Appendix 1. Details on the MAD criterion

A monochromatic image such as those produced by JIRAM can be considered as a single-valued function *f* of the two *integer* variables x and y representing the pixel coordinates, i.e.: the signal measured in the pixel with coordinates x_0 and y_0 is given by $f(x_0, y_0)$.

Let's consider a pair of images *f* and *p*, being *p* derived from *f* by some kind of motion in the observed scene. For each point (x_0, y_0) in *f* and each possible bi-dimensional shift vector (dx, dy) we define the mean absolute distortion (MAD) as

$$MAD(x_0, y_0, dx, dy) = \sum_{i=-\Delta x}^{+\Delta x} \sum_{j=-\Delta y}^{+\Delta y} |f(x_0+i, y_0+y) - p(x_0+i+dx, y_0+y+dy)| \quad [A1.1]$$

Here, $(2\cdot\Delta x+1)$ and $(2\cdot\Delta y+1)$ are the sides of a square neighborhood of (x_0, y_0) from the first image *f* that is translated over the second image *p* by varying the *integer* values of shift (dx, dy). The algorithm defines the motion of the point (x_0, y_0) from *f* to *p* as the shift vector (dx, dy) that minimize the MAD function. This criterion has largely been exploited in motion estimate, especially for the purpose of video signal compression [e.g. Li et al., 1994].

In practical terms, once (x_0, y_0) is fixed, our code performs $(2\cdot\max|dx|+1)\cdot(2\cdot\max|dy|+1)$ independent estimates of (A1.1) for different (dx, dy) values, selecting eventually the one providing the smaller MAD value. The estimates of best (dx, dy) for different points of *f* (i.e.:different x_0, y_0) are performed independently for each point.

Notably, the usage of the MAD criterion implicitly requires that *f* and *p* both refer to the very same bi-dimensional spatial grid. In our case, original JIRAM images were re-interpolated over an x,y plane on the basis of geographic coordinates of pixel centres (System III coordinates). We adopted an azimuthal equidistant projection centred over the pole of interest with a sampling step of 0.01 degrees of latitude. This value corresponds to about 11.6 km and final images are consequently oversampled of a factor between 1 (no oversampling) and 4 with respect to the original data. An oversampling up to a factor 4 is “typically” used when MAD criterion is adopted method (Gonzales & Woods, [2008]).

The size of the region to be explored in the second image *p* (i.e.: range of values for |dx| and |dy|) depends upon the expected displacement magnitude. With the re-sampling described above, a maximum shift of 20 points from nominal position along the vertical and horizontal directions (in both senses) would correspond to wind speeds up to 240 m/s for images pair acquired 16 minutes apart (as in the case of north pole observations listed in Table 1), to be compared against the maximum speeds of mid-latitude jets of 150m/s [Porco et al., 2003]. A pair with a longer time gap between members requires correspondingly larger regions to be explored over *p*, with computational times being proportional to $\max|dx|\cdot\max|dy|$. The values of Δx and Δy (i.e.: the size of neighbourhood for the points of the first image) is set by trial-and-error, but it should be of the order of maximum values for |dx| and |dy|: in our case a side of 20 points was again deemed adequate.

The MAD criterion allows one to estimate the motion from *f* to *p* for each individual point of *f*, and therefore to obtain full speed fields over the image. The method considers only rigid translations as

possible displacements for the neighborhoods of the *f* points and therefore possible rotations are not included. This limitation requires that time interval between the acquisitions of *f* and *p* shall remains small with respect to any expected rotation period in the observed scene. The sizes of Δx and Δy shall also remain reasonably small, since any rotation occurring at scales smaller than the size of the search neighborhood (but still greater than the sampling step) from the first image would hamper the effectiveness of MAD search.

Albeit from a formal perspective the use of MAD criterion can provide a speed vector for each point of *f*, in the practical implementation one shall consider how the intrinsic motion of the scene can bring the neighborhood outside *p*. For implementation simplicity in our code *f* and *p* were cut to the entire available overlap region and speed field were not computed for the points of *f* lying within $\max(\Delta x+|dx|, \Delta y+|dy|)$ from image borders.

Alternatively to MAD minimization, another approach may consist in maximizing the correlation between the neighborhood and the second image. This method however was eventually discarded since direct tests demonstrated its proneness to considerable systematic errors once applied to this specific subset of JIRAM data. The reasons for this behavior are beyond the scope of this paper, but are not surprising when dealing with turbulent motions where fractal-like structures are often observed.

To our knowledge this work represents the first instance of MAD criterion usage in cloud tracking from satellite images. The results of our implementation on a test image pair were firstly compared against the values provided by the software JPIV (https://www.jpiv.vennemann-online.de/settings.html), that works on the principle of correlation maximization and based on the concepts presented by Willert and Gharib [1991]. The comparison is presented in Fig. A4. In assessing the comparison, it is noteworthy to mention that while our implementation provides shifts in terms of integer values of pixels, JPIV provides shift values on much finer scale, based on the fit of the shape of the correlation function. Considering the differences in total magnitude of shift detected by the two methods, our implementation provides usually slightly higher values (mean difference = 0.75 pixels), while the random errors are found to be around 0.86 pixels.
Secondly, an independent validation was performed through a numerical simulation. Pixels from a test image are moved according a given arbitrary bi-dimensional (i.e.: along the x and y directions) shift field, to create a simulated time-projected image. Our algorithm is then applied to the pair formed by the original and the time-projected image. The retrieved shift values are eventually compared against those imposed initially, to study differences. Systematic differences are found to be negligible, while random differences are about 0.56 pixels.
For velocimetry algorithms based on the evaluation of a cost function over a finite size neighborhood (such as the MAD function given in A1.1), a smoothing of retrieved speed fields is implicitly applied with a size in the order of adopted neightborhood. Main criteria to identify suspicious results is therefore represented by sharp spatial fluctuations of retrieved speed, uncorrelated to obvious image features. In the cases described in this work, our algorithm found such instances only in the vicinity of image borders, where errors are expected to occur when the considered neighborhood has translated outside the overlap region of image pair.
Examples of cloud feature tracking derived by our algorithm are provided as Supporting Information.

Appendix 2. Mitigation of pointing errors

Systematic speed errors are more difficult to assess and may arise from a variety of sources. Notably, JIRAM pointing determination and – consequently – accuracy of geometric parameters of the observations (pixel coordinates) are limited by the precision to which the Juno spacecraft can determine its own phase over the rotation around the spin axis. Briefly, this translates into errors on estimated pixel spatial positions up to 240 μrad (i.e.: the size if 1 pixel) for the original JIRAM images and into errors up to 4 sampling points for the re-interpolated data (once the sampling size of 11.6 km and the maximum allowed spatial resolution of 40 km are considered).

In the estimate of speed fields from JIRAM sequences shall take into account a few specific issues.

1. Different images composing a sequence are not acquired simultaneously. Inside a given sequence, each image is acquired during a different Juno rotation and therefore there is a minimum time interval of 30 second between different acquisition (Figure A5)
2. The systematic pointing error (difference between estimated and actual pixel positions) is not merely unknown but – most important - different in different images.

These two facts preclude the simplistic approach of creating two mosaics of interpolated images (a mosaic from each sequence) and apply the MAD search over these two mosaics. The usage of a mosaic of images is prone to the ambiguity in the definition of acquisition time (that is different in different regions) as well as to *relative* unknown displacements between parts of the mosaic originating from different individual images.

Conversely, we decided to consider each possible pair of JIRAM images with the first image from the first sequence and the second image from the second sequence. The MAD search procedure was applied to each of these possible image pairs, to derive the speed field in the overlap region. For simplicity, the MAD search was limited to the cases where re-interpolated pairs of images have an overlap exceeding 10^4 sampling points.

Considering now a specific pair of images as described above (first image from first sequence, second image from second sequence), the speed field computed in the overlap region is affected by a systematic speed error (δv_x, δv_y), uniform over the entire overlap region, derived from the unknown (and different) systematic pointing errors - (δx_1, δy_1) and (δx_2, δy_2) - affecting the two images.

$$\delta v_x = \frac{(\delta x_1 + \delta x_2)\cdot s_x}{\Delta t} \qquad \delta v_y = \frac{(\delta y_1 + \delta y_2)\cdot s_y}{\Delta t} \qquad \text{[A2.1]}$$

Here s_x and s_y are the amplitudes of sampling steps along the x and y directions (11.6 km in both cases in our analysis) and Δt the time elapsed between the acquisitions of the two images of the pair, a quantity that – contrarily to mosaics - is defined unambiguously for individual images. For simplicity, we denote here as 'wind tile' the individual speed field derived from the analysis of a given pair of images as described above (first image from first sequence, second image from second sequence). Spatial coverage of each wind tile correspond to the overlap area between the two

images.

From now on we do not consider any longer the images or their overlap regions, but only wind tiles (speed fields over given areas). The mosaicking of wind tiles retains the ambiguity due to the unknown speed errors formally given by (2) that affects in different amount each wind tile. A simple stacking of wind tiles makes this issue well evident, with clear discontinuities of speed amplitude at tile borders.

However, spatial coverage of different wind tiles shows often large areas of overlap. In these cases it is possible to consider an arbitrary 'start tile' and to add - one by one - all available tiles exploiting the condition that in every point of overlap regions speed amplitudes must be identical in the two overlapping tiles (within the approximate random errors described in section 3). In order to met this condition, we assume initially as 'correct' the speeds of the start tile and we look for an additional speed vector that, once added to every point of the tile to be added, provides the best match of speed amplitudes in the overlap regions. This additional speed vector is the same for every point of the tile to be added and is eventually applied also to the points outside the overlap region with the start tile. Proceeding with all available tiles, the final 'wind mosaic' is still affected by an unknown systematic (i.e.: uniform over its spatial coverage) *absolute* errors on speed (namely, those affecting the 'start tile'), but virtually no *relative* errors between its different parts.

The method is however subject to some limitations. Each addition of a new tile to the wind mosaic requires the determination of a separate corrective speed vector. The precision at which this vector can be determined is limited by several factors, including the need to perform the search of the optimal correction over a grid of finite size (¼ of sampling step in our case) and the slightly different time gaps used to compute different wind tiles. This latter factor is strictly related to the fact that shifts are determined by the MAD algorithm in terms of integer sampling points: different time gaps lead therefore to slight different sets of possible speed values. In general terms, the corrective components applied during mosaicking are capable to reduce the amplitudes of speed discontinuities observed at tile borders to at least 25% of the original values.

All the speed fields discussed in this paper are actually wind mosaics creating according the procedure described above. In all cases, the selected start wind tile was the one exhibiting the most complete coverage on the central vortex of each pole.

Notably, we attempted also an alternative approach consisting in creating mosaics of images from a given sequence where, starting from an initial (re-interpolated) image, we added one-by-one other images introducing, at each addition, a corrective bi-dimensional shift in order to minimize differences in the overlap regions. These corrected image mosaics were eventual fed to the MAD algorithm. Discontinuities at borders in speed field at tile borders were found to be usually of comparable or *greater* magnitude that achieved without any correction, despite being computed for exact elapsed times.

The removal of relative errors between different parts allows one to compare distant spatial regions can not be observed in an individual JIRAM image at adequate spatial resolution. In the present study, this is particularly important for the central regions of different polar vortices. The procedure of mosaicking of 'wind tiles' described above was possible over the entire width of the investigated areas only for the N1, S1 and S3 fields described in table 2. In other cases, gaps exist between

different section of the considered mosaic.

Additional Figures

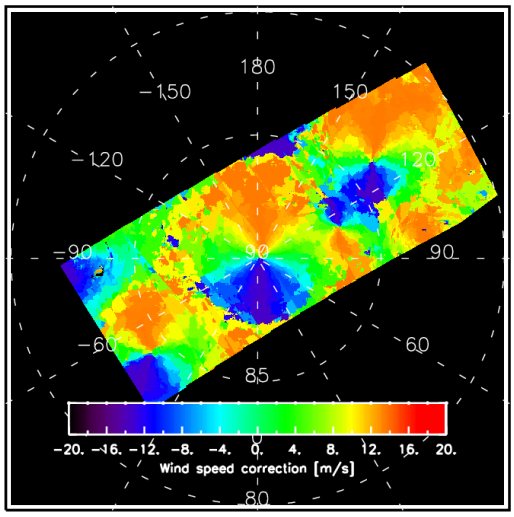
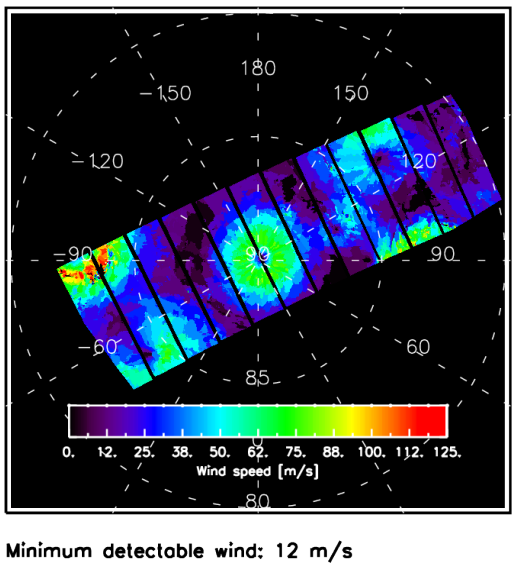
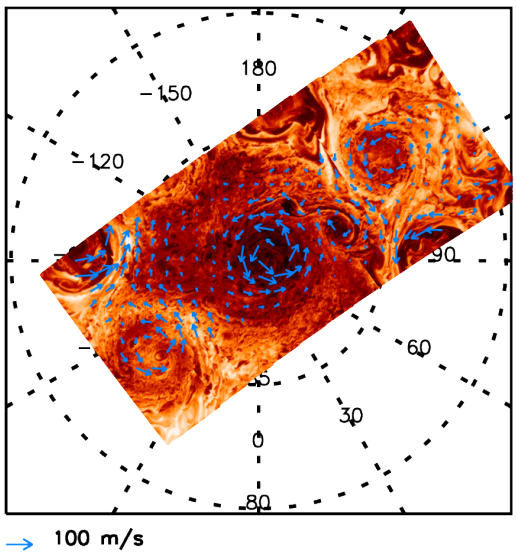


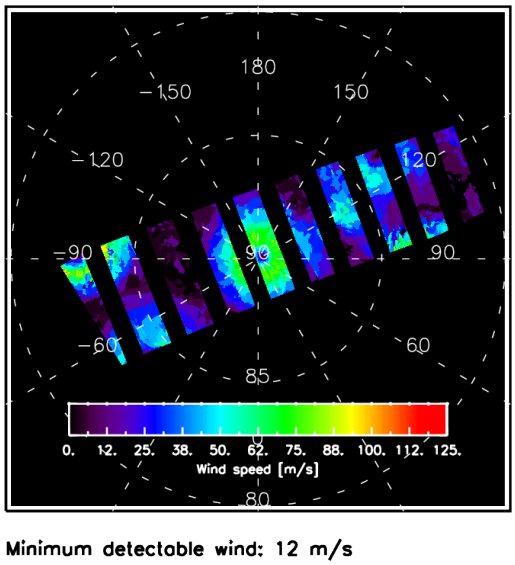
Figure A1: Net effect of symmetrization procedure on wind field once applied to case N1 of table 1. This figure represent the difference between panels 2a and 4a.



a

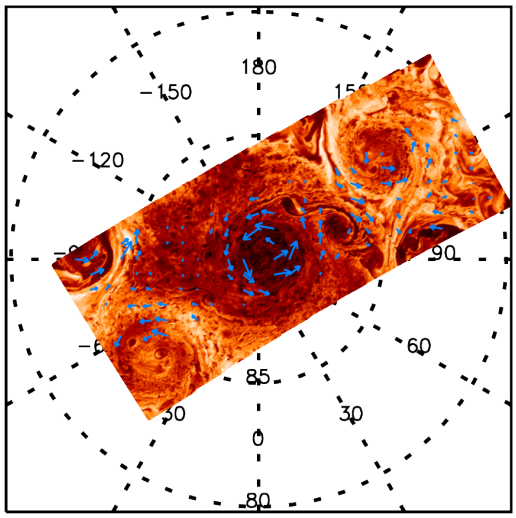


b



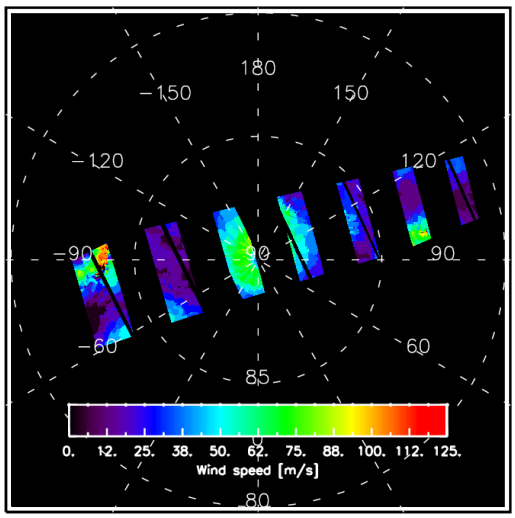
Minimum detectable wind: 12 m/s

c



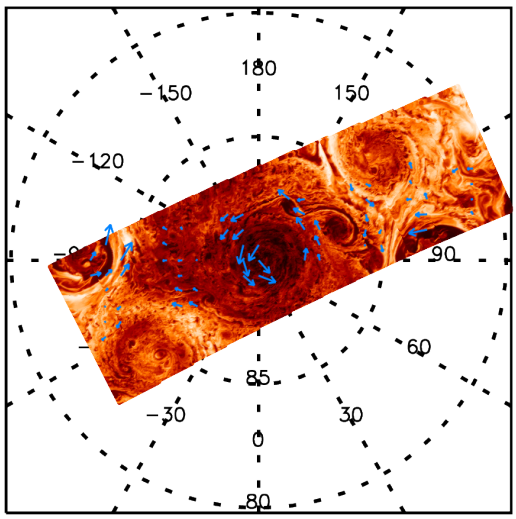
→ 100 m/s

d



Minimum detectable wind: 12 m/s

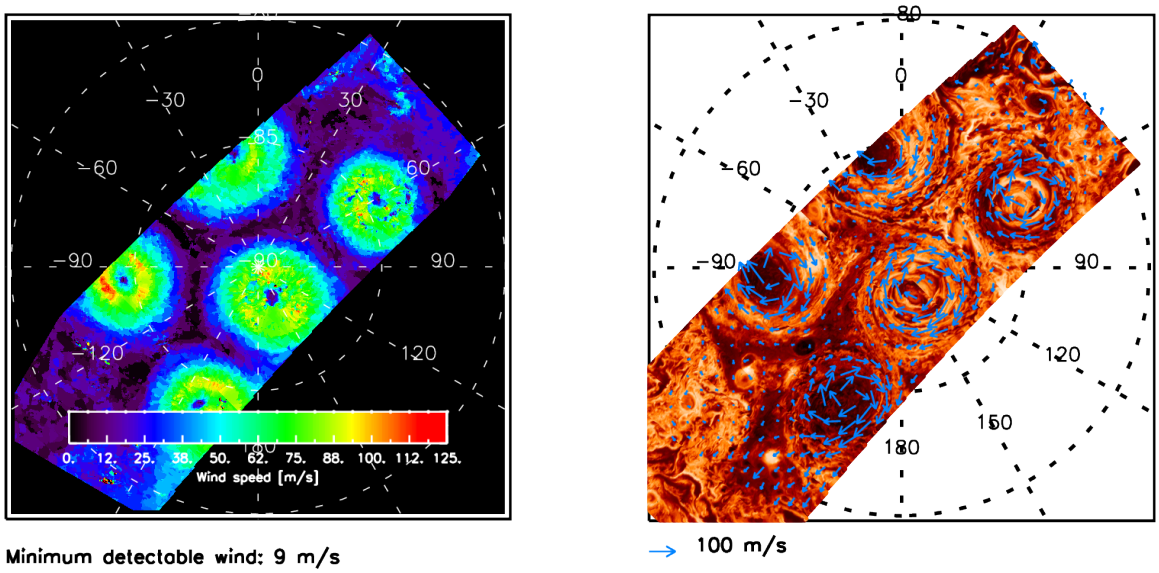
e



→ 100 m/s

f

Figure A2: Wind velocity amplitudes and directions over the north pole, as corrected to achieve maximum azimuthal symmetry of wind speed around the centre of NPC. Case N2: panels a and b; Case N3: panels c and d; Case N4: panels e and f.



a b

Figure A3: Wind velocity amplitudes and directions over the south pole, as corrected to achieve maximum azimuthal symmetry of wind speed around the centre of SPC. Case S1: panels a and b.

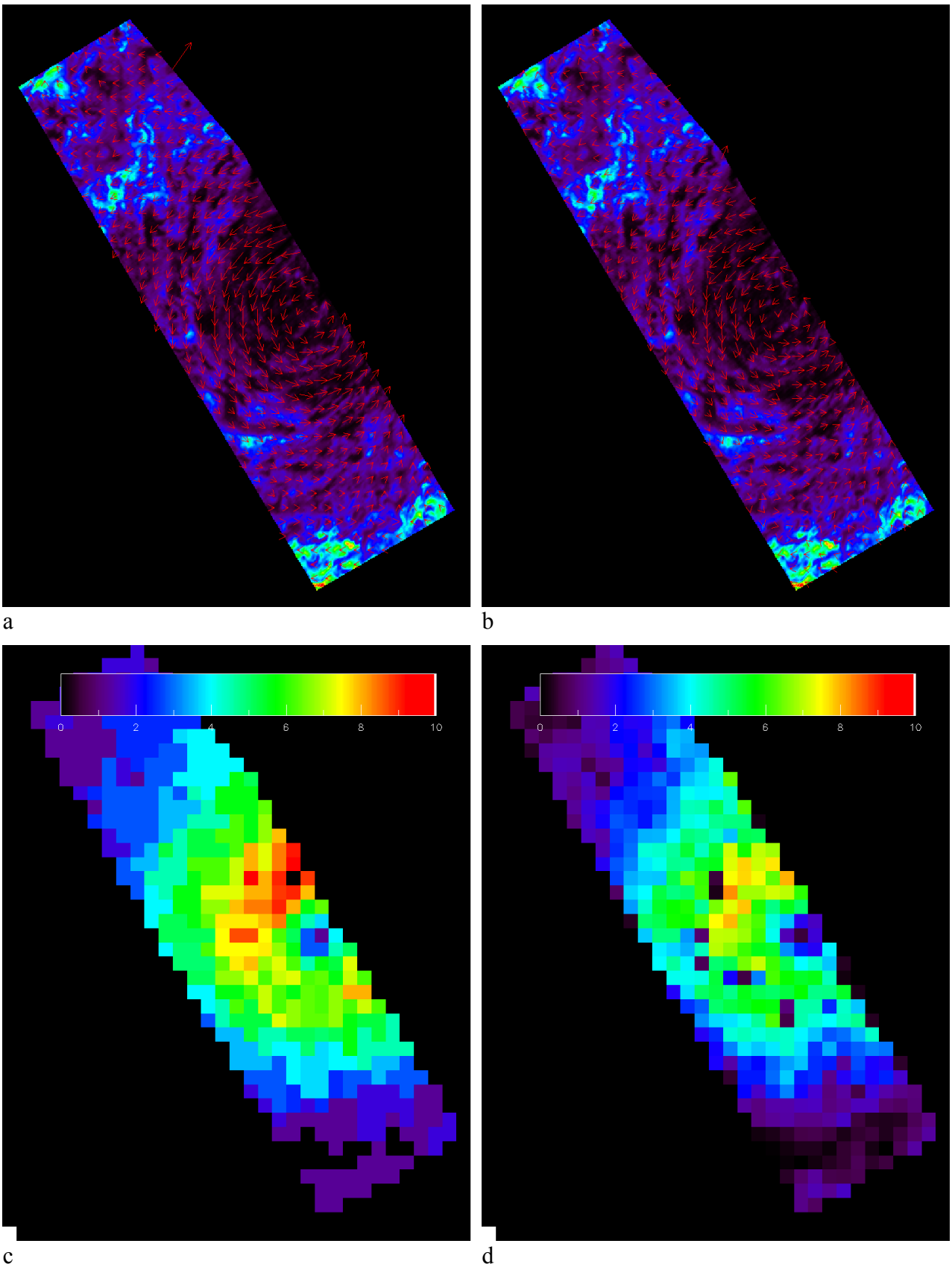


Fig. A4. Comparison between shifts retrieved from a test image pair by our implementation of MAD algorithm and the JPIV velocimetry code. a. MAD shift vectors b. JPIV shift vectors c.MAD shift magnitudes d. JPIV shift magnitudes. Shift magnitudes are given in terms of number of pixels.

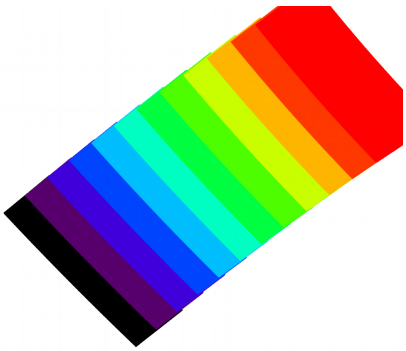


Figure A5: actual scheme of individual JIRAM images used to build the background image of figures 2b and 3b. Each colour corresponds to a different image inside the sequence 170202_114231. Acquisition times of images were spaced in intervals of 30 seconds.