Universitat Autònoma de Barcelona

Sounder Spectral Data Compression

Francisco Joaquín García Sobrino

Universitat Autònoma de Barcelona

Departament d'Enginyeria de la Informació i de les Comunicacions

Sounder Spectral Data Compression

Submitted to Universitat Autònoma de Barcelona
in partial fulfillment of the requirements for the
degree of Doctor of Philosophy in Computer Science

by Francisco Joaquín García Sobrino

Bellaterra, June 2018

Supervisor:
Dr. Joan Serra Sagristà

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Bellaterra, June 2018

Dr. Joan Serra Sagristà

Committee:

Dr. Víctor Francisco Sánchez Silva
Dr. Alaitz Zabala Torres
Dr. Vicente González Ruiz
(substitute) Dr. Michael W. Marcellin
(substitute) Dr. Jorge González Conejero
(substitute) Dr. Jordi Portell i de Mora

Abstract

The Infrared Atmospheric Sounding Interferometer (IASI) is a Fourier Transform Spectrometer implemented on the MetOp satellite series. The instrument, which represents a significant technological step forward with respect to previous instruments, is intended to measure infrared radiation emitted from the Earth. IASI produces data with unprecedented accuracy and spectral resolution. Notably, the sounder harvests spectral information to derive temperature and moisture profiles, as well as concentrations of trace gases, essential for the understanding of weather, for climate monitoring, and for atmospheric forecasts. In addition, the instrument collects data on sea-surface temperature and land-surface emissivity in cloud free conditions.

The large spectral, spatial, and temporal resolution of the data collected by the instrument involves generating products with a considerably large size, about 16 Gigabytes per day by each of the IASI-A and IASI-B instruments currently operated. The amount of data produced by IASI demands efficient compression techniques to improve both the transmission and the storage capabilities. This thesis supplies a comprehensive analysis of IASI data compression and provides effective recommendations to produce useful reconstructed spectra. The study analyzes data at different processing stages. Specifically, we use data transmitted by the instrument to the reception stations (IASI L0 products) and end-user data disseminated to the Numerical Weather Prediction (NWP) centres and the scientific community (IASI L1C products).

In order to better understand the nature of the data collected by the instrument, we analyze the information statistics and the compression performance of several coding strategies and techniques on IASI L0 data, which has been processed only on-board the satellite and is the input of the on-ground processing chain. The order-0 entropy and the order-1, order-2, and order-3 context-based entropies are analyzed in several IASI L0 products provided by CNES and EUMETSAT. This study reveals that the size of the data could be considerably reduced by exploiting the order-0 entropy. More significant gains could be achieved if contextual models were used. We also investigate the performance of several state-of-the-art lossless compression techniques. Experimental results suggest that a compression ratio of 2.6:1 can be achieved, which involves that more data could be transmitted at the original transmission rate or, alternatively, the transmission rate of the instrument could be further decreased.

A comprehensive study of IASI L1C data compression is performed as well. Several state-of-the-art spectral transforms and compression techniques are evaluated on IASI L1C spectra. Extensive

experiments, which embrace lossless, near-lossless, and lossy compression, are carried out over a wide range of IASI-A and IASI-B orbits. We also evaluate the computational cost of several spectral transforms, which may be critical in a real scenario when the number of spectral components is large (this is the case of IASI). For lossless compression, compression ratios over 2.5:1 can be achieved. For near-lossless and lossy compression, higher compression ratios can be achieved, while producing useful reconstructed spectra.

Even though near-lossless and lossy compression produce higher compression ratios compared to lossless compression, the usefulness of the reconstructed spectra may be compromised because some information is removed during the compression stage. Therefore, we investigate the impact of near-lossless and lossy compression on end-user applications. Specifically, the impact of compression on IASI L1C data is evaluated when statistical retrieval algorithms are later used to retrieve physical information. Experimental results reveal that the reconstructed spectra can enable competitive retrieval performance, improving the results achieved for the uncompressed data, even at high compression ratios.

We extend the previous study to a real scenario, where spectra from different disjoint orbits are used in the retrieval stage. Experimental results suggest that the benefits produced by compression are still significant. The retrieval performance is improved when moderate to high compression ratios (between 32:1 and 160:1) are employed. When the compression ratio is further increased, the same performance as with uncompressed data can be achieved. Of course, if the compression ratio is extremely high (above 1,600:1), the retrieval performance decreases. We also investigate the origin of the benefits produced by compression. On the one hand, results illustrate that compression performs signal filtering and denoising, which benefits the retrieval methods. On the other hand, compression is an indirect way to produce spectral and spatial regularization, which helps pixel-wise statistical algorithms.

In summary, we have carried out a comprehensive study of IASI data compression. We have analyzed the nature of the data and proposed several compression schemes and recommendations to alleviate the large size of IASI products. Moreover, we have provided strategies to improve the performance of linear and nonlinear statistical retrieval algorithms when IASI L1C spectra is employed to retrieve physical variables.

Contents

Chapter 1

Introduction

Remote sensing refers to the use of aerial sensor technologies to detect and classify objects on either the land surface, the oceans or in the atmosphere by means of electromagnetic radiation. Usually, space satellites or aircrafts harvest and transmit data from different parts of the electromagnetic spectrum providing valuable information to monitor and manage several application fields.

Remote sensing products have shown to be useful in several application areas including environmental measurements, weather, biology, defence, and mineralogy. The information derived from the data collected allows to analyze different scenarios, providing support for defining relevant actions in a given field. Information supplied by remote sensing products is essential to evaluate the impact of a particular phenomenon and has become the basis for decision making.

In recent decades, development of remote sensing technologies have allowed to improve thermal infrared remote sensing measurements, leading to a new generation of Earth observation instruments, which are able to yield improved products in terms of spatial, spectral, and temporal resolution, as well as an enhanced accuracy of the measurements. The Infrared Atmospheric Sounding Interferometer (IASI) belongs to the new generation of remote sensing instruments developed in recent years.

1.1 IASI Instrument

IASI is a calibrated Fourier Transform Spectrometer based on a Michelson Interferometer associated with an integrated imaging system. The instrument is intended to collect infrared radiation emitted by the Earth's atmosphere.

IASI is the most advanced instrument implemented on the MetOp satellite series, which are operated in the framework of the EUMETSAT Polar System (EPS) [1]. The main objective of the program is to collect and disseminate atmospheric information to the Numerical Weather Prediction (NWP) centres and the scientific community for the purpose of supporting global climate monitoring [2].

The EPS comprises both ground and space components. The space component consists of the MetOp-A, MetOp-B, and MetOp-C satellites. MetOp-A and MetOp-B were launched in October 2006 and September 2012, respectively. MetOp-C is scheduled to be launched in October 2018 [3]. The ground component consists of several reception stations responsible for collecting, processing and distributing the data captured by the IASI instruments [4, 5].

IASI implements an across track scanning system over a horizontal swath width of, approximately, 2,200 km. Each scanned line produces 30 Fields of Regard (FOR), which corresponds to 30 mirror positions. Each FOR consists of four Instantaneous Fields of View (IFOV) arranged as a 2×2 circular pixel matrix. The instrument produces 14 orbits everyday, which ensures 99% global coverage of the Earth two times per day.

The instrument yields data with high spectral resolution between 645 cm^{-1} and 2,760 cm^{-1} wavelength leading to over 8,000 components. Products obtained from IASI data represent a significant improvement in both the quality and the accuracy of the measurements used in meteorological models. Notably, the instrument provides infrared soundings of unprecedented accuracy to derive humidity and atmospheric temperature profiles, as well as some of the chemical components of the atmosphere [6, 7].

Data collected by IASI are processed both on-board the satellite and on-ground throughout several processing stages, which yield products at different levels [8, 9,

10, 11]. First, a significant part of the IASI data processing is performed on-board the satellite. The data rate is reduced from 45 Mb/s to 1.5 Mb/s to meet the data rate allocated to the instrument. The on-board processing chain generates IASI L0 products, which are the inputs of the on-ground processing chain.

The on-ground processing chain leads from IASI L0 to IASI L2 products delivering L1A, L1B, and L1C intermediate products. Each processing stage carries out the following processes:

- IASI L1A spectra are derived from L0 after decoding, spectral calibration, co-registration with the Advanced Very High Resolution Radiometer (AVHRR), geolocation, and dating.
- IASI L1B spectra are derived from L1A after performing spectral resampling.
- IASI L1C products are obtained from L1B after apodisation.
- IASI L2 products involve derivation of physical parameters from radiance measurements.

Only IASI L1C and IASI L2 products are considered useful for end users.

1.2 IASI Data Compression and Dissemination

IASI L1C and IASI L2 are the main products disseminated through EUMETCast [12], the primary dissemination mechanism of EUMSETSAT to distribute IASI data. Alternatively, other institutions may request and disseminate products from EUMETSAT.

The large size of the products, composed of thousands of spectra with over 8,000 components, demands efficient compression strategies to improve the capabilities of storage and transmission of such amount of information. In the IASI community, Principal Component Compression (PCC) [13] is the common strategy used for compression of IASI products. PCC is a lossy compression technique, which yields a truncated principal components representation of the original data and achieves data

compaction by dimensionality reduction [14]. PCC is able to attain a compression ratio of, approximately, 50:1 on IASI products.

The main interest in the use of PCC focuses on the ability to remove certain level of the noise from the data. However, PCC presents some drawbacks. A major concern refers to the use of a training set of spectra to produce the principal components. If the spectra used in the training stage are not representative enough or present deficiencies, some features may be considered noise and, consequently, discarded in the compression process. Moreover, some features may not be properly retained in the recovered spectra when the original signal is weak [15, 16].

Despite the large size of the IASI products, IASI data compression has been scarcely studied beyond PCC. Although IASI is a sounder, coding techniques intended for images can achieve competitive performance in the compression of IASI products.

This thesis aims at providing a comprehensive study of IASI data compression. Advantages and disadvantages of state-of-the-art compression techniques and spectral transforms are evaluated. A wide range of coding schemes for lossless, near-lossless, and lossy compression are proposed to achieve both competitive compression performance and accurate reconstructed spectra. In addition, the impact of lossy compression on end-user applications is investigated. Specifically, the performance of statistical retrieval algorithms is assessed when reconstructed IASI spectra is employed to predict atmospheric parameters.

1.3 Contributions and Thesis Organization

The contributions of this thesis consist of a compendium of publications addressed to the compression of IASI data and published in the relevant journals on this area:

- J. García-Sobrino, I. Blanes, M. Albinet, R. Camarero, and J. Serra-Sagristà, **"Proposal for Infrared Atmospheric Sounding Interferometer on-board data compression,"**, SPIE Journal of Applied Remote Sensing, vol. 9, num. 1, pp. 097498, May 2015. [17]

- J. García-Sobrino, J. Serra-Sagristà, and J. Bartrina-Rapesta, **"Hyperspectral**

IASI L1C Data Compression," MDPI Sensors, vol. 17, num. 6, pp. 1404, June 2017. [18]

- J. García-Sobrino, J. Serra-Sagristà, V. Laparra, X. Calbet, and G. Camps-Valls, **"Statistical Atmospheric Parameter Retrieval Largely Benefits From Spatial–Spectral Image Compression,"** IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 9, pp. 2213-2224, April 2017. [19]

- J. García-Sobrino, V. Laparra, J. Serra-Sagristà, X. Calbet, and G. Camps-Valls, **"Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI data,"** MDPI Remote Sensing, vol. xx, no. xx, pp. xx, 2018. Under Review. [20]

In Chapter 2, the first publication [17] analyzes the IASI L0 products. We study the information statistics of the data, where both the spectral and the spatial joint information are investigated. A theoretical study of order-0 entropy, order-1 conditional entropy, order-2 conditional entropy, and order-3 conditional entropy is performed on IASI L0 products. This initial analysis reveals that a simple variable-length code might save at least 1.6 bits per pixel per component (bpppc). Saving increases when contextual models are considered. The best results are produced by order-3 models, which might be able to save at least 6.7 bpppc. The performance of different state-of-the-art lossless compression techniques is also investigated on IASI L0 data, namely, the CCSDS-123 [21], the JPEG-LS [22], and the JPEG 2000 [23] standards and the M-CALIC [24] coding technique. To exploit the high spectral redundancy present in the IASI spectrum, the Pairwise Orthogonal Transform (POT) [25] is paired along with JPEG-LS and JPEG-2000. Experimental results suggest that a compression ratio over 2.6:1 can be achieved.

The second publication [18], discussed in Chapter 3, studies the compression performance of several coding strategies and techniques on IASI L1C products. A comprehensive analysis of lossless, near-lossless, and lossy compression is reported for 96 IASI L1C orbits acquired over a full year by the IASI-A and IASI-B instruments. Several state-of-the-art coding techniques and spectral transforms are analyzed. The

computational cost of the selected spectral transforms, which may be critical in scenarios like the IASI near-real time dissemination, is also assessed. Experimental results illustrate that, for lossless compression, a compression ratio over 2.5:1 can be achieved. For near-lossless and lossy compression, the reconstructed spectra is compared to the reconstructed spectra produced by PCC, which is the common lossy compression strategy used in the IASI community. Results reveal that the proposed compression schemes produce competitive performance compared to PCC.

In Chapter 4, the third publication [19] investigates the impact of near-lossless and lossy compression of IASI L1C products on end-user applications. Specifically, the performance of statistical retrieval algorithms is studied when reconstructed spectra is employed to retrieve physical information. We propose several compression schemes according to the characteristics of the IASI data. Experimental results illustrate that both near-lossless and lossy compression benefit the retrieval performance even at high compression ratios. The improvements are especially significant for lossy compression.

In the fourth publication [20], discussed in Chapter 5, the study reported in the third publication [19] is extended. Experiments are carried out in a realistic scenario, where data from different IASI orbits are used to define the training, the validation, and the testing subsets of spectra used in the retrieval stage. Moreover, additional compression schemes are proposed. Experimental results illustrate that the retrieval algorithms benefit from compression also in a realistic scenario. In this contribution, the origin of the improvement produced by the compression stage is also analyzed. On the one hand, we show that certain level of noise is removed from the data during the compression stage, which benefits the retrievals. On the other hand, compression is a simple way to exploit spectral and spatial relations between neighboring samples, which improves the accuracy of the retrieval results.

Chapter 6 describes the IASI orbits used in the experiments and summarizes the experimental results. Chapter 7 draws some conclusions and provides some insights of the future work.

Chapter 2

Proposal for Infrared Atmospheric Sounding Interferometer On-Board Data Compression

Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Proposal for Infrared Atmospheric Sounding Interferometer on-board data compression

Joaquín García-Sobrino
Ian Blanes
Mathieu Albinet
Roberto Camarero
Joan Serra-Sagristà

SPIE.

Proposal for Infrared Atmospheric Sounding Interferometer on-board data compression

Joaquín García-Sobrino,[a,*] Ian Blanes,[a] Mathieu Albinet,[b] Roberto Camarero,[b] and Joan Serra-Sagristà[a]

[a]Universitat Autònoma de Barcelona, Department of Information and Communications Engineering, Edifici Q, UAB, 08193 Cerdanyola del Vallès (Barcelona), Spain

[b]Centre National D'études Spatiales–CNES, 18 av. Edouard Belin, 31401 Toulouse Cedex 9, France

Abstract. The Infrared Atmospheric Sounding Interferometer (IASI) system provides infrared soundings of moisture and temperature profiles, as well as soundings of chemical components. These measurements play a key role in atmospheric chemistry, global change, and climate monitoring. The instrument, developed by a cooperating agreement between European Organisation for the Exploitation of Meteorological Satellites and Centre National d'Études Spatiales, is implemented on the Metop satellite series. The instrument data production rate is 45 Mb/s while the transmission rate allocated to IASI measurements is 1.5 Mb/s. It is thus necessary to implement a significant part of the IASI data processing on-board the instrument. We investigate the information statistics of IASI L0 data once the on-board processing chain is finished. We analyze order-0 entropy, and order-1, order-2 and order-3 conditional entropies, where conditional entropies assess both the spectral and the spatial joint information. According to the simple order-0 entropy, at least one bit per sample could be spared if a variable-length code was employed. We also investigate the actual performance of different lossless compression techniques on IASI L0 data. The CCSDS-123, JPEG-LS, and JPEG2000 standards, as well as M-CALIC coding technique are evaluated. Experimental results reveal that IASI Level 0 data can be coded by a compression ratio above 2.6:1. [DOI: 10.1117/1.JRS.9.097498]

Keywords: remote sensing compression; lossless data coding; predictive coding; entropy analysis; multi- and hyperspectral imagery; CCSDS 123.

Paper 14769SS received Dec. 12, 2014; accepted for publication Apr. 28, 2015; published online May 20, 2015.

1 Introduction

IASI is an Infrared Atmospheric Sounder Interferometer that provides atmospheric spectra to the scientific and meteorological communities. It is a key element of the payload on the Metop series of European meteorological polar-orbiting satellites. IASI represents a significant scientific and technological step forward that provides meteorologists with atmospheric emission spectra to derive temperature and humidity profiles with a vertical resolution of 1 km and an accuracy of within 1 K for temperature and 10% precision for humidity measurements.[1]

IASI instrument scans the Earth's surface in step and stare mode, harvesting atmospheric soundings on both sides of the vertical line, producing fields of view (FOVs). Each FOV consists of four full spectra, where each spectrum represents a single circular pixel of the scanned image;[2,3] all pixels are captured in the same way. Complete coverage of the Earth orbit is carried out by scanning lines at regular intervals. Each scanned line consists of 30 FOVs. Figure 1(a), reproduced from Ref. 4, shows the modus operandi of the IASI instrument. Each scanned pixel represents 12 km of the Earth's surface.[3] Pixels are spaced by approximately 12 km so that each FOV represents nearly 50 km of the Earth's surface at the nadir position. Figure 1(b) shows the

*Address all correspondence to: Joaquín García-Sobrino, E-mail: joaquin.garcia.sobrino@deic.uab.cat

1931-3195/2015/$25.00

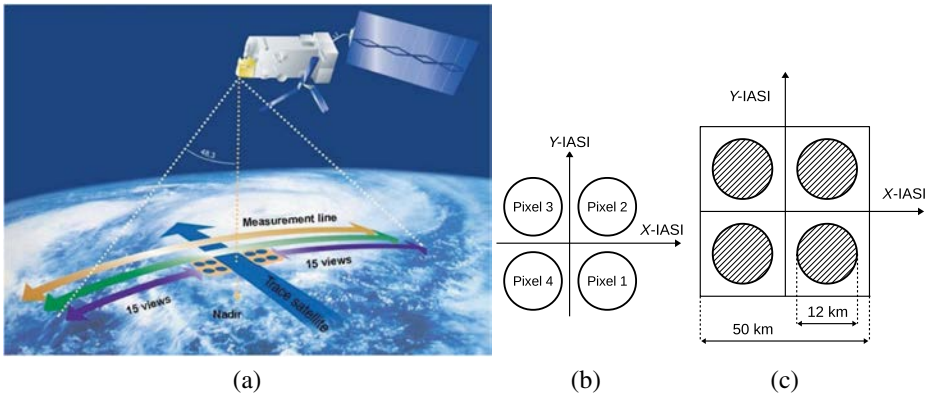


Fig. 1 Details of the IASI instrument. (a) Modus operandi of IASI instrument, (b) FOV pixel numbering, and (c) FOV size and pixel size.

structure and the pixel numbering of an FOV,[5] while Fig. 1(c) reports the size of the Earth's surface scanned by an FOV.

The IASI instrument provides infrared spectra with high resolution between wavelengths 645 and 2760 cm^{-1}. The spectral resolution required from IASI after apodisation is equal to 0.5, i.e., 0.25 cm^{-1} before apodisation. This represents a challenge because today's technology does not produce detectors having the required performance. It is necessary to implement three different detectors per pixel—providing three different sub-bands (B1, B2, and B3)—to supply the required spectral range.[2] The nominal limits of each sub-band are reported in Fig. 2(a). However, this approach produces high noise levels in sub-band edges, as shown in Fig. 2(b), reproduced from Ref. 2, which leads to a band merging process to reduce the increased noise. This way, edges on sub-bands B1/B2 and B2/B3 are combined, resulting in a spectra less noisy than the original, but keeping the required spectral resolution (645 to 2760 cm^{-1}).

While the IASI data production rate is 45 Mb/s, the transmission rate allocated to these instrument measurements is 1.5 Mb/s.[6] Accordingly, it is necessary to implement a significant part of the IASI data processing on-board the instrument. An inverse Fourier transform and a radiometric calibration are performed on-board to reduce the size of the data to be transmitted. The spectral data are then encoded before being sent to the reception stations. This part of the IASI processing chain is known as Level 0.

The sensor captures 8359 spectral bands for each pixel, producing 8359 samples. These samples are split into 522 ranges of 16 bands each one and 1 range of 7 bands (the last one). Each sample is quantized then coded as a natural number using a fixed number of bits (from 6 to 10 bits, depending on the range) before its transmission to ground. The number of bits for

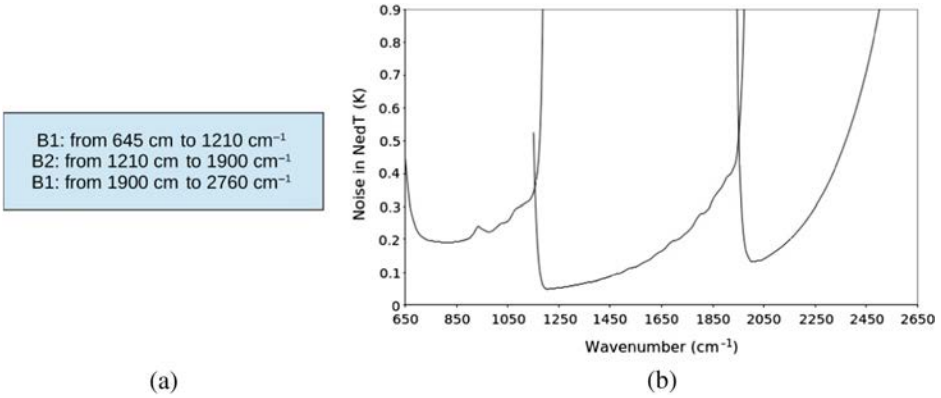


Fig. 2 Sub-bands in IASI spectral range. (a) Nominal limits of sub-bands and (b) level noise in sub-bands.

each sample was chosen to respect the NedT (Noise Equivalent Delta Temperature) degradation specification. In the end, spectral data are transmitted in the form of a bit-stream.

This method provides simplicity and a fixed bit-rate output, but presents risks of overflow and its performance is suboptimal. This choice was implemented in the instrument because most of the processing power in the dedicated electronic unit was devoted to perform the Fourier transform.

The different processes performed on-board the satellite are described in Tournier et al.,[6] where mathematical and physical contents of the IASI Level 0 processing algorithms are detailed. Different algorithms are run in real-time in order to complete the on-board processing chain.[2]

Currently, the IASI instrument is able to transmit an average of 8.2 bits per spectral sample without loss of useful information.[6] However, information statistics of IASI L0 data once the on-board processing chain is finished suggests that a smaller number of bits per spectral sample might be needed in practice.

IASI L0 data are not considered as end-user data, therefore, they are not publicly available. To the best of our knowledge, no compression study beyond those related to the design of the instrument has been carried out.

This paper investigates the modification of the Level 0 processing chain such that improved data transmission rate could be provided. We analyze the information statistics of IASI L0 data before being sent to the ground. We study order-0 entropy, and order-1, order-2, and order-3 conditional entropies, where conditional entropies assess both the spectral and the spatial joint information. According to the studied contextual models, theoretical results suggest that between one and seven bits per sample might be spared if a variable-length code was employed.

The rest of the paper is organized as follows: Sec. 2 presents the theoretical results about different order-n entropies for two sets of IASI L0 images. Section 3 presents the experimental results of IASI L0 data compression using different techniques: CCSDS-123, JPEG-LS, and JPEG2000 standards, and M-context-based adaptive lossless image coding (M-CALIC) coding technique. Section 4 concludes the paper.

2 Entropy Analysis on IASI L0 Data

The information statistics analysis is performed on a corpus of 16 three-dimensional images consisting of a set of eight images from European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) and a set of eight images from Centre National d'Études Spatiales (CNES). Since raw data produced by the instrument are not available on ground, all the analyzed images are mounted using spectral data once the on-board processing chain is finished. All volumes are mounted using data sent from satellite to reception station before any further processing on ground. Table 1 provides the technical names of the employed IASI L0 products along with their identifier (id.) used in the remainder of the paper.

All IASI L0 images have 8359 bands, 60 columns, and a different number of rows. Images from EUMETSAT have 172, 158, 110, 180, 196, 154, 110, and 132 rows, respectively. Images from CNES have a larger number of rows than images from EUMETSAT, with, respectively, 1484, 1484, 1528, 1482, 1528, 1528, 1528, and 1528 rows; this number of rows approximates a full Earth orbit (1530 rows).

As explained, the on-board processing chain encodes each sample using a fixed number of bits, from 6 to 10 bits depending on the range, i.e., all samples belonging to a given range—and, thus, all samples from the same spectral component—are encoded using the same number of bits.

The entropy analysis presented in this paper comprises order-0 entropy, order-1 conditional entropy, order-2 conditional entropy, and order-3 conditional entropy. Order-0 entropy is computed as described in Eq. (1), where $p(x)$ represents the probability of occurrence of symbol x:

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x). \tag{1}$$

The conditional entropy assesses the spatial and/or spectral joint information. Equation (2) provides the definition for order-1 conditional entropy. Order-2 conditional entropy and order-3 conditional entropy are defined accordingly.

Table 1 Analyzed images.

Id.	IASI L0 product
EUMETSAT 1	IASI_HRP_00_M01_20130813184200Z_20130813185335Z_N_O_20130813184205Z
EUMETSAT 2	IASI_HRP_00_M01_20130813202226Z_20130813203309Z_N_O_20130813202229Z
EUMETSAT 3	IASI_HRP_00_M01_20130814070242Z_20130814071311Z_N_O_20130814070245Z
EUMETSAT 4	IASI_HRP_00_M01_20130814084228Z_20130814085440Z_N_O_20130814084231Z
EUMETSAT 5	IASI_HRP_00_M01_20130814101020Z_20130814102504Z_N_O_20130814101026Z
EUMETSAT 6	IASI_HRP_00_M01_20130814120024Z_20130814121046Z_N_O_20130814120030Z
EUMETSAT 7	IASI_HRP_00_M02_20130813192744Z_20130813193516Z_N_O_20130813192748Z
EUMETSAT 8	IASI_HRP_00_M02_20130814092926Z_20130814093921Z_N_O_20130814092930Z
CNES 1	IASI_xxx_00_M02_20091007112100Z_20091007130000Z_N_O_20091007125723Z
CNES 2	IASI_xxx_00_M02_20091007130000Z_20091007143900Z_N_O_20091007143543Z
CNES 3	IASI_xxx_00_M02_20091007143900Z_20091007162100Z_N_O_20091007161509Z
CNES 4	IASI_xxx_00_M02_20091017125400Z_20091017143300Z_N_O_20091017142943Z
CNES 5	IASI_xxx_00_M02_20100319050300Z_20100319064500Z_N_O_20100319064052Z
CNES 6	IASI_xxx_00_M02_20120718075700Z_20120718093900Z_N_O_20120718084400Z
CNES 7	IASI_xxx_00_M02_20130116133300Z_20130116151500Z_N_O_20130116142103Z
CNES 8	IASI_xxx_00_M02_20130916080300Z_20130916094500Z_N_O_20130916093859Z

$$H(X|Y) = -\sum_{x\in X, y\in Y} p(x,y)\log_2 \frac{p(x,y)}{p(y)}. \tag{2}$$

Figure 3 illustrates the different contexts analyzed in the study. Order-0 entropy computes the statistics of the current pixel [Fig. 3(a)]. Six different contextual models are taken into consideration for order-1 conditional entropy: the left pixel [Fig. 3(b)], the top-left pixel [Fig. 3(c)], the top pixel [Fig. 3(d)], the top-right pixel [Fig. 3(e)], the colocated pixel in a previous component [Fig. 3(f)], and the colocated pixel in the next to previous component [Fig. 3(g)] are used as contexts. In the case of order-2 conditional entropy, two different contexts are considered, the left and the top pixels [Fig. 3(h)] and the colocated pixels in the previous and in the next to previous component [Fig. 3(g)]. Three different order-3 conditional entropies are also investigated: the left, the top, and the colocated pixel in a previous component [Fig. 3(f)]; the left and the colocated pixels in the two previous components [Fig. 3(k)]; and the top and the colocated pixels in the two previous components [Fig. 3(l)].

To get consistent results, only pixels having all the required contexts shall be analyzed. In this way, the first row and the first and the last columns in a component are not considered for entropy computation as their pixels lack some neighboring context. Regarding the spectral dependencies, only the last 14 components in each range are analyzed, again because of some lacking neighboring contexts for the first two components.

Figure 4 provides an example of the entropy distribution for a single range (range 11) for two of the analyzed volumes (EUMETSAT 1 and CNES 1). Range 11 comprises bands from 161 to 176 (both included), with all components coded with 9 bits per pixel (bpp). Columns in the graph correspond to different order-n entropies (contexts). Marks in columns represent the entropy of a single component. As explained, only 14 marks per column are plotted.

Because of space constraints, only plots corresponding to two images are reported. This same analysis is performed for all ranges (523) for each of the 16 images (EUMETSAT and CNES).

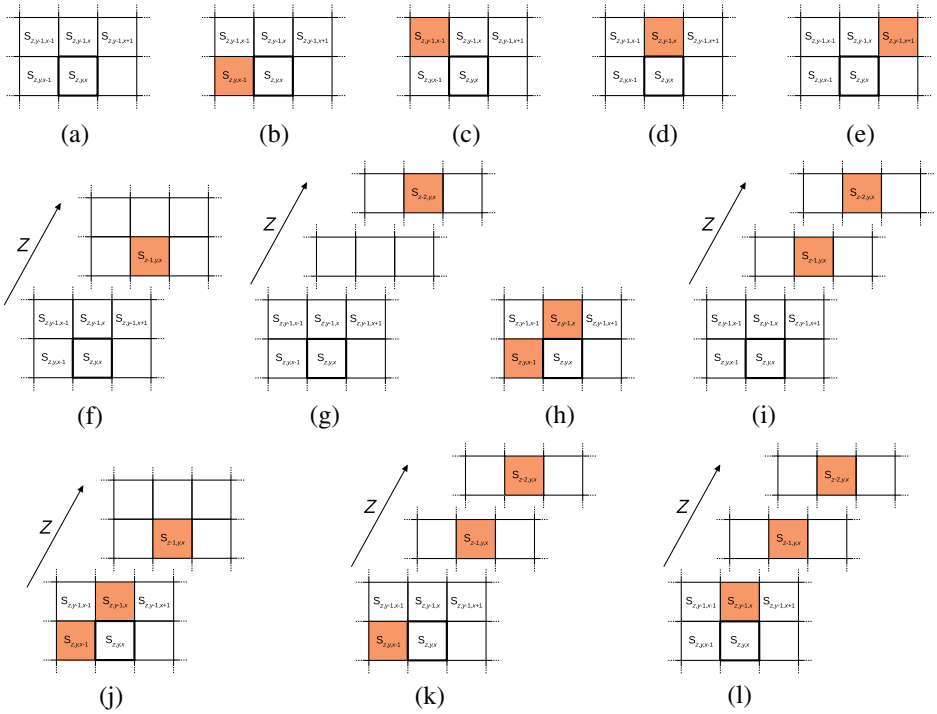


Fig. 3 Defined contexts in the context-based entropy analysis. (a) order-0, (b) order-1 left, (c) order-1 top-left, (d) order-1 top, (e) order-1 top-right, (f) order-1 one previous, (g) order-1 next to previous, (h) order-2 left and top, (i) order-2 two previous, (j) order-3 left and top and one previous, (k) order-3 left and two previous, and (l) order-3 top and two previous

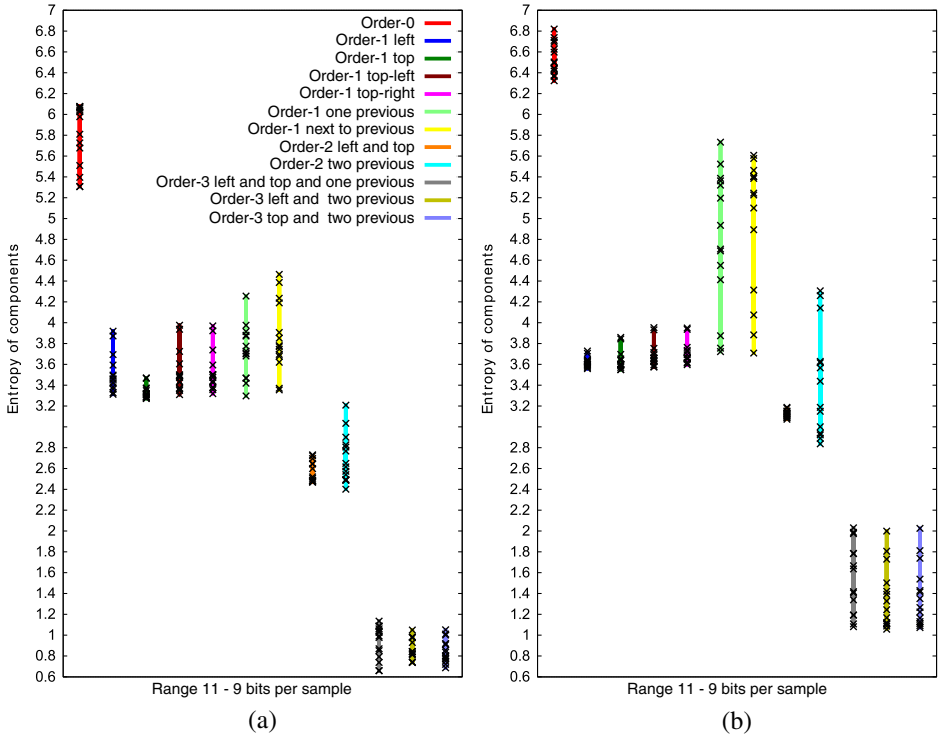


Fig. 4 Entropy distribution of range 11 in bits per pixel (bpp). It is originally encoded with 9 bpp. (a) EUMETSAT 1 and (b) CNES 1.

The results indicate that the entropy distribution follows a very similar pattern in each range in all images. This behavior is common to images from both EUMETSAT and CNES corpus, the latter having a much larger number of rows. Even though the particular entropy values may differ, the entropy distribution of the different studied contexts is very similar, suggesting that the performance of the selected contextual model is similar for all IASI L0 data.

Tables 2 and 3 report the results of the entropy analysis in the studied images. The different entropy columns provide the average entropy for all components in an image. Entropy is computed on a per component basis. The average entropy is calculated to determine the theoretical average number of bits per pixel needed to encode the whole image.

Currently, the on-board processing chain implemented in the instrument allows the encoding of the IASI L0 data using, on average, 8.2 bpp before being transmitted to reception stations. The theoretical order-0 entropy analysis illustrates that the images from EUMETSAT and CNES could be coded using, on average, 5.96 and 6.35 bpp, respectively. This implies that 2.23 and 1.84 bpp might be spared in images from EUMETSAT and CNES, respectively, if a simple variable-length code was employed. In the worst case, for EUMETSAT images, which have approximately 150 rows, at least 1.73 bpp could be spared; for CNES images, which have approximately 1500 rows, at least 1.62 bpp could be saved.

These gains could be higher if contextual models were used. When one pixel belonging to the same spatial component acts as the context, on average, 4.45 bpp (images from EUMETSAT) and 3.98 bpp (images from CNES) might be saved. If spectral contexts, i.e., the colocated pixel in a previous component, were used, it would be possible to spare, on average, 5.05 bpp in the case of images from EUMETSAT and 4.90 bpp for CNES. In the worst case, considering any order-1 context, 4.13 and 3.78 bpp at least might be spared in images from EUMETSAT and CNES, respectively.

These results are further improved when order-2 conditional contexts are used. Taking as context the colocated pixels in the previous and the next to previous component allows us to spare, on average, 6.18 bpp in images from EUMETSAT and 5.59 bpp in images from CNES. The savings are greater when order-2 left and top pixel context are used. In this case, on average, 6.42 and 5.63 bpp could be saved in images from EUMETSAT and CNES, respectively. In the worst case, using any order-2 conditional context, 5.96 bpp might be saved in images from EUMETSAT and 5.57 bpp might be spared for CNES.

Nevertheless, the best results are obtained when order-3 conditional contexts are used. When one spatial neighboring pixel (the left or the top) and two spectral neighboring pixels (the colocated pixels in the previous and in the next to previous components) are used as context, it would be possible to save, on average, 7.36 and 6.76 bpp in images from EUMETSAT and CNES, respectively. When two spatial neighboring pixels (the left and the top) and one spectral neighboring pixel (the colocated pixel in the previous component) are used as context, on average, 7.46 bpp (images from EUMETSAT) and 6.94 bpp (images from CNES) could be saved.

In general, using any contextual model improves the entropy results. Figure 5(a) compares order-0, order-1 conditional entropy, order-2 conditional entropy, and order-3 conditional entropy in the image EUMETSAT 1, along with the original bit-depth. Figure 5(b) shows the same comparison for image CNES 1. To ease the visual interpretation, only one curve for order-1, order-2 and order-3 entropy is plotted, namely the one providing the best performance on average, i.e., one previous component context for order-1, left and top context for order-2, and left and top and one previous component context for order-3. These plots show the average entropy per range, i.e., the X-axis represents the 523 ranges in which a spectrum is divided, and the Y-axis represents the average entropy per range. Due to space constraints only two graphics are plotted. The results in the other images are almost identical to those shown here.

One can see how order-0 entropy already outperforms the current bit-depth in IASI L0 data. Comparing the different entropies, as expected, order-3 conditional entropy yields the best results in terms of bit rate in all ranges, followed by order-2, then order-1, and finally order-0. We also note that, for ranges between 398 and 431, order-0 is very similar to order-1, order-2, and order-3 conditional entropies.

Although, in general, order-2 left and top pixel context outperforms order-1 one previous component context in most ranges, this order-1 does yield lower entropy for ranges between

Table 2 Context-based entropy analysis of IASI L0 data from EUMETSAT. Reported results in bpp (lower is better).

IASI L0 product	Order-0	Order-1 left	Order-1 top-left	Order-1 top	Order-1 top-right	Order-1 One Previous	Order-1 next to previous	Order-2 left and top	Order-2 two previous	Order-3 left and top and one previous	Order-3 left and two previous	Order-3 top and two previous
EUMETSAT 1	6.092	3.709	3.851	3.609	3.822	3.176	3.337	1.836	2.042	0.737	0.819	0.826
EUMETSAT 2	5.848	3.671	3.829	3.642	3.848	3.100	3.256	1.827	2.083	0.779	0.883	0.875
EUMETSAT 3	6.469	3.710	3.793	3.625	3.834	3.165	3.319	1.391	1.747	0.495	0.568	0.570
EUMETSAT 4	6.364	3.831	3.976	3.688	3.897	3.188	3.357	1.750	2.025	0.687	0.770	0.782
EUMETSAT 5	5.363	3.843	3.927	3.759	3.966	2.875	2.998	2.032	2.232	0.964	1.100	1.110
EUMETSAT 6	5.784	3.935	4.062	3.796	3.976	2.918	3.051	1.826	2.073	0.753	0.872	0.892
EUMETSAT 7	5.550	3.271	3.461	3.289	3.458	3.034	3.176	1.902	2.033	0.861	0.910	0.933
EUMETSAT 8	6.270	3.652	3.836	3.655	3.766	3.091	3.248	1.655	1.929	0.640	0.716	0.725
Average	5.967	3.702	3.841	3.632	3.820	3.068	3.217	1.777	2.020	0.739	0.829	0.839

Table 3 Context-based entropy analysis of IASI L0 data from CNES. Reported results in bpp (lower is better).

IASI L0 product	Order-0	Order-1 left	Order-1 top-left	Order-1 top	Order-1 top-right	Order-1 one previous	Order-1 next to previous	Order-2 left and top	Order-2 two previous	Order-3 left and top and one previous	Order-3 left and two previous	Order-3 top and two previous
CNES 1	6.297	4.051	4.211	3.986	4.218	3.213	3.395	2.598	2.627	1.304	1.462	1.487
CNES 2	6.303	4.217	4.382	4.154	4.386	3.169	3.340	2.570	2.590	1.227	1.409	1.430
CNES 3	6.354	4.144	4.297	4.101	4.322	3.188	3.360	2.577	2.592	1.241	1.416	1.434
CNES 4	6.335	4.153	4.297	4.104	4.332	3.210	3.387	2.557	2.607	1.243	1.415	1.435
CNES 5	6.282	4.064	4.219	4.014	4.230	3.172	3.350	2.548	2.590	1.292	1.458	1.477
CNES 6	6.573	4.217	4.357	4.164	4.407	3.255	3.442	2.539	2.603	1.193	1.364	1.386
CNES 7	6.315	4.057	4.196	3.987	4.220	3.206	3.384	2.600	2.609	1.290	1.455	1.478
CNES 8	6.376	4.256	4.408	4.196	4.420	3.215	3.396	2.574	2.616	1.225	1.403	1.425
Average	6.354	4.144	4.295	4.088	4.316	3.203	3.381	2.570	2.604	1.251	1.422	1.444

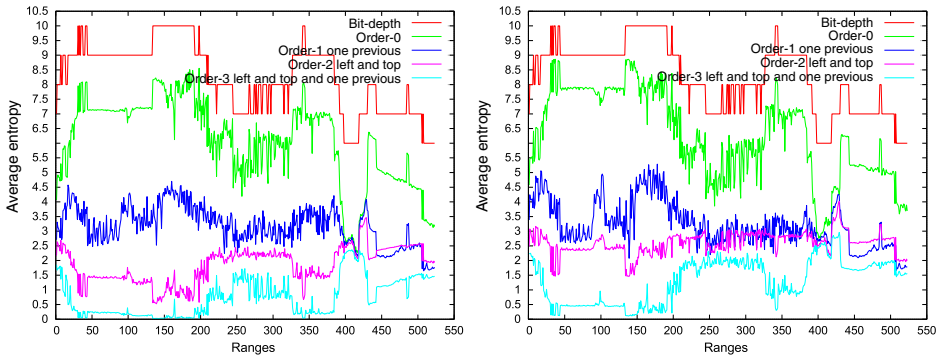


Fig. 5 Average entropy analysis by ranges. (a) EUMETSAT 1 and (b) CNES 1.

430 and 523; in these ranges, spectral contexts yield better performance than spatial contexts (notice that we are comparing order-1 spectral against order-2 spatial).

For these same ranges, from 430 to 523, order-2 spectral context also outperforms order-2 spatial context. Obviously, order-2 spectral context is better than order-1 spectral context. For all the other ranges, order-2 spatial context is better than order-2 spectral context.

3 Compression of IASI L0 Data

In this section, we investigate the actual performance of different lossless coding techniques and standards on the two sets of IASI L0 data introduced earlier. The performance of the new CCSDS-123 standard, the M-CALIC technique, the well-known JPEG-LS, and JPEG2000 standards for IASI L0 lossless data compression is tested.

The recently approved CCSDS-123 standard[7,8] provides a competitive performance with a low computational cost for a large variety of multi-, hyper-, and ultraspectral data. CCSDS-123 is a prediction-based coding technique specially designed to operate on-board satellites. CCSDS-123 operates in a two-stage mode. In the first stage, an estimate of the current pixel is performed and used to compute a prediction residual. In the second stage, this prediction residual is first mapped to an integer and then encoded with a Golomb code.

M-CALIC[9] is a lossless and near-lossless compression technique for hyperspectral images based on CALIC.[10] The algorithm uses a multiband spectral predictor along with optimized model parameters and optimization thresholds. Correlation between bands by employing the two previous bands of the current line is exploited.

JPEG-LS[11] is a two-dimensional low-complexity lossless and near-lossless compression technique that operates in a two-stage mode, modeling, and encoding. It is specially designed to be simple and fast.

JPEG2000[12,13] is an international standard developed by the Joint Photographic Experts Group (JPEG). It is intended as the successor of JPEG in many of its application areas due to its superior compression performance.

Here, JPEG-LS and JPEG2000 are tested together with pairwise orthogonal transform (POT) for spectral image coding.[14] POT is based on the application of a divide-and-conquer strategy for the KLT,[15,16] where the resulting transform is a composition of smaller KLT transforms. In a full KLT, all components are decorrelated with each other independently of how much energy they share. In contrast, the proposed POT has a structure that decorrelates parts with high shared energy while ignoring the other parts, as parts with low energies have a lower influence in the coding performance.[14] POT is used in the tests to perform a spectral transform before applying JPEG-LS or JPEG2000.

Although samples have an actual bit-depth precision from 6 to 10 bits, for storing purposes, each sample is stored in 2 bytes (16 bits) without a sign. Each volume has 8359 bands, 60 columns, and a variable number of rows. Images from EUMETSAT have between 110 and 196 rows. Images from CNES are larger, with between 1482 and 1528 rows.

As for the parameters employed in the experiments, CCSDS-123 lossless coding technique includes several user-specified configuration parameters. Here, we resort to default parameters as suggested in Ref. 8. In the last approach, once POT has been applied for spectral transform, JPEG2000 performs two levels of DWT spatial transform.

Tables 4 and 5 report the performance of CCSDS-123, M-CALIC, POT + JPEG-LS, and POT + JPEG2000 for lossless compression of IASI L0 images.

The performance of CCSDS-123, M-CALIC, and POT + JPEG2000 is very similar, while POT + JPEG-LS yields worse results. Although not reported here, applying a POT before JPEG-LS helps improve the performance of JPEG-LS by about 0.9 bpp.

On average, a compression ratio of 2.656:1 and 2.846:1 can be achieved for, respectively, EUMETSAT and CNES images using either CCSDS-123, M-CALIC, or POT + JPEG2000. Compression ratio has been computed considering the length of the original image bit-stream and the length of the compressed codestream.

Related to entropy results reported earlier in Tables 2 and 3, order-2 entropy analysis indicated that, on average, images from EUMETSAT and CNES might be coded using about 1.89 and 2.58 bpp, respectively. Here, we see that CCSDS-123, M-CALIC, and POT + JPEG2000 allows one to compress these images using, on average, 3.08 and 2.87 bpp, respectively, which is very close to the theoretical optimal compression when order-2 context is used—at least for CNES images, which have a larger row size and may better represent the average performance.

Table 4 Lossless compression performance of images from EUMETSAT. Reported results in compression ratio (higher is better) and bits per pixel (lower is better).

IASI L0 product	CCSDS-123	M-CALIC	POT+JPEG-LS	POT+JPEG2000
		Compression ratios		
EUMETSAT 1	2.617	2.693	2.145	2.653
EUMETSAT 2	2.627	2.701	2.116	2.650
EUMETSAT 3	2.483	2.537	1.902	2.489
EUMETSAT 4	2.581	2.637	2.140	2.617
EUMETSAT 5	2.805	2.862	2.308	2.801
EUMETSAT 6	2.761	2.783	2.176	2.706
EUMETSAT 7	2.618	2.746	1.964	2.611
EUMETSAT 8	2.588	2.623	2.010	2.567
Average	2.635	2.697	2.095	2.636
		Bits per pixel		
EUMETSAT 1	3.13	3.04	3.82	3.09
EUMETSAT 2	3.12	3.03	3.87	3.09
EUMETSAT 3	3.30	3.23	4.31	3.29
EUMETSAT 4	3.17	3.10	3.83	3.13
EUMETSAT 5	2.92	2.86	3.55	2.92
EUMETSAT 6	2.96	2.94	3.76	3.03
EUMETSAT 7	3.13	2.98	4.17	3.14
EUMETSAT 8	3.16	3.12	4.07	3.19
Average	3.11	3.03	3.92	3.11

Table 5 Lossless compression performance of images from CNES. Reported results in compression ratio (higher is better) and bits per pixel (lower is better).

IASI L0 product	CCSDS-123	M-CALIC	POT+JPEG-LS	POT+JPEG2000
		Compression ratios		
CNES 1	2.839	2.874	2.614	2.858
CNES 2	2.846	2.858	2.596	2.834
CNES 3	2.857	2.873	2.597	2.847
CNES 4	2.840	2.859	2.585	2.832
CNES 5	2.843	2.871	2.605	2.857
CNES 6	2.835	2.834	2.569	2.819
CNES 7	2.843	2.883	2.605	2.852
CNES 8	2.832	2.823	2.574	2.812
Average	2.841	2.859	2.593	2.838
		Bits per pixel		
CNES 1	2.88	2.85	3.13	2.86
CNES 2	2.88	2.86	3.15	2.89
CNES 3	2.87	2.85	3.15	2.88
CNES 4	2.88	2.86	3.17	2.89
CNES 5	2.88	2.85	3.14	2.87
CNES 6	2.89	2.89	3.19	2.90
CNES 7	2.88	2.84	3.14	2.87
CNES 8	2.89	2.90	3.18	2.91
Average	2.88	2.86	3.15	2.88

CCSDS-123 is especially designed to operate on-board satellites. It has low computational requirements and acceptable performance. The transmission rate for IASI Level 0 data could be further decreased, or even some more data could be transmitted at the original transmission rate if CCSDS-123 was considered as a candidate for on-board lossless compression of IASI L0 data. M-CALIC and POT+JPEG2000 also achieve good performances, but their computational requirements are higher.

A brief discussion follows on the potential of integrating CCSDS-123 in the IASI on-board processing chain. The IASI instrument has a data production rate of 45 Mbit/s at varying sample bit depths, and it has a total power consumption of 210 W.[17] The fast lossless compressor, on which the CCSDS-123 standard is based, has been implemented by Aranki et al. in XILINX VIRTEX IV and VIRTEX V FPGAs.[18–20] The authors report an implementation capable of operating at 40 Msample/s with a power consumption of 700 mW, which seems to fit the requirements of the IASI instrument.

4 Conclusion

The IASI is a key element of the payload on the Metop series of EUMETSAT. The instrument provides atmospheric spectra to derive temperature and humidity profiles to the scientific and meteorological communities. The data production rate in the instrument is 45 Mb/s while the transmission rate allocated to IASI measurements is 1.5 Mb/s. Consequently, a significant part

of the IASI data processing is performed on-board the instrument before transmitting data to reception stations.

In this paper, we analyzed the potential of entropy coding in IASI L0 data. A theoretical analysis of order-0 entropy, order-1 conditional entropy, order-2 conditional entropy, and order-3 conditional entropy, where conditional entropies assess both the spectral and the spatial joint information, was performed. This analysis yielded that at least 1.6 bpp might be saved if a simple variable-length code was employed. The best results were obtained when order-3 contextual models were used, saving at least 6.7 bpp.

We also investigated the performance of different lossless compression techniques on IASI L0 data. The recent CCSDS Recommended Standard for Lossless Multispectral & Hyperspectral image compression (CCSDS-123), the JPEG-LS, and JPEG2000 standards and the M-CALIC coding technique were evaluated. Experimental results suggested that, on average, a compression ratio of 2.6:1 could be achieved for images with approximately 150 rows. The performance was even better for larger images (approximately 1500 rows), reaching a compression ratio of 2.8:1.

Given that CCSDS-123 is a competitive and efficient coding technique whose hardware implementation is amenable to satellite probes, a particularly interesting point to consider in the future could be the potential of CCSDS-123 on IASI raw data, i.e., even before the on-board processing chain has taken place, as the on-board processing introduces loss because a quantization is carried out. It might be worth investigating whether CCSDS-123 lossless compression could be applied on original IASI raw data, thus enabling transmission of data that has not gone through any irreversible processing.

Acknowledgments

This work was supported in part by FEDER, the Spanish Government (MINECO), the Catalan Government, and the Centre National d'Études Spatiales (CNES), under grants TIN2012-38102-C03-03 and 2014SGR-691.

An earlier version of this work was presented at the ESA CNES On-Board Payload Data Compression Workshop 2014.

We thank the reviewers for their comments, which have helped to improve the quality of the paper.

References

1. G. Chalon, F. Cayla, and D. Diebel, "IASI: an advanced sounder for operational meteorology," in *52nd Proc. Int. Astronautical Congress (IAF)* (2001).
2. CNES, "Dossier de définition des algorithmes IASI," REF. IA-DF-0000-2006-CNE, http://smsc.cnes.fr/IASI/index.htm (2009).
3. CNES, "Spécification Technique de Besoin du logiciel opérationnel IASI," REF. IA-SB-2100-9462-CNE, http://smsc.cnes.fr/IASI/index.htm (2006).
4. EUMETSAT, "Metop design—IASI," 2014, http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/IASI/index.html (14 May 2015).
5. EUMETSAT, "IASI measurement and verification data," REF. IA-ID-1000-6477-AER, www.eumetsat.int (2010).
6. B. Tournier et al., "IASI level 0 and 1 processing algorithms description," in *Proc. ISTCXII Conf.* (2002).
7. Consultative Committee for Space Data Systems (CCSDS), "Lossless Multispectral & Hyperspectral Image Compression CCSDS 123.0-B-1," ser. Blue Book, CCSDS (2012), http://public.ccsds.org/publications/archive/123x0b1ec1.pdf (14 May 2015).
8. E. Augé et al., "Performance impact of parameter tuning on the CCSDS-123 lossless multi- and hyperspectral image compression standard," *J. Appl. Remote Sens.* **7**(1), 074594 (2013).
9. E. Magli, G. Olmo, and E. Quacchio, "Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC," *IEEE Geosci. Remote Sens. Lett.* **1**(1), 21–25 (2004).

10. X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Trans. Commun.* **45**(4), 437–444 (1997).
11. M. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS," *IEEE Trans. Image Process.* **9**(8), 1309–1324 (2000).
12. D. Taubman and M. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice: Image Compression Fundamentals, Standards, and Practice*, Springer, US (2002).
13. JPEG-Committee, "Standard JPEG2000," 2014, http://www.jpeg.org/jpeg2000 (14 May 2015).
14. I. Blanes and J. Serra-Sagristà, "Pairwise orthogonal transform for spectral image coding," *IEEE Trans. Geosci. Remote Sens.* **49**(3), 961–972 (2011).
15. J. Saghri, S. Schroeder, and A. Tescher, "An adaptive two-stage KLT scheme for spectral decorrelation in hyperspectral bandwidth compression," *Proc. SPIE* **7443**, 744313(2009).
16. R. Dony, "Karhunen-Loeve transform," Chapter 1 in *The Transform and Data Compression Handbook*, K. R. Rao and P. C. Yip, Eds., CRC Press (2001).
17. ESA, "Meteorological operational satellite program of Europe," https://directory.eoportal.org/web/eoportal/satellite-missions/m/metop (14 May 2015).
18. N. Aranki et al., "Fast and adaptive lossless on-board hyperspectral data compression system for space applications," *IEEE Aerospace Conf.*, IEEE (2009).
19. N. Aranki et al., "Hardware implementation of lossless adaptive and scalable hyperspectral data compression for space," *NASA/ESA Conf. Adaptive Hardware and Systems*, pp. 315–322, IEEE (2009).
20. N. Aranki et al., "FPGA provides speedy data compression for hyperspectral imagery," *Xilinx Newslett.* (2012).

Joaquín García-Sobrino received his BS and MS degrees in computer science from the Universidad de Almería, Spain, in 2009 and 2010, respectively. From 2010 to 2013, he was with the Group on Supercomputing—Algorithms of the Universidad de Almería. Since 2013, he has been with the group on Interactive Coding of Images of the Universitat Autònoma de Barcelona, where he currently undertakes a PhD. His current research interests focus on data compression.

Ian Blanes received his BS, MS, and PhD degrees in computer science from the Universitat Autònoma de Barcelona, in 2007, 2008, and 2010, respectively. Since 2003, he has been with the group on Interactive Coding of Images of the Universitat Autònoma de Barcelona, where he currently holds a postdoctoral position. He was a visiting researcher at CNES. He was the second-place finisher as best computer-science student of Spanish.

Mathieu Albinet received his engineering degree (BS) in electronics and signal processing from the École nationale suprieure d'électronique, d'électrotechnique, d'informatique, d'hydraulique et des télécommunications (ENSEEIHT) in 2003. He has been employed with the French Space Agency (CNES) since 2005, where his main area of work was on-board payload data processing. He is currently the head of the On-Board Data Handling Department.

Roberto Camarero received his engineering degree (BS) in electronics and communications from the University of Zaragoza, Spain, in 2005 and an MS degree from SUPAERO, France, in 2006. He has been employed with the French Space Agency (CNES) since 2006 in the On-Board Data Handling Department. His main area of work is on-board image coding and processing for remote sensing missions.

Joan Serra-Sagristà received his PhD in computer science from Universitat Autònoma de Barcelona (UAB), Spain, in 1999. From 1997 to 1998, he was at University of Bonn, Germany. He is currently an associate professor at UAB. His research interests focus on source coding and data transmission. He serves as an associate editor for *IEEE Transactions on Image Processing*. He has coauthored over 125 publications. He was the recipient of the Spanish Intensification Young Investigator Award in 2006.

Chapter 3

Hyperspectral IASI L1C Data Compression

sensors

MDPI

Article

Hyperspectral IASI L1C Data Compression

Joaquín García-Sobrino *, Joan Serra-Sagristà and Joan Bartrina-Rapesta

Department of Information and Communications Engineering, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain; joan.serra@uab.cat (J.S.-S.); joan.bartrina@uab.cat (J.B.-R.)

* Correspondence: joaquin.garcia.sobrino@deic.uab.cat; Tel.: +34-93-5811861

Received: 20 March 2017; Accepted: 9 June 2017; Published: 16 June 2017

Abstract: The Infrared Atmospheric Sounding Interferometer (IASI), implemented on the MetOp satellite series, represents a significant step forward in atmospheric forecast and weather understanding. The instrument provides infrared soundings of unprecedented accuracy and spectral resolution to derive humidity and atmospheric temperature profiles, as well as some of the chemical components playing a key role in climate monitoring. IASI collects rich spectral information, which results in large amounts of data (about 16 Gigabytes per day). Efficient compression techniques are requested for both transmission and storage of such huge data. This study reviews the performance of several state of the art coding standards and techniques for IASI L1C data compression. Discussion embraces lossless, near-lossless and lossy compression. Several spectral transforms, essential to achieve improved coding performance due to the high spectral redundancy inherent to IASI products, are also discussed. Illustrative results are reported for a set of 96 IASI L1C orbits acquired over a full year (4 orbits per month for each IASI-A and IASI-B from July 2013 to June 2014) . Further, this survey provides organized data and facts to assist future research and the atmospheric scientific community.

Keywords: IASI instrument; hyperspectral remote sensing; data compression; lossless; near-lossless and lossy compression

1. Introduction

The Infrared Atmospheric Sounding Interferometer (IASI) is a new generation of nadir viewing instruments for obtaining atmospheric measurements with unprecedented quality. The acquired data provides useful information for many application areas such as meteorology, climate monitoring or atmospheric chemistry. IASI data is recorded with high spectral accuracy, producing more than 8000 spectral channels that need be stored and transmitted.

The literature on IASI data is extensive. IASI products provide qualitative data for a wealth of possibilities such as numerical weather prediction (NWP) [1]; for studying the essential climate variables as cloud properties, greenhouse gases, or the hydrological cycle evaluation [2]; for predicting temperature and water vapor profiles [3,4]; or for analyzing several chemical atmospheric components (CO, CO_2, CH_4, SO_2, N_2O, HNO_3, NH_3, OCS, and CF_4) [5–11]. The high resolution of the data also allows to examine the composition of the lowest part of the atmosphere, enabling the research of specific events. For instance, Coheur et al. [12] and Turquety et al. [13] use IASI data to study the chemical composition deep in the troposphere to track the emission and movement of pollution from wildfires.

The high definition of the sensor in terms of spectral, spatial, and temporal resolution produce collected data with a considerably large size: about 16 Gigabytes per day in Binary Universal Form (BUFR) for the Representation of meteorological data format. IASI covers the spectral range between 645 and 2760 cm^{-1}. In each acquisition, 8359 spectral channels are acquired on the satellite, the IASI processing chain [14] leads to 8461 channels distributed on Earth, thus yielding a large volume of information, which is costly to manage in an operational context, i.e., for transmission and storage.

An effective way to alleviate the large amount of data produced by the instrument is to compress the IASI products according to the specific needs of the final users.

In the IASI community, Principal Component Compression (PCC) is an accepted approach for compression of IASI data. PCC is a lossy compression strategy intended to produce a truncated Principal Components (PC) representation; additionally, it allows to reduce the dimensionality of the data [15–20]. Although PCC is a mature field in the scope of IASI dimensionality-reduction and of IASI compression, other data compression techniques can also produce competitive performance for compression of IASI spectra.

In the framework of remote sensing data compression, three data coding paradigms can be adopted: lossless, near-lossless, or lossy compression. Lossless compression allows perfect reconstruction but achieves low compression ratios. Lossy compression introduces distortion in the reconstructed data while achieving high compression ratios. Near-lossless compression introduces a restricted distortion and achieves moderate compression ratios [21,22].

In some remote sensing applications, lossy compression techniques are still appropriate because several application-oriented processes do not reduce their performance even for large levels of distortion [23–27]. Furthermore, lossy compression is acknowledged in new remote sensing missions because the inherent data acquisition noise is usually larger than the distortion introduced during the coding process [28].

Near-lossless paradigm is a particular kind of lossy compression. The data quality is controlled by selecting a maximum acceptable distortion error, usually the Peak Absolute Error (PAE), between each sample of the original and the reconstructed data. Near-lossless coding is convenient when efficient data transmission or storage is intended and preserving a specific accuracy of the recovered data is requested.

This paper puts forward a comprehensive review of the compression performance of several lossless, near-lossless, and lossy coding techniques for IASI L1C products, which are the most common format for distribution of IASI data. In particular, the following six recent coding techniques and standards will be considered and their performance assessed: JPEG-LS [29], JPEG 2000 [30], M-CALIC [31], CCSDS-122.0 [32], CCSDS-123.0 [33] and HEVC [34]. Also, four spectral transform will be paired along with these coding techniques to exploit the high spectral redundancy inherent to IASI data (over 8000 channels); specifically, we will look upon Karhunen-Loève Transform (KLT) [35], Wavelet Transform (WT) [36], Pairwise Orthogonal Transform (POT) [37], and Regression Wavelet Analysis Transform (RWA) [38].

To provide a quantitative and qualitative comparison and an accurate analysis, a representative set of 96 IASI L1C products has been thoroughly selected over a full year, from July 2013 to June 2014. The same number of orbits have been selected from each instrument, 48 orbits from MetOp-A and 48 orbits from MetOp-B, one orbit per week. The selection criteria have considered different areas, seasons, and acquisition time.

This investigation reviews effective strategies and furnishes instructions and recommendations to improve the transmission and storage of IASI L1C products, which can benefit the development of prevailing and upcoming high spectral resolution infrared instruments.

The remaining part of the paper is organized as follows. Section 2 briefly introduces the space program operating IASI, details of the instrument, and the processing performed since the data are acquired by the instrument until they are disseminated to end-users. Section 3 introduces the basic scheme of a data coding system, the characteristics of the coding techniques assessed, puts forward the setting and parameters of each technique, and states the benefits of applying a spectral transform along the spectral dimension. Section 4 reports the experimental results and provides analysis and an extensive discussion. Finally, Section 5 draws some conclusions.

2. IASI Instrument

This section reviews the operational structure of the IASI instrument. First, the basic structure of the space program is outlined. Then, details of the architecture and the operating mode of the instrument are described. Finally, the main stages of the processing chain are summarized, explaining how interferograms captured by the sensor are transformed into end-user products and disseminated to data centers.

2.1. Space Program of IASI Instrument

IASI instrument is implemented on the MetOp satellite series, which is part of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Polar System (EPS). The mission is led by EUMETSAT in cooperation with American scientific agency National Oceanic and Atmospheric Administration (NOAA). Both organizations hold close collaboration through the Initial Joint Polar System (IJPS). The MetOp satellites carry a set of instruments provided by NOAA and a new generation of European instruments, IASI among them. The main objective of the program is to harvest and exchange environmental data between EUMETSAT and NOAA and disseminate the collected information to the scientific community in support of global climate monitoring and NWP [39,40], where IASI represents the principal instrument of the mission.

The EPS comprises both space and ground components. The space component consists of the MetOp satellite series (MetOp-A, MetOp-B, and MetOp-C), which are being jointly developed by EUMETSAT and the European Space Agency (ESA). While MetOp-A and MetOp-B were launched in October 2006 and September 2012, respectively, MetOp-C is planed to be launched in October 2018 [41]. The recent extension of MepOp-A useful lifetime to 2022 [42] will enable joint operation of MetOp-A, MetOp-B, and MetOp-C from 2019 onwards. The operation of the three IASI instruments on-board of the MetOp satellite series will provide over 20 years of continuous observations, which represents a statistically significant series of climate variables.

The ground component of the program consists of several reception and operating stations responsible for collecting, operating, processing and distributing the collected data. Figure 1 (courtesy of EUMETSAT) illustrates the operational mode of the EPS program. Specific and more detailed elements of the program can be found in [43,44].

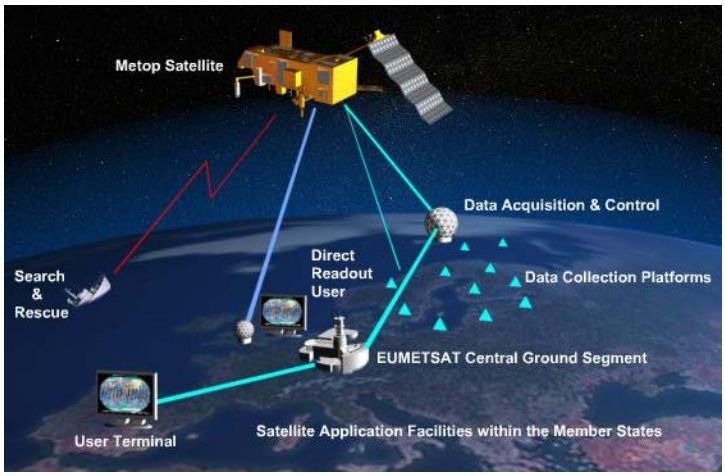


Figure 1. EPS program elements. The space component comprises the MetOp-A, MetOp-B, and MetOp-C satellites, while the ground component includes reception and operating stations.

2.2. IASI Instrument Details

IASI instrument is the result of a cooperating agreement between EUMETSAT and the French Centre National d'Études Spatiales (CNES). CNES is responsible for the IASI instrument development

and the data processing software, while EUMETSAT has the responsibility of storage, distribution, and exploitation of IASI data.

The instrument collects data over a horizontal swath width of, approximately, 2200 km. This ensures 99% global coverage of the Earth's surface performed every 12 h (2 times per day), which means 14 daily orbits in a sun-synchronous mid-morning orbit [39]. IASI observes the Earth's surface in a step and stare mode with fast movements between different look locations and stop during the acquisition of interferograms (see Table 1). The scanning process gathers atmospheric soundings on both sides of the vertical line along 30 look positions spaced by approximately 3.3 degrees. The optical axis moves from -47.85 degrees to $+47.85$ degrees with respect to the nadir position [45]. The scanning process takes 8 seconds per line and produces 30 elementary fields of regard (FOR) that correspond to 30 mirror positions. Each FOR consists of a 2×2 matrix of instantaneous fields of view (IFOV), matching to four circular pixels of the captured orbit. Each IFOV represents a full spectrum that is acquired in 3 bands: 645–1240 cm^{-1}, 1200–2040 cm^{-1}, and 1960–2760 cm^{-1} [14,46]. Each collected IFOV spreads 12 km of the Earth's surface and is separated from another IFOV by 12.5 km, so that each FOR covers, approximately, 50 km at nadir position. Figure 2 (courtesy of EUMETSAT) illustrates the modus operandi of the instrument. Figure 3 displays FOR characteristics and IFOV numbering [46,47]. Table 1 summarizes the main characteristics of IASI instrument. Further IASI technical descriptions can be found in [3,45,48,49].

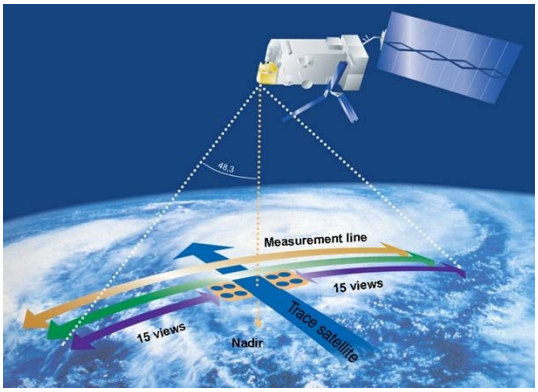


Figure 2. Modus operandi of IASI instrument. The instrument scans the Earth's surface at regular intervals producing 30 FORs per line. Each FOR consists of 4 IFOVs, each of which represents a full spectrum.

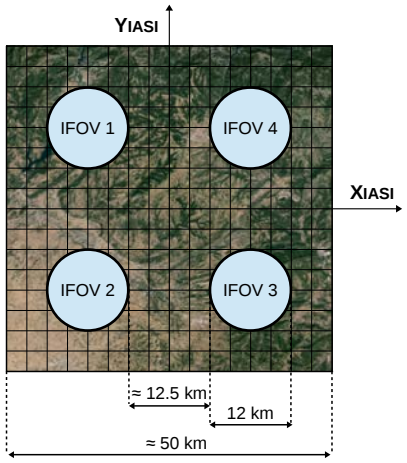


Figure 3. FOR and IFOV details. A single FOR consists of 4 IFOVs. Each IFOV spreads 12 km of the Earth's surface and is separated from its neighboring IFOVs by 12.5 km. Each FOR corresponds to, approximately, 50 km of the Earth's surface.

Table 1. Main characteristics of IASI instrument [50].

Characteristics of IASI instrument	
Orbit	Polar sun-synchronous
Time for one orbit	101 min
Global Earth coverage	2 times per day
Repeat cycle	29 days (412 orbits)
Altitude	~819 km
Scan type	Step and stare
Interferograms	30 per scan line
	151 ms per interferogram
	taken in equally spaced time intervals every 8/37 s
FOR	30 per line
	50 km (3.33°) at nadir position
	4 simultaneous IFOVs of 12 km
Full swath width	~2200 km (±48.3°)
Data production	120 spectra every 8 s
	~1,300,000 observations per day
Data acquisition rate	45 Mbps
Data transmission rate	1.5 Mbps
Spectral range	Band-1: 645–1240 cm^{-1}
	Band-2: 1200–2040 cm^{-1}
	Band-3 :1960–2760 cm^{-1}
Spectral sampling	0.25 cm^{-1} (0.5 cm^{-1} apodized)

2.3. IASI Processing Chain

The data collected by the instrument are processed through an on-board and an on-ground processing chain until they are considered end-user products. The processing chain comprises different stages that yield products at various levels. The first data on ground are IASI L0 products: these have gone only through the on-board processing chain. The on-ground processing chain leads from IASI L0 to IASI L2 yielding intermediate products such as L1A, L1B and L1C. Figure 4 illustrates the main stages in the IASI data processing chain.

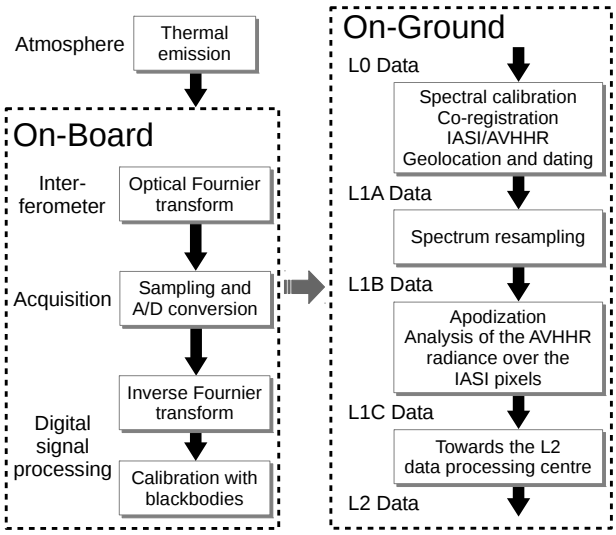


Figure 4. Main stages in the IASI data processing.

2.3.1. On-Board Processing Chain

Data collected by the instrument are processed on board the satellite to produce calibrated atmospheric spectra from raw interferograms. The IASI data production rate is 45 Megabits/s, while the transmission rate allocated to IASI measurements is 1.5 Megabits/s. In order to reach the allocated data transmission rate, it is necessary to implement a significant part of the IASI data processing on board the satellite [51]. The main objective of the on-board processing chain is to convert raw interferograms into complex spectra meeting the allocated data rate to be transmitted to the ground station. Extensive details of the different processes performed by the instrument are described in [14,51].

2.3.2. On-Ground Processing Chain

Once the data are received from the satellites, they are further processed until converted into end-user products. The on-ground processing chain comprises different stages, yielding products at different processing levels. L1A products are unapodized calibrated and geolocated spectra with corresponding Integrated Imaging Subsystem (IIS) images. L1B products are obtained from L1A after spectral resampling. L1C products are obtained from L1B after apodisation. In L1C level, the data are sampled every 0.25 cm^{-1} and the results of the analysis of Advanced Very High Resolution Radiometer (AVHRR) radiance over the IASI FOR are appended. The IASI L2 processing involves derivation of geophysical parameters from radiance measurements. This stage is performed in synergy with measurements from other instruments [50,52,53].

2.4. Data Dissemination

EUMETCast [54] is EUMETSAT's primary dissemination mechanism for the near real-time delivery of satellite data and products. IASI products are thus mostly disseminated through EUMETCast to the NWP centres and scientific community. In turn, other institutions may request and distribute IASI data from EUMETSAT, for instance, the Physical Oceanography Distributed Active Archive Center (PODAAC) [55] and the Centre for Environmental Data Analysis (CEDA) [56] disseminate IASI L1C and L2 products.

In the case of EUMETCast, the number of registered users by June 2016 [57] (last available report) was over 4500 stations, with over 1,200,000 items delivered, and distributing more than 60 Terabytes (TB) per month. In the case of CEDA, 90.82 TB of IASI data were disseminated through 636,453 accesses during the last 12 months, and 466.74 TB through 1,188,507 accesses during the last 5 years [58].

Regarding IASI data, one of the most popular collections of distributed data is IASI Regional Data Service Level 1 [59], where 500 selected original IASI Channels and 300 Principal Component Scores (PCS) are combined in products with an average file size of 10 MB. These products discard many spectral channels and/or components due to the difficulties of transmitting files of larger size.

PCC is the common strategy used in the IASI community to reduce the large size of the data [60]. This technique is able to achieve a compression ratio of, approximately, 50:1 [16,61]. PCC is a lossy approach that reduces large correlated spectra, composed by thousands of channels, into some truncated pieces of information—the PCS—[62], which represent the most of the variance observed in the data. The most part of the atmospheric information is typically contained within the first few hundred of PCS, thus the most important information present in the spectrum can be preserved by retaining only the most significant PCS. The EUMETSAT Advanced Retransmission Service (EARS) provides a total of 290 PCS for the three bands of the IASI spectrum: 90 for Band 1, 120 for Band 2, and 80 for Band 3. This number of PCS allow to retain the atmospheric signal with negligible loss of information [16,63].

PCC exploits the high level of correlation between channels to achieve data compression [19]. The main advantage of PCC is the potential to remove part of the noise present in the original data [16]. However, atmospheric information is also lost. A reduction of the random instrument noise between

4 and 6 is achievable, while values of the reconstructed noise and the atmospheric information loss get close to the optimal ones proposed by the linear estimation theory [64].

A major concern in the use of PCC is that the PCS are determined from a training set. If the data used for training do not contain specific events, like volcanic eruptions, heavy biomass burning, wildfires, etc., these characteristics will not be present in the leading PCS and might be considered noise, reducing the usefulness of the data [62]. In order to minimize this drawback, a global training set, adequate to represent most of the atmospheric situations, should be employed and updated periodically to include rare events [61].

Another concern in the use of PCC is that some features associated with trace gases may not be properly retained in the reconstructed spectra, which is specially critical when the trace gas signal is weak. This may be caused when the number of PCS used in the reconstruction of the signal is not large enough or if the training set from which the reference eigenvectors were derived presents some deficiency [17,65].

PCC is a transform-based approach widely accepted for IASI dimensionality-reduction and for data compression, whose benefits and drawbacks are well known. Although IASI is not an imager but a sounder, coding techniques intended for images may also be employed to compress IASI data. In this paper we will analyze the performance of a wide range of coding techniques for lossless, near-lossless, and lossy compression of IASI data, including current standards and state-of-the-art coding techniques. We will review the performance of coding techniques that follow a different approach than PCC, such as prediction-based techniques, which allow lossless and near-lossess compression. We also evaluate transform-based coding techniques and the current video coding standard (HEVC), which include a rate-distortion optimisation stage to determine which contribution from each transformed channel should be included in the final compressed file, instead of applying the transform and selecting a subset of the transformed channels as PCC does.

We have observed in recent studies [26,27] that some of the proposed lossy compression schemes produce reconstructed radiances that are suitable for statistic retrieval algorithms, achieving competitive performance compared to retrievals performed over the original radiances.

One of the goals of our contribution is to report the performance of several compression schemes for IASI data, which allow different features in the reconstructed data as compared to PCC, for instance, compression of the whole spectra, specific accuracy in the recovered data, etc.

Figure 5 illustrates the proposed coding approach. Once the original data (e.g., radiance data) have been compressed and transmitted, they must be decompressed to produce the reconstructed data (e.g., radiance data too). The dimensions and size of the reconstructed data are identical to that of the original data. If a near-lossless or a lossy compression paradigm is selected, the quality of the reconstructed data will be different than the quality of the original data.

IASI L1C products are the most common format for dissemination of IASI data [54–56,59] and will be the considered data in this research.

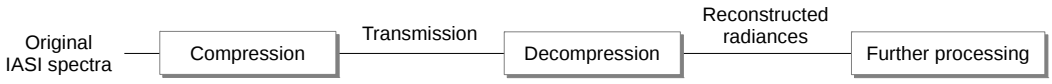


Figure 5. Sequential approach for IASI data compression.

3. Data Compression

This section reports schematically the main characteristics of the compression techniques employed in this paper. Essentially, we introduce first the basic scheme of a data coding system, then we outline the skilled characteristics of the six examined coding techniques, next we provide the setting and parameters used for each tested coding technique, and finally we discuss the benefits of applying a spectral transform along the spectral dimension.

3.1. Data Coding System Pipeline

A data coding system usually comprises three main stages: (1) pre-processing, (2) coding, and (3) post-processing, as illustrated in Figure 6.

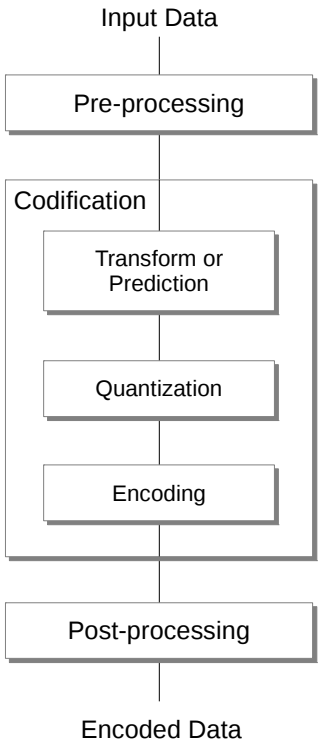


Figure 6. Data compression systems are usually composed of three main stages: pre-processing, coding, and post-processing. The coding stage may, in turn, comprise three steps: either transform or prediction, quantization, and encoding. Only the encoding process is displayed; decoding proceeds in reverse order.

The *pre-processing* stage is intended to prepare the data for the compression process. In some scenarios, like remote sensing, it may become a very important stage, having high influence in the later stages. Partitioning, denoising or segmentation are common processes performed during the pre-processing stage. The *coding* stage encodes the data resulting from the pre-processing stage. Different steps can be carried out in this stage. First, either a transform or a prediction step is applied to modify the representation space. The transform approach aims at providing a more decorrelated and compact representation of the signal. An example of this approach is the wavelet transform, providing a spatial-frequency domain representation. In its turn, the prediction approach aims at exploiting the correlation among neighbouring—causal—coefficients by guessing the next coefficient and incurring in a prediction error, which is expected to facilitate a better performing entropy encoding. The second step is a quantization step, applied in the case of near-lossless or lossy coding, as it entails a loss of information. The third step is an entropy encoding step. Common approaches include Huffman [36], Golomb [66] and Arithmetic encoding [67]. Depending on the compression technique employed, the *post-processing* stage can manipulate either the final codestream or the data recovered. In JPEG 2000, this stage organizes the final codestream to minimize the error between the original data and the reconstructed data at a desired target bit-rate. In HEVC, it defines some parameters for smoothing artifacts in the reconstructed data to improve its quality.

3.2. Characteristics of the Coding Techniques

Focusing on the coding stage, in this paper we screen two *transform-based* coding techniques, namely JPEG 2000 [30] and CCSDS-122.0 [32], three *prediction-based* coding techniques, namely JPEG-LS [29], M-CALIC [31] and CCSDS-123.0 [33], and the most recent video coding standard that includes both a transform and a prediction step, HEVC [34].

For each of the six considered compression techniques, Tables 2 and 3 provide the following information: (1) *Year*: reports when that particular coding technique or standard was published; (2) *Compression paradigm*: indicates which of the three different coding paradigms are deployed, i.e., lossless, near-lossless, or lossy compression and whether that technique is prediction-based or transform-based; (3) *Reference*: cites the main reference; (4) *Pre-processing*: indicates what type of pre-processing stage is performed, if any; (5) *Post-processing*: indicates what type of post-processing stage is performed, if any; (6) *Spatial transform*: for the case of *transform-based* coding techniques, it provides information about what transform is employed to exploit the spatial redundancy; (7) *Prediction*: for the case of *prediction-based* coding techniques, it provides information about what type of prediction is employed to exploit redundancy and whether it is applied in the spatial direction (intra), in the spectral direction (inter), or in both; (8) *Quantization*: indicates what type of quantization is performed, if any; (9) *Bitplane encoding*: explains how a bitplane coding strategy [36] is applied, if any; and (10) *Entropy coder*: provides information about the type of entropy encoder used.

3.3. Setting and Parameter Configuration

All compression techniques allow different parameter and configuration options: on one hand, selecting appropriate settings has a significant influence on the compression performance; on another hand, these settings may determine the computational cost. Careful analysis has to be carried out to disclose appropriate settings.

To account for the reproducibility of the research, for the six evaluated coding techniques, and for each coding paradigm (as most coding techniques allow more than one coding paradigm), Table 4 provides the configurations of setting and mode, spatial transform and spectral transform used in our experiments. *Setting and mode* column refers to particular characteristics of each compression technique; *Spatial transform* column and *Spectral transform* column indicate what type of spatial or spectral transform is applied, if any.

Since JPEG-LS [29] and CCSDS-122.0 [32] coding techniques are devised to encode mono-channel data (2D data), for encoding data scenes with more than one channel (3D data) we used the following methodology: (1) split the data scenes into mono-channel data (in the case of IASI data, each spectral channel with a different wave-number shall be a mono-channel data), (2) each of those mono-channel data are individually encoded, and (3) the total bit-rate is the sum of the bit-rates for each mono-channel data. When JPEG-LS is paired with a spectral transform (see below), the scenes are first spectrally transformed and then the splitting procedure above is applied. In the case of CCSDS-122.0, the upcoming standard CCSDS-122.1 [68] is employed.

Table 2. Technical characteristics of the considered compression techniques (year, compression paradigm, reference, pre-processing, and post-processing).

	JPEG-LS	**JPEG 2000**	**M-CALIC**	**CCSDS-122.0**	**CCSDS-123.0**	**HEVC**
Year	1999	2000	2004	2005	2012	2013
Compression Paradigm	Lossless and near-lossless	Lossless and lossy	Lossless and near-lossless	Lossless and lossy	Lossless	Lossless and lossy
	Prediction-based	Transform-based	Prediction- based	Transform-based	Prediction-based	Prediction- and Transform-based
Reference	[29]	[30]	[31]	[32]	[33]	[34]
PRE-PROCESSING						
	✗	Possibility of multi-channel transform, tile partitioning, and level-shift for unsigned data	✗	✗	✗	Possibility of tiles. channels are partitioned into Coding Tree Units (CTUs).
POST-PROCESSING						
	✗	Bit-stream organization (bit-allocation, data ordering, error resilience, and file format)	✗	✗	✗	Deblock Filtering (DBF) and Sample-Adaptive Offset (SAO). Both stages are optional.

Table 3. Technical characteristics of the considered compression techniques (coding).

	JPEG-LS	**JPEG 2000**	**M-CALIC**	**CCSDS-122.0**	**CCSDS-123.0**	**HEVC**
CODING						
Spatial transform	✗	Wavelet transform (up to 32 levels of IWT 5/3 or DWT 9/7)	✗	Wavelet transform (3 levels of 9/7 Integer DWT or 9/7 Float DWT)	✗	Discrete cosine transform (DCT) and discrete sine transform (DST)
Prediction	Intra: using 3 neighbor samples	✗	Inter: using 2 channels for spectral prediction	✗	• Intra: using 1 or 4 neighbor samples • Inter: up to 15 channels for spectral prediction	• Intra: using adjacent blocks as reference, 33 directional plus 2 special modes supported. • Inter: up to 15 frames
Quanti- zation	Uniform scalar quantization	• Uniform scalar deadzone quantization (Part-1 of standard) • Variable scalar deadzone quantization, and Trellis coded quantization (Part-2 of standard)	Uniform scalar quantization	Uniform scalar quantization	✗	Uniform scalar quantization
Bitplane coding	✗	Each bitplane is encoded with three coding passes: (1) significance propagation pass, (2) magnitude refinement pass, and (3) clean-up pass. For the first bitplane only clean-up pass is used	✗	First, the first bits of the quantized DC coefficients are encoded. Then, the remaining DC coefficients bit planes are encoded along with the bit planes of AC coefficients using several refinement passes	✗	✗
Entropy coder	Golomb Coder and Run Length Coder	MQ Arithmetic Coder. Contextual binary arithmetic coder. Contexts are defined using the 8 adjacent neighbors	Contextual Arithmetic Coder using up to 1024 contexts	Variable Length Coder and Fixed Length Coder	Golomb Coder	Arithmetic Coder (CABAC with 154 contexts) and Variable Length Coder (CAVLC)

Table 4. For each coding technique, the configuration used in the experiments is reported. Default option is employed for the parameters not specified in the table.

Coding Technique	Paradigm	Setting and Mode	Spatial Transform	Spectral Transform
JPEG-LS	Lossless	Plane-interleaved mode	—	• Multilevel Clustering RKLT (200 clusters in first level and multilevel mode) • IWT 5/3 (5 levels) • RPOT • Maximum RWA (Exogenous variant)
	Near-lossless	Plane-interleaved mode	—	—
JPEG 2000	Lossless	Code-blocks of 64 × 64 size and 1 quality layer	IWT 5/3 (5 levels)	• Multilevel Clustering RKLT (200 clusters in first level and multilevel mode) • IWT 5/3 (5 levels) • RPOT • Maximum RWA (Exogenous variant)
	Lossy	Code-blocks of 64 × 64 size and 1 quality layer	DWT 9/7 (5 levels)	• Multilevel Clustering KLT (200 clusters in first level and multilevel mode) • DWT 9/7 (5 levels) • POT
M-CALIC	Lossless	Default	—	• Multilevel Clustering RKLT (200 clusters in first level and multilevel mode) • IWT 5/3 (5 levels) • RPOT • Maximum RWA (Exogenous variant)
	Near-lossless	Default	—	—
CCSDS-122.0	Lossless	Default	Default	• Multilevel Clustering RKLT (200 clusters in first level and multilevel mode) • IWT 5/3 (5 levels) • RPOT • Maximum RWA (Exogenous variant)
	Lossy	Default	Default	• Multilevel Clustering KLT (200 clusters in first level and multilevel mode) • DWT 9/7 (5 levels) • POT
CCSDS-123.0	Lossless	Default	—	• Multilevel Clustering RKLT (200 clusters in first level and multilevel mode) • IWT 5/3 (5 levels) • RPOT • Maximum RWA (Exogenous variant)
HEVC	Lossless	Intra and inter prediction	Default	• Multilevel Clustering RKLT (200 clusters in first level and multilevel mode) • IWT 5/3 (5 levels) • RPOT • Maximum RWA (Exogenous variant)

3.4. Spectral Transforms

To achieve competitive compression performance in hyperspectral data like IASI L1C, which are composed of more than 8000 channels, it is of paramount importance to exploit its high spectral redundancy. Spectral transforms are a commonly adopted approach. Among all the spectral transforms in the literature, we focused on four of them: (1) Wavelet Transform (WT), due to its extended use and low complexity; (2) the Karhunen-Loève Transform (KLT), because it is the optimal decorrelating transform for gaussian sources despite its high computational cost; (3) the Pairwise Orthogonal Transform (POT), as it is a low complexity approximation to KLT and is at the basis of upcoming standard CCSDS-122.1; and (4) the Regression Wavelet Analysis Transform (RWA) because of its highly competitive performance and bearable computational cost.

Depending on the coding paradigm used, a reversible (RKLT, RPOT, IWT 5/3, RWA) or an irreversible (KLT, POT, DWT 9/7) version of the spectral transforms must be employed. For RWA, two different estimation models [38] could be adopted: *Maximum model* and *Restricted model*. Here, we use a variant of the Maximum model, the *Exogenous variant*, which considerably reduces the computational cost and does not entail transmission of any side-information.

All of these spectral transforms are used in combination with the six coding techniques evaluated here. Although M-CALIC, CCSDS-123.0 and HEVC already exploit themselves the spectral redundancy by employing prediction techniques, we also pair them with the different spectral transforms and evaluate their coding performance.

A particular aspect to consider when applying a spectral transform is the computational complexity, because it may be critical in many scenarios. The computational cost in floating-point operations (FLOPs) of applying lossless forward and backward transforms on a typical IASI L1C orbit for RKLT, IWT, RPOT, and RWA is shown in Figure 7. In this particular case, the RKLT cost is over 2400 times higher than that of RPOT or IWT, and approximately 6 times higher than that of Maximum RWA (Exogenous variant).

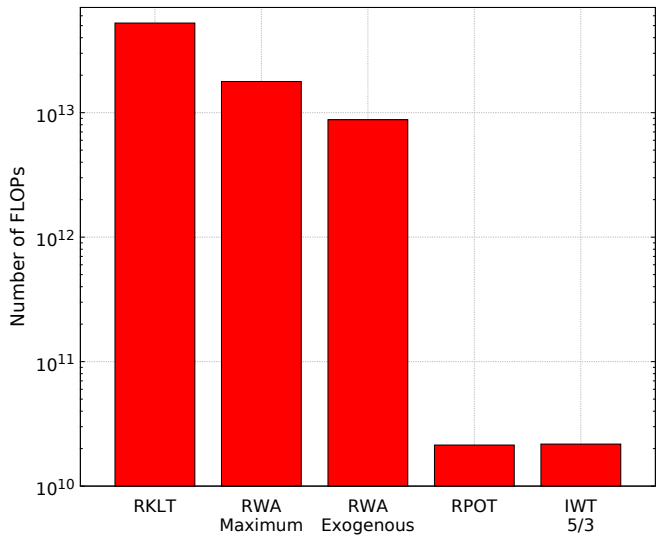


Figure 7. Cost comparison in FLOPs for the different spectral transforms used in the experiments applied to an IASI L1C orbit with 8461 spectral channels and a spatial resolution of $765 \times 30 \times 4$ (number of scan lines $\times$ number of FORs per line $\times$ number of IFOVs per FOR).

Table 5 reports a detailed analysis for the reversible version of each spectral transform, in terms of computational cost (in FLOPs). The computational cost of RKLT and Multilevel Clustering RKLT (see below) mainly depends on the squared number of spectral channels, which substantially increases the computational cost as the number of channels increases. The IWT 5/3 and RPOT transforms have

an approximately linear cost in relation to the spatial locations and the spectral channels. The cost of RWA is dominated by the estimation of the regression coefficients and the generation of the predictions [38]. In the case of the Exogenous variant, the estimation stage is performed offline and does not imply additional transform cost during the encoding process.

Table 5. Computational cost in FLOPs for IWT, RPOT, RWA Maximum, RWA Exogenous, RKLT, and Multilevel Clustering RKLT. z is the number of spectral channels, m is the number of spatial samples per channel, y is the number of rows, l is the number of wavelet decomposition levels, k is the number of detail channels employed in the prediction level i [38], s is the number of spectral channels per cluster ($s \ll z$), and C is the total number of clusters.

Transform	FLOPs
IWT	$2 \times 14(1 - \frac{1}{2^l})mz$
RPOT	$16mz + 26zy - 12m - 28y + 11mz + 5zy - 10m - 5y$
RWA Maximum	$8(1 - \frac{1}{2^l})mz + (\sum_{i=1}^{l}(2m-1)(k_i+1)^2 + (k_i+1)^3 + (\frac{z}{2^i})(k_i+1)\,[(2m-1)+(2k_i+1)]) + +(2\sum_{i=1}^{l}(2k_i-1)m\frac{z}{2^i}) + 2m(z-1)$
RWA Exogenous	$8(1 - \frac{1}{2^l})mz + (2\sum_{i=1}^{l}(2k_i-1)m\frac{z}{2^i}) + 2m(z-1)$
RKLT	$m(4z^2 + 3z + 1) + \frac{32}{3}z^3 + \frac{1}{2}z^2 - \frac{37}{6}z + 5 + m(3z^2 + z - 3)$
Multilevel Clustering RKLT	$\sum_{c \in C} m(4s^2 + 3s + 1) + \frac{32}{3}s^3 + \frac{1}{2}s^2 - \frac{37}{6}s + 5 + m(3s^2 + s - 3)$

As seen, a case of very high computational cost and memory requirements transform is KLT/RKLT, which renders it unusable in situations where the number of spectral channels is large. To alleviate its high computational cost, there exist a number of strategies. Here, we use a divide-and-conquer strategy, the Multilevel Clustering KLT/RKLT [69], as described in Section 3.5.

Some spectral transforms may produce data with more than 16 bits per pixel per channel (bpppc), whereas software implementations used for JPEG-LS, M-CALIC, CCSDS-122.0, CCSDS-123.0, and HEVC deal with input data of at most 16 bpppc. In these cases, each transformed channel is split into two different channels: a channel formed by the 16 most significant bits (MSB), and a channel formed by the 16 least significant bits (LSB). The MSB channels and the LSB channels are grouped in two different volumes and encoded separately. The total bit-rate of the compressed data is the sum of the bit-rates of the two compressed volumes.

3.5. Divide-and-Conquer Strategy for KLT/RKLT

The Classical Clustering divide-and-conquer strategy divides a large transform in several clusters and applies a smaller transform to each cluster. This approach significantly reduces the overall computational cost, but only provides local decorrelation within each cluster. Global decorrelation can be achieved by applying a Multilevel Clustering strategy, where the most important parts of each local transform are further decorrelated in the next levels. In Figure 8, the structure of a plain KLT/RKLT transform, of a Classical Clustering KLT/RKLT and of a Multilevel Clustering KLT/RKLT are displayed.

The computational complexity of KLT/RKLT stems from the number of spectral channels to be transformed. The complexity of Multilevel Clustering KLT/RKLT depends thus on the employed cluster size. An appropriate configuration for applying Multilevel Clustering KLT/RKLT on IASI L1C orbits is found by assessing three different criteria: *computational cost*, related to the *execution time*, and *transform coding performance*. To perform this assessment, we consider an IASI L1C orbit with 2^{13} (8192) spectral channels (discarding the last 269 channels of the 8461 channels spectrum).

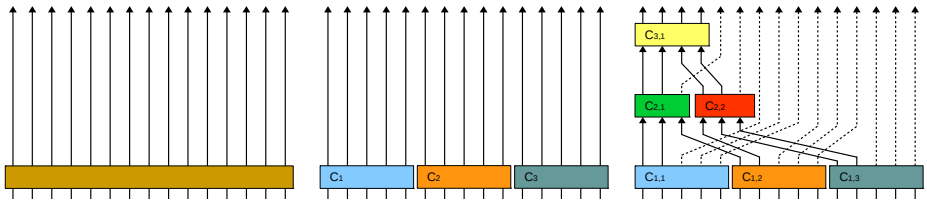


Figure 8. Structure of plain KLT/RKLT, Classical Clustering KLT/RKLT, and Multilevel Clustering KLT/RKLT. This example decorrelates 15 spectral channels. Each arrow denotes a channel and each coloured rectangle represents the computation of a KLT/RKLT transform. In the case of Classical Clustering KLT/RKLT, three clusters are employed. In the case of Multilevel Clustering KLT/RKLT, 3 levels of Multilevel Clustering are applied. (**a**) Plain KLT/RKLT; (**b**) Classical Clustering KLT/RKLT; (**c**) Multilevel Clustering KLT/RKLT.

3.5.1. Computational Cost

The computational cost of different cluster sizes for Multilevel Clustering RKLT is illustrated in Figure 9. The computational cost rapidly increases as the number of clusters defined in the first level decreases, which is equivalent to increase the cluster size. Notice that using 1 cluster is identical to not using any clustering strategy.

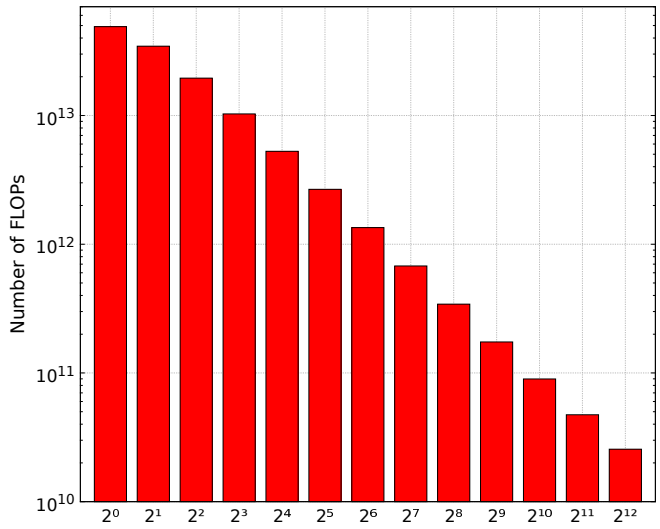


Figure 9. Cost comparison in FLOPs for different cluster sizes of Multilevel Clustering RKLT applied to an orbit with 2^{13} spectral channels and a spatial resolution of $765 \times 30 \times 4$ (number of scan lines $\times$ number of FORs per line $\times$ number of IFOVs per FOR).

3.5.2. Execution Time

The forward and backward execution times of different cluster sizes when Multilevel Clustering RKLT is applied to the proposed 2^{13} IASI L1C orbit are compared in Figure 10. The longest runtimes are required when less than 2^5 clusters (2^8 channels per cluster) are defined in the first level. Execution times are not provided for 1, 2, 4, and 8 clusters defined in the first level due to the high computational complexity. All experiments have been performed on an Intel Xeon CPU E3-1230 V2 @ 3.30 GHz processor.

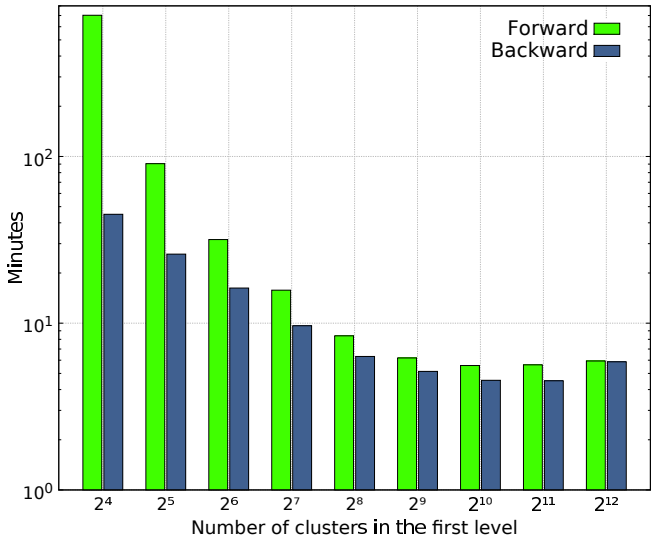


Figure 10. Runtime comparison in minutes for different cluster sizes of Multilevel Clustering RKLT applied to an IASI L1C orbit with 2^{13} spectral channels and a spatial resolution of $765 \times 30 \times 4$ (number of scan lines $\times$ number of FORs per line $\times$ number of IFOVs per FOR). The dissemination granularity of the data is 3 min for Level 1c [70].

3.5.3. Transform Coding Performance

The performance of Multilevel Clustering RKLT in terms of both computational cost and transform coding performance is illustrated in Table 6. The best trade-off between computational cost and entropy of the transformed orbit is obtained when 2^7 (128) or 2^8 (256) clusters are defined in the first level.

Table 6. Computational cost (in FLOPs) and transform performance (entropy) for different cluster sizes of Multilevel Clustering RKLT. Transform performance results are not provided when 2^0, 2^1, 2^2, and 2^3 clusters are defined in the first level. For these cases, applying the spectral transform would require several days due to the high computational cost, which results impractical in a real scenario.

Number of Clusters Defined in the First Level	Cluster Size	Total Number of Clusters	FLOPs	Entropy
2^0	2^{13}	1	4.90×10^{13}	-
2^1	2^{12}	3	3.45×10^{13}	-
2^2	2^{11}	7	1.95×10^{13}	-
2^3	2^{10}	15	1.03×10^{13}	-
2^4	2^9	31	5.27×10^{12}	5.20
2^5	2^8	63	2.67×10^{12}	5.14
2^6	2^7	127	1.35×10^{12}	5.14
2^7	2^6	255	6.78×10^{11}	5.13
2^8	2^5	511	3.42×10^{11}	5.13
2^9	2^4	1023	1.74×10^{11}	5.17
2^{10}	2^3	2047	8.98×10^{10}	5.25
2^{11}	2^2	4095	4.74×10^{10}	5.48
2^{12}	2^1	8191	2.56×10^{10}	5.94

Based on the previous analysis of Multilevel Clustering RKLT, 200 clusters in the first level and multilevel mode has been used in all the experiments of this manuscript.

For the irreversible case, a Multilevel Clustering KLT with the same configuration has been selected too, because the coding performance for KLT and Multilevel Clustering KLT is almost equivalent, because Multilevel Clustering KLT requires much less side-information than KLT, and because although KLT has a lower computational complexity than RKLT, its application on a 8461 IASI L1C orbit may take over 30 h.

4. Experimental Results

This section presents a set of experiments aimed at the analysis and evaluation for lossless, near-lossless, and lossy compression of IASI L1C products. First, a description of the IASI L1C products and the software employed in the experiments is provided. Then, we will focus on the compression results produced by the different coding techniques.

4.1. Data Collection and Software

To obtain sound conclusions, the experiments are conducted with a set of 96 IASI L1C orbits granted by EUMETSAT (http://catalogue.ceda.ac.uk/), representing more than 148 Gigabytes. 48 orbits belong to IASI-A and 48 orbits belong to IASI-B. These orbits are acquired with, respectively, MetOp-A and MetOp-B satellites. To get a representative data set, orbits acquired throughout a full year are selected for each sensor: from July 2013 to June 2014, 4 orbits per month, 1 per week. For the sake of conciseness, details and results will be grouped by instrument, computing the average of the 48 orbits. Results for each individual orbit are very similar. All data are 16 bpppc and are stored as signed integers. For each product, Table 7 provides the sizes and the average zero-order entropy, which is the smallest number of bits, on average, required to represent a sample without considering any dependency among pixels within or between channels.

Table 7. IASI L1C products used in the experiments. Sizes and averaged zero-order entropies per instrument are provided (48 orbits per instrument). *M* is the number of spectral channels, *Ns* is the number of scan lines, *N-FORs* is the number of FORs per line, and *N-IFOVs* is the number of IFOVs per FOR.

Instrument	Size (M × Ns × N-FORs × N-IFOVs)	Average Entropy
IASI-A Products	8461 × (630-787) × 30 × 4	12.84
IASI-B Products	8461 × (742-788) × 30 × 4	12.83
Average	8461 × (761) × 30 × 4	12.83

All software used to produce the experimental results is public. The implementations employed are the following: JPEG-LS software [71], Kakadu software [72] for JPEG 2000, M-CALIC software [73], TER software [74] for CCSDS-122.0, EMPORDA software [75] for CCSDS-123.0, and HEVC software [76]; Spectral Transform software [77] for Multilevel Clustering KLT/RKLT and WT, Pairwise Orthogonal Transform software [78] for POT/RPOT, and Regression Wavelet Analysis software [79] for RWA.

4.2. Lossless Compression Results

Lossless compression of IASI L1C products is evaluated for the suggested approach: spectral transform followed by coding technique. Four spectral transforms have been tested: Multilevel Clustering RKLT, IWT, RPOT and RWA. All six coding techniques are assessed: JPEG-LS, JPEG 2000, M-CALIC, CCSDS-122.0, CCSDS-123.0 and HEVC. Table 8 reports the average lossless coding performance (compression ratio). Results suggest that:

Table 8. Lossless compression of IASI L1C products. Results are reported in compression ratio (higher is better). Percent savings (higher is better) with respect to original technique are provided within brackets.

IASI-A—Lossless Compression Ratio & Percent Savings					
Tra. / Tech	No Transform	IWT	RPOT	RWA	Multilevel Clustering RKLT
JPEG-LS	1.78:1	2.26:1 (21.24%)	2.26:1 (21.24%)	2.44:1 (27.05%)	2.46:1 (27.64%)
JPEG 2000	1.73:1	2.24:1 (22.77%)	2.24:1 (22.77%)	2.43:1 (28.81%)	2.47:1 (29.96%)
M-CALIC	2.32:1	2.32:1 (0.00%)	2.34:1 (0.85%)	2.48:1 (6.45%)	2.54:1 (8.66%)
CCSDS-122.0	1.68:1	2.13:1 (21.13%)	2.13:1 (21.13%)	2.29:1 (26.64%)	2.33:1 (27.90%)
CCSDS-123.0	2.42:1	2.42:1 (0.00%)	2.39:1 (−1.24%)	2.46:1 (1.63%)	2.47:1 (2.02%)
HEVC	2.23:1	2.29:1 (2.62%)	2.28:1 (2.19%)	2.45:1 (8.98)	2.50:1 (10.80%)
IASI-B—Lossless Compression Ratio & Percent Savings					
Tra. / Tech	No Transform	IWT	RPOT	RWA	Multilevel Clustering RKLT
JPEG-LS	1.79:1	2.28:1 (21.49%)	2.27:1 (21.15%)	2.45:1 (26.94%)	2.48:1 (27.82%)
JPEG 2000	1.74:1	2.25:1 (22.67%)	2.25:1 (22.67%)	2.44:1 (28.69%)	2.49:1 (30.12%)
M-CALIC	2.34:1	2.33:1 (−0.43%)	2.35:1 (0.43%)	2.50:1 (6.40%)	2.56:1 (8.59%)
CCSDS-122.0	1.69:1	2.14:1 (21.03%)	2.14:1 (21.03%)	2.30:1 (26.52%)	2.34:1 (27.78%)
CCSDS-123.0	2.44:1	2.44:1 (0.00%)	2.40:1 (−1.64%)	2.48:1 (1.61%)	2.48:1 (1.61%)
HEVC	2.24:1	2.30:1 (2.61%)	2.29:1 (2.18%)	2.47:1 (9.31%)	2.52:1 (11.11%)

- *Coding performance for IASI-A and IASI-B products is nearly the same.* Lossless compression of IASI-B products is, on average, only 0.75% better than for IASI-A. This negligible difference happens for all IASI-A and IASI-B products and for all compression schemes.
- *IASI L1C data present high spectral redundancy.* M-CALIC, CCSDS-123.0 and HEVC, which originally exploit the spectral redundancy, achieve better outcomes than JPEG-LS, JPEG2000 or CCSDS-122.0, which do not exploit this redundancy. For the latter techniques, taking advantage of this redundancy through a spectral transform yields significantly better compression performance, bridging the gap with the former techniques.
- *Compression techniques that already exploit the spectral redundancy by themselves also benefit from applying a spectral transform.* When paired with a spectral transform, M-CALIC, CCSDS-123.0, and HEVC usually achieve better coding performance too (except for IWT + M-CALIC and RPOT + CCSDS-123.0). This effect is specially significant in the case of HEVC, where up to 11.11% can be improved, but also for M-CALIC, where gains are close to 9%. Gains for CCSDS-123.0, which was the coding technique providing the best performance, are less meaningful.
- *Multilevel Clustering RKLT or RWA yield the best coding performance.* Multilevel Clustering RKLT brings the largest improvements, closely followed by RWA. As compared to original CCSDS-123.0, which is the coding technique providing the best performance when no spectral transform is applied, the improvements for Multilevel Clustering RKLT and for RWA when combined with M-CALIC are, respectively, of 4.7% and 2.4%.
- *Compression ratios over 2.5:1 (bit-rates close to 6.3 bpppc) can be achieved for lossless compression of IASI L1C products.* The best results are obtained by *Multilevel Clustering RKLT + M-CALIC*, which achieves, on average, a compression ratio of 2.54:1 for IASI-A products and 2.56 for IASI-B products.

4.3. Near-Lossless Compression Results

Two different coding techniques, M-CALIC and JPEG-LS, are used for near-lossless compression of IASI L1C products. Eight different Peak Absolute Errors (PAE) ($\delta \in \{1, 3, 7, 15, 31, 63, 127, 255\}$) are considered. Results are reported in Table 9. Three main observations can be drawn:

- *As expected, compression ratio increases as PAE increases.*

- *Competitive compression performance is achieved even by allowing small errors.* Large savings over 17% and 30% with respect to lossless compression are already achieved for such small PAE as 1 and 3.
- *M-CALIC yields higher compression ratio than JPEG-LS.* M-CALIC uses an arithmetic coder, while JPEG-LS uses Golomb codes, for which bit-rates below 1 bpppc are not achievable.

Figure 11 illustrates the rate-distortion performance of near-lossless compression in terms of Signal Noise Ratio (SNR) Energy vs. PAE. Using as small PAE as 1 and 3, SNR Energy over 65 dB can already be achieved.

Table 9. Near-lossless compression of IASI L1C products. Results are reported in compression ratio (higher is better). Results for lossless compression (PAE = 0) are included. Percent savings (higher is better) with respect to lossless compression are provided within brackets.

	IASI-A		**IASI-B**	
PAE	**JPEG-LS**	**M-CALIC**	**JPEG-LS**	**M-CALIC**
0	1.78	2.32	1.79	2.34
1	2.17 (17.97%)	3.02 (23.18%)	2.18 (17.89%)	3.05 (23.28%)
3	2.60 (31.54%)	3.90 (40.51%)	2.61 (31.42%)	3.95 (40.76%)
7	3.15 (43.49%)	5.21 (55.47%)	3.18 (43.71%)	5.28 (55.68%)
15	3.93 (54.71%)	7.34 (68.39%)	3.98 (55.03%)	7.48 (68.72%)
31	5.11 (65.17%)	11.11 (79.18%)	5.18 (65.44%)	11.35 (79.38%)
63	6.99 (74.54%)	18.39 (87.38%)	7.08 (74.72%)	18.82 (87.57%)
127	10.00 (82.20%)	33.33 (93.03%)	10.19 (82.43%)	34.04 (93.13%)
255	15.09 (88.20%)	61.54 (96.23%)	15.38 (88.36%)	64.00 (96.34%)

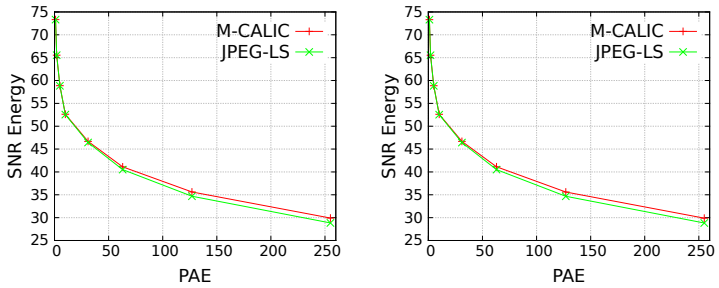


Figure 11. Rate-distortion performance of near-lossless compression of IASI L1C products. Results report SNR Energy (in dB, higher is better) vs. PAE. (**a**) IASI-A; (**b**) IASI-B.

4.4. Lossy Compression Results

Lossy compression of IASI L1C products is evaluated using JPEG 2000 and CCSDS-122.0 standards along with three spectral transforms: Multilevel Clustering KLT, DWT and POT. All schemes are evaluated using nine target bit-rates between 0.01 and 2 bpppc.

Figure 12 illustrates the lossy compression performance of IASI L1C products for JPEG 2000 and CCSDS-122.0. Several conclusions can be drawn:

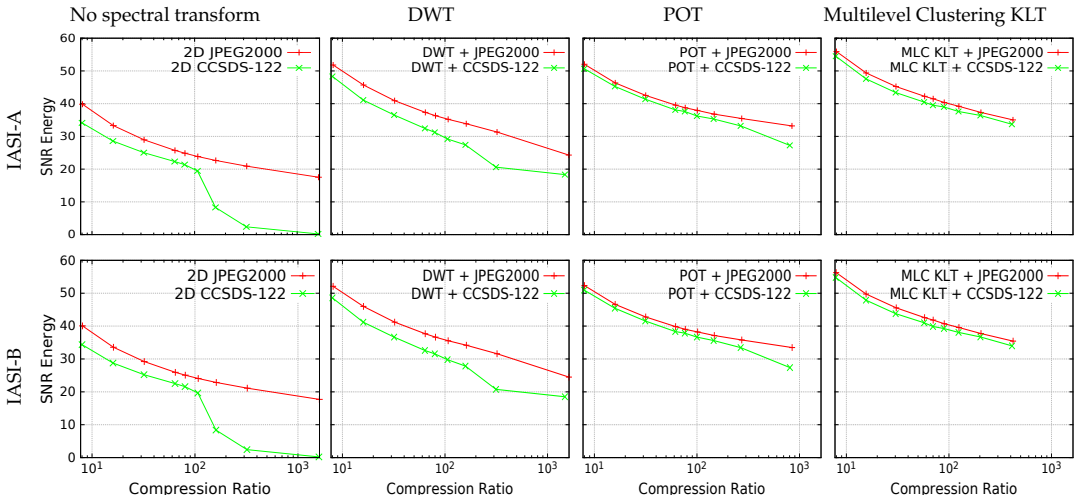


Figure 12. Rate-distortion performance of lossy compression of IASI L1C products. Results report SNR Energy (in dB, higher is better) vs. compression ratio. Results for different spectral transforms are plotted in the columns. In each plot, curves for JPEG 2000 and CCSDS-122.0 performance are displayed. Ranges are the same in all the plots to ease the comparison. Top row: IASI-A products; Bottom row: IASI-B products. POT and Multilevel Clustering KLT are not able to reach such high compression ratios (over 1000:1) as DWT because side-information needs to be transmitted besides the compressed data.

- *Exploiting the spectral redundancy is essential to achieve competitive performance.* Applying a spectral transform always outperforms the scheme that does not exploit the spectral redundancy. Performance difference is more apparent as the compression ratio decreases, growing from 5 to over 15 dB.
- *Multilevel Clustering KLT yields the best coding performance.* As happened for lossless compression, also in the case of lossy compression, Multilevel Clustering KLT furnishes the highest results, followed by POT and DWT. At high compression ratios (higher than 20:1), POT yields almost equivalent performance, mostly because of the larger size of the side-information needed by Multilevel Clustering KLT.
- *JPEG 2000 outperforms CCSDS-122.0.* JPEG 2000 is a more complex coding technique that is able to produce more competitive results.
- *Plain 2D CCSDS-122.0 yields low performance at high compression ratios.* This standard starts achieving good results for compression ratios lower than 100:1.

4.5. Comparison between Near-Lossless and Lossy Compression

A comparison of the two analyzed compression paradigms that introduce distortion in the reconstructed data, i.e., near-lossless and lossy compression, is performed in Figure 13. The best coding scheme for near-lossless (M-CALIC) and for lossy (Multilevel Clustering KLT + JPEG 2000) compression are compared from the point of view of PAE and SNR Energy. Bit-rates between 0.1 and 2 bpppc are compared (very large PAE—higher than 1023—are requested to achieve bit-rates lower than 0.1 for near-lossless compression). Some conclusions can be drawn:

- *Near-lossless outperforms lossy compression in terms of PAE.* Near-lossless compression introduces lower maximum errors in the data than lossy compression.
- *Lossy compression outperforms near-lossless compression in terms of SNR Energy.* Lossy compression yields larger results, especially at large compression ratios.

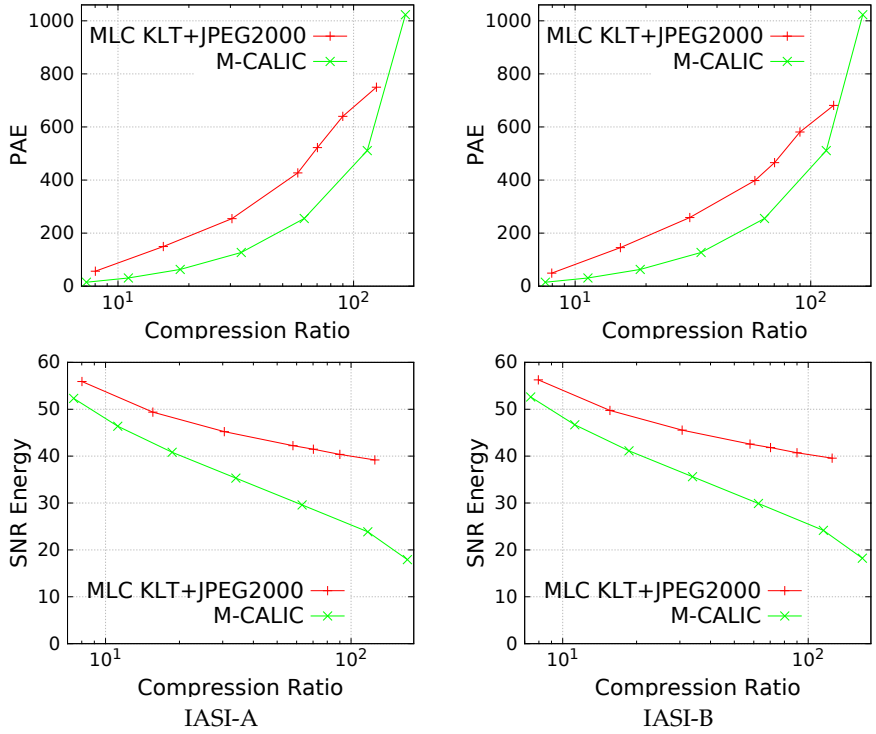


Figure 13. Performance comparison between near-lossless (M-CALIC) and lossy compression (Multilevel Clustering KLT + JPEG 2000). Top row: PAE (lower is better); Bottom row: SNR Energy (in dB, higher is better).

4.6. Compression and Decompression Runtimes

IASI Level 1 products are distributed to users in different dissemination modes and formats. While the timeliness for Near-Real Time dissemination through EUMETCast is 2 h 15 min, the timeliness for products on the EUMETSAT Data Centre retrieval is approximately 8–9 h [70]. Table 10 summarizes the compression runtimes for the coding schemes that provide the best performance for lossless, near-lossless and lossy compression. The decompression runtimes, which are applicable at the receiver side, are also provided. All experiments have been performed on an Intel Xeon CPU E3-1230 V2 @ 3.30 GHz processor.

Table 10. Compression and decompression runtimes for the coding schemes that produce the best performance for lossless, near-lossless, and lossy compression. The PAE employed for near-lossless compression is 1. The target bit-rate used for lossy compression is 2 bpppc. All times are expressed in minutes.

Runtimes (in Minutes)	Lossless	Near-Lossless	Lossy
Compression	81.7	15	13.4
Decompression	41.4	11.3	6.2

Compression schemes that involve lossy coding achieve competitive runtimes and might be considered in a near-real time scenario. Both near-lossless and lossy compression require less than 15 min in the compression stage. At the receiver side, 11 and 6 min for near-lossless and lossy coding, respectively, would be required to decompress the codestream. Longer runtimes are required for lossless compression mainly due to the computation of Multilevel Clustering RKLT. Lossless compression would be appropriate only in scenarios where the delivery time is not critical.

4.7. Analysis of the Reconstructed Radiances

To evaluate the usefulness of the reconstructed radiances, M-CALIC and Multilevel Clustering KLT + JPEG 2000, which are the compression schemes that produce the best performance for, respectively, near-lossless and lossy compression, are compared with Principal Component Compression (PCC).

The experiments are conducted using the product IASI_xxx_1C_M02_20140305023859Z_20140305042058Z_N_O_20140305042027Z (details about file naming convention can be found at [50]). To simplify the comparison, the first 1800 channels of this IASI L1C orbit are used. All of them belong to Band-1. Two compression ratios are compared by retaining a different number of eigenvectors in PCC, either 150 or 200, which is common in practical scenarios. For M-CALIC and Multilevel Clustering KLT + JPEG 2000, the PAE and target bit-rate that produce, respectively, the same compression ratio as compared to PCC are employed. Table 11 summarizes the settings for each experiment.

Table 11. Compression setting for PCC, M-CALIC, and Multilevel Clustering + JPEG 2000 comparison.

		PCC	**M-CALIC**	**Multilevel Clustering KLT + JPEG 2000**
	Compression ratio	PC scores	PAE	Target bit-rate
Experiment 1	9:1	200	19	1.78
Experiment 2	12:1	150	29	1.33

The noise covariance matrix of the original radiances and of the reconstructed radiances after Principal Component Compression for experiment 1 and experiment 2 are illustrated in Figure 14. It is known that the noise covariance matrix of the original radiances is diagonal, while the noise covariance matrix of the reconstructed radiances is quite similar when 200 or 150 PCS are employed.

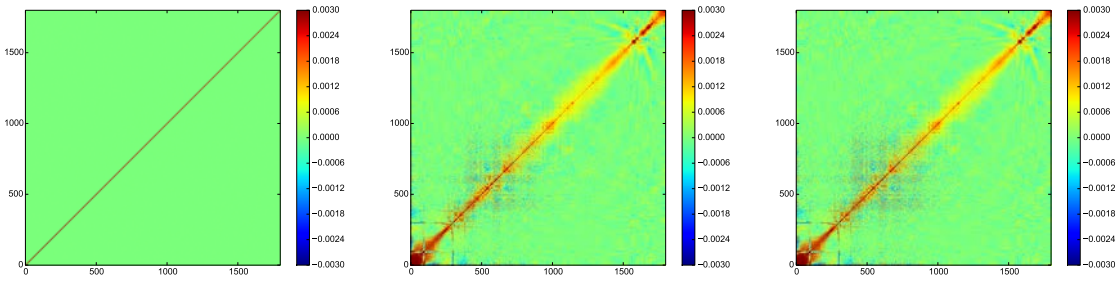


Figure 14. Noise covariance matrix of the original radiances and noise covariance matrix of the reconstructed radiances after Principal Component Compression when 200 and 150 PCS are employed.

In Figure 15 the normalized radiance residual statistics as a function of component number for experiment 1 and experiment 2 are shown. The normalized reconstructed radiances are subtracted from the normalized original radiances. Normalization takes into account the noise covariance matrix inherent to IASI. The average of the normalized radiance residuals, the standard deviation and the maximum and minimum values per channel are reported.

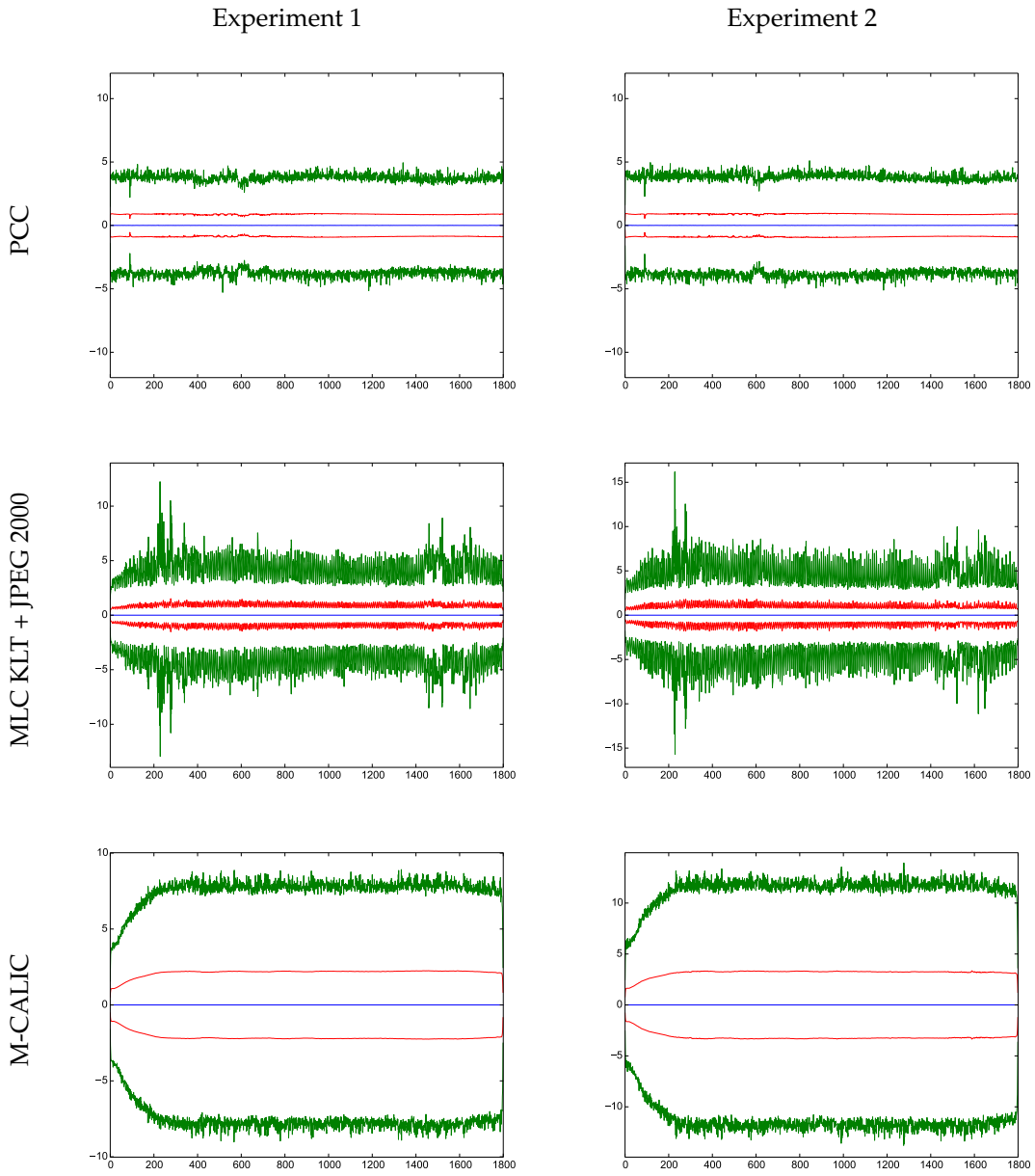


Figure 15. Normalized radiance residuals statistics. The average of the normalized radiance residuals is shown in blue, standard deviation in red, and maximum and minimum values in green.

The average of the normalized radiance residual and the standard deviation are very similar for all compression schemes. As for the maximum and minimum differences, PCC and Multilevel Clustering KLT + JPEG 2000 produce smaller values as compared to M-CALIC. The magnitude of the maximum and minimum values is slightly lower for PCC than for Multilevel Clustering KLT + JPEG 2000.

The covariance matrix of the original and reconstructed radiances is illustrated in Figure 16 for the three different coding techniques at the proposed compression ratios. The nature of the original data and of the reconstructed data has a similar nature.

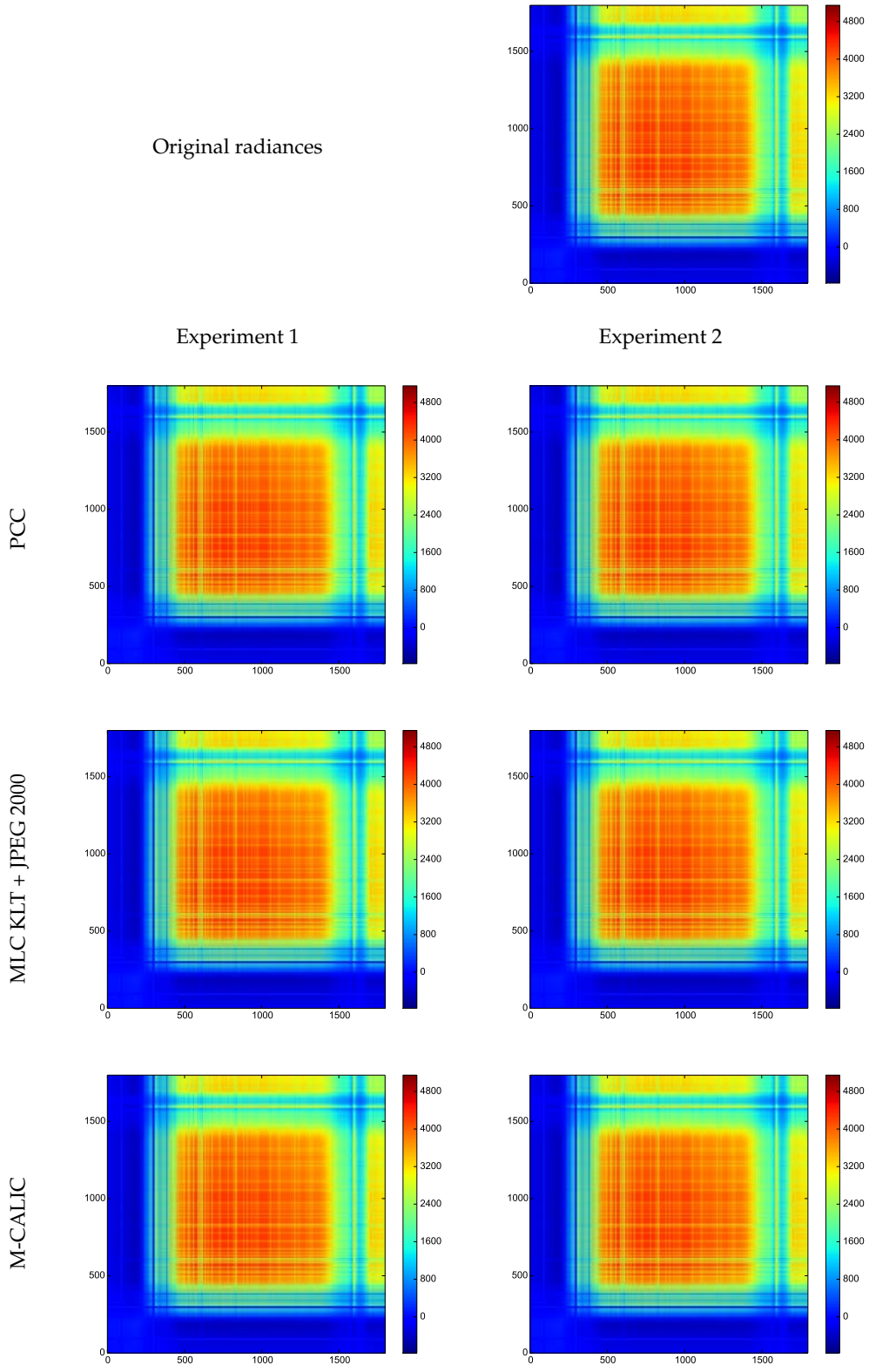


Figure 16. Covariance matrix of the original radiances and covariance matrix of the reconstructed radiances.

To analyze the impact of the compression, Figure 17 reports the differences between the covariance matrix of the original radiances and the covariance matrix of the reconstructed radiances.

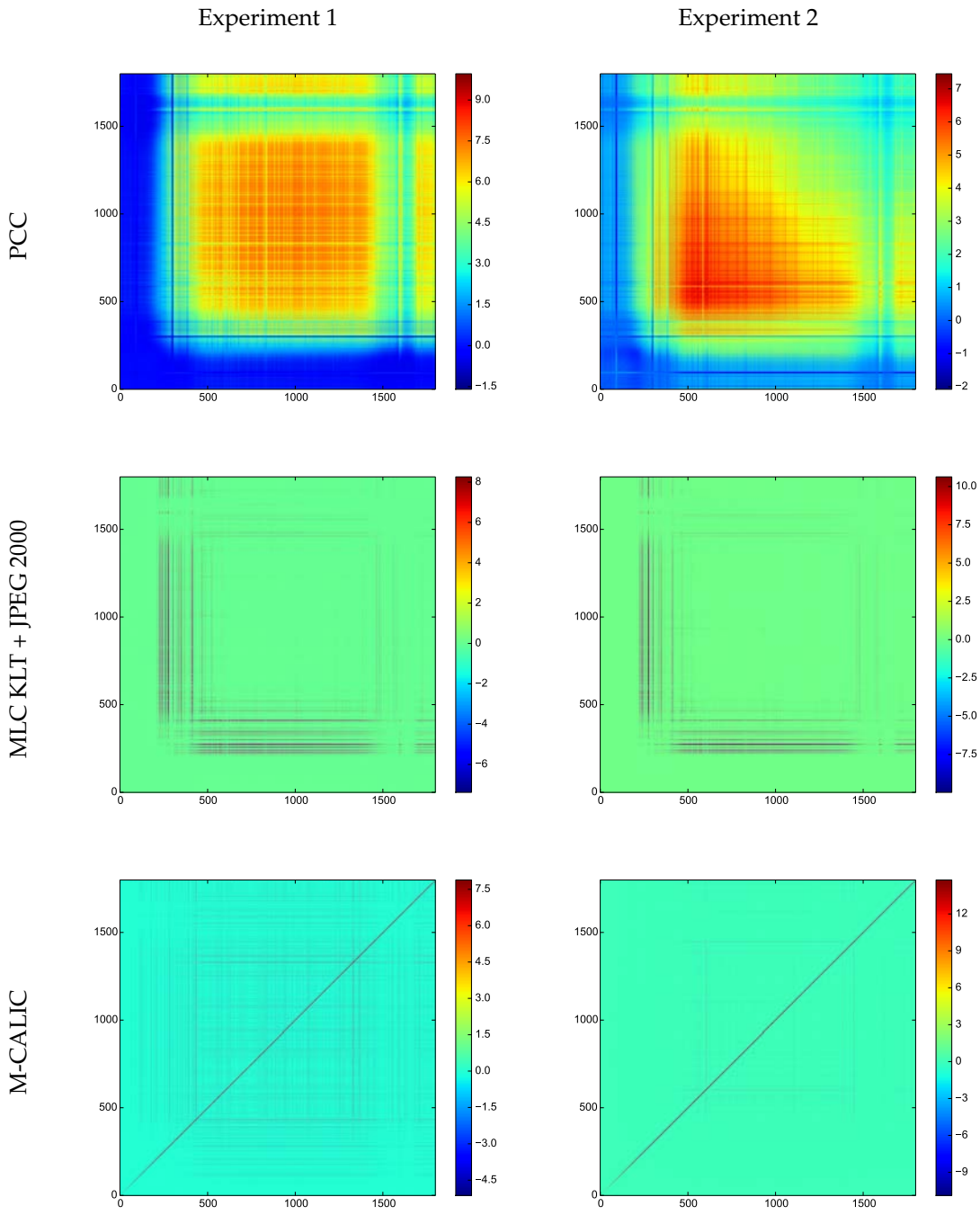


Figure 17. Differences between the covariance matrix of the original radiances and the covariance matrix of the reconstructed radiances.

The differences are very similar for the two compression ratios analyzed in all coding schemes. The covariance matrices of the reconstructed spectra from Multilevel Clustering KLT + JPEG 2000 and M-CALIC are very similar to the covariance matrix of the original data. For M-CALIC, the difference is focused in the main diagonal, while for Multilevel Clustering KLT + JPEG 2000 the differences are clear in a small set of channels. For PCC, the differences are more apparent.

4.8. Discussion

The performance of IASI L1C data compression has been investigated for lossless, near-lossless, and lossy compression. Lossless compression is sometimes a demanded requirement in remote sensing applications because introducing some amount of distortion in the reconstructed data may compromise the quality of derived products. For IASI orbits, lossless compression can contribute to alleviate the large size of the data. As reported in Table 8, lossless compression can reduce the data size to less than half the original size, achieving compression ratios of 2.5:1.

Experimental results reveal that compression techniques that originally exploit the spectral redundancy such as M-CALIC, CCSDS-123.0 and HEVC produce better performance than JPEG-LS, JPEG2000 and CCSDS-122.0, which do not exploit the spectral dimension. It is acknowledged that exploiting the spectral redundancy present in hyperspectral data is of paramount importance to achieve competitive compression performance [80]. This is especially critical in the case of IASI L1C products due to the large number of highly correlated spectral channels. CCSDS-123.0, which is the coding technique providing the best compression performance, is superior to JPEG-LS, JPEG 2000, and CCSDS-122.0 in 26.5%, 28.5%, and 30.5%, respectively. When JPEG-LS and JPEG 2000 are paired with a spectral transform to exploit the spectral redundancy, they surpass CCSDS-123.0. For CCSDS-122.0 prepended with a spectral transform, the difference with respect to CCSDS-123.0 decreases to less than 4%.

However, the large spectral dimension of IASI data imposes a careful selection to exploit the spectral redundancy. Some spectral transforms, such as RKLT, may be unusable when the number of spectral channels is large due to its expensive computational complexity. In this case, a divide-and-conquer strategy like Multilevel Clustering RKLT may be a very effective approach [81], both in terms of coding performance and of computational cost. If lower computational requirements are demanded, RPOT and IWT are two alternatives, although they yield lower compression performance. RWA spectral transform provides also a competitive coding performance, improving on average, 7.1% and 6.9% with respect to RPOT and IWT.

We observed that compression techniques that already exploit the spectral redundancy by themselves also improve the coding performance when a spectral transform is applied. This is due to the large spectral dimension and the high redundancy present in IASI data. For HEVC and M-CALIC up to 11.11% and 8.59%, respectively, can be improved. In the case of CCSDS-123.0, gains are less significant.

Lossless compression can be an appropriate approach to compress IASI L1C products because all information is preserved, however the achieved compression ratios are limited. When larger compression ratios are requested, near-lossless or lossy compression is needed. Although using these approaches prevents a perfect reconstruction, it has been observed that some applications do not reduce their performance when certain level of distortion is introduced [26,27]. Near-lossless and lossy compression might be two reasonable compression approaches for IASI L1C products. In fact, data disseminated today through IASI Regional Data Service is not either the complete original scene, since only 800 (or less) out of the original 8461 channels are distributed for near-real time dissemination.

If a specific accuracy must be preserved in the reconstructed data, near-lossless compression is a proper strategy since the quality of the recovered data can be controlled by bounding the peak absolute error per pixel. In our experiments, we observed that using as small PAEs as 1 and 3, compression ratios of, respectively, 3:1 and 4:1 can already be achieved, while the data quality, measured in SNR Energy, still exceeds 65 dB. Table 9 illustrates how the compression ratio increases as the allowed PAE increases. Among the two compression techniques evaluated for near-lossless compression, M-CALIC produces larger compression ratios than JPEG-LS, achieving similar SNR Energy. The performance improvement of M-CALIC increases as PAE value increases, being, approximately, 25% more competitive for small PAEs and 70% for large PAEs. M-CALIC produces improved performance mainly due to the ability of M-CALIC to exploit the spectral redundancy present in the data, which is essential to achieve competitive coding performance. JPEG-LS is a 2D

data compression standard and is not able to exploit the spectral dimension. In addition, M-CALIC implements an arithmetic coder, while JPEG-LS uses Golomb codes, not being able to produce bit-rates below 1 bpppc. In this case, pairing JPEG-LS with a spectral transform would not be an appropriate approach, because the spectral transform forestalls the precise control over the peak absolute error.

For lossy compression, as happened for lossless compression, exploiting the spectral redundancy yields improved outcomes. Experimental results reveal that applying a spectral transform always produce better performance, with differences increasing for smaller compression ratios, growing from 5 to over 15 dB. Multilevel Clustering KLT also produces always the best coding performance among the spectral transforms. Applying POT achieves similar performance at compression ratios higher than 20:1, mainly due to the larger size of the side-information produced by Multilevel Clustering KLT. In turn, employing POT produces more competitive performance than DWT at large compression ratios, but the performance gets closer at small compression ratios.

With regard to comparison between coding standards JPEG 2000 and CCSDS-122.0, the former always yields improved coding performance, for both lossless and lossy. When paired with Multilevel Clustering KLT or POT spectral transforms, these two coding techniques show very similar behaviour for lossy compression, while there is a difference of about 5.5% for lossless.

Applying a compression process through either M-CALIC or Multilevel Clustering KLT + JPEG 2000, the compression schemes that produce the best coding performance for, respectively, near-lossless and lossy compression, takes less than 15 min, which may be acceptable in a near-real time scenario. The decompression stage, required to retrieve the reconstructed spectra at the receiver side, takes, approximately, 11 and 6 min for, respectively, near-lossless and lossy compression. Multilevel Clustering RKLT, the lossless scheme that produces the largest compression ratios, requires longer runtimes mainly due to the application of the Multilevel Clustering RKLT.

The analysis of the recovered data indicated that the covariance matrix of the reconstructed radiances for both Multilevel Clustering KLT + JPEG 2000 and M-CALIC is very similar to the covariance matrix of the original radiances. Although promising, more experiments are needed to determine whether these two coding schemes could become an alternative to Principal Component Compression for IASI data near-real time dissemination.

Finally, we note that the compression performance for IASI-A and IASI-B products is almost equivalent for lossless, near-lossless, and lossy compression.

5. Concluding Remarks

Infrared Atmospheric Sounding Interferometer (IASI) data acquired from MetOp-A and MetOp-B satellites are mostly disseminated as IASI L1C products. These products have provided over 10 years of continuous observations, and with the foreseen launch of MetOp-C satellite in October 2018, this long time series of climate variables will be further extended. So far, distribution of IASI L1C data has been mostly conducted through IASI Regional Data Service Level 1, where 800 (or less) channels out of the original 8461 spectral channels are disseminated in near-real time, however, a number of applications have recently identified the need to operate with the complete range of spectral channels. Transmission and storage of complete scenes with such large size pose a challenge, which might be alleviated thanks to data compression.

In this paper we put forward a comprehensive study of IASI L1C data compression. Lossless, near-lossless and lossy compression paradigms have been investigated on a representative set of 96 orbits selected over a full year, 48 orbits from each MetOp-A and MetOp-B satellite, 4 orbits per month. Two wavelet-based coding standards, JPEG 2000 and CCSDS-122.0, three prediction-based techniques, JPEG-LS, M-CALIC and CCSDS-123.0, and the most recent video coding standard, HEVC, have been evaluated. To account for the large spectral redundancy in IASI products, four spectral transforms, RKLT/KLT, IWT/DWT, RPOT/POT and RWA, have been combined with the six coding techniques and their performance assessed.

Experimental results suggest that Multilevel Clustering RKLT/KLT is an efficient approach in terms of both coding performance and computational complexity, providing the best outcome for lossless and lossy compression when paired with, respectively, M-CALIC and JPEG 2000. For near-lossless compression, M-CALIC is the best performing technique.

The covariance matrix of the reconstructed radiances for Multilevel Clustering KLT + JPEG 2000 and M-CALIC, the compression schemes that provides the best coding performance for lossy and near-lossless compression, respectively, are very similar to the covariance matrix of the original radiances, which suggests that the quality of the recovered data is still adequate for further processings. Although promising, more experiments are needed to determine whether the proposed compression schemes could become an alternative to Principal Component Compression for IASI data near-real time dissemination.

The reported analysis can contribute to deploy new methodologies to manage data from current and upcoming high spectral resolution infrared instruments and improve the quality of disseminated products as demanded in several application areas. It is important to note that the selected compression scheme must preserve the atmospheric information content and reduce the level of noise contained in the data, while achieving competitive compression ratios.

Acknowledgments: This work was supported in part by the Spanish Ministry of Economy and Competitiveness (MINECO), by the European Regional Development Fund (FEDER) and by the European Union (EU) under Grants TIN2012-38102-C03-03 and TIN2015-71126-R, and by the Catalan Government under Grant 2014SGR-691. The authors are indebted to Xavier Calbet (AEMET, Spain) for invaluable help with the assessment of the quality of the reconstructed data. The authors would like to thank the anonymous reviewers, whose constructive feedback was critical to improve the contents of the manuscript.

Author Contributions: All co-authors of this manuscript significantly contributed to all phases of the investigation. They contributed equally to the preparation, analysis, review and editing of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Prunet, P.; Thépaut, J.N.; Cassé, V. The information content of clear sky IASI radiances and their potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.* **1998**, *124*, 211–241.
2. Hilton, F.; Armante, R.; August, T.; Barnet, C.; Bouchard, A.; Camy-Peyret, C.; Capelle, V.; Clarisse, L.; Clerbaux, C.; Coheur, P.F.; et al. Hyperspectral Earth observation from IASI: Five years of accomplishments. *Bull. Am. Meteorol. Soc.* **2012**, *93*, 347–370.
3. Amato, U.; Cuomo, V.; Serio, C. Assessing the impact of radiometric noise on IASI performances. *Remote Sens.* **1995**, *16*, 2927–2938.
4. Pougatchev, N.; August, T.; Calbet, X.; Hultberg, T.; Oduleye, O.; Schlüssel, P.; Stiller, B.; Germain, K.S.; Bingham, G. IASI temperature and water vapor retrievals-error assessment and validation. *Atmos. Chem. Phys.* **2009**, *9*, 6453–6458.
5. George, M.; Clerbaux, C.; Hurtmans, D.; Turquety, S.; Coheur, P.F.; Pommier, M.; Hadji-Lazaro, J.; Edwards, D.P.; Worden, H.; Luo, M.; et al. Carbon monoxide distributions from the IASI/METOP mission: Evaluation with other space-borne remote sensors. *Atmos. Chem. Phys.* **2009**, *9*, 8317–8330.
6. Clarisse, L.; Coheur, P.F.; Prata, A.J.; Hurtmans, D.; Razavi, A.; Phulpin, T.; Hadji-Lazaro, J.; Clerbaux, C. Tracking and quantifying volcanic SO_2 with IASI, the September 2007 eruption at Jebel at Tair. *Atmos. Chem. Phys.* **2008**, *8*, 7723–7734.
7. Clerbaux, C.; Boynard, A.; Clarisse, L.; George, M.; Hadji-Lazaro, J.; Herbin, H.; Hurtmans, D.; Pommier, M.; Razavi, A.; Turquety, S.; et al. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmos. Chem. Phys.* **2009**, *9*, 6041–6054.
8. Wespes, C.; Hurtmans, D.; Clerbaux, C.; Santee, M.L.; Martin, R.V.; Coheur, P.F. Global distributions of nitric acid from IASI/MetOp measurements. *Atmos. Chem. Phys.* **2009**, *9*, 7949–7962.
9. Pommier, M.; Law, K.S.; Clerbaux, C.; Turquety, S.; Hurtmans, D.; Hadji-Lazaro, J.; Coheur, P.F.; Schlager, H.; Ancellet, G.; Paris, J.D.; et al. IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns. *Atmos. Chem. Phys.* **2010**, *10*, 10655–10678.

10. Grieco, G.; Masiello, G.; Matricardi, M.; Serio, C. Partially scanned interferogram methodology applied to IASI for the retrieval of CO, CO_2, CH_4 and N_2O. *Opt. Express* **2013**, *21*, 24753–24769.
11. Liuzzi, G.; Masiello, G.; Serio, C.; Venafra, S.; Camy-Peyret, C. Physical inversion of the full IASI spectra: Assessment of atmospheric parameters retrievals, consistency of spectroscopy and forward modelling. *J. Quant. Spectrosc. Radiat. Transf.* **2016**, *182*, 128–157.
12. Coheur, P.F.; Clarisse, L.; Turquety, S.; Hurtmans, D.; Clerbaux, C. IASI measurements of reactive trace species in biomass burning plumes. *Atmos. Chem. Phys.* **2009**, *9*, 5655–5667.
13. Turquety, S.; Hurtmans, D.; Hadji-Lazaro, J.; Coheur, P.F.; Clerbaux, C.; Josset, D.; Tsamalis, C. Tracking the emission and transport of pollution from wildfires using the IASI CO retrievals: Analysis of the summer 2007 Greek fires. *Atmos. Chem. Phys.* **2009**, *9*, 4897–4913.
14. CNES. *Dossier de Définition des Algorithmes IASI*; REF. IA-DF-0000-2006-CNE; CNES: Toulouse, France, 2009.
15. Masiello, G.; Serio, C. Dimensionality-reduction approach to the thermal radiative transfer equation inverse problem. *Geophys. Res. Lett.* **2004**, *31*, doi: 10.1029/2004GL019845.
16. Hultberg, T. IASI Principal Component Compression (IASI PCC) FAQ. Available online: http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=pdf_ipcc_faq&RevisionSelectionMethod=LatestReleased&Rendition=Web (accessed on 14 June 2017).
17. Atkinson, N.C.; Hilton, F.I.; Illingworth, S.M.; Eyre, J.R.; Hultberg, T. Potential for the use of reconstructed IASI radiances in the detection of atmospheric trace gases. *Atmos. Meas. Tech.* **2010**, *3*, 991–1003.
18. Camps-Valls, G.; Muñoz-Marí, J.; Gómez-Chova, L.; Guanter, L.; Calbet, X. Nonlinear statistical retrieval of atmospheric profiles from MetOp-IASI and MTG-IRS infrared sounding data. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1759–1769.
19. Masiello, G.; Serio, C.; Antonelli, P. Inversion for atmospheric thermodynamical parameters of IASI data in the principal components space. *Q. J. R. Meteorol. Soc.* **2012**, *138*, 103–117.
20. Serio, C.; Masiello, G.; Liuzzi, G. Demonstration of random projections applied to the retrieval problem of geophysical parameters from hyper-spectral infrared observations. *Appl. Opt.* **2016**, *55*, 6576–6587.
21. Motta, G.; Rizzo, F.; Storer, J.A. *Hyperspectral Data Compression*; Springer Science & Business Media: New York, NY, USA, 2006.
22. Huang, B. *Satellite Data Compression*; Springer Science & Business Media: New York, NY, USA, 2011.
23. Mercier, G.; Mouchot, M.; Cazuguel, G. Joint classification and compression of hyperspectral images. *IEEE Int. Remote Sens. Symp.* **1999**, *4*, 2035–2037.
24. Blanes, I.; Serra-Sagristà, J. Quality evaluation of progressive lossy-to-lossless remote-sensing image coding. In Proceedings of the 16th IEEE International Conference in Image Processing (ICIP), Cairo, Egypt, 7–9 November 2009; pp. 3709–3712.
25. García-Vílchez, F.; Muñoz-Marí, J.; Zortea, M.; Blanes, I.; González-Ruiz, V.; Camps-Valls, G.; Plaza, A.; Serra-Sagristà, J. On the impact of lossy compression on hyperspectral image classification and unmixing. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 253–257.
26. García-Sobrino, J.; Blanes, I.; Laparra, V.; Camps-Valls, G.; Serra-Sagristà, J. Impact of Near-Lossless Compression of IASI L1C data on Statistical Retrieval of Atmospheric Profiles. In Proceedings of the On-Board Payload Data Compression Workshop (OBPDC), Venice, Italy, 23–24 October 2014.
27. García-Sobrino, J.; Serra-Sagristà, J.; Laparra, V.; Calbet, X.; Camps-Valls, G. Statistical Atmospheric Parameter Retrieval Largely Benefits from Spatial-Spectral Image Compression. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2213–2224.
28. Qian, S.E.; Bergeron, M.; Cunningham, I.; Gagnon, L.; Hollinger, A. Near lossless data compression onboard a hyperspectral satellite. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 851–866.
29. ISO/IEC. *JPEG-LS Lossless and Near-Lossless Compression for Continuous-Tone Still Images*; ITU: Geneva, Switzerland, 1999.
30. JPEG-Committee. Standard JPEG2000, Document ISO/IEC 15444. Available online: http://www.jpeg.org/jpeg2000/workplan.html (accessed on 14 June 2017).
31. Magli, E.; Olmo, G.; Quacchio, E. Optimized onboard lossless and near-Lossless compression of hyperspectral data using CALIC. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 21–25.
32. Consultative Committee for Space Data Systems (CCSDS). Image Data Compression CCSDS 122.0-B-1; Blue Book, CCSDS, 2005. Available online: https://public.ccsds.org/Pubs/122x0b1c3.pdf (accessed on 14 June 2017).

33. Consultative Committee for Space Data Systems (CCSDS). Lossless Multispectral & Hyperspectral Image Compression CCSDS 123.0-B-1; Blue Book, CCSDS, 2012. Available online: https://public.ccsds.org/Pubs/123x0b1ec1.pdf (accessed on 14 June 2017).
34. ISO/IEC. High Efficiency Coding and Media Delivery in Heterogeneous Environments—Part 2: High Efficiency Video Coding, 2013. Available online: http://www.iso.org/iso/catalogue_detail.htm?csnumber=67660 (accessed on 14 June 2017).
35. Chang, L.; Cheng, C.M.; Chen, T.C. An efficient adaptive KLT for multispectral image compression. In Proceedings of 4th IEEE Southwest Symposium on Image Analysis and Interpretation, Austin, TX, USA, 2–4 April 2000.
36. Salomon, D. *Data Compression: The Complete Reference*; Springer Science & Business Media: New York, NY, USA, 2004.
37. Blanes, I.; Serra-Sagristà, J. Pairwise orthogonal transform for spectral image coding. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 961–972.
38. Amrani, N.; Serra-Sagristà, J.; Laparra, V.; Marcellin, M.W.; Malo, J. Regression Wavelet Analysis for Lossless Coding of Remote-Sensing Data. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5616–5627.
39. EUMETSAT. IASI Mission. Available online: http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/IASI/index.html (accessed on 14 June 2017).
40. The World Data Center for Remote Sensing of the Atmosphere (WDC-RSAT). IASI (Infrared Atmospheric Sounding Interferometer). Available online: http://andromeda.caf.dlr.de/sensors/iasi (accessed on 14 June 2017).
41. EUMETSAT. MetOp Mission. Available online: http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/index.html (accessed on 14 June 2017).
42. EUMETSAT. EUMETSAT News. Available online: http://www.eumetsat.int/website/home/News/DAT_3304789.html (accessed on 14 June 2017).
43. Step, A.M. The EUMETSAT Polar System. *ESA Bull.* **2006**, *127*, 19.
44. Klaes, K.D.; Cohen, M.; Buhler, Y.; Schlëssel, P.; Munro, R.; Engeln, A.; Clérigh, E.; Bonekamp, H.; Ackermann, J.; Schmetz, J.; et al. An introduction to the EUMETSAT polar system. *Bull. Am. Meteorol. Soc.* **2007**, *88*, 1085–1096.
45. Chalon, G.; Cayla, F.; Diebel, D. IASI: An advanced sounder for operational meteorology. In Proceedings of the 52nd International Astronautical Congress (IAF), Toulouse, France, 1–5 October 2001.
46. CNES. *Spécification Technique de Besoin du Logiciel Opérationnel IASI*; REF. IA-SB-2100-9462-CNE; CNES: Toulouse, France, 2006.
47. EUMETSAT. *IASI Measurement and Verification Data*; REF. IA-ID-1000-6477-AER; EUMETSAT: Darmstadt, Germany, 2010.
48. Hébert, P.; Blumstein, D.; Buil, C.; Carlier, T.; Chalon, G.; Astruc, P.; Clauss, A.; Siméoni, D.; Tournier, B. IASI instrument: Technical description and measured performances. In Proceedings of the 5th International Conference on Space Optics, Toulouse, France, 30 March–2 April 2014; Volume 554, pp. 49–56.
49. Simeoni, D.; Astruc, P.; Miras, D.; Alis, C.; Andreis, O.; Scheidel, D.; Degrelle, C.; Nicol, P.; Bailly, B.; Guiard, P.; et al. Design and development of IASI instrument. In Proceedings of the SPIE 49th Annual Optical Science and Technology Meeting, Denver, Colorado, 2–6 August 2004; pp. 208–219.
50. EUMETSAT. *IASI Level 1: Product Guide*; REF. EUM/OPS-EPS/MAN/04/0032; EUMETSAT: Darmstadt, Germany, 2012.
51. Tournier, B.; Blumstein, D.; Cayla, F.; Chalon, G. IASI level 0 and 1 processing algorithms description. In Proceedings of the 12th International TOVS Study Conference (ITSC-XII), Lorne, Victoria, Australia, 27 February–5 March 2002.
52. EUMETSAT. *IASI Level 2: Product Guide*; REF. EUM/OPS-EPS/MAN/04/0033; EUMETSAT: Darmstadt, Germany, 2012.
53. ESA. IASI Data Processing Chain. Available online: http://www.esa.int/Our_Activities/Observing_the_Earth/The_Living_Planet_Programme/Meteorological_missions/MetOp/Data_processing_chain (accessed on 14 June 2017).
54. EUMETCast. EUMETCast Website. Available online: http://www.eumetsat.int/website/home/Data/DataDelivery/EUMETCast/index.html (accessed on 14 June 2017).
55. PODAAC. PODAAC Website. Available online: https://podaac.jpl.nasa.gov/ (accessed on 14 June 2017).

56. CEDA. CEDA Website. Available online: http://catalogue.ceda.ac.uk/ (accessed on 14 June 2017).
57. EUMETSAT. Central Operations Report for the Period January to June 2016. EUM/OPS/REP/16/866335, v1A, 2016. Available online: http://www.eumetsat.int/website/home/Data/ServiceStatus/CentralOperationsReports/index.html (accessed on 14 June 2017).
58. CEDA Support, 2017. Private correspondence.
59. EUMETCast. IASI Regional Data Service Level 1 Website. Available online: http://navigator.eumetsat.int/discovery/Start/DirectSearch/Extended.do?f(r0)=EO:EUM:DAT:METOP:EARS-IASI (accessed on 14 June 2017).
60. EUMETSAT. IASI PCA-Based Compression Package. Available online: https://nwpsaf.eu/site/software/iasi-pca/ (accessed on 14 June 2017).
61. Hultberg, T.; August, T.; Atkinson, N.C.; Smith, F. IASI PC compression—Searching for signal in the residuals. In Proceedings of the ECMWF/EUMETSAT NWP-SAF Workshop on Efficient Representation of Hyper-Spectral Infrared Satellite Observations, Exeter, UK, 5–7 November 2013.
62. Hilton, F.; Collard, A.D. *Recommendations for the Use of Principal Component-Compressed Observations from Infrared Hyperspectral Sounders*; Met Office Forecasting R&D Technical Report; Met Office, Exeter, UK, 2009; Volume 536.
63. Atkinson, N.; Ponsard, C.; Hultberg, T. AAPP Enhancements for the EARS-IASI Service. Available online: https://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=PDF_CONF_P55_S8_39_ATKINSON_P&RevisionSelectionMethod=LatestReleased&Rendition=Web&usg=AFQjCNHm0O2USellr5iHpsff7Y0l17EBRQ&sig2=Ad6FE7ZHHZUpxV03AJZQTA (accessed on 14 June 2017).
64. Antonelli, P.; Revercomb, H.E.; Sromovsky, L.A.; Smith, W.L.; Knuteson, R.O.; Tobin, D.C.; Garcia, R.K.; Howell, H.B.; Huang, H.L.; Best, F.A. A principal component noise filter for high spectral resolution infrared measurements. *J. Geophys. Res. Atmos.* **2004**, *109*, D23102.
65. Hultberg, T. IASI Principal Components—Experiences at EUMETSAT, 2011. Available online: ftp://193.17.11.194/pub/EPS/out/Hultberg/IasiPCC/EUMETSAT_PCC_Presentation_at_ECMWF_pdf.pdf (accessed on 14 June 2017).
66. Golomb, S.W. Run-length encodings. *IEEE Trans. Inf. Theory* **1966**, *12*, 399–401.
67. Witten, I.H.; Neal, R.M.; Cleary, J.G. Arithmetic coding for data compression. *Commun. ACM* **1987**, *30*, 520–540.
68. Consultative Committee for Space Data Systems (CCSDS). Spectral Pre-Processing Transform for Multispectral & Hyperspectral Image Compression, 2017. Available online: http://cwe.ccsds.org/fm/Lists/Projects/DispFormDraft.aspx?ID=562&Source=http://cwe.ccsds.org/fm/Lists/Projects/AllOpenChartersWithDraftProjects.aspx (accessed on 14 June 2017).
69. Blanes, I.; Serra-Sagristà, J. Cost and Scalability Improvements to the Karhunen-Loêve Transform for Remote-Sensing Image Coding. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2854–2863.
70. EUMETSAT. IASI Level 1 Product Formats and Dissemination. EUM/OPS-EPS/MAN/04/0032, 2012. Available online: http://oiswww.eumetsat.org/WEBOPS/eps-pg/IASI-L1/IASIL1-PG-6ProdFormDis.htm (accessed on 14 June 2017).
71. Clunie, D.A. JPEG-LS Software. Available online: http://www.dclunie.com/jpegls.html (accessed on 14 June 2017).
72. Taubman, D.S. Kakadu Software. Available online: http://www.kakadusoftware.com/ (accessed on 14 June 2017).
73. Magli, E. M-CALIC Software. Available online: http://www1.tlc.polito.it/oldsite/sas-ipl/download.php (accessed on 14 June 2017).
74. GICI-UAB. TER Software. Available online: http://gici.uab.cat/GiciWebPage/downloads.php#ter (accessed on 14 June 2017).
75. GICI-UAB. EMPORDA Software. Available online: http://gici.uab.cat/GiciWebPage/downloads.php#emporda (accessed on 14 June 2017).
76. Fraunhofer-HHI. HEVC Software. Available online: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/ (accessed on 14 June 2017).
77. GICI-UAB. Spectral Transform Software. Available online: http://gici.uab.cat/GiciWebPage/downloads.php#spectral (accessed on 14 June 2017).

78. GICI-UAB. Pairwise Orthogonal Transform (POT) Software. Available online: http://gici.uab.cat/GiciWebPage/downloads.php#pot (accessed on 14 June 2017).
79. GICI-UAB. Regression Wavelet Analysis (RWA) Transform Software. Available online: http://gici.uab.cat/GiciWebPage/downloads.php#RWA (accessed on 14 June 2017).
80. Magli, E. Multiband lossless compression of hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1168–1178.
81. Blanes, I.; Serra-Sagristà, J.; Marcellin, M.W.; Bartrina-Rapesta, J. Divide-and-Conquer Strategies for Hyperspectral Image Processing: A Review of Their Benefits and Advantages. *IEEE Signal Process. Mag.* **2012**, *29*, 71–81.

Chapter 4

Statistical Atmospheric Parameter Retrieval Largely Benefits From Spatial-Spectral Image Compression

Statistical Atmospheric Parameter Retrieval Largely Benefits From Spatial–Spectral Image Compression

Joaquín García-Sobrino, Joan Serra-Sagristà, *Senior Member, IEEE*, Valero Laparra, Xavier Calbet, and Gustau Camps-Valls, *Senior Member, IEEE*

***Abstract*— The infrared atmospheric sounding interferometer (IASI) is flying on board of the Metop satellite series, which is part of the EUMETSAT Polar System. Products obtained from IASI data represent a significant improvement in the accuracy and quality of the measurements used for meteorological models. Notably, the IASI collects rich spectral information to derive temperature and moisture profiles, among other relevant trace gases, essential for atmospheric forecasts and for the understanding of weather. Here, we investigate the impact of near-lossless and lossy compression on IASI L1C data when statistical retrieval algorithms are later applied. We search for those compression ratios that yield a positive impact on the accuracy of the statistical retrievals. The compression techniques help reduce certain amount of noise on the original data and, at the same time, incorporate spatial–spectral feature relations in an indirect way without increasing the computational complexity. We observed that compressing images, at relatively low bit rates, improves results in predicting temperature and dew point temperature, and we advocate that some amount of compression prior to model inversion is beneficial. This research can benefit the development of current and upcoming retrieval chains in infrared sounding and hyperspectral sensors.**

***Index Terms*— Infrared atmospheric sounding interferometer (IASI), JPEG 2000, kernel methods, lossy compression, M-CALIC, near-lossless compression, spectral transforms, statistical retrieval.**

I. Introduction

TEMPERATURE and water vapor atmospheric profiles are essential meteorological parameters for weather forecasting and atmospheric chemistry studies. Observations from high spectral resolution infrared sounding instruments on board of satellites provide unprecedented accuracy and vertical resolution of temperature and water vapor profiles. However, it is not trivial to retrieve the full information content from radiation measurements. Accordingly, improved retrieval algorithms are desirable to achieve optimal performance for existing and future infrared sounding instrumentation.

A. Atmospheric Parameter Retrieval With IASI

The use of Metop data in numerical weather prediction (NWP) accounts for 40% of the impact of all space-based observations in NWP forecasts. The infrared atmospheric sounding interferometer (IASI) sensor is implemented on the Metop satellite series. Products obtained from IASI data are a significant improvement in the quality of the measurements used for meteorological models. In particular, IASI collects rich spectral information to derive temperature and moisture profiles, which are essential to the understanding of weather and to derive atmospheric forecasts. The sensor provides infrared spectra with high resolution between 645 and 2760 cm^{-1}, from which temperature and humidity profiles with high vertical resolution and accuracy are derived. Additionally, these spectra are used for the determination of trace gases such as ozone, nitrous oxide, carbon dioxide, and methane, as well as land and sea surface temperatures and emissivities and cloud properties [1], [2].

EUMETSAT, NOAAA, NASA, and other operational agencies are continuously developing product-processing facilities to obtain L2 atmospheric profile products from infrared hyperspectral radiance instruments, such as the IASI. One of the retrieval techniques commonly used in L2 processing is based on linear regression (LR), which is a valuable and very computationally efficient method. It consists of performing a canonical least squares LR on top of the data projected onto the first principal components or empirical orthogonal functions (EOFs) of the measured brightness temperature spectra (or radiances) and the atmospheric state parameters. To further improve the results of this scheme for retrieval, nonlinear statistical retrieval methods can be applied as an efficient alternative to more costly optimal estimation (OE) schemes. These methods have proved to be valid in retrieval of temperature, dew point temperature (humidity), and ozone atmospheric profiles when the original data are used [3].

Manuscript received January 10, 2016; revised June 16, 2016 and October 18, 2016; accepted November 29, 2016. Date of publication January 16, 2017; date of current version February 24, 2017. This work was supported in part by the Spanish Ministry of Economy and Competitiveness, in part by the European Regional Development Fund under Grant TIN2015-71126-R and Grant TIN2012-38102-C03-00, in part by the Catalan Government under Grant 2014SGR-691, and in part by the European Research Council under Consolidator Grant SEDAL ERC-2014-CoG 647423.

J. García-Sobrino and J. Serra-Sagristà are with the Group on Interactive Coding of Images, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain (e-mail: joaquin.garcia.sobrino@deic.uab.cat; joan.serra@uab.cat).

V. Laparra and G. Camps-Valls are with the Image Processing Laboratory, Universitat de València, 46980 València, Spain (e-mail: valero.laparra@uv.es; gustau.camps@uv.es).

X. Calbet is with AEMET, 28071 Madrid, Spain (e-mail: xcalbeta@aemet.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2016.2639099

B. Impact of Hyperspectral Image Compression

Given the orbit time of Metop satellites (101 min), the large spectral resolution (8461 components), and spatial resolution (60×1530 samples) of IASI L1C products and that there are two active Metop satellites (A and B), about 41.3 GB of IASI L1C data are produced daily. This large volume of remote sensing data asks for efficient compression systems for both storage and transmission. Remote sensing data compression is a mature field attracting interest of space administrations, public bodies, and private companies. Lossless, near-lossless, and lossy coding techniques are already in use in on-going satellite missions and have been adopted in several current standards. However, the *impact* of the (near-lossless and lossy) compression on the radiances can compromise the quality of posterior products, such as classification and detection maps, or biophysical –geophysical parameter estimates, as is the case of this paper.

1) Impact on Image Classification: Effects of lossy data compression have been analyzed mostly in the scope of classification applications. Mercier *et al.* [4] proposed a vector quantization for joint classification and compression of hyperspectral data. This approach proved very competitive for compressing images, achieving compression ratios of approximately 70:1 and classification performance was not markedly reduced. Analogous results were achieved in [5]–[7], where high compression ratios yielded high classification accuracy.

Especially interesting is the observation that this behavior is more apparent when a spatial–spectral wavelet transform is applied. Penna *et al.* [8] reported a stimulating consideration: classification performance was not closely linked to the distortion levels introduced in the image, meaning that, in general, the best classification accuracy was not achieved after applying the coding algorithms/settings that yielded the best rate–distortion performance.

Du and Fowler [9] proposed a scheme based on principal component analysis (PCA) deployed in JPEG 2000 to provide spectral decorrelation. The proposed scheme produced competitive data analysis performance in terms of information preservation in an anomaly detection task. Blanes *et al.* [10] conducted a study on compression of hyperspectral images through 3-D JPEG 2000, where supervised and unsupervised classification of the recovered images was evaluated. The results reflect that classification accuracy is still reliable after the compression stage. Similar results were achieved in [11], where the proposed lossy and near-lossless compression algorithms for hyperspectral images yielded good results for hard classification, spectral unmixing, and anomaly detection.

2) Impact on Image Unmixing: The impact of lossy compression on linear spectral unmixing and nonlinear hyperspectral image classification using support vector machines (SVMs) was investigated in [12]. The experimental results suggest that for some compression techniques, a higher compression ratio may produce more accuracy in classification results. Chang [13] and Du and Chang [14] have reported that lossy compression can preserve the needed information to estimate endmember fractional abundances in linear spectral unmixing even at low bit rates (BRs). Particularly interesting is the observation reported in [15] and [16], where it is noted that the joint use of spectral and spatial information in some stages of the linear spectral unmixing chain can produce smaller image reconstruction errors.

3) Impact on Biophysical–Geophysical Parameter Retrieval: It is worth noting that to the authors' knowledge, there are no works analyzing the impact of spatial–spectral data compression on the bio physical–geophysical parameter retrieval. We reported preliminary results in [17], which are further extended here with more data and in-depth analysis. The experimental results revealed an interesting issue: the performance of statistical retrieval methods substantially improved after near-lossless compression. This observation leads us to define the main goal of this paper.

C. Goal and Main Findings

This paper is concerned with the performance of different (near-lossless and lossy) compression techniques and different (linear and nonlinear) statistical retrieval algorithm on IASI L1C data. We will touch upon M-CALIC [18] coding technique and upon JPEG 2000 standard [19], [20]. In this latter case, two spectral transforms will be employed to better exploit the high spectral redundancy inherent in IASI data, namely, pairwise orthogonal transform (POT) [21], [22] and discrete wavelet transform (DWT), which have proved to achieve good data decorrelation in multi-, hyper-, and ultraspectral images. Then, two different statistical retrieval algorithms, LR and kernel ridge regression (KRR) [23], will be evaluated in the retrieval of temperature and dew point temperature (humidity) profiles from the recovered decoded images.

In short, the objective of this paper is to provide an outlook of the effects of near-lossless and lossy compression of IASI L1C data when statistical retrieval methods are employed to retrieve physical information from the reconstructed images. An important conclusion of the work is that retrieval methods may operate on data that do not have the same quality precision as that originally provided by IASI L1C products. Actually, though it may appear counterintuitive at first glance, some amount of compression can improve the accuracy in atmospheric parameter profile estimation. This analysis may have a deep impact on other infrared sounding instruments and hyperspectral sensors, both currently flying (e.g., AIRS) or upcoming (e.g., MTG-IRS). The observed effect has been widely exploited in signal and image processing problems (see [24]–[27]), and can be explained by noting that compression performs a sort of signal denoising and it generally constitutes an indirect way to include spatial–spectral feature relations, which ultimately help to pixelwise retrieval algorithms. This turns to be an extremely simple yet an effective way to improve retrieval results and comes at the price of including a compression step before retrievals, where the compression ratio needs to be properly adjusted.

D. Outline of This Paper

The remainder of this paper is organized as follows. Section II details the proposed approach, describing the different compression paradigms and statistical retrieval methods.

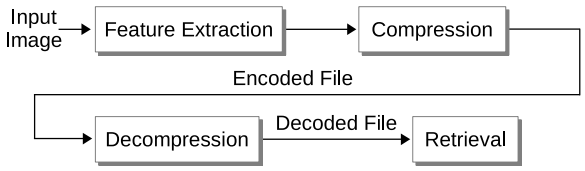


Fig. 1. Proposed sequential approach: after a feature extraction process, compression is applied first, followed by the statistical parameter retrieval.

Section III provides the experimental results. Section IV provides some discussion about the reasons behind the observed phenomena. Conclusions and further work are drawn in Section V.

II. Methods

This section introduces the experimental setting and methods. The proposed sequential approach includes first a compression stage and then a stastistical parameter retrieval. Fig. 1 illustrates the adopted scheme. Two different compression approaches will be considered. On the one hand, the maximum absolute error per pixel introduced in the reconstructed image shall be bounded using a near-lossless paradigm. On the other hand, the overall BR shall be controlled using a lossy approach. Similarly, two different statistical retrieval algorithms will be studied: a standard least squares LR and a KRR method, which have provided a very good performance in [3]. At a very early stage, a feature extraction/selection process is conducted to discard some less useful components, leading from the original 8461 spectral components on IASI L1C products to 4699 spectral components [3].

A. Near-Lossless and Lossy Compression Techniques

Lossless compression techniques allow one to completely reconstruct all the original data in the reconstructed image after the decoding process is performed. However, compression ratios for lossless coding are limited, usually not going beyond 2:1. Near-lossless and lossy compression removes some information during the coding process, preventing the recovery of all the original data, but allowing one to improve the compression ratios as the introduction of distortion becomes larger. These approaches rely on the assumption that the quality of the data after the coding process is still appropriate for the intended specific use.

Two recognized coding techniques, M-CALIC for near-lossless and JPEG 2000 for lossy compression, are discussed in this section.

1) Near-Lossless Compression: M-CALIC [18] has been selected for near-lossless compression since it provides a competitive performance when spectral redundancy is high, which is the case of IASI L1C products. M-CALIC is a lossless and near-lossless compression technique based on context-based adaptive lossless image coding (CALIC) [28]. CALIC was designed as a proposal for the ISO standard for lossless and near-lossless compression of 2-D images, and although it was not finally selected because of its higher computational complexity compared with LOCO-I [29] algorithm, CALIC provides a higher coding performance.

Fig. 2. M-CALIC basic architecture.

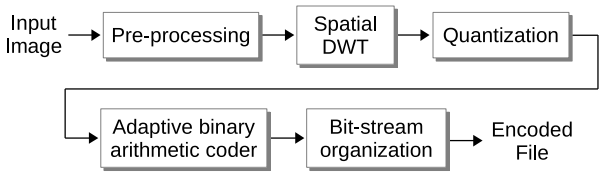


Fig. 3. JPEG 2000 basic architecture.

M-CALIC uses a multicomponent spectral predictor, along with optimized model parameters and optimization thresholds. The algorithm exploits correlation among components by employing the two previous components of the current one in the prediction, which allows capturing most of the spectral correlation. The intensity of the pixel to be encoded is estimated based mostly on the intensity of colocated pixels in previous components, i.e., pixels at the same spatial position but in components with a lower index. In fact, it is the prediction error that is encoded. For near-lossless compression, these prediction errors are quantized, ensuring that all distortion peak absolute errors (PAEs) fall below a given threshold. Fig. 2 illustrates the basic architecture of M-CALIC.

2) Lossy Compression: JPEG 2000 is an international standard developed by the *Joint Photographic Experts Group (JPEG)*. It was intended as the successor of JPEG due to the JPEG limitations in many application areas, such as mobile communications, medical diagnostic imaging, enhanced Internet browsing, digital cinema, and multimedia. JPEG 2000 provides a wide range of features in a single compressed bitstream for a large amount of applications. Fig. 3 illustrates the basic architecture of JPEG 2000.

Two spectral transforms, namely, POT and DWT, are applied along with JPEG 2000 to account for the large correlation among the spectral components.

POT [21], [22] relies on the implementation of a divide-and-conquer strategy to the Karhunen–Loève transform (KLT), where the composition of smaller KLT transforms produces the resulting transform. Each composition of the KLT is computed from only two image components. In a classic KLT, every component is decorrelated with each other, irrespective of how much energy they share. In its turn, POT implements a mechanism that decorrelates portions with high shared energy while ignoring the other portions. POT works in a multilevel mode, where a two-component KLT transform is computed for every pair of consecutive components at each level. This allows accumulating most of the image energy in the first components since each composition is computed as a classic KLT. Most of the signal energy is grouped into one of the two resulting components allowing most of the image energy to flow across the composition of transforms up to the last level.

TABLE I
EXAMPLES OF IASI L1C IMAGES USED IN THE EXPERIMENTS. TECHNICAL NAMES, IDENTIFIERS, AND SIZES ARE PROVIDED

Technical name	Identifier	Size ($x \times y \times z$)
IASI_xxx_1C_M01_20130817004753Z_20130817022952Z_N_O_20130817013849Z	IASI_20130817004753Z	60×1530×8461
IASI_xxx_1C_M01_20130817041457Z_20130817055656Z_N_O_20130817050513Z	IASI_20130817041457Z	60×1530×8461
IASI_xxx_1C_M01_20130817055657Z_20130817073856Z_N_O_20130817064707Z	IASI_20130817055657Z	60×1530×8461
IASI_xxx_1C_M01_20130817073857Z_20130817092056Z_N_O_20130817082957Z	IASI_20130817073857Z	60×1530×8461

The DWT can be seen as successive levels of decomposition that allow decorrelating the processed data. Every time the transform is applied, the signal is decomposed into two sets of coefficients, the low frequencies (L) and the high frequencies (H) sets. The low frequency set is a representation of the input signal at a coarser resolution, while the high frequency set represents details.

B. Statistical Parameter Retrieval

We aim at studying the impact of image compression on the retrieval of atmospheric profiles using standard and modern statistical retrieval algorithms. In particular, we will pay attention to the KRR algorithm [23], [30], which generalizes least squares LR. KRR has shown good performance in the prediction of such variables using IASI L1 data in [3]. We analyze the performance for both LR and KRR at different compression ratios.

Let us now fix the notation and review the formulation of KRR for regression. The KRR, also known as least squares SVM, is the kernel version of a regularized LR. Let $\mathbf{x}_i \in \mathbb{R}^N$ (spectra) and $y_i \in \mathbb{R}^M$ (state vectors), where $i = 1, \ldots, n$ indicates the index of the n training samples. We want to perform a linear least squares regression in a Hilbert space $\mathcal{H}$, of very high (possibly infinite) dimensionality $D_{\mathcal{H}}$, where samples have been mapped through a mapping $\boldsymbol{\phi}(\mathbf{x}_i)$. In the matrix notation, the model is given by $\mathbf{Y} = \boldsymbol{\Phi}\mathbf{W}+\mathbf{b}$, and under the assumption of an additive $i.i.d.$ noise model, $\hat{\mathbf{Y}} = \mathbf{Y} + \mathbf{E}$ with Gaussian noise $\mathbf{E} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2\mathbf{I})$ of zero mean and standard deviation σ_n.

Then, as in the regularized LR setting, we want to minimize the regularized squared loss function $\mathcal{L}_p = \|\mathbf{Y} - \boldsymbol{\Phi}\mathbf{W}\|^2 + \lambda\|\mathbf{W}\|^2$ with respect to model weights $\mathbf{W}$. In what follows, we deliberately drop the bias term. Therefore, taking derivatives with respect to $\mathbf{W}$ and equating them to zero gives $\mathbf{W} = (\boldsymbol{\Phi}^\top\boldsymbol{\Phi}+\lambda\mathbf{I})^{-1}\boldsymbol{\Phi}^\top\mathbf{Y}$, where $\boldsymbol{\Phi}$ is the matrix of mapped samples, $[\boldsymbol{\phi}(\mathbf{x}_1)^\top, \boldsymbol{\phi}(\mathbf{x}_2)^\top, \ldots, \boldsymbol{\phi}(\mathbf{x}_n)^\top]$, whose size is $n \times D_{\mathcal{H}}$. Note that this problem is not solvable as the inverse runs on matrix $\boldsymbol{\Phi}\boldsymbol{\Phi}^\top$, which is of size $D_{\mathcal{H}} \times D_{\mathcal{H}}$, and $\boldsymbol{\Phi}$ is in principle unknown. Here, we apply the representer theorem by which we can express the solution as a linear combination of mapped samples, $\mathbf{W} = \boldsymbol{\Phi}^\top\boldsymbol{\alpha}$, and then the solution is expressed as a function of the dual weights $\boldsymbol{\alpha}$ (one per sample), $\boldsymbol{\alpha} = (\boldsymbol{\Phi}\boldsymbol{\Phi}^\top + \lambda\mathbf{I})^{-1}\mathbf{Y}$. Note that now the problem is solvable as we need only to compute the inverse of the (regularized) Gram matrix $\mathbf{K} = \boldsymbol{\Phi}\boldsymbol{\Phi}^\top$ of size $n \times n$. Even though the mapping is unknown, one can replace this inner product matrix with a similarity matrix between samples, which is known as the *kernel matrix* $\mathbf{K}$.

We finally need to show that we never actually require access to the mapped feature vectors, which could be of infinite dimension. What we need in practice is the predicted value for a new matrix of test examples, $\mathbf{X}_*$. This is computed by projecting it onto the solution $\mathbf{W}$

$$\hat{\mathbf{Y}}_* = \boldsymbol{\Phi}_*\mathbf{W} = \boldsymbol{\Phi}_*\boldsymbol{\Phi}^\top\boldsymbol{\alpha} = \mathbf{K}_*\boldsymbol{\alpha} \tag{1}$$

where the matrix $\mathbf{K}_*$ contains the similarities between all test and training samples, whose entries are $K(\mathbf{x}_i, \mathbf{x}_j) = \boldsymbol{\phi}(\mathbf{x}_i)\boldsymbol{\phi}(\mathbf{x}_j)^\top$. The important message here is of course that we only need access to the kernel function K. Examples of typical kernel functions are the linear $K(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{x}_i^\top\mathbf{x}_j$, the polynomial $K(\mathbf{x}_i,\mathbf{x}_j) = (\mathbf{x}_i^\top\mathbf{x}_j + 1)^d$, and the one used in our implementation, the Gaussian function [radial basis function (RBF)] kernel $K(\mathbf{x}_i,\mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2))$. Therefore, in KRR, two free parameters are tuned: the regularization parameter λ and the kernel parameter σ. We used a cross-validation strategy for their optimization. For the interested reader, a MATLAB implementation of KRR and other regression algorithms can be found at http://isp.uv.es/soft_regression.html.

III. EXPERIMENTAL RESULTS

This section is devoted to report the experimental results. First, we will describe the IASI L1C data used in the experiments. Then, we will pay attention to the compression results obtained by a near-lossless (M-CALIC) and a lossy (JPEG 2000) coding technique, and finally, we will evaluate the impact of compression on linear (LR) and nonlinear (KRR) retrieval of atmospheric parameters. Due to restrictions in the space, we show results for only one image and one biophysical variable in most of the experiments (see Supplementary Material for the results for all the other images).[1]

A. Data Collection

To conduct the experiments, several IASI L1C images were used. Table I provides characteristics for four images of this data set. All the four images were obtained from the IASI-B instrument (implemented on MetOp-B satellite), are 16 bits per pixel per component (bpppc), and are stored as signed integers.

B. Compression Results

All the considered IASI L1C images have gone through a feature (component) selection process before carrying out the compression stage. We selected a subset of spectral components based on the minimization of measurement errors [3], which discards components with negative

[1] http://isp.uv.es/coding_retrieval.html

TABLE II
M-CALIC COMPRESSION PERFORMANCE IN BR (MEASURED IN BITS PER PIXEL PER COMPONENT, LOWER IS BETTER) AND SNR ENERGY (MEASURED IN dB, HIGHER IS BETTER)

IASI L1C	$\delta = 1$		$\delta = 3$		$\delta = 7$		$\delta = 15$		$\delta = 31$		$\delta = 63$		$\delta = 127$		$\delta = 255$		$\delta = 511$		$\delta = 1023$		$\delta = 2047$	
	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR	BR	SNR
IASI_20130817004753Z	5.07	75.02	3.89	67.24	2.88	60.59	2.01	54.38	1.31	48.37	0.78	42.68	0.41	37.36	0.22	31.74	0.13	25.70	0.09	19.63	0.07	13.75
IASI_20130817041457Z	5.07	75.49	3.88	67.71	2.87	61.06	2.00	54.85	1.30	48.84	0.77	43.16	0.40	37.83	0.21	32.18	0.13	26.11	0.08	20.05	0.07	14.12
IASI_20130817055657Z	5.05	76.14	3.86	68.36	2.85	61.71	1.98	55.50	1.29	49.49	0.76	43.81	0.39	38.48	0.21	32.82	0.12	26.73	0.08	20.62	0.07	14.73
IASI_20130817073857Z	5.08	75.75	3.89	67.97	2.88	61.31	2.01	55.10	1.31	49.09	0.78	43.40	0.40	38.07	0.21	32.44	0.13	26.42	0.08	20.34	0.07	14.35
Average	5.06	75.55	3.88	67.77	2.87	61.12	2.00	54.91	1.30	48.90	0.77	43.22	0.40	37.90	0.21	32.26	0.13	26.20	0.08	20.12	0.07	14.18

radiance and high levels of estimated noise by applying thresholds on the estimated noise standard deviation per component.

For near-lossless compression, we used the M-CALIC software [31]. The tested images have been compressed using the default parameters configuration. Eleven different PAEs, $\delta \in \{1, 3, 7, 15, 31, 63, 127, 255, 511, 1023, 2047\}$, are used in our experiments. Table II shows the performance for M-CALIC compressor on IASI L1C products. Of course, the higher the PAE, the lower the BR (or, equivalently, the higher the compression ratio).

For lossy compression, we used the well-known JPEG 2000 Part 2 standard with two spectral transforms, DWT and POT. We analyze three different compression schemes: 1) only a spatial transform (DWT); 2) only a spectral transform (either DWT or POT); and 3) both a spectral transform (either DWT or POT) and a spatial transform (DWT). When applicable, schemes are evaluated considering five and ten levels of DWT spectral transform, and one, three, and five levels of DWT spatial transform. Seventeen target BRs are analyzed, uniformly distributed from 0.001 to 2 bpppc. Kakadu software [32] was used for JPEG 2000 experiments. The standard provides a multicomponent extension in its Part 2 [33]. In order to compute the POT spectral transform, the POT software [34] was used.

Fig. 4 shows the performance of the different lossy compression schemes for the product IASI_20130817055657Z. The results for the other images are practically identical and hence not reported here for the sake of brevity. Actually, due to space constraints, only the best configuration for each scheme is plotted: 1) *spatial DWT five levels*; 2) *spectral POT*; and 3) both *spectral POT + spatial DWT five levels* and *spectral DWT ten levels + spatial DWT five levels*. To facilitate the performance comparison between near-lossless and lossy compression, a curve reporting the performance of M-CALIC is also plotted in Fig. 4. For M-CALIC, only compression results for PAEs between 15 and 2047 are plotted, because lower PAEs imply BRs larger than those commonly employed for lossy compression.

One can see that, as expected, the best performance is achieved when a spectral transform followed by a spatial transform is applied. The performance comparison between POT and DWT as spectral transforms is very similar, and selecting one or the other may depend not only on the coding performance but also on complexity issues. In addition,

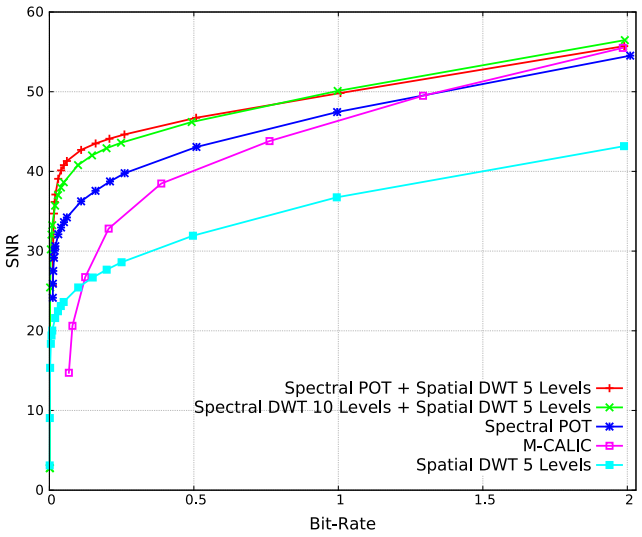


Fig. 4. Rate–distortion performance for near-lossless and lossy compression. The results show SNR energy (in db, higher is better) versus BR (in bits per pixel per component) for image IASI_20130817055657Z.

as expected, M-CALIC starts achieving competitive coding performance only at medium to high BRs.

C. Retrieval Assessment

This section presents the results of estimating physical variables (temperature and dew point temperature) using compressed IASI L1C images. Results are reported for image IASI_20130817055657Z for moisture prediction (related dew point temperature). Once more, results for the other images are practically identical, and hence not reported here for the sake of brevity. Similarly, reciprocal performance is obtained for temperature estimation; these results are not reported either for the sake of conciseness. For Supplementary Material, we encourage the reader to browse through http://isp.uv.es/coding_retrieval.html, which reports detailed results for four IASI L1C products, for both near-lossless and lossy compression and for both temperature and dew point temperature.

The acquisition conditions vary depending on several factors, and globally, this affects the values of the radiances from overpass to overpass. We can identify three different strategies to mitigate the effect of training and testing on statistically different data sets: 1) adapt the regression method

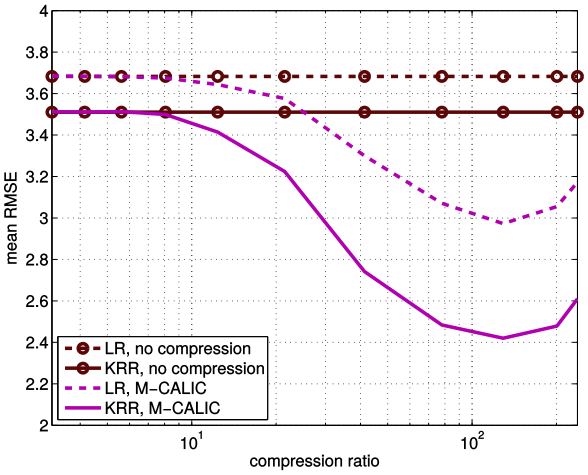


Fig. 5. Estimation using LR or KRR for M-CALIC compression. The horizontal axis represents the compression ratio and the vertical axis represents the mean RMSE (dew point temperature in kelvin) over the different pressure levels. The plot shows the results for M-CALIC compression when using LR (dashed line) and KRR (solid line) for predicting moisture as compared with the retrieval results on original images.

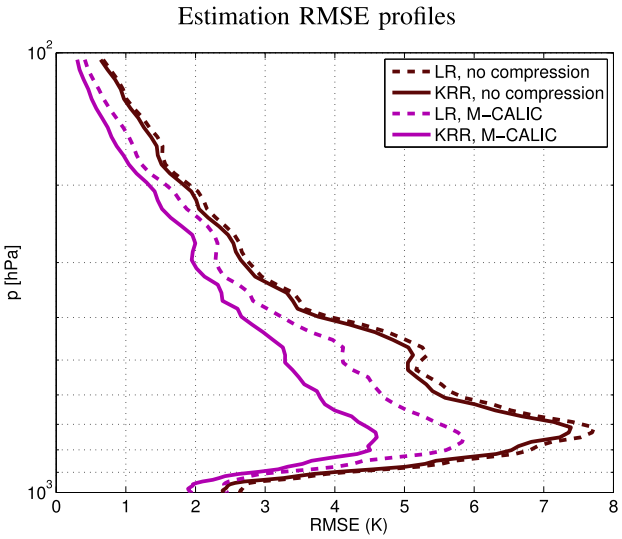


Fig. 6. Moisture RMSE (dew point temperature in kelvin) profiles for near-lossless compression. We chose the PAE with the best average RMSE. For LR (dashed line) and KRR (solid line), the minimum averaged RMSE is obtained at a compression ratio of 129:1. Results when using original images are shown as well for comparison purposes.

in each overpass; 2) build a sufficiently large and representative training data set; and 3) adapt the feature representation (relative normalization) in each overpass. In this paper, we decided to start with the simpler first approach, and hence the statistical retrieval algorithms (LR and KRR) are trained (and cross validated) in each overpass, as previously done in [3]. The approach is simple and efficient. The second approach implies building a representative big data set, but this is challenging, suboptimal, and makes training methods more complex. The last approach would be certainly ideal, but still unsolved in signal processing and machine learning, where this is commonly referred to as *manifold alignment or domain adaptation*. Domain adaptation assumes that the test examples come from a slightly different distribution to those from the training set. Several works trying to accommodate such statistical differences between train–test distributions have been proposed in [35]–[37]. Such complex methods are not considered in this paper. In order to avoid the effect of this *data set shift problem*, we assume an ideal situation: the acquisition conditions are exactly the same for both train and test (data points inside the same image are used for training and testing), but both data sets are disjoint, i.e., we report results in a test set never seen by the regression algorithms for training. This is the most common case in [38].

In the experiments, we follow a similar procedure as in [3]. We apply LR and KRR on IASI data hyperpixels to predict temperature and humidity at different pressure levels, i.e., all the spectral components at a particular spatial position are considered for the regression. The experiments are conducted on images compressed at different target BRs (different compression ratios). For each image, coding scheme, and compression ratio, we follow the same procedure. A set of 2000 samples is used for training and a different set of 2000 samples is used for testing. Note that the test samples are not used in any moment in the training process and that test samples are used only for the evaluation of the performance. Samples are first drawn following a random distribution; then, to allow a fair comparison, the position of the training samples and the position of the test samples are kept constant for all coding schemes. The procedure consists of two steps. First, the dimensionality of the data is reduced to 260 components using the classical PCA/EOF transformation in the spectral domain. Then, LR and KRR models are applied to predict the temperature and the humidity profiles (i.e., the values at each pressure level) at the spatial position corresponding to each particular sample. LR and KRR are trained on input samples of 260 dimensions to predict output samples of 137 dimensions (one for each pressure level). We employ regularization for both methods. In addition, we train the parameters of both methods using cross validation for minimizing the average root-mean-square error (RMSE) of the predictions for all the pressure levels, one parameter for the LR method (the regularization parameter) and two parameters for the KRR method (the regularization parameter and the kernel RBF sigma parameter). Everything that needs to be trained (PCA and the regression parameters) is trained using just the training set. Once everything is learned, the same procedure is applied to the test data. To assess the performance, we compute in the test set the RMSE between the predictions given by the regression models and the ground-truth values of temperature and humidity in each pressure level.

1) Retrieval Assessment for Near-Lossless Compressed Images: Figs. 5 and 6 summarize the prediction results that can be obtained using the selected near-lossless coding technique, M-CALIC, and the investigated regression methods, LR and KRR. Fig. 5 shows the average RMSE across the whole atmospheric column as a function of the compression ratio and the RMSE for each specific pressure level, respectively. One can see that while small compression ratios [left of Figs. 5 and 6] keep the results almost unchanged, larger compression ratios benefit the retrieval performance.

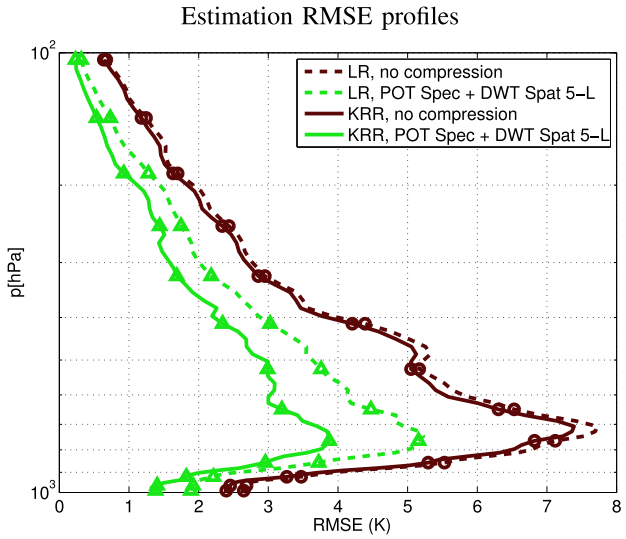


Fig. 7. Moisture RMSE (dew point temperature in kelvin) profiles for lossy compression. The selected configuration for multicomponent JPEG 2000 lossy compression is POT as spectral transform and five levels of spatial DWT. We chose the compression ratio with the best average RMSE. For LR, the minimum averaged RMSE is obtained at a compression ratio of 3200:1. For KRR, the minimum averaged RMSE is obtained at a compression ratio of 1600:1. The results when using original images are shown as well for comparison purposes.

Nevertheless, when the level of distortion inserted—due to the excessive compression—is increased [right of Figs. 5 and 6], the prediction results start to deteriorate. The effect is the same for both LR and KRR, but with an improved accuracy for KRR. These results, although counterintuitive, are consistent with the facts and results reported before in the literature. We refer the reader to Section IV for a lengthy discussion providing explanations and examples for this behavior.

Fig. 6 reports the RMSE results for the different pressure levels (measured in hectopascals) achieved after a near-lossless compression at compression ratio 129:1; again, we can observe that compression benefits the retrieval estimation and that KRR yields a higher performance than LR.

2) Retrieval Assessment for Lossy Compressed Images: As Fig. 6 in the case of near-lossless compression, Fig. 7 reports the RMSE results for the whole range of pressure levels. The above conclusions apply here too: introducing compression benefits the retrieval estimation and KRR is superior to LR.

Next, Fig. 8 summarizes the prediction results that can be obtained using the selected lossy compression approach (JPEG 2000 with several compression configurations) and LR or KRR. Fig. 8 shows the average of the RMSE prediction over the different pressure levels. Several relevant conclusions can be extracted from Fig. 8.

1) *KRR Always Yields Better Results Than LR:* It is clear from the individual plots that for the same configuration of spatial and spectral compression methods, KRR always produces less error than LR.
2) *Using Compressed Images Produces Better Results Than Using Original Images:* This statement is derived by comparing the top leftmost plot of Fig. 8 (original images) with the other plots of Fig. 8. Actually, this is confirmed within a wide BR range. Although it could seem counterintuitive at first glance (compressed images carry less information than original ones), compression algorithms capture and convey the important information, which allows the prediction algorithms to focus on just *this spatial–spectral compact information*, and thus they can yield a better performance. Of course, when the BR is very small (i.e., a high compression is applied), the remaining information is very small and the prediction performance decreases regardless of the capacity of the regression method (either LR or KRR). We argue that the positive influence is an effect of the spatial–spectral compression applied, which can be seen as an efficient (yet indirect) way to include spatial–spectral relations in the retrieval algorithm (see Section IV).
3) *Spatial Transform Is More Important Than Spectral Transform:* If we look at the plots within the same row, going from left to right (from a lower to a larger number of levels of spatial wavelet transform), the performance of the retrieval algorithms increases. This effect is less notable across columns, i.e., using different spectral transforms or a different number of transform levels for the same wavelet-based spectral transform does not have significantly different effects on the retrieval performance. Note that prediction is performed using single hyperpixels, and therefore, the prediction models are not using spatial information directly, although they are indirectly using it through the compressed data. Actually, spatial transform fuse the spatial information in a way that the compressed hyperpixels carry some information about the close-by hyperpixels, while original hyperpixels carry only the spectral information. As such, spatial transform helps to introduce some spatial information in the prediction (see Section IV).
4) *Using Both Spectral and Spatial Transforms Yield the Best Retrieval Performance for Low BRs:* Using an extremely low BR (0.005 bpppc) produces a very good performance in retrieval accuracy [see the rightmost plot in the second row (green triangles) and in the last row (gray diamonds) of Fig. 8]. These would be the best configurations if the interest is to optimize the performance and compression ratio simultaneously.

Finally, Fig. 9 compares the performance of the different compression schemes considering the compression ratio that yields the lower mean RMSE. Again, the conclusion is that using both spectral and spatial transforms yield the best retrieval performance.

IV. Discussion

As reported in Figs. 5–8, compressing the images before carrying out the statistical retrieval for weather forecast brings a significant improvement. In order to understand this observed positive effect, we should first review some well-known facts in signal and image processing. First, it is acknowledged that signal coding is a way of performing signal filtering and denoising. This is a known fact reported elsewhere [24]–[27]. The connections between denoising and compression have been actually widely studied theoretically [39] and experimentally in general and for wavelet transforms in

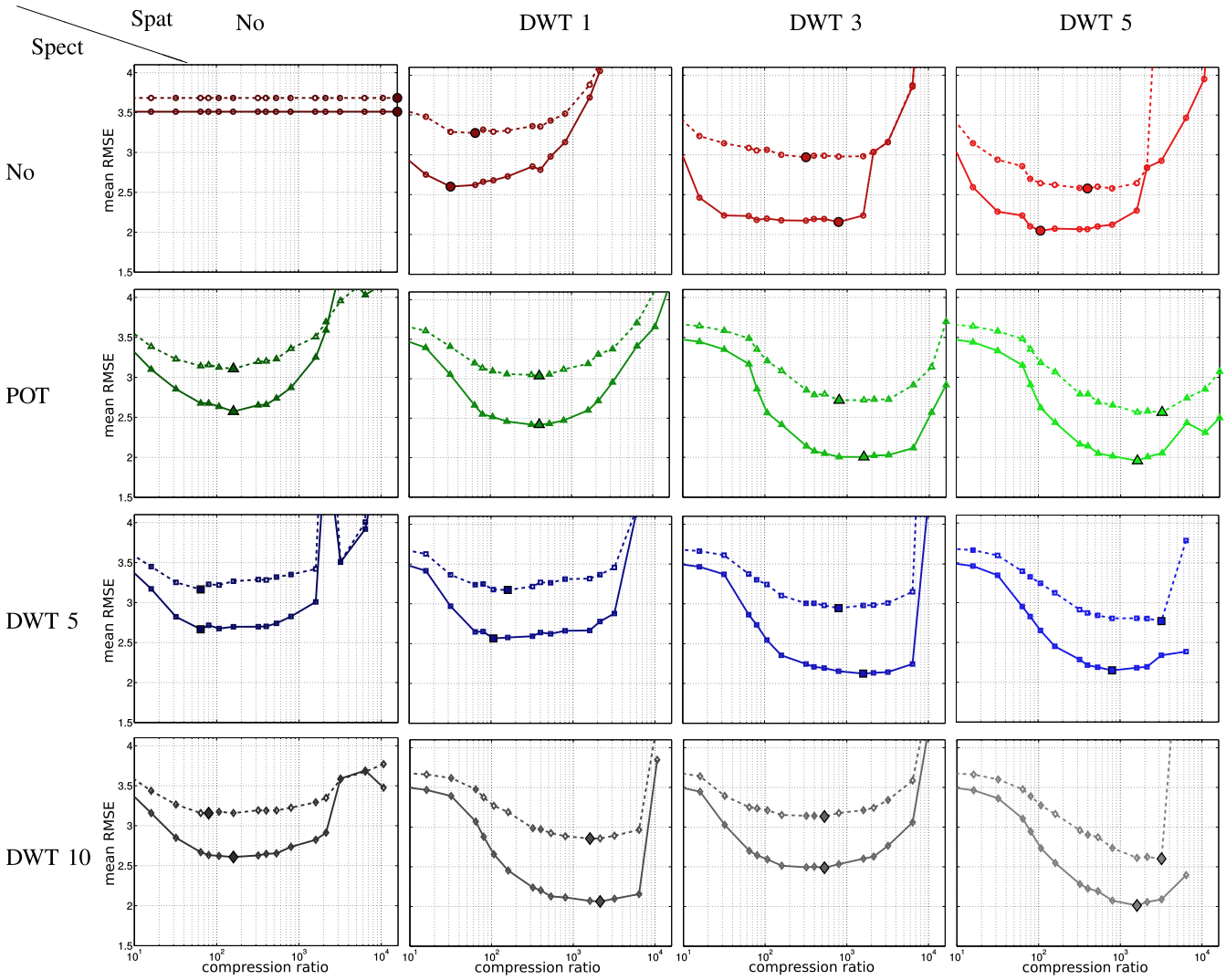


Fig. 8. Estimation using LR or KRR for different spatial and spectral configurations for lossy compression. In all the plots, the horizontal axis represents the compression ratio, and the vertical axis represents the averaged RMSE (dew point temperature in kelvin) over the different pressure levels. Ranges are the same in all the plots to ease the comparison. Each plot shows the results for a particular compression configuration when using LR (dashed lines) and KRR (solid lines) for predicting moisture. Configurations using the same number of levels for the spatial wavelet transform are in the same column and configurations using the same spectral transform are in the same row (same color and multiple color shades). "No" refers to no transform (neither spatial nor spectral). The minimum averaged RMSE in each curve is indicated with a larger marker point.

particular [26], [27], [40]. Therefore, it goes without saying that when one performs lossy signal compression (coding), the obtained signal is denoised.

An additional important observation, relevant to our discussion, is the fact that using a filtered (denoised) signal typically improves regression and function approximation results. The presence of noise in the observations obviously hampers estimation of the underlying signal. This is why very often when performing regression, one aims at estimating the noise (or its covariance) where the signal was buried. Such a noise estimate is then used to discount the uninformative noise contributions in the observation, which definitely helps in recovering the signal. There is a vast literature in both noise (covariance) estimators and how one can embed these priors in generalized least squares [41], which has been extended for nonlinear regression under a Bayesian framework in [42] and the kernel framework in [43], [44]. However, noise estimation is typically difficult, especially when signal-dependent noise (heteroscedasticity), structured domains (as in time series or images), or outlying samples are present. A much simpler and more practical approach is to just filter out the observed signal before applying the regression algorithm. Such filtering can be done with standard PCA projection on the top eigenvectors (thus assuming that noise is related to the higher frequency components), via wavelets (thresholding coefficients in particular scales or orientations whenever some knowledge about the signal is available) or via iterative thresholding methods (that iteratively refine the signal estimation). In conclusion, as mentioned before, compression schemes constitute an alternative indirect simple way to filter signals.

Actually, using wavelet-based and M-CALIC coding schemes yields an extremely useful by-product: they reinforce/encode spatial relations in the generate product. Coding schemes such as the ones used in this paper

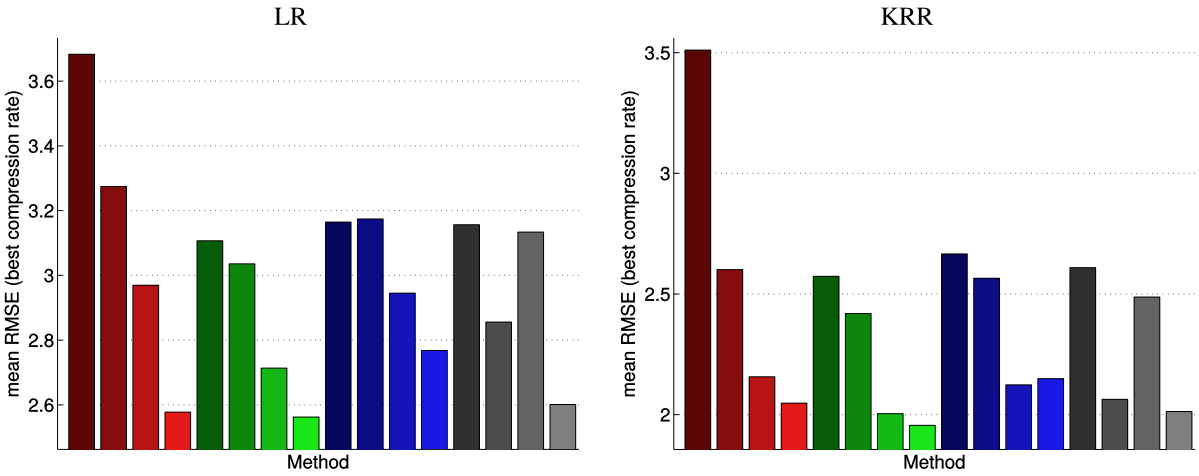


Fig. 9. Performance of the different lossy compression schemes for predicting moisture. The vertical axis reports the minimum averaged RMSE (dew point temperature in kelvin) achieved for each compression scheme. The horizontal axis represents the different lossy compression configuration schemes; the ordering follows that of Fig. 8, first by type of spectral transform (same color) and then by number of levels for the spatial wavelet transform (color shade).

(wavelet-based multicomponent JPEG 2000 for lossy compression and M-CALIC for near-lossless compression) not only typically reduce the noise variance but also encode spatial smoothness in the compressed image. Compression algorithms aim to concentrate information in a reduced set of coefficients, which is done through "decorrelation" of the observation space (spatial, spectral, or both), and very often assume stationarity (hence smoothness) implicitly. This is obviously an indirect simple way to enforce local (spatial) relations between nearby pixels. Image compression therefore enforces pixel correlations and smoothness in the image plane, which can be exploited in pixelwise regression algorithms to improve the spatial homogeneity/consistency of the prediction maps too. In fact, it is widely acknowledged that model inversion and parameter retrieval in remote sensing applications largely improve when including spatial information (see [45], [46], and the references therein). Moreover, in recent years, it is being widely accepted that using the spectral information only (as done by pixelwise algorithms), is not optimal and that it is important to find a tradeoff between the spatial and the spectral information [47].

Finally, we should mention that these observations are not new in signal and image processing. Several applications benefit from a coding-first scheme. There is a wide evidence that compressing the signal before any further processing is beneficial. We have shown this effect previously for several compression algorithms in hyperspectral image classification [12], and it has been a common strategy in other applications such as genomics,in particular for next-generation sequencing [48], [49], seismic signal processing [50], [51], bioengineering [52], and communications [53]. Therefore, we posit that what we report is yet another case of an effect previously observed in signal and image processing, here transported to the field of biophysical–geophysical parameter estimation in atmospheric applications using infrared sounding data.

V. Conclusion

This paper was concerned with how to include spatial–spectral information for retrieving atmospheric profiles of temperature and humidity using infrared sounders and statistical regression. While many sophisticated approaches can be deployed for this purpose (e.g., contextual and spatial–spectral feature extraction or filtering at a postprocessing stage), they typically involve handcrafted features and engineering new and complicated retrieval algorithms, nonautomatic processes, and tedious user intervention. Last but not least, such processes raise important issues about parameter tuning and the little control one has on the retrieval generalization capabilities. Alternatively, we approached the problem in an indirect way by compressing the hyperspectral cubes before performing retrieval. Compression techniques are well understood and the impact on the quality of the radiances can be easily controlled by prescribing a compression ratio. A wide range of spatial, spectral, and spatial–spectral image compression approaches were evaluated and two linear and nonlinear regression algorithms were compared. Results for the prediction of both temperature and humidity can be found at http://isp.uv.es/coding_retrieval.html.

Near-lossless compression was carried out through the M-CALIC coding technique, while lossy compression was based on the multicomponent JPEG2000 standard. Even though both compression approaches introduce certain level of distortion on the original data, they simultaneously incorporate the spatial–spectral feature relations, so there is no need to design specific retrieval algorithms that fuse such information. In our experiments, we searched for an acceptable compression ratio on IASI L1C products such that it has a positive impact on the accuracy of the statistical retrievals, i.e., an optimal tradeoff between the positive effect of the compression, i.e., the introduction of spatial relations and some sort of denoising on the data, and its negative effect, i.e., the reduction of information compared with the original image. Our first conclusion is interesting enough, and we think that the message can be of high interest to the parameter retrieval community in general and to scientists and organizations dealing with atmospheric retrieval models in particular. We show that compression is a very easy, cheap, and consistent way to perform denoising and to introduce spatial–spectral image relations in the compressed

product that largely benefits the current (pointwise) retrieval. We observed that high compression ratios yielded improved results in predicting dew point temperature (similar results for temperature were obtained; see Supplementary Material) in all the experiments for moderate and high compression ratios. Of course when the compression ratio is extremely large, this effect vanishes, since the amount of information removed by the compression is too large.

Our second message we want to convey in this communication is that a spatial transform revealed itself to be more important than a spectral transform. This observation is related to the fact that the retrieval algorithms work in a pixelwise manner, and hence the spatial component is missing. This is why the spectral–spatial image compression configuration results in such large gains in spectrum-based retrieval. We will explore in future works the tradeoff between the gains obtained when using a scheme based on *spectral–spatial compression plus spectral retrieval* versus a more complete yet challenging *spectral–spatial compression plus spectral–spatial retrieval*. This second approach would require, however, redesigning retrieval algorithms, which in most of the cases are pixel based.

As the main conclusion, we want to highlight that compression prior to model inversion is largely beneficial in retrieval problems in general. The benefit is not only of practical use but also computationally convenient and more accurate. It is *practical* because following the presented methodology does not change current operational approaches much.

1) Current retrieval schemes would be still valid since they would be fed with efficiently compressed data before the retrieval instead of the original data.
2) There are moderate-to-big savings in data storage and transmission because of the compression step.

The benefit is also in terms of prediction accuracy, as we have observed consistent gains in all cases and images. We believe that the proposed methodology may benefit the development of current and upcoming infrared sounding and hyperspectral sensors to advance in biophysical–geophysical parameter estimation schemes. It has not escaped our notice that the same procedure can be actually applied to other algorithms for retrieval, such as the OE scheme or Bayesian approaches, and to other problems dealing with high resolution data, such as spectral unmixing. These issues are the subject of further research studies.

ACKNOWLEDGMENT

The authors thank Dr. T. August (EUMETSAT, Germany) for the provided data and the fruitful discussions on retrievals and spatial–spectral coding.

REFERENCES

[1] *IASI Level 1: Product Guide*, document EUM/OPS-EPS/MAN/04/0032, EUMETSAT, 2014.
[2] B. Tournier, D. Blumstein, and F.-R. Cayla, "IASI level 0 and 1 processing algorithms description," in *Proc. 12th ISTC Conf.*, 2002.
[3] 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 Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1759–1769, May 2012.
[4] G. Mercier, M. Mouchot, and G. Cazuguel, "Joint classification and compression of hyperspectral images," in *Proc. IEEE Int. Remote Sens. Symp. (IGARSS)*, vol. 4. May 1999, pp. 2035–2037.
[5] M. D. Pal, C. M. Brislawn, and S. R. Brumby, "Feature extraction from hyperspectral images compressed using the JPEG-2000 standard," in *Proc. 5th IEEE Southwest Symp. Image Anal. Interpretation*, Apr. 2002, pp. 168–172.
[6] A. Kaarna, P. Toivanen, and P. Keränen, "Compression and classification methods for hyperspectral images," *Pattern Recognit. Image Anal.*, vol. 16, no. 3, pp. 413–424, 2006.
[7] I. Blanes and J. Serra-Sagristà, "Quality evaluation of progressive lossy-to-lossless remote-sensing image coding," in *Proc. 16th IEEE Int. Conf. Image Process. (ICIP)*, Apr. 2009, pp. 3709–3712.
[8] B. Penna, T. Tillo, E. Magli, and G. Olmo, "Transform coding techniques for lossy hyperspectral data compression," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1408–1421, May 2007.
[9] Q. Du and J. E. Fowler, "Hyperspectral image compression using JPEG2000 and principal component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 2, pp. 201–205, Feb. 2007.
[10] I. Blanes, A. Zabala, G. Moré, X. Pons, and J. Serra-Sagristà, "Classification of hyperspectral images compressed through 3D-JPEG2000," in *Knowledge-Based Intelligent Information and Engineering Systems*. Berlin, Germany: Springer, 2008, pp. 416–423.
[11] G. Carvajal, B. Penna, and E. Magli, "Unified lossy and near-lossless hyperspectral image compression based on JPEG 2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 593–597, Oct. 2008.
[12] F. Garcia-Vilchez *et al.*, "On the impact of lossy compression on hyperspectral image classification and unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 253–257, Feb. 2011.
[13] C. I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Norwell, MA, USA: Kluwer, 2003.
[14] Q. Du and C.-I. Chang, "Linear mixture analysis-based compression for hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 875–891, Apr. 2004.
[15] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2025–2041, Sep. 2002.
[16] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.
[17] J. García-Sobrino, I. Blanes, V. Laparra, G. Camps-Valls, and J. Serra-Sagristà, "Impact of near-lossless compression of IASI L1C data on statistical retrieval of atmospheric profiles," in *Proc. On-Board Payload Data Compress. Workshop (OBPDC)*, Venice, Italy, Oct. 2014.
[18] E. Magli, G. Olmo, and E. Quacchio, "Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 1, pp. 21–25, Jan. 2004.
[19] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. New York, NY, USA: Springer, 2012.
[20] *Standard JPEG2000*, document ISO/IEC 15444, JPEG-Committee, accessed on 2016. [Online]. Available: http://www.jpeg.org/jpeg2000/documentation.html
[21] I. Blanes and J. Serra-Sagristà, "Pairwise orthogonal transform for spectral image coding," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 961–972, Mar. 2011.
[22] I. Blanes, J. Serra-Sagristà, M. W. Marcellin, and J. Bartrina-Rapesta, "Divide-and-conquer strategies for hyperspectral image processing: A review of their benefits and advantages," *IEEE Signal Process. Mag.*, vol. 29, no. 3, pp. 71–81, Mar. 2012.
[23] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[24] B. K. Natarajan, "Filtering random noise via data compression," in *Proc. IEEE Data Compress. Conf.*, Jun. 1993, pp. 60–69.
[25] B. K. Natarajan, "Filtering random noise from deterministic signals via data compression," *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2595–2605, Nov. 1995.
[26] P. L. Dragotti and M. Vetterli, "Footprints and edgeprints for image denoising and compression," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2. Oct. 2001, pp. 237–240.
[27] P. L. Dragotti and M. Vetterli, "Wavelet and footprint sampling of signals with a finite rate of innovation," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2004, pp. 941–944.
[28] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 437–444, Apr. 1997.

[29] M. J. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1309–1324, Aug. 2000.
[30] B. Schölkopf and A. Smola, *Learning With Kernels—Support Vector Machines, Regularization, Optimization and Beyond*. Cambridge, MA, USA: MIT Press, 2002.
[31] E. Magli. *M-CALIC Software*, accessed on Jan. 2017. [Online]. Available: http://www1.tlc.polito.it/oldsite/sas-ipl/download.php
[32] D. Taubman. *Kakadu Software*, accessed on Jan. 2017. [Online]. Available: http://www.kakadusoftware.com/
[33] *JPEG2000 Part 2—Extensions*, document ISO/IEC 15444-2, 2004.
[34] GICI-UAB. *Spectral Transform*, accessed on Jan. 2017. [Online]. Available: http://gici.uab.cat/GiciWebPage/downloads.php
[35] E. Izquierdo-Verdiguier, V. Laparra, L. Gómez-Chova, and G. Camps-Valls, "Encoding invariances in remote sensing image classification with SVM," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 1–5, Apr. 2013.
[36] M. U. Gutmann, V. Laparra, and A. H. J. Malo, "Spatio-chromatic adaptation via higher-order canonical correlation analysis of natural images," *PLoS ONE*, vol. 9, no. 2, p. e86481, 2014.
[37] D. Tuia and G. Camps-Valls, "Kernel manifold alignment for domain adaptation," *PLoS ONE*, vol. 11, no. 2, p. e86481, 2016.
[38] S. Srivastava, P. K. Thapliyal, M. V. Shukla, J. S. H. Bisht, and D. Mitra, "Statistical retrieval of ozone and meteorological parameters using SHADOZ observations and radiative transfer model," *J. Indian Soc. Remote Sens.*, vol. 44, no. 6, pp. 1–8, Feb. 2016.
[39] T. Weissman and E. Ordentlich, "The empirical distribution of rate-constrained source codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3718–3733, Nov. 2005.
[40] S. G. Chang, B. Yu, and M. Vetterli, "Image denoising via lossy compression and wavelet thresholding," in *Proc. Int. Conf. Image Process.*, vol. 1. Oct. 1997, pp. 604–607.
[41] J. Johnston. (1972). *Econometric Methods*. [Online]. Available: https://books.google.es/books?id=BZtvwZAGyV0C
[42] M. Lazaro-Gredilla, M. K. Titsias, J. Verrelst, and G. Camps-Valls, "Retrieval of biophysical parameters with heteroscedastic Gaussian processes," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 4, pp. 838–842, Apr. 2014.
[43] G. Camps-Valls and L. Bruzzone, Eds., *Kernel Methods for Remote Sensing Data Analysis*. Hoboken, NJ, USA: Wiley, Dec. 2009.
[44] L. Gomez-Chova and G. Camps-Valls. (2012). *Learning With the Kernel Signal to Noise Ratio*. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84870658481&partnerID=40&md5=702f25c382d3d87f1d5c9b3509995474
[45] S. Jimenez and J. Malo, "The role of spatial information in disentangling the irradiance–reflectance–transmittance ambiguity," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4881–4894, Aug. 2014.
[46] G. Camps-Valls, D. Tuia, L. Gómez-Chova, S. Jiménez, and J. Malo, *Remote Sensing Image Processing* (Synthesis Lectures on Image, Video, and Multimedia Processing). San Mateo, CA, USA: Morgan Kaufmann, 2011.
[47] V. Laparrcr and R. Santos-Rodriguez, "Spatial/spectral information trade-off in hyperspectral images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 1124–1127.
[48] Y. W. Yu, D. Yorukoglu, and B. Berger, "Traversing the k-mer landscape of NGS read datasets for quality score sparsification," in *Research in Computational Molecular Biology* (Lecture Notes in Computer Science). Cham, Switzerland: Springer International Publishing, 2014, pp. 385–399.
[49] I. Ochoa, M. Hernaez, R. Goldfeder, T. Weissman, and E. Ashley, "Denoising of quality scores for boosted inference and reduced storage," in *Proc. IEEE Data Compress. Conf.*, Oct. 2016, pp. 1–10.
[50] M. F. Khene and S. H. Abdul-Jauwad, "Adaptive seismic compression by wavelet shrinkage," in *Proc. 10th IEEE Workshop Statist. Signal Array Process.*, Aug. 2000, pp. 544–548.
[51] J. Ramirez, Jr., and F. G. Meyer, "Machine learning for seismic signal processing: Phase classification on a manifold," in *Proc. 10th Int. Conf. Mach. Learn. Appl. Workshops (ICMLA)*, vol. 1. Dec. 2011, pp. 382–388.
[52] A. A. Petrosian and F. G. Meyer, Eds., *Wavelets in Signal and Image Analysis: From Theory to Practice* (Computational Imaging and Vision). Boston, MA, USA: Kluwer, 2001. [Online]. Available: http://opac.inria.fr/record=b1098905
[53] L. Hanzo, F. C. Somerville, and J. Woodard, *Voice and Audio Compression for Wireless Communications*. Hoboken, NJ, USA: Wiley, 2007. [Online]. Available: http://dx.doi.org/10.1002/9780470516034.fmatter

Joaquín García-Sobrino received the B.S. and M.S. degrees in computer science from the University of Almería, Almería, Spain, in 2009 and 2010, respectively. He is currently pursuing the Ph.D. degree with the Universitat Autònoma de Barcelona, Barcelona, Spain.

Since 2013, he has been with the Group on Interactive Image Coding, Universitat Autònoma de Barcelona. His research interests include remote sensing data compression.

Joan Serra-Sagristà (S'97–M'05–SM'11) received the Ph.D. degree in computer science from the Universitat Autònoma de Barcelona (UAB), Barcelona, Spain, in 1999.

From 1997 to 1998, he was at the University of Bonn, Bonn, Germany, funded by DAAD. He is currently an Associate Professor (hab. Full Professor) with the Department of Information and Communications Engineering, UAB. He has co-authored over 100 papers. His research interests include data compression, with a focus on image coding for remote sensing applications.

Dr. Serra-Sagristà was a recipient of the Spanish Intensification Young Investigator Award in 2006. He serves or has served as an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE TRANSACTIONS ON IMAGE PROCESSING, and as the Committee Chair of the Data Compression Conference.

Valero Laparra was born in València, Spain, in 1983. He received the B.Sc. degree in telecommunications engineering and the B.Sc. degree in electronics engineering from the Universitat de València, València, in 2005 and 2007, respectively, the B.Sc. degree in mathematics from the Universidad Nacional de Educación a Distancia, Madrid, Spain, in 2010, and the Ph.D. degree in computer science and mathematics from the Universitat de València in 2011.

He is currently a Post-Doctoral Researcher with the Image Processing Laboratory, Universitat de València.

Xavier Calbet received the M.S. degree in physics from the University of Valencia, Valencia, Spain, in 1989, the M.S. degree in meteorology from the University Complutense of Madrid, Madrid, Spain, in 1993, and the Ph.D. degree in physics from the University of La Laguna, La Laguna, Spain, in 1993.

In 1996, he joined the Spanish Meteorological Service (currently known as State Meteorological Agency, AEMET), Madrid, as a State Meteorologist. From 2003 to 2015, he was with the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany, specializing in atmospheric profile retrievals obtained from hyperspectral sounding instruments such as the infrared atmospheric sounding interferometer and planning for the future geostationary hyperspectral sounding mission MTG-IRS. Since 2015, he has been with the Spanish Meteorological Service as a Science Coordinator of the Nowcasting Satellite Application Facility from EUMETSAT.

Gustau Camps-Valls (M'04–SM'07) received the B.Sc. degree in physics, the B.Sc. degree in electronics engineering, and the Ph.D. degree in physics from the Universitat de València, València, Spain, in 1996, 1998, and 2002, respectively.

He was a Visiting Researcher at the Remote Sensing Laboratory, University of Trento, Trento, Italy, in 2002, and at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany, in 2009. He was an Invited Professor at the Laboratory of Geographic Information Systems, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, in 2013. He is currently an Associate Professor (hab. Full Professor) with the Department of Electronics Engineering and a Research Coordinator with the Image and Signal Processing Group, Universitat de València. He is interested in the development of machine-learning algorithms for geoscience and remote sensing data analysis. He has authored 120 journal papers, more than 150 conference papers, and 20 international book chapters. He is the Editor of the books *Kernel Methods in Bioengineering, Signal, and Image Processing* (IGI, 2007), *Kernel Methods for Remote Sensing Data Analysis* (Wiley, 2009), and *Remote Sensing Image Processing* (MC, 2011). He is the Co-Editor of the book *Digital Signal Processing with Kernel Methods* (Wiley, 2015). He holds an H-index of 48, entered the ISI list of Highly Cited Researchers in 2011, and Thomson Reuters ScienceWatch identified one of his papers on kernel-based analysis of hyperspectral images as a Fast Moving Front Research.

Dr. Camps-Valls has been a member of the Data Fusion Technical Committee of the IEEE GRSS since 2007, and the Machine Learning for Signal Processing Technical Committee of the IEEE SPS since 2009. He is currently a member of the MTG-IRS Science Team of EUMETSAT, and a Referee and Program Committee Member of many international journals and conferences. In 2015, he obtained the prestigious European Research Council consolidator grant on statistical learning for earth observation data analysis. He is currently an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE SIGNAL PROCESSING LETTERS, and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and an Invited Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING in 2012 and the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE in 2015.

Chapter 5

Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI data

MDPI

Article

Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI data

Joaquín García-Sobrino [1,*], Valero Laparra[2], Joan Serra-Sagristà [1], Xavier Calbet [3], and Gustau Camps-Valls [2]

[1] Group on Interactive Coding of Images, Universitat Autònoma de Barcelona, Bellaterra, Spain. Web: http://www.gici.uab.cat. joaquin.garcia.sobrino@deic.uab.cat (J.G.); joan.serra@uab.cat (J.S.)
[2] Image Processing Laboratory (IPL), Universitat de València, València, Spain. Web: http://isp.uv.es. valero.laparra@uv.es (V.L.); gustau.camps@uv.es (G.C)
[3] AEMET, C/ Leonardo Prieto Castro, 8, 28071 Madrid, Spain. Web: http://www.aemet.es. xcalbeta@aemet.es (X.C)
* Correspondence: joaquin.garcia.sobrino@deic.uab.cat; Tel.: +34-93-5811861

Academic Editor: name
Version June 12, 2018 submitted to Remote Sens.

Abstract: In this paper we analyze the effect of spatial and spectral compression on the performance of statistically based retrieval. Although the quality of the information is not completely preserved during the coding process, experiments reveal that a certain amount of compression may yield a positive impact on the accuracy of retrievals. We unveil two strategies, both with interesting benefits: either to apply a very high compression, which still maintains the same retrieval performance as that obtained for uncompressed data; or to apply a moderate to high compression, which improves the performance. As a second contribution of this work, we focus on the origins of these benefits. On the one hand, we show that a certain amount of noise is removed during the compression stage, which benefits the retrievals performance. On the other hand, we analyze the effect of compression on spectral/spatial regularization (smoothing). We quantify the amount of information shared among the spatial neighbours for the different methods and compression ratios. We also propose a simple strategy to specifically exploit spectral and spatial relations and find that, when these relations are taken into account beforehand, the benefits of compression are reduced. These experiments suggest that compression can be understood as an indirect way to regularize the data and exploit spatial neighbours information, which improves the performance of pixel-wise statistics based retrieval algorithms.

Keywords: Infrared Atmospheric Sounding Interferometer (IASI), Lossy Compression, Spectral Transforms, JPEG 2000, Statistically Based Retrieval, Kernel Methods, Regression.

1. Introduction

In recent decades, advances in remote sensing technology have made it possible to collect information from the electromagnetic spectrum with unprecedented accuracy and resolution [1]. Infrared sounder instruments usually produce large volumes of data, which is costly to manage in an operational context, i.e., for transmission, processing, and storage. An effective strategy to alleviate the problems derived from the data size is to compress the data according to the specifics needs of the final users. However, in order to achieve high compression ratios, it is typically needed to go through a lossy compression stage. Lossy compression entails to lose some information in the reconstructed data, which may compromise the quality of the products in later processing stages. Therefore, a careful evaluation of the impact of the compression process applied to the data is needed to determine the quality of the information and of the derived products.

Although lossy compression implies going through a distortion process, the quality of the recovered data can be still adequate for the intended specific use. If the amount of the inherent data acquisition noise is large, the signal removed during the coding process is mostly noise for certain compression ratios [2,3]. Therefore, the

amount of useful signal lost is usually low and the quality of the recovered products may be yet appropriate, or even better suited, to feed a subsequent information extraction stage. Of course, if the compression ratio is very high, a lot of useful information is removed and the quality of the reconstructed data is seriously compromised.

The impact of lossy compression on information extraction systems has been investigated in several application areas. Ryan and Arnold [4] discussed the lossy compression of remotely sensed images using vector quantization in the context of image classification. The impact of compression was analyzed on the maximum likelihood classification of the recovered data concluding that the loss in the classification accuracy was not significant (less than 8%). Sánchez and Perronnin [5] evaluated two lossy compression techniques, one based on product quantizers and another based on dimensionality reduction, to assess large-scale image classification. Results reported that compression ratios between 64:1 and 128:1 produced little loss in the classification performance. Similar results were obtained in Minguillon *et al.* [6], Kaarna *et al.* [7], Zabala *et al.* [8], Blanes and Serra-Sagristà [9], Álvarez-Cortés *et al.* [10], where lossy compression did not markedly reduce classification accuracy.

Besides, several studies have analyzed the effect of near-lossless and lossy compression techniques based on JPEG 2000 standard within the framework of feature extraction, classification, and anomaly detection tasks. Pal *et al.* [11] analyzed the performance of supervised, unsupervised, and hybrid classification processes. Results reflected that classification accuracy was still reliable even at low bit-rates (high compression ratio). A compression scheme based on Principal Component Analysis (PCA) was proposed in Du and Fowler [12] producing competitive results in terms of information preservation in anomaly detection tasks. Supervised and unsupervised classification of reconstructed data was evaluated in Blanes *et al.* [13]. The experimental results showed competitive classification performance after the compression stage. Analogous results were reported in Carvajal *et al.* [14], where the proposed compression algorithm for hyperspectral images did not significantly reduce the performance of hard classification, linear spectral unmixing, and anomaly detection.

An interesting observation was reported in Penna *et al.* [15], where the impact of lossy compression of hyperspectral data was analyzed on multiclass classification, classification of mixed pixels via spectral unmixing, binary hard classification, and anomaly detection. Experimental results revealed that for multiclass classification and spectral unmixing, the Signal-to-Noise ratio (SNR) was a reasonable indicator of classification performance; for binary hard classification, the performance was little dependent on the SNR; and for anomaly detection, the compression algorithms that produced the best rate-distortion performance were not the best choice. Actually, the effect of lossy compression on supervised classification and spectral unmixing using support vector machines was quantified in García-Vílchez *et al.* [2]. The assessment reported that lossy compression can produce accurate results even at high compression ratios (above 16:1).

The estimation of atmospheric parameters from remote sensing data is an *inverse problem* where we retrieve the physical parameters given a set of observations. Inverting the physical (radiative transfer) model through look-up-tables or optimal estimation models (OEM) [16] are standard approaches, but they lead to much higher computational costs than those required for statistical approaches [17]. In the last decade, statistical model inversion based on machine learning has provided excellent performance in accuracy and efficiency terms. Statistically based atmospheric parameter retrieval from IASI and AIRS data was first conducted using vanilla neural networks trained with backpropagation [18–20]. In a set of previous comparisons [17,21–24] we used other forms of nonlinear and nonparametric machine learning regression like more advanced neural nets, kernel ridge regression and Gaussian processes. Other strategies to plain regression consider introducing a dimensionality reduction before the regression. Smart approaches consider spatially and noise-aware transforms like in [23,25], or more advanced *nonlinear* dimensionality reduction approaches based on kernel machines [26].

The functional sliced inverse regression (FSIR) in [27] is a dimensionality reduction method that generalizes PCA by designing a spectral information driven by the parameters to be retrieved. FSIR is actually related to linear discriminant analysis (LDA) and partial least squares (PLS) approaches, being a kind of supervised dimensionality reduction approach, and provides noise-aware feature components (hence being related to minimum noise fraction -MNF- methods). Nevertheless, the standard PCA approach is still the most widely adopted and studied for dimensionality reduction in atmospheric parameter retrieval. In [28] the impact of spectral compression on statistically based retrievals was analyzed, investigating the use of Principal Components Analysis (PCA) based

methods for compressing high-resolution infrared measurements before performing linear regression. It was shown that a compression ratio of 15:1 can be performed with low degradation of the temperature and water vapour retrievals. In [29,30] we observed that data that had gone through a near-lossless and lossy compression process enabled achieving, for certain compression ratios, improved statistically based retrieval performance compared to the retrieval on the original products. It was suggested that the spectral/spatial regularization and the noise filtering produced by the compression stage benefit the performance of the statistical retrieval algorithms. However, these counterintuitive observations were scarcely studied. In this paper, we provide appropriate responses to the observations reported in [29,30], investigate the actual impact of compression in the retrievals, and supply results and discussion to understand why lossy compression may give rise to improved retrievals. This led us to lay out the main contributions of this paper:

1. While experiments in [29,30] were carried out in an ideal scenario, where samples to define the training, validation and testing subsets were acquired under the same conditions, here, we assume a more realistic scenario, where samples used to define the training, the validation and the testing subsets come all from disjoint orbits. Therefore, conclusions are not biased by an eventual risk of model overfitting.
2. Three different multi-component transforms have been employed in the compression process, all bringing forth improved performance.
3. The causes of the benefits on the retrieval performance due to a compression stage are thoroughly investigated and discussed, and reasons of the statistically based retrieval improvement are provided.

In this paper, a wide spread lossy compression technique and two different statistical retrieval algorithms are evaluated on Infrared Atmospheric Sounding Interferometer (IASI) L1C data [31]. For lossy compression, JPEG 2000 standard [32] is paired along with three different spectral transforms: the Discrete Wavelet Transform (DWT) [33], the Pairwise Orthogonal Transform (POT) [34]), and the Multilevel Clustering Karhunen-Loève Transform (Multilevel Clustering KLT) [35]. For atmospheric variables prediction, a linear and a nonlinear statistically based retrieval algorithm, i.e., the Linear Regression (LR) and the Kernel Ridge Regression (KRR) [36], respectively, are employed in the retrieval of physical information (temperature and dew point temperature profiles) from the reconstructed data.

To investigate the origin of the improved retrieval performance produced by lossy compression, two approaches are followed. On the one hand, we analyze the noise level remaining in the recovered spectra. On the other hand, we compare the performance of the retrieval algorithms when a compressed (and reconstructed) data is employed, with a simple method which specifically takes into account information from the spatial and spectral neighbours. Experiments reveal that the improvement in retrieval accuracy is motivated by: 1) lossy compression performing some noise filtering, which typically improves regression and function approximation results; and 2) compression being an indirect way to exploit spatial feature relations, which, generally, helps in pixel-wise retrieval algorithms.

The remainder of the paper is organized as follows. Section 2 introduces the proposed sequential approach detailing the compression scheme and the statistically based retrieval methods used in the experiments. Section 3 reports the data collection used in the experiments. Section 4 reports and analyzes the experimental results. Section 5 provides an extensive discussion. Finally, Section 6 draws some conclusions.

2. Methodology

This section describes the methods and techniques used in the experiments. For the compression stage, the JPEG 2000 standard is paired along with three spectral transforms to exploit the high spectral redundancy present in IASI L1C data. For atmospheric prediction, two different statistically based retrieval algorithms, which have provided competitive performance in atmospheric parameter retrieval [37], are employed: a standard least squares LR and a kernel ridge regression (KRR) method.

Figure 1 illustrates the proposed sequential scheme. In a first stage, the data is lossy compressed. At the receiver side, data is decompressed to produce the reconstructed data. Finally, the statistical parameter retrieval is carried out on the reconstructed data.

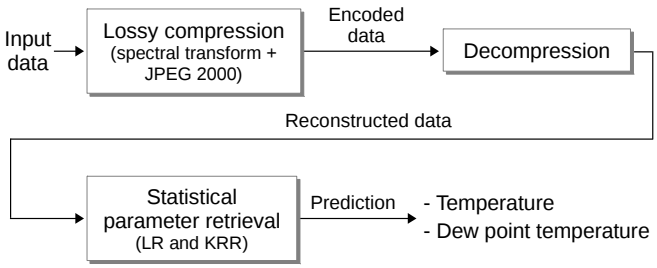


Figure 1. Proposed sequential chain.

2.1. Data Compression

Compression of remote sensing products is an efficient strategy to reduce the large size of the collected data. In recent years, various compression strategies have been adopted in several standards and have been implemented both on board the satellites and on-the-ground processing stations [38–42].

Three compression approaches can be considered, namely, lossless, lossy, and near-lossless. Lossless compression is the mandatory approach when the original signal must be fully preserved, however, the achievable compression ratios are rather low (below 5:1). Lossy compression can be a desirable approach in scenarios where high compression ratios are required and losing some information in the signal might be admissible. Near-lossless compression is also a lossy compression approach, where some information is removed from the original data, but the compression ratios achieved are closer to lossless compression than to typical lossy compression. Near-lossless compression is an adequate approach when higher compression ratios compared to lossless compression are demanded and, at the same time, a specific fidelity criterion (usually the peak absolute error) must be preserved in the recovered data.

Here we will focus on *lossy compression* because it is a widely accepted strategy when large volumes of remote sensing data need be compressed, as witnessed in the today-in-use pipeline for IASI data [43,44] (recall that we will use IASI data for the experiments).

2.1.1. Lossy Compression

Lossy compression approach allows to achieve high compression ratios at the expense of recovering a reconstructed data not identical to the original ones. It is expected that the reconstructed data preserve enough information to be used for the intended specific purpose, while alleviating problems derived from the transmission, handling, and storage of large volumes of information.

The international standard JPEG 2000 [32] is used to carry out the lossy compression stage. JPEG 2000 was published by the *Joint Photographic Experts Group (JPEG)* in year 2000 as the successor of classical JPEG coding standard, and provides a wide diversity of features and functionalities in a single compressed code-stream: among others, lossless and lossy compression, progressive lossy-to-lossless coding, robustness to the presence of errors, region-of-interest coding, and progressive transmission. JPEG 2000 is employed in a wide range of applications, e.g., remote sensing, medical imagery, mobile applications, digital library, digital photography, etc.

To achieve competitive compression performance in data formed by thousands of spectral components, like IASI L1C products, it is of paramount importance to exploit the redundancy present in the spectral dimension. Three spectral transforms are paired along with JPEG 2000: the Discrete Wavelet Transform (DWT), the Pairwise Orthogonal Transform (POT), and the Multilevel Clustering Karhunen-Loève Transform (Multilevel Clustering KLT).

DWT decomposes the processed signal into different subbands, which allows to decorrelate the data to be encoded. When the transform is applied, the signal is decomposed into two set of coefficients: the low and high frequency subbands. The low frequency subbands (L) have a coarse frequency resolution, while the high frequency subbands (H) represent fine details of the data. The high frequency subbands contain coefficients that can be discarded or quantized to achieve efficient lossy compression. Often, competitive data quality can be achieved from a small amount of transform coefficients, which makes DWT a suitable transform

for lossy compression [45]. DWT can be applied on the transform coefficients (low frequency subbands) in successive levels producing further decorrelation, although the optimal number of transform levels shall depend on the particularities of the data. DWT can be used too to transform two dimensional signals by applying a one-dimensional transform in the vertical direction and a one-dimensional transform in the horizontal direction, which usually improves the coding results.

Both POT and Multilevel Clustering KLT are affordable approximations of the computationally demanding Karhunen-Loève Transform (KLT) [46], which is optimal decorrelating Gaussian sources. A critical consideration when applying a spectral transform on data volumes with large spectral dimension is the computational complexity [47]. Some spectral transforms like the KLT may prove unusable for large dimensions such as in IASI data. POT and Multilevel Clustering KLT rely on the implementation of a divide-and-conquer strategy to the KLT. While a classic KLT decorrelates all components with each other regardless of how much energy they share, a divide-and-conquer strategy implements a mechanism to decorrelate several spectral components with high shared energy and ignore the other components. The resulting transform is the composition of smaller KLT transforms, where each composition is processed as a classic KLT. If a multilevel mode is used, most of the data energy flows across the composition of transforms up to the last level, because most of the signal energy is grouped in the first few decorrelated components, which are further decorrelated in upper levels. POT uses a two-component KLT transform for every pair of consecutive components and works in a multilevel mode. In the case of Multilevel Clustering KLT, each composition of KLT can be formed by a number of spectral components between two and all the spectral components. Figure 2 illustrates the structure of a classical KLT, of a POT and of a Multilevel Clustering KLT.

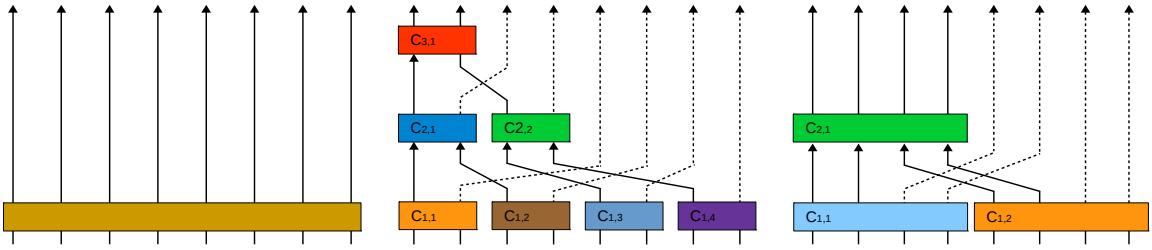


Figure 2. Structure of a classical KLT, POT, and Multilevel Clustering KLT. This example decorrelates eight spectral components. Each arrow denotes a component and each coloured rectangle represents the computation of a KLT. In the case of POT, three levels of transform are applied, with a different number of clusters per level, and all clusters transforming two components. In the case of Multilevel Clustering KLT, two levels of multilevel clustering are applied and two clusters are defined in the first level.

2.2. Statistically Based Parameter Retrieval

In the last decades, statistically based retrieval has proved to be very useful to solve inverse problems using remote sensing data. Different algorithms have been considered, but mainly restricted to neural networks [18–20, 23] and kernel machines and Gaussian processes [17,21,22,24]. In particular, kernel methods are well suited due to its ability to deal with high dimensional data [48]. Kernel methods generalize linear algorithms while still relying on linear algebra. This is why we focus on two approaches for (first guess) parameter estimation: least squares (regularized) linear regression and its generalization using kernels. We will use classical linear regression (LR) and kernel methods in this work to implement the step of retrieving physical parameters from the data. We will use the Kernel Ridge Regression (KRR) algorithm [36,49], which has shown very good performance in prediction when used on IASI L1 data on different problems [37,50,51]. KRR was the method employed in [29,30] to analyze the effect of compression on physical parameters retrieval at different compression ratios. The KRR has the advantage of generalizing least squares linear regression to the nonlinear regression case.

2.2.1. Statistically Based Retrieval Methods

Here we define the notation and review the formulation of KRR from the LR solution. We will denote $\mathbf{x}_i \in \mathbb{R}^{D_x}$ as the inputs and $\mathbf{y}_i \in \mathbb{R}^{D_y}$ as the desired outputs. In matrix notation, we will denote the training input samples as $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_N] \in \mathbb{R}^{N \times D_x}$, the training desired outputs as $\mathbf{Y} = [\mathbf{y}_1 \dots \mathbf{y}_N] \in \mathbb{R}^{N \times D_y}$, the test input samples as $\mathbf{X}^* = [\mathbf{x}_1^* \dots \mathbf{x}_N^*] \in \mathbb{R}^{M \times D_x}$, and the test desired outputs $\mathbf{Y}^* = [\mathbf{y}_1^* \dots \mathbf{y}_N^*] \in \mathbb{R}^{M \times D_y}$. Using the matrix notation, the application of the linear model is:

$$\hat{\mathbf{Y}}^* = \mathbf{X}^* \mathbf{W}_L, \tag{1}$$

where we have discarded the bias term and $\mathbf{W}_L \in \mathbb{R}^{D_x \times D_y}$ are the regression weights. We fit the weights using classical least squares solution, which depends on the inversion of $(\mathbf{X}^\top \mathbf{X})$. In order to ensure inversion we use the classical Tikhonov regularized solution:

$$\mathbf{W}_L = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^\top \mathbf{Y}. \tag{2}$$

Kernel methods are based on defining a mapping function $\boldsymbol{\phi}(\mathbf{x})$ for the input samples to a Hilbert space, $\mathcal{H}$, of very high (possibly infinite) dimensionality $D_{\mathcal{H}}$. KRR can be defined following the same procedure as in the LR case as a linear least squares regression in the Hilbert space. If we map the samples using the mapping function we have $\boldsymbol{\Phi} = \boldsymbol{\phi}(\mathbf{X}) \in \mathbb{R}^{N \times D_{\mathcal{H}}}$, and $\boldsymbol{\Phi}^* = \boldsymbol{\phi}(\mathbf{X}^*) \in \mathbb{R}^{M \times D_{\mathcal{H}}}$. Therefore the prediction model is given by:

$$\hat{\mathbf{Y}}^* = \boldsymbol{\Phi}^* \mathbf{W}. \tag{3}$$

As in the LR case, we can apply the Tikhonov regularized solution to find the weights:

$$\mathbf{W} = (\boldsymbol{\Phi}^\top \boldsymbol{\Phi} + \lambda \mathbf{I})^{-1} \boldsymbol{\Phi}^\top \mathbf{Y}. \tag{4}$$

Note that this problem is not solvable as the inverse runs on matrix $\boldsymbol{\Phi}\boldsymbol{\Phi}^\top$, which is of size $D_{\mathcal{H}} \times D_{\mathcal{H}}$, and $\boldsymbol{\Phi}$ is in principle unknown. However, if we use a λ that ensures that the matrix can be inverted, the solution is equivalent to:

$$\mathbf{W} = \boldsymbol{\Phi}^\top (\boldsymbol{\Phi}\boldsymbol{\Phi}^\top + \lambda \mathbf{I})^{-1} \mathbf{Y}. \tag{5}$$

If we summarize the right part of the equation by $\boldsymbol{\alpha} = (\boldsymbol{\Phi}\boldsymbol{\Phi}^\top + \lambda \mathbf{I})^{-1} \mathbf{Y}$, the predictions for the text input samples is:

$$\mathbf{Y}_* = \boldsymbol{\Phi}_* \boldsymbol{\Phi}^\top \boldsymbol{\alpha}. \tag{6}$$

Note that even though the mapping $\boldsymbol{\Phi}$ is unknown, one can replace this inner product matrix with a similarity matrix between samples, which is known as the *kernel matrix* $\mathbf{K}$. In this case we can replace the Gram matrix as $\mathbf{K} = \boldsymbol{\Phi}\boldsymbol{\Phi}^\top$, and equivalently $\mathbf{K}_* = \boldsymbol{\Phi}_* \boldsymbol{\Phi}^\top$. Therefore the predictions can be computed by:

$$\hat{\mathbf{Y}}_* = \mathbf{K}_* \boldsymbol{\alpha}. \tag{7}$$

In KRR we only need a kernel function, $k(\mathbf{x}_i, \mathbf{x}_j)$, according to the Mercer's theorem [52]. Kernel methods literature is full of examples of proper kernel matrices. Here we will use the most standard one, the Gaussian Function (Radial Basis Function, RBF) kernel $k(\mathbf{x}_i,\mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / (2\sigma^2))$, which has only one free parameter, σ. Therefore, in LR one parameter is tuned: the regularization parameter λ. In KRR, two free parameters are tuned: the regularization parameter λ and the kernel parameter σ.

In both cases we used a cross-validation strategy for their optimization using one half of the data for training and the other half for validating. A MATLAB implementation of KRR and other regression algorithms can be found at http://isp.uv.es/soft_regression.html.

3. Data and Experimental Setting

This section introduces the data collection used in the experiments and the parameter configuration employed in the compression and in the statistically based regression stages. First, a description of the main characteristics of the IASI instrument is given, along with the specific products used in the experiments.

3.1. Data Collection

IASI is the main payload instrument carried on the MetOp satellite series [53]. Data provided by the instrument represent a significant improvement in the retrieval accuracy and vertical resolution of atmospheric parameters (temperature and water vapour concentration mostly) with respect to previous lower spectral resolution instruments, such as HIRS, SEVIRI, etc. This has in turn improved the output from numerical weather prediction models and atmospheric chemistry studies.

IASI produces rich spectral information to derive temperature and moisture profiles. The spectra is also used for the determination of trace gases such as CO, CO_2, CH_4, SO_2, N_2O, HNO_3, NH_3, OCS, and CF_4 [54–59], land and sea surface temperatures, and cloud properties.

The instrument collects data with high spectral, spatial, and temporal resolution producing large volumes of information (about 16 Gigabytes per day generated by each of the IASI-A and IASI-B instruments). IASI covers the spectral range between 645 and 2760 cm^{-1} yielding 8461 spectral components. Such an amount of information is costly to manage, hence the need to search for efficient strategies to reduce the large size of the data for improved processing, transmission, and storage.

An effective strategy to alleviate the large volume of information produced by the IASI instrument is to compress the data according to the specific requirements of the end-user applications. As outputs of the models, we use the physical variables (temperature and dew point temperature) given by the analyses of the model of the European Centre for Medium-Range Weather Forecasts (ECMWF). The model provides estimations for 137 different pressure levels between $[10^{-2} \cdots 10^{3}]$ hPa in the atmosphere and spatial resolution of 0.5 degrees. We co-registered the predictions supplied with the scenes acquired by IASI instrument with the analyses of the ECMWF.

The experiments have been conducted in a realistic context, where data from disjoint orbits are used for training and testing. The training set is composed by seven orbits and the atmospheric parameter retrieval test is carried out on six different orbits. All data have been produced by the IASI-B instrument, implemented on the MetOp-B satellite. Table 1 reports the characteristics of the 13 IASI orbits used in the experiments. In order to isolate the results for cloud free and cloudy areas we use the fractional cloud cover data provided in the L2 level [60].

Usually, the Band 3 of the IASI spectrum and some channels from the Band 1 and from the Band 2 are not used for temperature and dew point temperature retrieval because they are influenced by solar radiation and trace gases such as CO, CH_4, etc. Following the feature selection in [37] we performed feature selection removing the most noisy bands and keeping 4699.

3.2. Setting and Parameter Configuration for Lossy Compression

Lossy compression is carried out through JPEG 2000 standard. It is meaningful to consider that, in hyperspectral data with such a large number of spectral components as IASI products, it is of utmost importance to exploit the high spectral redundancy inherent to the data. As mentioned, to achieve improved coding performance, JPEG 2000 is paired along with three spectral transforms: for DWT, 10 levels of 9/7 DWT are applied in the spectral dimension; POT is run using the default parameter settings; and for Multilevel Clustering KLT, 100 clusters are defined in the first level. To account for the spatial redundancy, a 9/7 DWT with five levels is applied in the spatial dimension. Nine target bit-rates are analyzed, distributed from 2 to 0.0025 bit per pixel per component (bpppc), corresponding to compression ratios from 8:1 to 6,400:1. Table 2 summarizes the main characteristics of the three compression settings studied in the experiments.

In the experiments, Kakadu software [61] has been used for JPEG 2000, Pairwise Orthogonal Transform software [62] has been run for POT, and Spectral Transform software [63] has been employed for Multilevel

Table 1. IASI L1C products used in the experiments. Identifiers and sizes are provided. M is the number of spectral components, Ns is the number of scan lines, N-FORs is the number of elementary fields of regard (FOR) per line, and N-IFOVs is the number of instantaneous fields of view (IFOV) per FOR.

Orbit identifiers	Size (M × Ns × N-FORs × N-IFOVs)
Training set	
IASI_xxx_1C_M01_20131017012656Z_20131017030856Z_N_O_20131017021701Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017045352Z_20131017063552Z_N_O_20131017054413Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017081752Z_20131017095656Z_N_O_20131017090901Z	(8461 × 743 × 30 × 4)
IASI_xxx_1C_M01_20131017113856Z_20131017131752Z_N_O_20131017123024Z	(8461 × 742 × 30 × 4)
IASI_xxx_1C_M01_20131017145656Z_20131017163856Z_N_O_20131017155030Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017181752Z_20131017195952Z_N_O_20131017190946Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017214152Z_20131017232352Z_N_O_20131017223236Z	(8461 × 765 × 30 × 4)
Test set	
IASI_xxx_1C_M01_20131017030856Z_20131017045352Z_N_O_20131017035958Z	(8461 × 787 × 30 × 4)
IASI_xxx_1C_M01_20131017063552Z_20131017081752Z_N_O_20131017072637Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017095656Z_20131017113856Z_N_O_20131017104905Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017131752Z_20131017145656Z_N_O_20131017141023Z	(8461 × 743 × 30 × 4)
IASI_xxx_1C_M01_20131017163856Z_20131017181752Z_N_O_20131017173059Z	(8461 × 742 × 30 × 4)
IASI_xxx_1C_M01_20131017195952Z_20131017214152Z_N_O_20131017205049Z	(8461 × 765 × 30 × 4)

Table 2. Configuration of the compression setting used in the experiments.

Spectral transform		Spatial transform	Compression technique	Target bit-rates (bpppc)	Compression ratios
DWT	10 levels	5 levels of DWT	JPEG 2000	2, 0.5, 0.25, 0.1, 0.05, 0.03, 0.01, 0.005, and 0.0025	8, 32, 64, 160, 320, 533, 1,600, 3,200, and 6,400
POT	Default parameter settings				
Multilevel Clustering KLT	100 clusters in the first level and multilevel mode				

Clustering KLT. The JPEG 2000 standard provides a multicomponent extension in its Part 2 [64], which has been used to apply the DWT.

It is worth mentioning that, in our lossy compression approach, we apply a decoding stage prior to the statistically based retrieval, so that retrieval works with the reconstructed data and not in the compressed domain. These reconstructed data have exactly the same spectral and spatial dimensions as the original data, i.e., no spectral or spatial dimensionality reduction occur. It is the quality of the observations that is being modified, not the size of the scenes.

3.3. Settings of Statistically Based Regression

The retrieval experiments have been carried out in a realistic scenario, i.e., data from disjoint IASI orbits have been used for training and testing. The statistical retrieval algorithms are applied on IASI data hyperpixels at different pressure levels considering all the spectral components at a particular spatial position in the regression.

In the training stage 8,000 samples are randomly selected from each of the training orbits (seven products), which produces a training set of 56,000 samples. To allow a fair comparison, the positions of the samples used for training are kept constant for all the experiments, i.e., all three multi-component transforms and all compression ratios. Six different orbits are used for testing. The prediction is carried out on the whole orbit. To assess the retrieval performance, the RMSE between the ground-truth values of temperature/dew point temperature and the predictions produced by the regression models is computed in each pressure level.

We trained one model for each coding setting and each compression ratio. Note that the input data to the retrieval process is the reconstructed spectra, i.e., compressed and decompressed data. The models were trained using a simple cross-validation scheme with half of the data used for training and the other half used for validating the parameters.

4. Experimental Results

This section presents the experimental results. We report results for atmospheric parameter retrieval over the reconstructed spectra and investigate the origin of the improvement in the retrievals produced using the reconstructed spectra. Results are evaluated in terms of noise removal and shared spatio/spectral information to assess the impact of compression on the retrievals. This study is conducted through several experiments:

1. We evaluate the performance of the retrievals produced using reconstructed spectra. Different compression settings and retrieval algorithms are used. The experiment is carried out in a realistic scenario where data used for testing comes from orbits different from those used for training.
2. We analyze the amount of noise left in the data after being processed with each compression setting and at different compression ratios.
3. We analyze the spectral and spatial information in the compression settings. First we quantify the amount of information shared among the spatial neighbours. We then propose an scheme which specifically enforces different levels of spatial and spectral relations in a simple way, and compare the performance of the retrieval algorithms when using data processed following this scheme and when using reconstructed data from the compression algorithms.

Results are reported for one IASI L1C orbit and one physical variable (dew point temperature) in most of the experiments due to restrictions in page length. Conclusions from the results for the remainder orbits and for temperature are similar. We report all these results in the supplementary material at http://isp.uv.es/spatio_spectral_compression.html.

4.1. Retrieval Assessment

This section reports the results of estimating atmospheric physical variables through LR and KRR using reconstructed IASI L1C spectra. Results are presented for the orbit IASI_xxx_1C_M01_20131017030856Z_20131017045352Z_N_O_20131017035958Z (see table 1) and dew point temperature prediction. Results for all the other orbits and variables are very similar and reported as supplementary material.

A first assessment of the reconstructed radiances is reported in the Appendix, where the normalized radiance residuals statistics and the spectral signature of the reconstructed radiances are analyzed. Next, Figure 3 shows the retrieval performance of dew point temperature prediction for the lossy compression settings proposed. Plots show the average of the RMSE prediction over the different pressure levels between 1100 and 100 hPa. We analyze four different scenarios, i.e., land and cloud free, land and cloudy, ocean and cloud free, and ocean and cloudy conditions. Cloud free conditions is equivalent to cloud fraction equals to 0%. Otherwise, it is considered cloudy conditions. Ocean is equivalent to land fraction equals to 0%. Otherwise, it is considered land. Table 3 reports the number of spectra analyzed in each scenario. The compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data.

Table 3. Number of spectra in each of the scenarios analyzed. The percentage of the orbit in each scenario is reported in brackets.

Land - Cloud free	Land - Cloudy	Ocean - Cloud free	Ocean - Cloudy
10,218 (10.8%)	28,613 (30.3%)	6,658 (7.1%)	48,951 (51.8%)

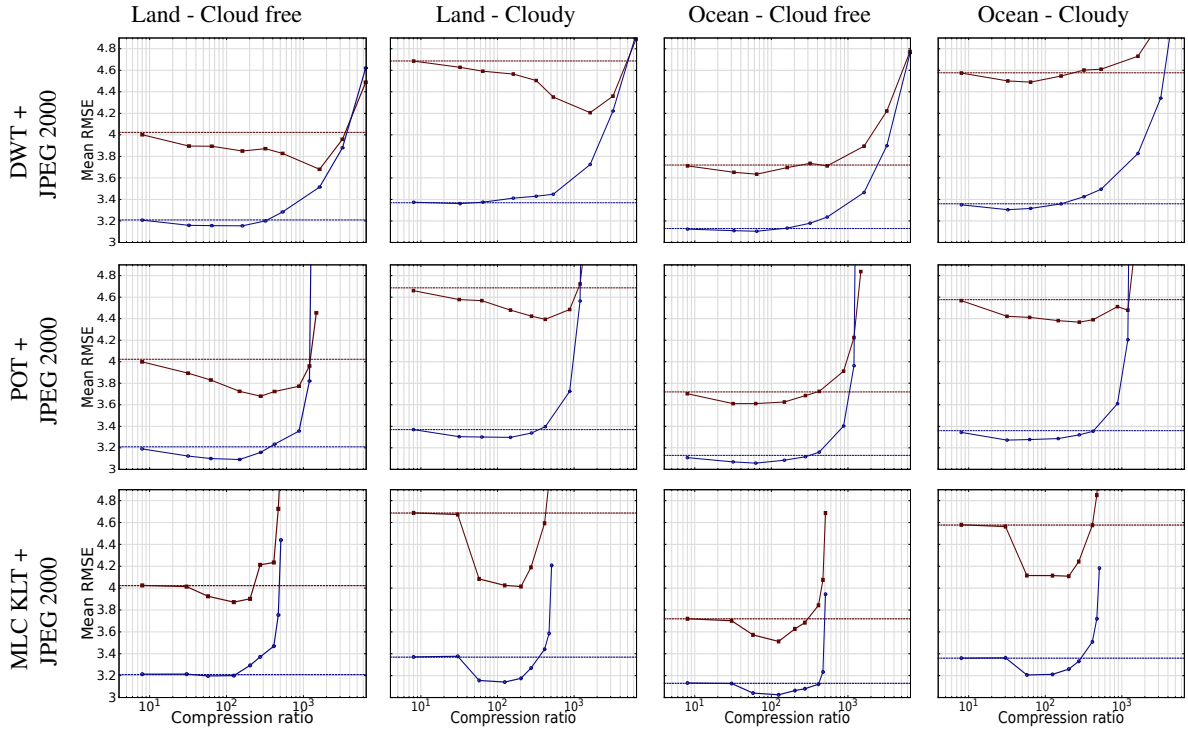


Figure 3. Dew point temperature (in kelvin) retrieval performance for different lossy compression settings using LR (solid red lines) and KRR (solid blue lines). In all the plots, the vertical axis represents the averaged RMSE over the different pressure levels between 1100 and 100 hPa and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Each row shows the results for a particular compression setting and each column shows the results for a particular scenario. Results using uncompressed spectra (original data) for LR (dashed red lines) and KRR (dashed blue lines) are plotted for comparison purposes.

One can see that while low compression ratios do not reduce the accuracy in the retrievals, the prediction results are improved as the compression ratio increases. Nonetheless, when the compression ratio is very high, the retrieval performance begins to deteriorate because too much distortion is introduced in the reconstructed data. This behaviour occurs for both LR and KRR, although the retrieval improvements are larger for LR (in the sense that higher compression ratios can be applied before the retrieval performance starts to deteriorate). These results are consistent with the observations reported in García-Sobrino *et al.* [30]. However, note that here the situation is more realistic than in García-Sobrino *et al.* [30], since we use different orbits for training, validation, and testing. By doing so we ensure that no information about the true values due to the blurring introduced by the compression settings can be used in the predictions. Results are also consistent for the new compression settings evaluated.

An interesting analysis was introduced in [25] about the correlation between the dimensions of the predictions. Figure 4 shows the value of the i_D index for the different compression configurations and the different regression algorithms taken into consideration in this paper. We provide the index without the normalization by the number of output dimensions. We can see that given a particular regression method, the differences in the correlation between the outputs for different compression ratios (and also for different compression settings) are negligible for the compression ranges with low prediction errors, i.e., until compression ratios about 300:1.

Several conclusions can be drawn:

1. *Statistical atmospheric parameter retrieval benefits from spectral/spatial lossy compression*. It is clear from Fig. 3 that using reconstructed data after a compression process allows to achieve improved statistically based retrieval performance. While moderate and medium compression ratios enable to achieve at least the same performance than the original spectra, when the distortion level introduced in the data is high, the accuracy in the retrievals decreases. In all the scenarios analyzed, a compression ratio of, approximately,

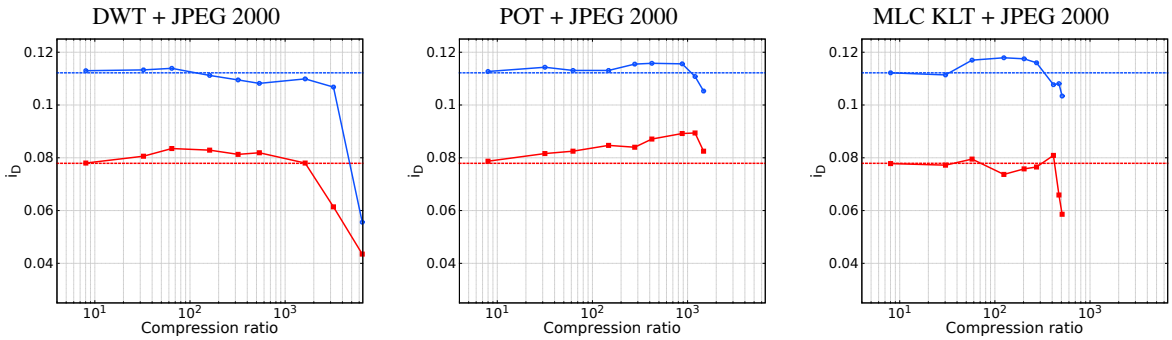


Figure 4. Index i_D for different lossy compression settings using LR (solid red lines) and KRR (solid blue lines). In all the plots, the vertical axis represents the i_D index and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Results using uncompressed spectra (original data) for LR (dashed red lines) and KRR (dashed blue lines) are plotted for comparison purposes.

100:1 allows to achieve the same or improved retrieval results compared to retrieval results obtained from uncompressed spectra (original data).

2. *KRR produces more accurate predictions than LR*. For all plots, KRR consistently yields better retrieval performance than LR. These results are constant for all the data (original and reconstructed spectra) and scenarios analyzed. The difference is specially significant for cloudy conditions.
3. *Lossy compression enables more significant retrieval improvements for LR compared to KRR*. The retrieval improvement is more evident for LR than for KRR when a compression process has taken place, which suggests that the spectral/spatial regularization produced by the compression stage particularly benefits LR.
4. *LR achieves competitive performance when the spectra have been compressed at higher compression ratios compared to KRR*. Fig. 3 illustrates that while KRR is able to achieve improved retrieval results when data compressed at a compression ratio of, approximately, 124:1 are used, LR still produces competitive performance when spectra compressed at a compression ratio of, approximately, 203:1 is employed.
5. *Multilevel Clustering KLT + JPEG 2000 leads to the best prediction results*. The compression setting Multilevel Clustering KLT + JPEG 2000 leads to the best retrieval improvements for land and cloudy, ocean and cloud free, and ocean and cloudy conditions. For land and cloud free conditions, both DWT + JPEG 2000 and POT + JPEG 2000 slightly overcome Multilevel Clustering KLT + JPEG 2000. If the interest is to optimize the retrieval performance in all the scenarios, Multilevel Clustering KLT + JPEG 2000 is the best compression setting.
6. *POT + JPEG 2000 allows to achieve higher compression ratios leading at least to the same retrieval performance than the original spectra*. If the interest is to achieve very high compression ratios, POT + JPEG 2000 would be the best compression setting. A compression ratio of 421:1 allows to achieve at least the same prediction results compared to the original data for all the scenarios analyzed. Higher compression ratios decrease the retrieval performance because useful information is missing in the recovered spectra.
7. *The effect of lossy compression is alike for all the compression settings and scenarios analyzed*. One can see in Fig. 3 that, depending on the compression setting applied and the scenario analyzed, the results can slightly vary, but the impact of lossy compression on retrieval performance is similar for all cases.
8. *The effect of the compression on the predictions correlation is negligible for useful compression ratios*. If we take in consideration the range of compression ratios for low prediction error (i.e. [0-300], see Fig. 3), the difference in correlation between the predicted outputs when using and when not using compression is very low (Fig. 4). The same effect can be observed if we compare this difference with the difference inferred by using different prediction methods.

Figure 5 shows the bias and the RMSE results for the range of pressure levels corresponding to the troposphere (between 1100 and 100 hPa). For each compression setting and each scenario, results for the compression ratio that produces the best prediction performance are reported.

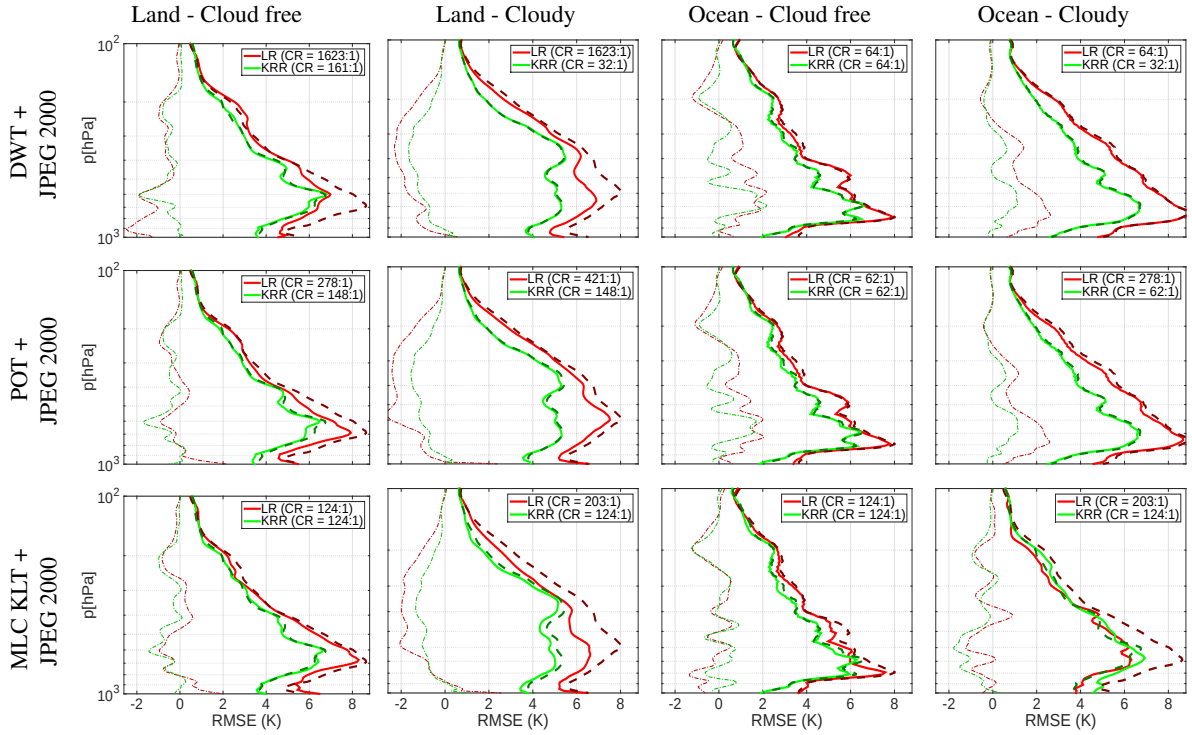


Figure 5. Dew point temperature (in kelvin) RMSE profiles and bias. The compression ratio (CR) with the best average RMSE is reported for LR (thick, solid and red lines) and KRR (thick, solid and green lines). Results using the original data are shown as well for comparison purposes when LR (thick, dashed and red lines) and KRR (thick, dashed and green lines) are used for the predictions. The bias of the reconstructed data are plotted with thin and dash-dot lines for LR (red) and KRR (green).

Reconstructed spectra produce better (or at least the same) results compared to the original data for all pressure levels and all compression settings. Generally, improvements are more evident in the mid-low troposphere. As illustrated in Fig. 3, the improved prediction is more significant for LR. Results are usually better over ocean areas than over land, and better on cloud free situations rather than in cloudy scenarios.

The Mean Absolute Error (MAE) between the dew point temperature values provided by the ECMWF analyses (and taken here as the ground truth), and the predicted values produced by LR and KRR when uncompressed and reconstructed spectra are employed in the retrievals is illustrated in Fig. 6. Maps plot the MAE for all the analyzed orbits. Results are reported only for the original data and Multilevel Clustering KLT + JPEG 2000, which is the compression technique that produces the most competitive performance, due to restrictions in page length. As expected, KRR produces smaller errors than LR (more bluish areas), and maps show that the MAE is very similar when the original data and the reconstructed spectra are used. Errors occur in the same geographic areas and the error pattern is similar for both the original and the reconstructed data.

4.2. Noise Reduction (denoising) through Compression

This section is specifically dedicated to analyze the effect of the compression process on the noise present in the data. Lossy compression is known to be an indirect way of performing signal filtration and denoising. Compression aims for a compact representation of the data by eliminating redundant information, preserving as much as possible the characteristics of the observation needed by the intended data user.

Figure 7 reports an estimate of the noise level present in the reconstructed spectra for the compression settings analyzed. In the experiments it is assumed that the noise model in the IASI data comes from an additive white Gaussian distribution, and also independent and identically distributed. This is a realistic assumption for an interferometer and is the model adopted in EUMETSAT [44], Hultberg [65], Aires *et al.* [66], Calbet *et al.* [67]. To compute the noise standard deviation, the Anisotropic Nonparametric Image Restoration toolbox [68] was

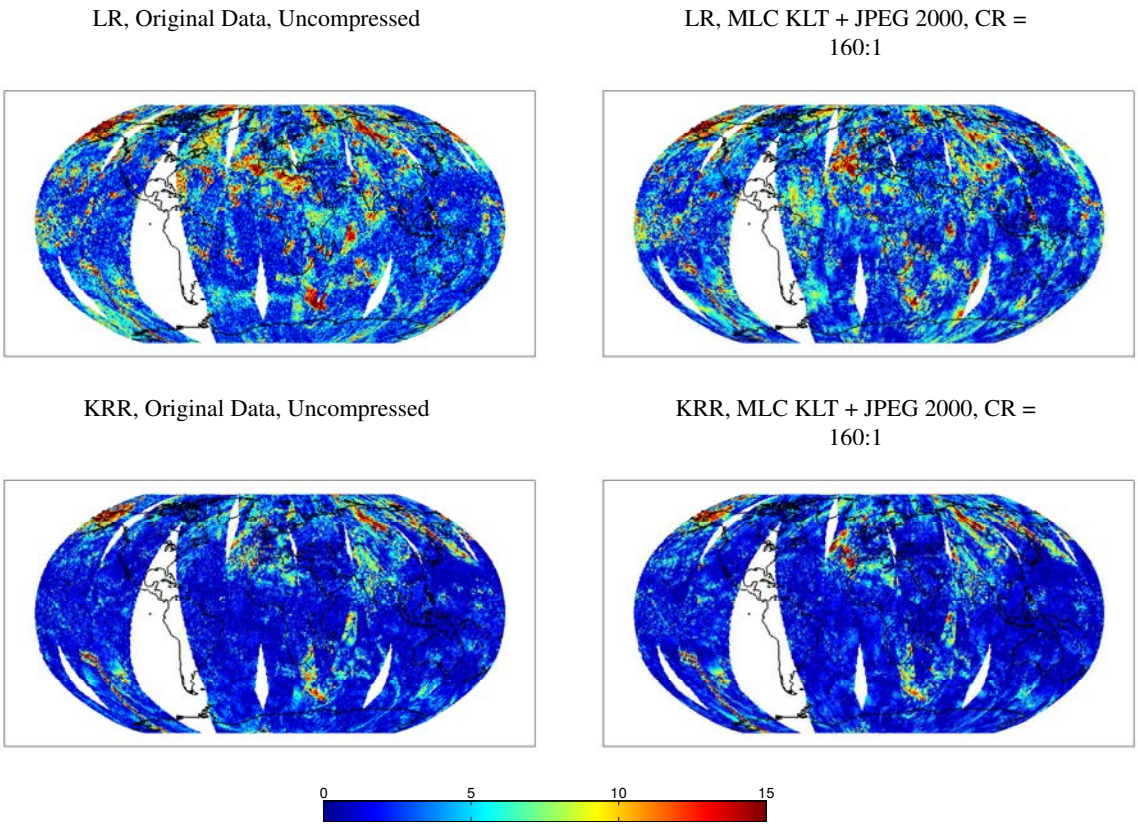


Figure 6. MAE for dew point temperature over the whole profile. Plots report MAE for uncompressed and reconstructed spectra and LR and KRR algorithms. Larger errors are represented by red color and smaller errors by blue color. CR reports the compression ratio achieved in the compression stage.

used. This method provides a robust estimate of the standard deviation based on the well-known median of the absolute deviation (MAD) [69]. This technique is used along with an orthogonal wavelet transform in such a way that the median estimator is applied on the high frequency subbands (fine details) of the transform domain to reduce the impact of the features present in the signal. In the experiments, a Daubechies wavelet transform was used. Specific and more detailed description of the adopted strategy can be found in [70,71].

We can see how part of the improvement in the retrievals is due to the filtering of noise performed by the compression process. It is interesting to compare Fig 7 with Fig 3. It is clear from Fig. 7 that, as the compression ratio increases, the noise level in the reconstructed data decreases. While most of the noise is removed at moderate compression ratios, the noise level keeps unchanged for extremely high compression ratios because almost all the noise has already been removed from the data. When data is compressed in moderate to high compression ratios (i.e., [10:1-300:1]), most of the content removed by the compression process is noise, keeping the retrieval performance competitive or even getting better results than using uncompressed data. However, if the compression ratio is extremely high (i.e., [300:1-2,000:1]), useful information is also eliminated, degrading the accuracy of the retrievals. We can suppose that noise filtering is one of the reasons why reconstructed compressed spectra from high compression ratios can result in better performance compared to uncompressed data. It is also clear that increasing the compression ratio beyond a certain point has no effect on the elimination of noise and therefore the only content removed from now on is relevant information. Therefore, the retrievals performance decreases beyond this point. Note that the compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data.

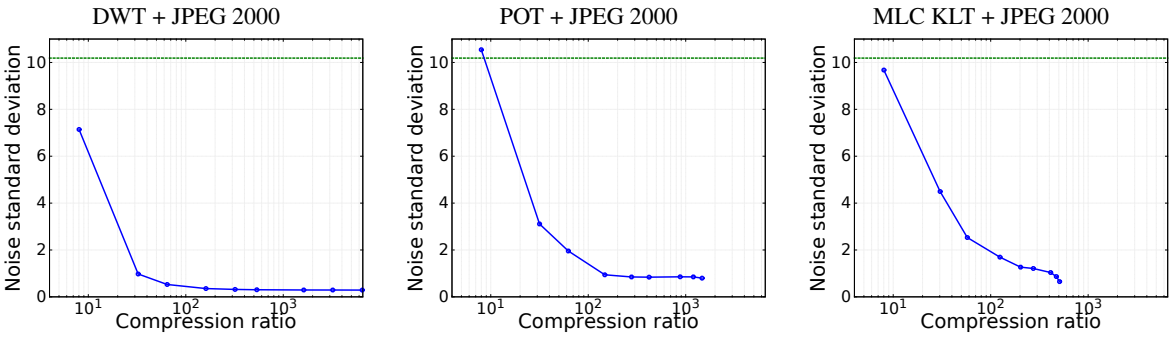


Figure 7. For each compression setting, the Noise Standard Deviation is plotted (solid blue lines). In all the plots, the vertical axis represents the noise level (noise standard deviation of IASI raw data) in the reconstructed spectra and the horizontal axis represents the compression ratio. Noise estimation for uncompressed data (dashed green lines) is reported as well for comparison purposes.

4.3. Exploiting Spectral/Spatial Relations through Compression

It is acknowledged that parameter retrieval and model inversion applications largely improve when exploiting spatial information [72–74]. Compression is an indirect way to stress relevant feature relations and to enforce smoothness in the reconstructed data, which can be exploited by pixelwise retrieval methods.

In this section we analyze the impact that the analyzed compression approaches have on sharing spectral and spatial information about the neighboring pixels and the effect it has in the retrieval performance. We first asses how much information is shared between spatial neighbours. Then we propose a simple method which specifically takes into account information from the spatial and spectral neighbours, and we compare these results with the results obtained when we do the retrieval over the recovered coefficients after compression. We also investigate how the benefits of compression (sec. 4.1) are affected when we have already exploited specifically the spatial/spectral information in the data.

4.3.1. Assessing the amount of Spectral/Spatial Information shared through Compression

Here we use information theory [75] measures in order to quantify the amount of information that spatial neighbours share when use different compression approaches and different compression ratios. In particular we compute the mutual information between one pixel and its neighbours. The same procedure has been used to measure the amount of information shared between spatial, spectral and orientation neighbours in the wavelet domain [76] and in the normalized domain [77].

In Fig. 8 we show the results when the spatial neighbours are taken into account. We computed the mutual information between each pixel and its neighbours in a neighbouring of 11×11. For the experiment we used the data contained in the whole orbit, i.e., for the computation of the mutual information we used $94,440$ samples for each component and compute the average among all the components. Particular values are different for the three compression settings analyzed, however, the trend is the same for all of them. For low to high compression ratios (i.e. [30:1-500:1]) each coefficient contains more information about its neighbours than the original scenes. When the compression ratio is extremely high (i.e., higher than 500:1), the shared information decreases dramatically. This behavior is also easy to see in the detailed squares below the graphs. In these squares we show the spatial pattern of the mutual information. It is easy to see that closer coefficients are more related (as expected), and that only the global intensity of the relations changes (i.e., the pattern of the relations is similar for different compression ratios).

4.3.2. A Simple Way to Exploit Spectral/Spatial Relations

In the following we investigate the impact that specifically exploiting spectral and spatial information has in the retrievals. Elaborated models that take into account the dependence of the the spatial smoothness with the spectral channel [78], or that avoid the neighbour pixels corrupted by clouds could be used. However, for

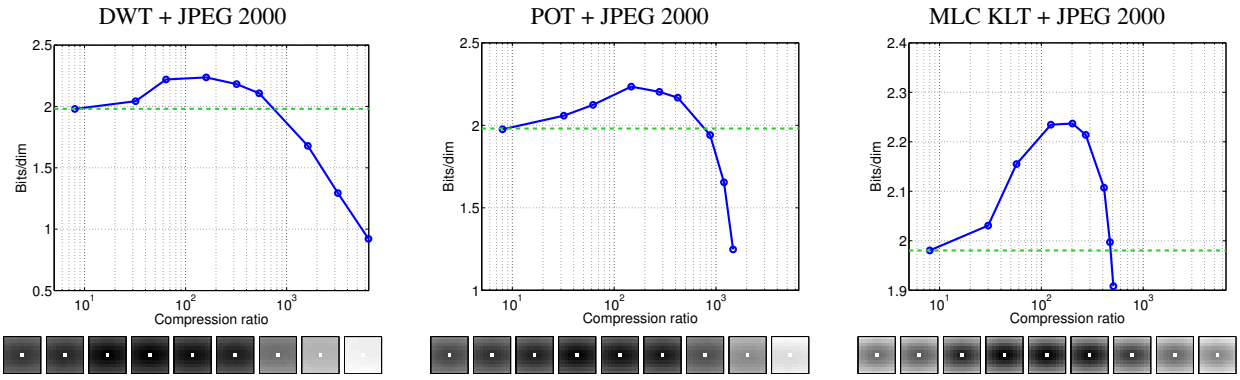


Figure 8. Spatial information shared for each compression approach and different compression ratios. Blue solid curve shows the average amount of mutual information (in number of bits) shared between one coefficient and its spatial neighbours. Green dashed line represents the same quantity for the uncompressed image. Below each plot, the detailed shared information between the central pixel and each neighbour is reported. In each square, the amount of black corresponds to the mutual information between this spatial position and the central pixel. Each square corresponds to a different compression ratio in the plot above, and are sorted left to right as the dots in the blue solid curve in the plot.

simplicity we choose here to simply convolve the data with two Gaussian filters, one over the spectral domain and one with spatial neighboring pixels. We define the σ_{sc} (spectral) and the σ_{st} (spatial) parameters for the spatial deviation of the Gaussian filter in the spectral and the spatial dimension, respectively. Combinations of five different σ_{sc} and seven different σ_{st} are evaluated. The minimum values, $\sigma_{sc} = 0.01$ and $\sigma_{st} = 0.01$, correspond to a identity function, i.e., no spectral/spatial features are taken into account.

Table 4 reports the retrieval performance for dew point temperature for different combinations of the parameters values. While exploiting both spectral and spatial features benefits the retrieval performance for LR, exploiting spectral features does not improve the results for KRR. In the case of LR, the performance is slightly improved when σ_{sc} is large and σ_{st} is small. However, for KRR the retrieval results are generally degraded when σ_{sc} is increased.

Table 4. RMSE of dew point temperature predictions (in kelvin) for different combinations of σ_{st} and σ_{sc}. The improvement over the original data is reported in percentage in brackets. Best (green) and worst (red) performance are reported.

LR

$\sigma_{st}(spatial)$ \ $\sigma_{sc}(spectral)$	0.01	0.25	1	5	10
0.01	4.515	4.515 (0%)	4.540 (-0.6%)	4.370 (3.2%)	4.362 (3.4%)
0.25	4.343 (3.8%)	4.343 (3.8%)	4.400 (2.6%)	4.254 (5.8%)	4.233 (6.2%)
1	4.061 (10.1%)	4.061 (10.1%)	4.097 (9.3%)	4.036 (10.6%)	3.996 (11.5%)
5	3.834 (15.1%)	3.835 (15.1%)	3.880 (14.1%)	3.809 (15.6%)	3.771 (16.5%)
10	3.742 (17.1%)	3.740 (17.2%)	3.795 (15.9%)	3.724 (17.5%)	3.711 (17.8%)
20	3.684 (18.4%)	3.683 (18.4%)	3.737 (17.2%)	3.697 (18.1%)	3.681 (18.4%)
30	3.669 (18.7%)	3.668 (18.8%)	3.736 (17.3%)	3.711 (17.8%)	3.694 (18.2%)

KRR

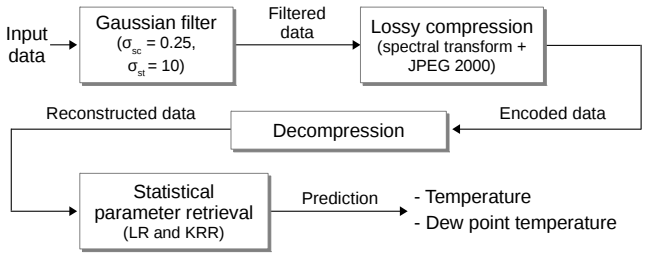
$\sigma_{st}(spatial)$ \ $\sigma_{sc}(spectral)$	0.01	0.25	1	5	10
0.01	3.345	3.345 (0%)	3.376 (-0.9%)	3.389 (-1.3%)	3.379 (-1.0%)
0.25	3.248 (2.9%)	3.248 (2.9%)	3.276 (2.1%)	3.305 (1.2%)	3.266 (2.4%)
1	3.082 (7.9%)	3.082 (7.9%)	3.109 (7.1%)	3.113 (6.9%)	3.109 (7.1%)
5	2.994 (10.5%)	2.995 (10.5%)	3.017 (9.8%)	3.028 (9.5%)	3.021 (9.7%)
10	3.007 (10.1%)	3.007 (10.1%)	3.025 (9.6%)	3.035 (9.3%)	3.031 (9.4%)
20	3.069 (8.3%)	3.070 (8.2%)	3.108 (7.1%)	3.114 (6.9%)	3.095 (7.5%)
30	3.138 (6.2%)	3.138 (6.2%)	3.139 (6.2%)	3.160 (5.5%)	3.152 (5.8%)

It is clear that spatial features are more important than spectral features, as shown in Laparra and Santos-Rodríguez [74] for hyperspectral data. It is interesting to note that, while the best prediction results for LR are yielded for large σ_{st}, the best choice for KRR is significantly smaller σ_{st}. Improvements for LR are more significant than for KRR. Results confirm the suggestion reported in García-Sobrino *et al.* [30]: exploitation of spectral and spatial regularization (smoothing) improves the retrieval results.

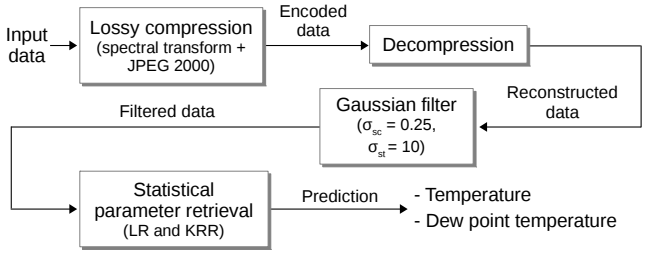
4.3.3. Combining Compression Settings with the Specific Exploitation of Spectral/Spatial Relations

In this section we show how the benefits due to compression are drastically reduced when we have already taken into account spatial/spectral information when pre-processing the data. We will show that compression is an effective, consistent, and cheap way to exploit important spectral and spatial data relations and removing noise in the reconstructed spectra. However, it is done in an indirect way and the benefits are limited.

In order to analyze the performance of compression paired along with the exploitation of spectral/spatial feature relations, we analyze several configurations. On the one hand we present results for the already explored strategies: retrieval over the original data, retrieval over the compressed and recovered data, and retrievals over the data processed using the method presented in the previous section. On the other hand we follow two extra strategies: 1) first, the Gaussian filter is applied to the original data and then, the filtered data are compressed; 2) first, the original data are compressed and then, the reconstructed spectra are filtered. Figure 9 illustrates the sequential chain of the last two strategies.



a) Spectral/spatial regularization + compression. First, we filter the data to take into account the spatial and spectral relations. Then, the filtered data are compressed and decompressed, finally statistical parameter retrieval is carried out over the reconstructed spectra.



b) Compression + spectral/spatial regularization. First, a compression and decompression stage is performed. Then, feature relations are taken into account in the reconstructed spectra using a Gaussian filter, finally statistical parameter retrieval is carried out over the filtered data.

Figure 9. Adopted sequential chain when a spectral/spatial regularization stage and a coding process are carried out.

Figure 10 and Fig. 11 illustrate the retrieval performance of dew point temperature for LR and KRR, respectively. Plots show the average of the RMSE prediction over the different pressure levels. A configuration that produces competitive performance for both LR and KRR is used in the filtering stage, i.e., $\sigma_{sc} = 0.25$ and $\sigma_{st} = 10$. Results are reported for four different scenarios: land and cloud free, land and cloudy, ocean and cloud

free, and ocean and cloudy. The compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data. Several conclusions can be extracted from this experiment:

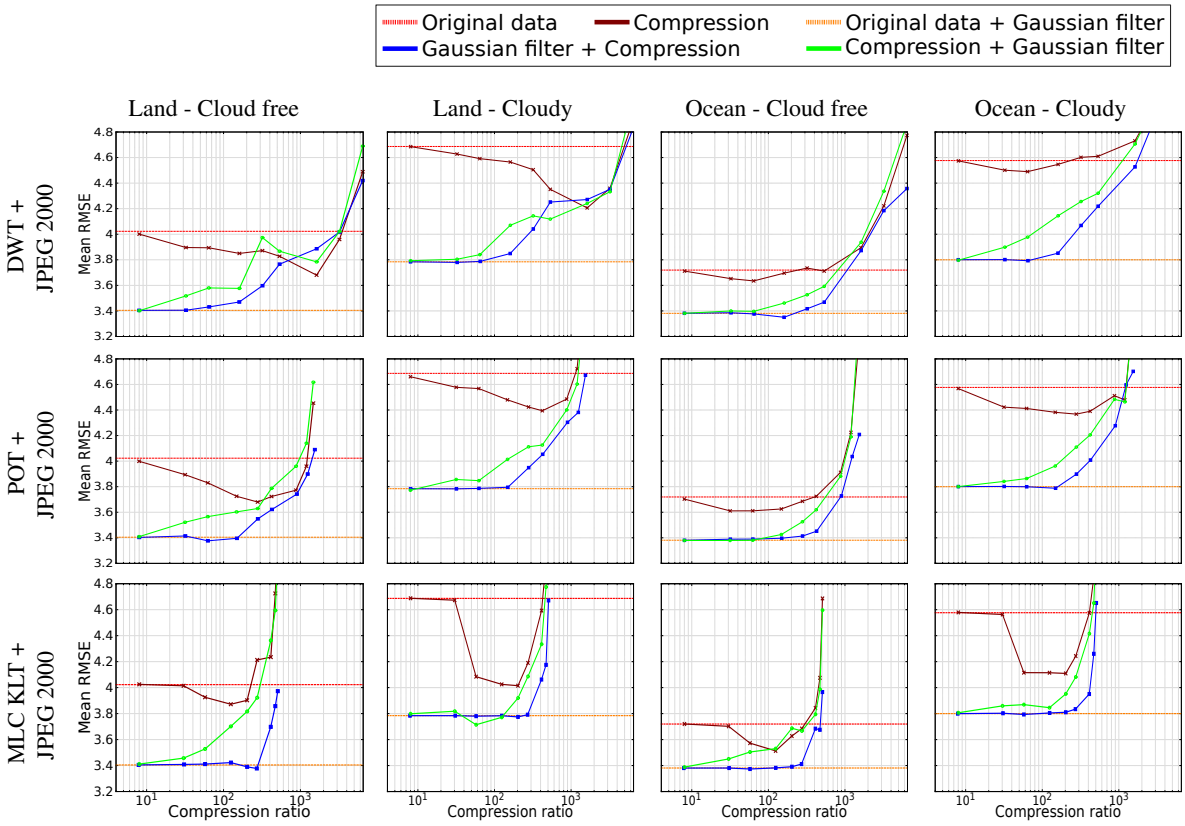


Figure 10. Dew point temperature (in kelvin) retrieval performance for different lossy compression settings using LR. In all the plots, the vertical axis represents the averaged RMSE over the different pressure levels and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Each row shows the results for a particular compression setting and each column shows the results for a particular scenario. Each plot compares five different approaches, i.e., original data, compression of the original data, original data + Gaussian filter, Gaussian filter + compression, and compression + Gaussian filter.

1. *Pre-process the data to take into account the spatial and the spectral information about the neighbours compensates the effect of using recovered data from compression.* While using recovered data (maroon curve) improves the results obtained for the original data (red curve), this improvement vanishes when spectral/spatial information is exploited, i.e., using a compression method over data that have been already pre-processed to take into account spectral/spatial information (blue line) does not improve the performance (orange line). This observation indicates that when spectral/spatial regularization is carried out before compression, the spectral/spatial transform applied in the compression stage is not able to exploit extra features.
2. *Exploiting feature relations achieves improved retrieval performance.* As we have already seen in section 4.3.2, exploiting extra features about the neighbouring pixels in the original spectra clearly benefits the prediction results. The consistency and homogeneity of the data is improved by pre-processing the data to take into account local relations between neighboring pixels, which benefits the regression algorithms. The improvement is more significant for cloudy conditions than for cloud free conditions.
3. *LR benefits more significantly from spectral/spatial regularization than KRR.* According to the results reported in the supplementary material at http://isp.uv.es/spatio_spectral_compression.html, it is clear from

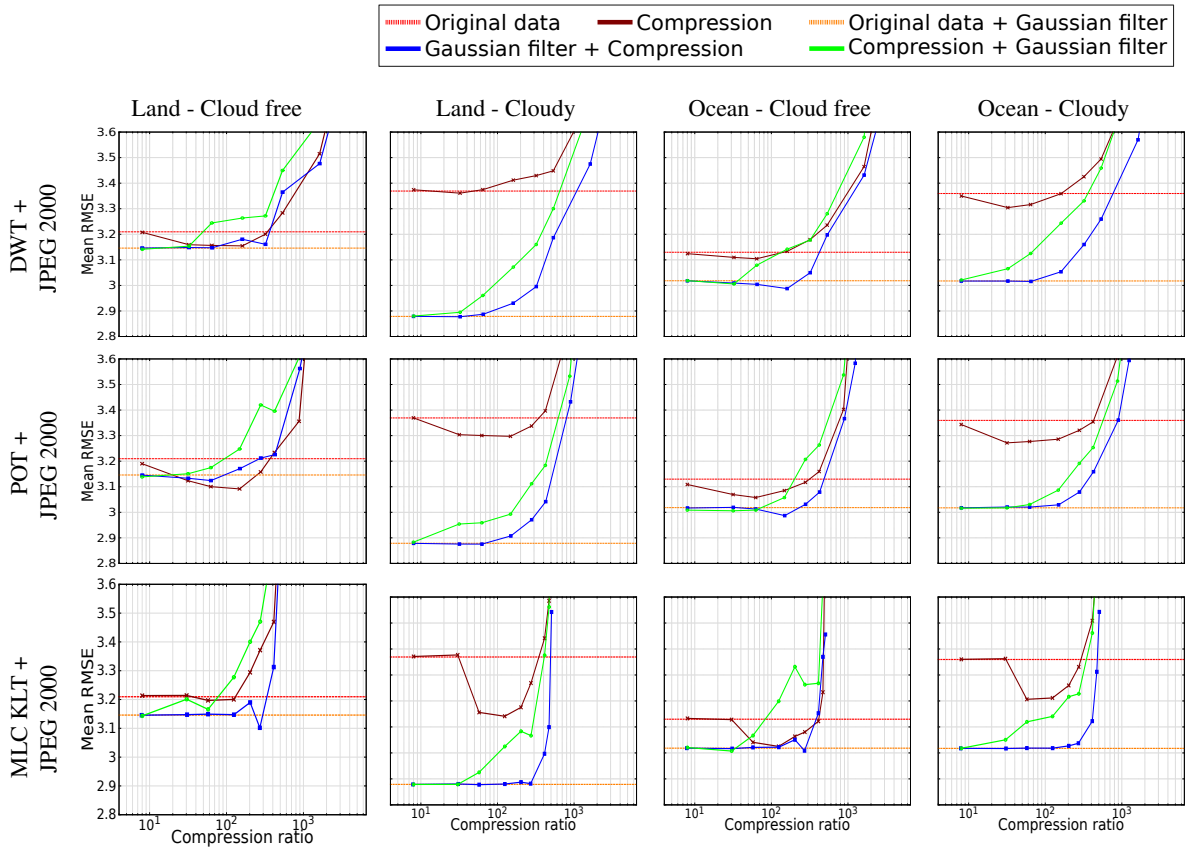


Figure 11. Dew point temperature (in kelvin) retrieval performance for different lossy compression settings using KRR. In all the plots, the vertical axis represents the averaged RMSE over the different pressure levels and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Each row shows the results for a particular compression setting and each column shows the results for a particular scenario. Each plot compares five different approaches, i.e., original data, compression of the original data, original data + Gaussian filter, Gaussian filter + compression, and compression + Gaussian filter.

the individual plots that, when spectral/spatial regularization is added, larger improvements in the retrieval performance are achieved for LR compared to KRR. Gains produced for LR are, approximately, twice the gains produced for KRR.

4. *Filtering + compression yields better results than compression + filtering*. When the exploitation of features relations is paired along with compression, it is more efficient to perform first the spectral/spatial regularization and then compress the filtered data. When the compression stage is first carried out and the reconstructed spectra are filtered, the prediction performance is degraded as the compression ratio increases. The loss of performance is clear even at low compression ratios.
5. *Results are consistent for all the approaches and scenarios analyzed*. Conclusions are similar for all the plots, which suggests that the observed behaviour is consistent for all the approaches, compression settings, and scenarios analyzed.

Figure 12 illustrates the bias and the RMSE results for the range of pressure levels between 1100 and 100 hPa. The figure illustrates that the benefits of compression are compensated by pre-processing the data to take into account neighbouring spatial and spectral features. This supports the idea that one important advantage of using recovered compressed data for retrieval is that the process exploits information about the neighbours in the data.

Both approaches achieve the similar performance for all pressure levels and scenarios, which indicates that the benefits obtained from compression were mostly due to the indirect exploitaiton of spatial and spectral neighbouring pixels information.

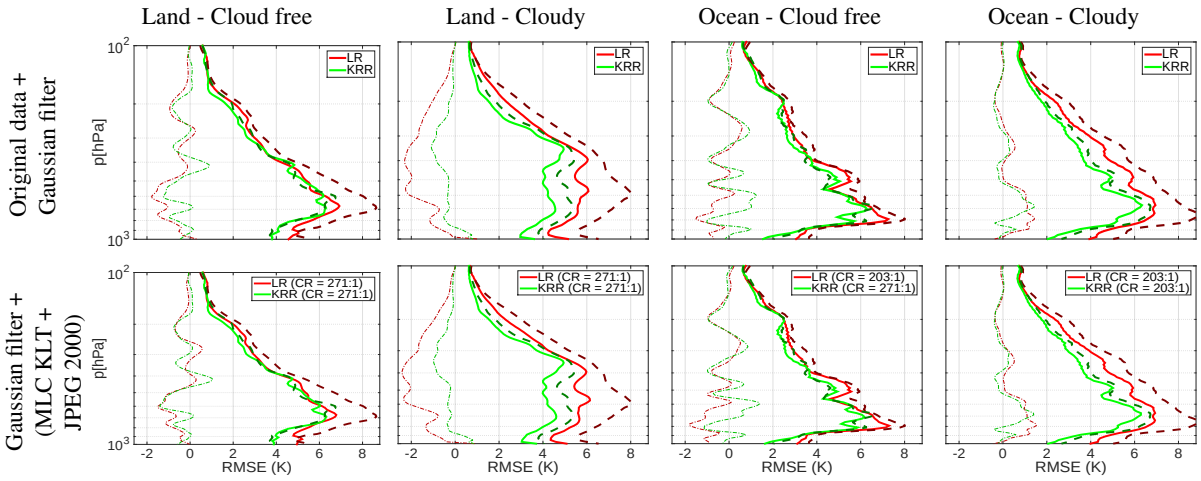


Figure 12. Dew point temperature (in kelvin) RMSE profiles and bias. Two approaches are compared: original data + Gaussian filter, and Gaussian filter + Multilevel Clustering KLT + JPEG 2000. LR (thick, solid and red lines) and KRR (thick, solid and green lines). The bias of the each approach is plotted with thin and dash-dot lines for LR (red) and KRR (green). For the compression setting, only the compression ratio (CR) with the best average RMSE is reported. Results using the original data are also displayed for comparison purposes, LR (thick, dashed and red lines) and KRR (thick, dashed and green lines).

5. Discussion

Lossy compression techniques aim at reducing the size of the transmitted/stored information keeping the most informative part of the signal. As a consequence, the data information becomes more compact, which helps the subsequent information extraction steps. The multi-component transforms employed in the compression stage in this paper are based either in the Karhunen-Loeve transform (KLT) or the wavelet transform. Both transformations have the effect of translating the data to a domain where the statistical relations (or correlations) are reduced [33,79,80]. When compressing in these domains, the less relevant features are reduced, thus stressing the important ones. When transforming back to the original domain, the reconstructed data is a version of the original data where the less relevant features have also been reduced. From the signal processing point of view, it can be seen as a filtering process.

As an example, when using KLT on images, the transformation becomes very similar to the Fourier transformation [81], i.e., the signal is decomposed in different frequencies. The high frequencies of the images have low magnitude in this domain, which in general translates into a poor signal to noise ratio –this serves to explain why most coding techniques aim at reducing the energy of these frequencies–. As a consequence, the reconstructed image has less energy in the high frequencies, which stresses the contribution of the low frequencies.

We investigated how much each of these aspects affects the performance of the statistically-based retrieval methods for sounder data. In particular, we analyzed the retrieval of physical variables (temperature and dew point temperature) from IASI L1C data.

We first verified that applying a compression process to remote sensing scenes can improve the performance of the statistically based retrieval methods. We analyzed different compression settings combined with different retrieval methods in a realistic scenario. This was illustrated in Figure 3 and Figure 5. While low compression ratios kept the results almost unchanged, moderate and high compression ratios generally enabled improved retrievals. However, when the compression ratios were very high, the retrieval performance was notably decreased, as expected.

Then we analyzed the first effect: how certain compression ratios help to remove useless information while keeping the relevant one. We analyzed the compression approaches as if they were denoising methods. When a signal is lossy compressed, the reconstructed (decompressed) signal has lost information. When this removed

559 information is mostly noise and only a small amount of useful information is removed, compression has a positive
560 impact on the regression results. We showed that for certain compression ratios and compression settings, most
561 of the removed information was noise. Figure 7 reported the noise level present in the reconstructed spectra for
562 different compression ratios and different compression settings. Plots illustrated that low compression ratios of
563 approximately 8:1 removed small amounts of noise, specially in the case of POT + JPEG 2000 and Multilevel
564 Clustering KLT + JPEG 2000. Moderate and high compression ratios were able to filter most of the noise present
565 in the data. However, when extremely high compression ratios (i.e., above 500:1) were used in the compression
566 stage, useful information was also removed and the quality of the reconstructed data was degraded. This explains
567 why reconstructed spectra produce accurate retrieval performance when the original data have been compressed
568 at moderate and high compression ratios (between 32:1 and 420:1, approximately, depending on the compression
569 setting). When extremely high compression ratios were used, the removed signal was informative and therefore
570 retrieval performance decreased.

571 The last part of the paper was focused on the second important effect of the lossy compression techniques.
572 We posit that the compaction of the information introduced by the compression has the effect of exploiting spatial
573 and spectral information about the neighbours in the reconstructed signal.

574 First, we specifically computed the amount of information about the spatial neighbours is introduced
575 when compressing the data. After that, we showed that exploiting certain spectral/spatial regularization on
576 hyperspectral data had a positive effect in the retrievals (Table 4). In particular, the spatial relations were more
577 important than the spectral relations. Then we evaluated the impact of compression in the retrievals when the
578 data was already pre-processed to take into account spectral/spatial feature relations. Results in Fig. 10 and
579 Fig. 11 showed that when the data was pre-processed the effect of compression in retrieval performance are
580 compensated. In other words, an important advantage of using recovered compressed data for retrieval is that the
581 process exploits information about the data neighbours.

582 6. Conclusions

583 This paper studied the impact of lossy compression on statistically based regression algorithms for retrieving
584 atmospheric profiles of temperature and dew point temperature using infrared sounder data (IASI L1C). We
585 analyzed the reasons behind the benefits produced by compression and provided recommendations for effective
586 prediction performance and data compression.

587 Several compression settings were evaluated and a linear and a nonlinear regression algorithm were assessed
588 in a realistic training/testing scheme. Results for the prediction of two physical variables (temperature and dew
589 point temperature) and four different scenarios (land and cloud free, land and cloudy, ocean and cloud free, and
590 ocean and cloudy) were reported.

591 Lossy compression was carried out through different compression settings, always within the scope of
592 international standard JPEG 2000. To achieve competitive compression performance in hyperspectral data with a
593 large number of spectral components (such as IASI data), three spectral transforms were paired along with JPEG
594 2000, i.e., DWT, POT, and Multilevel Clustering KLT. In the experiments, nine compression ratios were analyzed
595 to search for an optimal trade-off between retrieval performance and data compression.

596 Experimental results revealed that reconstructed spectra have enough quality to achieve competitive retrieval
597 performance because noise filtering is carried out in the compression stage, which allows to achieve high
598 compression ratios while retaining as much features as possible. Another positive effect arises from the ability of
599 compression to exploit spectral and spatial feature relations in an indirect way, which benefits the retrieval methods
600 (on average and using ECMWF analysis as ground truth). We observed that moderate-to-high compression
601 ratios produced improved results in predicting temperature and dew point temperature in the different scenarios
602 analyzed. As expected, when the compression ratio is extremely high, the benefits disappear because large
603 amounts of useful information are removed from the data.

604 Results reported that exploiting spatial relations between neighboring pixels is more productive than
605 exploiting spectral relations. While spectral regularization kept the results almost unchanged, spatial
606 regularization improved the predictions by almost 20%. Spatial regularization is, hence, a key element to
607 be exploited in pixelwise retrieval algorithms, where the spatial component is missing.

Experiments revealed that when the data is pre-processed to take into account the spectral and spatial feature relations before compression, the retrieval performance was improved compared to the results yielded by the original data even at high compression ratios. The compression setting Multilevel Clustering KLT + JPEG 2000 was able to significantly improve the atmospheric predictions at compression ratios higher than 200:1.

The proposed methodology can be applied to other retrieval methods and benefits the development of current and upcoming infrared sounding instruments. While current retrieval methods would benefit from efficiently compressed spectra, savings in data transmission and storage would involve operational improvements.

This study may have a deep impact on both currently flying infrared sounding instruments (e.g., IASI and CrIS) or upcoming (e.g., MTG-IRS).

Acknowledgments: This work was supported in part by the Spanish Ministry of Economy and Competitiveness and by the European Regional Development Fund under Grants TIN2015-71126-R, by the Catalan Government under Grant 2017SGR-463, and by the European Research Council under Consolidator Grant SEDAL ERC-2014-CoG 647423.

Author Contributions: All co-authors of this manuscript significantly contributed to all phases of the investigation. They contributed equally to the preparation, analysis, review and editing of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Bibliography

1. Chang, C. *Hyperspectral Data Exploitation: Theory and Applications*; John Wiley & Sons, 2007.
2. García-Vílchez, F.; Muñoz-Marí, J.; Zortea, M.; Blanes, I.; González-Ruiz, V.; Camps-Valls, G.; Plaza, A.; Serra-Sagristà, J. On the impact of lossy compression on hyperspectral image classification and unmixing. *IEEE Geoscience and Remote Sensing Letters* **2011**, *8*, 253–257.
3. Qian, S.E.; Bergeron, M.; Cunningham, I.; Gagnon, L.; Hollinger, A. Near lossless data compression onboard a hyperspectral satellite. *IEEE Transactions on Aerospace and Electronic Systems* **2006**, *42*, 851–866.
4. Ryan, M.J.; Arnold, J.F. Lossy compression of hyperspectral data using vector quantization. *Remote Sensing of Environment* **1997**, *61*, 419–436.
5. Sánchez, J.; Perronnin, F. High-dimensional signature compression for large-scale image classification. Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011, pp. 1665–1672.
6. Minguillon, J.; Pujol, J.; Serra-Sagristà, J.; Ortuno, I.; Guitart, P. Adaptive lossy compression and classification of hyperspectral images. Proc. of Image and Signal Processing for Remote Sensing VI, 2001, pp. 214–226.
7. Kaarna, A.; Toivanen, P.; Keränen, P. Compression and classification methods for hyperspectral images. *Pattern Recognition and Image Analysis* **2006**, *16*, 413–424.
8. Zabala, A.; Pons, X.; Díaz-Delgado, R.; García, F.; Aulí-Llinàs, F.; Serra-Sagristà, J. Effects of JPEG and JPEG 2000 lossy compression on remote sensing image classification for mapping crops and forest areas. Proc. of IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS), 2006, pp. 790–793.
9. Blanes, I.; Serra-Sagristà, J. Quality evaluation of progressive lossy-to-lossless remote-sensing image coding. 16th IEEE International Conference in Image Processing (ICIP), 2009, pp. 3709–3712.
10. Álvarez-Cortés, S.; Amrani, N.; Hernández-Cabronero, M.; Serra-Sagristà, J. Progressive Lossy-to-Lossless Coding of Hyperspectral Images through Regression Wavelet Analysis. *Taylor and Francis International Journal of Remote Sensing* **2018**, *39*.
11. Pal, M.; Brislawn, C.; Brumby, S. Feature extraction from hyperspectral images compressed using the JPEG 2000 standard. Fifth IEEE Southwest Symposium in Image Analysis and Interpretation, 2002, pp. 168–172.
12. Du, Q.; Fowler, J.E. Hyperspectral image compression using JPEG 2000 and principal component analysis. *IEEE Geoscience and Remote Sensing Letters* **2007**, *4*, 201–205.
13. Blanes, I.; Zabala, A.; Moré, G.; Pons, X.; Serra-Sagristà, J. Classification of hyperspectral images compressed through 3D-JPEG 2000. *Knowledge-Based Intelligent Information and Engineering Systems. Springer Berlin Heidelberg* **2008**, pp. 416–423.
14. Carvajal, G.; Penna, B.; Magli, E. Unified lossy and near-lossless hyperspectral image compression based on JPEG 2000. *IEEE Geoscience and Remote Sensing Letters* **2008**, *5*, 593–597.
15. Penna, B.; Tillo, T.; Magli, E.; Olmo, G. Transform coding techniques for lossy hyperspectral data compression. *IEEE Transactions on Geoscience and Remote Sensing* **2007**, *45*, 1408–1421.
16. Rodgers, C.D. *Inverse methods for atmospheric sounding : theory and practice*; World Scientific Publishing, 2000.

17. Camps-Valls, G.; Muñoz-Marí, J.; Gómez-Chova, L.; Guanter, L.; Calbet, X. Nonlinear statistical retrieval of atmospheric profiles from MetOp-IASI and MTG-IRS infrared sounding data. *IEEE Transactions on Geoscience and Remote Sensing* **2012**, *50*, 1759–1769.
18. Aires, F. A regularized neural net approach for retrieval of atmospheric and surface temperatures with the IASI instrument. *Journal of Applied Meteorology* **2002**, *41*, 144–159.
19. Blackwell, W. A neural-network technique for the retrieval of atmospheric temperature and moisture profiles from high spectral resolution sounding data. *IEEE Transactions on Geoscience and Remote Sensing* **2005**, *43*.
20. Blackwell, W.; Pieper, M.; Jairam, L. Neural network estimation of atmospheric profiles using AIRS/IASI/AMSU data in the presence of clouds. Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques, and Applications II, Proceedings of SPIE Vol. 7149; Suzuki, A.M.L.M.J.L.M., Ed.; , 2008.
21. Laparra, V.; Malo, J.; Camps-Valls, G. Dimensionality reduction via regression in hyperspectral imagery. *IEEE Journal on Selected Topics in Signal Processing* **2015**, *9*, 1026–1036. cited By 1.
22. Garcia Sobrino, J.; Serra-Sagrista.; Laparra, V.; Calbet, X.; Camps-Valls, G. Statistical Atmospheric Parameter Retrieval Largely Benefits from Spatial-Spectral Image Compression. *IEEE Transactions on Geoscience and Remote Sensing* **2017**, *PP*, 1–12.
23. Malmgren-Hansen, D.; Laparra, V.; Nielsen, A.; Calbet, X.; Camps-Valls, G. Temperature Retrieval with Spatial Noise-aware dimensionality reduction. 2017 EUMETSAT Meteorological Satellite Conference; , 2017.
24. Camps, G.; Laparra, V.; Muñoz, J.; Gómez, L.; Calbet, X.; Hultberg, T.; August, T. Advances in Non-linear Retrievals for IASI and MTG-IRS Hyperspectral Infrared Sounders. EUMETSAT Meteorological Satellite Conference; , 2014; pp. 22–26.
25. Amato, U.; Antoniadis, A.; De Feis, I.; Masiello, G.; Matricardi, M.; Serio, C. Technical Note: Functional sliced inverse regression to infer temperature, water vapour and ozone from IASI data. *Atmospheric Chemistry and Physics* **2009**, *9*, 5321–5330.
26. Arenas-García, J.; Petersen, K.; Camps-Valls, G.; Hansen, L. Kernel multivariate analysis framework for supervised subspace learning: A tutorial on linear and kernel multivariate methods. *IEEE Signal Processing Magazine* **2013**, *30*, 16–29. cited By 17.
27. Amato, U.; Antoniadis, A.; De Feis, I.; Masiello, G.; Matricardi, M.; Serio, C. Technical Note: Functional sliced inverse regression to infer temperature, water vapour and ozone from IASI data. *Atmospheric Chemistry and Physics* **2009**, *9*, 5321–5330.
28. Huang, H.L.; Antonelli, P. Application of principal component analysis to high-resolution infrared measurement compression and retrieval. *J. Appl. Met.* **2001**, *40*, 365–388.
29. García-Sobrino, J.; Blanes, I.; Laparra, V.; Camps-Valls, G.; Serra-Sagristà, J. Impact of Near-Lossless Compression of IASI L1C data on Statistical Retrieval of Atmospheric Profiles. Proc. of On-Board Payload Data Compression Workshop (OBPDC); , 2014.
30. García-Sobrino, J.; Serra-Sagristà, J.; Laparra, V.; Calbet, X.; Camps-Valls, G. Statistical Atmospheric Parameter Retrieval Largely Benefits From Spatial-Spectral Image Compression. *IEEE Transactions on Geoscience and Remote Sensing* **2017**, *55*, 2213–2224.
31. EUMETSAT. IASI Instrument, 2018. Available: https://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/IASI/index.html (accessed on January 2018).
32. Taubman, D.; Marcellin, M. *JPEG 2000 image compression fundamentals, standards and practice: image compression fundamentals, standards and practice*; Springer Science & Business Media, 2002.
33. Daubechies, I. *Ten Lectures on Wavelets*; Society for Industrial and Applied Mathematics, 1992.
34. Blanes, I.; Serra-Sagristà, J. Pairwise orthogonal transform for spectral image coding. *IEEE Transactions on Geoscience and Remote Sensing* **2011**, *49*, 961–972.
35. Blanes, I.; Serra-Sagristà, J. Cost and Scalability Improvements to the Karhunen-Loêve Transform for Remote-Sensing Image Coding. *IEEE Transactions on Geoscience and Remote Sensing* **2010**, *48*, 2854–2863.
36. Shawe-Taylor, J.; Cristianini, N. *Kernel Methods for Pattern Analysis*; Cambridge University Press, 2004.
37. Camps-Valls, G.; Munoz-Mari, J.; Gomez-Chova, L.; Guanter, L.; Calbet, X. Nonlinear statistical retrieval of atmospheric profiles from MetOp-IASI and MTG-IRS infrared sounding data. *IEEE Transactions on Geoscience and Remote Sensing* **2012**, *50*, 1759–1769.
38. Surrey Satellite Technology Limited. BILSAT-1: The Mission. Available: https://www.sstl.co.uk/Missions/BILSAT-1-Launched-2003/BILSAT-1/BILSAT-1-The-Mission (accessed on March 2018).

39. Sharing Earth Observation Resources. THEOS. Available: https://directory.eoportal.org/web/eoportal/satellite-missions/t/theos (accessed on March 2018).
40. Ismailoglu, N.; Benderli, O.; Yesil, S.; Sever, R.; Okcan, B.; Oktem, R. GEZGIN-2: an advanced image processing subsystem for earth-observing small satellites. Proceedings of 2nd IEEE International Conference on Recent Advances in Space Technologies, RAST 2005, 2005, pp. 605–610.
41. Sharing Earth Observation Resources. PICARD. Available: https://directory.eoportal.org/web/eoportal/satellite-missions/p/picard (accessed on March 2018).
42. Geospatial Data Abstraction Library. SENTINEL-2 Products. Available: http://www.gdal.org/frmt_sentinel2.html (accessed on March 2018).
43. EUMETSAT. IASI Level 1 PCC Product Generation Specification, EUM/OPS-EPS/SPE/08/0199, 2010.
44. EUMETSAT. EPS Product Validation Report: IASI L1 PCC PPF, EUM/OPS-EPS/REP/10/0148, 2010.
45. Salomon, D. *Data compression: the complete reference*; Springer Science & Business Media, 2004.
46. Chang, L.; Cheng, C.M.; Chen, T.C. An efficient adaptive KLT for multispectral image compression. Proceedings of 4th IEEE Southwest Symposium on Image Analysis and Interpretation; , 2000.
47. García-Sobrino, J.; Serra-Sagristà, J.; Bartrina-Rapesta, J. Hyperspectral IASI L1C Data Compression. *Sensors* **2017**, *17*, 1404.
48. Gómez-Chova, L.; Muñoz-Marí, J.; Laparra, V.; Malo-López, J.; Camps-Valls, G., A Review of Kernel Methods in Remote Sensing Data Analysis. In *Optical Remote Sensing: Advances in Signal Processing and Exploitation Techniques*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011; pp. 171–206.
49. Schölkopf, B.; Smola, A. *Learning with Kernels – Support Vector Machines, Regularization, Optimization and Beyond*; MIT Press Series, 2002.
50. Camps-Valls, G.; Verrelst, J.; Munoz-Mari, J.; Laparra, V.; Mateo-Jimenez, F.; Gomez-Dans, J. A Survey on Gaussian Processes for Earth-Observation Data Analysis: A Comprehensive Investigation. *IEEE Geoscience and Remote Sensing Magazine* **2016**, *4*, 58–78.
51. Laparra, V.; Muñoz-Marí, J.; Gómez-Chova, L.; Calbet, X.; Camps-Valls, G. Nonlinear Statistical Retrieval of Surface Emissivity from IASI data. 2017 IEEE International Geoscience and Remote Sensing Symposium; , 2017.
52. Aizerman, M.A.; Braverman, E.M.; Rozoner, L. Theoretical Foundations of the Potential Function Method in Pattern Recognition Learning. *Automation and remote Control* **1964**, *25*, 821–837.
53. Hilton, F.; R. Armante, T.A.; Barnet, C.; Bouchard, A.; Camy-Peyret, C.; Capelle, V.; Clarisse, L.; Clerbaux, C.; Coheur, P.; Collard, A.; Crevoisier, C.; Dufour, G.; Edwards, D.; Faijan, F.; Fourrié, N.; Gambacorta, A.; Goldberg, M.; Guidard, V.; Hurtmans, D.; Illingworth, S.; Jacquinet-Husson, N.; Kerzenmacher, T.; Klaes, D.; Lavanant, L.; Masiello, G.; Matricardi, M.; McNally, A.; Newman, S.; Pavelin, E.; Payan, S.; Péquignot, E.; Peyridieu, S.; Phulpin, T.; Remedios, J.; Schlüssel, P.; Serio, C.; Strow, L.; Stubenrauch, C.; Taylor, J.; Tobin, D.; Wolf, W.; Zhou, D. Hyperspectral Earth Observation from IASI: Five Years of Accomplishments. *Bulletin of the American Meteorological Society* **2012**, *93*, 347–370.
54. George, M.; Clerbaux, C.; Hurtmans, D.; Turquety, S.; Coheur, P.F.; Pommier, M.; Hadji-Lazaro, J.; Edwards, D.P.; Worden, H.; Luo, M.; Rinsland, C.; McMillan, W. Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors. *Atmospheric Chemistry and Physics* **2009**, *9*, 8317–8330.
55. Clarisse, L.; Coheur, P.F.; Prata, A.J.; Hurtmans, D.; Razavi, A.; Phulpin, T.; Hadji-Lazaro, J.; Clerbaux, C. Tracking and quantifying volcanic SO2 with IASI, the September 2007 eruption at Jebel at Tair. *Atmospheric Chemistry and Physics* **2008**, *8*, 7723–7734.
56. Clerbaux, C.; Boynard, A.; Clarisse, L.; George, M.; Hadji-Lazaro, J.; Herbin, H.; Hurtmans, D.; Pommier, M.; Razavi, A.; Turquety, S.; Wespes, C.; Coheur, P.F. Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder. *Atmospheric Chemistry and Physics* **2009**, *9*, 6041–6054.
57. Wespes, C.; Hurtmans, D.; Clerbaux, C.; Santee, M.L.; Martin, R.V.; Coheur, P.F. Global distributions of nitric acid from IASI/MetOp measurements. *Atmospheric Chemistry and Physics* **2009**, *9*, 7949–7962.
58. Pommier, M.; Law, K.S.; Clerbaux, C.; Turquety, S.; Hurtmans, D.; Hadji-Lazaro, J.; Coheur, P.F.; Schlager, H.; Ancellet, G.; Paris, J.D.; Nédélec, P.; Diskin, G.S.; Podolske, J.R.; Holloway, J.S.; Bernath, P. IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns. *Atmospheric Chemistry and Physics* **2010**, *10*, 10655–10678.
59. Liuzzi, G.; Masiello, G.; Serio, C.; Venafra, S.; Camy-Peyret, C. Physical inversion of the full ASI spectra: Assessment of atmospheric parameters retrievals, consistency of spectroscopy and forward modelling. *Journal of Quantitative Spectroscopy and Radiative Transfer* **2016**, *182*, 128–157.

60. EUMETSAT. *IASI Level 2: Product Guide*, 2017. EUM/OPS-EPS/MAN/04/0033,.
61. Taubman, D.S. Kakadu software, 2018. Available: http://www.kakadusoftware.com/ (accessed on January 2018).
62. GICI-UAB. Pairwise Orthogonal Transform (POT) software, 2018. Available: http://gici.uab.cat/GiciWebPage/downloads.php#pot (accessed on January 2018).
63. GICI-UAB. Spectral Transform software, 2018. Available: http://gici.uab.cat/GiciWebPage/downloads.php#spectral (accessed on January 2018).
64. JPEG. *JPEG 2000 Part 2 - Extensions*, 2004. Document ISO/IEC 15444-2, 2004.
65. Hultberg, T. IASI Principal Component Compression (IASI PCC) FAQ, 2009. Available: http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=pdf_ipcc_faq&RevisionSelectionMethod=LatestReleased&Rendition=Web (accessed on January 2018).
66. Aires, F.; Rossow, W.B.; Scott, N.A.; Chedin, A. Remote sensing from the infrared atmospheric sounding interferometer instrument 1. Compression, denoising, and first-guess retrieval algorithms. *Journal of Geophysical Research: Atmospheres* **2002**, *107*.
67. Calbet, X.; Sanz, J.; Martínez-Rubio, M.A. Scientific Report on noise characteristics dependency of the decompressed radiances when using different principal components compression methods and its application to MTG-IRS, 2017. Technical Report: NWC/CDOP3/MTG/AEMET/SCI/RP/PCnoise, AEMET.
68. LASIP. Anisotropic Nonparametric Image Restoration Toolbox, 2018. Available: http://www.cs.tut.fi/~lasip/2D/ (accessed on January 2018).
69. Hampel, F.R. The influence curve and its role in robust estimation. *Journal of the american statistical association* **1974**, *69*, 383–393.
70. Mosteller, F.; Tukey, J. *Data analysis and regression: a second course in statistics*; Addison-Wesley Series in Behavioral Science: Quantitative Methods, 1977.
71. Foi, A. Anisotropic nonparametric image processing: theory, algorithms and applications. PhD thesis, Dip. di Matematica, Politecnico di Milano, 2005.
72. Camps-Valls, G.; Tuia, D.; Gómez-Chova, L.; Jiménez, S.; Malo, J. *Remote Sensing Image Processing. Synthesis Lectures on Image, Video, and Multimedia Processing*; Morgan & Claypool Publishers, 2011.
73. Jiménez, S.; Malo, J. The Role of Spatial Information in Disentangling the Irradiance-Reflectance-Transmittance Ambiguity. *IEEE Transactions on Geoscience and Remote Sensing* **2014**, *52*, 4881–4894.
74. Laparra, V.; Santos-Rodríguez, R. Spatial/spectral information trade-off in hyperspectral images. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 1124–1127.
75. Shannon, C.E. A Mathematical Theory of Communication. *Bell Systems Technical Journal* **1948**, *27*, 623–656.
76. Laparra, V.; Gutiérrez, J.; Camps-Valls, G.; Malo, J. Image denoising with kernels based on natural image relations. *Journal of Machine Learning Research* **2010**, *11*, 873–903. cited By 9.
77. Malo, J.; Laparra, V. Psychophysically tuned divisive normalization approximately factorizes the PDF of natural images. *Neural Computation* **2010**, *22*, 3179–3206. cited By 14.
78. Plokhenko, Y.; Menzel, W.P.; Revercomb, H.E.; Borbas, E.; Antonelli, P.; Weisz, E. Analysis of multispectral fields of satellite IR measurements: Using statistics of second spatial differential of spectral fields for measurement characterization. *International Journal of Remote Sensing* **2008**, *29*, 2105–2125.
79. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psych.* **1933**, *24*.
80. Hyvärinen, A.; Oja, E. Independent Component Analysis: Algorithms and Applications. *Neural Netw.* **2000**, *13*, 411–430.
81. Hyvrinen, A.; Hurri, J.; Hoyer, P.O. *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision.*, 1st ed.; Springer Publishing Company, Incorporated, 2009.

Appendix Quality of Reconstructed Radiances

The quality of reconstructed radiances has been analyzed from two points of view. On the one hand, we examine the residuals computed as the difference between the original and the reconstructed radiances. On the other hand, the spectral signature of the reconstructed radiances is compared to the spectral signature of the original radiances.

Figure A1 illustrates the mean and the standard deviation of the noise normalized residuals for different compression settings. This strategy is used in [44,67] to validate reconstructed radiances from Principal Component Compression (PCC). One can see that as the compression ratio increases, both the mean and the

standard deviation increase. For Multilevel Clustering KLT + JPEG 2000, the mean is close to zero and the standard deviation is close to one even at high compression ratios (i.e., 533:1). Very high compression ratios (3,200:1) increase both the mean and the standard deviation. For DWT + JPEG 2000 and POT + JPEG 2000, the mean is close to zero at low compression ratios (8:1). When the compression ratio is higher than 533:1, the mean increases for most components. The standard deviation is only close to one at a compression ratio of 8:1.

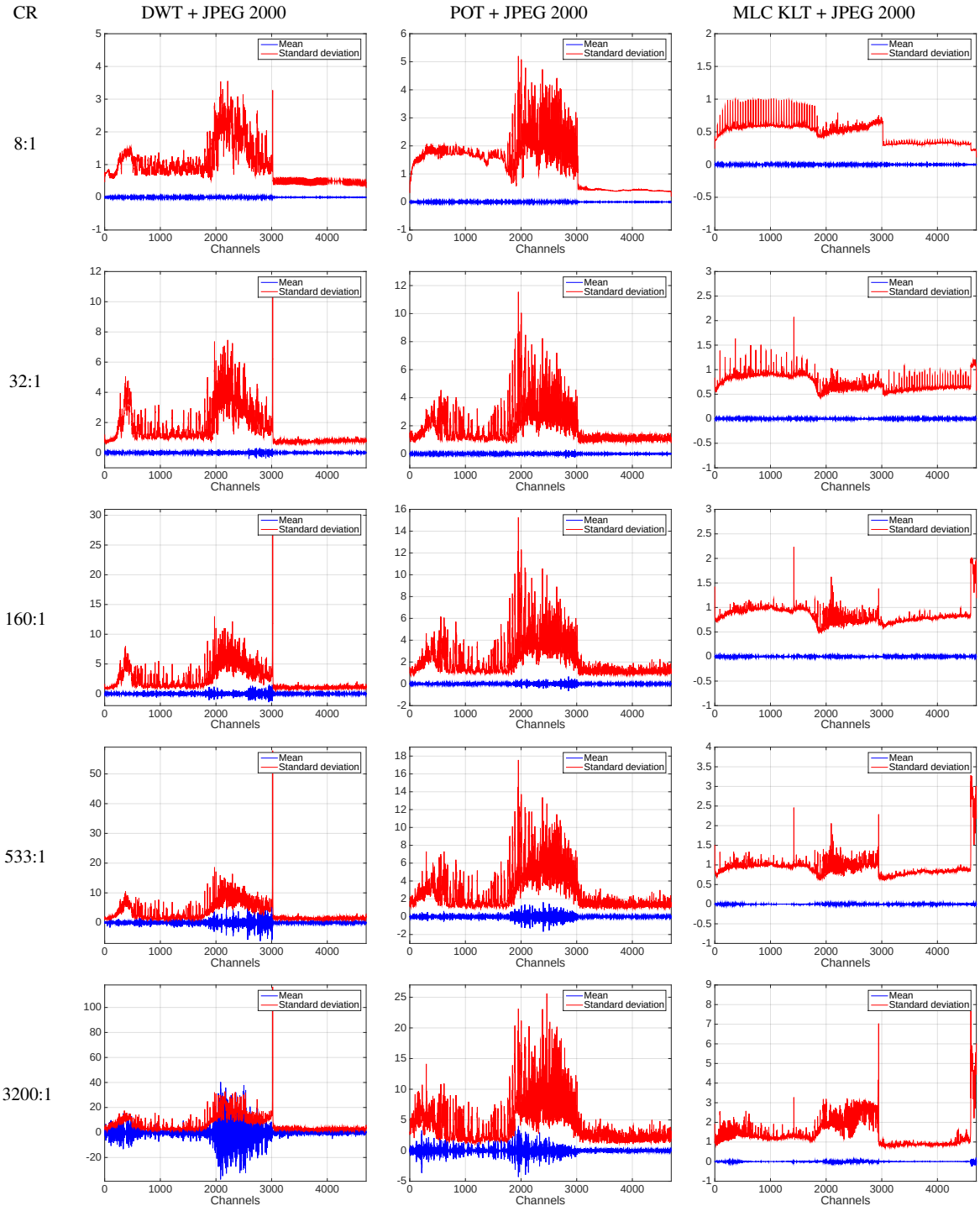


Figure A1. Radiance residuals statistics as a function of channel number. The mean of the radiance residuals of the normalized radiances is plotted in blue and the standard deviation in red. Following the feature selection in [37], 4699 channels are used. Results for different compression ratios (CR) are reported.

Figure A2 reports the average of the standard deviation and the mean (absolute value) for each compression setting. While DWT + JPEG 2000, significantly increases both the standard deviation and the mean as the compression ratio increases, the standard deviation is close to one and the mean is close to zero, even at high compression ratios of 3,200:1, when Multilevel Clustering KLT + JPEG 2000 is used.

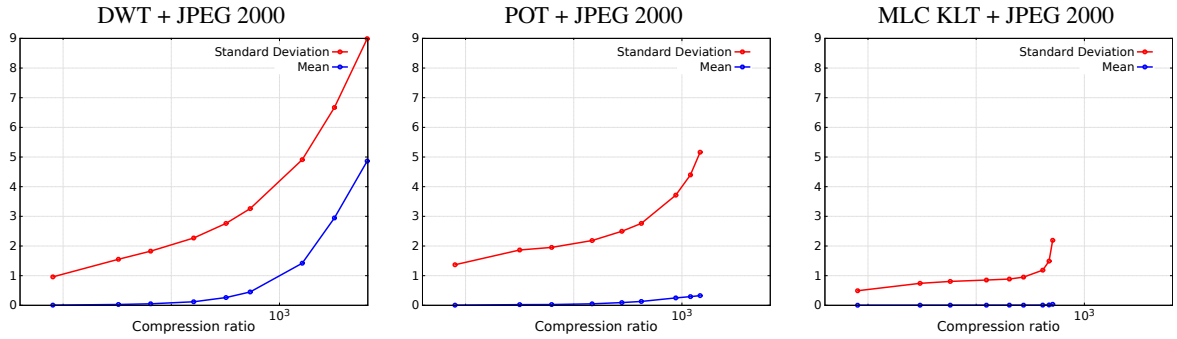


Figure A2. Average radiance residual statistics.

The difference between the spectral signature of the reconstructed radiances and the spectral signature of the original radiances (in $mW/m^2/sr/cm^{-1}$) is reported in FigureA3. Results are reported for a randomly selected pixel at coordinates ($y = 50$, $x = 10$) from a 3D volume representing an IASI orbit. Is is clear that the differences are particularly small at compression ratios lower than 533:1. When the compression ratio increases, the differences also increase. It is interesting to note that Multilevel Clustering KLT + JPEG 2000 produces smaller differences than DWT + JPEG 2000 and POT + JPEG 2000 at high compression ratios (above 533:1).

Figure A4 illustrates the average of the differences (absolute value) between the spectral signature of the reconstructed radiances and the spectral signature of the original radiances for the compression settings analyzed. It is clear that Multilevel Clustering KLT + JPEG 2000 produces the smallest errors.

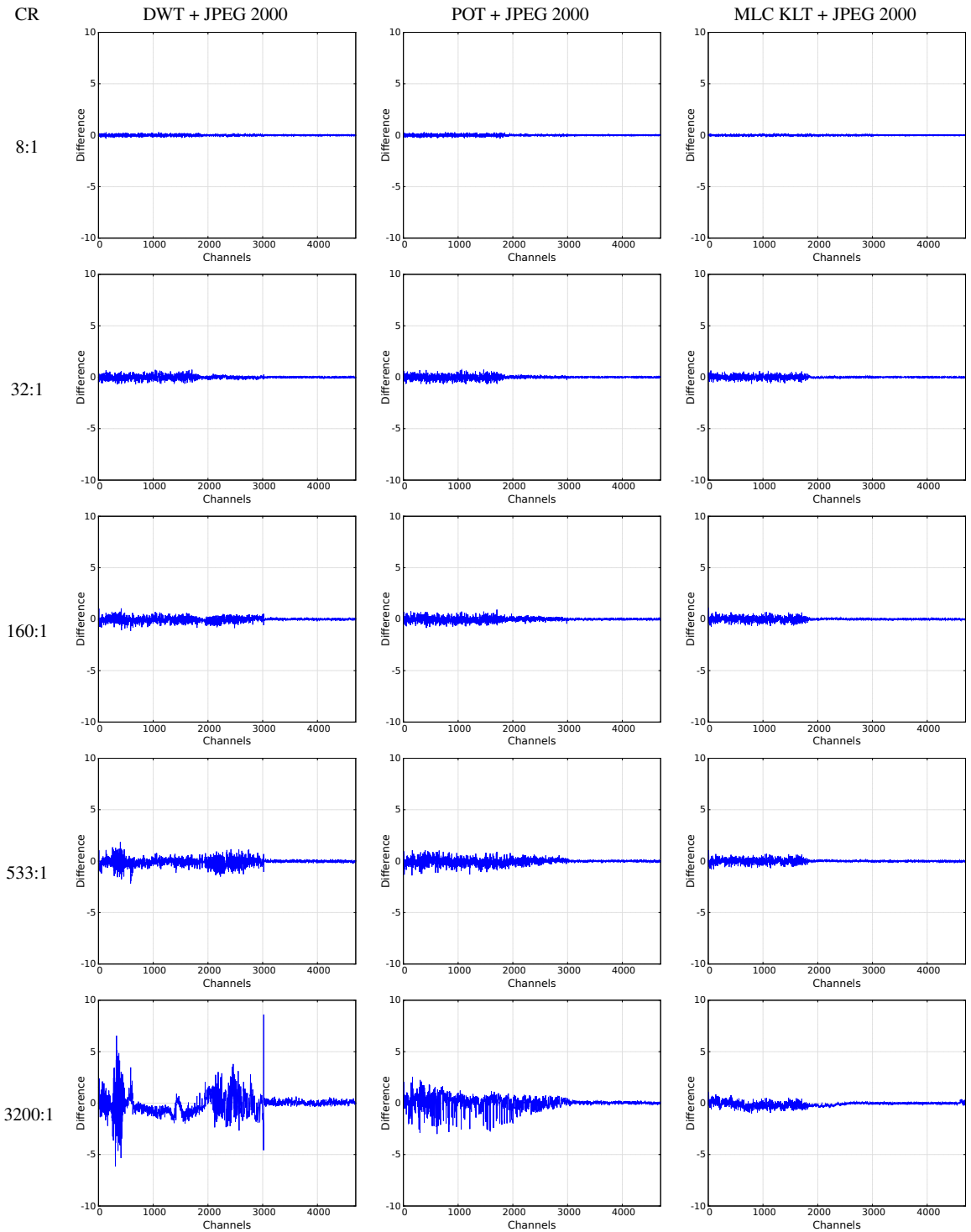


Figure A3. Difference between the spectral signature of the original data and the spectral signature of the reconstructed data as a function of channel number. Following the feature selection in [37], 4699 channels are used. Results for different compression ratios (CR) are reported.

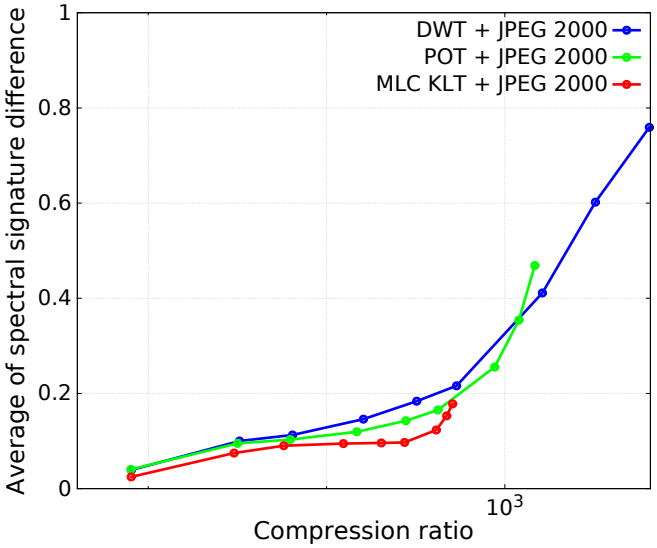


Figure A4. Average of the spectral signature difference.

Chapter 6

Results summary

Experimental results of IASI data compression are summarized in this chapter. First, we present the data collection used in the experiments. IASI L0 and IASI L1C products acquired with IASI-A and IASI-B instruments are used to conduct the tests. Then, we report the compression performance of lossless, near-lossless, and lossy compression. Several state-of-the-art compression techniques and spectral transforms are evaluated on different scenarios. Finally, we investigate the impact of near-lossless and lossy compression on statistical retrieval algorithms. The origin of the positive impact of compression in the performance of the statistical retrieval methods is also studied.

6.1 Data Collection

The experiments are conducted with a set of IASI L0 products reported in Table 6.1 and three sets of IASI L1C products reported in Tables 6.2, 6.3, and 6.4. All products have been provided by CNES and EUMETSAT. IASI L0 products are stored without a sign and have a bit-depth precision from 6 to 10 bits. Each volume has 8,359 components, 30 FORs, and a variable number of lines. While products from CNES contain between 741 and 764 lines, products from EUMETSAT contain between 55 and 98 lines. To evaluate the compression performance of several coding strategies and techniques on IASI L1C products, a set of 96 IASI L1C orbits (48 orbits per

instrument) is employed. Due to space constraints, averaged sizes and zero-order entropies per instrument are reported in Table 6.2. The IASI L1C products reported in Table 6.3 are employed to evaluate the impact of lossy compression on statistical retrieval algorithms. Table 6.4 illustrates the products used to investigate the origin of the benefits produced by compression when reconstructed spectra is used in statistical retrieval methods. All IASI L1C products have 16 bpppc and are stored as signed integers. Each volume contains 8,461 components, 30 FORs, and a variable number of lines between 630 and 788.

Table 6.1: IASI L0 data used in the experiments. Identifiers, which indicate the provider of each product, and technical names are provided.

Identifier	**IASI L0 Product**
EUMETSAT 1	IASI_HRP_00_M01_20130813184200Z_20130813185335Z_N_O_-20130813184205Z
EUMETSAT 2	IASI_HRP_00_M01_20130813202226Z_20130813203309Z_N_O_-20130813202229Z
EUMETSAT 3	IASI_HRP_00_M01_20130814070242Z_20130814071311Z_N_O_-20130814070245Z
EUMETSAT 4	IASI_HRP_00_M01_20130814084228Z_20130814085440Z_N_O_-20130814084231Z
EUMETSAT 5	IASI_HRP_00_M01_20130814101020Z_20130814102504Z_N_O_-20130814101026Z
EUMETSAT 6	IASI_HRP_00_M01_20130814120024Z_20130814121046Z_N_O_-20130814120030Z
EUMETSAT 7	IASI_HRP_00_M02_20130813192744Z_20130813193516Z_N_O_-20130813192748Z
EUMETSAT 8	IASI_HRP_00_M02_20130814092926Z_20130814093921Z_N_O_-20130814092930Z
CNES 1	IASI_xxx_00_M02_20091007112100Z_20091007130000Z_N_O_-20091007125723Z
CNES 2	IASI_xxx_00_M02_20091007130000Z_20091007143900Z_N_O_-20091007143543Z
CNES 3	IASI_xxx_00_M02_20091007143900Z_20091007162100Z_N_O_-20091007161509Z
CNES 4	IASI_xxx_00_M02_20091017125400Z_20091017143300Z_N_O_-20091017142943Z
CNES 5	IASI_xxx_00_M02_20100319050300Z_20100319064500Z_N_O_-20100319064052Z
CNES 6	IASI_xxx_00_M02_20120718075700Z_20120718093900Z_N_O_-20120718084400Z
CNES 7	IASI_xxx_00_M02_20130116133300Z_20130116151500Z_N_O_-20130116142103Z
CNES 8	IASI_xxx_00_M02_20130916080300Z_20130916094500Z_N_O_-20130916093859Z

Table 6.2: IASI L1C products used to investigate the compression performance of several coding strategies and techniques on IASI L1C data. Sizes and averaged zero-order entropies per instrument are provided (48 orbits per instrument). *M* is the number of spectral channels, *Ns* is the number of scan lines, *N-FORs* is the number of elementary fields of regard (FOR) per line, and *N-IFOVs* is the number of instantaneous fields of view (IFOV) per FOR.

Instrument	**Size (M × Ns × N-FORs × N-IFOVs)**	**Average Entropy**
IASI-A Products	8461 × (630-787) × 30 × 4	12.84
IASI-B Products	8461 × (742-788) × 30 × 4	12.83
Average	8461 × (761) × 30 × 4	12.83

Table 6.3: IASI L1C orbits used to investigate the impact of near-lossless and lossy compression on statistical retrieval algorithms. Technical names, identifiers and sizes are provided.

Technical name	**Identifier**	**Size ($x \times y \times z$)**
IASI_xxx_1C_M01_20130817004753Z_-20130817022952Z_N_O_20130817013849Z	IASI_20130817004753Z	60×1530×8461
IASI_xxx_1C_M01_20130817041457Z_-20130817055656Z_N_O_20130817050513Z	IASI_20130817041457Z	60×1530×8461
IASI_xxx_1C_M01_20130817055657Z_-20130817073856Z_N_O_20130817064707Z	IASI_20130817055657Z	60×1530×8461
IASI_xxx_1C_M01_20130817073857Z_-20130817092056Z_N_O_20130817082957Z	IASI_20130817073857Z	60×1530×8461

Table 6.4: IASI L1C products used to investigate the origin of the benefits produced by lossy compression on statistical retrieval algorithms. Identifiers and sizes are provided. *M* is the number of spectral channels, *Ns* is the number of scan lines, *N-FORs* is the number of elementary fields of regard (FOR) per line, and *N-IFOVs* is the number of instantaneous fields of view (IFOV) per FOR.

Orbit identifiers	**Size (M × Ns × N-FORs × N-IFOVs)**
Training set	
IASI_xxx_1C_M01_20131017012656Z_20131017030856Z_N_-O_20131017021701Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017045352Z_20131017063552Z_N_-O_20131017054413Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017081752Z_20131017095656Z_N_-O_20131017090901Z	(8461 × 743 × 30 × 4)
IASI_xxx_1C_M01_20131017113856Z_20131017131752Z_N_-O_20131017123024Z	(8461 × 742 × 30 × 4)
IASI_xxx_1C_M01_20131017145656Z_20131017163856Z_N_-O_20131017155030Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017181752Z_20131017195952Z_N_-O_20131017190946Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017214152Z_20131017232352Z_N_-O_20131017223236Z	(8461 × 765 × 30 × 4)
Test set	
IASI_xxx_1C_M01_20131017030856Z_20131017045352Z_N_-O_20131017035958Z	(8461 × 787 × 30 × 4)
IASI_xxx_1C_M01_20131017063552Z_20131017081752Z_N_-O_20131017072637Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017095656Z_20131017113856Z_N_-O_20131017104905Z	(8461 × 765 × 30 × 4)
IASI_xxx_1C_M01_20131017131752Z_20131017145656Z_N_-O_20131017141023Z	(8461 × 743 × 30 × 4)
IASI_xxx_1C_M01_20131017163856Z_20131017181752Z_N_-O_20131017173059Z	(8461 × 742 × 30 × 4)
IASI_xxx_1C_M01_20131017195952Z_20131017214152Z_N_-O_20131017205049Z	(8461 × 765 × 30 × 4)

6.2 Compression of IASI L0 Products

Tables 6.5 and 6.6 report the order-0 entropy and the order-1, order-2, and order-3 context-based entropies of IASI L0 data. Each column provides the average entropy for all samples in a product. One can see that, since the IASI on-board processing chain achieves an average bit-rate of 8.2 bpppc, at least 1.63 bpppc might be saved if a variable-length code were used. Results improve when contextual models are employed. Order-1 context-based entropy allows to gain at least 3.78 bpppc. Order-2 and order-3 context-based models further improve the results, allowing to spare at least 5.57 and 6.71 bpppc, respectively.

Tables 6.7 and 6.8 illustrate the lossless compression performance of several coding strategies and techniques on IASI L0 data. Results indicate that all the compression schemes employed produce similar results. On average, a compression ratio of 2.75:1 can be achieved using either CCSDS-123, M-CALIC, or POT + JPEG 2000. Slightly lower compression ratios are yielded by POT + JPEG-LS.

Table 6.5: Context-based entropy analysis of IASI L0 data from EUMETSAT. Results are reported in bpppc (lower is better).

IASI L0 Product	Order-0	Order-1 Left	Order-1 Top-Left	Order-1 Top	Order-1 Top-Right	Order-1 One Previous	Order-1 Next to Previous	Order-2 Left and Top	Order-2 Two Previous	Order-3 Left and Top and One Previous	Order-3 Left and Two Previous	Order-3 Top and Two Previous
EUMETSAT 1	6.092	3.709	3.851	3.609	3.822	3.176	3.337	1.836	2.042	0.737	0.819	0.826
EUMETSAT 2	5.848	3.671	3.829	3.642	3.848	3.100	3.256	1.827	2.083	0.779	0.883	0.875
EUMETSAT 3	6.469	3.710	3.793	3.625	3.834	3.165	3.319	1.391	1.747	0.495	0.568	0.570
EUMETSAT 4	6.364	3.831	3.976	3.688	3.897	3.188	3.357	1.750	2.025	0.687	0.770	0.782
EUMETSAT 5	5.363	3.843	3.927	3.759	3.966	2.875	2.998	2.032	2.232	0.964	1.100	1.110
EUMETSAT 6	5.784	3.935	4.062	3.796	3.976	2.918	3.051	1.826	2.073	0.753	0.872	0.892
EUMETSAT 7	5.550	3.271	3.461	3.289	3.458	3.034	3.176	1.902	2.033	0.861	0.910	0.933
EUMETSAT 8	6.270	3.652	3.836	3.655	3.766	3.091	3.248	1.655	1.929	0.640	0.716	0.725
Average	5.967	3.702	3.841	3.632	3.820	3.068	3.217	1.777	2.020	0.739	0.829	0.839

Table 6.6: Context-based entropy analysis of IASI L0 data from CNES. Results are reported in bpppc (lower is better).

IASI L0 Product	**Order-0**	**Order-1 Left**	**Order-1 Top-Left**	**Order-1 Top**	**Order-1 Top-Right**	**Order-1 One Previous**	**Order-1 Next To Previous**	**Order-2 Left and Top**	**Order-2 Two Previous**	**Order-3 Left and Top and One Previous**	**Order-3 Left and Two Previous**	**Order-3 Top and Two Previous**
CNES 1	6.297	4.051	4.211	3.986	4.218	3.213	3.395	2.598	2.627	1.304	1.462	1.487
CNES 2	6.303	4.217	4.382	4.154	4.386	3.169	3.340	2.570	2.590	1.227	1.409	1.430
CNES 3	6.354	4.144	4.297	4.101	4.322	3.188	3.360	2.577	2.592	1.241	1.416	1.434
CNES 4	6.335	4.153	4.297	4.104	4.332	3.210	3.387	2.557	2.607	1.243	1.415	1.435
CNES 5	6.282	4.064	4.219	4.014	4.230	3.172	3.350	2.548	2.590	1.292	1.458	1.477
CNES 6	6.573	4.217	4.357	4.164	4.407	3.255	3.442	2.539	2.603	1.193	1.364	1.386
CNES 7	6.315	4.057	4.196	3.987	4.220	3.206	3.384	2.600	2.609	1.290	1.455	1.478
CNES 8	6.376	4.256	4.408	4.196	4.420	3.215	3.396	2.574	2.616	1.225	1.403	1.425
Average	6.354	4.144	4.295	4.088	4.316	3.203	3.381	2.570	2.604	1.251	1.422	1.444

Table 6.7: Lossless compression performance of several coding strategies and techniques on IASI L0 products from EUMETSAT. Results are reported in compression ratio (higher is better) and bpppc (lower is better).

Compression ratios				
IASI L0 Product	**CCSDS-123**	**M-CALIC**	**POT + JPEG-LS**	**POT + JPEG2000**
EUMETSAT 1	2.617	2.693	2.145	2.653
EUMETSAT 2	2.627	2.701	2.116	2.650
EUMETSAT 3	2.483	2.537	1.902	2.489
EUMETSAT 4	2.581	2.637	2.140	2.617
EUMETSAT 5	2.805	2.862	2.308	2.801
EUMETSAT 6	2.761	2.783	2.176	2.706
EUMETSAT 7	2.618	2.746	1.964	2.611
EUMETSAT 8	2.588	2.623	2.010	2.567
Average	2.635	2.697	2.095	2.636
Bits per pixel per component				
IASI L0 Product	**CCSDS-123**	**M-CALIC**	**POT + JPEG-LS**	**POT + JPEG2000**
EUMETSAT 1	3.13	3.04	3.82	3.09
EUMETSAT 2	3.12	3.03	3.87	3.09
EUMETSAT 3	3.30	3.23	4.31	3.29
EUMETSAT 4	3.17	3.10	3.83	3.13
EUMETSAT 5	2.92	2.86	3.55	2.92
EUMETSAT 6	2.96	2.94	3.76	3.03
EUMETSAT 7	3.13	2.98	4.17	3.14
EUMETSAT 8	3.16	3.12	4.07	3.19
Average	3.11	3.03	3.92	3.11

Table 6.8: Lossless compression performance of several coding strategies and techniques on IASI L0 products from CNES. Results are reported in compression ratio (higher is better) and bpppc (lower is better).

Compression ratios				
IASI L0 Product	**CCSDS-123**	**M-CALIC**	**POT + JPEG-LS**	**POT + JPEG2000**
CNES 1	2.839	2.874	2.614	2.858
CNES 2	2.846	2.858	2.596	2.834
CNES 3	2.857	2.873	2.597	2.847
CNES 4	2.840	2.859	2.585	2.832
CNES 5	2.843	2.871	2.605	2.857
CNES 6	2.835	2.834	2.569	2.819
CNES 7	2.843	2.883	2.605	2.852
CNES 8	2.832	2.823	2.574	2.812
Average	2.841	2.859	2.593	2.838
Bits per pixel per component				
IASI L0 Product	**CCSDS-123**	**M-CALIC**	**POT + JPEG-LS**	**POT + JPEG2000**
CNES 1	2.88	2.85	3.13	2.86
CNES 2	2.88	2.86	3.15	2.89
CNES 3	2.87	2.85	3.15	2.88
CNES 4	2.88	2.86	3.17	2.89
CNES 5	2.88	2.85	3.14	2.87
CNES 6	2.89	2.89	3.19	2.90
CNES 7	2.88	2.84	3.14	2.87
CNES 8	2.89	2.90	3.18	2.91
Average	2.88	2.86	3.15	2.88

6.3 Compression of IASI L1C Products

Table 6.9 reports the average lossless compression performance of several coding strategies and techniques on IASI L1C data. Results are computed using six state-of-the-art compression standards and techniques, namely, JPEG-LS [22], JPEG 2000 [23], M-CALIC [24], CCSDS-122.0 [26], CCSDS-123.0 [21], and HEVC [27]. To exploit the high spectral redundancy present in IASI data, four widely accepted spectral transforms are paired along with the compression methods, namely, the Multilevel Clustering Reversible Karhunen-Loève Transform (Multilevel Clustering RKLT) [28], the Integer Wavelet Transform (IWT) [29], the Reversible Pairwise Orthogonal Transform (RPOT) [25], and the Regression Wavelet Analysis Transform (RWA) [30].

One can see that exploiting the high spectral redundancy present in IASI L1C data is of paramount importance to produce competitive compression results. Compression ratios over 2.5:1 can be achieved for lossless compression of IASI L1C products. The best performance is achieved by Multilevel Clustering RKLT + M-CALIC, which yields, on average, a compression ratio of 2.55:1.

Experimental results for near-lossless compression are reported in Table 6.10. As expected, when the Peak Absolute Error (PAE) increases, higher compression ratios are produced. It is interesting to note that competitive coding performance can be achieved by allowing small errors. PAEs equal to 1 and 3 allow to improve the coding results by 17% and 30%, respectively, compared to lossless compression. Figure 6.1 reports the rate-distortion performance of near-lossless compression as a function of the PAE and the SNR Energy. Plots indicate that small PAEs as 1 and 3 allow to achieve SNR Energy over 65 dB.

Figure 6.2 illustrates the lossy compression performance of several coding strategies and techniques on IASI L1C products. Lossy compression is evaluated employing the JPEG 2000 and the CCSDS-122.0 standards. Three spectral transforms are used to exploit the spectral dimension of the data, namely, the Multilevel Clustering Karhunen-Loève Transform (Multilevel Clustering KLT), the Discrete Wavelet Transform (DWT), and the Pairwise Orthogonal Transform (POT). All schemes are assessed using nine target bit-rates distributed between 0.01 and 2 bpppc.

Results reveal that, as happened for lossless compression, exploiting the redundancy present in the spectral dimension is of utmost importance to achieve competitive coding performance. The best results are produced by Multilevel Clustering KLT + JPEG 2000.

Table 6.9: Lossless compression of IASI L1C products. Results are reported in compression ratio (higher is better). Percent savings (higher is better) with respect to original technique are provided within brackets.

IASI-A—Lossless Compression Ratio & Percent Savings

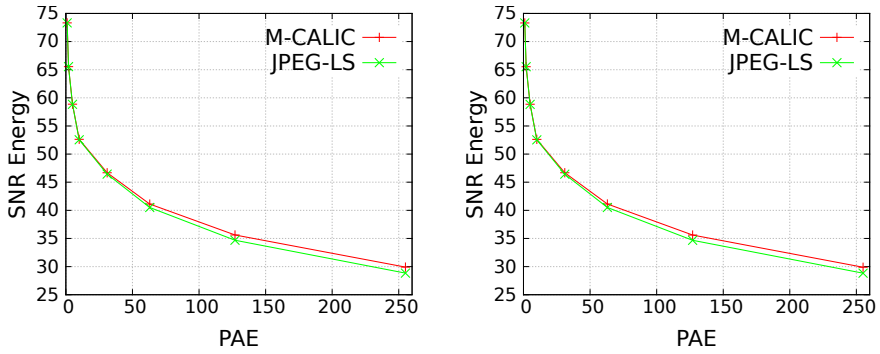
Tech. \ Tra.	No Transform	IWT	RPOT	RWA	Multilevel Clustering RKLT
JPEG-LS	1.78:1	2.26:1 (21.24%)	2.26:1 (21.24%)	2.44:1 (27.05%)	2.46:1 (27.64%)
JPEG 2000	1.73:1	2.24:1 (22.77%)	2.24:1 (22.77%)	2.43:1 (28.81%)	2.47:1 (29.96%)
M-CALIC	2.32:1	2.32:1 (0.00%)	2.34:1 (0.85%)	2.48:1 (6.45%)	2.54:1 (8.66%)
CCSDS-122.0	1.68:1	2.13:1 (21.13%)	2.13:1 (21.13%)	2.29:1 (26.64%)	2.33:1 (27.90%)
CCSDS-123.0	2.42:1	2.42:1 (0.00%)	2.39:1 (−1.24%)	2.46:1 (1.63%)	2.47:1 (2.02%)
HEVC	2.23:1	2.29:1 (2.62%)	2.28:1 (2.19%)	2.45:1 (8.98)	2.50:1 (10.80%)

IASI-B—Lossless Compression Ratio & Percent Savings

Tech. \ Tra.	No Transform	IWT	RPOT	RWA	Multilevel Clustering RKLT
JPEG-LS	1.79:1	2.28:1 (21.49%)	2.27:1 (21.15%)	2.45:1 (26.94%)	2.48:1 (27.82%)
JPEG 2000	1.74:1	2.25:1 (22.67%)	2.25:1 (22.67%)	2.44:1 (28.69%)	2.49:1 (30.12%)
M-CALIC	2.34:1	2.33:1 (−0.43%)	2.35:1 (0.43%)	2.50:1 (6.40%)	2.56:1 (8.59%)
CCSDS-122.0	1.69:1	2.14:1 (21.03%)	2.14:1 (21.03%)	2.30:1 (26.52%)	2.34:1 (27.78%)
CCSDS-123.0	2.44:1	2.44:1 (0.00%)	2.40:1 (−1.64%)	2.48:1 (1.61%)	2.48:1 (1.61%)
HEVC	2.24:1	2.30:1 (2.61%)	2.29:1 (2.18%)	2.47:1 (9.31%)	2.52:1 (11.11%)

Table 6.10: Near-lossless compression of IASI L1C products. Results are reported in compression ratio (higher is better). Results for lossless compression (PAE = 0) are included. Percent savings (higher is better) with respect to lossless compression are provided within brackets.

	IASI-A		**IASI-B**	
PAE	**JPEG-LS**	**M-CALIC**	**JPEG-LS**	**M-CALIC**
0	1.78	2.32	1.79	2.34
1	2.17 (17.97%)	3.02 (23.18%)	2.18 (17.89%)	3.05 (23.28%)
3	2.60 (31.54%)	3.90 (40.51%)	2.61 (31.42%)	3.95 (40.76%)
7	3.15 (43.49%)	5.21 (55.47%)	3.18 (43.71%)	5.28 (55.68%)
15	3.93 (54.71%)	7.34 (68.39%)	3.98 (55.03%)	7.48 (68.72%)
31	5.11 (65.17%)	11.11 (79.18%)	5.18 (65.44%)	11.35 (79.38%)
63	6.99 (74.54%)	18.39 (87.38%)	7.08 (74.72%)	18.82 (87.57%)
127	10.00 (82.20%)	33.33 (93.03%)	10.19 (82.43%)	34.04 (93.13%)
255	15.09 (88.20%)	61.54 (96.23%)	15.38 (88.36%)	64.00 (96.34%)



(**a**) (**b**)

Figure 6.1: Rate-distortion performance of near-lossless compression of IASI L1C products. Results report SNR Energy (in dB, higher is better) vs. PAE. (**a**) IASI-A; (**b**) IASI-B.

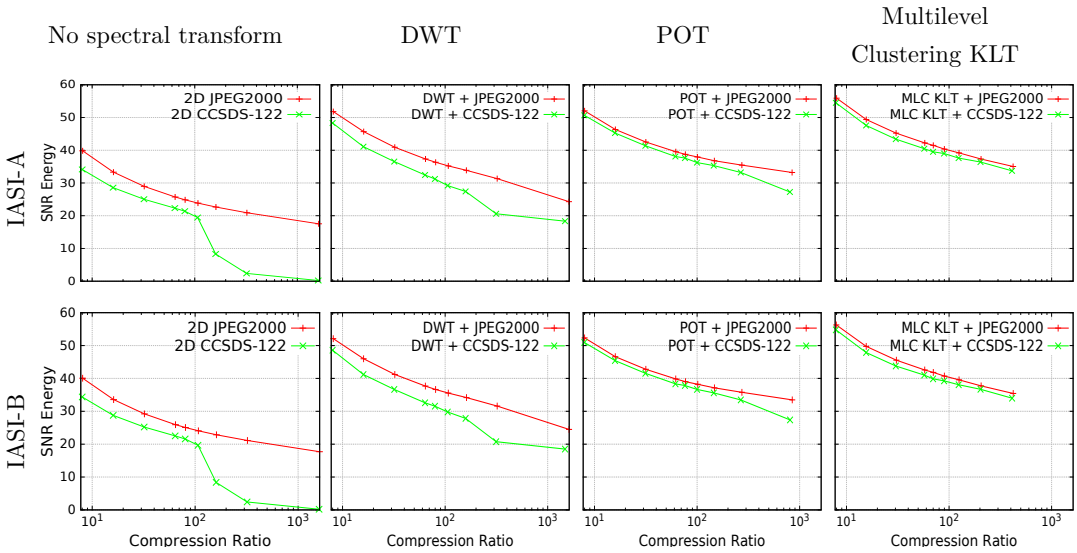


Figure 6.2: Rate-distortion performance of lossy compression of IASI L1C products. Results report SNR Energy (in dB, higher is better) vs. compression ratio. Results for different spectral transforms are plotted in the columns. In each plot, curves for JPEG 2000 and CCSDS-122.0 performance are displayed. Ranges are the same in all the plots to ease the comparison. Top row: IASI-A products; Bottom row: IASI-B products. POT and Multilevel Clustering KLT are not able to reach such high compression ratios (over 1,000:1) as DWT because side-information needs to be transmitted besides the compressed data.

Figure 6.3 reports a comparison between lossy and near-lossless compression. The coding schemes that yield the best performance for lossy compression (Multilevel Clustering KLT + JPEG 2000) and near-lossless compression (M-CALIC) are compared from the perspective of PAE and SNR Energy.

Two main conclusions can be drawn from this comparison. On the one hand, plots reveal that near-lossless compression produces smaller errors in the reconstructed data than lossy compression. On the other hand, lossy compression achieves better performance than near-lossless compression in terms of SNR Energy.

Figures 6.4, 6.5, and 6.6 analyze the quality of the reconstructed spectra when some distortion is introduced in the data, i.e., when near-lossless o lossy compression schemes are used. Reconstructed radiances previously compressed with the coding strategies that achieve the best compression performance for near-lossless (M-CALIC) and lossy (Multilevel Clustering KLT + JPEG 2000) compression are compared with

reconstructed spectra from PCC, which is the common strategy used to compress IASI L1C data. Two target compression ratios are compared. Configuration for each compression scheme is summarized in Table 6.11.

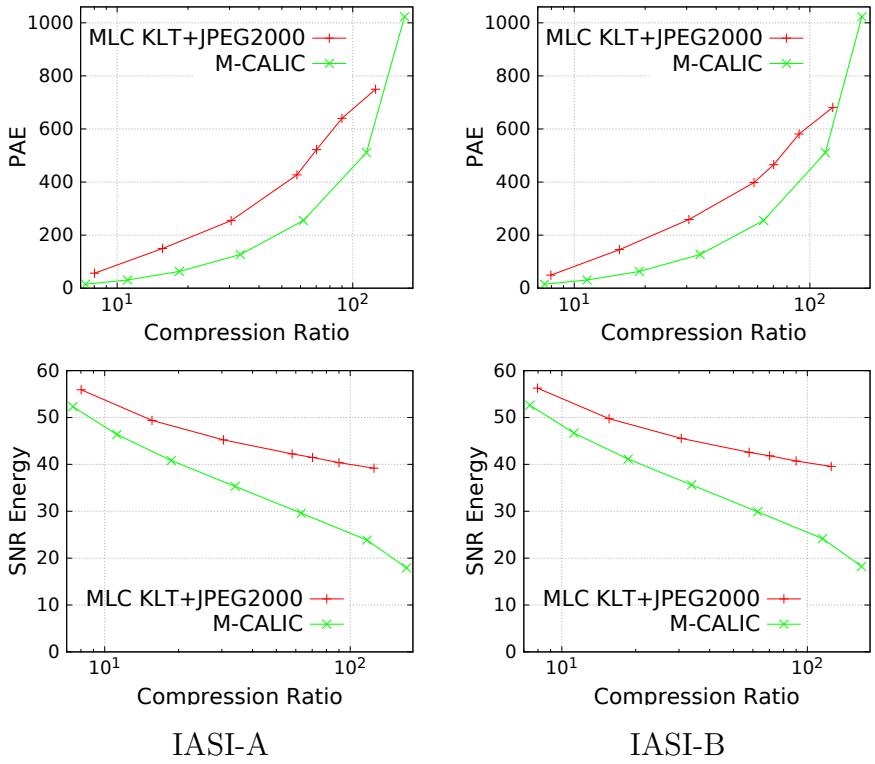


Figure 6.3: Performance comparison between near-lossless (M-CALIC) and lossy compression (Multilevel Clustering KLT + JPEG 2000). Top row: PAE (lower is better); Bottom row: SNR Energy (in dB, higher is better).

The maximum and minimum values, the standard deviation and the average of the normalized radiance residuals are compared in Figure 6.4. Note that only the first 1,800 components of the IASI spectrum are compared to simplify the comparison. Plots illustrate that the standard deviation and the average are similar for all compression schemes analyzed. However, Multilevel Clustering KLT + JPEG 2000 and PCC yield smaller maximum and minimum values than M-CALIC.

Table 6.11: Compression setting for PCC, M-CALIC, and Multilevel Clustering KLT + JPEG 2000 comparison.

		PCC	**M-CALIC**	**Multilevel Clustering KLT + JPEG 2000**
	Compression ratio	PC scores	PAE	Target bit-rate
Experiment 1	9:1	200	19	1.78
Experiment 2	12:1	150	29	1.33

Figure 6.5 reports the covariance matrix of the original and the reconstructed radiances for the compression strategies analyzed. One can see that both the original and the reconstructed data have a similar nature in all cases.

The differences between the covariance matrix of the original data and the covariance matrix of the reconstructed data are reported in Figure 6.6. The differences produced by PCC are more apparent than the differences produced by Multilevel Clustering KLT + JPEG 2000 and M-CALIC, which suggest that the reconstructed radiances from Multilevel Clustering KLT + JPEG 2000 and M-CALIC can retain at least the same quality than the reconstructed radiances from PCC.

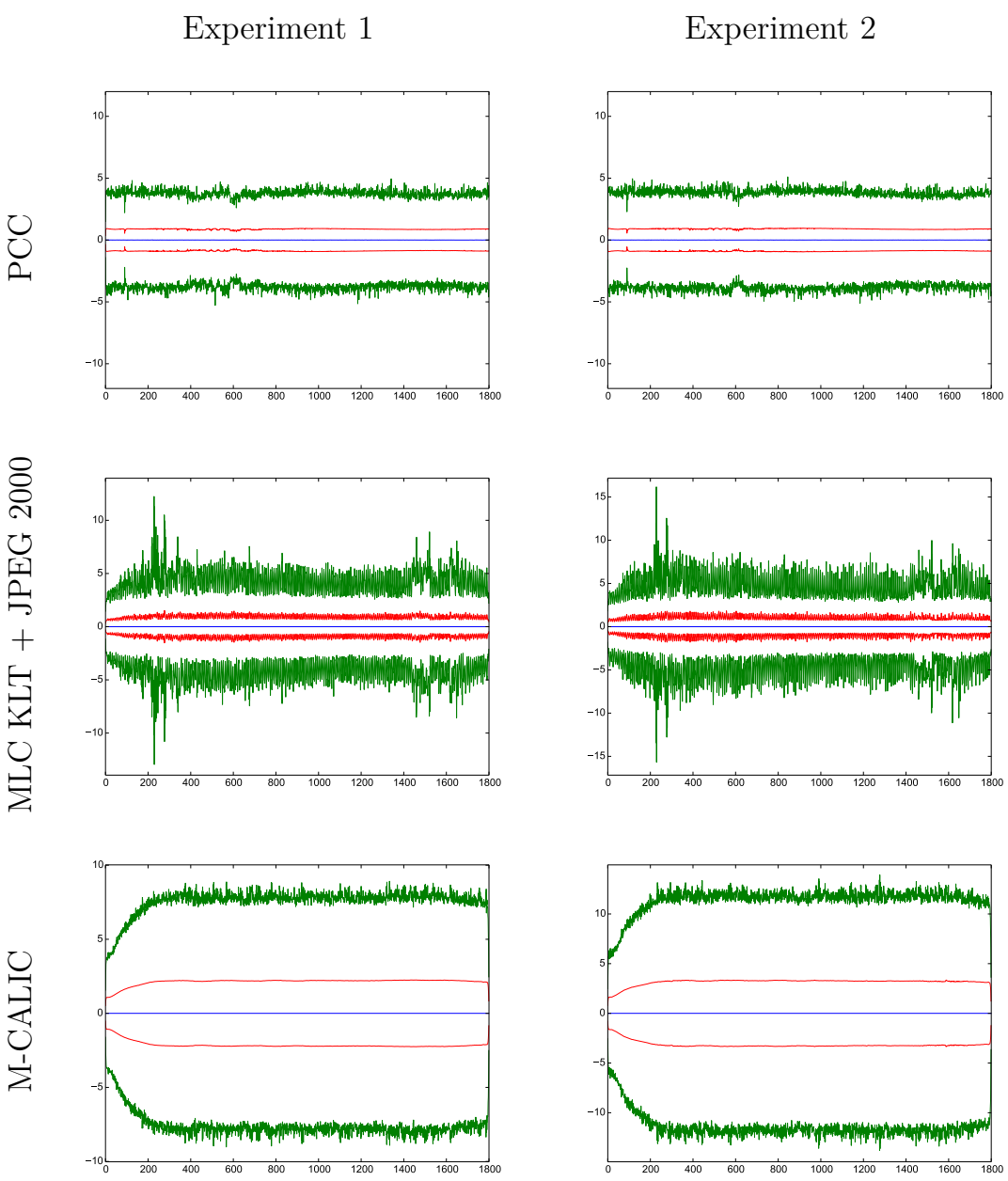


Figure 6.4: Normalized radiance residuals statistics. The average of the normalized radiance residuals is shown in blue, standard deviation in red, and maximum and minimum values in green.

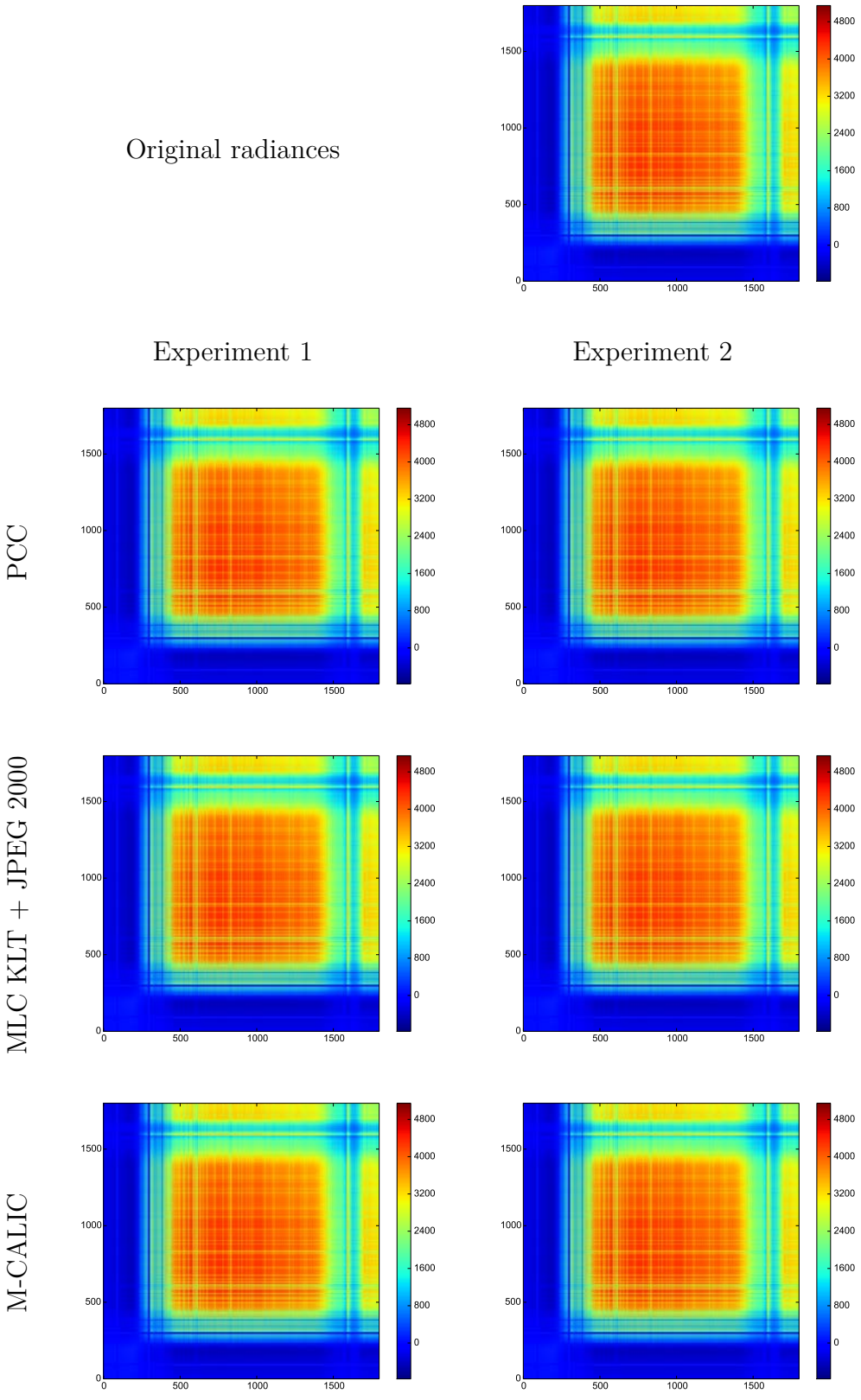


Figure 6.5: Covariance matrix of the original radiances and covariance matrix of the reconstructed radiances.

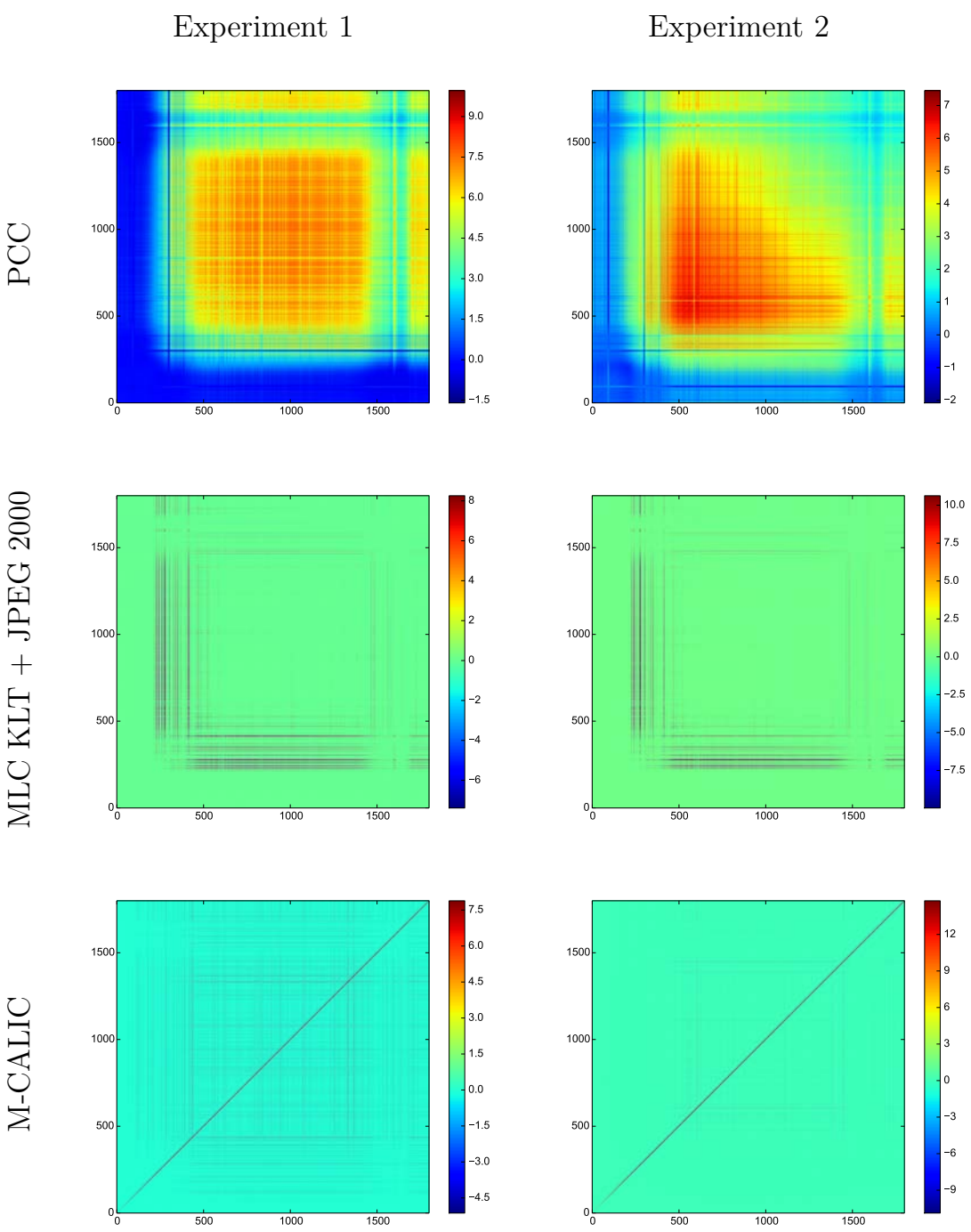


Figure 6.6: Differences between the covariance matrix of the original radiances and the covariance matrix of the reconstructed radiances.

6.4 Impact of Near-lossless and Lossy Compression on Statistical Retrieval Algorithms

Figures 6.7, 6.8, 6.9, and 6.10 illustrate the performance of statistical retrieval algorithms when reconstructed IASI L1C spectra is employed to retrieve physical information. Specifically, the retrieval of dew point temperature profiles is reported.

The average RMSE across the whole atmospheric column when reconstructed spectra from near-lossless compression (M-CALIC) is used in the retrieval stage is reported in Fig. 6.7. It is clear that low compression ratios achieve the same retrieval performance than the original data (uncompressed). As the compression ratio increases the retrieval performance increases. However, if the compression ratio is high (approximately 100:1) results start to deteriorate. Figure 6.8 illustrates the RMSE results for different pressure levels achieved for reconstructed data from near-lossless compression. The compression ratio with the best average RMSE is reported, namely, 129:1. The results are consistent with the average RMSE reported in Fig. 6.7. One can see that compression improves the retrieval performance throughout the whole atmospheric column.

As Fig. 6.7 and Fig. 6.8 in the case of near-lossless compression, Fig. 6.9 and Fig. 6.10 illustrate the average RMSE across the whole atmospheric column and the RMSE results for the whole range of pressure levels, respectively, for lossy compression schemes (spectral transform + JPEG 2000). Conclusions are similar to those drawn for near-lossless compression. However, the retrieval improvements are more significant in the case of lossy compression. The best performance is produced when both spectral and spatial transforms are used. Moreover, higher compression ratios compared to near-lossless compression allow to achieve the best retrieval results. The benefits of compression are present in all compression schemes and retrieval algorithms analyzed. The improvements are clear throughout the whole atmospheric column.

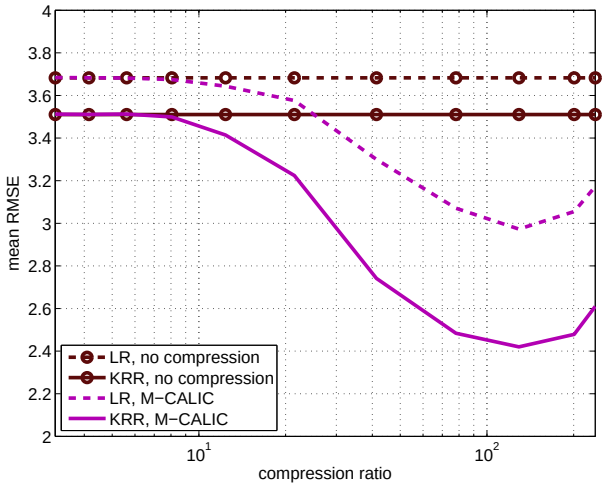


Figure 6.7: Estimation using LR or KRR for M-CALIC compression. The horizontal axis represents the compression ratio, and the vertical axis represents the mean RMSE (dew point temperature in K) over the different pressure levels. The plot shows the results for M-CALIC compression when using LR (dashed line) and KRR (solid line) for predicting moisture as compared to retrieval results on original data.

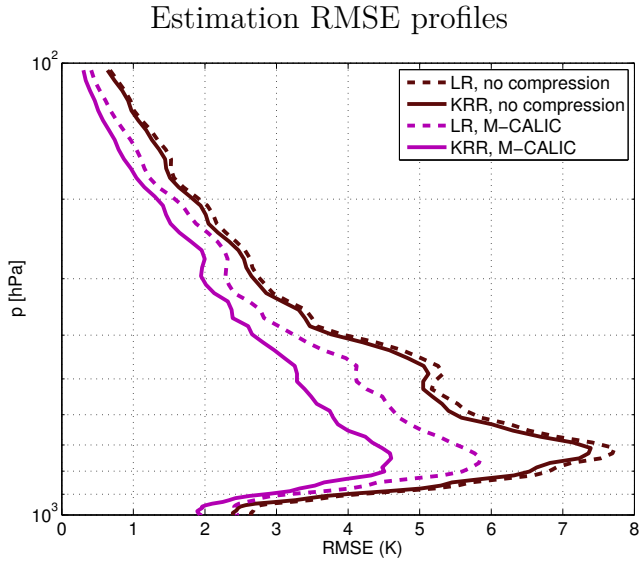


Figure 6.8: Moisture RMSE (dew point temperature in K) profiles for near-lossless compression. We chose the PAE with the best average RMSE. For LR (dashed line) and KRR (solid line), the minimum averaged RMSE is obtained at a compression ratio of 129:1. Results when using original data are shown as well for comparison purposes.

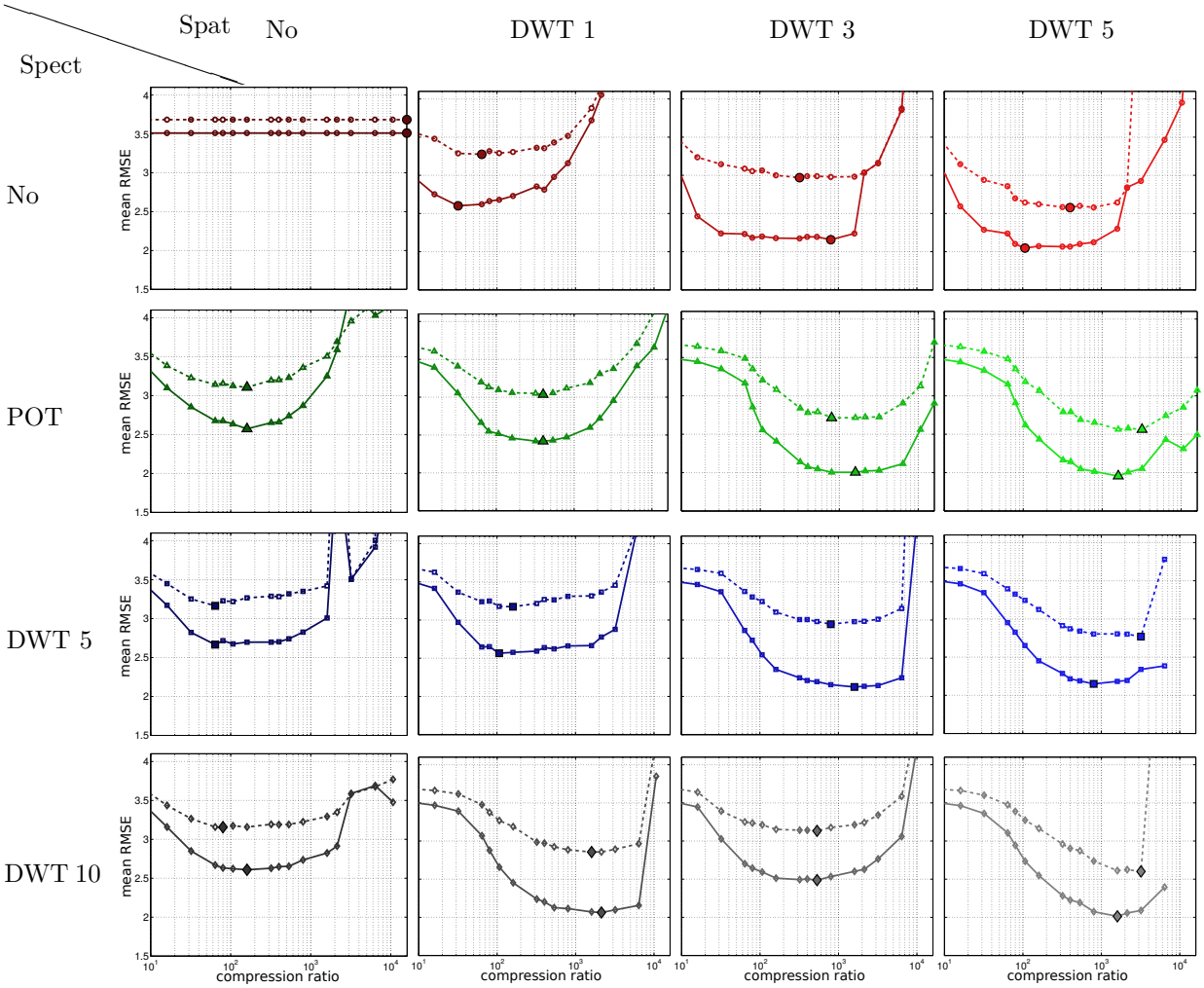


Figure 6.9: Estimation using LR or KRR for different spatial and spectral configurations for lossy compression. In all the plots, the horizontal axis represents the compression ratio, and the vertical axis represents the averaged RMSE (dew point temperature in K) over the different pressure levels. Ranges are the same in all the plots to ease the comparison. Each plot shows the results for a particular compression configuration when using LR (dashed lines) and KRR (solid lines) for predicting moisture. Configurations using the same number of levels for the spatial wavelet transform are in the same column and configurations using the same spectral transform are in the same row (same color, multiple color shades). 'No' refers to no transform (neither spatial nor spectral). The minimum averaged RMSE in each curve is indicated with a larger marker point.

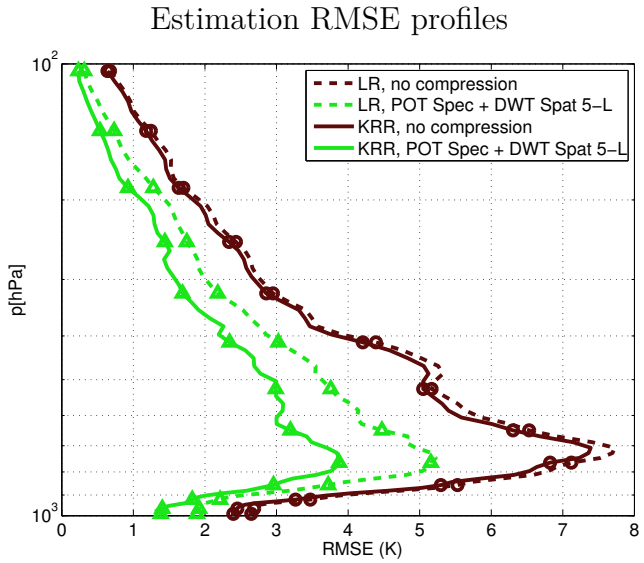


Figure 6.10: Moisture RMSE (dew point temperature in K) profiles for lossy compression. The selected configuration for multi-component JPEG 2000 lossy compression is POT as spectral transform and 5 levels of spatial DWT. We chose the compression ratio with the best average RMSE. For LR, the minimum averaged RMSE is obtained at a compression ratio of 3,200:1. For KRR, the minimum averaged RMSE is obtained at a compression ratio of 1,600:1. Results when using original data are shown as well for comparison purposes.

6.5 Origin of the Benefits Produced by Compression

Figure 6.11 illustrates the performance of atmospheric parameter retrieval methods when reconstructed spectra from lossy compression are used in a realistic scenario, i.e., samples from different orbits are employed to define the training, validation and testing subsets used by the statistical retrieval algorithms. Plots show the average RMSE of the predictions over different pressure levels. Results are reported for four different scenarios: land and cloud free, land and cloudy, ocean and cloud free, and ocean and cloudy conditions.

It is clear that statistical atmospheric parameter retrieval benefits from lossy compression. Low compression ratios keep the results almost unchanged compared to uncompressed data. However, the retrieval performance is improved when the compression ratio is increased. As expected, when the compression ratio is extremely high (above 300:1), the retrieval results start to deteriorate because too much information have been removed. It is interesting to note that while Multilevel Clustering KLT + JPEG 2000 produces the best prediction results, POT + JPEG 2000 enables competitive retrieval performance at higher compression ratios.

Figure 6.12 plots the RMSE results for different pressure levels. For each compression scheme, the compression ratio that achieves the best prediction performance is reported. One can see that the reconstructed spectra improve (or at least equal) the results achieved for the uncompressed data in all pressure levels. Improvements are more apparent in the mid-low troposphere.

Experiments suggest that the origin of the benefits produced by lossy compression relies on two consequences of the coding stage. On the one hand, lossy compression removes certain amount of noise from the original data, which benefits statistical retrieval algorithms. On the other hand, compression is an indirect way to exploit spectral and spatial relations between neighboring samples, which helps pixelwise statistical methods.

Figure 6.13 analyzes the noise level remaining in the reconstructed data when different compression schemes and compression ratios are used. Plots indicate that

the amount of noise in the reconstructed data decreases as the compression ratio increases. These results are consistent with the retrieval performance reported in Fig. 6.11. Moderate to high compression ratios (i.e., [10:1-300:1]) remove most of the noise present in the data, which benefits the retrieval results. Of course, if the compression ratio is extremely high (i.e., [300:1-2,000:1]) compression does not have any effect on the noise removal beyond a certain point and only useful information is removed.

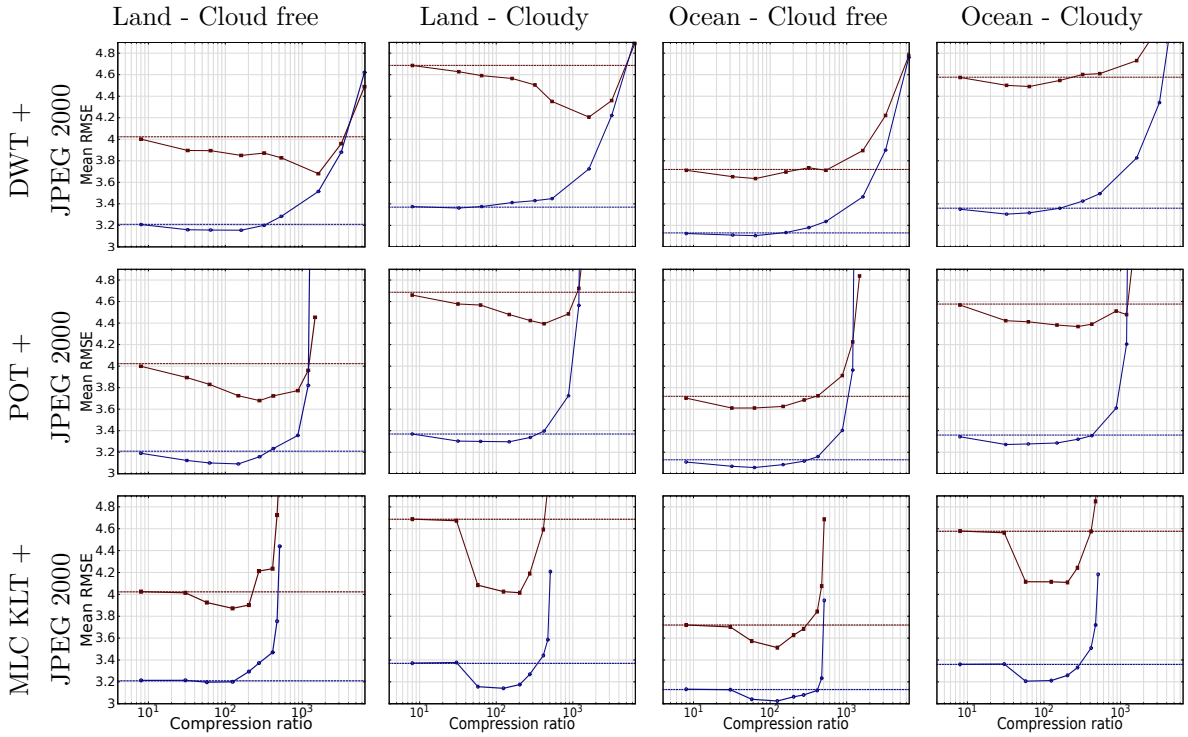


Figure 6.11: Dew point temperature (in kelvin) retrieval performance for different lossy compression settings using LR (solid red lines) and KRR (solid blue lines). In all the plots, the vertical axis represents the averaged RMSE over the different pressure levels between 1,100 and 100 hPa and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Each row shows the results for a particular compression setting and each column shows the results for a particular scenario. Results using uncompressed spectra (original data) for LR (dashed red lines) and KRR (dashed blue lines) are plotted for comparison purposes. The compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data.

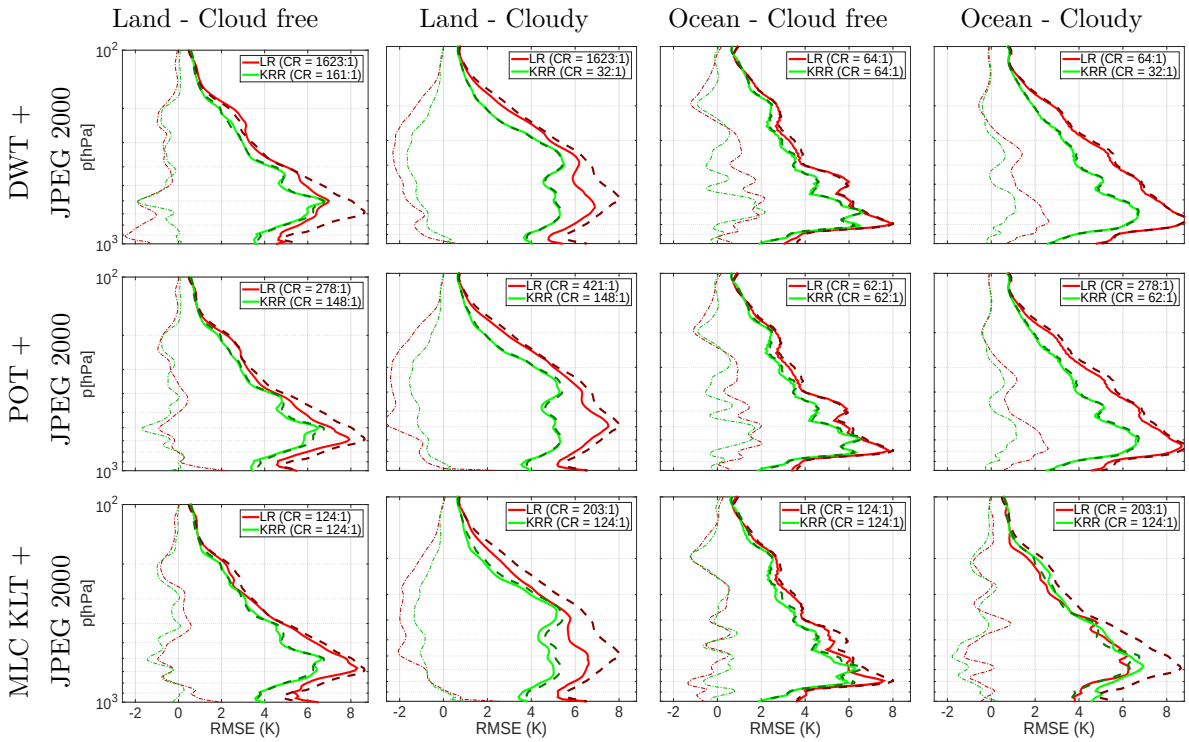


Figure 6.12: Dew point temperature (in kelvin) RMSE profiles and bias. The compression ratio (CR) with the best average RMSE is reported for LR (thick, solid and red lines) and KRR (thick, solid and green lines). Results using the original data are shown as well for comparison purposes when LR (thick, dashed and red lines) and KRR (thick, dashed and green lines) are used for the predictions. The bias of the reconstructed data are plotted with thin and dash-dot lines for LR (red) and KRR (green).

Table 6.12, Fig. 6.14, and Fig. 6.15 illustrate the impact of the spectral and spatial regularization produced by the compression stage on the retrieval performance.

First, different levels of spectral and spatial feature relations are exploited in the original data between neighboring samples. The objective is to assess the impact of spectral and spatial regularization on the retrieval performance. A simple strategy based on a data convolution with two Gaussian filters is applied to the spectral (σ_{sc}) and the spatial (σ_{st}) domain. Table 6.12 reports the retrieval performance for the different combinations analyzed. It is clear that exploiting spectral and spatial features improves the retrieval results. Nonetheless, more significant improvements are achieved by spatial regularization.

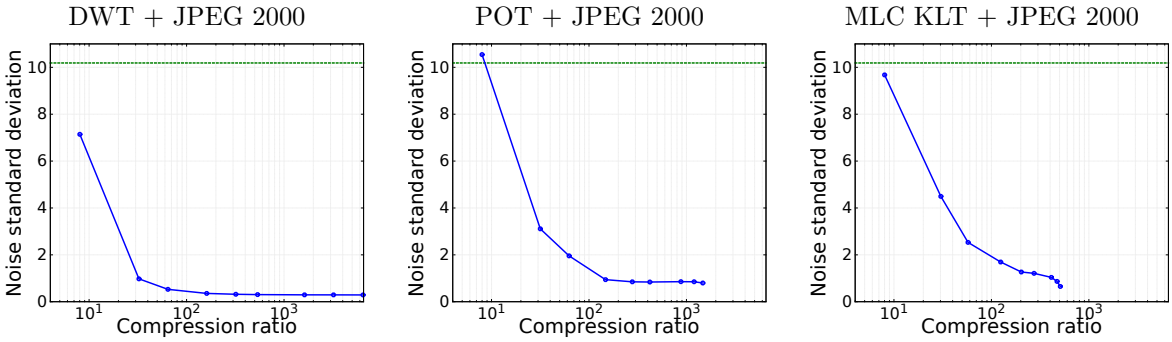


Figure 6.13: For each compression setting, the Noise Standard Deviation is plotted (solid and blue lines). In all the plots, the vertical axis represents the noise level (noise standard deviation of IASI raw data) in the reconstructed spectra and the horizontal axis represents the compression ratio. Noise estimation for uncompressed data (dashed and green lines) is reported as well for comparison purposes. The compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data.

Table 6.12: RMSE of dew point temperature predictions (in kelvin) for different combinations of σ_{st} and σ_{sc}. The improvement over the original data is reported in percentage in brackets. Best (green) and worst (red) performance are reported.

LR

		$\sigma_{sc}(spectral)$				
		0.01	0.25	1	5	10
$\sigma_{st}(spatial)$	0.01	4.515	4.515 (0%)	4.540 (-0.6%)	4.370 (3.2%)	4.362 (3.4%)
	0.25	4.343 (3.8%)	4.343 (3.8%)	4.400 (2.6%)	4.254 (5.8%)	4.233 (6.2%)
	1	4.061 (10.1%)	4.061 (10.1%)	4.097 (9.3%)	4.036 (10.6%)	3.996 (11.5%)
	5	3.834 (15.1%)	3.835 (15.1%)	3.880 (14.1%)	3.809 (15.6%)	3.771 (16.5%)
	10	3.742 (17.1%)	3.740 (17.2%)	3.795 (15.9%)	3.724 (17.5%)	3.711 (17.8%)
	20	3.684 (18.4%)	3.683 (18.4%)	3.737 (17.2%)	3.697 (18.1%)	3.681 (18.4%)
	30	3.669 (18.7%)	3.668 (18.8%)	3.736 (17.3%)	3.711 (17.8%)	3.694 (18.2%)

KRR

		$\sigma_{sc}(spectral)$				
		0.01	0.25	1	5	10
$\sigma_{st}(spatial)$	0.01	3.345	3.345 (0%)	3.376 (-0.9%)	3.389 (-1.3%)	3.379 (-1.0%)
	0.25	3.248 (2.9%)	3.248 (2.9%)	3.276 (2.1%)	3.305 (1.2%)	3.266 (2.4%)
	1	3.082 (7.9%)	3.082 (7.9%)	3.109 (7.1%)	3.113 (6.9%)	3.109 (7.1%)
	5	2.994 (10.5%)	2.995 (10.5%)	3.017 (9.8%)	3.028 (9.5%)	3.021 (9.7%)
	10	3.007 (10.1%)	3.007 (10.1%)	3.025 (9.6%)	3.035 (9.3%)	3.031 (9.4%)
	20	3.069 (8.3%)	3.070 (8.2%)	3.108 (7.1%)	3.114 (6.9%)	3.095 (7.5%)
	30	3.138 (6.2%)	3.138 (6.2%)	3.139 (6.2%)	3.160 (5.5%)	3.152 (5.8%)

Figures 6.14 and 6.15 report the retrieval performance when reconstructed spectra from compression and reconstructed spectra from compression paired along with the specific exploitation of spectral and spatial feature relations are used. Plots illustrate that the specific exploitation of spectral and spatial information between neighbouring samples before compression (blue curve) produces the same retrieval results than uncompressed filtered spectra (orange line). This observation suggests that the transforms applied in the compression stage are behind the spectral/spatial regularization, which ultimately benefits the retrieval performance.

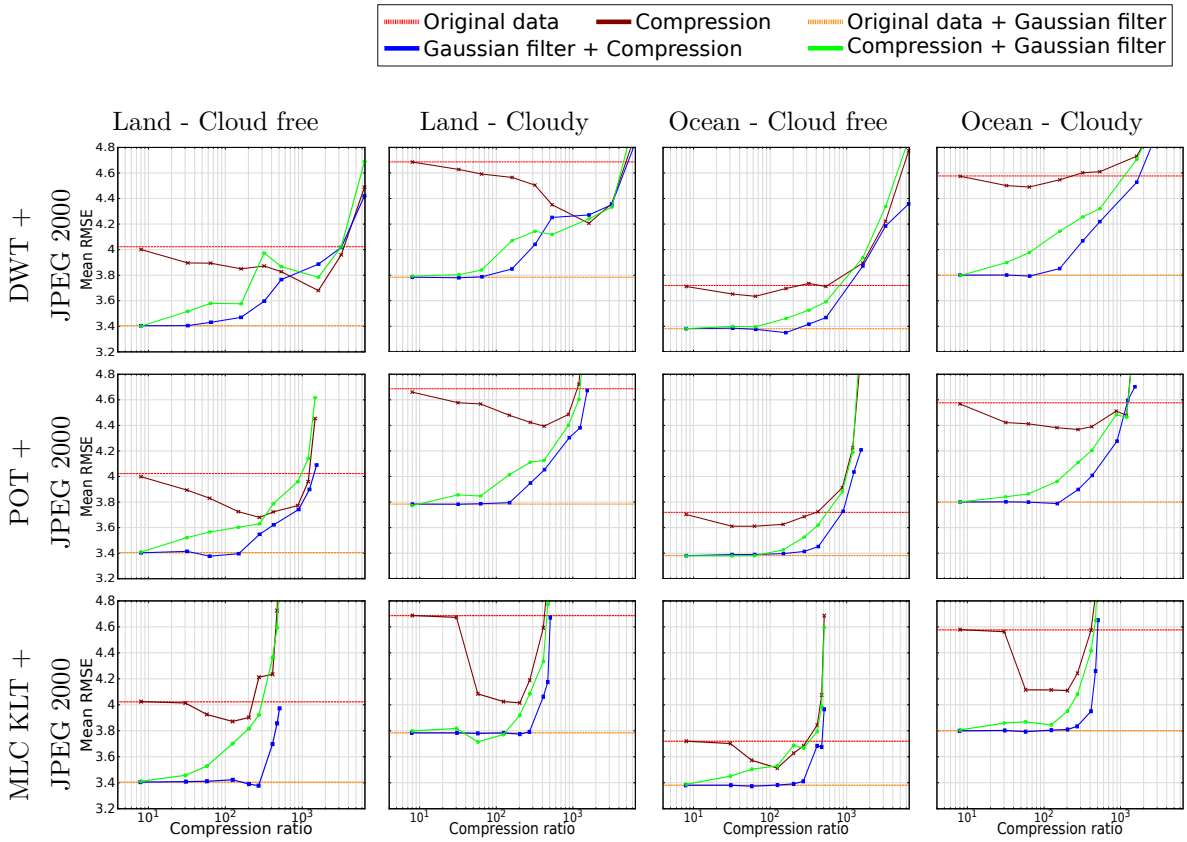


Figure 6.14: Dew point temperature (in kelvin) retrieval performance for different lossy compression settings using LR. In all the plots, the vertical axis represents the averaged RMSE over the different pressure levels and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Each row shows the results for a particular compression setting and each column shows the results for a particular scenario. Each plot compares 5 different approaches, i.e., original data, compression of the original data, original data + Gaussian filter, Gaussian filter + compression, and compression + Gaussian filter. The compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data.

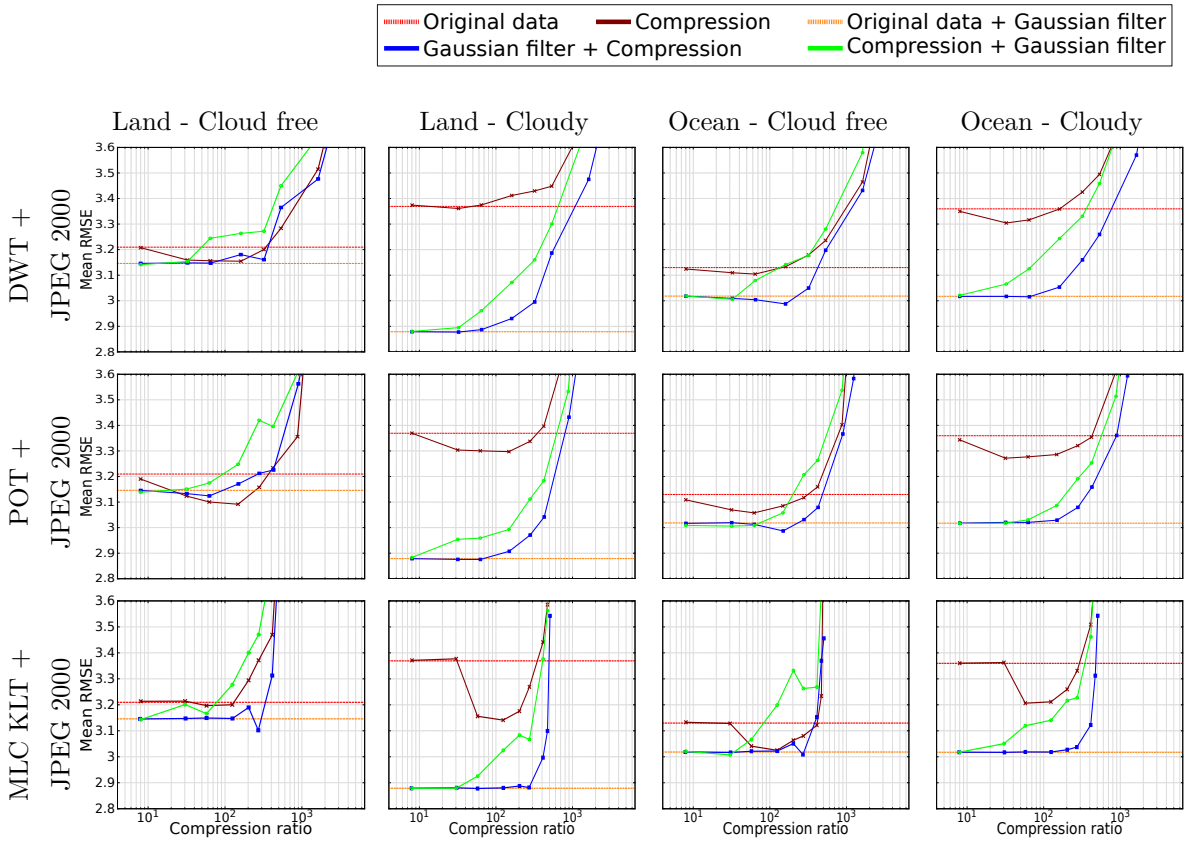


Figure 6.15: Dew point temperature (in kelvin) retrieval performance for different lossy compression settings using KRR. In all the plots, the vertical axis represents the averaged RMSE over the different pressure levels and the horizontal axis represents the compression ratio. Ranges are the same in all the plots to ease the comparison. Each row shows the results for a particular compression setting and each column shows the results for a particular scenario. Each plot compares 5 different approaches, i.e., original data, compression of the original data, original data + Gaussian filter, Gaussian filter + compression, and compression + Gaussian filter. The compression settings POT + JPEG 2000 and Multilevel Clustering KLT + JPEG 2000 achieve lower maximum compression ratios compared to DWT + JPEG 2000 because side-information needs to be transmitted in addition to the compressed data.

Chapter 7

Conclusions

7.1 Summary

The Infrared Atmospheric Sounding Interferometer (IASI) represents a significant progress in the accuracy and quality of the measurements injected into meteorological models. The particular characteristics of the instrument lead to generate a large amount of data every day. Therefore, efficient strategies are demanded to improve the capabilities of transmission and storage of such an amount of information. Compression is an effective way to reduce the size of the data produced. However, specific compression schemes should be designed to adapt to the particular nature of the IASI data. Currently, Principal Component Compression (PCC) is an accepted strategy to compress IASI products. Nonetheless, PCC entails several drawbacks related with the use of a training set to obtain a truncated principal components representation of the original signal. Accordingly, image compression techniques may be an adequate alternative to PCC. In this thesis we present a comprehensive study of IASI data compression based on state-of-the-art image compression techniques.

Our first analysis focused on the study of IASI L0 products, which are data sent from the instrument to the reception stations and the inputs of the on-ground processing chain. We investigated the order-0 entropy and the order-1, order-2, and order-3 context-based entropies on several IASI L0 products. Our analysis in Chapter 2 suggested that, using a simple variable-length code, at least 1.62 bpppc might be

saved. However, gains could be even higher if contextual models were used. Order-1 conditional contexts revealed that at least 3.78 bpppc might be spared. Order-2 conditional contexts allowed to save over 5.50 bpppc. The best results were achieved by order-3 conditional contexts, which yielded gains higher than 6.70 bpppc. Then, we translated these observations to a real compression scenario, where the compression performance of different state-of-the-art lossless compression techniques was investigated on IASI L0 products. We observed that compression ratios over 2.6:1 could be achieved.

On the basis of the aforementioned study, in Chapter 3 we extended the analysis to IASI L1C products, which are considered useful for end users. The assessment of IASI L0 data compression revealed that exploiting the spectral dimension is of utmost importance to achieve competitive compression performance. Therefore, we performed a computational complexity analysis of several spectral transforms and provided specific recommendations for IASI data. Based on this study, we proposed several compression schemes for lossless, near-lossless, and lossy compression of IASI L1C products. For lossless compression, we observed that compression techniques that already exploit the spectral redundancy by themselves like M-CALIC and CCSDS-123 also benefit from applying a spectral transform. Multilevel Clustering RKLT and RWA were the spectral transforms that achieved the best coding performance. Compression ratios over 2.5:1 were achieved for IASI L1C orbits. The near-lossless compression analysis revealed that small PAEs such as 1 and 3 achieved gains over 17% and 30%, respectively, compared to lossless compression, while a SNR Energy performance over 65 dB was obtained. For lossy compression, Multilevel Clustering KLT + JPEG 2000 yielded the best compression performance for all bit-rates analyzed. We also compared the reconstructed spectra from Multilevel Clustering KLT + JPEG 2000 and M-CALIC, which are the coding schemes that achieved the more competitive lossy and near-lossless compression performance, respectively, with reconstructed spectra from PCC and with the original spectra. This analysis suggested that the proposed compression schemes are able to produce useful reconstructed spectra to be used in further processings.

The promising coding performance of the proposed compression schemes for IASI

products led us to analyze the impact of compression in end-user applications. In Chapter 4 we investigated the impact of IASI L1C data compression when the reconstructed radiances were later used by statistical retrieval algorithms to predict physical parameters. We analyzed both near-lossless and lossy compression. Near-lossless compression produced significant improvements in the prediction performance at compression ratios of, approximately, 100:1. However, lossy compression achieved even better results. The retrieval results were further improved at higher compression ratios. We observed that, for lossy compression, the retrieval performance was more competitive when a spatial transform had been applied, which suggests that the spatial regularization produced by the spatial transform is behind the improvements. As a conclusion, the proposed lossy compression schemes revealed themselves as an effective strategy to exploit spectral and spatial feature relations in an indirect way, which benefits the retrieval methods.

On the basis of the previous analysis, the origin of the benefits produced by compression was investigated in Chapter 5. We focused on two consequences of compression. On the one hand, the noise filtering produced by compression was analyzed. We observed that the noise level in the recovered spectra decreased as the compression ratio increased. When the retrieval results were compared to the level of noise in the reconstructed data, we observed that the retrieval performance improved at low noise levels. If the compression ratio was higher than some point, the noise filtering effect vanished, which yielded degraded retrieval performance because relevant information was removed beyond this point. On the other hand, we analyzed the exploitation of spectral and spatial feature relations through the spectral and spatial transforms applied in the compression stage. To this end, we proposed a simple technique to specifically include spectral and spatial regularization consisting of a spectral and a spatial Gaussian filter to convolve neighboring pixels. We observed that the improvements in the retrieval performance achieved by compression decreased when the Gaussian filters was used, which suggests that the benefits produced by compression are mainly caused by the exploitation of spectral and spatial relations.

As a general conclusion of this thesis, we have carried out a comprehensive analysis of IASI data compression. Several lossless, near-lossless, and lossy compression

schemes have been proposed for IASI L0 and IASI L1C products, which are able to produce useful reconstructed spectra at high compression ratios. The quality of the reconstructed radiances have been analyzed on statistical retrieval algorithms. Experimental results suggest that compression benefits the retrieval results at moderate and high compression ratios. We have investigated the origin of the benefits and provided recommendations for competitive compression of IASI L1C products.

7.2 Future work

Even though competitive compression strategies for IASI data compression have been proposed in this thesis, the reported conclusions lead to new ideas for further research and contributions.

The theoretical entropy study carried out on IASI L0 data yielded gains over 80% when order-3 contextual models were employed. It seems interesting to extend this analysis to IASI L1C products. The large size of the data has led EUMETSAT to adopt lossy compression strategies, like PCC, to produce appropriate products for the near real-time dissemination. Effective lossless compression strategies might help to disseminate the whole products in such scenarios. As a future work, we aim to explore the potential of contextual entropy coding within the framework of lossless compression of IASI L1C data.

Regarding lossy compression, the spectral/spatial regularization produced during the compression stage has revealed itself to be the main reason of the retrieval improvements. Exploitation of spectral/spatial feature relations is an indirect consequence of using spectral and spatial transforms. Therefore, it is worth to further investigate the impact of transforms on the retrieval methods. Analyzing additional transforms, testing another parameter configuration and settings, and adapting state-of-the-art transforms to enhance the retrieval results and the compression performance are areas of interest.

Another area of interest is to investigate the impact of lossy compression on other application fields. For instance, neural networks, radiative transfer models, trace gases measurements, etc.

Finally, the IASI-C instrument is planned to be launched in October 2018. The nature of the IASI next-generation (IASI-NG) products shall be different to the nature of the IASI-A and IASI-B products. Both the instrument characteristics and the processing chain shall evolve with respect to the previous IASI instruments. Therefore, IASI-NG data compression is an interesting research field. Lossless, near-lossless, and lossy compression of IASI-NG products and the impact of lossy compression on end-user applications are two attractive areas that are still unexplored.

Bibliography

[1] EUMETSAT, "EUMETSAT Polar System," Available: https://www.eumetsat.int/eps_webcast/eps/index.htm.

[2] ——, "IASI instrument," Available: https://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/IASI/index.html.

[3] ——, "MetOp mission," Available: http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/index.html.

[4] A. M. Step, "The EUMETSAT Polar System," ESA bulletin, 127, 19.

[5] K. D. Klaes, M. Cohen, Y. Buhler, P. Schlëssel, R. Munro, A. Engeln, E. Clérigh, H. Bonekamp, J. Ackermann, J. Schmetz, and J. P. Luntama, "An introduction to the EUMETSAT polar system," *Bulletin of the American Meteorological Society*, vol. 88, no. 7, pp. 1085–1096, 2007.

[6] N. Pougatchev, T. August, X. Calbet, T. Hultberg, O. Oduleye, P. Schlüssel, B. Stiller, K. St. Germain, and G. Bingham, "IASI temperature and water vapor retrievals-error assessment and validation," *Atmospheric Chemistry and Physics*, vol. 9, no. 17, pp. 6453–6458, 2009.

[7] M. George, C. Clerbaux, D. Hurtmans, S. Turquety, P. F. Coheur, M. Pommier, J. Hadji-Lazaro, D. P. Edwards, H. Worden, M. Luo, C. Rinsland, and W. McMillan, "Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors," *Atmospheric Chemistry and Physics*, vol. 9, no. 21, pp. 8317–8330, 2009.

[8] B. Tournier, D. Blumstein, F. Cayla, , and G. Chalon, "IASI Level 0 and 1 processing algorithms description," in *Proc. of ISTCXII Conference*, 2002.

[9] ESA, "IASI Data Processing Chain," Available: http://www.esa.int/Our_Activities/Observing_the_Earth/The_Living_Planet_Programme/Meteorological_missions/MetOp/Data_processing_chain (accessed on June 2017).

[10] EUMETSAT, *IASI Level 1: Product Guide*, REF. EUM/OPS-EPS/MAN/04/0032, 2012.

[11] ——, *IASI Level 2: Product Guide*, REF. EUM/OPS-EPS/MAN/04/0033, 2012.

[12] EUMETCast, "EUMETCast website," Available: http://www.eumetsat.int/website/home/Data/DataDelivery/EUMETCast/index.html (accessed on June 2017).

[13] T. Hultberg, "IASI Principal Component Compression (IASI PCC) FAQ," Available: http://www.eumetsat.int/website/wcm/idc/idcplg?IdcService=GET_FILE&dDocName=pdf_ipcc_faq&RevisionSelectionMethod=LatestReleased&Rendition=Web.

[14] T. Hultberg, T. August, N. C. Atkinson, and F. Smith, "IASI PC compression - searching for signal in the residuals," in *ECMWF/EUMETSAT NWP-SAF Workshop Efficient Representation of Hyper-spectral Infrared Satellite Observations*, 2013.

[15] F. Hilton and A. D. Collard, "Recommendations for the use of principal component-compressed observations from infrared hyperspectral sounders," Met Office Forecasting R&D Technical Report, 536, 2009.

[16] N. C. Atkinson, F. I. Hilton, S. M. Illingworth, J. R. Eyre, and T. Hultberg, "Potential for the use of reconstructed IASI radiances in the detection of atmospheric trace gases," *Atmospheric Measurement Techniques*, vol. 3, no. 4, pp. 991–1003, 2010.

[17] J. García-Sobrino, I. Blanes, M. Albinet, R. Camarero, J. Serra-Sagristà, V. Laparra, X. Calbet, and G. Camps-Valls, "Proposal for Infrared Atmospheric Sounding Interferometer on-board data compression," *SPIE Journal of Applied Remote Sensing*, vol. 9, no. 1, p. 097498, 2015.

[18] J. García-Sobrino, J. Serra-Sagristà, and J. Bartrina-Rapesta, "Hyperspectral IASI L1C Data Compression," *Sensors*, vol. 17, no. 6, p. 1404, 2017.

[19] J. García-Sobrino, J. Serra-Sagristà, V. Laparra, X. Calbet, and G. Camps-Valls, "Statistical Atmospheric Parameter Retrieval Largely Benefits From Spatial-Spectral Image Compression," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 4, pp. 2213–2224, 2017.

[20] J. García-Sobrino, V. Laparra, J. Serra-Sagristà, X. Calbet, and G. Camps-Valls, "Improved Statistically Based Retrievals via Spatial-Spectral Data Compression for IASI data," *Remote Sensing*, vol. xx, no. xx, p. xx, 2018, Under Review.

[21] Consultative Committee for Space Data Systems (CCSDS), *Lossless Multispectral & Hyperspectral Image Compression CCSDS 123.0-B-1*, ser. Blue Book. CCSDS, May 2012, Available: https://public.ccsds.org/Pubs/123x0b1ec1.pdf.

[22] ISO/IEC, "JPEG-LS Lossless and Near-Lossless Compression for Continuous-Tone Still Images," 1999.

[23] D. Taubman and M. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice*, ser. Vol. 642. Springer Science & Business Media, 2012.

[24] E. Magli, G. Olmo, and E. Quacchio, "Optimized onboard lossless and near-lossless compression of hyperspectral data using CALIC," *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 1, pp. 21–25, January 2004.

[25] I. Blanes and J. Serra-Sagristà, "Pairwise Orthogonal Transform for Spectral Image Coding," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 961–972, March 2011.

[26] Consultative Committee for Space Data Systems (CCSDS), *Image Data Compression CCSDS 122.0-B-1*, ser. Blue Book. CCSDS, 2005, Available: https://public.ccsds.org/Pubs/122x0b1c3.pdf (accessed on June 2017).

[27] ISO/IEC, *High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding*, 2013, Available: https://www.iso.org/standard/69668.html.

[28] I. Blanes and J. Serra-Sagristà, "Clustered Reversible-KLT for Progressive Lossy-to-Lossless 3d Image Coding," in *IEEE Data Compression Conference, DCC'09*, March 2009, pp. 233–242.

[29] D. Salomon, *Data Compression: the complete reference*. Springer Science & Business Media, 2004.

[30] N. Amrani, J. Serra-Sagristà, V. Laparra, M. W. Marcellin, and J. Malo, "Regression Wavelet Analysis for Lossless Coding of Remote-Sensing Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 9, pp. 5616–5627, 2016.