Marked-up manuscript of "Toward the assimilation of images" by F.-X. Le Dimet et al.

F.-X. Le Dimet and I. Souopgui and O. Titaud and V. Shutyaev

December 9, 2014

This document combines the author's responses to reviews with the marked-up revised manuscript.

Author's reponses to comments from referees on "Toward the assimilation of images" by F.-X. Le Dimet et al.

F.-X. Le Dimet and I. Souopgui and O. Titaud and V. Shutyaev

December 9, 2014

Anonymous Referee #1

1 Comment

1.1 Comment from Referee

The paper constitutes a review of state of the art of data assimilation of images. It is well written by the leading expert in the field and addresses itself to a broad audience.

I would appreciate if the topics of data assimilation in medicine namely medical Imaging that investigates processes in the brain by techniques such as MRI, EEG, MEG and many more could be briefly addressed. Usually dynamical models based on finite element discretisation approaches are coupled with data by inversion and data assimilation.

The same relates to image reconstruction from noisy data that is an important inverse problem. where Electrical Impedance Tomography (EIT)can be used.

Otherwise a very good review that should be published.

See for instance : D. Chapelle, M. Fragu, V. Mallet, P. Moireau: Fundamental principles of data assimilation underlying the Verdandi library: applications to biophysical model personalization within euHeart. Medical & Biological Engineering & Computing Vol. 51 (2013) 1221-1233

1.2 Author's response

The Editor gave a direction to answer this comment based on the scope of the Journal and we thank him for his answer. Since this journal is focused on geophysics, the text does not address data assimilation in other domains like medecine. However, the introduction mentions that the thechnique can be applied to other problems. As an illustration of the application to other problems, we added the reference suggested by the referee in the introduction.

1.3 Author's changes in manuscript

See the introduction of the revised manuscript for Author's cganges.

Referee #2 J. Ma

1 Comment

1.1 Comment from Referee

Data assimilation is the science of coupling information coming from different sources: model, statistics and observations. Data assimilation has been successfully applied to meteorology and oceanography. It was also used for fields such as agronomy, economy, medicine, and oil/gas reservoir description in exploration geophysics (see attached references). Variational method proposed by Le Dimet and Talagrand in 1986 plays important role in the field of data assimilation. In last decade, the researchers pay attentions to the use of quantitative information rather than qualitative analysis from the observed image sequences. In this paper, the authors made a nice review on the variational data assimilation with the use of quantitative image information, and described several possibilities for such assimilation and identify associated difficulties. The paper is well organized and written.

Eq. (26) can be defined as scale-dependent hard thresholding. The thresholding rules in 1 and 3 in the same page are actually the special cases of Eq. (26). The motivation of the thresholding is to extract the edge structures of images. The edge structures and motion vectors are extracted simultaneously by the curvelet transform, in the reference "J. Ma, A. Antoniadis, F.-X. Le Dimet, Curvelets-based multiscale de- tection and tracking for geophysical fluids, IEEE Transactions on Geoscience and Remote Sensing, 2006, 44 (12), 3626-3638". The description could be added in the final version. The paper is recommended to be published. 1. P. Chen, Full-wave seismic data assimilation: theoretical background and recent Advances, Pure and Applied Geophysics, 168(10):1527-1552. 2. Y. Dong, Y. Gu, D. Oliver, Sequential assimilation of 4D seismic data for reservoir description using the ensemble Kalman filter, J. Petroleum Science and Engineering, 2006, 53, 83-99

1.2 Author's response

We agree with the Referee that the rule in 1 is the special case of Eq. (26) where the threshold is set to the same value for all scales. The rule in 3 is the special case of Eq. (26) where the threshold is set to zero for the coarse scale and constant for all other scales.

1.3 Author's changes in manuscript

See the paragraph 6.1.5 of the revised paper.

Anonymous Referee #3

1 Comment 1

1.1 Comment from Referee

This paper is a very interesting review paper on the problem of assimilation of image data. It is well written and nicely documented. I particularly appreciated the work of formalization of this new type of data in order to incorporate it in the traditional formalism of 4D VAR, that is used for conventional data. This is performed in section 5 and illustrated by numerical experiments in section 6. The paper is written in a nice way, so that reading is really pleasant. My comments, both on the theoretical aspects as well as on the applications to image assimilation, is that this paper is worth being published. I suggest a certain number of improvements so that the paper becomes excellent :

I did not understand the order of the figures. Usually the figures have numbers that correspond to the appearance in the paper. This is not the case here. For instance figure 6 appears very early in the text, before some figures that have smaller numbers. Why ? I think that it is better to change the numbers of the figures to respect the general rule of appearance in the text.

1.2 Author's response

This is an important point for the reader and we thank the Referee for his attention. We will address this problem of figure ordering in the revision.

1.3 Author's changes in manuscript

The figures are now reordered as follows:

- 6 becomes 3
- 3 becomes 4
- 7 becomes 5
- 4 becomes 6

- 5 becomes 7
- No change for other figures

2 Comment 2

2.1 Comment from Referee

At the end of paragraph 3.1, there is a comment on the two main types of data assimilation methods (Kalman methods and variational ones). Is it true that Kalman methods are not implemented in operational centers ? It seems to me that there is a kind of non-objectivity from the authors.

2.2 Author's response

This question is of crucial importance and we thank the Referee for raising it. Our assertion referred to the traditional Kalman Filter and should not be associated with all the Kalman Filter approaches as it is the case in this version of the paper. The Ensemble Kalman Filter (Evensen, 1994) method includes an elegant definition of the approximation of the covariance matrix and its practical use as long as the infrastructures allow to run the ensemble forecast.

We also thank the Editor for his comment on this question, as well as O. Talagrand for pointing to various sources to answer this comment.

2.3 Author's changes in manuscript

The revised text makes it clear that the problem with the evolution of the covariance matrix is relevant only for traditional Kalman Filter. It also makes it clear that the ensemble approach solves that problem; see the section 3.1.

3 Comment 3

3.1 Comment from Referee

The threshholding procedure in 6.1.5 could be more clearly explained for non specialists of curvelets.

3.2 Author's response

The Referee raises here an important question of the difficulty associated with the topic of curvelet in general. Paragraph 6.1.2 has a short description of the curvelet decomposition. To help the non specialist of curvelet understanding the thresholding procedure, and based on the parallel comment of Referee #2 we rewrote the paragraph.

3.3 Author's changes in manuscript

See our response to the comment of Referee #2.

4 Comment 4

4.1 Comment from Referee

The paragraph 6.2 is not written as well as the other ones. I suggest to rewrite it in a style that is consistent with the rest of the paper and that is easier to understand.

4.2 Author's response

The remark of the Referee is capital for this paragraph that focus on the topic of Lyapunov exponents as observation operator for images in Data Assimilation. We rewrote the paragraph to get more uniformity.

4.3 Author's changes in manuscript

The paragraph 6.2 is rewritten, see the revised manuscript.

W. Hsieh (Editor)

1 Comment 1

1.1 Comment from the Editor

I am pleased that all three reviews are generally positive about this manuscript, and the requested revisions are not major.

in regard to Review #1, since this journal is focused on geophysics, addressing data assimilation in medicine will be interesting but not essential for this review paper.

1.2 Author's response

We thank the Editor's for answering the comment from Referee #1;

2 Comment 2

2.1 Comment from the Editor

Review #3, point 2) wondered if the authors were non-objective in claiming Kalman filters were not used in operational centers. In defense of the authors, I believe their claim is indeed true. Maybe the authors can give an order of magnitude estimate on the computing power needed to do data assimilation with a Kalman filter as compared to using variational assimilation for a typical operation model, which should demonstrate that the Kalman filter is far too expensive to be practical.

2.2 Author's response

We thank the Editor's for giving indications to answer the comment from Referee #3.

3 Comment 3

3.1 Comment from the Editor

It seems Sects. 6.1.5 and 6.2 need to be made clearer, based on Reviews #3 and #2.

3.2 Author's response

We rewrote section 6.1.5 and section 6.2, see our response to Reviews #2 and #3.

O. Talagrand

4 Comment

4.1 Comment from O. Talagrand

I pop up into the discussion on the question of operational use of Kalman Filter (KF) and Variational Assimilation (4D-Var). Yes, Kalman Filter is used in operational Numerical Weather Prediction. It is used at the Canadian Meteorological Center (CMC) in the form of Ensemble Kalman Filter (EnKF). And, unless it has changed recently, the US National Centers for Environmental Prediction uses a sequential assimilation algorithm which is essentially a Kalman Filter where the updating of the background with new observations is achieved through minimization of an appropriate objective function (3D-Var).

However, I do not remember how the prediction error covariance matrix is evolved in time (if at all). That should be eas y to check. As to the compared numerical cost of KF and 4D-Var, it can depend very much on the numerical implementation of the algorithms (for instance, number of elements in EnKF, or degree of parallelization). Buehner et al. (2010a, b) have made a rather clean comparison at CMC, which possesses both a lgorithms in operational order (but uses at present EnKF). They found that, for the same overall computational cost, the two algorithms produce results that can be different , but are of globally similar quality.

4.2 Author's response

We Thank O. Talagrand for pointing to various sources to answer the question of the referees. We used those informations to answer the comments from Referee #3

Additional change in the text

Author added few lines of text in the introduction, and rewrote the acknowledgement section.

Manuscript prepared for Nonlin. Processes Geophys.
with version 5.0 of the LaTeX class copernicus.cls.
Date: 9 December 2014

Toward the Assimilation of Images

F.-X. Le Dimet [1], I. Souopgui[2], O. Titaud[3], and V. Shutyaev[4]

[1]INRIA, LJK and Université de Grenoble Alpes, 38041, Grenoble CEDEX, France
[2]Department of Marine Science, University of Southern Mississippi,
1020 Balch Blvd, Stennis Space Center, MS 39529, USA
[3]CERFACS, 31057 Toulouse Cedex 01, France
[4]Institute of Numerical Mathematics, Russian Academy of Sciences,
119333 Gubkina 8, Moscow

Correspondence to: F.-X. Le Dimet (ledimet@imag.fr)

Abstract. The equations that govern geophysical fluids (namely atmosphere, ocean and rivers) are well known but their use for prediction requires the knowledge of the initial condition. In many practical cases, this initial condition is poorly known and the use of an imprecise initial guess is not sufficient to perform accurate forecasts because of the high sensitivity of these systems to small perturbations. As every situation is unique, the only additional information than can help to retrieve the initial condition are observations and statistics. The set of methods that combine these sources of heterogeneous information to construct such an initial condition are referred to as data assimilation. More and more images and sequences of images, of increasing resolution, are produced for scientific or technical studies. This is particularly true in the case of geophysical fluids that are permanently observed by remote sensors. However, the structured information contained in images or image sequences is not assimilated as regular observations: images are still (under) utilized to produce *qualitative* analysis by experts. This paper deals with the *quantitative* assimilation of information provided in an image form into a numerical model of a dynamical system. We describe several possibilities for such assimilation and identify associated difficulties. Results from our ongoing research are used to illustrate the methods. The assimilation of image is a very general framework that can be transposed in several scientific domains.

1 Introduction

For more than six decades, following the works of J. Von Neumann and J. Charney, the fluid envelope of the Earth has been described by mathematical models giving the evolution of its *state variables*: wind, temperature, pressure and moisture for the atmosphere, current, temperature, salinity and surface elevation for the sea. Models are routinely used for prediction and the level of prediction has been dramatically improved over the last few years.

For more than five decades, the fluid envelope of the Earth has been observed by satellite providing a long time and total coverage of the ocean and of the atmosphere. Billions of images have been produced, some of which are exhibited in art galleries showing the beauty of our Earth. These images and their dynamics show complex structures in different areas: tropical depressions, storms at mid-latitudes but also temperature, salinity and phytoplankton blooming in the ocean. These images are often used in meteorological bulletins on TV to illustrate the evolution of the weather. Thus they are important for a qualitative understanding of the evolution of the weather.

Images and models describe the same objects but with different tools. Images are often used to verify models – in general in fluid dynamics and turbulence – but it is done in a qualitative way rather than in a quantitative one. Both models and experiments display, for instance, images of Kelvin waves showing that models can mimic nature. But to what extent? How is it possible to quantitatively compare images of Kelvin waves observed from experiments and images from numerical models?

Over more than two decades, data assimilation has progressed into a very important development which is considered as the main reason for the improvement of forecasts. By data assimilation, we mean all the methods able to link together all the available information on geophysical fluids:

1. mathematical information provided by models

2. physical information provided by in-situ or remote observations
3. statistical information issued both from observations and from past predictions
4. a priori information, e.g. the regularity of the fields.

For many years, numerical models and images have both been used for a qualitative prediction. However, the structured information borne by images and by models are not presently used together in a quantitative framework.

The purpose of this paper is *Direct Image Assimilation*, which can be summarized as: *How to couple the information provided by numerical models and the information provided by images?* And the dual problem: how to validate mathematical models of flows using images and their dynamics. By validating, we mean a quantitative validation and not just a qualitative one. We must deal with a very general problem arising in many other fields out of geophysical problems. In nature, every situation is unique; steady state or asymptotic solutions do not exist. Most of the time, this assumption of uniqueness is implicitly used in modeling. Because of the enhancement of modeling, this fact will become crucial. In many cases, the initial condition and/or boundary conditions can not be experimentally controlled and consequently mathematical models are not sufficient to give an accurate representation of the situation. More information must be added and inserted into models. Images are, in some situations, candidates for this purpose. We would like to point out the ambiguity of the nature of some sequences of images. Some elements are clearly lagrangian, this is the case of small cumulus humilis clouds drifted with the wind, under the tropics they are used, by operational centers, as lagrangian markers. As such, they give an estimation of the wind that can be used in a data assimilation scheme. Some elements are clearly eulerian: this is the case of lenticularis clouds, they seem to be quasi steady, but in fact they are the signature of a strong wind. Estimating wind velocity from the shift of these clouds would lead to erroneous data. In betweens these examples, many visual elements in geophysical fluids dynamics have both an eulerian and a lagrangian character. The methods developed for assimilating images could be helpful for a better understanding of the underlying physics.

For engineering problems, the unknown conditions may be some parameters which have to be identified as a solution of an inverse problem, a methodology which can be included in data assimilation as it is~~.~~, see for example Chapelle et al. (2013) for application in Biological Engineering problems. Nevertheless, this paper will be more oriented towards applications to geophysical fluids. The remaining part of the paper is organized as follows: Sect. 2 gives a brief description of the observation by satellites. Section 3 is devoted to a brief introduction to data assimilation using variational methods. It also describes the characteristics of satellite observations in the sense of data assimilation. Section 4 describes the use of images as the source of pseudo-observations in data assimilation. Section 5 gives a methodology for direct assimilation of images and introduces the notion of an observation operator for images. Examples of such operators are given in Sect. 6 as well as associated numerical results. Section 7 concludes the paper.

2 Satellite observations

At the present time, more than forty satellites are continuously scanning the atmosphere and the ocean. As an illustration, Fig. 1a gives the number of observations provided by satellites and its evolution from 1996 to 2010.

2.1 Classification of satellites

Satellites can be classified according to many criteria: usage and orbit characteristics are the major criteria. In terms of usage, we are interested in earth observation and weather satellites. The next section gives a brief description of data provided by those satellites. In terms of orbit characteristics, the most important are the altitude and the inclination (in reference to the equatorial plane). The altitude and the inclination define the resolution, the coverage, and the acquisition conditions (local solar time at the acquisition point) of the measurement instruments on board the satellite. Most of the earth observation and weather satellites can be classified as geostationary or polar orbiting:

1. Geostationary satellites. They are synchronous with Earth rotation, consequently, because of the Coriolis force, they are necessarily located above the equator at an altitude of $\sim 35\,786$ km. Their position above the equator makes it almost impossible to observe polar regions and the high altitude does not allow acquisition in the microwave band. The spatial resolution of measurements is fine at the equator and degrades gradually as one moves away. Their stationary position above the earth makes it possible to get frequent measurements at the same point. At the current time, most of the operational weather geostationary satellites provide a full image of their coverage area every 15 min. The coverage area for such a satellite is about the quarter of the surface of the Earth. There are presently around ten geostationary satellites, each one observing a part of the Earth. Most of the visual information displayed on weather bulletins on TV are issued from these satellites. They clearly show the evolution of the large scale air masses, the birth of tropical depressions and hurricanes and, even at a local scale, the development of thunderstorms. Figure 1b shows the coverage of the Earth by observations from geostationary satellites.

2. Polar orbiting satellites. These satellites have an altitude between 400 and 800 km. At that elevation, they

can cover the broad spectrum of radiation including the microwave range as opposed to geostationary satellites. Also, the spatial resolution of measurements is very fine thanks to the low altitude. However, the geographical coverage is quite narrow at a given time. Also, the time resolution is coarse. In some orbits, it takes several days for the satellite to travel above the same point. They need several orbits to cover the entire globe. These satellites pass over the poles at each revolution, making it possible to get information more frequently in those areas not covered by geostationary satellites and almost inaccessible by conventional instruments. According to their inclination, they can be divided into subclasses. The subclass that draws attention is the subclass of sun-synchronous orbit. The polar sun-synchronous satellites pass the equator at the same local time on every pass. Those satellites are useful for imaging and weather. Figure 1c shows the distribution of observations from polar orbiting satellites equipped with the AMSU-A sensor.

2.2 Content of satellite measurements

Satellite sensors measure radiation reflected or emitted by the Earth, the seas or the atmosphere. The measured radiations are reflected light for visible channels and radiance for infrared channels. Depending on the wavelength employed, the measured radiations quantify a variable or a set of variables of the studied system. They can therefore be considered as observations in the sense of data assimilation.

In visible channels, satellites measure the reflective properties of the observed system (see Fig. 3). This is often limited to the upper layer of clouds. If the atmosphere is not cloudy, the observed surface can be extended to the Earth and sea surface. The observation of the sea in the visible channel produces the sea surface color (see Fig. 2). It shows the concentration of the phytoplankton in a thin upper layer of the sea. In infrared channels, satellites measure an integration over a certain thickness of the emissive properties of the observed system. Examples are water vapor images in the atmosphere and the sea surface temperature (SST) images. The integration thickness is highly dependent on the observed system. The sea is impervious to electromagnetic waves making measurements to be limited to a very thin layer of the sea surface. SST measurements, for example, are limited to only a few millimeters of the sea surface. This thickness is negligible compared to the thickness of the top layer in numerical models of the sea, which can extend to some hundred meters. Data like the SST are thus more related to interactions between the sea and the atmosphere than to the sea state variables. For the atmosphere, the probed layer can extend to its full thickness under the satellite. In the case of water vapor, the thickness of the probed layer is more important than in the case of the SST, but depends on the distribution of moisture in the atmosphere. Some specialized satellites provide more complex data, an example is the Jason type satellites that give the sea surface elevation with a precision of some centimeters.

As opposed to in-situ measurement devices that usually acquire observations at a single point at each time, satellite measurements provide observation of a large area at the same time. Thus, these satellites provide the image of their coverage area by the radiation function at each time. For this reason, satellite observations are usually called images. Subsequently, the expression *"satellite image"* means satellite observation.

3 Data assimilation

3.1 Definition of data assimilation

By data assimilation, we mean the methods permitting the best retrieval of the state of the environment, a mandatory step prior to prediction. From the formal point of view, the problem is to link heterogeneous sources of information, the heterogeneity bearing on the nature, the quality and density. Basically we have:

1. Mathematical information. This is the model which is used to describe the flow.

2. Physical information. It is given by data issued from in situ or remote measurements such as images.

3. Statistical information. It could be produced from statistics on the observations as well as statistics on the outputs of the model.

4. A priori information, for instance on the regularity of the fields or the existence of singularities. More generally, qualitative information used in the analysis.

Basically, there are two approaches of data assimilation methods for combining all the previously mentioned information:

1. Approaches derived from the Kalman filter. They are based on Bayesian estimation and are of great theoretical importance~~, but having~~. Having to deal with a huge covariance matrix, ~~they~~ the traditional Kalman filters are not implemented in operational centers. The Ensemble Kalman filter of Evensen (1994) overcomes that limitation by introducing an ensemble approximation of the covariance matrix.

2. Variational approaches. They are based on optimal control and the calculus of variations. These methods are presently used by most important operational centers for weather prediction. They seem well adapted for the assimilation of images and, in the sequel, we will only consider the variational approaches.

3.2 Variational data assimilation

The ingredients of variational data assimilation are:

- A model describing the evolution of the state variable $\boldsymbol{X} \in \mathcal{X}$. The model is usually given as a system of partial differential equations (PDE) of the form:

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \mathcal{M}(\boldsymbol{X}), \quad t \in [0,T], \\ \boldsymbol{X}(0) = \boldsymbol{U}, \end{cases} \tag{1}$$

  where the initial condition $\boldsymbol{U} \in \mathcal{X}$ is unknown, $\mathcal{X}$ is the state space and $\mathcal{M}$ is the model operator. For illustration, in the case of atmospheric systems, the state variable $\boldsymbol{X}$ represents variables such as wind, temperature, pressure, etc., and the dynamic model $\mathcal{M}$ describes the set of physical laws that the variables must respect over time. These laws are: thermodynamics laws, conservation laws, etc.

- A set of observations $\boldsymbol{Y}^{\mathrm{o}}$ given by physical measurement (direct or indirect) of the system state. For the sake of simplicity, we will assume observations to be continuous in time:

$$\begin{aligned} &\boldsymbol{Y}^{\mathrm{o}} : \mathbf{R}^+ \to \mathcal{O}, \\ &\qquad t \mapsto \boldsymbol{Y}^{\mathrm{o}}(t); \end{aligned} \tag{2}$$

- An operator of observation $\mathcal{H}$: observations are usually made up of partial or indirect measurements of the state variables. The observation space $\mathcal{O}$ is not necessarily the same as the state space $\mathcal{X}$. The *observation operator* $\mathcal{H}$ is defined as the mapping operator from $\mathcal{X}$ onto $\mathcal{O}$:

$$\begin{aligned} &\mathcal{H} : \mathcal{X} \to \mathcal{O} \\ &\qquad \boldsymbol{X}(t) \mapsto \boldsymbol{Y}(t) = \mathcal{H}[\boldsymbol{X}(t)]. \end{aligned} \tag{3}$$

- A background estimation $\boldsymbol{U}^{\mathrm{b}}$ of the initial state $\boldsymbol{U}$. In operational meteorology, this background estimation can be deduced from previous forecasts.

- Statistical information, for instance the error covariance matrix $\mathbf{Q}$ of the observation error and the covariance matrix $\mathbf{B}$ of the background estimation.

Variational data assimilation (VDA) defines the optimal initial condition $\boldsymbol{U}^a$ as:

$$\boldsymbol{U}^a = \text{argmin } J(\boldsymbol{U}), \tag{4}$$

where the so-called cost function J is defined as:

$$J(\boldsymbol{U}) = \frac{1}{2}\int_0^T \| \mathcal{H}[\boldsymbol{X}(t)] - \boldsymbol{Y}^{\mathrm{o}}(t) \|^2_{\mathbf{Q}^{-1}} \,\mathrm{d}t + \frac{1}{2} \| \boldsymbol{U} - \boldsymbol{U}^{\mathrm{b}} \|^2_{\mathbf{B}^{-1}}, \tag{5}$$

with the norms $\|\boldsymbol{Y}\|_{\mathbf{Q}^{-1}} = \|\mathbf{Q}^{-\frac{1}{2}}\boldsymbol{Y}\|$ and $\|\boldsymbol{U}\|_{\mathbf{B}^{-1}} = \|\mathbf{B}^{-\frac{1}{2}}\boldsymbol{U}\|$.

The cost function contains two terms: the first one is the discrepancy between observations and the solution of the model associated with the initial condition $\boldsymbol{U}$. The second one is the background term. It will require the solution to be located in the vicinity of $\boldsymbol{U}^{\mathrm{b}}$. It is also a regularization term in the sense of Tikhonov (1963). This term is mandatory due to the ill-posedness of the problem. In operational meteorology, the dimension of the state vector, and consequently the dimension of the initial condition, is of the order of one billion while the number of daily observations is of the order of ten millions. Therefore we would have to deal with a severely ill-posed problem (as defined by Hadamard) if the regularization were not introduced. A necessary condition for the optimality is given by the Lagrange–Euler equation:

$$\nabla J(\boldsymbol{U}^a) = 0, \tag{6}$$

this is also a sufficient condition if J is strictly convex and coercive. This is the case if we have a linear model but realistic models are nonlinear in general. Solving the Lagrange–Euler equations requires the gradient ∇J of the cost function. The main difficulty is that J is an implicit function of $\boldsymbol{U}$. In VDA, ∇J is computed through the *adjoint variable* $\boldsymbol{P}$, which is defined as the solution of the adjoint model:

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} + \left[\dfrac{\partial \mathcal{M}}{\partial \boldsymbol{X}}\right]^* \cdot \boldsymbol{P} = \left[\dfrac{\partial \mathcal{H}}{\partial \boldsymbol{X}}\right]^* .\mathbf{Q}^{-1}(\mathcal{H}(\boldsymbol{X}) - \boldsymbol{Y}^o), \; t \in [0,T] \\ \boldsymbol{P}(T) = 0, \end{cases} \tag{7}$$

where the * denotes the adjoint operator. The adjoint model is deduced from the direct model Eq. (1) using calculus of variations based on the Gateaux derivatives, see Le Dimet and Talagrand (1986) for details. The gradient of the cost function $\nabla J(\boldsymbol{U})$ is given by

$$\nabla J(\boldsymbol{U}) = -\boldsymbol{P}(0) + \mathbf{B}^{-1}(\boldsymbol{U} - \boldsymbol{U}^{\mathrm{b}}). \tag{8}$$

The gradient is then used in an optimization algorithm (Truncated or Quasi-Newton methods, L-BFGS) to compute an estimate of the optimal solution.

3.3 Satellite observations and data assimilation

Observations that are quantitatively used in data assimilation are usually limited to measurements of the state variables, such as wind, moisture and pressure given by terrestrial centers for meteorological prediction. These observations will be named conventional in the subsequent. Apart from conventional observations, there exists another class of observations that is mainly used only for qualitative purposes: these are images. Among the various sources of images, satellites plays an important role for the observation of the atmosphere and seas.

As we mention in Sect. 2.2, satellite observations or satellite images are indirect measurements of the state variables of observed systems like the atmosphere or the sea. Thanks to post processing, they can be converted into observations of the associate variables. As an example, Fig. 2 shows the SST and the chlorophyll concentration derived from MODIS (on board satellite Aqua) observation of the Gulf Stream. MODIS stand from Moderate-resolution Imaging Spectroradiometer; it is a 36-channels scientific instruments that equips NASA satellites Terra and Aqua.

Thanks to their high resolution and their spatial coverage, satellite images also provide information on structures ranging from mesoscale to synoptic scale. Structure refers to the spatial organization of individual measurements. A sequence of images shows the dynamical evolution of the structures. As example, in addition to the SST and the chlorophyll concentration, Fig. 2 shows a couple of large Gulf Stream eddies. The similarity of observed structures between SST (infrared channel) and Chlorophyll concentration (visible channel) shows that such information can be obtained from different measurements. The example of Fig. 3 shows a depression over western Europe and its evolution from 28 April to 29 April 2008. The observed structures (eddies for the sea, depression for the atmosphere, etc.) represent Lagrangian information and are clearly useful for the prediction of the observed system. From the above description, we can distinguish two major types of satellites observations:

1. indirect measurements of the state variables of observed systems;
2. characteristic structures of the observed system and their dynamics.

3.3.1 Satellite observations as indirect measurements of the state variables

From this point of view, one can derive two approaches for using satellite observations. The first approach consists in extracting the variables that are indirectly observed and use them as conventional observations in a model that contains those variables in the system state. The second approach consists in modeling an appropriate observation operator that computes radiance from the system state given by the model. In both cases, satellite observations are used as conventional observations; this consideration will not be taken into account in the rest of this paper. The two cases are subject to some problems including: the difficulty of extracting variables from indirect measurements or modeling the appropriate observation operator; the sensitivity of satellite measurements to acquisition conditions. For example, a substantial cloud cover makes the error rate prohibitive in the observations of temperature and moisture of the atmosphere. In these cases, measurements are used to derive other products such as velocity fields (atmospheric motion vector or AMV). However the combination of four Dimensional Variational Data Assimilation (4D-VAR) and the use of satellite measurements has significantly improved the forecasts as shown by Fig. 4. This figure shows the anomaly correlation at 500 hPa height for 3, 5 and 7 days forecast between the years 1992 to 2007. Before the year 2000, there was a significant difference between the forecast in the Northern Hemisphere (high performance) and Southern Hemisphere (poor performance). The difference was due to the lack of conventional observation in the Southern Hemisphere. In the early 2000s, the introduction of satellite observations in data assimilation made it possible to get the same performance in both hemispheres.

3.3.2 Satellite observations as characteristic structures of the studied system and their dynamics

This is the approach that will be developed in the rest of this paper. In this case, satellite measurements can not simply be used as conventional observations. In fact, as structures refer to the spatial organization of individual measurements, a single measurement is useless. Similarly, as the dynamics refers to the evolution of measurements in time, a single image is not sufficient. However, the observed structures are indirectly present in the model output, provided with appropriate initial conditions and other parameters of the model. The question that arises is: how to use such information in data assimilation? The answer to this question is the *assimilation of images*. This is a concept that emerged recently with the aim of using images as observations in data assimilation. There are two basic approaches:

1. Assimilation of pseudo-observations. In a first step, the images are analyzed. The results of the process is a field of velocities obtained by the comparison of two or several successive images. In a second step, these velocities are assimilated as conventional observations in a classical method of data assimilation.
2. Direct assimilation of images. Images (image structures) are considered as conventional observations and assimilated as such. To do so, depending on the application, we need to define an adequate mathematical space in which images or image structures will be modeled. Corresponding observation operators that map the control space into the structure space should be constructed. The structure space must conserve and extract the most pertinent information of the images. If we want to remain in the framework of optimal control methods, then the space must be defined in such a way that the rules of differential calculus can be applied. It is also important to underline that, for computing purposes, the space dimension should not be too large.

In both cases, a preliminary step for using images in data assimilation is the identification of the underlying process. However, this paper does not focus on this preliminary step;

instead, it focuses on the mathematical aspects of the use of images in data assimilation. Before moving forward, let us give some illustrations of the importance of that preliminary step. Lenticular clouds may be observed under the wind over a mountain; they are an Eulerian property of an area where there is condensation of water vapor. These clouds appear to be quasi-stationary, consequently if they were used as a Lagrangian tracer, they would lead to a small wind velocity. Such an analysis would be a misinterpretation of reality as these clouds are actually the signature of a strong wind. This is also the case of some small cumulus clouds that can appear at the vertical edge of some crops with strong radiative properties. They are the signature of a local vertical convection and therefore are not useful for retrieving horizontal velocities. It is important to mention that phase errors and joint phase-amplitude should be considered in the assimilation of remote measurements. This issue is not addressed in this paper as it is not the main topic of this paper. However, there is a significant literature on the topic that can be of interest to the reader (Hoffman et al., 1995; Hoffman and Grassotti, 1996; Brewster, 2003a,b; Ravela et al., 2007).

4 Images as source of pseudo-observations in data assimilation

4.1 Principle of assimilation of pseudo-observations

Since the early 80's with the works of Horn and Schunck (1981), research has been carried out to derive velocity fields from images sequences, with applications to fluid dynamics mainly (and very recently to movie compression and medical imagery). The velocity field derived from the image processing techniques can be used as pseudo-observations of wind in an assimilation system, for instance in a regular VDA scheme.

The left panel of Fig. 5 shows the principle of assimilation of pseudo-observations of velocity fields. From a sequence of images, a velocity fields is estimated and used as observation of the velocity field in a regular scheme of data assimilation. There are several methods to extract a velocity field from a sequence of images. They can be divided into two categories: the frame-to-frame motion estimation and the so-called image model technique. Motion estimation or frame-to-frame motion estimation is a technique from image processing that aims at estimating the velocity field that transports an image to another. A mathematical description of this technique is given in Sect. 4.2. From frame-to-frame motion estimation, one gets a velocity field between each pair of successive images of a sequence, but there is no guarantee of consistency in the resulting sequence of fields if it is applied to many pairs of images of the same sequence. In such cases, the image model technique can be more appropriate. It couples the frame-to-frame technique with an evolution model for the velocity field. For details on this technique, the reader is referred to Herlin et al. (2006); Huot et al. (2006); Korotaev et al. (2007).

4.2 Frame-to-frame motion estimator

The description of motion estimation in this paper is limited to *optical flow*. It is a variational method and is well suited for image sequences in geophysics. There also exist statistical methods based on the correlation between successive images. For more information on those methods, the reader is referred to Adrian (1991) that describes the commonly used one: Particle Image Velocimetry (PIV). *Optical flow* is a classical method of motion estimation. It is based on the conservation of the global luminance between two images (Horn and Schunck, 1981). Let $I : \Omega \times \mathbf{R} \to \mathbf{R}$ be the luminance function defined on the pixel grid $\Omega \subset \mathbf{R}^2$ and the time $t \in \mathbf{R}$, the optical flow is the vector field $\boldsymbol{V}(x,y)$ that satisfies the luminance conservation given by the following equation:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\partial I}{\partial t} + \nabla I \cdot \boldsymbol{V} = 0. \tag{9}$$

According to the nature of the images, the law of conservation of the luminance Eq. (9) can be replaced by a specific law. For example, with images of the ocean's color, conservation of chlorophyll (with source and sink terms) can be considered; with images of Sea Surface Temperature (SST), the Boussinesq approximation can be used, etc. In many cases, the validity of these laws, from the optical point of view, is dubious. For instance, between two satellite images, the enlightenment will have changed and some corrective term will have to be added to the equation. As a consequence, it is necessary to carry out a preliminary study of the images to detect structures on which the information borne by the equation of conservation and the images is maximized. For instance, if we are working with Sea Surface Temperature (SST), filaments are important structures which have to be identified in the analysis. They are characterized by: elongated structures, constant temperature, significant contrast with the surrounding area, and motion by translation. To identify filaments, it is necessary to use the tools of mathematical morphology (Serra, 1988; Najman and Talbot, 2010). In the images, it will also be necessary to discard points with a weak spatial gradient or with a weak temporal evolution. Detecting and/or eliminating structures from the images requires the application of a thresholding operator (e.g. on the norm of the gradient of the SST). Of course, the analysis will be sensitive to the threshold value chosen. The choice of the threshold is usually empirical.

For a two dimensional problem, the velocity field $\boldsymbol{V} = (u,v)$ is determined as the solution of an optimization problem. To this end, one defines a cost function J to be minimized as follows:

$$J(u,v) = \frac{1}{2}\int_0^T \int_\Omega \left[\frac{\partial I}{\partial t} + u\frac{\partial I}{\partial x} + v\frac{\partial I}{\partial y}\right]^2 \mathrm{d}x\mathrm{d}y\mathrm{d}t. \tag{10}$$

A necessary condition for optimality is expressed by the Euler–Lagrange equations that involve the gradient of J with respect to u and v. For the cost function of Eq. (10), the Euler–Lagrange equations give the solution $\boldsymbol{V}^* = (u^*, v^*)$ as the solution of the linear system:

$$u\left[\frac{\partial I}{\partial x}\right]^2 + v \cdot \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} = -\frac{\partial I}{\partial t} \cdot \frac{\partial I}{\partial x} \tag{11}$$

$$u \cdot \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} + v \cdot \left[\frac{\partial I}{\partial y}\right]^2 = -\frac{\partial I}{\partial t} \cdot \frac{\partial I}{\partial y}. \tag{12}$$

The determinant det of this system and the determinants $\det_u$ and $\det_v$ relative to unknowns u and v are all zero, meaning that the solution is not unique. The problem is ill-posed. In fact, let $\boldsymbol{V} = (u,v)$ be a solution of the non regularized problem, and $\boldsymbol{W} = (w_1, w_2)$ a vector fields, orthogonal to the image gradient, i.e. ($\langle \boldsymbol{W}, \nabla I \rangle = 0$). We have:

$$J(\boldsymbol{V} + \alpha \boldsymbol{W}) = J(\boldsymbol{V}), \forall \alpha \in \mathbf{R}. \tag{13}$$

As a consequence, it is impossible to determine the motion in the direction orthogonal to the image gradient: this is the *aperture problem* that is well known in computer vision. It is a source of ill-posedness. To address the ill-posedness, regularization techniques are used. The literature on the regularization for image processing is very large. The references Tikhonov (1963), Horn and Schunck (1981), Alvarez et al. (1999), Nagel (1983), Schnörr (1994), Suter (1994), Weickert and Schnörr (2001), Black and Anandan (1991), Hinterberger et al. (2002), and Mémin and Perez (1998) give a start point for the interested reader.

5 Direct assimilation of images

5.1 Mathematical formulation

By direct assimilation of images, we mean using image observations directly in the cost function of variational data assimilation. In this case, image observations are jointly used with conventional observations to compute the optimal control variable of numerical models. The right panel of Fig. 5 shows a schematic representation of direct assimilation of images. Images are used directly in the optimality system jointly with conventional observations. This direct use of images in the optimality system requires the definition of a mathematical space for the images with adequate topology and the associated *images observation operator*. An *images observation operator* is a mapping from the space of the numerical solution of the model toward the space of images. No prior step to extract pseudo-observations of state variables is needed. Direct assimilation of images requires the modification of the cost function in order to take into account the image observations. The cost function that takes into account images can be written as follows:

$$J(\boldsymbol{X}_0) = \underbrace{\frac{1}{2}\int_0^T \|\mathcal{H}(\boldsymbol{X}(t)) - \boldsymbol{Y}^{\mathrm{o}}(t)\|^2_{\mathbf{Q}^{-1}}\,\mathrm{d}t}_{\text{conventional cost}} + \underbrace{\frac{1}{2}\int_0^T \|\mathcal{H}_{\mathcal{X}\to\mathcal{F}}(\boldsymbol{X}(t)) - f(t)\|^2_{\mathcal{F}}\,\mathrm{d}t}_{\text{image cost}}, \tag{14}$$

where $f(t)$ is the image function at time t, $\|.\|_{\mathcal{F}}$ is the appropriate norm in the image space $\mathcal{F}$, and $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ is the observation operator for images; subsequently, it will be called the *model to image operator*. In Eq. (14), the background and regularization terms are omitted for sake of clarity. The regularization term will be the canonical one in VDA.

5.2 Observation operators for images

A first consideration of the image cost function is to take the norm of the left hand side of the equation of the optical flow Eq. (9). The image is considered as a passive tracer moving with respect to the dynamics of the system, and more precisely with the motion field $\boldsymbol{V}$. This approach, proposed in Béréziat and Herlin (2011); Papadakis and Mémin (2008); Gorthi et al. (2011) leads to the following image cost function:

$$\frac{1}{2}\int_0^T \int_\Omega \left\| \frac{\partial I}{\partial t} + \nabla I \cdot \boldsymbol{V} \right\|^2_{\mathbf{Q}^{-1}} \mathrm{d}x\mathrm{d}y\mathrm{d}t. \tag{15}$$

This cost function can not be turned easily into the form suggested by Eq. (14). The covariance matrix $\mathbf{Q}$ is defined with respect to the image gradients ∇I in order to restrict image information to pertinent areas containing discontinuities. If the model $\mathcal{M}$ is monotone and ensures the spatio-temporal continuity of the state $\boldsymbol{X}(t)$, the regularization of the flow $\boldsymbol{V}$ at time t now only depends on the regularity of the background condition $\boldsymbol{V}_0$.

Due to the characteristics of images, they should not be used directly as an array of pixels in the cost function. Specific structures of the image, such as lenticular clouds mentioned above, may have their own dynamics. In such cases, image observations can not simply be considered as a passive tracer moving under the dynamics of the studied system. It is also important to point out that from a dynamical point of view, information in an image sequence are located in discontinuities and the dynamics of those discontinuities. Even with advanced covariance matrices, the pixel representation of images is not suitable to describe such phenomena in data assimilation. Additional operators should be used to isolate structures of interest from the image. In this case, the cost function Eq. (14) takes the form:

$$J(\boldsymbol{X}_0) = \underbrace{\frac{1}{2}\int_0^T \|\mathcal{H}(\boldsymbol{X}(t)) - \boldsymbol{Y}^{\mathrm{o}}(t)\|^2_{\mathbf{Q}^{-1}}\,\mathrm{d}t}_{\text{conventional cost}} + \underbrace{\frac{1}{2}\int_0^T \|\mathcal{H}_{\mathcal{X}\to\mathcal{S}}(\boldsymbol{X}(t)) - \mathcal{H}_{\mathcal{F}\to\mathcal{S}}(f(t))\|^2_{\mathcal{S}}\,\mathrm{d}t}_{\text{image cost}}, \tag{16}$$

where $\mathcal{S}$ is the space of features of interest to be isolated from the image. This notation is borrowed from Titaud et al. (2009) where such a space is called the "space of structures". $\mathcal{H}_{\mathcal{F}\to\mathcal{S}}$ is the *image to structure operator* and $\mathcal{H}_{\mathcal{X}\to\mathcal{S}}$ is the *model to structure operator*.

Image to structure operator: The goal of such an operator is to extract features of interest from image observations. As we said previously, the main information obtained by human vision from the image is located in the discontinuities. A definition of $\mathcal{S}$ must be related to the discontinuities in the image function. Discontinuities are well characterized in spectral spaces. Thus, the basic definition of $\mathcal{S}$ may be based on a spectral decomposition such as Fourier, wavelet or curvelet.

Model to structure operator: This operator extracts features of interest from the system state given by the model. It can be defined as:

$$\mathcal{H}_{\mathcal{X}\to\mathcal{S}} = \mathcal{H}_{\mathcal{F}\to\mathcal{S}} \circ \mathcal{H}_{\mathcal{X}\to\mathcal{F}}, \tag{17}$$

where $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ is the *model to image operator* previously defined. Setting the *image to structure operator* to be the identity ($\mathcal{H}_{\mathcal{F}\to\mathcal{S}} = Id$), we get the cost function given by the Eq. (14). Another approach of using image observation in Laboratory data assimilation can be found in Ravela et al. (2010). The authors used computer vision system to extract measurements from the physical simulation with parallel computing and decomposition to account for observation in real time as well as using the numerical model to adapt the observing system.

5.3 Adjoint model in direct image assimilation

When the image term is added to the cost function, Eq. (16), the adjoint model of variational data assimilation, Eq. (7) becomes:

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t} + \left[\dfrac{\partial\mathcal{M}}{\partial\boldsymbol{X}}\right]^* \cdot \boldsymbol{P} \;=\; \underbrace{\left[\dfrac{\partial\mathcal{H}}{\partial\boldsymbol{X}}\right]^* .\mathbf{Q}^{-1}(\mathcal{H}(\boldsymbol{X}) - \boldsymbol{Y}^{\mathrm{o}})}_{\text{conventional forcing term}} \\ + \underbrace{\left[\dfrac{\partial\mathcal{H}_{\mathcal{X}\to\mathcal{S}}}{\partial\boldsymbol{X}}\right]^* .(\mathcal{H}_{\mathcal{X}\to\mathcal{S}}[\boldsymbol{X}] - \mathcal{H}_{\mathcal{F}\to\mathcal{S}}[f])}_{\text{image forcing term}} \\ \boldsymbol{P}(T) \;=\; 0, \end{cases} \tag{18}$$

The expression

$$\left[\frac{\partial\mathcal{H}_{\mathcal{X}\to\mathcal{S}}}{\partial\boldsymbol{X}}\right], \tag{19}$$

is the Jacobian of the *model to structure operator*. The presence of this expression means that the *model to structure operator* must be differentiable. Then we can compute its Jacobian and the gradient of the cost function in order to be able to carry out an optimization algorithm and identify the optimal initial condition.

6 Examples of direct image assimilation techniques

In this section, we describe two tools that can be used to construct observation operators. The first method uses the advection of a passive tracer whose concentration map is considered as the image. This method is well adapted for assimilating a sequences of images. The second method uses the computation of Lagrangian Coherent Structures (LCS) of the flow. This method exploits the *integrated information* contained in tracer images and is well suited for single image assimilation. We will also discuss different examples of mathematical spaces for image structures. All the associated topologies will be of $\mathcal{L}^2$-type. The main purpose of this choice is its convenience; other choices, e.g $\mathcal{L}^1$ that is commonly used in image processing, could be considered. Also the question of introducing some covariance matrix into the definition of the topology remains open. The choices shown below are not exhaustive. Many other potential spaces could be considered. The choice of the mathematical space for images defines the *image to structure operator* that has been introduced in the previous section.

6.1 Observation operators based on the advection of passive tracer

In this subsection, we consider the case where the *model to structure operator* can be decomposed into a *model to image operator* and an *image to structure operator*. We focus only in the *image to structure operator* that is the most important as stated in Sect. 5. A simple example of the *model to image operator* can be defined by considering images as observations of a passive tracer. Image evolution is then modeled by a transport equation, the initial distribution of the passive tracer being given by the first image. Interpolation from the grid points of the numerical model toward the grid points of the image can be necessary. In this case, an image is considered as the concentration of the passive tracer.

6.1.1 Pixel representation of image

The pixel representation of a 2-D image is a discretization (numerical representation) of a mathematical function of two variables that defines the image. It is usually given as a 2-D array, each entry of the array being the value of the image at the associated grid cell of the discretization. The simple case of image assimilation is to consider the identity *image to structure operator*. In this case, the cost function associated with the image will take the form:

$$J(\boldsymbol{U}) = \frac{1}{2}\int_0^T \|\mathcal{H}_{\mathcal{X}\to\mathcal{F}}(\boldsymbol{X}(t)) - f(t)\|_{\mathcal{F}}^2 \mathrm{d}t \tag{20}$$

where, $\mathcal{H}_{\mathcal{X}\to\mathcal{F}}$ defines the concentration of the passive tracer from the system state and f is the observed concentration (image) at time t. In this case, the image is considered as an array of Eulerian observations of the tracers and the features of the dynamics (fronts, vortices, etc.) are not explicitly taken into account.

6.1.2 Multiscale analysis of images: curvelets

Recent years have seen a rapid development of new tools for harmonic analysis. For general fluid dynamics and also for geophysical flows, there are coherent structures evolving in an incoherent random background. If the flow is considered as an ensemble of structures, then the geometrical representation of flow structures might seem to be restricted to a well-defined set of curves along the singularities in the data. The first step in using images as observations in data assimilation is to separate the resolved structures, which are large, coherent and energetic from the unresolved ones, which are supposed to be small, incoherent and bearing little energy. One of the first studies in this sense can be found in Farge (1992). It shows that the coherent flow component is highly concentrated in wavelet space. Wavelet analysis is a particular space-scale representation of signals which has found a wide range of applications in physics, signal processing and applied mathematics in the last few years. The literature is rich regarding wavelets. The interested reader can be referred to Mallat (1989), Coifman (1990), and Cohen (1992) for example. A major inconvenience of wavelets is that they tend to ignore the geometric properties of the structure and do not take into account the regularity of edges. This issue is addressed by the curvelet transform. The curvelet transform is a multiscale directional transform that allows an almost optimal nonadaptive sparse representation of objects with edges. It has been introduced by Candès and Donoho (Candes and Donoho, 2004, 2005a,b; Candes et al., 2006). In $\mathbf{R}^2$, the curvelet transform allows an optimal representation of structures with C^2-singularities. As curvelets are anisotropic, they have a high directional sensitivity and are very efficient in representing vortex edges.

A function $f \in \mathcal{L}^2(R^2)$ is expressed in terms of curvelets as follows:

$$f = \sum_{j,l,\boldsymbol{k}} \langle f, \Psi_{l,j,\boldsymbol{k}}\rangle \Psi_{l,j,\boldsymbol{k}}. \tag{21}$$

where $\Psi_{j,l,\boldsymbol{k}}$ is the curvelet function at scale j, orientation l and spacial position $\boldsymbol{k}$ ($\boldsymbol{k} = (k_1, k_2)$). The orientation parameter is the one that makes the major difference with the wavelet transform. The set of curvelet functions $\Psi_{j,l,\boldsymbol{k}}$ does not form an orthonormal basis as it is the case for some families of wavelets. However, the curvelet transform satisfies the Parseval relation so that the L_2-norm of the function f is given by:

$$\|f\|^2 = \sum_{j,l,k} |c_{j,l,\boldsymbol{k}}|^2, \tag{22}$$

where $c_{j,l,\boldsymbol{k}}$ are the curvelet coefficients given by:

$$c_{j,l,\boldsymbol{k}} = \langle f, \Psi_{l,j,\boldsymbol{k}}\rangle. \tag{23}$$

In Fig. 6 from Ma et al. (2009), the supports of some wavelets and curvelets are presented. The figure shows the strong anisotropy curvelets and suggest that curvelet representation will give a better adjustment for 2-D-curves.

Figure 7 shows an illustrative comparison of the approximation of a circle by wavelets and by curvelets. The curvelets provide a better approximation of this perfectly anisotropic object. The convergence of curvelets is also better: the best m-term approximation f_m of a function f has the representation error

$$\|f - f_m\| \approx m^{-1}$$

for wavelets and

$$\|f - f_m\| \approx C m^{-2}(\ln m)^3$$

for curvelets. Another interesting property of curvelets in the framework of variational data assimilation is that the adjoint of the curvelet transform is the inverse of the curvelet transform. Therefore, to represent an image, we will consider the truncation of its expression in a curvelet frame.

6.1.3 Numerical experiments

In this subsection, we present numerical experiments of direct image assimilation with observation operators based on the advection of a passive tracer. We used images from experimental physics: the drift of a vortex is studied through physical experiment in the Coriolis platform: it is a circular rotating tank with a diameter of 14 m, located at Laboratoire des Écoulements Géophysiques et Industriels (LEGI), Grenoble, France. The rotation of the tank recreates the effect of the Coriolis force in a thin layer of fluid. The vortex is generated by stirring the fluid and made visible for optical images thanks to the addition of the fluorescein that is a passive tracer. Pictures are taken from above the turntable at regular time intervals to study the evolution of the vortex. A full description of a similar experiment can be found in Flór and Eames (2002). A sequence of two images from that experiment is used for the motion estimation experiment in this paper. This sequence is named Coriolis sequence after the name of the platform.

6.1.4 Experimental framework

In the configuration of the Coriolis platform as described above, the state variable is $\boldsymbol{X} = (u, v, h)$, which satisfy the shallow-water equations

$$\begin{cases} \partial_t u - (f+\zeta)v + \partial_x B = -ru + \nu\Delta u, \\ \partial_t v + (f+\zeta)u + \partial_y B = -rv + \nu\Delta v, \\ \partial_t h + \partial_x(hu) + \partial_y(hv) = 0. \end{cases} \tag{24}$$

Unknowns are the zonal component $u(t,x,y)$ and meridional component $v(t,x,y)$ of the current velocity and the surface elevation $h(t,x,y)$. They depend on time t and the two horizontal directions x and y. We define the relative vorticity $\zeta = \partial_x v - \partial_y u$ and the Bernoulli's potential $B = gh + \frac{1}{2}(u^2+v^2)$, where g is the gravity. The Coriolis parameter on the β-plane is given by $f = f_0 + \beta y$, ν is the diffusion coefficient and r the bottom friction coefficient. In this paper, the following numerical values are used for the parameters: $r = 0.9 \times 10^{-7}\,\mathrm{s}^{-1}$, $\nu = 0\,\mathrm{m}^2\,\mathrm{s}^{-1}$, $f_0 = 0.25\,\mathrm{s}^{-1}$, $g = 9.81\,\mathrm{m\,s}^{-2}$ and $\beta = 0.0406\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$. The simulation is performed on a rectangular domain $\Omega =]0, L[\times]0, H[$ representing a subdomain of the turntable with $L = H = 2.525\,\mathrm{m}$. The domain is discretized on a $N \times N = 128 \times 128$ uniform Arakawa C-type square grid. A finite difference scheme is used for space discretization. Time integration is performed using a fourth order Runge–Kutta scheme. The time step is set to 0.01 s in the turntable experiment, which corresponds to 14.4 s in the atmosphere.

6.1.5 Assimilation procedure

We consider the problem of recovering the ~~*initial state*~~ initial state of the fluid ~~$\boldsymbol{U}(x,y) = \boldsymbol{X}_0(x,y) = (u,v,h)(0,x,y)$~~ $U(x,y) = X_0(x,y) = (u,v,h)(0,x,y)$ which constitutes our control variable. Only images are used as observations. We use image to structure operators based on pixels~~and the thresholding of the curvelet decomposition. Three examples of the thresholding operator are considered: let $c_{j,l,\boldsymbol{k}}$~~. Edge structures of images are extracted by applying a threshold operator on their curvelet coefficients. More precisely, let $c_{j,l,\mathbf{k}}$ be the curvelet coefficients of the expression of a given function f in the frame of curvelets, see Eq. (23)~~. The following thresholding functions are considered~~: we consider the scale-dependent hard thresholding operator τ defined as:

$$\tau(c_{j,l,\mathbf{k}}) = \begin{cases} c_{j,l,\mathbf{k}} & \text{if } |c_{j,l,\mathbf{k}}| \geq \sigma_j, \\ 0 & \text{if } |c_{j,l,\mathbf{k}}| < \sigma_j, \end{cases} \tag{25}$$

where σ_j is the threshold value for the scale j. The σ_j are predefined and depend on the problem and on the data. We mention two particular cases:

1. hard thresholding ~~τ_1~~

$$\tau_1(c_{j,l,\boldsymbol{k}}) = \begin{cases} c_{j,l,\boldsymbol{k}} & \text{if } |c_{j,l,\boldsymbol{k}}| \geq \sigma, \\ 0 & \text{if } |c_{j,l,\boldsymbol{k}}| < \sigma, \end{cases}$$

τ_h with the same threshold for all scales: $\sigma_j = \sigma$ for a given σ,

$$\tau_h(c_{j,l,\mathbf{k}}) = \begin{cases} c_{j,l,\mathbf{k}} & \text{if } |c_{j,l,\mathbf{k}}| \geq \sigma, \\ 0 & \text{if } |c_{j,l,\mathbf{k}}| < \sigma, \end{cases} \tag{26}$$

2. ~~scale by scale thresholding τ_2~~

$$\tau_2(c_{j,l,\boldsymbol{k}}) = \begin{cases} c_{j,l,\boldsymbol{k}} & \text{if } |c_{j,l,\boldsymbol{k}}| \geq \sigma_j, \\ 0 & \text{if } |c_{j,l,\boldsymbol{k}}| < \sigma_j, \end{cases}$$

hard thresholding zeroing ~~the coarsest scale coefficients τ_3~~coarse scale coefficients τ_z; this is similar to the hard thresholding with the exception that the coefficient associated with each curvelet function of the coarsest scale is set to zero.

~~With the thresholding operator τ, the function f is approximated by:~~

$$\tilde{f} = \sum_{j,l,\boldsymbol{k}} \tau(c_{j,l,\boldsymbol{k}})\Psi_{l,j,\boldsymbol{k}}.$$

~~σ and σ_j in Eqs. (26) and (25) are predefined threshold that depend on the problem and on the data. For numerical experiments presented in this section, σ and σ_j are defined such that at much 10of the total number of coefficients are used.~~ Curvelet thresholding for edge extraction can also be found in Ma et al. (2006)

6.1.6 Numerical results

Figure 8 shows the initial analyzed velocity field with different observation operators. With the identity observation operator (pixels), the analyzed velocity field shows a non symmetric vortex and large motion where there must be no dynamics. With the hard thresholding of the curvelet decomposition, the problem of parasitic motion is solved. On the other hand, the order of magnitude is underestimated. Using hard thresholding with the coarsest scale coefficients set to zero, the problem of order of magnitude is solved, although the problem of parasitic motion arises again with less significance. Using scale by scale thresholding of the curvelet decomposition, the main problems (parasitic motion, underestimation of order of magnitude) encountered with other operators are solved. The result of this set of experiments illustrates the importance of an adequate observation operator in direct image assimilation.

6.2 Observation operators based on finite-time Lyapunov exponents and vectors computation

Ocean tracer images (Sea Surface Temperature and Ocean Color for instance) show patterns, like fronts and filaments, that characterize the flow dynamics. They are closely related to the underlying flow dynamics and are referred to as *Lagrangian Coherent Structures* (LCS). ~~Their location and shape are the signature of integrated dynamic information that should be exploited in a data assimilation scheme. For that, one needs to quantify the relation between the fluid flow and these patterns. First, Haller and Yuan (2000) defines an LCS as a material curve (more precisely a material surface in an extended phase space)~~ They are material curve which exhibits locally the strongest attraction, repulsion or shearing in the flow over a finite-time interval ~~. A rigorous mathematical theory that fits with this physical concept was recently developed in Haller (2011) where quantitative and robust criteria are given to identify hyperbolic (i. e. repelling and attracting) LCSs . However, and despite some caveats, LCSs~~ (Haller and Yuan, 2000; Haller, 2011) . Their location and shape are the signature of integrated dynamic information that should be exploited in a data assimilation scheme. LCSs are usually identified in a practical manner as maximizing ridges in Finite-Time Lyapunov Exponents (FTLE) fields (Haller, 2001). FTLE is a scalar local notion that represents the rate of separation of initially neighboring particles over a finite-time window $[t_0, t_0+T]$, $T \neq 0$. It is defined as the largest eigenvalue of the Cauchy–Green strain tensor of the flow map. The corresponding eigenvector is called the Finite-Time Lyapunov Vector (FTLV). Let $\boldsymbol{X}(t) = \boldsymbol{X}(t; \boldsymbol{X}_0, t_0)$ be the position of a Lagrangian particle at time t, which started at $\boldsymbol{X}_0$ at $t = t_0$ and was advected by the time-dependent fluid flow ~~$\boldsymbol{U}(\boldsymbol{X},t)$, $\boldsymbol{X} \in \Omega$,~~ $\mathbf{U}(\boldsymbol{X},t)$, $t \in [t_0, t_0+T]$. An infinitesimal perturbation $\delta\boldsymbol{X}(t)$ started at $t = t_0$ from $\delta_0 = \delta\boldsymbol{X}(t_0)$ around $\boldsymbol{X}_0$ then satisfies, for all $t \in [t_0, t_0+T]$,

$$\begin{cases} \cancel{\dfrac{D\delta\boldsymbol{X}(t)}{Dt} = \nabla\boldsymbol{U}(\boldsymbol{X}(t),t).\delta\boldsymbol{X}(t),} \\ \cancel{\delta\boldsymbol{X}(t_0) = \delta_0, \boldsymbol{X}(t_0) = \boldsymbol{X}_0.} \end{cases}$$

$$\frac{D\delta\boldsymbol{X}(t)}{Dt} = \nabla\mathbf{U}(\boldsymbol{X}(t),t).\delta\boldsymbol{X}(t), \tag{27}$$

$$\delta\boldsymbol{X}(t_0) = \delta_0, \boldsymbol{X}(t_0) = \boldsymbol{X}_0. \tag{28}$$

Let $\lambda_{\max}$ be the largest eigenvalue of the Cauchy–Green strain tensor

$$\Delta = \left[\nabla\phi_{t_0}^{t_0+T}(\boldsymbol{X}_0)\right]^* \left[\nabla\phi_{t_0}^{t_0+T}(\boldsymbol{X}_0)\right], \tag{29}$$

where $\phi_{t_0}^{t} : \boldsymbol{X}_0 \mapsto \boldsymbol{X}(t; \boldsymbol{X}_0, t_0)$ represents the flow map of the system (it links the location $\boldsymbol{X}_0$ of a Lagrangian particle at $t = t_0$ to its position $\boldsymbol{X}(t; \boldsymbol{X}_0, t_0)$ at time $t \neq t_0$). The *forward* FTLE at the point $\boldsymbol{X}_0 \in \Omega$ and for an advection time T starting at $t = t_0$ is defined as

$$\sigma_{t_0}^{t_0+T}(\boldsymbol{X}_0) = \frac{1}{|T|}\ln\sqrt{\lambda_{\max}(\Delta)}. \tag{30}$$

FTLV is the eigenvector associated with $\lambda_{\max}$. The FTLE thus represents the growth factor of the norm of the perturbation $\delta\boldsymbol{X}_0$ started around $\boldsymbol{X}_0$ and advected by the flow after the finite advection time T. Maximal stretching occurs when $\delta\boldsymbol{X}_0$ is aligned with the FTLV. *Backward* FTLE and FTLV (BFTLE and BFTLV) are similarly defined, with the time direction being inverted, in Eq. (27).

BFTLE (BFTLV) is a scalar (vector) that is computed at a given location $\boldsymbol{X}_0$. Seeding a domain with particles initially located on a grid leads to the computation of discretized scalar (BFTLE) and vector (BFTLV) fields. Ridges of the BFTLE field approximate attracting LCSs (Haller, 2011). An example of a BFTLE and corresponding BFTLV orientation maps, computed from a mesoscale (1/4°) time-dependent *surface* velocity field coming from a simulation of the North-Atlantic ocean, is given in Fig. 9. The BFTLE field shows contours that correspond reasonably well to the main structures such as filaments, fronts and spirals that appear in the Sea Surface Temperature (SST) field of the same simulation (see Fig. 10 left panel). Note that this field can be distinguished by spatial observations. Also the BFTLVs align with the gradients of this tracer field: Figs. 9 and 10 (right panels) show that BFTLVs and SST gradients have similar orientations. These similarities illustrate the strong link between the tracer field patterns and the underlying flow dynamics. ~~In order to exploit the properties of~~ BFTLE and BFTLV ~~in a~~ properties may thus be exploited to identify appropriate structure space to be used in a direct image assimilation framework ~~, one needs to quantify this link by identifying the appropriate structure space.~~ (Titaud et al., 2011) .

~~The~~ First, thanks to its almost-lagrangian nature ~~of FTLE (Lekien et al., 2005) permits the interpretation of the BFTLE field as a tracer field which in turn~~ (Lekien et al., 2005) BFTLE field can be considered as an image ~~. In other words, the BFTLE can be used to define a~~ of tracer field. Then $\mathcal{H}_{\mathcal{X}\to\mathcal{F}} : \mathbf{U} \mapsto \mathrm{BFTLE}(\mathbf{U})$ defines a model-to-image operator ~~. This thus makes possible the comparison between BFTLE and the corresponding ocean tracer field in the *structure space* once the~~ which is composed with an image-to-structure operator ~~has been defined. Note that this model-to-image operator~~ $\mathcal{H}_{\mathcal{F}\to\mathcal{S}}$:

$$\mathcal{H}_{\mathcal{X}\to\mathcal{S}}(\mathbf{U}) = \mathcal{H}_{\mathcal{F}\to\mathcal{S}}(\mathrm{BFTLE}(\mathbf{U})). \tag{31}$$

Note that BFTLE produces images with stronger discontinuities than operators based on passive tracer advection: in this later case, the numerical diffusion softens the discontinuities which makes the comparison with high-resolution satellite images less accurate.

~~The~~ Secondly, the alignment of the BFTLV with the tracer gradients ~~allows the direct link of the structured information contained in the image with the flow dynamics: if~~ (Lapeyre, 2002; d'Ovidio et al., 2009) allows the construction of a *strict* model-to-structure operator: structures are identified as the orientation of the gradient of the image ~~, then~~ and the observation operator is ~~a strict model-to-structure operator: it is not defined as a composition of a model-to-image operator with an image-to-structure operator.~~ simply defined as

$$\mathcal{H}_{\mathcal{X}\to\mathcal{S}}(\mathbf{U}) = \mathrm{BFTLV}(\mathbf{U}). \qquad (32)$$

~~Using observation~~ Several studies showed that model-to-structure operators based on BFTLE-V ~~computation supposes that the corresponding image cost functions Eq. (16)~~ are sensitive to perturbations ~~on the control variable $\boldsymbol{X}_0$. We also expect that the cost functions admit a minimum value at the reference (i. e. non-perturbed) state. Such prerequisites have been verified in Titaud et al. (2011) on simulated ocean tracer fields (SST and Mixed Layer Phytoplankton). Figure 11 shows the behaviour of the image cost functions, for each BFTLE and BFTLVbased observation operator,~~ of the velocity field in a direct image assimilation framework: Fig. 11 left panel (resp. right panel) shows a set of misfits in the image ridges space (resp. in the image gradient orientation space) bewteen BFTLE (resp. BFTLV) and SST fields with respect to the amplitude of ~~nine random perturbationsapplied to a reference velocity field. The left panel corresponds to a BFTLE-based observation operator that maps the control variable (time-dependent surface velocity field) to the space of binary images where the image ridges are modelled (see Fig. 12). The right panel corresponds to a BFTLV-based observation operator: the misfit is defined as an angular error. These results clearly indicates that tracer images can be exploited to reverse a velocity field using a~~ random perturbations: each misfit admits a unique minimum close to the non-perturbed state. Moreover BFTLV shows a more robust behavior than BFTLE: misfts are smoother and minima are identical. These studies clearly illustrate the feasability of the use of such operators in direct image assimilation ~~scheme and~~ to control surface velocity fields. For more details about theoretical framework and experimental setup about the use of BFTLE-V ~~computations. See Titaud et al. (2011) for more details on the experimental setup and the analysis of the results.~~ as model-to-structure operator see Gaultier et al. (2013); Titaud et al. (2011) .

7 Conclusions

Data assimilation is the science of coupling heterogeneous information coming from different sources: model, statistics, observations. During the last two decades, data assimilation has shown a dramatic development, mainly in meteorology and oceanography. It is beginning to be used in many other fields like agronomy, economy or medicine. Data assimilation is a universal problem if we want to understand and predict the evolution of a system governed by a corpus of deterministic or random equations. This is especially true if, in reality, any realization of the system is unique. More and more, information is available as images or image sequences of the observed system. Their dynamics often permit a better understanding of the system. However the information contained in images is still mainly used in a qualitative way by experts of the application domain.

In this paper, we described two frameworks where data assimilation schemes can deal with image information. First, images and sequence of images may be post-processed in order to extract some indirect (pseudo) observations that are related to the state variables of the model. The most common example is the motion vector field which can be inferred using motion estimation techniques. The result of post-processing is then used as a conventional observation in the data assimilation scheme. This approach has several limitations which should be overcome by the Direct Image Assimilation approach. In this framework, we consider the image or the image sequence as regular observations which must be linked to the control variable using an appropriate observation operator. For dynamical systems, the pertinent information that should be observed is brought by the structures of the image (e.g. the discontinuities). The observation operator must then map the control space into the image structure space. We show some examples of direct assimilation techniques. The corresponding results are very encouraging.

However, we are still far from an operational use of the assimilation of images. We need to keep in mind that almost two decades were necessary to make variational data assimilation operational in the primary meteorological centers worldwide. Many questions and difficulties remain both from the theoretical and practical points of view:

1. What are the most adapted structure spaces defining images? From the computational point of view, images have to live in a reduced space with respect to the trivial definition as an ensemble of pixels.

2. What topology should be used in the space of images? In this paper we have used $\mathcal{L}^2$ type metrics which tend to regularize the estimated control variable. We have to keep in mind that the information in images is borne by their singularities, so that other metrics, such as $\mathcal{L}^1$, have to be considered.

3. How to use images to guide nesting of models?

Outside of geophysics, there are many fields of application: aeronautics, especially for non-stationary flows, medicine and other fields for which images are an important source of information.

Acknowledgements. The Authors thank A. Vidard for helpful discussions that led to this paper. ~~The authors~~ F.-X. Le Dimet and I. Souopgui ~~thank~~would like to express their thanks and gratitude for encouragement and support to Doctor Hussaini, Sir James Lighthill Distinguished Professor, ~~of~~ Florida State University, Tallahassee, Florida, USA ~~for the encouragement and support~~ . The Authors are also thankful to O. Talagrand, three reviewers and the Editor for providing helpful comments that resulted in substantial improvements of the manuscript. This work was partly supported by Région Rhône-Alpes and Agence Nationale de la Recherche, France. V. Shutyaev acknowledges the support by the Russian Science Foundation (project 14-11-00609).

References

Adrian, R. J.: Particle imaging techniques for experimental fluid mechanics, Annu. Rev. Fluid Mech., 23, 261–304, 1991.

Alvarez, L., Esclarin, J., Lefbure, M., and Snchez, J.: A PDE model for computing the optical flow, in: XVI congreso de ecuaciones diferenciales y aplicaciones (C.E.D.Y.A. XVI), Las Palmas de Gran Canaria, 21–24 September, 1349–1356, 1999.

Béréziat, D. and Herlin, I.: Solving ill-posed Image Processing problems using data assimilation, Numer. Algorithms, 56, 219–252, 2011.

Black, M. J. and Anandan, P.: Robust dynamic motion estimation over time, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, 3–6 June, IEEE Computer Society Press, 296–302, 1991.

Brewster, K. A.: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part I: Method description and simulation testing, Mon. Weather Rev., 131, 480–492, 2003a.

Brewster, K. A.: Phase-correcting data assimilation and application to storm-scale numerical weather prediction. Part II: Application to a severe storm outbreak, Mon. Weather Rev., 131, 493–507, 2003b.

Candes, E. and Donoho, D.: New tight frames of curvelets and optimal representations of objects with piecewise-C2 singularities, Commun. Pur. Appl. Math., 57, 219–266, 2004.

Candes, E. and Donoho, D.: Continuous curvelet transform. I. Resolution of the wavefront set, Appl. Comput. Harm. A., 19, 162–197, 2005a.

Candes, E. and Donoho, D.: Continuous curvelet transform. II. Discretization and frames, Appl. Comput. Harm. A., 19, 198–222, 2005b.

Candes, E., Demanet, L., Donoho, D., and Ying, L.: Fast discrete curvelet transforms, Multiscale Model. Sim., 5, 861–899, 2006.

Chapelle, D., Fragu, M., Mallet, V., Moireau, P.:Fundamental principles of data assimilation underlying the Verdandi library: applications to biophysical model personalization within euHeart, Medical & Biological Engineering & Computing, 51, 1221-1233, 2013.

Cohen, A.: Ondelettes et traitement numérique du signal, MASSON, Paris, France, 1992.

Coifman, R. R.: Wavelet analysis and signal processing, in: Signal Processing, Part I: Signal Processing Theory, edited by: Auslander, L., Kailath, T., and Mitter, S. K., Springer-Verlag, New York, NY, 59–68, 1990.

d'Ovidio, F., Taillandier, V., Taupier-Letage, I., and Mortier, L: Lagrangian validation of the mediterranean mean dynamic topography by extraction of tracer frontal structures, Mercator Ocean Quaterly Newsletter 32, 24–32, 2009.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10143–10162, 1994.

Farge, M.: Wavelet transforms and their applications to turbulence, Annu. Rev. Fluid Mech., 24, 395–457, 1992.

Flór, J.-B. and Eames, I.: Dynamcis of monopolar vortices on a topographic beta-plane, J. Fluid Mech., 456, 353–376, 2002.

Gaultier, L., Verron, J., Brankart, J.-M., Titaud, O., and Brasseur, P.: On the inversion of submesoscale tracer fields to estimate the surface ocean circulation, Journal of Marine Systems, 26, 33–42, 2013.

Gorthi, S., Beyou, S., Corpetti, T., and Memin, E.: Multiscale weighted ensemble Kalman filter for fluid flow estimation, in: International Conference on Scale Space and Variational Methods in Computer Vision (SSVM'11), Vol. 6667, Ein-Gedi, Israel, 29 May–2 June 2011, Springer Verlag, Ein-Gedi, Israël, 749–760, 2011.

Haller, G.: Distinguished material surfaces and coherent structures in three-dimensional fluid flows, Physica D, 149, 248–277, 2001.

Haller, G.: A variational theory of hyperbolic Lagrangian Coherent Structures, Physica D, 240, 574–598, 2011.

Haller, G. and Yuan, G.: Lagrangian coherent structures and mixing in two-dimensional turbulence, Physica D, 147, 352–370, 2000.

Herlin, I., Huot, E., Berroir, J.-P., Le Dimet, F.-X., and Korotaev, G.: Estimation of a motien field on satellite images from a simplified ocean circulation model, in: ICIP International Conference on Image Processing, Atlanta, USA, 1077–1080, 8–11 October, 2006.

Hinterberger, W., Scherzer, O., Schnörr, C., and Weickert, J.: Analysis of optical flow models in the framework of calculus of variations, Numer. Func. Anal. Opt., 23, 69–89, 2002.

Hoffman, R. N. and Grassotti, C.: A technique for assimilating SSM/I observations of marine atmospheric storms: tests with ECMWF analyses, J. Appl. Meteorol., 35, 1177–1188, 1996.

Hoffman, R. N., Liu, Z., Louis, J.-F., and Grassoti, C.: Distortion representation of forecast errors, Mon. Weather Rev., 123, 2758–2770, 1995.

Horn, B. K. and Schunck, B. G.: Determining optical flow, Artif. Intell., 17, 185–203, 1981.

Huot, E., Isambert, T., Herlin, I., Berroir, J.-P., and Korotaev, G.: Data assimilation of satellite images within an oceanographic circulation model, in: IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2, ICASSP 2006 Proceedings, Toulouse, France, 265–268, 14–19 May, 2006.

Korotaev, G., Huot, E., Le Dimet, F.-X., Herlin, I., Stanichny, S., and Solovyev, D.: Analysis of the black sea surface currents retrieved from space imagery, in: Rapport du 38ème Congrès de la Commission Internationale pour l'Exploration Scientifique de la mer Méditerranée, Vol. 38, Commission Internationale pour l'Eploration Scientifique de la mer Mediterranee, Istanbul, Turkey, 2007.

Lapeyre, G.: Characterization of finite-time lyapunov exponent and vectors in two-dimensional turbulences, Chaos, 12(3), 688–698, 2002.

Le Dimet, F.-X. and Talagrand, O.: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus A, 38, 97–110, 1986.

Lekien, F., Coulliette, C., Mariano, A. J., Ryan, E. H., Shay, L. K., Haller, G., and Marsden, J.: Pollution release tied to invariant manifolds: a case study for the coast of Florida, Physica D, 210, 1–20, 2005.

Ma, J., Antoniadis, A., and Le Dimet, F.-X.: Curvelets-based multiscale detection and tracking for geophysical fluids, IEEE Transactions on Geoscience and Remote Sensing, 44, 3626-3638, 2006.

Ma, J., Hussaini, M. Y., Vasilyev, O. V., and Le Dimet, F.-X.: Multiscale geometric analysis of turbulence by curvelets, AIP, 21, 075104, 2009.

Mallat, S. G.: A theory for multiresolution signal decomposition: the wavelet representation, IEEE T. Pattern Anal., 11, 674–693, 1989.

Mémin, E. and Perez, P.: Dense estimation and object-based segmentation of the optical flow with robust techniques, IEEE T. Image Process., 7, 703–719, 1998.

Nagel, H.: Constraints for the estimation of displacement vector elds from image sequences, in: Eighth International Joint Conference on Articial Intelligence, Karlsruhe, West Germany, 8–2 August, 945–951, 1983.

Najman, L. and Talbot, H.: Mathematical Morphology, John Wiley & Sons, London, UK, 2010.

Papadakis, N. and Mémin, E.: Variational assimilation of fluid motion from image sequences, J. Imaging Sci., 1, 343–363, 2008.

Ravela, S., Emanuel, K., and McLaughlin, D.: Data assimilation by field alignment, Physica D, 230, 127–145, 2007.

Ravela, S., Marshall, J., Hill, C., Wong, A., and Stransky, S.: A realtime observatory for laboratory simulation of planetary flows, Exp. Fluids, 48, 915–925, 2010.

Schnörr, C.: Segmentation of visual motion by minimizing convex non-quadratic functionals, in: 12th IAPR International Conference on Pattern Recognition, Vol. 1, 9–13 October, IEEE Computer Society Press, Jerusalem, Israel, 661–663, 1994.

Serra, J.: Image Analysis and Mathematical Morphology: Theoretical Advances, Image Analysis and Mathematical Morphology, Academic Press, Orlando, FL, USA, 1988.

Suter, D.: Motion estimation and vector splines, in: Proc. Conf. Comp. Vision Pattern Rec, Seattle, WA USA, 21–23 June, IEEE, 939–942, 1994.

Tikhonov, A. N.: Regularization of incorrectly posed problems, Soviet Math., 4, 1624–1627, 1963.

Titaud, O., Vidard, A., Souopgui, I., and Le Dimet, F.-X.: Assimilation of image sequences in numerical models, Tellus A, 62, 30–47, 2009.

Titaud, O., Brankart, J.-M., and Verron, J.: On the use of finite-time Lyapunov exponents and vectors for direct assimilation of tracer images into ocean models, Tellus A, 65, 1038–1051, 2011.

Weickert, J. and Schnörr, C.: A theoretical framework for convex regularizers in PDE-based computation of image motion, Int. J. Comput. Vision, 45, 245–264, 2001.

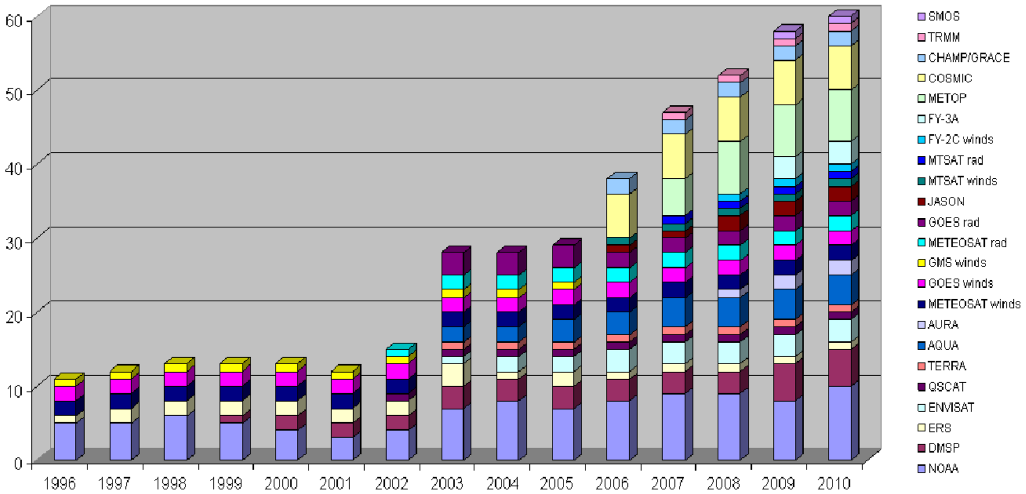

Fig. 1a. Satellite observations: evolution from 1996 to 2010 (Courtesy of ECMWF).

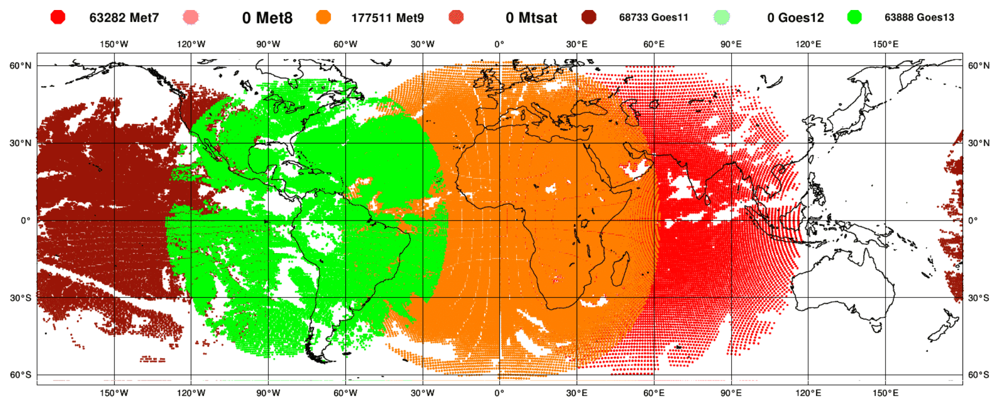

Fig. 1b. Atmospheric Motion Vector (AMV) coverage by geostationary satellites.

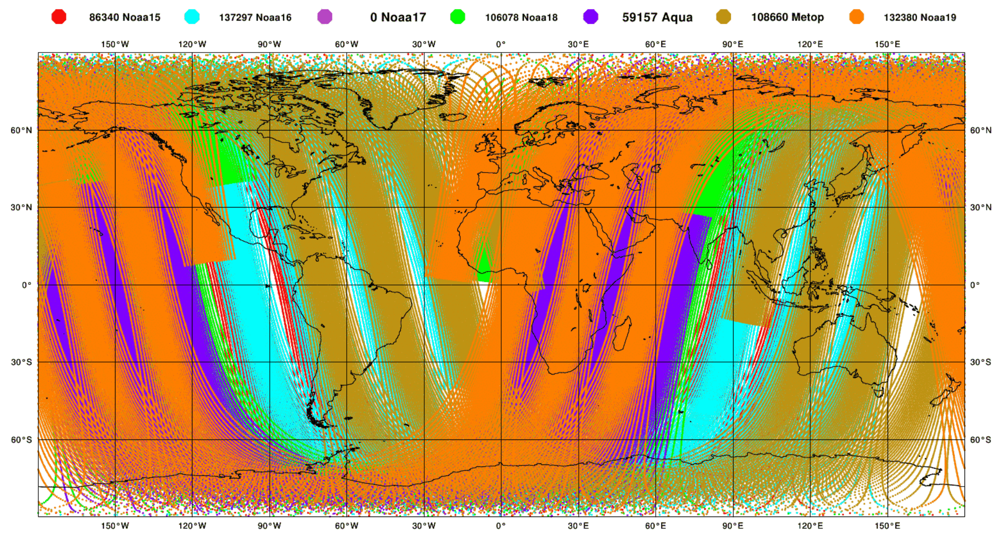

Fig. 1c. Data coverage by polar orbiting satellites equipped with AMSU-A radiometer.

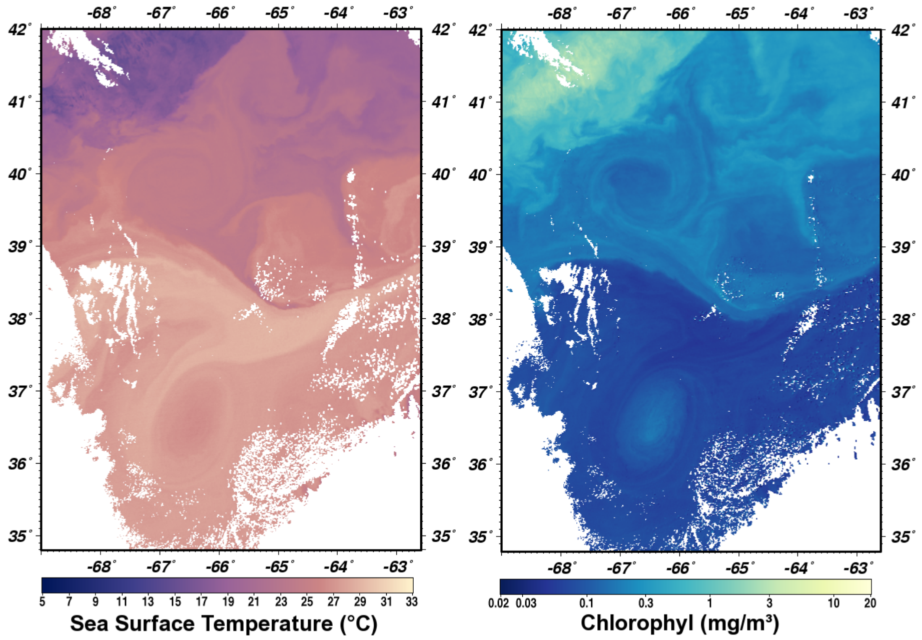

Fig. 2. Image of Sea Surface Temperature and chlorophyl (courtesy of NASA for research and educational use, http://oceancolor.gsfc.nasa.gov).

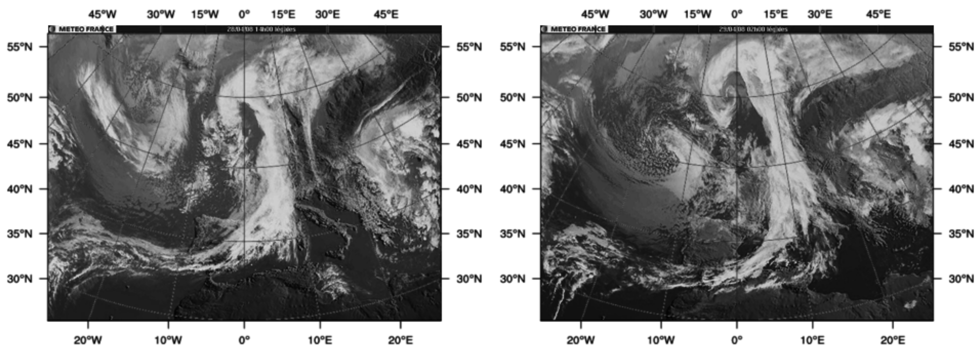

Fig. 3. Evolution of a storm on western Europe: 28 April 2008 (left) and 29 April 2008 (right).

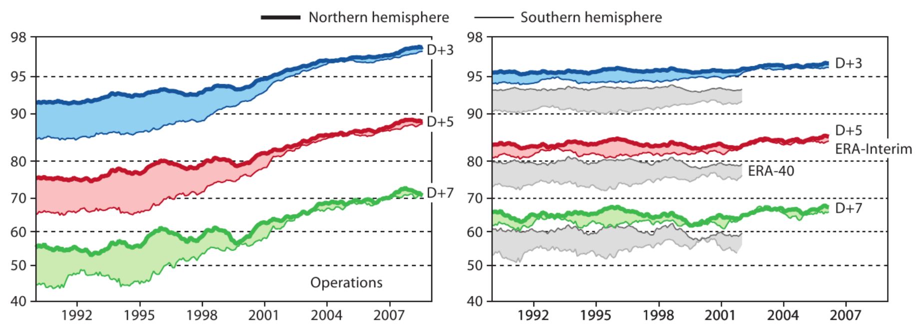

Fig. 4. Performance of the forecast: anomaly correlation at 500 hPa height forecast (courtesy of ECMWF).

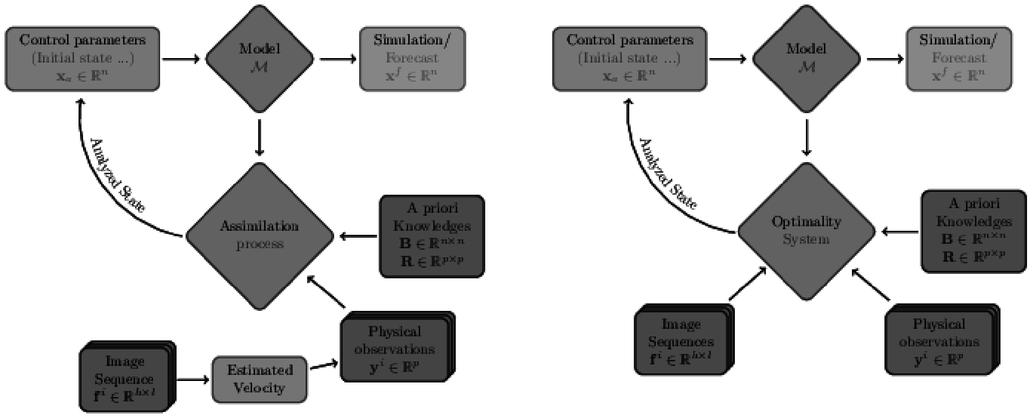

Fig. 5. Schematic representation of the use of images in data assimilation: assimilation of pseudo-observations (left); direct assimilation of images (right).

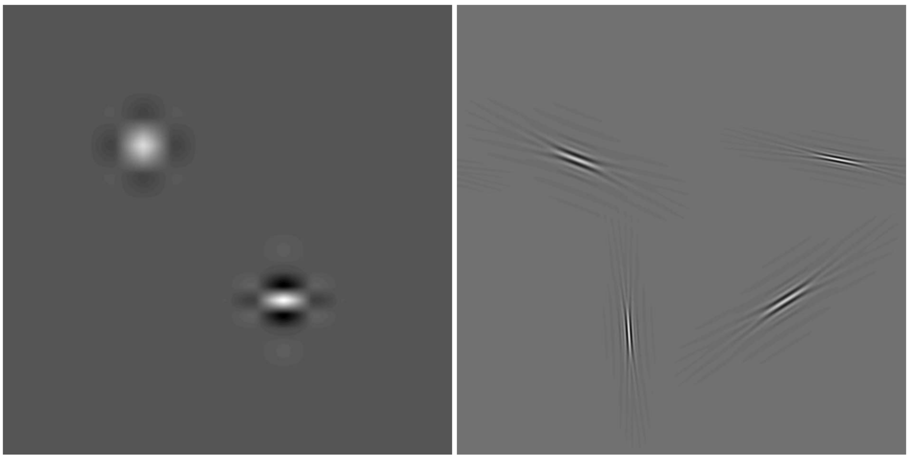

Fig. 6. Support of atoms of multiscale decomposition: wavelet (left) and curvelet (right).

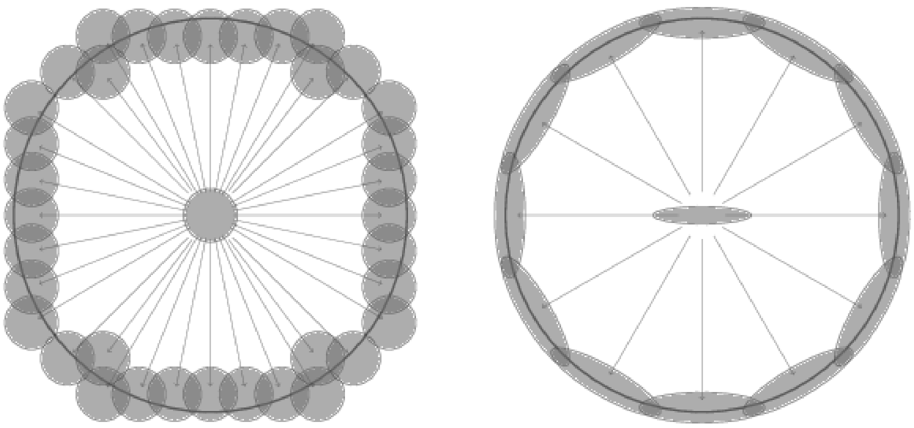

Fig. 7. Schematic view of a single scale approximation of a circle with multiscale decomposition wavelet (left) and curvelet (right).
~~Evolution of a storm on western Europe: 28 April 2008 (left) and 29 April 2008 (right).~~
~~Schematic representation of the use of images in data assimilation: assimilation of pseudo-observations (left); direct assimilation of images (right).~~

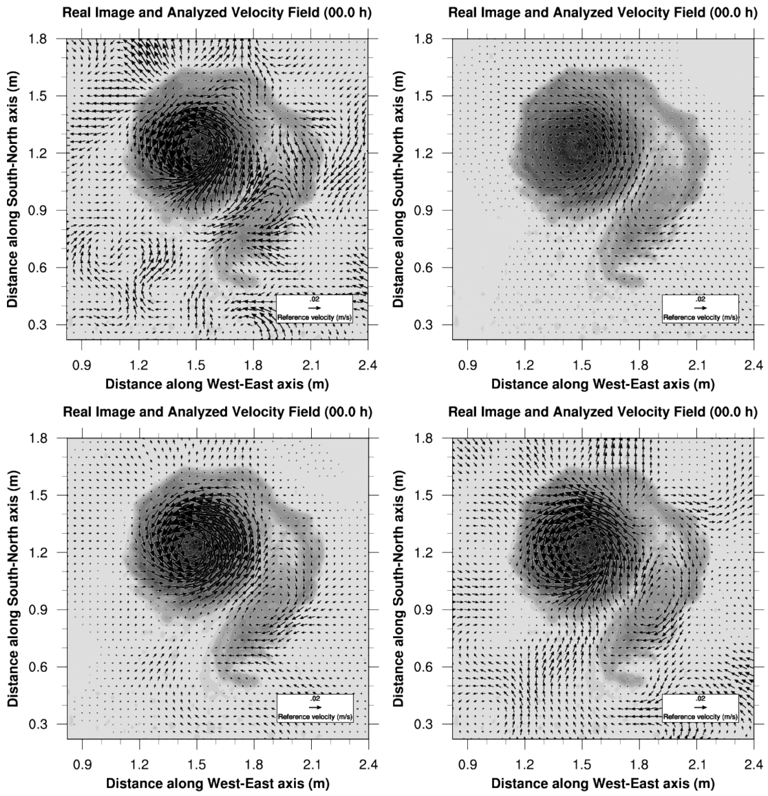

Fig. 8. Analysed initial velocity field computed by direct image sequence assimilation with different image observation operators: identity operator (top left); curvelet decomposition and hard thresholding (top right); curvelet decomposition and scale by scale thresholding (bottom left); curvelet decomposition and hard thresholding zeroing coarsest scale (bottom right).

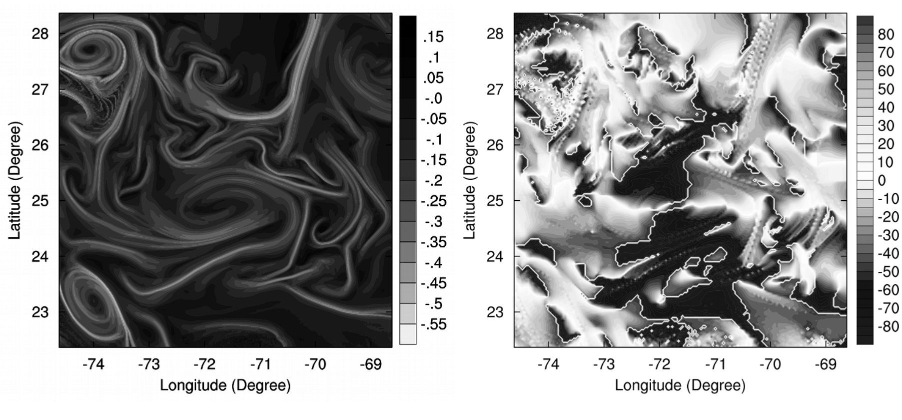

Fig. 9. Backward FTLE (day^{-1}) (left) and corresponding backward FTLV orientations (angular degree) (right) computed from the surface velocity of a simulation of the North Atlantic Ocean.

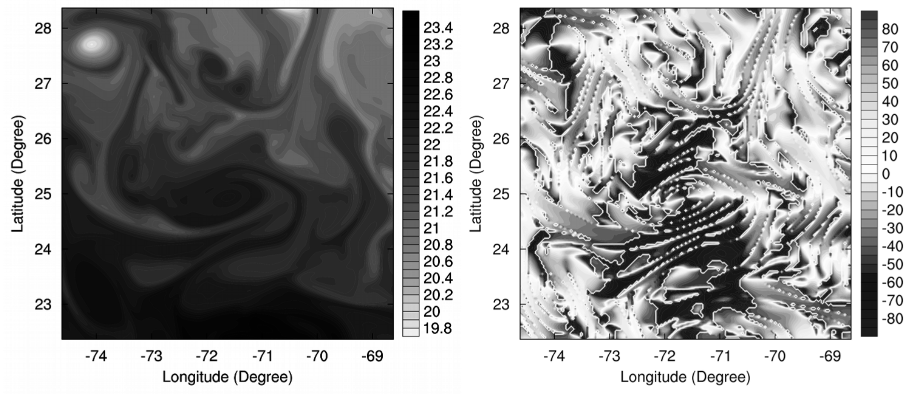

Fig. 10. Sea Surface Temperature field (left) and the corresponding orientations (angular degree) of the gradients (right). SST field comes from the same ocean simulation from which the BFTLE-Vs were computed to produce fields in Fig. 9.

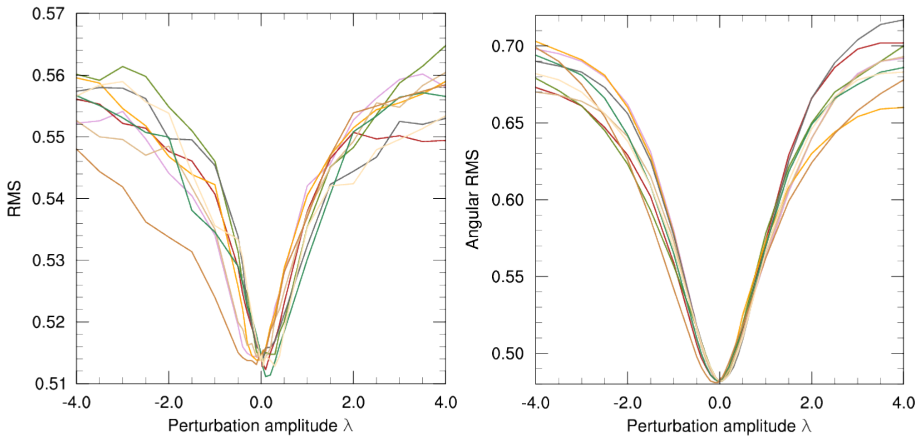

Fig. 11. Sensitivity of the misfit between BFTLEV and SST fields with respect to the amplitude of nine random perturbations applied to a reference velocity field. Left: misfit between BFTLE and SST fields computed in the space of binary images (after the application of the image-to-structure operator). Right: angular misfit between BFTLV and SST gradients.

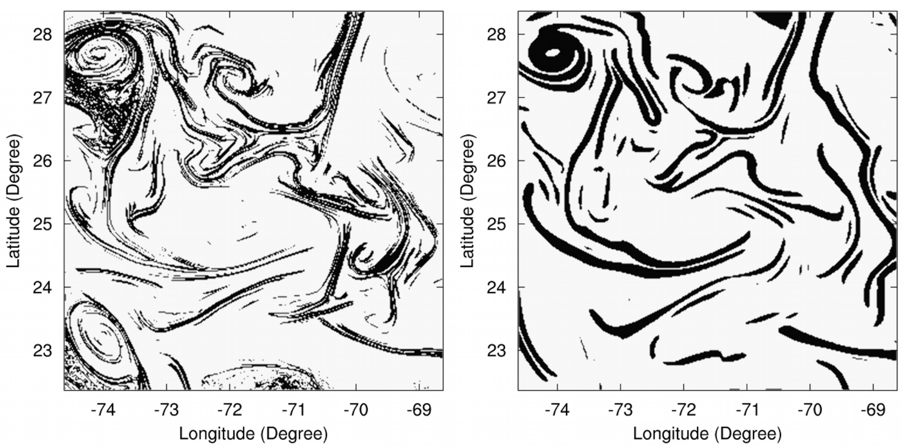

Fig. 12. Structures extraction: binarization of the FTLE (left) and SST (right) gradient fields of Figs. 9 (left) and 10 (right) using a basic threshold technique.