Journal of Quantitative Spectroscopy & Radiative Transfer 206 (2018) 8–21

Contents lists available at ScienceDirect

Journal of Quantitative Spectroscopy & Radiative Transfer

journal homepage: www.elsevier.com/locate/jqsrt

PCA determination of the radiometric noise of high spectral resolution infrared observations from spectral residuals: Application to IASI

C. Serio [a,*], G. Masiello [a], C. Camy-Peyret [b], E. Jacquette [c], O. Vandermarcq [c], F. Bermudo [c], D. Coppens [d], D. Tobin [e]

[a] Scuola di Ingegneria, Universitá della Basilicata, Potenza, Italy
[b] Institut Pierre-Simon Laplace (IPSL), UPMC/UVSQ, Paris, France
[c] CNES, Toulose, France
[d] EUMETSAT, Darmstadt, Germany
[e] CIMMS/SSEC, University of Wisconsin, Madison, USA

ARTICLE INFO

Article history:
Received 22 September 2017
Revised 27 October 2017
Accepted 27 October 2017
Available online 31 October 2017

Keywords:
Instrument characterization
Infrared
Satellite
Spectroscopy
Radiometric noise

ABSTRACT

The problem of characterizing and estimating the instrumental or radiometric noise of satellite high spectral resolution infrared spectrometers directly from Earth observations is addressed in this paper. An approach has been developed, which relies on the Principal Component Analysis (PCA) with a suitable criterion to select the optimal number of PC scores. Different selection criteria have been set up and analysed, which is based on the estimation theory of Least Squares and/or Maximum Likelihood Principle. The approach is independent of any forward model and/or radiative transfer calculations. The PCA is used to define an orthogonal basis, which, in turn, is used to derive an optimal linear reconstruction of the observations. The residual vector that is the observation vector minus the calculated or reconstructed one is then used to estimate the instrumental noise. It will be shown that the use of the spectral residuals to assess the radiometric instrumental noise leads to efficient estimators, which are largely independent of possible departures of the true noise from that assumed *a priori* to model the observational covariance matrix. Application to the Infrared Atmospheric Sounder Interferometer (IASI) has been considered. A series of case studies has been set up, which make use of IASI observations. As a major result, the analysis confirms the high stability and radiometric performance of IASI. The approach also proved to be efficient in characterizing noise features due to mechanical micro-vibrations of the beam splitter of the IASI instrument.

1. Introduction

The Infrared Atmospheric Sounder Interferometer (IASI) [1] has become a reference for space infrared instrumentation, for climate change studies (e.g. see [2]) and for the assessment of forward model and spectroscopy errors (e.g., [3,4]). Thus, the independent characterization of the instrumental noise and of its stability is highly desirable, as an alternative technique to the two point radiometric calibration (warm blackbody and cold space), which could be not available at all times and locations.

This paper addresses the problem of characterizing and estimating the instrumental or radiometric noise of satellite high spectral resolution infrared spectrometers directly from observations whenever and wherever they are available. Toward this objective, we have developed an algorithm, which relies on the Principal Component Analysis (PCA) and a suitable criterion to select the optimal number of PC scores (dimension of the useful eigenvector basis) for the linear reconstruction of observations. Once the linear reconstructed radiances have been computed, the instrumental noise is recovered from the covariance matrix of the residuals i.e. Observations-Calculations or Obs-Calc. The scheme we have developed is fully analytical and includes an automatic choice of the dimensionality for PCA.

Principal Component Analysis and the related Singular Value Decomposition (SVD) theorem (see [5] among many others), have been recognized as the most valuable tools from applied linear algebra. Their combination forms an ubiquitous technique for spectrometric data analysis and processing.

In the wide general context of Meteorology, Empirical Orthogonal (EOF) Functions were first introduced by Lorenz [6] and then by Obukhov [7]. These two authors coined the term Empirical Orthogonal Functions, while the term PCA has been mostly used in

* Corresponding author.
E-mail address: carmine.serio@unibas.it (C. Serio).

https://doi.org/10.1016/j.jqsrt.2017.10.022
0022-4073/

the context of Statistics, where the basic idea was introduced by Pearson in the early 1900's [8]. However, the first work on PCA dates back to 1870's with the paper on Singular Value Decomposition by Beltrami [9]. Formal treatment of the PCA method is due to Hotelling [10] and Rao [11].

In the context of Theoretical Physics, the PCA method is equivalent to the discrete version of the Karhunen-Loéve transform [12,13], which applies to stochastic processes. The link of this transform to PCA is provided by the Singular Value Decomposition theorem applied to the covariance matrix of the given process. We will show that SVD also provides the link between PCA and its probabilistic interpretation through the Maximum Likelihood principle (e.g. see [14]). The link will be exploited to set up suitable criteria for the optimal choice of the number of PCA scores.

Practical methods for computing the SVD were unknown until 1965. The gap was filled by Golub and Kahan [15] and later by Golub and Reinsch [16] who published a variant of the Golub/Kahan algorithm, which is still used today. Since then, the application of PCA/EOF to practical retrieval problems has flourished in many applied research fields.

In the context of Satellite Meteorology, EOF decomposition was first exploited by Wark and Flemming [17] for the estimation of temperature profiles from infrared spectral radiances. The application to high spectral resolution infrared observations have been considered by Huang and co-workers [18–20], who provided a general discussion of PCA as a tool for the analysis and de-noising of infrared spectra. A quite general EOF approach for dimensionality reduction of the data space for application to the radiative transfer equation inverse problem has been provided by Masiello et al. [21,22]. Furthermore, the application of EOF *regression* to high spectral resolution infrared observations for the estimation of geophysical parameters and trace gases have been explored (among many others) by [23–27].

To our knowledge, a systematic analysis and use of PCA combined with SVD decomposition, for the assessment of instrumental noise dates back to Malinowski [28]. Later Tobin and co-workers [29,30] have applied SVD decomposition to high spectral resolution infrared observations for the assessment of calibration artefacts, instrument line shape distortion introduced by inhomogeneous scenes and assessment of radiometric noise. However, one general problem in PCA analysis is the determination of the appropriate number of PCA scores, or the optimal dimension of a linear model, which explains the total variance of the underlying signal.

It appears that within the context of hyper-spectral satellite sensors, the problem has not yet received a rigorous treatment, e.g., through Information and/or Bayes statistical methods. This paper is aiming at filling this gap. As said, we resort to the probabilistic interpretation of PCA given by Tipping et al. [14] and develop appropriate analytical formulas for both the Akaike Information Criterion (AIC) [31] and the Bayesian Information Criterion (BIC) [32]. Both AIC and BIC are common criteria for the estimation of the dimension of a model, however their correct formulation depends on the given problem, and cannot be simply derived in analogy with formulas normally used, e.g., in time series analysis or multivariate Least Squares regression.

The approach will be exemplified through application to IASI [1], for which the methodology has been specifically developed. However, the tool is quite general and can be applied to any spectral radiometer or spectrometer. In fact, the methodology has been developed also in perspective of the IASI follow on instrument or IASI Next Generation (IASI-NG), which is expected to fly in 2022 on board the Meteorological Operational Second Generation (Metop-SG) satellite platform.

Concerning IASI, the analysis has been performed for one year of observations, which has allowed us to follow the time evolution of the radiometric noise and check for its stability.

It will be shown that our observation-based errors agree with blackbody in-flight calibration. A comparison will be also provided with a method [33], which computes the spectral residual on the basis of a forward model. The comparison will allow us to assess the potential advantages of a methodology, such as PCA, which is independent of any radiative transfer calculations and to exemplify how the methodology can be used to check the consistency of forward-inverse tools for the retrieval of geophysical parameters and gas species.

The paper is organized as follow. Section 2 will describe data and methods used in the analysis. Section 3 will describe the results: Section 3.1 is mostly dealing with numerical exercises based on simulations; Section 3.2 will deal with applications to real observations and comparison with blackbody in-flight calibration. Comparison with a forward-based methodology will be performed in Section 3.2, as well. The problem of noise features due to mechanical micro-vibrations of the beam splitter of the IASI instrument is covered in Section 3.3. Finally conclusions will be drawn in Section 4.

2. Instrument, data and methods

2.1. Instrument and data

IASI [1] was developed in France by CNES and is flying on board the Metop platforms. These are satellites of the EUMETSAT European Polar System (EPS). The instrument has a spectral coverage extending from 645 to 2760 cm^{-1}, with a sampling interval $\Delta\sigma = 0.25\ \mathrm{cm}^{-1}$ providing 8461 data points or channels for each single spectrum. In the normal mode of operation, IASI is a cross-track scanner, with 30 fields of regard (FOR) per scan, spanning an off-nadir angle range of $\pm 48.33°$ on either side of nadir; the two symmetric near-nadir FORs are at angles of $\pm 1.67°$. Each FOR consists of a 2×2 matrix of four so-called instantaneous fields of view (IFOV). Each IFOV has a diameter of 14.65 mrad, which corresponds to a ground resolution of 12 km at nadir and for a satellite altitude of 819 km. The 2×2 matrix is centred on the viewing direction. At nadir the FOR of 4 IASI pixels project at ground a square area of $\approx 50 \times 50\ \mathrm{km}^2$. More details on IASI and its mission objectives can be found in [1].

The data we use in this analysis have been obtained with IASI operated in the so-called *external calibration mode* or ExtCal. In this mode, each scanning line, which in normal operation mode spans an angle range of $\pm 48.33°$ on either side of nadir, is squeezed into nadir view. In other words, in the same time as that taken for a complete $\pm 48.33°$ scan, IASI is operated to look exactly at nadir. This result in two sub-satellite lines, which correspond to 27 overlapping nadir FORs, each one with its 4 IFOVs. IFOV-1/IFOV-2 form the first sub-staellite line and IFOV-3/IFOV-4 form the second sub-satellite line. In this way, the dual measurement line consists of a total of $4 \times 27 = 108$ overlapping nadir IFOVs, which sense an area of $\approx 70 \times 50\ \mathrm{km}^2$ (see Figs. 1 and 2). In the ExtCal mode the horizontal spatial resolution along the satellite track is increased at the expense of that along the cross-track direction. It is expected that the set of data acquired in the ExtCal mode will show a high redundancy because along each sub-satellite line IASI is sensing almost the same target area. This is even more the case, if we consider sea surface and clear sky conditions alone.

The complete set of data cover the whole year 2016, one daily orbit for each month. More details about the data set are shown in Table 1. This set of data has been considered because it simplifies the assessment of observation errors as will be discussed in Sections 2.2 and 3.2.

In addition, we have also considered a set of IASI data from January 2014 to September 2015, before the configuration of the IASI cube corner mechanism was modified by switching off