



**HAL**  
open science

# Detection and Correction of Glitches in a Multiplexed Multichannel Data Stream-Application to the MADRAS Instrument

Herwig Wendt, Nicolas Dobigeon, Jean-Yves Tourneret, Mathieu Albinet,  
Christophe Goldstein, Nadia Karouche

► **To cite this version:**

Herwig Wendt, Nicolas Dobigeon, Jean-Yves Tourneret, Mathieu Albinet, Christophe Goldstein, et al.. Detection and Correction of Glitches in a Multiplexed Multichannel Data Stream-Application to the MADRAS Instrument. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, vol. 54 (n° 5), pp. 2803-2811. 10.1109/TGRS.2015.2505902 . hal-01338814

**HAL Id: hal-01338814**

**<https://hal.science/hal-01338814>**

Submitted on 29 Jun 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>  
Eprints ID : 15770

**To link to this article** : DOI:10.1109/TGRS.2015.2505902  
URL : <http://dx.doi.org/10.1109/TGRS.2015.2505902>

**To cite this version** : Wendt, Herwig and Dobigeon, Nicolas and Tourneret, Jean-Yves and Albinet, Mathieu and Goldstein, Christophe and Karouche, Nadia *Detection and Correction of Glitches in a Multiplexed Multichannel Data Stream—Application to the MADRAS Instrument*. (2016) IEEE Transactions on Geoscience and Remote Sensing, vol. 54 (n° 5). pp. 2803-2811. ISSN 0196-2892

Any correspondence concerning this service should be sent to the repository administrator: [staff-oatao@listes-diff.inp-toulouse.fr](mailto:staff-oatao@listes-diff.inp-toulouse.fr)

# Detection and Correction of Glitches in a Multiplexed Multichannel Data Stream—Application to the MADRAS Instrument

Herwig Wendt, *Member, IEEE*, Nicolas Dobigeon, *Senior Member, IEEE*,  
Jean-Yves Tournet, *Senior Member, IEEE*, Mathieu Albinet,  
Christophe Goldstein, and Nadia Karouche

**Abstract**—This paper presents a new strategy to correct the Earth data corrupted by spurious samples that are randomly included in the multiplexed data stream provided by the MADRAS instrument. The proposed strategy relies on the construction of a trellis associated with each scan of the multichannel image, modeling the possible occurrences of these erroneous data. A specific weight that promotes the smooth behavior of the signals recorded in each channel is assigned to each transition between trellis states. The joint detection and correction of the erroneous data are conducted using a dynamic programming algorithm for minimizing the overall cost function throughout the trellis. Simulation results obtained on synthetic and real MADRAS data demonstrate the effectiveness of the proposed solution.

**Index Terms**—Destriping, dynamic programming, MADRAS, multiband imaging, multiplexing.

## I. INTRODUCTION

**B**ORN from a close collaboration between the Indian and French space agencies (namely, ISRO and CNES, respectively), the MEGHA-Tropiques mission aims at developing a monitoring system dedicated to the study of the tropical atmosphere [1]. The measurements collected over the intertropical belt by multiple sensors embedded on the spacecraft platform allow various ocean and atmospheric parameters of interest (e.g., rain rate, profile of water vapor content, and sea surface wind) to be determined with high spatial and temporal sampling [2]. These climate and atmospheric parameters are disseminated over the scientific community through academic institutions and national agencies, whose objectives are, e.g., climate research, weather forecasting, and prediction of major events (i.e., monsoons) [3].

H. Wendt, N. Dobigeon, and J.-Y. Tournet are with the IRIT/INP-École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEIH), Centre National de Recherche Scientifique (CNRS), University of Toulouse, 31071 Toulouse, France (e-mail: Herwig.Wendt@irit.fr; Nicolas.Dobigeon@enseeiht.fr; Jean-Yves.Tournet@enseeiht.fr).

M. Albinet, C. Goldstein, and N. Karouche are with the Centre National d'Études Spatiales, 31401 Toulouse, France (e-mail: Mathieu.Albinet@cnes.fr; Christophe.Goldstein@cnes.fr; Nadia.Karouche@cnes.fr).

The satellite payload is composed of four instruments: GPS-ROSA, a GPS occultation sensor designed to provide atmospheric temperature and humidity profiles; SCARAB, an optical radiometer retrieving the radiation parameters; SAPHIR, a microwave sensor for vertical humidity profiling; and MADRAS, a microwave imager used to provide rain and cloud properties [4]. This latter sensor, jointly developed by ISRO (for the scan mechanisms) and CNES (for the radio-frequency subsystems), is a passive conical microwave imager measuring the radiation at nine frequency bands, at vertical and horizontal polarizations. The exploitation of the scientific data collected by MADRAS has already motivated several studies to retrieve rainfall parameters, demonstrating the interest in these parameters by the science data users [5]–[9]. However, after a few weeks in orbit, an anomaly in the communication chain between two electronic devices was detected [10]. This anomaly leads to a mixing of the channels that compose the images provided by MADRAS. More precisely, additional data can be randomly inserted into the main data streams associated with each column of the MADRAS images. Visually, these corruptions result in the occurrence of vertical stripe noise, i.e., vertically and contiguously distributed erroneous pixels in the columns that compose the MADRAS images. Stripe noise, which generally comes from undesirable gain and/or offset variations of the sensors, is a common and well-known degradation that affects, for instance, images acquired by push-broom scanners. Thus, destriping has motivated numerous research works for several decades, not only for Earth remote sensing images [11]–[13] but also for biomedical images [14], [15] and astronomical data [16], [17]. Most of these destriping methods consist of locating the affected pixels in the image domain or using an appropriate representation (e.g., subspace, wavelet, or histogram), and then replacing them by spatially interpolated or more probable values. However, in the case of the MADRAS applicative context, such interpolation-like techniques remain prohibited to maintain the highest integrity of the scientific data and also to guarantee the confidence that the scientists may have in their results. This constraint makes inapplicable all the destriping methods proposed in the literature. Fortunately, after thorough analysis of the corruption process that affects the MADRAS images, it appears that the corrupted data streams still contain most of the measurements of interest, but in a wrong order. By removing the spurious extra data, one may expect that the correct order of the measurements can be

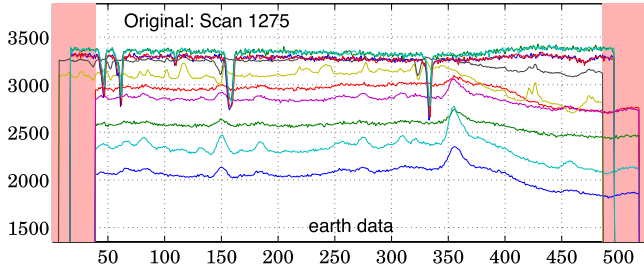


Fig. 1. Example of a scan without corruption. The Earth data of interest are those not masked by pink areas.

reestablished to obtain exploitable scientific data. This finding opens the door for a correction method that fulfills the initial requirement of avoiding any creation of new pixel values, e.g., by interpolation, which is precisely the objective of this paper.

In this paper, we focus on the correction of the anomalies in the scientific data, called Earth data, collected by the MADRAS instrument. The problem is formulated as the detection and the removal of spurious data in a data stream resulting from a cyclic multiplexing of several individual signals. This problem may be encountered in various applicative contexts where physical data are measured in several channels, such as multiband (e.g., multispectral or hyperspectral) imaging. The proposed solution relies on the construction of an oriented graph, or *trellis*, modeling the possible occurrences of abnormal samples in the data stream. Application-driven weights are proposed and associated with transitions between trellis states. A similar approach has been adopted in [18] to detect and correct errors encountered in automatic identification systems benefiting from a cyclic redundancy check. Finally, a Viterbi-like dynamic programming algorithm [19]–[24] is designed to recover the optimal path of minimal cumulative weight through the trellis.

This paper is organized as follows. The MADRAS multi-channel images and the problem to be solved are described in Section II. The strategy proposed to detect and correct possible anomalies in the MADRAS data is introduced in Section III. Section IV reports experimental results. Finally, some conclusions are drawn in Section V.

## II. PROBLEM STATEMENT

### A. MADRAS Data

The scientific data acquired by MADRAS take the form of a multichannel image, as depicted in Fig. 5 (first panel) using an arbitrary composition color. This image, composed of  $M = 11$  individual channels, consists of a set of  $P$  contiguous *scans*, where one given scan corresponds to a unique column of this image. After sampling correction, each scan is composed of  $T$  multivalued pixels, called *frames*. Each frame is thus a vector of  $M$  individual samples and corresponds to a given pixel observed in the  $M$  channels. The number  $T$  of frames depends on the type of acquired data: Earth data, which are considered in this paper, are composed of  $T = 526$  frames. A typical example of a scan is depicted in Fig. 1, where the signals recorded in the  $M$  channels are depicted in distinct colors.

To summarize, a scan can be given in the form of a matrix of size  $M$  (channels)  $\times$   $T$  (frames), where the first (last,

TABLE I  
NOTATIONS AND NOMENCLATURE

Notation	Definition
$P$	number of scans (i.e., image columns)
$T$	number of frames (i.e., pixels per image column)
$M$	number of channels (i.e., pixel dimension)
$N$	total number of samples per scan, i.e., in the data stream ( $= TM$ )
$x(j)$	sample # $j$ in the data stream
$\mathbf{x}(j)$	first $j$ samples of the data stream ( $= [x(1), \dots, x(j)]$ )
$\mathbf{x}(N)$	full data stream composed of $N$ samples ( $= [x(1), \dots, x(N)]$ )
$S$	number of trellis states ( $=$ max. number of glitches per scan)
$c_{k,j}$	trellis node associated with sample $x(j)$
$v_{k,j}^0$	branch connecting $c_{k,j-1}$ and $c_{k,j}$ (" $x(j)$ is not a glitch")
$v_{k,j}^1$	branch connecting $c_{k-1,j-1}$ and $c_{k,j}$ (" $x(j)$ is a glitch")
$d_0(k,j)$	weight associated with branch $v_{k,j}^0$
$d_1(k,j)$	weight associated with branch $v_{k,j}^1$
$D(k,j)$	cumulative weights along the unique path reaching the node $c_{k,j}$
$N_{k,j}$	number of accepted samples along the path reaching $c_{k,j}$
$\hat{\mathbf{x}}_{k,j}$	sequence of the $N_{k,j}$ samples along the path reaching $c_{k,j}$

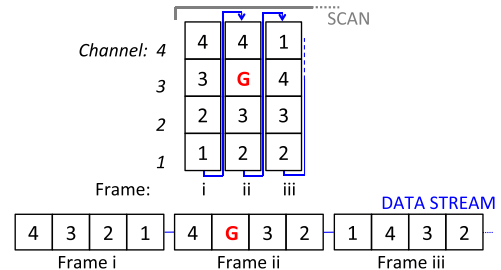


Fig. 2. Schematic view of anomaly in the data stream. The symbol "G" is used to represent the glitch that has appeared in channel #3.

respectively) row corresponds to samples assigned to channel # $M$  (#1, respectively) and the first (last, respectively) column contains the frame #1 (# $T$ , respectively). However, a given scan of  $T$  frames of  $M$  channels actually results from the reordering of a unique data stream. This data flow contains a cyclic sequence of  $N = MT$  samples that are sequentially and periodically acquired in the  $M$  channels. These notations are gathered in Table I. The relation between a scan and the corresponding data stream is schematically illustrated in Fig. 2.

### B. Anomalies

The anomalies considered in this paper consist of additions of extra samples in the data stream associated with each scan. In the sequel of this paper, these extra samples will be called *glitches*, while the valid samples will be called *measurements*. The term *sample* will thus now stand for undifferentiated data that could be either a glitch or a measurement. These multiple and random valued insertions in a given scan result in the following:

- 1) the presence of erroneous samples in the multiplexed data stream (the glitches);
- 2) cyclic permutations of the channels after recombining the data stream by demultiplexing (due to the presence of glitches).

More precisely, assume that the successive samples are periodically assigned to channels  $M, M-1, \dots, 1$ . If a glitch appears when acquiring a measurement for channel 3, then

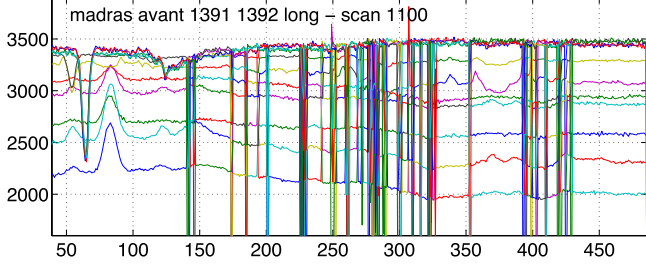


Fig. 3. Example of Earth data corrupted by several anomalies.

this channel will receive an outlier (this glitch), channel 2 will receive the measurement initially intended for channel 3, channel 1 will receive the measurement initially intended for channel 2, and the shift is propagated into the next frame and until the end of the scan. Note that, due to the  $M$ -periodic cyclic reordering of the data stream samples into the  $M$  channels, the insertion of  $M$  glitches between the sample  $\#n_1$  (belonging to the frame  $\#f_1$ ) and the sample  $\#n_2$  (belonging to the frame  $\#f_2 > \#f_1$ ) leads to an assignment of the samples after the sample  $\#n_2$  to the correct channels, subjected to a simple delay of one frame. In other words, in this case, the ideal and corrupted scans only differ by a single shift of all the frames after the frame  $\#f_1$ .

A schematic view of the corruption affecting a scan is depicted in Fig. 2, where, for conciseness, only  $M = 4$  channels have been considered and the data flow is split into  $T = 3$  successive frames.

A typical example of a scan corrupted by several glitches is depicted in Fig. 3, and typical examples of resulting MADRAS images are plotted in Figs. 5 and 6 (second panels).

### III. PROPOSED ALGORITHM

This section describes a new algorithm that is proposed for the detection and removal of glitches in the MADRAS data. The algorithm operates on each scan individually, and the technical developments that follow are related to the analysis of the corresponding data stream composed of  $N$  samples denoted as  $\mathbf{x}(N)$ . Moreover, for clarity, the set of the first  $j$  observed samples of the data stream is denoted by  $\mathbf{x}(j) \triangleq [x(1), \dots, x(j)]$ .

#### A. Trellis Design

A trellis is an oriented graph whose nodes are organized into vertical stacks that identify all the possible states of a given system at the same discrete time step. Each node is connected to at least one node from the previous time step and at least one node from the next time step. In this paper, a given discrete time step corresponds to a given sample of the data stream, and the states are defined by the number of glitches that have already been detected in the data stream preceding a given sample.

More precisely, the trellis is defined by the following characteristics.

- 1) The trellis contains  $S$  nodes per received sample  $x(j)$ , which are denoted by  $c_{k,j}$  ( $k = 0, \dots, S - 1$ ), where  $S$  is the maximum number of glitches. The node  $c_{k,j}$  corresponds to the presence of  $k$  glitches (modulo  $S - 1$ ) in the set of received samples  $\mathbf{x}(j)$ .

- 2) Each node  $c_{k,j}$  is connected to the nodes  $c_{k-1,j-1}$  and  $c_{k,j-1}$  associated with the previous sample, and  $c_{k,j+1}$  and  $c_{k+1,j+1}$  associated with the next sample. The trellis is circular in the sense that node  $c_{0,j}$  is connected to node  $c_{S-1,j-1}$  and node  $c_{S-1,j}$  is connected to node  $c_{0,j+1}$ .
- 3) The vertices connecting the nodes are referred to as branches or transitions. The branch connecting nodes  $c_{k,j-1}$  and  $c_{k,j}$  is denoted by  $v_{k,j}^0$ , and the branch connecting nodes  $c_{k-1,j-1}$  and  $c_{k,j}$  is denoted by  $v_{k,j}^1$ .
- 4) The branch  $v_{k,j}^1$  corresponds to the proposition “ $x(j)$  is a glitch” for state  $k$ , whereas the branch  $v_{k,j}^0$  corresponds to the proposition “ $x(j)$  is not a glitch” for state  $k$  (i.e.,  $x(j)$  is a valid measurement).
- 5) The branches  $v_{k,j}^0$  and  $v_{k,j}^1$  are assigned weights  $d_0(k, j)$  and  $d_1(k, j)$ , respectively. These weights can be interpreted as the (inverse) probabilities of reaching node  $c_{k,j}$  from node  $c_{k,j-1}$  or  $c_{k-1,j-1}$ , respectively, given the new observed sample  $x(j)$ . Consequently, these weights should penalize the transitions  $v_{k,j}^0$  and  $v_{k,j}^1$  according to their respective likelihoods. The choice of the weights  $d_0(k, j)$  and  $d_1(k, j)$  is discussed in Section III-C.

Based on this trellis, the most likely configuration of glitch occurrences in the set of samples  $x(1), \dots, x(j)$  can be identified by the path connecting the series of  $j$  successive nodes of minimum cumulative costs. This optimal path can be recovered by a Viterbi-like dynamic programming algorithm described in the next section.

#### B. Viterbi Algorithm

The Viterbi algorithm removes at each time instant  $j$  all branches but one reaching the states  $c_{k,j}$  such that each state  $c_{k,j}$  can be reached by only one unique path through the trellis. More specifically, the following rules are applied to sequentially prune the trellis.

- 1) At time  $j - 1$ , each node  $c_{k,j-1}$  has been assigned the cumulative weight  $D(k, j - 1)$  defined as the sum of the weights of the branches of the unique path reaching it.
- 2) At node  $c_{k,j}$ , if the sum of  $D(k, j - 1)$  and  $d_0(k, j)$  is smaller than the sum of  $D(k - 1, j - 1)$  and  $d_1(k, j)$ :
  - a) The branch  $v_{k,j}^1$  is removed from the trellis, and the branch  $v_{k,j}^0$  is kept.
  - b) The sample  $x(j)$  is accepted as a valid measurement at state  $k$ .
  - c) The sequence of  $N_{k,j}$  samples that have been accepted along the unique path reaching the node  $c_{k,j}$  is denoted by  $\hat{\mathbf{x}}_{k,j}$ , with final value  $\hat{x}_{k,j}(N_{k,j}) = x(j)$ . Note that  $N_{k,j} = j - k$  since  $k$  glitches have been detected at state  $k$  among the  $j$  already analyzed samples.
- 3) Otherwise, if the sum of  $D(k, j - 1)$  and  $d_0(k, j)$  is larger than the sum of  $D(k - 1, j - 1)$  and  $d_1(k, j)$  at node  $c_{k,j}$ :
  - a) The branch  $v_{k,j}^0$  is removed from the trellis, and the branch  $v_{k,j}^1$  is kept.

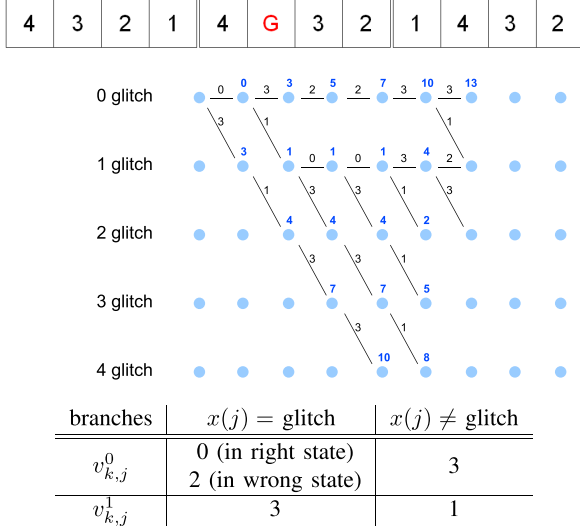


Fig. 4. Illustration of the Viterbi algorithm on a toy example with frames consisting of  $M = 4$  channels, using a trellis with  $S = 5$  states (0 to 4 glitches). The blue disks correspond to the nodes of the trellis (from left to right, the first column is the initial node before receiving data, and the last two columns have not yet received data). The blue numbers next to the nodes provide the cumulative distance along the single path through the trellis that reaches them. Here, for simplicity, artificial constant weights are used (given in the adjacent table), which correspond to the black numbers on the branches of the trellis. The algorithm has been initialized at the state  $c_{0,4}$  before receiving  $x(5)$ . The trellis is plotted during the reception of the sample  $x(10)$  nominally destined to channel #1, while one glitch has previously corrupted the data stream. State  $c_{1,10}$  needs to decide on which of the two paths reaching it (from states  $c_{0,9}$  and  $c_{1,9}$ ) has the smallest cumulative weight. In this example, the branch to be preserved is  $v_{1,10}^1$  (cumulative weight of  $D(0,9) + d_1(1,10) = 6$ ), and the branch  $v_{1,10}^0$  will hence be pruned from the trellis (cumulative weight of  $D(1,9) + d_0(1,10) = 11$ ), leading to  $D(1,10) = 11$ .

- b) The sample  $x(j)$  is identified as a glitch and is not accepted as a valid measurement by state  $k$ .
- c) The resulting sequence of  $N_{k,j}$  valid measurements  $\hat{\mathbf{x}}_{k,j}$  is hence given by  $\hat{\mathbf{x}}_{k,j} = \hat{\mathbf{x}}_{k-1,j-1}$ , with final value  $\hat{x}_{k,j}(N_{k,j}) = \hat{x}_{k-1,j-1}(N_{k-1,j-1})$ . Note also that, in this case,  $N_{k,j} = N_{k-1,j-1}$ .

After receiving the last sample  $x(N)$ , the path through the trellis with the smallest cumulative weight  $D(k, N)$ , terminating at the optimal node, denoted by  $c_{\hat{k},N}$ , is chosen. The corrected set of  $N_{\hat{k},N}$  samples is given by the vector  $\hat{\mathbf{x}}_{\hat{k},N}$ .

The design of the trellis and that of the Viterbi algorithm are illustrated in Fig. 4 on a toy example with a data stream composed of 12 samples, including 1 glitch and 11 valid measurements, distributed into  $T = 3$  frames of  $M = 4$  channels (the toy example corresponds to that shown in Fig. 2), using a trellis of  $S = 5$  possible states with constant and artificial branch weights. The trellis is depicted at the instant of reception of the tenth sample  $x(10)$ .

### C. Branch Weights

As previously stated, the choice of weights must promote the most likely transition from a node at time instant  $j - 1$  to a connected node at time instant  $j$ , given the new sample  $x(j)$ . For the application considered in this paper, the weights

assigned to the trellis branches are based on local derivatives of the samples  $x(j)$  with future and past received samples. Indeed, since the MADRAS instrument records physical parameters, the evolution between two successive valid measurements in a given channel is expected to be rather smooth while the difference in value between distinct channels is large, as illustrated in Fig. 1. More precisely, the weights are defined as follows.

*Weights  $d_0(k, j)$ :* The weight assigned to the branch  $v_{k,j}^0$  connecting nodes  $c_{k,j-1}$  and  $c_{k,j}$  is given by the square root of the absolute difference between  $x(j)$  and the last valid measurement received by state  $k$  that is supposed to belong to the same channel as  $x(j)$ , namely,  $\hat{x}_{k,j}(N_{k,j} - M + 1)$

$$d_0(k, j) = (|x(j) - \hat{x}_{k,j}(N_{k,j} - M + 1)|)^p. \quad (1)$$

As expected, this weight will be small if the new sample  $x(j)$  is not a glitch and should be assigned to the same channel as  $\hat{x}_{k,j}(N_{k,j} - M + 1)$ .

*Weights  $d_1(k, j)$ :* The weights  $d_1(k, j)$  assigned to the branches  $v_{k,j}^1$  connecting nodes  $c_{k-1,j-1}$  and  $c_{k,j}$  are chosen equal for all the transitions  $v_{k,j}^1$ ,  $k = 0, \dots, S - 1$ ; hence,  $d_1(j) = d_1(k, j)$ . The weight  $d_1(j)$  is derived from a “robustified” mean of the absolute differences between the last valid measurement  $\hat{x}_{k,j}(N_{k,j} - M + 1)$  that the state  $k$  has received and the  $N_f$  future samples  $x(j + 1), \dots, x(j + N_f)$ . More precisely

$$d_1(j) = \frac{\alpha}{S} \sum_{k=0}^{S-1} \bar{\gamma}_{k,j}, \quad \text{where } \bar{\gamma}_{k,j} = \frac{2}{N_f} \sum_{i=1}^{\frac{N_f}{2}} \gamma_{k,j}(i) \quad (2)$$

$$\gamma_{k,j} = \text{sort}_{i=1, \dots, N_f}^{\uparrow} \{(|x(j+i) - \hat{x}_{k,j}(N_{k,j} - M + 1)|)^p\}. \quad (3)$$

Only the  $N_f/2$  smallest differences associated with each state are considered in the average to discard any distance that could correspond with the presence of another but not yet detected glitch in the future samples. The parameters  $N_f, p$ , and  $\alpha$  have been chosen as  $N_f = 10$ ,  $p = 1/2$ , and  $\alpha = 1.77$  after testing different possible values and keeping those providing the best results, i.e., according to a cross-validation technique, for the application to the MADRAS data considered here.

## IV. EXPERIMENTS

### A. Simulated Data

To assess the performance of the proposed algorithm, an anomaly-free MADRAS image has been artificially corrupted by simulated glitches. The simulated glitch corruptions are designed to closely resemble those observed on corrupted MADRAS images and consist of random values (drawn in the image dynamics range) that are inserted at (groups of) random positions in the data stream. Four simulated data sets with several degrees of anomaly severities have been considered: from scenario 1, which corresponds to relatively clean data, to scenario 4, which corresponds to highly corrupted data. The numbers of glitches and corrupted samples for the four

TABLE II  
ALGORITHM PERFORMANCE AS NUMBER OF GLITCHES NOT DETECTED (TOP) AND INCORRECTLY DETECTED (BOTTOM)

Sc.	#glitches	$\Delta=0$	$\Delta=1$	$\Delta=2$	$\Delta=3$	$\Delta=4$	$\Delta=5$	$\Delta=6$	$\Delta=7$	$\Delta=8$
#1	267	1 6	0 5	0 5	0 5	0 5	0 5	0 5	0 5	0 5
#2	23970	106 109	31 34	13 17	6 9	1 4	0 3	0 3	0 3	0 3
#3	47894	192 236	128 172	78 122	44 88	21 65	12 56	4 48	2 46	0 44
#4	136378	611 853	398 640	270 512	116 358	66 308	30 272	13 255	6 248	5 247

TABLE III  
ALGORITHM PERFORMANCE AS PERCENTAGE OF CORRUPTED SAMPLES AND PSNR ERRORS BEFORE AND AFTER CORRECTION

Sc.	corrupted data samples (%)				PSNR (dB)			
	original	proposed	wFFT	TV	original	proposed	wFFT	TV
#1	7.81	0.20	100	100	23.5	44.3	33.3	36.0
#2	55.28	0.13	100	100	15.1	46.1	19.8	21.8
#3	55.07	0.14	100	100	14.3	46.1	18.9	20.3
#4	84.67	0.29	100	100	12.1	42.7	15.2	15.6

scenarios are reported in Tables II and III. Three kinds of performance analysis have been conducted. In a first analysis, reported in Section IV-A1, the corrected image is visually compared with the original uncorrupted image. Then, in Section IV-A2, a quantitative analysis is conducted on the synthetic data sets to evaluate the ability of the proposed algorithm to detect glitches. More precisely, the  $\ell_0$ -norm of the correction error is computed to measure (i.e., to count) the number of badly corrected samples and frames (i.e., pixels) between the original (uncorrupted) image and the corrected one. Let  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_P]$  denote the matrix of the  $P$  data streams  $\mathbf{x}_p = [x_p(1), \dots, x_p(N)]^T$  associated with the  $P$  contiguous scans that compose the original image. Denote the corrected samples as  $\hat{x}_p(n)$  ( $p = 1, \dots, P$ ,  $n = 1, \dots, N$ ). The proposed error measure is

$$e_0 = \|\mathbf{X}\|_0 = \sum_{p=1}^P \|\mathbf{x}_p - \hat{\mathbf{x}}_p\|_0 = \sum_{p=1}^P \sum_{n=1}^N \delta(x_p(n) - \hat{x}_p(n))$$

where the Kronecker function  $\delta(\cdot)$  is defined as

$$\delta(x) = \begin{cases} 1, & \text{if } x \neq 0; \\ 0, & \text{if } x = 0. \end{cases}$$

Note that this  $\ell_0$ -measure is particularly drastic since it penalizes all errors equally, whatever the absolute difference between the original and corrected samples. However, in the MADRAS applicative context, preserving the integrity of uncorrupted samples is crucial, which can be assessed only by this samplewise comparison before and after corruption. Finally, Section IV-A3 compares the correction performance of the proposed algorithm with those obtained by two destriping methods. In addition to the  $\ell_0$ -norm-based quality measure, the peak signal-to-noise ratio (PSNR) which relies on an  $\ell_2$ -norm reconstruction error is considered.

1) *Visual Inspection*: The visual inspection of the corrupted and corrected images has been conducted for scenarios 1 and 4 (Figs. 5 and 6, respectively). The corrected image is visually

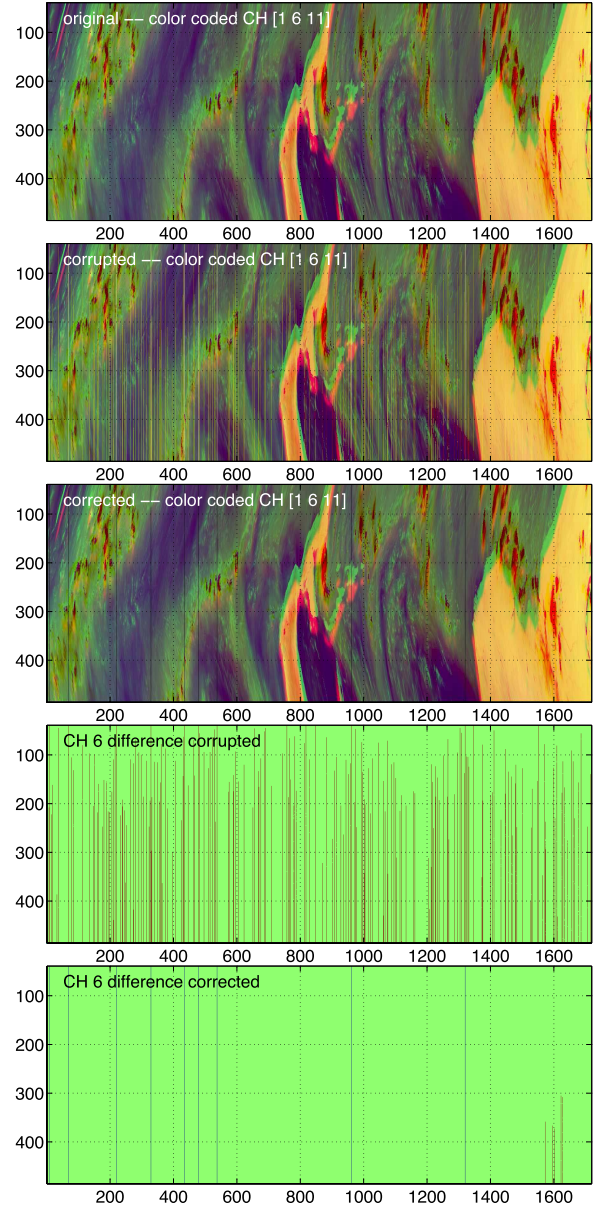


Fig. 5. Scenario 1. First panel: Original image. Second panel: Corrupted image. Third panel: Corrected image. Fourth panel:  $\ell_0$ -norm errors for corrupted image (channel 6). Fifth panel:  $\ell_0$ -norm errors for corrected image (channel 6).

compared with the original uncorrupted image: original data (top panel, synthetic color composition), corrupted data (second panel, synthetic color composition), corrected data (third panel, synthetic color composition),  $\ell_0$ -norm of the error between the original and corrupted data (fourth panel, displayed in channel 6), and  $\ell_0$ -norm of the error between the original and corrected data (bottom panel, displayed in channel 6). Note that the  $\ell_0$ -norm error takes two values (0 in green and 1 in dark red) that indicate which samples have been properly or wrongly corrected, respectively. Note also that some scans entirely appear as blue lines. This is due to another sensor anomaly that has been simulated, which is different from the one considered in this paper but easily detectable (it consists of scans with all constant values and will not be further discussed here). Moreover, note that, due to the insertion of glitches, the

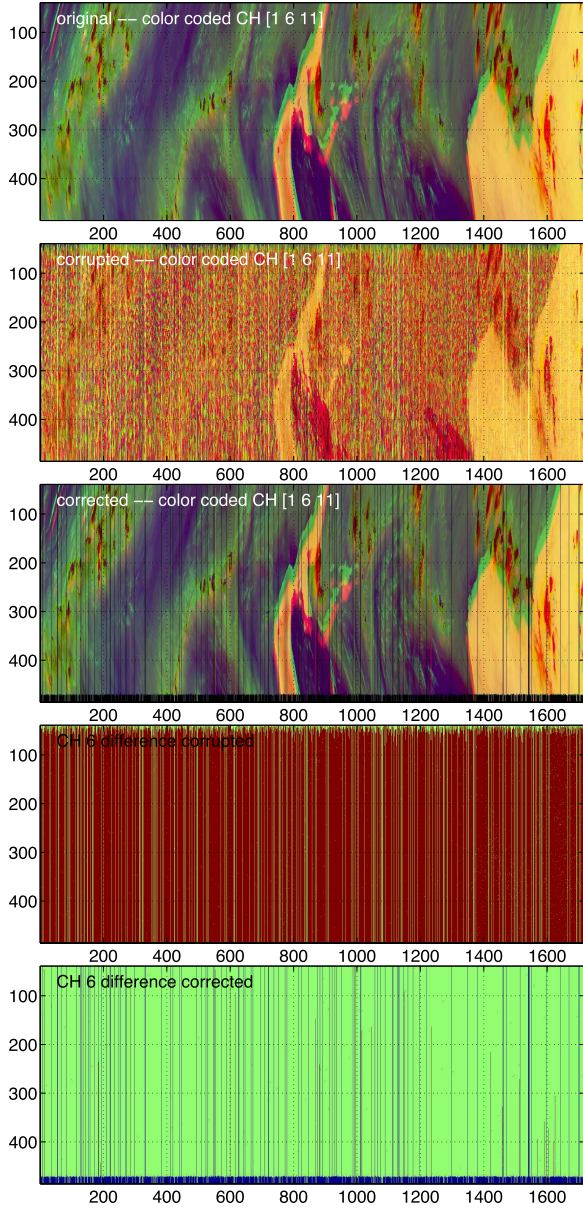


Fig. 6. Scenario 4. First panel: Original image. Second panel: Corrupted image. Third panel: Corrected image. Fourth panel:  $\ell_0$ -norm errors for corrupted image (channel 6). Fifth panel:  $\ell_0$ -norm errors for corrected image (channel 6).

corrupted scans contain less than  $N$  valid measurements. Thus, once these glitches have been detected and removed by the proposed algorithm, no additional measurement can be recovered at the end of the scan since these measurements are not contained in the corrupted data. These missing measurements appear as dark blue pixels in the corrected data of Figs. 5 and 6 (third panel).

The obtained corrected images are visually of remarkably good quality, even in the most perturbed case. In particular, for less corrupted data (scenario 1), almost all the glitches have been properly detected and removed. For this scenario, only some scans (five scans around scan #1600) appear as improperly corrected after frames/pixels #300. A thorough analysis of these scans allows two kinds of correction errors to be identified: First, for scan #1572 of scenario 1, an extra

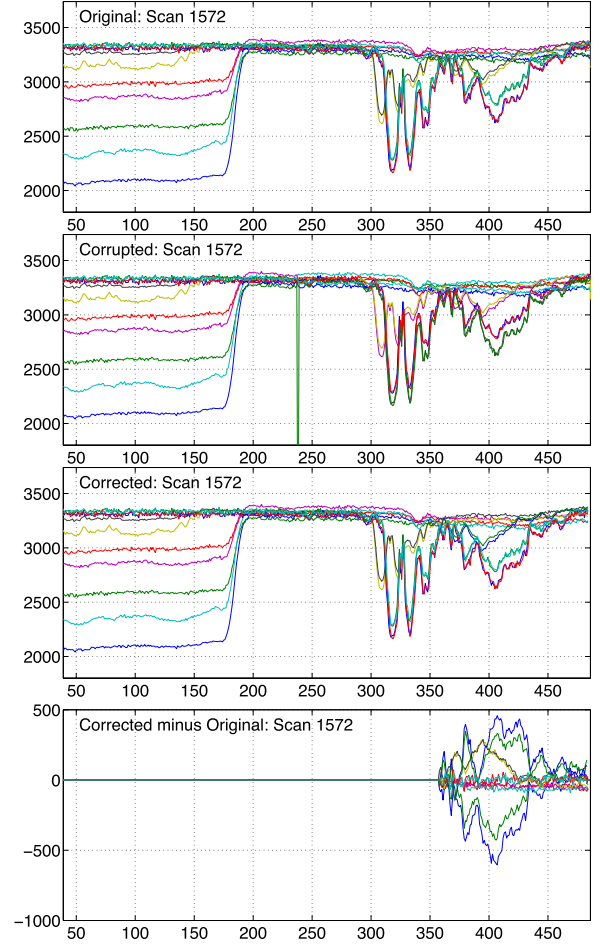


Fig. 7. Scenario 1: End of scan #1572 appears as badly corrected due to a wrongly detected glitch. Note that the corrected signals in the channels are subjected to a permutation illustrated by a change of curve colors between plots 1 and 3 after frame #358.

glitch has been detected, likely due to the fact that all channels contain similar values in the concerned frame. This leads to channel permutations of the remaining frames of the scan as illustrated by a color permutation after frame #358 in Fig. 7. Second, the algorithm sometimes removes an entire frame/pixel when several consecutive extra glitches have been detected in the data stream. This results in the deletion of a frame/pixel in all the channels. Consequently, it has a small impact on the visual inspection since spatial coherence has been preserved. This behavior is illustrated in Fig. 8. The  $\ell_0$ -norm of the error computed on scan #1627 of scenario 1 indicates that the proposed correction is completely wrong for all the frames/pixels after the frame #307. Even if the corrected scan (third plot) seems to be in very good agreement with the original data (first plot) since there is no change in curve colors, the residual error (pixelwise distance in each channel) is nonzero for all the frames/pixels after frame #307 (4th plot). A simple explanation is that the original and corrected scans only differ by an entire frame/pixel that has been incorrectly removed. Indeed, the residual error now computed with a shift of one frame/pixel is zero in all the frames/pixels for every channels (see the last plot).



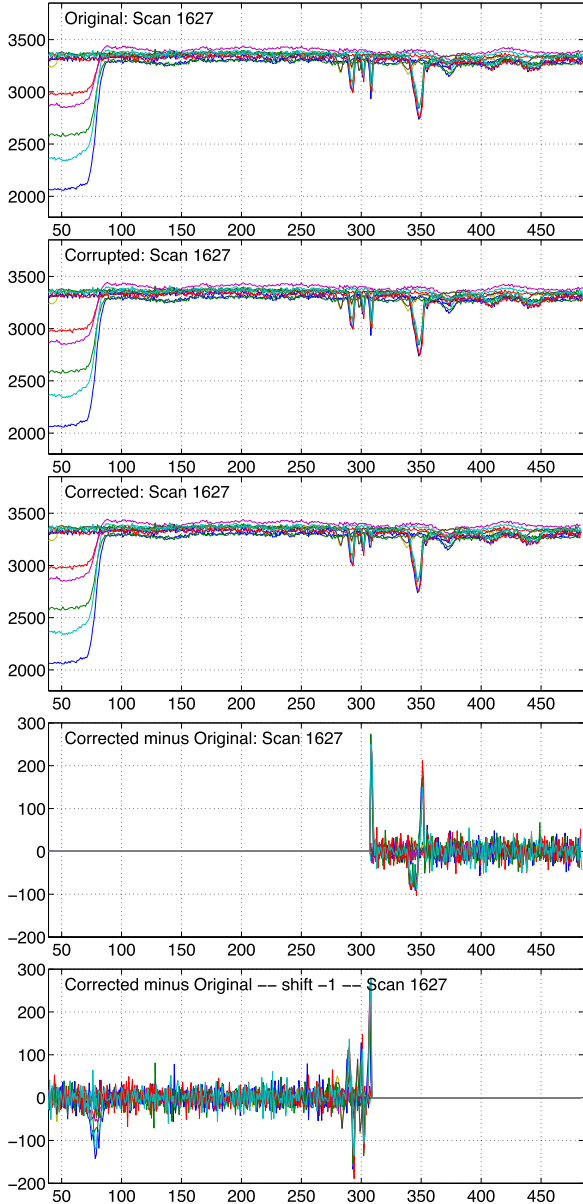


Fig. 8. Scenario 1: End of scan #1627 appears as badly corrected due to extra detected glitches, which results in the deletion of an entire frame. Note that the overall behavior of the individual signals in each channel has been preserved after correction (no channel permutation).

2) *Detection Performance Statistics*: A comprehensive quantitative performance analysis is conducted for evaluating the ability and accuracy of the proposed algorithm for detecting glitches. Two further scenarios with relatively severe corruptions (scenarios 2 and 3) are simulated in addition to those used in the previous paragraph for visual inspection. Scenario 3 contains twice as many glitches as scenario 2, yet affecting exactly the same scans (thus resulting in a similar number of corrupted samples). The total number of glitches in scans 1 to 1000 for the four scenarios is reported in Table II (second column).

Columns 3 to 11 of Table II report the number of nondetected (related to the probability of detection) and incorrectly detected glitches (related to the probability of false alarm) within a

local neighborhood ranging from  $\Delta = 0$  (exact localization) to  $\Delta = 8$  of the glitch locations. The results demonstrate that the proposed algorithm is capable of detecting and correcting a very large majority of the glitches originally present in the data stream at their precise location, and nearly all of them in a small neighborhood.

3) *Correction Performance Statistics*: The correction performance is further investigated by means of the PSNR (expressed in decibels), which is a well-admitted quality measure for image processing applications. As in the previous paragraph, the number of corrupted samples in the corrected data has also been counted using the  $\ell_0$ -norm of the reconstruction error (expressed as percentages). The proposed algorithm has been also compared with two standard destriping methods from the literature. The first one, denoted as wFFT, consists of a combined wavelet-FFT filtering [14]. The second method, denoted as TV, formulates the destriping task as a TV-regularized optimization problem [25], solved here using an alternating direction method of multipliers (ADMM). Quantitative results<sup>1</sup> are reported in Table III for all four scenarios. They demonstrate that the glitch detection and correction algorithm is highly effective and insensitive to the level of corruption, contrary to the two destriping methods. Even for scenario 4, for which, initially, almost 85% of samples were corrupted, less than 0.3% samples remain corrupted after correction with the proposed algorithm, i.e., more than 99.7% of the corrected samples are identical to the original measurements. Note also that both TV and wFFT methods slightly improve the PSNR measures when compared to the original image. However, these methods are unable to keep the correct samples unaltered. As a consequence, these techniques cannot maintain the integrity of the data since they slightly improve the (visual) quality of the corrupted images while failing to recover any of the valid measurements.

### B. Real MADRAS Data

The proposed algorithm has been applied on the real image #1391–1392 acquired by MADRAS, which is considered as being very strongly corrupted. Scans 1–1298 of the original image are plotted in Fig. 9 together with the corresponding corrected image (artificial color compositions). While the corrected image is not perfect, most of the glitches have been detected and corrected. A thorough analysis of the residual corruptions in the corrected image reveals that a large part of them are due to anomalies that are different from the one considered in this paper.

Fig. 10 summarizes as a histogram the number of detected and corrected glitches per scan. For this image, most of the scans are detected to be corrupted, with a large number of them severely corrupted. More precisely, the algorithm has detected 147 820 glitches, which roughly corresponds to 57 glitches per scan on average and a considerable number of scans with more than 100 glitches.

<sup>1</sup>A visual inspection of the results obtained by the wFFT and TV methods can be conducted from the figures in the companion technical report [26] available online.

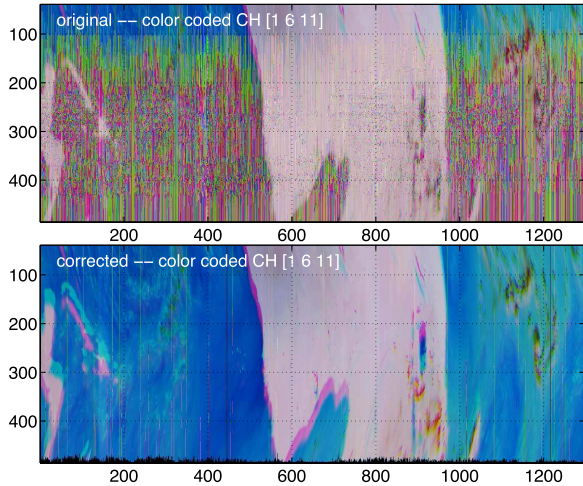


Fig. 9. Results on image #1391–1392. Top: Channel 6 of original (first 1298 scans). Bottom: Corrected image.

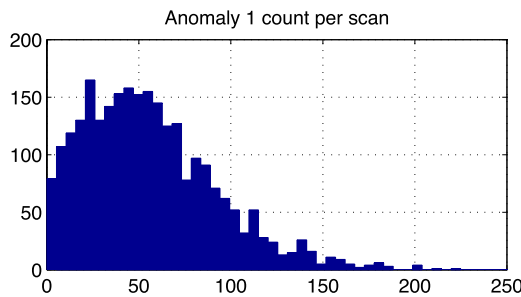


Fig. 10. Detection results for image #1391–1392: Histogram of number of detected and corrected anomalies.

## V. CONCLUSION

We proposed a simple and efficient Viterbi-like dynamic programming algorithm for the detection and removal of glitches from multiplexed data streams of multichannel measurements. Experiments conducted on simulated and real data provided by the MADRAS instrument demonstrated its excellent correction performance for both low and high corruption levels. The efficiency and performance of the algorithm were achieved by a concise modeling of the process corrupting the data. One of the specificities of the proposed correction algorithm was that no interpolation or approximation schemes were used and only valid original measurements were precisely recovered. The algorithm is operational [27] and has already been used to correct the data stream provided by the MADRAS instrument for its exploitation by the scientific community [28], [29]. Its versatility to correct data acquired by other modalities was further assessed by analyzing LANDSAT data, as reported in the companion report [26]. Future work will include other entities of geophysical multichannel data, notably hyperspectral images, and will consider different types of anomalies.

## ACKNOWLEDGMENT

The authors would like to thank Prof. H. Shen of Wuhan University for providing the alternating-direction-method-of-multipliers code used to solve the variational problem in [25].

## REFERENCES

- [1] J. P. Aguttes, J. Schiue, C. Goldstein, M. Rouze, and G. Raju, "MEGHA-TROPIQUES, a satellite for studying the water cycle and energy exchanges in the tropiques," in *Proc. IEEE IGARSS*, Honolulu, HI, USA, Jul. 2000, pp. 3042–3044.
- [2] G. Raju and S. K. Shivakumar, "MEGHA-TROPIQUES mission for climate and atmospheric applications," *IEEE Geosci. Remote Sens. Newslett.*, no. 165, pp. 33–37, Dec. 2012.
- [3] N. Karouche, C. Goldstein, A. Rosak, C. Malassingne, and G. Raju, "MEGHA-TROPIQUES satellite mission: In flight performances results," in *Proc. IEEE IGARSS*, Munich, Germany, Jul. 2012, pp. 4684–4687.
- [4] N. Karouche and G. Raju, "Megha-Tropiques satellite mission: Sensors performances," in *Proc. SPIE Sensors, Syst., Next-Gen. Satell. XIV*, R. Meynart, S. P. Neeck, and H. Shimoda, Eds., Toulouse, France, Sep. 2010, vol. 7826, p. 78260Q.
- [5] P. Chakraborty, A. Misra, T. Misra, and S. S. Rana, "Brightness temperature reconstruction using BGI," *IEEE Trans. Geosci. Remote Sens.*, vol. 24, no. 6, pp. 1768–1773, Jun. 2008.
- [6] K. Srinivasa Ramanujam and C. Balaji, "Radiative transfer simulations for the MADRAS imager of Megha-Tropiques," *J. Earth Syst. Sci.*, vol. 120, no. 1, pp. 1–17, Feb. 2011.
- [7] C. Balaji and K. Srinivasa Ramanujam, "Retrieval of rainfall from the MADRAS microwave imager of Megha-Tropiques," in *Proc. IEEE IGARSS*, Vancouver, BC, Canada, Jul. 2011, pp. 1283–1286.
- [8] F. Aires, F. Bernardo, and C. Prigent, "Atmospheric water-vapour profiling from passive microwave sounders over ocean and land. Part I: Methodology for the Megha-Tropiques mission," *Q. J. R. Meteorol. Soc.*, vol. 139, no. 673, pp. 852–864, Apr. 2013.
- [9] B. S. Gohil *et al.*, "Algorithms for retrieving geophysical parameters from the MADRAS and SAPHIR sensors of the Megha-Tropiques satellite: Indian scenario," *Q. J. R. Meteorol. Soc.*, vol. 139, no. 673, pp. 954–963, Apr. 2013.
- [10] C. Deniel and P. Veyre, *Report to COSPAR 2014*, Section: Atmosphere. Paris, France: CNES, 2014, pp. 88–93.
- [11] B. K. P. Horn and R. J. Woodham, "Destriping Landsat MSS images by histogram modification," *Comput. Graph. Image Process.*, vol. 10, no. 1, pp. 69–83, May 1979.
- [12] M. Wegener, "Destriping multiple sensor imagery by improved histogram matching," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 859–875, May 1990.
- [13] H. Shen and L. Zhang, "A MAP-based algorithm for destriping and inpainting of remotely sensed images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1492–1502, May 2009.
- [14] B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet—Fourier filtering," *Opt. Exp.*, vol. 17, no. 10, pp. 8567–8591, May 2009.
- [15] S.-w. W. Chen and J.-L. Pellequer, "DeStripe: Frequency-based algorithm for removing stripe noises from AFM images," *BMC Struct. Biol.*, vol. 11, no. 7, pp. 1–9, Feb. 2011.
- [16] H. Kurki-Suonio *et al.*, "Destriping CMB temperature and polarization maps," *Astron. Astrophys.*, vol. 506, no. 3, pp. 1511–1539, Nov. 2009.
- [17] M. Tristram *et al.*, "Iterative destriping and photometric calibration for Planck-HFI, polarized, multi-detector map-making," *Astron. Astrophys.*, vol. 534, no. 11, pp. 1–19, Oct. 2011.
- [18] R. Prévost *et al.*, "CRC-based detection algorithms for AIS signals received by satellite," *Int. J. Satell. Commun. Netw.*, vol. 31, no. 4, pp. 157–176, Jul. 2013.
- [19] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Belmont, MA, USA: Athena Scientific, 1995, ch. 2.
- [20] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 260–269, Apr. 1967.
- [21] C. Miller, B. R. Hunt, M. A. Neifeld, and M. W. Marcellin, "Binary image reconstruction via 2-D Viterbi search," in *Proc. IEEE ICIP*, pp. 181–184, vol. 1, 1997.
- [22] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.
- [23] C. Miller, B. R. Hunt, M. Marcellin, and M. A. Neifeld, "Image restoration with the Viterbi algorithm," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 17, no. 2, pp. 265–275, Feb. 2000.
- [24] B. Aminov, O. Nichtern, and S. R. Rotman, "Spatial and temporal point tracking in real hyperspectral images," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, p. 30, Jul. 2011.
- [25] M. Bouali and S. Ladjal, "Toward optimal destriping of MODIS data using a unidirectional variational model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2924–2935, Aug. 2011.

- [26] H. Wendt *et al.*, “Detection and correction of glitches in a multiplexed multi-channel data stream—Application to the MADRAS instrument—Complementary results and supporting materials,” Univ. Toulouse, IRIT/INP-ENSEEIH, Toulouse, France, Tech. Rep., Oct. 2015. [Online]. Available: [http://dobigeon.perso.enseeiht.fr/papers/Wendt\\_TechReport\\_2015.pdf](http://dobigeon.perso.enseeiht.fr/papers/Wendt_TechReport_2015.pdf)
- [27] H. Wendt, N. Dobigeon, and J.-Y. Tourneret, “*Procédé de détection et/ou de correction automatique deerreurs dans un flux de données multiplexées*,” FR Patent 3 007 232, Dec. 19, 2014.
- [28] R. G. Sivira, H. Brogniez, C. Mallet, and Y. Oussar, “A relative humidity profile retrieval from Megha-Tropiques observations without explicit thermodynamical constraints,” *Atmos. Meas. Tech. Discuss.*, vol. 7, no. 9, pp. 8983–9023, Dec. 2014.
- [29] E. Defer, V. S. Galligani, C. Prigent, and C. Jimenez, “First observations of polarized scattering over ice clouds at close-to-millimeter wavelengths (157 GHz) with MADRAS on board the Megha-Tropiques mission,” *J. Geophys. Res., Atmos.*, vol. 119, no. 21, pp. 12 301–12 316, Nov. 2014.



**Herwig Wendt** (M’13) received the M.S. degree in electrical engineering and telecommunications from the Vienna University of Technology, Vienna, Austria, in 2005. In September 2008, he received the Ph.D. degree in physics and signal processing from École Normale Supérieure de Lyon, Lyon, France.

From October 2008 to December 2011, he was a Postdoctoral Research Associate with the Department of Mathematics and with the Geomathematical Imaging Group, Purdue University, West Lafayette, IN, USA. Since 2012, he has been a tenured Research Scientist with the Centre National de Recherche Scientifique and with the Signal and Communications Group of the IRIT Laboratory, University of Toulouse, Toulouse, France. He is also an Affiliated Faculty Member of the TeSA Laboratory, Toulouse.



**Nicolas Dobigeon** (S’05–M’08–SM’13) was born in Angoulême, France, in 1981. He received the Engineering degree in electrical engineering from École Nationale Supérieure d’Électronique, d’Électrotechnique, d’Informatique, d’Hydraulique et des Télécommunications, Toulouse, France, and the M.Sc. degree in signal processing from the INP Toulouse, both in 2004, and the Ph.D. degree and Habilitation Diriger des Recherches in signal processing from INP Toulouse in 2007 and 2012, respectively.

From 2007 to 2008, he was a Postdoctoral Research Associate with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA. Since 2008, he has been with INP Toulouse, University of Toulouse, Toulouse, where he is currently an Associate Professor. He conducts his research within the Signal and Communications Group, IRIT Laboratory, and he is also an Affiliated Faculty Member of the TeSA Laboratory. His recent research activities have been focused on statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing, biomedical imaging, and genomics.



**Jean-Yves Tourneret** (SM’08) received the Ungénieur degree in electrical engineering from the École Nationale Supérieure d’Électronique, d’Électrotechnique, d’Informatique, d’Hydraulique et des Télécommunications (ENSEEIH) de Toulouse, France, in 1989 and the Ph.D. degree from the National Polytechnic Institute from Toulouse in 1992.

He is currently a Professor in the University of Toulouse (ENSEEIH) and a member of the IRIT laboratory [UMR 5505 of the Centre National de Recherche Scientifique]. His research activities are centered around statistical signal and image processing, with a particular interest in Bayesian and Markov chain Monte Carlo methods.

Dr. Tourneret has been involved in the organization of several conferences, including the European conference on signal processing EUSIPCO’02 (Program Chair), the international conference ICASSP’06 (plenaries), the statistical signal processing workshop SSP’12 (international liaisons), the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2013 (local arrangements), the statistical signal processing workshop SSP’2014 (special sessions), and the workshop on machine learning for signal processing MLSP’2014 (special sessions). He has been the General Chair of the CIMI workshop on optimization and statistics in image processing held in Toulouse in 2013 (with F. Malgouyres and D. Kouamé) and of the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2015 (with P. Djuric). He has been a member of different technical committees, including the Signal Processing Theory and Methods committee of the IEEE Signal Processing Society (2001–2007, 2010–present). He has been serving as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2008–2011, 2015–present) and for the *EURASIP Journal on Signal Processing* (2013–present).



**Mathieu Albinet** was born in Rodez, France, in 1980. He received the Engineering degree in electronics and signal processing from the École Nationale Supérieure d’Électronique, d’Électrotechnique, d’Informatique, d’Hydraulique et des Télécommunications de Toulouse, France, in 2003.

He has been employed with the French Space Agency (CNES) since 2005, where his main area of work was onboard payload data processing. He is currently the Head of the “On-Board Data Systems Office.”

**Christophe Goldstein**, photograph and biography not available at the time of publication.

**Nadia Karouche**, photograph and biography not available at the time of publication.