithm which exploits the dimensionality reduction of the spectral radiances

in our case is the radiative transfer model. In order to capture the essential features of the simulated data and to avoid “over-dimensionality” (the so-called Hughes effect [58]), the simulated spectral data are compressed using an appropriate dimensionality-reduction technique. The mapping between the dimensionality-reduced spectral simulations and the parameter of interest is captured via machine learning.

### 5.3. Machine Learning Based on the Linear Regression Schemes

In the retrieval algorithms based on linear regression, the following representation for the retrieved parameter  $x$  is exploited:

$$x = c + \sum_{w=1}^W l_w y(\lambda_w),$$

where  $c$  is the linear offset and  $l_w$  are the regression coefficients. The principal component regression (PCR) method employs the linear regression model between  $x$  and the principal component scores of the spectral radiance:

$$x = c + \sum_{k=1}^K l_k t_k;$$

As  $K \ll W$ , the dimension of the linear regression model (and the corresponding inverse problem) is reduced. Moreover, since the instrument noise does not affect PC scores of low order, the whole inversion scheme is more stable.

For noisy data, the set of eigenvectors  $\mathbf{F}$  must be computed for the matrix  $\mathbf{C}_Y + \mathbf{C}_e$ , rather than for  $\mathbf{C}_Y$ , where  $\mathbf{C}_e$  is the noise covariance matrix. In this case, the PC scores for the noisy data are cor-

related and are therefore called “projected principal components” [59]. If the statistics of the noise is unknown, the noise covariance matrix can be estimated by making some assumptions (e.g., Gaussian noise) or by using the following approximation  $\mathbf{C}_e \approx \alpha \mathbf{I}$ , where  $\mathbf{I}$  is the identity matrix and  $\alpha$  is the regularization parameter. This procedure reduces the impact of high-order principal components.

The kernel ridge regression (KRR) algorithm [60] generalizes the PCR method; KRR has been used for predicting atmospheric profiles from the IASI (the infrared atmospheric sounding interferometer) instrument [61]. One drawback of the PCR and KRR models is that the basis vectors  $\mathbf{F}$  characterize the measurements  $\mathbf{Y}$ , while information contained in  $\mathbf{X}$  is not taken into account. An alternative model that gets round this drawback is the partial least squares regression (PLSR) [62]. In [62, 64], it was shown that PLSR leads to model-fitting with fewer PCs than required with PCR. In its turn, the PLSR approach can be generalized to the case when we are retrieving a set of correlated parameters (e.g., the temperature profile) rather than a single variable  $x$ . The corresponding method is then referred to as canonical correlations [65]. The use of canonical correlations in atmospheric sciences applications is summarized in [66].

The approach based on the PCR has been successfully applied for solving the problems of volcanic plume-height retrieval from GOME-2 [67] and TROPOMI measurements [68], as well as CO<sub>2</sub> retrieval from GOSAT measurements [69, 70].

## 6. DIMENSIONALITY REDUCTION OF INPUT OPTICAL DATA

### 6.1. Spectra Simulation

An efficient technique using the dimensionality reduction of the optical data has been proposed in [71]. This method relies on the local linearization of the radiative transfer model with respect to input parameters using finite differences. To reduce the number of radiative transfer model calls for estimating finite difference values, the linearization is done in the reduced data space. The method can be summarized as follows. We introduce a correction function as follows:

$$Q(\lambda_w) = \ln[y(\lambda_w)/y_a(\lambda_w)]. \quad (7)$$

Here  $y$  is the radiance computed with a full radiative transfer model, while  $y_a$  is the radiance computed with an approximate model (e.g., the two-stream model). Then, for the atmosphere consisting of  $L$  layers, we consider a state vector  $\mathbf{x}_w \in R^{2L+1}$  containing optical parameters for all layers, i.e.,

$$\mathbf{x}_w^T = \begin{bmatrix} \sigma_{\text{abs},1}(\lambda_w), \dots, \sigma_{\text{abs},L}(\lambda_w), \\ \sigma_{\text{sct},1}(\lambda_w), \dots, \sigma_{\text{sct},L}(\lambda_w), \rho(\lambda_w) \end{bmatrix},$$

where  $\sigma_{\text{sct},i}$  and  $\sigma_{\text{abs},i}$  are the scattering coefficient and the absorption coefficient in the  $i^{\text{th}}$  layer, respectively, while  $\rho$  is the surface albedo. Thus, the wavelength variability of the optical parameters, representing the radiative transfer code input parameters, is encapsulated in the vector  $\mathbf{x}_w$ . Note, that the phase function is assumed to be constant within a given spectral interval and therefore not included in the vector  $\mathbf{x}_w$ . By applying the PCA to  $\{\mathbf{x}_w\}_{w=1}^W$ , we obtain

$$\mathbf{x}_w \approx \bar{\mathbf{x}} + \sum_{k=1}^K t_{wk} \mathbf{f}_k, \quad \bar{\mathbf{x}} = (1/W) \sum_{w=1}^W \mathbf{x}_w.$$

Fig. 7 shows the results of PCA for input optical data in the Huggins band (315–335 nm) and O2A band (755–775 nm). Optical data is taken from [72]. Note that four principal components are sufficient to capture 99.9 % variability of the datasets.

Now, let us assume that  $Q(\mathbf{x}_w)$  can be approximated sufficiently well by its Taylor expansion around  $\bar{\mathbf{x}}$ , that is,

$$\begin{aligned} Q(\mathbf{x}_w) &\approx Q(\bar{\mathbf{x}} + \Delta \mathbf{x}_w) \approx \\ &\approx Q(\bar{\mathbf{x}}) + \Delta \mathbf{x}_w^T \nabla Q(\bar{\mathbf{x}}) + \frac{1}{2} \Delta \mathbf{x}_w^T \nabla^2 Q(\bar{\mathbf{x}}) \Delta \mathbf{x}_w, \end{aligned} \quad (8)$$

where  $\nabla Q$  and  $\nabla^2 Q$  are the gradient and the Hessian of  $Q$ , respectively. By using central differences to approximate the first and the second-order directional derivatives in (8), we obtain

$$\begin{aligned} Q(\mathbf{x}_w) &\approx Q(\bar{\mathbf{x}}) + \frac{1}{2} \sum_{k=1}^K \left[ Q(\bar{\mathbf{x}} + \mathbf{f}_k) - Q(\bar{\mathbf{x}} - \mathbf{f}_k) \right] t_{wk} \\ &\quad + \frac{1}{2} \sum_{k=1}^K \left[ Q(\bar{\mathbf{x}} + \mathbf{f}_k) - 2Q(\bar{\mathbf{x}}) + Q(\bar{\mathbf{x}} - \mathbf{f}_k) \right] t_{wk}^2. \end{aligned} \quad (9)$$



From (9) and (7) it is apparent that the computation of the correction factor requires  $2K + 1$  calls of the full- and two-stream models. Note that if we estimated the correction function using finite differences in the initial data space, that would require  $2L + 1$  calls of the full- and two-stream models. As a result and taking into account that usually  $K \ll L$ , we are led to a substantial reduction of the computational time.

This approach has been applied for simulating the spectra in the O2A band [71], the Huggins band [34], and CO<sub>2</sub> bands [73, 74]. Kopparla et al [75] applied the similar approach for modelling the radiances in the UV/Vis/NIR spectral range (0.3–3000) nm. In all cases authors reported that the root mean square errors of the computed radiances are of order 0.01 %, yet achieving almost a 10-fold increase in speed. The big advantage of this method is that unlike previously considered techniques, this one does not require precomputed databases of spectra.

In [76], the efficiency of input and output space dimensionality reduction techniques was analyzed for simulating the Hartley-Huggins band. The hybrid usage of these techniques was proposed. The output space reduction and the spectral sampling methods are applied to the two-stream solution by using corresponding lookup tables, while multi-stream solution computations are performed within the input data reduction framework, described in this Section. It was found that the combined use of these techniques yields accuracy better than 0.05 % while the speedup factor is about 20.

## 6.2. Retrieval in the Reduced Input Data Space

Since the atmospheric retrieval problem is severely ill-posed, a physically correct result can be obtained only by using a regularization procedure. The latter takes into account some *a priori* information. In this context, dimensionality reduction of the input data space can be regarded as a special type of regularization, i.e. the retrieved parameters should obey a certain dependency reproduced by a chosen set of EOFs.

Timofeyev et al. [77] applied the dimensionality reduction technique to parameterize the aerosol extinction coefficient for incorporation into the inversion algorithm, in which the corresponding PC scores rather than aerosol extinction dependence were retrieved. The system of EOFs was defined for

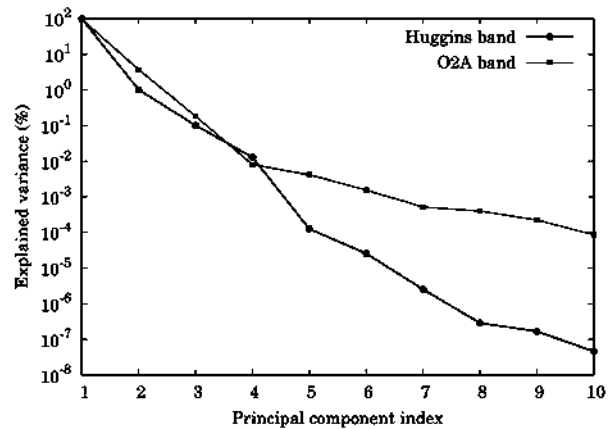


Fig. 7. The explained variance in percentage as a function of the principal component index in the Huggins band and the O2A band

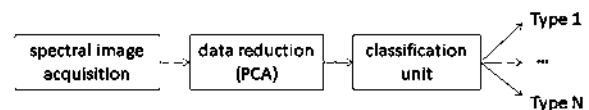


Fig. 8. A schematic representation of the PCA-based classification algorithms applied to the hyperspectral data

a dataset of aerosol extinction coefficients computed on the base of Mie theory [78] and algorithms for particle ensembles. Finally, in [79] the dimensionality reduction is performed in the input (temperature and humidity profiles) and output spaces (spectral radiances), while artificial neural networks are used to establish the interdependency between PC scores. Since the number of independent parameters is reduced, such scheme is more robust and efficient than the conventional one.

## 7. DIMENSIONALITY REDUCTION OF HYPER-SPECTRAL DATA IN CLASSIFICATION PROBLEMS

Dimensionality reduction plays an important role in classification algorithms applied to the hyper-spectral data. On-line anomaly detection and object recognition in remote sensing imagery is extremely important for forest fire and volcanic activity monitoring. In such kind of applications we are confronted with the classification problem. In a reduced data space, the number of variables used in classification is smaller, yet their value is higher than that in the original space. Therefore, the classification algorithms (e.g.  $K$ -nearest neighbour) are more robust and efficient [80, 81]. PCA can be used to visualize the hyper-spectral

data on a 2D plane, thereby identifying regions with certain features. The concept of combined usage of PCA with classifiers is illustrated in Fig. 8. Such an approach is used not only in space-borne data processing, but also in other fields, such as material science [82, 83], tobacco industry [84] and food production [85].

## 8. CONCLUSIONS

In this review, several techniques of hyper-spectral data processing have been considered. They are all based on the dimensionality reduction procedure. It has been shown that the principal component analysis can be utilized in several ways for hyper-spectral modelling. Therefore it seems that the nomenclature “PCA-based radiative transfer model” is not appropriate since it does not characterize the specific features of the algorithm (e.g., papers [40, 49, 75] present absolutely different models, although all of them are “PCA-based”).

It has been shown that the concept of dimensionality reduction gives the framework for formulating hyper-spectral RTM that directly takes into account a strong interdependency in the hyper-spectral data. Further research needs to examine more closely how to combine techniques outlined in this review. For instance in [86], a hybrid approach comprising the correlated- $k$  method and the dimensionality reduction of the input data has been described. Such models are extremely important for processing remote sensing Big Data in the current missions, and becoming a mainstream in the development of next generation atmospheric processors.

The discussed principles of data reduction of hyper-spectral data are generic and can be applied in various applications, including material science and electron spectroscopy due to the similar methodologies adopted in these fields [87]. In addition, PCA is a perspective tool for analysing hyper-spectral optical data in medicine. In particular, recent studies have shown that the dimensionality reduction of data cubes can improve the recognition and classification algorithms, which would be extremely important for optical early disease diagnostics [88].

## ACKNOWLEDGEMENTS

The authors are grateful to the Editor-in-Chief of the journal “Light and Engineering” professor

Vladimir P. Budak for inviting to write this survey, as well as to the anonymous reviewer for careful reading of the manuscript and valuable suggestions.

## REFERENCES

1. D.G. Loyola, S. Gimeno Garcia, R. Lutz, A. Argyrouli, F. Romahn, R.J.D. Spurr, M. Pedernana, A. Doicu, V. Molina Garcá, and O. Schüssler. The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 precursor. *Atmospheric Measurement Techniques*, 2018,11(1), pp.409–427.
2. C.A. Lee, S.D. Gasster, A. Plaza, C. Chang, and B. Huang. Recent developments in high performance computing for remote sensing: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 20114(3), pp.508–527.
3. J.P. Veefkind, I. Aben, K. McMullan, H. Forster, J. de Vries, G. Otter, J. Claas, H.J. Eskes, J.F. de Haan, Q. Kleipool, and et al. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 2012, 120, pp.70–83.
4. Y. Ma, H. Wu, L. Wang, B. Huang, R. Ranjan, A. Zomaya, and W. Jie. Remote sensing big data computing: Challenges and opportunities. *Future Generation Computer Systems*, 2015, 51, pp.47–60.
5. P. Liu. A survey of remote-sensing big data. *Frontiers in Environmental Science*, 2015, 3, 45p.
6. V.M. Roozendael, R. Spurr, D. Loyola, C. Lerot, D. Balis, J.C. Lambert, W. Zimmer, J. Gent, J. Van Geffen, M.E. Koukouli, J. Granville, A. Doicu, C. Fayt, and C. Zehner. Sixteen years of GOME/ERS2 total ozone data: the new direct-fitting GOME Data Processor (GDP) Version 5: I. algorithm description. *J Geophys Res: Atmospheres*, 2012,117: D03305, pp.1–18.
7. V. Natraj. A review of fast radiative transfer techniques. In A.A. Kokhanovsky, editor, *Light scattering reviews*, volume 8, pages 475–504. Springer Berlin Heidelberg, 2013.
8. D. Efremenko, A. Doicu, D. Loyola, and T. Trautmann. Acceleration techniques for the discrete ordinate method. *J Quant Spectrosc Radiat Transfer*, 2013,114, pp.73–81.
9. D.S. Efremenko, D. Loyola, A. Doicu, and T. Trautmann. Data-intensive computing in radiative transfer modelling. In P. Soille and P.G. Marchetti, editors, *Proc. of the 2016 conference on Big Data from Space (BiDS16)*, Santa Cruz de Tenerife, Spain, 2016, pp. 188–191.

10. V.P. Budak, G.A. Kaloshin, O.V. Shagalov, and V.S. Zheltov. Numerical modeling of the radiative transfer in a turbid medium using the synthetic iteration. *Opt. Express*, 23(15): A829, 2015.
11. D.S. Efremenko, D.G. Loyola, A. Doicu, and R.J.D. Spurr. Multi-core-CPU and GPU-accelerated radiative transfer models based on the discrete ordinate method. *Computer Physics Communications*, 2014, 185(12), pp.3079–3089.
12. D.S. Efremenko, D. Loyola, R.J.D. Spurr, and A. Doicu. Acceleration of radiative transfer model calculations for the retrieval of trace gases under cloudy conditions. *J Quant Spectrosc Radiat Transfer*, 2014,135, pp.58–65.
13. R.J.D. Spurr. VLIDORT: A linearized pseudo-spherical vector discrete ordinate radiative transfer code for forward model and retrieval studies in multilayer multiple scattering media. *J Quant Spectrosc Radiat Transfer*, 2006, 102(2), pp.316–342.
14. R.J.D. Spurr, T.P. Kurosu, and K.V. Chance. A linearized discrete ordinate radiative transfer model for atmospheric remote-sensing retrieval. *J Quant Spectrosc Radiat Transfer*, 2001, 68(6), pp.689–735.
15. K. Pearson. On lines and planes of closest fit to systems of points in space. *Phil Mag*, 19012, 6, pp.559–572.
16. R. D. Morris, A. Kottas, M. Taddy, R. Furfaro, and B.D. Ganapol. A statistical framework for the sensitivity analysis of radiative transfer models. *IEEE Transactions on Geoscience and Remote Sensing*, 2008,46(12), pp.4062–4074.
17. V.A. Ambartzumyan. The effect of the absorption lines on the radiative equilibrium of the outer layers of the stars. *Publ. Obs. Astron. Univ. Leningrad*, 1936, 6, pp.7–18.
18. R.M. Goody, R. West, L. Chen, and D. Crisp. The correlated k-method for radiation calculations in nonhomogeneous atmosphere. *J Quant Spectrosc Radiat Transfer*, 1989, 42, 6, pp.539–550.
19. W.J. Wiscombe and J.W. Evans. Exponential-sum fitting of radiative transmission functions. *J Comput Phys*, 1997, 24, 4, pp.416–444.
20. R. West, D. Crisp, and L. Chen. Mapping transformations for broadband atmospheric radiation calculation. *J Quant Spectrosc Radiat Transfer*, 1990, 43,3, pp.191–199.
21. E. Boesche, P. Stammes, R. Preusker, R. Benartz, W. Knap, and J. Fischer. Polarization of skylight in the O2A band: effects of aerosol properties. *Applied Optics*, 2008, 47, 19, p. 3467.
22. Christiane Helling and Uffe Grae Jorgensen. Optimizing the opacity sampling method. *Astronomy and Astrophysics*, 1998, 337, pp.477–486.
23. Frédéric André, Longfeng Hou, Maxime Roger, and Rodolphe Vaillon. The multispectral gas radiation modeling: A new theoretical framework based on a multidimensional approach to k-distribution methods. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2014,147, pp.178–195.
24. Boris A. Fomin. A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 1. FKDM, fast k-distribution model for the longwave. *Journal of Geophysical Research*, 109(D2), 2004.
25. Boris Fomin. A k-distribution technique for radiative transfer simulation in inhomogeneous atmosphere: 2. FKDM, fast k-distribution model for the shortwave. *Journal of Geophysical Research*, 110(D2), 2005.
26. V. A. Falaleeva and B.A. Fomin. Overcoming spectroscopic challenges in direct problems of satellite sounding of the atmosphere. *Atmospheric and Oceanic Optics*, 2017, 30(1), pp.1–6.
27. B.A. Fomin. Effective interpolation technique for line-by-line calculations of radiation absorption in gases. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 1995, 53(6), pp.663–669.
28. S. Najmabadi, P. Offenhäuser, M. Hamann, G. Jajnabalkya, F. Hempert, C. Glass, and S. Simon. Analyzing the effect and performance of lossy compression on aeroacoustic simulation of gas injector. *Computation*, 2017, 5(4):24.
29. S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000, 290(22), pp.2323–2326.
30. M.A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, 1991, 37(2), pp.233–243.
31. A.N. Gorban, B. Kégl, D.C. Wunsch, and A.Y. Zinovyev, editors. *Principal Manifolds for Data Visualization and Dimension Reduction*. Springer Berlin Heidelberg, 2008.
32. I.K. Fodor. *A Survey of Dimension Reduction Techniques*. Office of Scientific and Technical Information (OSTI), 2002.
33. L.J.P. van der Maaten, E.O. Postma, and H.J. van den Herik. *Dimensionality Reduction: A Comparative Review*. Tilburg University Technical Report, TiCC-TR2009–005, 2009.
34. D.S. Efremenko, A. Doicu, D. Loyola, and T. Trautmann. Optical property dimensionality reduction techniques for accelerated radiative transfer performance: Application to remote sensing total ozone re-

- trievals. *J Quant Spectrosc Radiat Transfer*, 2014, 133, pp.128–135.
35. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011,12, pp.2825–2830.
36. Robert H. MacArthur. On the relative abundance of bird species. *Proc Natl Acad Sci USA*, 1957, 43(3), pp.293–295.
37. X Liu, W.L. Smith, D.K. Zhou, and A. Larar. Principal component-based radiative transfer model for hyperspectral sensors: theoretical concept. *Applied Optics*, 2006, 45(1), pp.201–208.
38. D. G. Loyola, M. Pedernana, and S. Gimeno Garcia. Smart sampling and incremental function learning for very large high dimensional data. *Neural Networks*, 2016, 78, pp.75–87.
39. J. H. Halton. Algorithm 247: Radical-inverse quasi-random point sequence. *Commun ACM*, 1964, 7(12), pp.701–702.
40. M. Matricardi. A principal component based version of the RTTOV fast radiative transfer model. *Quarterly Journal of the Royal Meteorological Society*, 2010, 136, pp.1823–1835.
41. P. D. Hurley, S. Oliver, D. Farrah, L. Wang, and A. Efstathiou. Principal component analysis and radiative transfer modelling of Spitzer Infrared spectrograph spectra of ultraluminous infrared galaxies. *Monthly Notices of the Royal Astronomical Society*, 2012, 424(3), pp.2069–2078.
42. A. Hollstein and R. Lindstrot. Fast reconstruction of hyperspectral radiative transfer simulations by using small spectral subsets: application to the oxygen A band. *Atmospheric Measurement Techniques*, 2014, 7(2), pp.599–607.
43. U. Platt. Differential optical absorption spectroscopy (DOAS). *Chem Anal Series*, 1994, 127, pp.27–83.
44. U. Platt and J. Stutz. *Differential Optical Absorption Spectroscopy: Principles and Applications*. Springer-Verlag, Berlin, Heidelberg, 2008.
45. G.W. Kattawar, A.T. Young, and T.J. Humphreys. Inelastic scattering in planetary atmospheres. I. The Ring effect, without aerosols. *Astrophys J*, 1981, 243, pp.1049–1057.
46. M. Vountas, V.V. Rozanov, and J.P. Burrows. Ring effect: Impact of rotational Raman scattering on radiative transfer in earth's atmosphere. *J Quant Spectrosc Radiat Transfer*, 1998, 60(6), pp.943–961.
47. James Slusser, Kyle Hammond, Arve Kylling, Knut Stamnes, Lori Perliski, Arne Dahlback, Donald Anderson, and Robert DeMajistre. Comparison of air mass computations. *Journal of Geophysical Research: Atmospheres*, 1996,101(D5), pp. 9315–9321.
48. Lori M. Perliski and Susan Solomon. On the evaluation of air mass factors for atmospheric near-ultraviolet and visible absorption spectroscopy. *Journal of Geophysical Research*, 1993, 98(D6), p.10363.
49. Can Li, Joanna Joiner, Nickolay A. Krotkov, and Pawan K. Bhartia. A fast and sensitive new satellite SO<sub>2</sub> retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument. *Geophysical Research Letters*, 2013, 40(23), pp. 6314–6318.
50. R.J.D. Spurr. LIDORT and VLIDORT. Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems. In A.A. Kokhanovsky, editor, *Light scattering reviews*, 2008, V. 3, pp. 229–275.
51. A. Doicu and T. Trautmann. Two linearization methods for atmospheric remote sensing. *J Quant Spectrosc Radiat Transfer*, 2009, 110(8), pp. 477–490.
52. P.F. Levelt, G.H.J. van den Oord, M.R. Dobber, A. Malkki, Huib Visser, Johan de Vries, P. Stammes, J.O.V. Lundell, and H. Saari. The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(5), pp.1093–1101.
53. A.N. Tikhonov and V.Y. Arsenin. *Solution of Ill-Posed Problems*. Winston, 1977.
54. C.D. Rodgers. *Inverse methods for atmospheric sounding: Theory and Practice*. World Scientific Publishing, 2000.
55. A. Doicu, T. Trautmann, and F. Schreier. *Numerical Regularization for Atmospheric Inverse Problems*. Springer, 2010.
56. D. S. Efremenko, O. Schüssler, A. Doicu, and D. Loyola. A stochastic cloud model for cloud and ozone retrievals from UV measurements. *J Quant Spectrosc Radiat Transfer*, November 2016,184, pp.167–179.
57. J. Xu, O. Schussler, D.G. Loyola Rodriguez, F. Romahn, and A. Doicu. A novel ozone profile shape retrieval using full-physics inverse learning machine (FP-ILM). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, 10(12), pp.5442–5457.
58. G. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inform. Theory*, 1968, 14(1), pp.55–63.
59. W.J. Blackwell and F.W. Chen. *Neural Networks in Atmospheric Remote Sensing*. Lexington, 2009.

60. John Shawe-Taylor and Nello Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, New York, NY, USA, 2004.
61. G. Camps-Valls, J. Munoz-Mari, L. Gomez-Chova, L. Guanter, and X. Calbet. Nonlinear statistical retrieval of atmospheric profiles from MetOp-IASI and MTG-IRS infrared sounding data. *IEEE Transactions on Geoscience and Remote Sensing*, May 2012 50(5), pp.1759–1769.
62. Roman Rosipal and Nicole Krämer. Overview and recent advances in partial least squares. In *Subspace, Latent Structure and Feature Selection*, pages 34–51. Springer Science + Business Media, 2006.
63. Peter D. Wentzell and Lorenzo Vega Montoto. Comparison of principal components regression and partial least squares regression through generic simulations of complex mixtures. *Chemometrics and Intelligent Laboratory Systems*, 2003, 65(2), pp.257–279.
64. Saikat Maitra and Jun Yan. Principal component analysis and partial least squares: Two dimension reduction techniques for regression. In *Discussion Papers: 2008 Discussion Paper Program – Applying Multivariate Statistical Models*, 2008, pp.79–90, Casualty actuarial society.
65. H. Hotelling. The most predictable criterion. *Journal of Educational Psychology*, 1935, 26, pp.139–142.
66. D. Wilks. *Statistical Methods in the Atmospheric Sciences, 3rd Edition*. New York: Elsevier, 2011.
67. D.S. Efremenko, D.G. Loyola R., P. Hedelt, and R.J.D. Spurr. Volcanic SO<sub>2</sub> plume height retrieval from UV sensors using a full-physics inverse learning machine algorithm. *International Journal of Remote Sensing*, 2017, 38(sup1), pp.1–27.
68. Pascal Hedelt, Dmitry S. Efremenko, Diego G. Loyola, Robert Spurr, and Lieven Clarisse. SO<sub>2</sub> layer height retrieval from Sentinel-5 Precursor/TROPOMI using FP\_ILM. *Atmospheric Measurement Techniques Discussions*, Feb. 2019, pp.1–23.
69. M.Yu. Kataev, A.K. Lukyanov, and A.A. Bekerov. Modification of the empirical orthogonal functions method for solving the inverse task of retrieving of the CO<sub>2</sub> total content from satellite data. *Journal of Siberian Federal University. Engineering & Technologies*, 2018,11(1), pp.77–85.
70. M.Y. Kataev and A.K. Lukyanov. Empirical orthogonal functions and its modification in the task of retrieving of the total amount CO<sub>2</sub> and CH<sub>4</sub> with help of satellite Fourier transform spectrometer GOSAT (TANSO-FTS). In G.G. Matvienko and O.A. Romanovskii, editors, *22nd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*. SPIE-Intl Soc Optical Eng, 2016.
71. V. Natraj, X. Jiang, R.L. Shia, X. Huang, J.S. Margolis, and Y.L. Yung. Application of the principal component analysis to high spectral resolution radiative transfer: A case study of the O<sub>2</sub>A-band. *J Quant Spectrosc Radiat Transfer*, 2005, 95(4), pp.539–556.
72. I.E. Gordon, L.S. Rothman, C. Hill, R.V. Kochanov, Y. Tan, P.F. Bernath, M. Birk, V. Boudon, A. Campargue, K.V. Chance, B.J. Drouin, J.-M. Flaud, R.R. Gamache, J.T. Hodges, D. Jacquemart, V.I. Perevalov, A. Perrin, K.P. Shine, M.-A.H. Smith, J. Tennyson, G.C. Toon, H. Tran, V.G. Tyuterev, A. Barbe, A.G. Császár, V.M. Devi, T. Furtenbacher, J.J. Harrison, J.-M. Hartmann, A. Jolly, T.J. Johnson, T. Karman, I. Kleiner, A.A. Kyuberis, J. Loos, O.M. Lyulin, S.T. Massie, S.N. Mikhailenko, N. Moazzen-Ahmadi, H.S.P. Müller, O.V. Naumenko, A.V. Nikitin, O.L. Polyansky, M. Rey, M. Rotger, S.W. Sharpe, K. Sung, E. Starikova, S.A. Tashkun, J. Vander Auwera, G. Wagner, J. Wilzewski, P. Wcislo, S. Yu, and E.J. Zak. The HITRAN2016 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2017, 203, pp.3–69.
73. P. Somkuti, H. Boesch, V. Natraj, and P. Koppa. Application of a PCA-based fast radiative transfer model to XCO<sub>2</sub> retrievals in the shortwave infrared. *Journal of Geophysical Research: Atmospheres*, 122(19), pp.10,477–10,496.
74. V. Natraj, R.L. Shia, and Y.L. Yung. On the use of principal component analysis to speed up radiative transfer calculations. *J Quant Spectrosc Radiat Transfer*, 2010, 111(5), pp.810–816.
75. P. Koppa, V. Natraj, R. Spurr, R. Shia, D. Crisp, and Y.L. Yung. A fast and accurate PCA based radiative transfer model: Extension to the broadband shortwave region. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2016, 173, pp. 65–71.
76. Ana del Águila, Dmitry Efremenko, Víctor Molina García, and Jian Xu. Analysis of two dimensionality reduction techniques for fast simulation of the spectral radiances in the Hartley-Huggins band. *Atmosphere*, Mar. 2019,10(3), p.142.
77. Y. Timofeyev, A. Polyakov, H. Steele, and M. Newchurch. Optimal eigenanalysis for the treatment of aerosols in the retrieval of atmospheric composition from transmission measurements. *Appl. Opt.*, July 2003,42 (15), p.2635.
78. G. Mie. Beitrage zur optik trueber medien, speziell kolloidaler metalloesungen. *Annalen der Physik*, 1908, 330(3), pp.377–445.

79. A. V. Polyakov, Yu.M. Timofeev, and Ya.A. Viro-lainen. Using artificial neural networks in the temperature and humidity sounding of the atmosphere. *Izvestiya, Atmospheric and Oceanic Physics*, 2014, 50(3), pp.330–336.
80. Michael T. Eismann, Joseph Meola, and Rus-sell C. Hardie. Hyperspectral change detection in the presence of diurnal and seasonal variations. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(1), pp. 237–249.
81. Dandan Ma, Yuan Yuan, and Qi Wang. Hyper-spectral anomaly detection via discriminative feature learning with multiple-dictionary sparse representation. *Remote Sensing*, May 2018,10(5), p.745.
82. David B Brough, Daniel Wheeler, and Surya R. Kalidindi. Materials knowledge systems in python – a data science framework for accelerated development of hierarchical materials. *Integrating Mate-rials and Manufacturing Innovation*, Mar.2017, 6(1), pp. 36–53.
83. Sai Kiranmayee Samudrala, Prasanna Venkatara-man Balachandran, Jaroslaw Zola, Krishna Rajan, and Baskar Ganapathysubramanian. A software framework for data dimensionality reduction: application to chemical crystallography. *Integrating Materials and Manufacturing Innovation*, 3(1), Jun. 2014.
84. Pilar B. Garcá-Allende, Olga M. Conde, Ana M. Cubillas, César Jáuregui, and José M. López-Higuera. New raw material discrimination system based on a spatial optical spectroscopy technique. *Sensors and Actuators A: Physical*, Apr. 2007, 135(2), pp. 605–612.
85. Juan Xing, Cédric Bravo, Pál T. Jancsó, Herman Ramon, and Josse De Baerdemaeker. Detecting bruises on ‘golden delicious’ apples using hyperspectral imaging with multiple wavebands. *Biosystems Engineering*, Jan 2005, 90(1), pp.27–36.
86. V. Molina García, S. Sasi, D.S. Efremenko, A. Doicu, and D. Loyola. Radiative transfer models for retrieval of cloud parameters from EPIC/DSCOVR measurements. *Journal of Quantitative Spectroscopy and Ra-diative Transfer*, 2018, 213, pp. 228–240.
87. V.P. Afanas’ev, V.P. Budak, D.S. Efremenko, and P.S. Kaplya. Application of the photometric theory of the radiance field in the problems of electron scattering. *Light & Engineering*, 2019, 27(2), pp. 88–96.
88. F. Vasefi, N. MacKinnon, and D.L. Farkas. Hyper-spectral and multispectral imaging in dermatology. In *Imaging in Dermatology*, Elsevier, 2016, pp. 187–201.



### **Ana del Águila**

graduated in Physics from the Granada University (UGR) in 2015. From 2016-2018 she worked as an early-stage researcher at the National Institute for Aerospace Technology (INTA) in Spain. At present, she is doing the Ph.D. in the German Aerospace Centre (DLR) with a DAAD/DLR scholarship. Her scientific interests are in-situ atmospheric aerosols, lidar systems, remote sensing, radiate transfer and Big Data analysis



### **Dmitry S. Efremenko**

graduated from the Moscow Power Engineering institute (MPEI) in 2009. He received his Ph.D. degree from the Moscow State University in 2011 and the habilitation degree from MPEI in 2017. Since 2011 he works as a research scientist at the German Aerospace Centre (DLR). His scientific interests include radiate transfer, remote sensing, and Big Data analysis



### **Thomas Trautmann,**

Prof., Dr., graduated with a diploma in meteorology at the Johannes Gutenberg University at Mainz (JGU) in 1985. From the JGU he received a Ph.D. degree in 1989, a habilitation degree in 1997, and in 2003 he was appointed as applicant Prof. at the University of Leipzig. Since 2003 he works as head of department Atmospheric Processors at the German Aerospace Center (DLR). His scientific interests include radiate transfer, electromagnetic scattering and atmospheric remote sensing