BBD: A NEW BAYESIAN BI-CLUSTERING DENOISING ALGORITHM FOR IASI-NG HYPERSPECTRAL IMAGES

M. Colom[1], *G. Blanchet*[2], *A. Klonecki*[3], *O. Lezeaux*[3], *E. Pequignot*[2], *F. Poustomis*[3], *C. Thiebaut*[2], *S. Ythier*[3], *J.-M. Morel*[1]

[1] CMLA, ENS Cachan. CNRS, Université Paris-Saclay. 94235, Cachan, France.

[2] Centre National d'Études Spatiales, Toulouse, France.

[3] NOVELTIS, 153 rue du Lac, 31670 Labège, France.

ABSTRACT

We propose a new denoising method for 3D hyperspectral images for the future MetOp-Second Generation series satellite incorporating the new IASI-NG interferometer, to be launched in 2021. This adaptive method retrieves the data model directly from the input noisy granule, using the following techniques: dual clustering (spectral and spatial), dimensionality reduction by adaptive PCA, and Bayesian denoising. The use of dimensionality reduction by PCA has been already proven an effective denoising technique because of intrinsic data redundancy. We demonstrate here that by combining a local PCA dimensionality reduction with a dual clustering and a Bayesian denoising, it is possible to improve significantly the PSNR with respect to PCA reduction alone. This noise reduction hints at the possibility to multiply of the resolution of the satellite by factor 4, while keeping an acceptable SNR .

Index Terms— IASI-NG, denoising, clustering, Bayesian, PCA

1. INTRODUCTION

In 2021 EUMETSAT will launch the first MetOp-SG satellite carrying a new interferometer (IASI-NG) developed by CNES[1] to measure weather variables, pollution and climate monitoring, and determining atmospheric gas composition (including $H_2O, CO_2, O_3, N_2O, CO, CH_4$) [1, 2]. The interferometer is made of 4 captors (one 4×4 PC MCT detector for band B1, and MOVPE PV detectors for B2, B3 and B4) which provide 16923 spectral channels between 15.5 and 3.63 μm with a spectral resolution of 0.125 cm^{-1}, and a FOV of 4×4 pixels for a box of 100×100 km^2 at nadir [3]. The noise level arriving to the captors is very low with respect to the signal and in general for HS sounding applications lossless compression is preferred to avoid damaging the data [4].

Our goal is to denoise 3D hyperspectral IASI-NG images (also called *granules*) taking advantage of the large redundancy of the data. A good enough denoising performance would simply permit to increase the satellite resolution. As a rule of thumb, dividing the noise by a factor p would permit to increase the resolution by the same factor. Denoising of HS granules by using PCA and dimensionality reduction has been already explored [5, 6, 7, 8], as well as the problem of grouping similar signals according to their wavenumber [9]. Here we propose a Bayesian method to denoise IASI-NG granules combining dual (spectral and spatial) clustering, a dimensionality reduction, and an optimal Bayesian method. We call this method *BBD*.

2. THE NOISE MODEL

The IASI-NG instrument is made of four different HS captors with spectral overlapping of ± 40 cm^{-1}, each of them affected by additive Gaussian noise which depends on both the wavenumber and the intensity (photonic noise). We take into account the IASI-NG worst case for the photonic noise, and therefore our noise model only depends on the wavenumber (detector noise relying on the focal plane temperature, readout, and constant ADC noise).

Thus, it is possible to characterize the noise model as a function which gives the standard deviation of the Gaussian noise according to the wavenumber. Fig. 1 shows such a noise level function. Its peaks are caused by the performance decay at the extremity of bands. The amplitude of the signal is more than 4700 times higher than the noise.

3. IASI-NG SIMULATED NOISY DATA

The IASI-NG instrument is planned to be launched in 2021 and therefore there are no real granule data available. However, the noise model (Sec. 2) is known in advance and well characterized, and it is therefore possible to generate reliable simulations of the data. In the framework of an EUMETSAT study, NOVELTIS and University of Basilicata have simulated such IASI-NG data. This synthetic data are based on the radiative transfer code σ-IASI RT code [10], and takes into account the IASI-NG configuration. To initialize the Radiative Transfer computations several databases have been used

[1] The French aerospace agency.

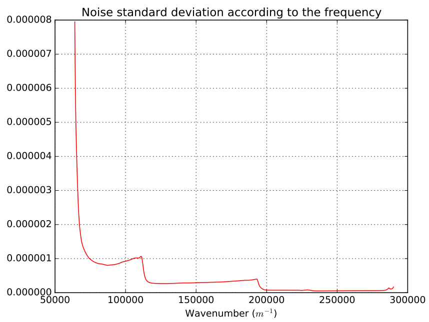


Fig. 1. Noise level curve giving the standard deviation of the noise according to the wavenumber. The noise depends only on the wavenumber, but not on the pixel intensity (since only the IASI-NG worst case for photonic noise is considered).

such as AVHRR products, ECMWF, and MODIS data. The spectra generated by the RT code were then processed to provide simulated L1C IASI-NG noise-free data, including the apodisation step. These synthetic data have been validated by performing a statistical analysis and comparing them with geophysical features[2]. We consider the usual compression technique known as *bit-trimming*, consisting in encoding the floating point values at each pixel with a fixed number of bits. The threshold is set to one quarter of the energy of the noise.

The procedure we use to simulate IASI-NG noisy data is the following: **(1)** Use noise free data simulated using RT code; **(2)** add wavenumber-dependent noise to the granule (Fig. 1); **(3)** apply a simple bit-trimming and obtain the input noisy granule, **(4)** denoise the input noisy granule and measure the denoising performance.

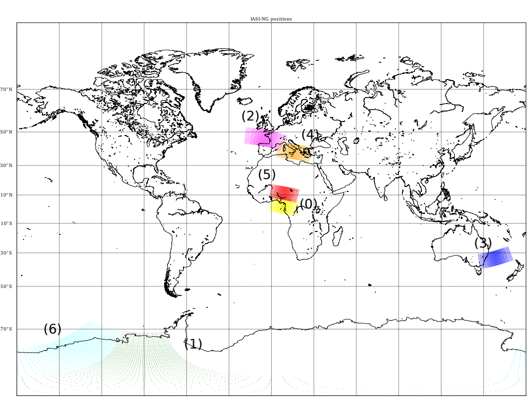


Fig. 2. Location of the seven IASI-NG granules.

[2]EUMETSAT internal report: NOV-7323-NT-3649_v4.2.

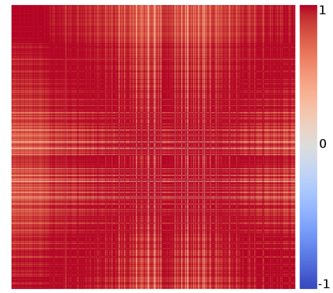
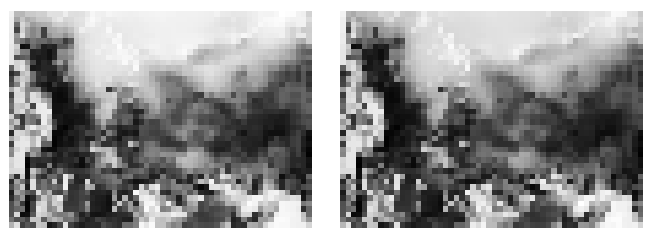


Fig. 3. Normalized Pearson's correlation matrix for band B2 (from 115000.0 m^{-1} to 195987.5 m^{-1}) in granule #1. The red color represents a high correlation, while white means low (*coolwarm* color map). Most of the channels are correlated with many of the others. Only a few frequencies are uncorrelated because of the presence of gases in the atmosphere at those particular frequencies.

4. THE DENOISING ALGORITHM

This section describes the denoising method. Its pseudo-code is given in Algo. 1. Most of the frequencies are highly correlated with the rest (see Fig. 3 and also the images acquired in Fig. 4), with the exception of a few of them. A fundamental principle of denoising is to exploit the auto-similarity of the data [11]. Thus, we build for each granule clusters of frequencies by grouping highly correlated frequencies, with the K-means clustering algorithm. We have fixed $Q = 32$ clusters of frequencies, denoted as Λ_i with $i \in [1, Q]$.



(a) 88362.5 m^{-1} (b) 96500.0 m^{-1}

Fig. 4. Simulation of the images acquired by the IASI-NG instrument at different wavenumbers (granule #0, band B1). The intensity of each pixel represents the energy measured in $\mathrm{Wm^{-2}sr^{-1}m}$ units.

Each Λ_i contains highly-correlated frequencies for each pixel. The next step is to reduce the dimensionality of each pixel to $N = 20$ dimensions by PCA[3]. We verified empirically that 20 PCs are enough the represent reliably the IASI-NG granules (Fig. 5). We denote by $\tilde{\mathbf{A}}$ the pixels in Λ_i after dimensionality reduction.

For each cluster Λ_i we apply an adaptation of the NL-Bayes algorithm [12], where P is the denoised pixel, and $\tilde{P}$ the noisy pixel after bit-trimming:

[3]We used the `prcomp` function in *R*.

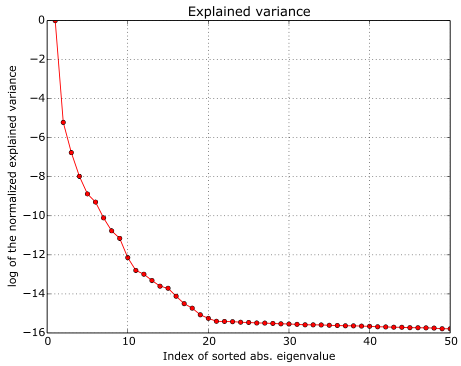


Fig. 5. Decay of the normalized explained variance according to the sorted eigenvalues, in $\log$ scale.

$$P = \overline{\tilde{P}} + \left[\mathbf{C}_{\tilde{P}} - \mathbf{C}_n\right] \mathbf{C}_{\tilde{P}}^{-1} \left(\tilde{P} - \overline{\tilde{P}}\right). \tag{1}$$

$\mathbf{C}_{\tilde{P}}$ is the empirical covariance matrix of the most similar pixels in the cluster. We use the Pearson's correlation coefficient as the similarity criterion[4]. The empirical covariance matrix $\mathbf{C}_n$ of the bit-trimmed noise after PCA transformation is required too. It needs to be computed only for the wavenumbers in Λ_i. Sec. 4.1 gives the details on how to compute $\mathbf{C}_n$. For each cluster Λ_i, the NL-Bayes algorithm is applied. Since each cluster contains only a part of the total set of wavenumbers, each pixel is denoised partly at each iteration. After all iterations, the pixel is completely denoised. Finally, the denoised pixels are unprojected (the data is expressed in the same axes as before the PCA rotation) and the whole denoised granule is obtained.

4.1. Calculation of $\mathbf{C}_n$

In this section we explain how to calculate the empirical covariance matrix of the noise $\mathbf{C}_n$ required in Eq. (1), under PCA axes rotation. Let us define $\mathbf{N}$ as an $n \times p$ matrix containing the noise. Each of the n rows corresponds to a particular noisy pixel observation and each of the p columns to a variable (wavenumber) of the noisy pixel. The PCA requires that the data is centered. In practice, the barycenter of the set of noisy pixels is subtracted from each column. In our case, since the noise has zero mean along any wavenumber, we can use directly $\mathbf{N}$ and compute its eigenvalues and eigenvectors. We shall call $\mathbf{W}$ the $p \times p$ matrix which contains in its rows the normalized eigenvectors of $\mathbf{N}$. Any wavenumber $\mathbf{F}_j$ with $j \in [1, n]$ can be written in the PCA rotated axes as $\mathbf{F}_j = \mathbf{W}\mathbf{N}_j$. Thus, the entries of the covariance matrix at $[j_1, j_2]$ are $\mathrm{cov}\left(\mathbf{F}_{j_1}, \mathbf{F}_{j_2}\right) = \mathbb{E}\left(\mathbf{F}_{j_1}\mathbf{F}_{j_2}^T\right) - \mathbb{E}\left(\mathbf{F}_{j_1}\right)\mathbb{E}\left(\mathbf{F}_{j_2}^T\right) = \mathbb{E}\left(\mathbf{F}_{j_1}\mathbf{F}_{j_2}^T\right) = \mathbb{E}\left(\mathbf{W}\mathbf{F}_{j_1}\mathbf{F}_{j_2}^T\mathbf{W}^T\right) = \mathbf{W}\mathbb{E}\left(\mathbf{F}_{j_1}\mathbf{F}_{j_2}^T\right)\mathbf{W}^T$.

We shall call $\mathbf{D} = \mathbb{E}\left(\mathbf{F}_{j_1}\mathbf{F}_{j_2}^T\right) = \begin{cases} \sigma_{j_1}^2 + \sigma_{\mathbf{F}_{j_1}}^2 & \text{if } j_1 = j_2 \\ 0 & \text{otherwise,} \end{cases}$

and thus $\mathrm{cov}\left(\mathbf{F}_{j_1}, \mathbf{F}_{j_2}\right) = \mathbb{E}\left(\mathbf{W}\mathbf{D}\mathbf{W}^T\right)$. The variance $\sigma_{\mathbf{F}_{j_1}}^2$ is added because a bit-trimming compression is applied to the noisy granule, which can be understood as adding a quantization noise. Let us define $\mathcal{B}\left(\mathbf{F}_j\right)$ as the bit-trimming operator applied to the pixels in a given wavenumber $\mathbf{F}_j$. Then,

$$\sigma_{\mathbf{F}_j}^2 = \mathrm{var}\left(\left|\mathbf{F}_j - \mathcal{B}\left(\mathbf{F}_j\right)\right|\right). \tag{2}$$

Finally, the empirical covariance matrix of the noise in the rotated PCA axes can be written as

$$\mathbf{C}_n = \mathbf{W}\mathbf{D}\mathbf{W}^T. \tag{3}$$

5. RESULTS

To measure the performance of our denoising method we defined the *MSNR*[5] metric as:

$$\mathrm{MSNR}_j(\mathbf{T}, \mathbf{G}) = 10 \log_{10}\left[\frac{\mathrm{median}(\mathbf{T}_j)^2}{\mathrm{MSE}(\mathbf{T}_j, \mathbf{G}_j)}\right], \tag{4}$$

where j is a given wavenumber, $\mathbf{T}$ the testing granule, $\mathbf{G}$ the reference noise-free granule, and MSE stands for the Mean Squared Error. Note that the formula (4) is the same as the usual PSNR for a wavenumber j, with the difference that we use the median of the noisy granule instead of its maximum value. The reason is to avoid the effect of outlier values which would bias the measurement of performance.

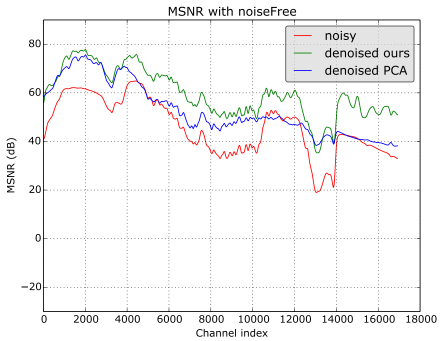
Fig. 6 shows the MSNR plots of two granules[6] and compares the MSNR of the noise (red), with our denoising method (green), and the results obtaining by simply performing PCA and keeping the $N = 20$ most significant PCs. As it can be observed, since the signal is highly correlated (see Fig. 3), simply applying PCA and keeping the most significant PCs is an effective strategy. However, combining dimensionality reduction with Bayesian denoising improves significantly the result.

Table 1 shows the square root of the ratio between the MSEs of the noisy and denoised granules with respect to the noise-free granule.

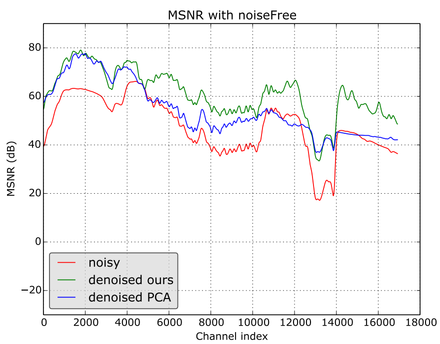
[4]It can be computed quickly with the `corrcoef` function in *NumPy*, for example.

[5]Median Signal-to-Noise Ratio.

[6]Similar results are obtained with the other granules.



(a) Granule #2



(b) Granule #5

Fig. 6. MSNR denoising performances of two of the seven granules analyzed. The reference is the noise-free granule *without* bit-trimming. In red: the MSNR corresponding to the noisy granule without bit-trimming. In green: the MSNR of our denoising method, with bit-trimming. In blue: the denoising result only applying PCA dimension reduction with $N = 20$ principal components, with bit-trimming. Since the curves oscillate much, they have been filtered with a Gaussian of $\sigma = 50$ for visualization and comparison purposes.

6. CONCLUSIONS

We have presented a new method to denoise hyperspectral 3D images. This method has been successfully applied to simulated IASI-NG spectra. IASI and IASI-NG data have been used as example to demonstrate the denoising capabilities and to illustrate the difficulties. However, the method is general enough to be applied to other types of HS images by choosing the right parameters $\{Q, K, N\}$ in Algo. 1. The method is unsupervised and uses several techniques to adapt to the data of each particular granule: first it creates clusters of frequencies to group data supposed to come from the same physical origin, then it reduces the dimensionality of the data by means of PCA based on the data of the granule, and finally it uses another kind of clustering (in this case based of inter-pixel correlation) to denoise the granule. The model for the data (in the form of the covariance matrix of similar pixels, $\mathbf{C}_{\tilde{P}}$) is learned from the input granule itself instead

Granule	#0	#1	#2	#3	#4	#5	#6
$\sqrt{\frac{\mathrm{MSE}(\mathcal{B}(\tilde{\mathbf{G}}),\mathbf{G})}{\mathrm{MSE}(\hat{\mathbf{G}},\mathbf{G})}}$	3.53	4.73	4.25	4.79	4.23	3.49	4.85

Table 1. Denoising gain as the square root of the ratio between the MSEs of the bit-trimmed noisy and denoised granules with respect to the noise-free granule. Average: **4.27**.

Algorithm 1 Denoise a IASI-NG hyperspectral image

```
Denoising
    Input G̃: acquired noisy granule
    Input Q = 32: number of spectral clusters
    Input K = 400: number of similar pixels
    Input N = 20: number of PCA PCs kept
    Output Ĝ: denoised granule

Ĝ = zeros(G̃.shape)                                  ▷ Placeholder
Split the spectrum into Q frequential clusters Λ_i, i ∈ [1, Q]
for each cluster Λ_i, i ∈ [1, Q] do
    Ã = PCA(Λ_i, N)   ▷ Reduce dimensionality to N variables
    Compute the covariance matrix C_n of the noise     ▷ Eq. (3)
    for each noisy pixel P̃ do
        For the K most correlated pixels compute: (i) their mean
        P̃̄ and (ii) their covariance matrix C_P̃
        Obtain the denoised pixel P using NL-Bayes     ▷ Eq. (1)
        Â ← P                                      ▷ Store denoised P
    end for
    Ĝ(Λ_i) ← UNPROJECT(Â)           ▷ Back to the original axes
end for
return Ĝ
```

of using a predefined model. The use of dimensionality reduction has two main benefits: first, after projection on the PCA basis, the underlying signal and the noise are much separated, which allows a better comparison of the pixels taking into account the signal in the first $N = 20$ PCs and not the noise. This can be understood as some kind of *oracle* technique to find similar pixels, in analogy with explicitly oracular methods like BM3D [13] . In our case, it is done implicitly. The second benefit is that the drastic dimensionality reduction (from 16923 spectral channels to $N = 20$ PCs) speeds up significantly the computations.

The quantitative results show that the combination of the bi-clustering, dimensionality reduction, and Bayesian denoising implies a gain in several dBs with respect to simply keeping the most significant PCA PCs. The reduction of the noise in a factor of 4 shown in Table 1 means that the surface of the captor could be reduced in a factor 16 while keeping the same performance. As future work, we aim to obtain denoising results for existing real IASI satellite images and use different metrics to evaluate the denoising performance, as for example the inter-channel correlation and the off-diagonal coefficients of the noise covariance matrix before and after denoising.

7. REFERENCES

[1] C. Crevoisier, C. Clerbaux, V. Guidard, T. Phulpin, R. Armante, B. Barret, Camy-Peyret C., J. P. Chaboureau, P. F. Coheur, L. Crépeau, et al., "Towards IASI-new generation (IASI-NG): impact of improved spectral resolution and radiometric noise on the retrieval of thermodynamic, chemistry and climate variables," *Atmospheric Measurement Techniques*, vol. 7, no. 12, pp. 4367–4385, 2014.

[2] Cyril Crevoisier, Cathy Clerbaux, Vincent Guidard, Eric Pequignot, and Frédérick Pasternak, "IASI-NG: un concentré d'innovations technologiques pour l'étude de l'atmosphère terrestre," *La Météorologie*, , no. 86, pp. 3–5, 2014.

[3] F. Bernard, B. Calvel, F. Pasternak, R. Davancens, C. Buil, E. Baldit, C. Luitot, and A. Penquer, "Overview of IASIw-NG the new generation of infrared atmospheric sounder," in *International Conference on Space Optics*, 2014, vol. 7, p. 10.

[4] Bormin Huang, Hung-Lung A Huang, Alok Ahuja, Timothy J. Schmit, and Roger W. Heymann, "Lossless data compression for infrared hyperspectral sounders: an update," in *Optical Science and Technology, the SPIE 49th Annual Meeting*. International Society for Optics and Photonics, 2004, pp. 109–119.

[5] Guangyi Chen and Shen-En Qian, "Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 3, pp. 973–980, 2011.

[6] Behnood Rasti, Johannes R Sveinsson, Magnus O Ulfarsson, and Jon Atli Benediktsson, "Hyperspectral image denoising using 3d wavelets," in *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*. IEEE, 2012, pp. 1349–1352.

[7] P. Antonelli, H. E. Revercomb, L. A. Sromovsky, W. L. Smith, R. O. Knuteson, D. C. Tobin, R. K. Garcia, H. B. Howell, H. L. Huang, and F. A. Best, "A principal component noise filter for high spectral resolution infrared measurements," *Journal of Geophysical Research: Atmospheres*, vol. 109, no. D23, 2004.

[8] A. D. Collard, A. P. McNally, F. I. Hilton, S. B. Healy, and N. C. Atkinson, "The use of principal component analysis for the assimilation of high-resolution infrared sounder observations for numerical weather prediction," *Quarterly Journal of the Royal Meteorological Society*, vol. 136, no. 653, pp. 2038–2050, 2010.

[9] Jean-Michel Gaucel, Carole Thiebaut, Romain Hugues, and Roberto Camarero, "On-board compression of hyperspectral satellite data using band-reordering," in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2011, pp. 81570W–81570W.

[10] Giuseppe Grieco, Carmine Serio, and Guido Masiello, "σ-IASI-β: a hyperfast radiative transfer code to retrieve surface and atmospheric geophysical parameters," *EARSeL eProceedings*, vol. 12, no. 2, pp. 149, 2013.

[11] M. Lebrun, M. Colom, A. Buades, and J. M. Morel, "Secrets of image denoising cuisine," *Acta Numerica*, vol. 21, no. 1, pp. 475–576, 2012.

[12] Marc Lebrun, Antoni Buades, and Jean-Michel Morel, "A nonlocal Bayesian image denoising algorithm," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1665–1688, 2013.

[13] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080–2095, 2007.

Acknowledgement: Centre National d'Etudes Spatiales (CNES, R&T action), EUMETSAT, European Research Council (Advanced Grant *Twelve Labours*), Office of Naval Research (Grant N00014-97-1-0839), Direction Générale de l'Armement (DGA).