



Assimilation of Image Sequences in Numerical Models

Olivier Titaud, Arthur Vidard, Innocent Souopgui, François-Xavier Le Dimet

► To cite this version:

Olivier Titaud, Arthur Vidard, Innocent Souopgui, François-Xavier Le Dimet. Assimilation of Image Sequences in Numerical Models. [Research Report] RR-6701, INRIA. 2008, pp.33. inria-00332815

HAL Id: inria-00332815

<https://inria.hal.science/inria-00332815>

Submitted on 21 Oct 2008

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

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



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Assimilation of Image Sequences in Numerical Models

O. Titaud — A. Vidard — I. Souopgui — F.-X. Le Dimet

N° 6701

Octobre 2008

Thème NUM

A large blue rectangle occupies the lower half of the page. Overlaid on the left side of this rectangle is a large, light grey stylized letter 'R'. To the right of the 'R', the words 'apport de recherche' are written in a white serif font, with 'apport' on the top line and 'de recherche' on the bottom line. A horizontal grey brushstroke is positioned below the text.

*Rapport
de recherche*

Assimilation of Image Sequences in Numerical Models

O. Titaud^{*}, A. Vidard^{*}, I. Souopgui^{*}, F.-X. Le Dimet^{*}

Thème NUM — Systèmes numériques
Équipe-Projet MOISE

Rapport de recherche n° 6701 — Octobre 2008 — 30 pages

Abstract: Understanding and forecasting the evolution of geophysical fluids is a major scientific and societal challenge. Forecasting algorithms should take into account all the available informations on the considered dynamical system. The Variational Data Assimilation (VDA) technique combines in a consistent way all these informations in an Optimality System in order to reconstruct the model inputs. VDA is currently used by the major meteorological centres. During the last two decades about thirty satellites were launched to improve the knowledge of the atmosphere and of the oceans. They continuously provide a huge amount of data that are still underused by numerical forecast systems. In particular, the dynamical evolution of some meteorological or oceanic features (such as eddies, fronts, ...) that a human vision may easily detect is not optimally taken into account in realistic applications of VDA. Image Assimilation in VDA framework can be performed using *pseudo-observation* techniques : they provide some apparent velocity fields which are assimilated as classical observations. These measurements are obtained by some external procedures which are decoupled with the considered dynamical system. In this paper, we suggest a more consistent approach which directly incorporates image sequences into the Optimality System.

Key-words: Data assimilation, Inverse problem, Image sequence, Curvelets

^{*} INRIA, Lab. Jean-Kuntzmann, BP 53, 38041 Grenoble Cedex 9, France. E-mail: Francois-Xavier.Le-Dimet@imag.fr, Innocent.Souopgui@imag.fr, Olivier.Titaud@imag.fr, Arthur.Vidard@imag.fr

Assimilation de séquence d'image dans les modèles numériques

Résumé : Comprendre et prévoir l'évolution des fluides géophysiques est un grand challenge scientifique à fort impact sociétal. Les méthodes de prévision doivent prendre en compte toutes les informations disponibles sur le système dynamique considéré. Les techniques d'Assimilation Variationnelle de Données (AVD) combinent de façon cohérente toutes ces informations au sein d'un système d'optimalité afin de reconstruire les entrées du modèle. L'AVD est actuellement utilisée par les grands centres de prévision météorologiques. Durant les deux dernières décennies une trentaine de satellites ont été lancés afin d'améliorer notre connaissance de l'atmosphère et de l'océan. Ils fournissent de façon continue une grande quantité de données qui est toujours sous-utilisée par les systèmes de prévision numériques. En particulier l'évolution dynamique de certains "objets" atmosphériques ou océaniques (tourbillons, fronts, ...) que l'œil humain peut détecter facilement, n'est pas prise en compte de façon optimale en AVD. L'assimilation de séquence d'image en AVD peut être effectuée en utilisant une approche de type pseudo-observation : elles fournissent des champs de vitesse apparente qui peuvent être assimilés comme des observations classiques. Ces pseudo-mesures sont obtenues par des procédures externes qui sont découplées du système dynamique considéré. Dans ce rapport, on propose une approche plus cohérente qui incorpore les séquences d'image directement dans le système d'optimalité.

Mots-clés : Assimilation de données, problèmes inverses, séquences d'image, courbelettes

1 Introduction: images and assimilation

Understanding and forecasting the evolution of geophysical fluids (ocean, atmosphere, continental water) is a major scientific and societal challenge with important applications (prediction of extreme meteorological events, estimation of climate change, droughts and floods forecast, etc.). It is currently the subject of an intensive research effort by the international scientific community. In order to achieve this goal scientists need to take into account all the available informations on the studied dynamical system. These informations can take several forms:

- Numerical models: they attempt to simulate the system dynamics on a wide range of time and space scales. Based on the Navier-Stokes equations, they are complicated because they must deal with the irregular shape of the domain, with the parameterisation of unresolved processes, and with highly nonlinear processes at the mesoscale and below. Numerical models represents mathematical and physical informations under the form of a set of coupled sets of non linear PDEs.
- Observations: By essence models simulate simplified real physical processes by a limited number of variables. Moreover, despite their regular improvement, some model parameters are poorly known (e.g. initial and boundary conditions). This is why it is important to also observe the system. Observations are now increasingly numerous as result of both satellite and in situ observing systems. However, their accuracy is not always satisfactory, and the processing of such a large quantity of heterogeneous data is difficult. Moreover, observations provide only a partial view of reality, localised in time and space, and some of them are only indirectly related to model variables.
- Statistical informations: On the first hand, there is an *a priori* knowledge about the error of the two above sources of information: some physical processes are simplified or even neglected and errors are introduced through numerical treatments (discretisation, computing errors, etc.). Most of the time there are known deficiencies of the model and therefore information about the structure of the needed correction needed is available. On the other hand observation errors are mainly driven by so called representativity errors i.e observation can represent phenomena that are not simulated by the model (subgrid processes). Most of the time the statistics of the representativity error can be estimated. All these informations are important and should be taken into account.
- Qualitative informations: forecasters use their knowledge on the system and additional materials such as satellite images to interpret the output of numerical prediction systems and provide corrected forecasts.

Since models and observations taken separately do not allow for a deterministic reconstruction of real geophysical flows, it is necessary to use simultaneously these heterogeneous

but complementary sources of information through so called Data Assimilation methods. These are inverse modelling techniques based on the mathematical theory of statistical estimation. Their aim is to combine observations and prior model estimates, taking into account their respective statistical accuracies, in such a way that the combined estimate is more accurate than either source of information taken individually. In the early 80's Le Dimet and Talagrand (1986) proposed to use optimal control techniques in order to perform this task. This method, commonly referred as Variational Data Assimilation (VDA), has now been adopted by most of the main meteorological operational centres (European Centre for Medium Range Weather Forecast, Japan Meteorological Agency, Météo France ...). In variational data assimilation, the *analysis problem* is defined by the minimisation of a cost function that measures the statistically weighted squared differences between the observational information and their model counterpart. The cost function is minimised with respect to the control variables and this is done iteratively using a gradient-type descent method.

Models are based on fluid dynamics equations and on thermodynamic laws. After discretisation in space (finite difference and spectral methods) the model is represented by a system of PDE's with about 500 millions of variables, and this number will increase further in the future. Data are from numerous sources (in situ, radiosondes, aircraft, geostationary and polar orbiting satellites, etc.) and they are heterogeneous in density, quality and nature.

During the last two decades many satellites have been launched for the observation of the Earth in order to improve the knowledge of the atmosphere and of the oceans. They provide, among other data, photographic images of the earth system. Image-type data could be used in a classical data assimilation scheme, using the grey levels as observations of radiances emitted or reflected by the system. But connecting model variables with this kind of measurements is often complex and finally pixel levels are not assimilated in practice. This approach is also too numerically expensive. However, satellites provide sequences of images that show the dynamical evolution of some meteorological or oceanic "objects": fronts, clouds, eddies, vortices, etc. The dynamics that the human vision easily detects in this kind of image sequences has clearly a strong predictive potential. Unfortunately this information is not currently optimally used in conjunction with numerical models. Furthermore, in some other domains, such as astrophysics, images are the only source of observation.

The purpose of this paper is to present an introduction to image sequences assimilation where two classes of techniques can be considered:

- Pseudo-Observation: an apparent velocity field is estimated from an image sequence using some image processing techniques. This estimated field can then be used as pseudo-observations in a classical variational assimilation scheme.
- Direct Image Sequence Assimilation (DISA): image sequences are assimilated through an appropriate observation operator directly into the Optimality System. This can

be done on a pixel by pixel basis, alternatively some features can be extracted from the image sequence and then assimilated.

Using image sequences together with numerical models may present several difficulties:

- The state variables of the numerical models (for instance: wind, temperature, pressure, humidity in atmosphere modelling and current velocities, temperature, surface elevation, salinity in ocean modelling) are not directly measured by satellites. What is observed is more or less linked to those variables by diagnostic equations which will have to be included in the assimilation procedure.
- The physical processes that are observed are not always taken into account in the model (e.g. local convection in atmospheric modelling, ocean colour in ocean modelling). This may happen for other kind of observation as well but are likely to be systematic in the case of image sequences.
- Images are bidimensionnal informations whereas physical processes of geophysical fluids are three dimensional. From this point of view a major difference between the ocean and the atmosphere comes from their radiative properties. In the first case images give informations about the ocean surface whereas in the second case images integrate all the radiative profile of the observed atmosphere column.
- As for human vision, we can presume that the pertinent informations coming from an image is mainly brought by its discontinuities or high gradients. Unfortunately, numerical models have a tendency to smooth these discontinuities out.

The paper is organised as follow: basic principles of variational data assimilation techniques are described in section 2. Section 3 briefly presents some mathematical considerations about images. Section 4 browses the different techniques currently available to assimilate – in a regular assimilation scheme – velocity fields derived from image sequences (pseudo-observations). A new method for direct variational image sequences assimilation is described in section 5: this forms the original part of this paper. Some preliminary results are presented in section 6.

2 Variational Data Assimilation

2.1 Principle of VDA

Let us briefly present the basic principle of Variational Data Assimilation (VDA). The state of the system is described by a state variable \mathbf{x} depending on time and space ; we have $\mathbf{x}(t) \in \mathcal{X}$. It represents the variables of the model (*e.g.* velocity, temperature, elevation of the free surface, salinity, concentration of biological or chemical species. . .). The evolution

of the state variable \mathbf{x} is governed by the differential system:

$$\begin{cases} \frac{d\mathbf{x}}{dt} &= \mathcal{M}(\mathbf{x}, U), \quad t \in [0, T], \\ \mathbf{x}(0) &= V, \end{cases} \quad (1)$$

where $U \in \mathcal{P}$ is some unknown parameters of the model: boundary conditions, model errors, parametrisation of subgrid effects... U may depend on space and time. The initial condition $V \in \mathcal{X}$ is unknown and depends on space. Let us assume that U and V being given, the problem (1) has a unique solution $\mathbf{x}(U, V, t)$, $t \in [0, T]$. $\mathbf{y}(t) \in \mathcal{O}$ are given observations of the system available between 0 and T . For the sake of simplicity of this description we suppose the continuity in time of the observations and that U depends on space only. The discrepancy between the observations and the state variable intervenes in the following *cost function*:

$$J(U, V) = \frac{1}{2} \int_0^T \|\mathcal{H}[\mathbf{x}(U, V, t)] - \mathbf{y}(t)\|_{\mathcal{O}}^2 dt + \frac{1}{2} \|U - U_0\|_{\mathcal{P}}^2 + \frac{1}{2} \|V - V_0\|_{\mathcal{X}}^2, \quad (2)$$

where \mathcal{H} is the so-called *Observation operator*: it is a mapping from the space of the state variable toward the space of observations \mathcal{O} where the comparison is carried out (for more details see section 2.2). The second and the third terms are regularisation terms in Tikhonov's sense (Tikhonov, 1963). They also allow to introduce some *a priori* informations. It is important to point out that J involves three different norms that define the topological structure of the functional spaces \mathcal{X} , \mathcal{P} and \mathcal{O} ; in practice norms come from an inner product and can take into account the statistical information by introducing the error covariance matrices. For instance a discretised in space version of the cost function (2) may involve

$$\|V - V_0\|_{\mathcal{X}}^2 := (V - V_0)^* \mathbf{B}^{-1} (V - V_0) \quad (3)$$

where \mathbf{B} is the background error covariance matrix.

The problem of VDA can be considered as the determination of the *analysed state* (U^a, V^a) minimising J , that is verifying

$$(U^a, V^a) = \operatorname{argmin} J(U, V). \quad (4)$$

A necessary condition for (U^a, V^a) to be a solution of (4) is that

$$\nabla J(U^a, V^a) = 0 \quad (5)$$

which becomes a sufficient condition if J is convex. As a first approximation one has to solve a problem of unconstrained optimisation despite the fact that, for instance, some quantities are known to be non-negative (such as humidity, salinity or concentration). From a practical and numerical point of view (U^a, V^a) is estimated by a descent-type algorithm *i.e* by finding the limit of the sequence

$$\begin{pmatrix} U_{k+1} \\ V_{k+1} \end{pmatrix} = \begin{pmatrix} U_k \\ V_k \end{pmatrix} + \lambda_k D_k, \quad (6)$$

where λ_k is the step size realizing the minimum of J along the direction of descent D_k deduced from the gradient of J . For computing the gradient ∇J with respect to the control variables (U, V) , one introduces the *adjoint variable* \mathbf{p} as the solution of the adjoint model:

$$\begin{cases} \frac{d\mathbf{p}}{dt} + \left[\frac{\partial \mathcal{M}}{\partial \mathbf{x}} \right]^* \cdot \mathbf{p} = \left[\frac{\partial \mathcal{H}}{\partial \mathbf{x}} \right]^* \cdot (\mathcal{H}(\mathbf{x}) - \mathbf{y}), & t \in [T, 0], \\ \mathbf{p}(T) = 0. \end{cases} \quad (7)$$

After a backward integration of the adjoint model, the gradient of J is given by:

$$\nabla J = \begin{pmatrix} \nabla_U J \\ \nabla_V J \end{pmatrix} = \begin{bmatrix} -\int_0^T \left[\frac{\partial \mathcal{M}}{\partial U} \right]^* \cdot \mathbf{p} \\ -\mathbf{p}(0) \end{bmatrix}. \quad (8)$$

The derivation of the system (7) can be found in Le Dimet and Talagrand (1986). The model (1) plus the adjoint model (7) along with the cost function (2) form the *Optimality System* (O.S.). The adjoint variable has the same dimension as the state variable of the direct model.

2.2 The observation space and operator

Measures are partial – in space and time – estimations of some continuous physical quantities. Moreover observations that are actually used in data assimilation schemes are in general a sampling of filtered measures. The space of observations \mathcal{O} is a subset of \mathbb{R}^p where p is the number of observations. Observations \mathbf{y} and state variables \mathbf{x} are linked through the *observation operator*

$$\begin{aligned} \mathcal{H} : \mathcal{X} &\rightarrow \mathcal{O} \\ \mathbf{x}(\mathbf{t}) &\mapsto \mathbf{y}(\mathbf{t}) = \mathcal{H}(\mathbf{x}(\mathbf{t})) \end{aligned} \quad (9)$$

which maps the state variable space \mathcal{X} onto the observation space \mathcal{O} . The observation space is equipped with an inner product whose definition takes into account some uncertainties on the measurements and some representativeness errors. We have

$$\|\mathcal{H}(\mathbf{x}) - \mathbf{y}\|_{\mathcal{O}}^2 = (\mathcal{H}(\mathbf{x}) - \mathbf{y})^* \mathbf{R}^{-1} (\mathcal{H}(\mathbf{x}) - \mathbf{y}), \quad (10)$$

where \mathbf{R} is the observation error covariance matrix.

We talk about direct observations when the observed variables belong to the set of model state variables, and about indirect observations otherwise. In the first case, \mathcal{H} may define a projection or an interpolation from \mathcal{X} onto \mathcal{O} . In the second case, evaluating $\mathcal{H}(\mathbf{x})$ may be difficult from a computational and mathematical point of view because \mathbf{y} and \mathbf{x} may be connected through complex physical laws. Observation operator should be regular

enough in order to ensure that the partial derivative $\frac{\partial \mathcal{H}}{\partial \mathbf{x}}$ involved in the adjoint model (7) exists. In the case of direct observations, this property is generally satisfied which is could be far from the case for indirect observations. For instance, let us take the case of an ocean model with state variables velocity current, surface elevation, temperature and salinity. Direct observations could be some *in situ* measures of temperature and salinity in some specific locations of the studied domain. Indirect observations could be measures of surface salinity inverted from microwave radiances coming from remote sensing captors. In that case \mathcal{H} should represent the relation between these microwave radiances and surface salinity.

3 Interpretation levels and mathematical definitions of an image

Our goal is to define a methodology based on the theoretical background of variational data assimilation (see section 2) to combine in a consistent way the informations coming from image sequences, and the ones coming from numerical model, a-priori knowledges and classical observations. A first step is to find a suitable level of interpretation of the image and a suitable mathematical framework to process it: this problematic is often called *Image modelling*.

Numerical raw images The acquisition device converts the radiations coming from the observed scene into an electric signal which is then digitalised to form the *raw image*. The device contains a sensor that is composed of an array of photo-sites that accumulate photons. A numerical raw image is an array of pixels whose values represent the quantity of photons the corresponding photo-site received during the acquisition procedure. Pixels of raw images can then be considered as radiance measurements. Pixels of black and white images are described by one single channel called *grey level* whereas pixels of coloured images are described with three channels (Red, Green, Blue). The intensity of a pixel is usually described by a number between 0 (black) and 1 (white) which is stored in a binary format by a string of digits (bits) whose length, the *pixel depth*, depends on the acquisition device and digitalisation procedure. For instance, METEOSAT images has a 8 bits pixel depth (256 grey levels) as it is the case for common grey-scale images. Some other remote sensing devices manage 10 or 12 bits pixel depth images. Note that computers usually store single precision floating point numbers with 32 bits (64 bits in double precision). If we consider that a pixel intensity is a measurement, then it is represented in a lower precision than single precision floating point numbers. Finally, the *spatial resolution* of a satellite image refers to the quantity of earth's surface a single pixel covers.

However, from a *mathematical* point of view, an image is commonly considered as a real valued function of two continuous real variables. This allows the use of powerful math-

ematical tools developed by the image processing scientific community. A *numerical image* is the discrete version of a mathematical one that is finally and practically manipulated by computers.

Levels of interpretations At the *pixel level* an image of a fluid in motion is a digital instantaneous expression of the state of the fluid through the radiation that it reflects (visible light) or emits (infrared light). It is then a snapshot of an *apparent state* which may involve other physical quantities than the model simulates, like tracers (e.g. nebulosity and aerosols in the atmosphere, chlorophyll in the ocean). The radiances are generally not part of the studied system and the dependence between this radiation and the state variables is often non trivial. For instance, *nebulosity* is an intricate function of the state variable of meteorological models: clouds reflectance and intrinsic infrared emission depend both on thermodynamics (temperature and humidity) and on the micro-physical processes that occur inside the clouds (involving water, ice, snow and the size of the particles). Satellite images that come from infrared sensors (see Fig 1 right) give some indirect informations about the temperature of the observed system. At this level, information contained in images are not structured enough and represent a huge amount of data: typical satellite images are about 5000×5000 size that is 25 millions of pixels. This number should be multiplied by the number of frames in the observed sequence. For this reason, this level will probably not be efficient in realistic image sequences assimilation. Note that low level vision problematics (denoising, deblurring, inpainting, segmentation, etc.) are usually performed at the pixel level.

At the *analysis level*, an image is symbolically described by its contents, typically by the objects and their characterisation (geometry, layout, etc.) that appear in the observed scene. This kind of interpretation is used for high level vision problematics as motion estimation, edge detection, pattern recognition, etc. The advantage of this level is that it may less depend on the effective radiance emission as suggests Figure 1. In these images, visible channel (left) corresponds to a chlorophyll concentration map whereas the infrared channel (right) corresponds to a surface temperature map. But in both of them, one may detect the same structures as eddies and filaments. Currently, going from pixel level to analysis level in fluid flow images processing is a hard and still challenging task but this is beyond the topic of this paper. In the following we make use of *structure* as a generic word for the different features that may compose a geophysical image at the analysis level of interpretation.

Mathematical definition There exists several mathematical definitions for images and it is not the aim of this section to describe them all. They depend essentially on the way the image need to be processed. One can cite the multi-scale decomposition approach which comes from signal processing theory (*wavelet* (Mallat, 1998) and the more recent

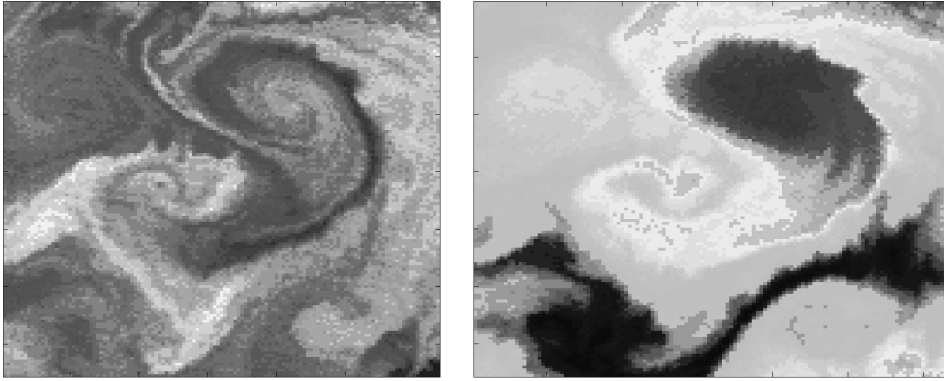


Figure 1: MODIS satellite images of the Black Sea in visible (left) and infrared (right) channels. In both image, one detects easily the same structures as eddies and filaments even though they are not linked to the same physical tracer radiance: chlorophyll concentration (left) and surface temperature (right). (Courtesy : MHI)

curvelet (Candès and Donoho, 2004) transforms). Another approach uses Partial Differential Equation framework (Aubert and Kornprobst, 2006) where the image is considered as a function of two variables verifying a PDE equation depending on the considered image processing problem. Some applications use stochastic approach as well to define images. For a good review of image analysis and the underlying mathematical definition see Chan and Shen (2005). At the end of this paper, we present an example where the anisotropic multi-scale decomposition approach by curvelet transform is used.

In the VDA framework, the level of interpretation together with the mathematical definition of the image may be represented by the image-type observation space and its corresponding observation operator (see section 5).

Sequence of images With a static image, one has only access to Eulerian informations. Sequences of images provide dynamical informations which can be interpreted as Lagrangian observations. The use of this type of information in variational data assimilation schemes was recently studied for ocean and hydrology problems (see Nodet (2005) and Honnorat (2007)). Fig 2 is an image sequence taken over the Atlantic Ocean. On this example, one can see the evolution of a front over western Europe and a cyclonic depression over the ocean. This illustrates the human vision capability for estimating the evolution of the atmosphere through the features that it is available to detect and track. Because it takes into account all the observations at once in the assimilation window, the VDA framework seems to be naturally well adapted to consistently process the dynamical information the image sequences provide. Processing and analysing image sequences may be done by extending the concepts listed above for static images using two different ap-

proaches: first, the sequence may be assimilated as a set of single two dimensional images (i.e. the frames) which are processed and analysed individually. An other approach is to consider the sequence as a three dimensional data : two spatial dimensions and time: the use of some extension of image processing techniques to the 3D case may reinforce the temporal consistency of the direct image sequences assimilation.

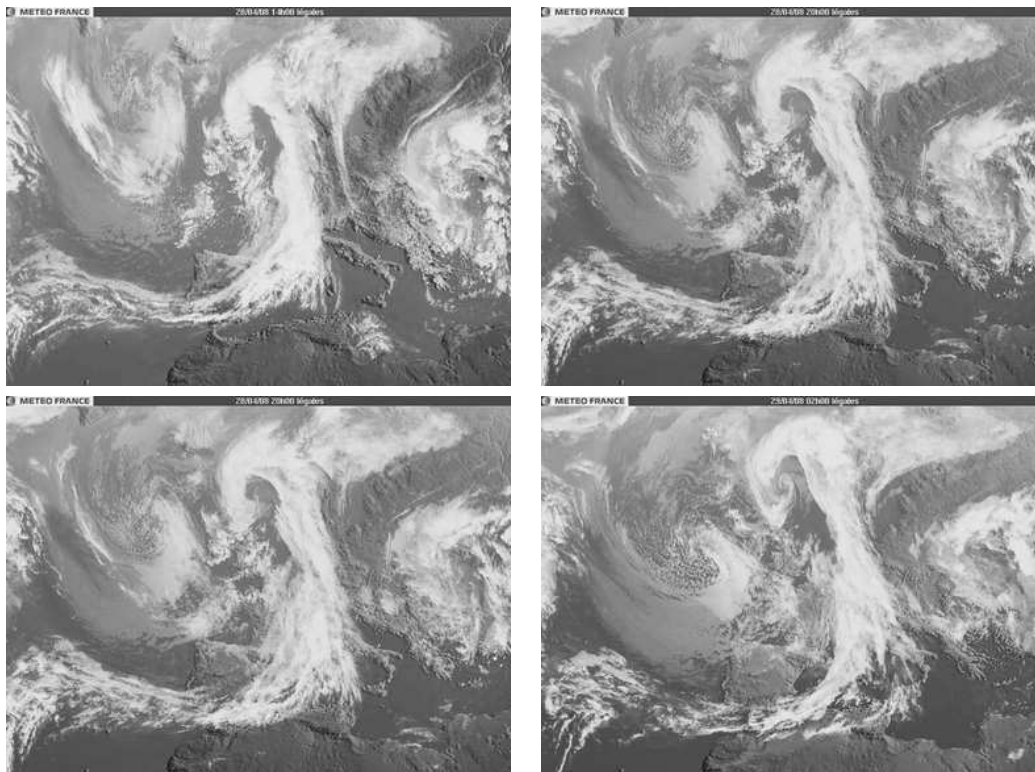


Figure 2: sequence of cloud coverage images above Europe from Meteosat (source Météo-France) from 28/04/2008 to 29/04/2008

4 Images as source of pseudo observation for data assimilation

Sequence of images are widely available in meteorology, oceanography, hydrology, astrophysics, and even medicine. Historically they have been mainly used for a "by eye" analysis from experts in the field. However since the mid 80' researches have been carried out to derive velocity fields from the sequences, with applications mainly for fluid dynamics (and very recently for movie compression and medical applications). One can use the velocity

field derived from the image processing techniques as pseudo-observation of the studied system, for instance in a regular VDA scheme (Herlin et al., 2004).

There are several ways to extract a velocity field from a sequence of image. They can be divided into two categories: the frame-to-frame estimators and the so-called Image Model (see section 4.2).

4.1 Frame-to-frame Motion Estimator

The first class of Motion Estimator, directly based on image processing techniques, aims at estimating the velocity field that would 'transport' one frame to another.

Among them one can cite the statistical methods (mostly based on correlations between 2 images). They are mainly used for PIV (Particle Image Velocities) and PTV (Particle Tracking Velocities) experiments i.e. for lab experiments in fluid mechanics where particles are added to the fluid in order to follow the flow (Adrian, 1991). They are easy to implement, but they can be quite expensive when the number of particles increases. Applications of such methods are not really in the scope of what we are studying here.

Variational methods are more adapted to assimilation of image sequences in geophysics. The most common one is usually called optical-flow; This classical approach in computer vision is based on the conservation of the global luminance between two images (Horn and Schunck, 1981): let $I : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be the luminance of the pixel, the optical flow is the vector $\mathbf{v}(x, y)$ verifying

$$\frac{dI}{dt} = \frac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} = 0. \quad (11)$$

This is not the only law of conservation which could be considered according to the nature of the image: with an image of the colour of the ocean an equation of conservation of chlorophyll (with source and sink) terms could be considered, with an image of Sea Surface Temperature (SST), the Boussinesq approximation could be used...

Problem (11) is an ill-posed problem therefore one need to add a regularisation constraint; the choice of the regularisation term depends on the application field (Vigan et al. (2000) for oceanography, Amodei and Benbourhim (1991) for wind field). This problem can be summarised by finding \mathbf{v}^* that minimises F defined by

$$F(\mathbf{v}) = \frac{1}{2} \int_{\Omega} \left\| \frac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} \right\|^2 dx dy + \frac{1}{2} \lambda \|S(\mathbf{v})\|^2, \quad (12)$$

where $S : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a spatial regularisation function (see Aurox and Fehrenbach (2008)) and $\lambda \in \mathbb{R}$ the regularisation factor. The minimisation problem (12) is treated either by optimal control or by vector spline (Amodei (1993), Suter (1994), Isambert et al. (2007)).

4.2 Image Model

Recently Herlin et al. (2006), Huot et al. (2006) and Korotaev et al. (2007) proposed an extension of the optical flow methods called "Image Model". This is a more advanced

motion estimator using techniques based on VDA that allows the use of several successive frames at once by adding a governing equation for \mathbf{v} to the optimality system that generates the pseudo observations.

The aim is to retrieve a 'continuous' in time motion vector field from a sequence of image $I_{obs}(t)$, $t \in [0, T]$. As in 4.1 it assumes the conservation of light and the image pixels are transported from one image to the other by the velocity field to be estimated. In order to improve the temporal consistency of the estimated velocities, this method requires a simple temporal evolution model \mathcal{N} , along with the transport of the pixels it forms the so-called Image Model:

$$\begin{cases} \frac{\partial I}{\partial t} + \nabla I \cdot \mathbf{v} = 0, \\ \frac{\partial \mathbf{v}}{\partial t} = \mathcal{N}(\mathbf{v}), \quad I(0) = I_0, \quad \mathbf{v}(0) = \mathbf{v}_0. \end{cases} \quad (13)$$

The pixels value of the images I_{obs} is then assimilated in the previous model using variational data assimilation (see 2.1) where the control vector is the initial condition \mathbf{v}_0 of \mathcal{N} .

$$F(\mathbf{v}_0) = \frac{1}{2} \int_0^T \|I - I_{obs}\|^2 dt + \frac{\lambda}{2} \|S(\mathbf{v}_0)\|^2. \quad (14)$$

The advantage of this method compared to the one described in 4.1 is that it takes into account all the images of the sequence at once and therefore can cope with missing data on some of the frames.

4.3 Limitations of the pseudo-observation technique

The velocity field produced by the motion estimation methods proposed in the two previous sections could be used as pseudo observation in a classical VDA scheme such as presented in section 2 (Herlin et al., 2004). However, despite their relative simplicity for implementation and their rapidity, pseudo-observations techniques based on motion estimation may suffer from some limitations. First, frame to frame motion estimator does not take into account a sequence of more than two images. Due to the lack of consistency in time, it cannot capture the dynamical evolution of the image during the entire assimilation window. One of the consequences is that it cannot deal with missing data which can be quite frequent in some of the applications targeted here. They can be due to a glitch in the observing system or an obstacle (*e.g.* clouds).

Image Model technique gives an answer to these two problems. However the model \mathcal{N} of the temporal evolution of the velocity only depends on \mathbf{v} and therefore is likely to be a somewhat poor representation of the real behaviour of the velocity.

Neither of these two techniques takes into account physical informations about the underlying processes observed in the images, moreover some assumptions like the conservation

of the luminance are clearly not valid for fluid flow images showing break up structures. Finally, velocity fields obtained through the pseudo-observation approach are apparent velocities that can be quite different from the actual velocities. For instance a tracer absence in a subset of an image sequence will lead to null apparent velocities in that area. Such techniques then need to provide error maps in order to account for this kind of error in the interpretation of the observation. Therefore the definition of the observation operators and corresponding observation error matrices required by the assimilation of such data is likely to be a very difficult task.

5 Direct Assimilation of Image Sequences

5.1 General Framework

In the pseudo observation approach, the information content is extracted from the images externally from the Optimality System equations (1, 2, 7). An other way to proceed is a direct assimilation of the image sequence in the O.S. This is more consistent with the VDA philosophy, in particular the information content of the images is extracted consistently with the physics of the model. This approach intrinsically avoids the problem of observation interpretation error mentioned above in the absence of tracer.

The principle is to add to the cost function a term directly linked to the images and their dynamics. Let us denote by $\mathcal{M}_t(U, V)$ the state variable $\mathbf{x}(t)$ when initial condition is $\mathbf{x}(0) = V$ and the model parameters are U . Then J would take the form:

$$J(U, V) = \underbrace{\int_0^T \|\mathbf{y} - \mathcal{H}[\mathcal{M}_t(U, V)]\|_{\mathcal{O}}^2 dt}_{\text{Classical term } J_o} + \int_0^T \|f(t) - \underbrace{\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}[\mathcal{M}_t(U, V)]}_{\substack{\text{Model to} \\ \text{Image} \\ \text{Operator}}}\|_{\mathcal{F}}^2 dt. \quad (15)$$

In this expression $f(t) \in \mathcal{F}$ is a frame of an image sequence of the observed dynamical system at time t . It belongs to the space of images \mathcal{F} equipped with a norm $\|\cdot\|_{\mathcal{F}}$. The operator $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}$ is a mapping from the space of the state variable toward the space of images: in a sense, it generates an image of the dynamical system at time t from the values of its state variable at this date: in the following $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}$ is called *Synthetic Image Operator*. Note that by the sake of simplicity, we omit in (15) the regularisation term which appears in the expression (2) of the classical VDA cost function.

Synthetic image operator works at the pixel level interpretation of the image (see section 3). Papadakis and Mémin (2008) showed on a simple case that assimilating directly pixel levels gives a better result than using an optical flow-based pseudo-observation approach. However, due to the size of satellite images (typically 5000×5000) such formalism would require the processing of a huge amount of data and would prove not feasible for realistic applications in geophysics.

At the analysis level (see section 3), an image is characterised by coherent structures (eddies, fronts, ...). If we suppose that these structures can be approximated in a functional space \mathcal{S} equipped with a suitable norm $\|\cdot\|_{\mathcal{S}}$, then the comparison between image-type observation and model output can then be done in that space and the cost function becomes:

$$J(U, V) = \int_0^T \|\mathbf{y} - \mathcal{H}[\mathcal{M}_t(U, V)]\|_{\mathcal{O}}^2 dt + \int_0^T \underbrace{\|\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}[f]\|_{\mathcal{S}}^2}_{\substack{\text{Image to} \\ \text{Structure} \\ \text{Operator}}} - \underbrace{\|\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}[\mathcal{M}_t(U, V)]\|_{\mathcal{S}}^2}_{\substack{\text{Model to} \\ \text{Structure} \\ \text{Operator}}} dt \quad (16)$$

where $\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}$ extracts structures from images (see next subsection): it represents the analysis level of the image. The operator $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}$ extracts the same type of structures from model state variables (model outputs). The norm $\|\cdot\|_{\mathcal{S}}$ depends on the mathematical definition of the image at the level \mathcal{S} . By choosing $\mathcal{S} = \mathcal{F}$ and $\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}} = f$ one go back to formulation (15).

With the usual variational approach described in section 2 a term in the form:

$$\left[\frac{\partial \mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}}{\partial \mathbf{x}} \right]^* \cdot (\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}[\mathbf{x}] - \mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}[f]) \quad (17)$$

is added in the right hand side of the adjoint model (7). In practical applications, we will then prefer working with differentiable Model to Structure operators. Note finally that one may consider assimilating several image sequences by adding corresponding terms in the DASI cost function (16).

5.2 Image to Structure operator $\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}$

As we previously said, the Image to Structure operator represents the level of analysis we consider to manipulate images. It could represent some transformations we apply to the image. The image space (in the mathematical sense) of this transformation is a space of another data type. The constraints which are imposed to choose \mathcal{S} is that it should have a structure of normed space in order to use simple rules for the differentiation of J . Moreover, its dimensionality must remain reasonable to store the extracted data.

It may exist many different approaches to define the space \mathcal{S} . For instance, one can consider

1. Frequency characteristics: in that case \mathcal{S} is the image space of a multi-scale transformation such as wavelet and curvelet transforms. This approach is interesting for its well known capability in data compression, denoising and edge extraction for example.
2. Geometric characteristics: some image analysis techniques can detect some geometrical feature in the image (e.g. by the mean of active contours). This features may take

the form of a parametrised curve in \mathbf{R}^2 which could be assimilated as Lagrangian observation. This approach will necessitate some high level image analysis.

3. Qualitative characteristics: in geophysics there are several recognisable structures connected to physical phenomena (*e.g.* cyclone, front, etc.). One could define \mathcal{S} as a dictionary of such couples observable structure / physical phenomenon.

5.3 Model to Structure operator $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}$

Features in geophysical images as eddies or filaments may correspond to some tracers in the observed system. Tracers can be defined as any quantities that are transported by the fluid flow. For instance, potential vorticity, temperature, salinity can be considered as atmospheric and ocean tracers. Modelling and simulate their evolution from the model state variables may help to define the Model to Structure operator. Indeed, one may use their description maps (*e.g.* concentration maps) as a *Synthetic Image Sequence* from which structures can be extracted using the Image to Structure operator. More precise correspondence between tracer description and measured radiances (*i.e.* pixel intensity) may also be considered by modelling the acquisition procedure. For instance, this may connect properly the three dimensional informations of the physical model with the two dimensional informations of an image. However, as suggests Figure 1, we may think that such precise modelling is not always useful. Using a tracer evolution modelling leads to a Model to Structure operator of the form

$$\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}} = \mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}} \circ \mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}, \quad (18)$$

where $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}$ is the Synthetic Image Sequence operator. Note that if tracers take part in the model the Synthetic Image operator is just some kind of projection (to be composed with an image acquisition model if necessary).

However some features of interest in geophysical fluid images do not correspond to any tracers (such as waves or dry intrusions in atmosphere images). In that case, other techniques for extracting structures from model outputs avoiding synthetic image sequence production should be developed.

6 Numerical example

6.1 Experimental framework

Laboratory experiment can be a good testbed for the new technique suggested in 5. Indeed it offers the complexity of true physical processes while keeping it manageable, and more importantly it opens the way to realistic applications with images.

As a first step to study the numerical feasibility of this new approach, one will focus on so called twin-experiments where observations do not come from real world but are

generated by the numerical model. In this configuration, the "true" state is known and this allows us to evaluate and understand the behaviour of a new method.

The experimental framework is that of the study of a drift of a vortex on a turntable. The evolution of a vortex in the atmosphere is simulated at the CORIOLIS experimental turntable (Grenoble, France) which re-create the effect of the Coriolis force on a thin layer of water. A complete rotation of the tank takes 60 seconds which corresponds to one Earth rotation. The vortex is created by stirring the water and made visible thanks to the addition of a passive tracer (fluorescein). Photographs of the vortex are taken from above the turntable which constitute the observed image sequence. For more details on the statements and the motivations of this experiment see Flór and Eames (2002).

In this configuration, the evolution of the fluid can be modelled by the state variable $\mathbf{x} = (u, v, h)$ whose components verify the Shallow-Water equations

$$\begin{cases} \partial_t u - (f + \zeta)v + \partial_x B &= -r^*u + \nu\Delta u, \\ \partial_t v + (f + \zeta)u + \partial_y B &= -r^*v + \nu\Delta v, \\ \partial_t h + \partial_x(hu) + \partial_y(hv) &= 0. \end{cases} \quad (19)$$

Unknowns are the zonal component $u(t, x, y)$ and meridional component $v(t, x, y)$ of the current velocity and the water elevation $h(t, x, y)$. They depend on time t and the two horizontal directions x and y . One defines the relative vorticity $\zeta = \partial_x v - \partial_y u$ and the Bernoulli's potential $B = g^*h + \frac{1}{2}(u^2 + v^2)$, where g^* is the reduced 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: $r^* = 0.9, 10^{-7}$, $\nu = 0$, $f_0 = 0.25$, $g^* = 0.02$ and $\beta = 0.0406$.

The simulation is performed on a rectangular domain $\Omega =]0, L[\times]0, H[$ representing a sub-domain of the turntable with $L = H = 3\text{m}$. The domain is discretised on a 128×128 uniform Arakawa C-type square grid. A finite differences scheme is used for space discretisation. Time integration is performed using Leap-Frog scheme. Time step is set to 0.05s of the turntable experiment which may correspond to 72s in the atmosphere. Image sequence acquisition is simulated using synthetic images whose grey levels correspond to the concentration of a passive tracer advected by the velocity field (see section 6.2.1).

6.2 Assimilation procedure

In our test, we consider the problem of recovering the *initial state* of the fluid $V = \mathbf{x}_0 = (u, v, h)(0, \cdot, \cdot)$ which constitute our control variable. The observations are images and we do not assimilate any measurements of some components of the state variable. We use an Images to Structure operator of the type "frequency characteristics" by the mean of a multi-scale and multi-orientation transformation of the images: the only source of information we assimilate is a subset of the coefficients of the decomposition (see below for more details). We consider an assimilation window of 100s (2000 time steps) which corresponds approximatively to 40 hours in the Earth atmosphere. Images are captured

every 5s (100 time steps) which then simulates a real capture frequency of one image per two hours. Observations then consist of a set of 20 images of 128×128 resolution.

The application of VDA to the above test case using section 5 formalism relies on the minimisation of the following cost function:

$$J(\mathbf{x}_0) = \underbrace{\alpha_o \int_0^T \|\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}[f] - \mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}[\mathcal{M}_t(\mathbf{x}_0)]\|_S^2 dt}_{J_o} + \underbrace{\alpha_b \|\mathbf{x}_0 - \mathbf{x}_b\|_{\mathcal{X}}^2}_{J_b}, \quad (20)$$

where observation and regularisation terms J_o and J_b are balanced by the mean of two adjustable coefficients whose ratio α_b/α_o may represent the confidence in the background \mathbf{x}_b . The following subsections are devoted to a detailed description of these terms, for this experiment.

6.2.1 Observation term J_o

Synthetic and true images generation As mentioned in 5.3, one way to define the Model to Structure operator $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}}$ is to first create an image from the model outputs using a Synthetic Images Operator $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}$ and then apply the Image to Structure Operator $\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}$. In the current experimental framework $\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}$ is obtained by simulating the advection of the monochromatic passive tracer into the fluid. To that purpose, we define

$$\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}}[u, v, h] = q, \quad (21)$$

where $q = q(t, x, y)$ is the passive tracer concentration verifying the following conservation equation posed in Ω and for $t \geq 0$

$$\partial_t q + u \partial_x q + v \partial_y q = 0, \quad (22)$$

where (u, v) is given by (19). One obtains the synthetic image at a given date t using the concentration map $(x, y) \mapsto q(t, x, y)$ as grey levels. True images are generated from a given true state (u_t, v_t, h_t) using the same technique, since we are in a twin experiments framework.

Image to Structure operator: Curvelet transform and Thresholding As mentioned in 5.2, one way to extract consistent structures from the images is to work in the frequency domain by applying a multi-scale transformation. In this paper we use a curvelet transform (Candès and Donoho, 2004): An image f can be decomposed as

$$f = \sum_{j,k,l} \langle \varphi_{j,k,l}, f \rangle \varphi_{j,k,l}, \quad (23)$$

where $(\varphi_{j,k,l})_{j,k,l}$ are the elements of the curvelet frame. The inner product $\langle \cdot, \cdot \rangle$ is the usual $L^2(\mathbb{R}^2)$ one. Curvelet transform is a linear isometry operator (see Appendix) and then its adjoint needed in (7) is given by the inverse transformation (reconstruction). This is a convenient feature for practical use in the VDA framework. For more details on curvelet transform see the Appendix. In order to extract the structures from the images, we need to apply a threshold on the curvelet coefficients $\langle \varphi_{j,k,l}, f \rangle$. We then define the Image to Structure operator by the threshold of some coefficients of the curvelet transform of the image:

$$\mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}}[f] = \mathcal{T}(\text{DCT}(f)), \quad (24)$$

where \mathcal{T} denotes a threshold operator applied on the curvelet coefficients of the image f computed by the Discrete Curvelet Transform (DCT) algorithm (see Appendix). By the sake of simplicity, we use in our experiments the following hard threshold function

$$\tau(x) = \begin{cases} x & \text{if } |x| \geq \sigma, \\ 0 & \text{if } |x| < \sigma, \end{cases} \quad (25)$$

where x represents a coefficient in the curvelet series (23) and $\sigma > 0$ is a given threshold value. The space \mathcal{S} is then the image space of the curvelet transform. We assimilate a fixed subset of the curvelet coefficients of the image decomposition. In our experiments, we keep only 10 percents of these coefficients. Using hard thresholding is more or less equivalent to assimilate images at the pixel level with this difference that it uses the efficient compression ability of this transformation (see Appendix): this feature is critical for future realistic applications. A large choice of threshold function is available depending on the applications. Some classical threshold functions are listed in Antoniadis (2007). More sophisticated thresholding using a rule of Total Variation (TV) minimisation can reduce the pseudo-Gibbs and element-like artifacts that appears with classical threshold functions (Durand and Froment, 2003), and may be used in this framework. DASI may give a starting point to very interesting investigations in the area of the application of curvelet and other multi-resolution transformations (Ma et al., 2008).

Model to Structure operator Finally our Model to Structure operator is the composition of the previous operator (24) with the Synthetic Image operator (21):

$$\mathcal{H}_{\mathcal{X} \rightarrow \mathcal{S}} = \mathcal{H}_{\mathcal{F} \rightarrow \mathcal{S}} \circ \mathcal{H}_{\mathcal{X} \rightarrow \mathcal{F}} = \mathcal{T} \circ \text{DCT}. \quad (26)$$

6.2.2 Regularisation term J_b

The regularisation term

$$J_b(\mathbf{x}_0) = \|\mathbf{x}_0 - \mathbf{x}_b\|_{\mathcal{X}}^2 \quad (27)$$

involves the background state \mathbf{x}_b and some nondimensionnalisation of the equations through the definition of the norm $\|\cdot\|_{\mathcal{X}}$. This norm is also used to define a pre-conditioning of the minimisation problem.

Background term In our experiments, the background is set to the constant value $\mathbf{x}_b = (u_b, v_b, h_b) = (0, 0, h_m)$ where h_m is the mean water elevation in the tank. This approximatively corresponds to a fluid at rest when the turntable speed is stationary. We then set the ratio α_b/α_o to a small value (about 10^{-6}).

Preconditioning The control state variable space \mathcal{X} is equipped with the following norm

$$\|\mathbf{x}\|_{\mathcal{X}} = \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x} \quad (28)$$

where \mathbf{B} is the so-called background error co-variances matrix. Its diagonal coefficients represent the standard deviations error on each variable with respect to itself: they are a-priori estimated from model outputs. The velocity field is regularised using a method based on the Gradient Vector Flow technique (Souopgui, 2008). The off-diagonal blocs are the cross-covariances between variables and come from balance relations between the observed variables and the others: this is a classical technique in VDA. Our observation operator connects images with the velocity field and that is why we decided to balance the water elevation h with respect to (u, v) by “inverting” the geostrophic equilibrium relation $u = -\frac{g}{f} \frac{\partial h}{\partial y}$, $v = \frac{g}{f} \frac{\partial h}{\partial x}$. Finally, we use the classical change of variable $\mathbf{x}_0' = \mathbf{B}^{-1/2} \mathbf{x}_0$ for preconditioning. For more details about this strategy see Vidard et al. (2004).

6.2.3 Optimisation

The cost function J is successively minimised on assimilation sub-windows of increasing sizes using the strategy described in Luong et al. (1998). This technique avoids some stagnation of the optimisation algorithm around a local minima which may occur when J is not convex. The minimisation procedure is initialised with the previous background. In each successive minimisation, we use the N1QN3 algorithm of the MODULOPT library (Gilbert and Lemaréchal, 1995).

6.3 Numerical Results

Following the experimental set-up described above, one try to reconstruct the true initial state assimilating only the image sequence and starting from a background at rest. The true state represents the evolution over the time window of a single vortex subject to the Coriolis force. A passive tracer is added into the vortex and its concentration is used to generate the images, only one snapshot every 100 time steps is retained as observation. Figure 3(left) shows a subset of the assimilated image sequence (*i.e.* the evolution of the passive tracer). The background state being the system at rest it does not allow for the transport of the tracer, as shown in Figure 3(centre). By comparison, figure 3(right) shows the reconstructed image sequence generated by the transport of the passive tracer by the model starting from the analysed (after assimilation) initial state. The sequence of images is relatively well reconstructed, at least at the beginning of the sequence. However, it seems

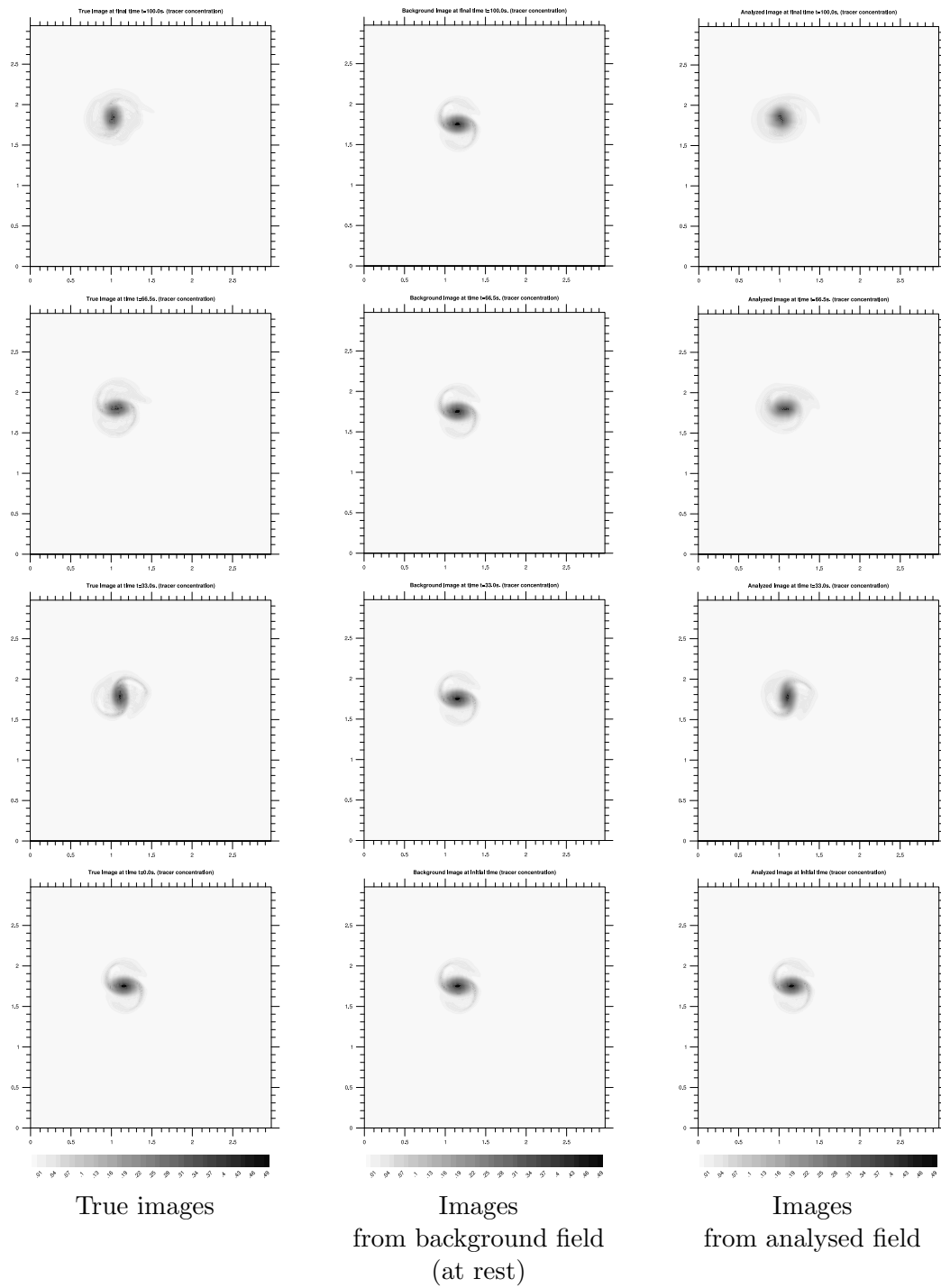


Figure 3: True images (true tracer concentration) (left) and concentration of the tracer advected by the background (centre) and by the retrieved velocity field (right) at initial time (bottom), $t=33s$ and $t=66.5s$ (centre) and $t=100s$ (top).
RR n° 6701

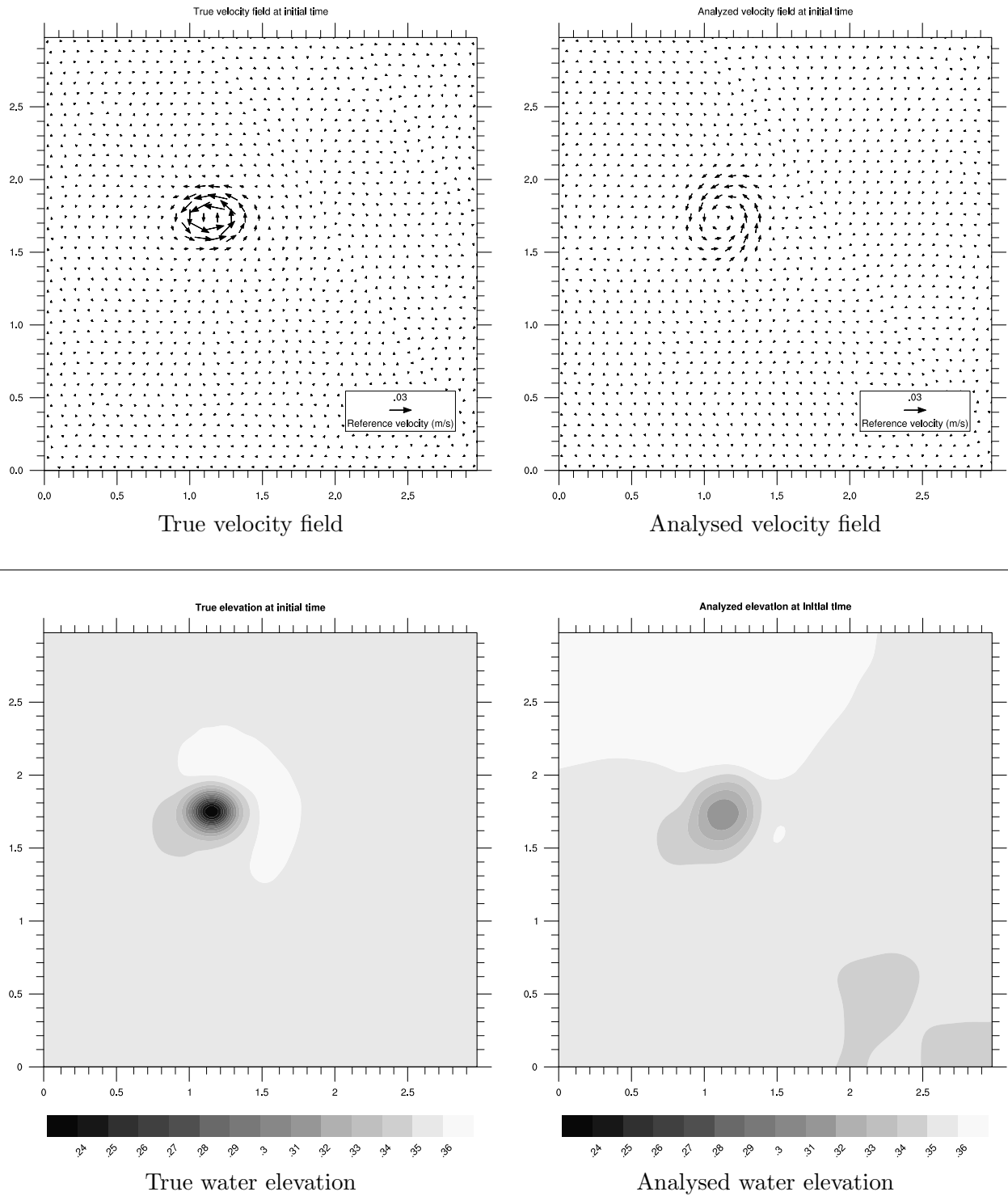


Figure 4: Initial true (left) and retrieved (right) control variables, i.e. initial velocity field (top) and initial water elevation (bottom). Background state variables correspond to a system at **INRIA**

that the estimated vortex is not sufficient to properly generate the images at the end of the sequence. Figure 4 shows the true velocity field (top left) that was used to generate the observation (left sequence of images of Figure 3) and the estimated velocity field (top right) that is retrieved thanks to the minimisation of J . It also shows the true water elevation (bottom left) and the estimated elevation retrieved at the same time (bottom right). The estimated vortex is at the right position, but its intensity is a bit weak compared to the true one, this explain why the reconstructed images at the end of the sequence of figure 3(right) are not perfect. Note that, in this example, the background \mathbf{x}_b (*i.e.* the starting point of the minimisation) was taken to be the system at rest. Additionally, the assimilated image sequence (figure 3(left)) does not show very contrasted images, these are smoother than real ones. Moreover, in realistic applications, one would be able to assimilate other kind of data. This is not a very favourable case and therefore it is encouraging to be able to recreate a vortex with about the same characteristics. A more realistic background, more contrasted images and/or a better threshold, should improve the efficiency of this test-case.

7 Conclusions

Predicting the evolution of natural dynamical systems requires to link together heterogeneous sources of information. There is a strong social demand for the improvement of the prediction of geophysical fluids evolution and VDA is a powerful tool to achieve this goal. The observation of Earth by satellites is very expensive but unfortunately dynamical information that image sequences contain are still underused by forecasting numerical systems. However they represent a large amount of structured data and it is then of great interest to suggest some way to process them in conjunction with numerical forecast systems. Some previous and still in development works could lead to the assimilation of data coming from image sequences into a VDA scheme: but this so-called pseudo-observation technique suffer from the lack of physical consistency with the observed system. The Direct Image Sequences Assimilation approach presented in this paper aims at processing the image sequences within the Optimality System of the VDA scheme: it then combines image-type information consistently with the underlying physical model. In practice, DASI is based on structures extraction from both image sequences and model outputs. One may use some already existing processing and analysing image techniques to perform the structures extraction. Features in geophysical images usually corresponds to some tracers presence in the observed system. A natural way to extract structure from the model output is then to simulate the evolution of these tracers. This approach builds a synthetic image sequence from which the structures are effectively extracted. Other techniques for extracting structures from model outputs avoiding synthetic image sequence production should be developed. Our first result shows that direct image sequence assimilation approach using structure extraction with anisotropic mutli-scale transformation of the images is achievable.

There is no doubt that investigating the way to optimally integrate image sequence in VDA framework will bring up some very interesting questions and gives some improvements in both areas of VDA and image analysis. For instance, because resolution of a satellite images is often higher than the spatial resolution of numerical models, we can expect that satellite image assimilation gives some improvements in the subgrid parametrisation problematic in VDA. DASI may also initiate some new definition of image interpretations and improve special features tracking in geophysical fluid image sequences.

Appendix: Curvelet transform

In 1999, an anisotropic geometric wavelet called *ridgelet* was proposed by Candès and Donoho (1999). It is optimal at representing straight-line singularities. However its application is limited to objects with global straight-line singularities. In order to take into account images with local line or curve singularities, the so-called first generation of *Curvelet transform* (Candès and Donoho, 2000) was developed using ridgelet transforms applied on sub-partitions of the image. A second generation of curvelet transform based on a frequency partition technique was introduced by the same authors in 2004 (Candès and Donoho, 2004). A 2D discrete curvelet transform was implemented in 2006: the Matlab and C++ source codes are freely available for non-commercial research purposes at <http://www.curvelet.org>. It also exists a parallel version of this algorithm. The following succinct description of the continuous curvelet transform is mainly inspired from Candès et al. (2006) and Ma and Plonka (2007).

We first fix some notations: Let $W : [1/2, 2] \mapsto \mathbb{R}$ and $V : [-1, 1] \mapsto \mathbb{R}$ be respectively a *radial* and an *angular* windows – in the spatial domain –, supported on $[1/2, 2]$ and $[-1, 1]$ respectively and obeying the admissibility condition $\sum_{l=-\infty}^{+\infty} V^2(t-l) = 1$, $t \in \mathbb{R}$ and $\sum_{j=-\infty}^{+\infty} W^2(2^{-j}r) = 1$, $r > 0$. Examples of such functions are given in Ma and Plonka (2007). Note that curvelet transform efficiency few depends on the choice of these functions. As in wavelet theory, we define the low-pass window $W_0(r) = 1 - \sum_{j \geq 0} W^2(2^{-j}r)$, for all $r > 0$. Finally, let us denote by \hat{f} the Fourier transform of an element f of $L^2(\mathbb{R}^2)$: it is a function of the frequency variables $\xi = (\xi_1, \xi_2)$ whose polar coordinates are denoted by (r, θ) . Curvelets are multi-scale, multi-orientation transformations with atoms $\varphi_{j,k,l}$ indexed by three integers $j \geq -1$, $k \in \mathbb{Z}$ and $l \in \mathbb{N}$ which respectively stand for *scale*, *location* and *orientation*. At the coarsest scale ($j = -1$) curvelets $\varphi_{-1,k}$ are translations of a “mother” curvelet φ_{-1} :

$$\varphi_{-1,k}(x) = \varphi_{-1}(x - k), \quad x \in \mathbb{R}^2, \quad k \in \mathbb{Z}^2. \quad (29)$$

They are then isotropic. The mother curvelet φ_{-1} is the inverse Fourier transform of the above low-pass window W_0 . We have

$$\hat{\varphi}_{-1}(\xi) = W_0(|\xi|). \quad (30)$$

For each scale 2^{-j} , $j \geq 0$, (finest scales), curvelets are defined through the inverse Fourier transform of compactly supported windows: we define a window U_j in the frequency domain for all $\xi = (\xi_1, \xi_2)$ with polar coordinates (r, θ) by

$$U_j(\xi) = 2^{-3j/4} W(2^{-j}r) V\left(\frac{2^{\lfloor j/2 \rfloor} \theta}{2\pi}\right), \quad (31)$$

where $\lfloor \cdot \rfloor$ denotes the integer part operator. The support of U_j is a polar 'wedge' determined by $\text{supp}(r \mapsto W(2^{-j}r)) = [2^{j-1}, 2^{j+1}]$ and $\text{supp}(\theta \mapsto V(2^{\lfloor j/2 \rfloor} \theta)) = [-2^{-\lfloor j/2 \rfloor}, 2^{-\lfloor j/2 \rfloor}]$. Figure 5 (left) shows the support of the first window U_0 in the frequency domain (see also Ma and Plonka (2007)). We consider a uniform discretisation of the angle coordinate $\theta \in [0, 2\pi[$

$$\theta_l = 2\pi \times l \times 2^{-\lfloor j/2 \rfloor}, \quad 0 \leq l < 2^{\lfloor j/2 \rfloor}, \quad (32)$$

and the rotation by θ radians matrix

$$R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (33)$$

Note that the number of discrete orientations θ_l depends on the scale index j : this number is doubled when scale change from an odd index $2j+1$ to the larger next even one $2j+2$. It is unchanged when scale index is changing from $2j$ to $2j+1$ (see Figure 5 (right)). Finally, for $j \geq 0$, curvelets are defined for all $x \in \mathbb{R}^2$ by

$$\varphi_{j,k,l}(x) = \varphi_j \left(R_{\theta_l}(x - x_k^{(j,l)}) \right), \quad (34)$$

where the waveform φ_j is the inverse Fourier transform of the frequency window U_j (i.e. $\hat{\varphi}_j = U_j$) and the translation vector $x_k^{(j,l)}$, $k = (k_1, k_2) \in \mathbb{Z}^2$ is defined by

$$x_k^{(j,l)} = R_{\theta_l}^{-1}(2^{-j}k_1, 2^{-j/2}k_2). \quad (35)$$

It is proved in Candès and Donoho (2004) that the above family of curvelets defines a *tight frame* (see Mallat (1998) for a detailed definition of this term): for all $f \in L^2(\mathbb{R}^2)$ we have (Parseval equality)

$$\sum_{j,k,l} |\langle f, \varphi_{j,k,l} \rangle|^2 = \|f\|_{L^2(\mathbb{R}^2)}^2, \quad (36)$$

and the reconstruction formula

$$f = \sum_{j,k,l} \langle \varphi_{j,k,l}, f \rangle \varphi_{j,k,l} \quad (37)$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product in $L^2(\mathbb{R}^2)$: $\langle f, g \rangle = \int_{\mathbb{R}^2} \bar{f}g$.

Remarks:

- Curvelet coefficients in (37) are determined by

$$c_{j,k,l} := \langle \varphi_{j,k,l}, f \rangle = \frac{1}{(2\pi)^2} \int \hat{f}(\xi) \overline{U_j(R_{\theta_l} \xi)} e^{i\langle x_k^{(j,l)}, \xi \rangle} d\xi \quad (38)$$

- In practice we use Discrete Curvelet Transform (DCT) which is designed for digital inputs f defined on a Cartesian grid $\{(n_1, n_2) \in \mathbb{N}, 0 \leq n_1, n_2 < n \in \mathbb{N}\}$. Digital coefficients are defined

$$c_{j,k,l}^D = \sum_{n_1, n_2} f(n_1, n_2) \overline{\varphi_{j,k,l}^D(n_1, n_2)} \quad (39)$$

where the construction of the digital curvelet waveforms $\varphi_{j,kl}^D$ is an adaptation of the above continuous transform for Cartesian grids. This lead to a non uniform discretisation of the orientations. For more details and C++ implementation see Candès et al. (2006).

- Let f_n^F (resp. f_n^W, f_n^C) denotes the partial reconstruction of a C^2 image away from C^2 edges (i.e. containing curve-line singularities) obtained after a hard thresholding by selecting the n largest terms in the Fourier series (resp. in an efficient wavelet basis, in the above curvelet decomposition), then we have (Candès and Donoho, 2004)

$$\|f - f_n^F\|_2^2 = \mathcal{O}(n^{-1/2}), \quad (40)$$

$$\|f - f_n^W\|_2^2 = \mathcal{O}(n^{-1}), \quad (41)$$

$$\|f - f_n^C\|_2^2 = \mathcal{O}((\log n)^3 \cdot n^{-2}), \quad (42)$$

where $\|\cdot\|_2$ denotes the $L^2(\mathbb{R}^2)$ norm. This can be explained by the oriented tight shape of the curvelets supports. Curvelet support tiling is then more efficient for curve-line singularities than wavelet supports do: this efficiency is illustrated in Figure 6.

References

- R. J. Adrian. Particle imaging techniques for experimental fluid mechanics. *Annual Review of Fluid Mechanics*, 23:261–304, 1991.
- L. Amodèi. *Étude d'une classe de fonctions splines vectorielles en vue de l'approximation d'un champ de vitesse. Application à la météorologie*. PhD thesis, Université Paul Sabatier, Toulouse, France, 1993.
- L. Amodèi and M. N. Benbourhim. A vector spline approximation. *J. Approx. Theory*, 67(1):51–79, 1991. ISSN 0021-9045. doi: 10.1016/0021-9045(91)90025-6.

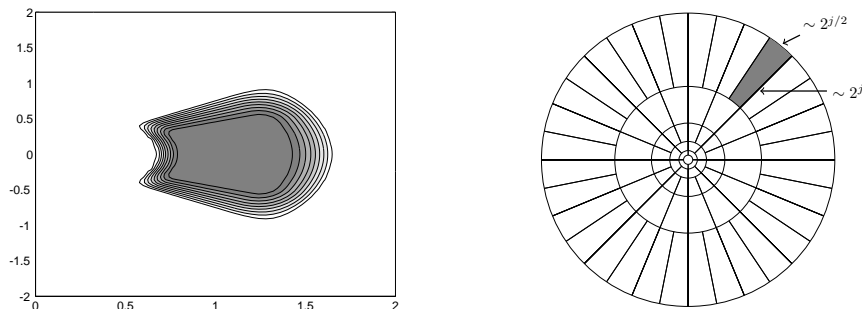


Figure 5: **Left:** support of the frequency domain window U_0 . **Right:** partition of the frequency plane: the shaded area shows an example of a “wedge” on which a curvelet is supported in the frequency plane. At scale j , this wedge has a width of order $2^{j/2}$ and a length of order 2^j .

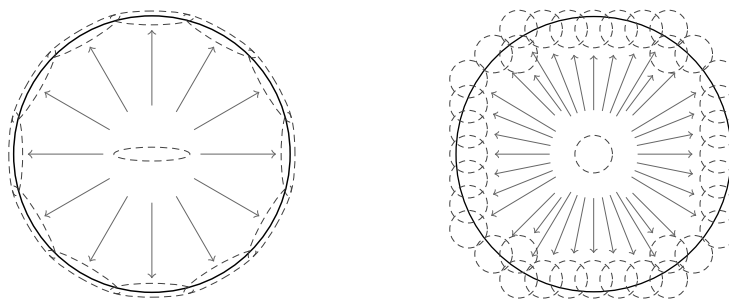


Figure 6: Reconstruction of a circle (solid) at a given scale. Curvelet supports have a tight shape and can be oriented. They fit better with curve-line singularities (left) than wavelet supports do (right). That is why curvelet transform is more efficient in terms of compression and edge enhancement for this kind of images than wavelet transform.

- Anestis Antoniadis. Wavelet methods in statistics: some recent developments and their applications. *Statistics Surveys*, 1:16–55, 2007.
- Gilles Aubert and Pierre Kornprobst. *Mathematical problems in image processing*, volume 147 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2006. ISBN 978-0387-32200-1; 0-387-32200-0.
- Didier Auroux and Jérôme Fehrenbach. Identification of velocity fields for geophysical fluids from a sequence of images. *Int. J. Computer Vision*, 2008.
- Emmanuel Candès and David Donoho. Ridgelets: A key to higher-dimensional intermitency ? *R. Soc. Lond. Philos. Trans. Ser. A.*, 357:2495–2509, 1999.
- Emmanuel Candès, Laurent Demanet, David Donoho, and Lexing Ying. Fast discrete curvelet transforms. *Multiscale Model. Simul.*, 5(3):861–899 (electronic), 2006. ISSN 1540-3459.
- Emmanuel J. Candès and David L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004. ISSN 0010-3640.
- Emmanuel J. Candès and David L. Donoho. A surprisingly effective nonadaptive representation for objects with edges. In L. Schumaker A. Cohen, C. Rabut, editor, *Curves and Surface Fitting*, pages 105–120, Saint-Malo, 2000. Vanderbilt Univ. Press, Nashville.
- Tony F. Chan and Jianhong Shen. *Image processing and analysis*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005. ISBN 0-89871-589-X.
- Sylvain Durand and Jacques Froment. Reconstruction of wavelet coefficients using total variation minimization. *SIAM Journal of Scientific Computing*, 24(5):1754–1767, 2003.
- J.-B. Flór and I. Eames. Dynamics of monopolar vortices on a topographic beta-plane. *J. Fluid Mech.*, 456:353–376, 2002.
- J.-C. Gilbert and C. Lemaréchal. The modules mlqn3 and nlqn3. version 2.0. Technical report, INRIA, June 1995.
- I. Herlin, F.-X. Le Dimet, E. Huot, and J.-P. Berroir. Coupling models and data : which possibilities for remotely-sensed images. In P. Prastacos, U. Cortés, J.-L. Diaz De León, and M. Murillo, editors, *e-Environment: Progress and Challenge of Research on Computing Science*, volume 11, pages 365–383, Mexico, November 2004. Instituto Politécnico Nacional.
- I. Herlin, E. Huot, J.-P. Berroir, F.-X. Le Dimet, and G. Korotaev. Estimation of a motion field on satellite images from a simplified ocean circulation model. In *ICIP International*

- Conference on Image Processing*, pages 1077–1080, Atlanta, USA, October, 8-11 2006. doi: 10.1109/ICIP.2006.312742.
- Marc Honnorat. *Assimilation de données Lagrangiennes pour la simulation numérique en hydraulique fluviale*. PhD thesis, INP Grenoble, 2007.
- B.K.P Horn and B.G. Schunck. Determining optical flow. *AI*, 17:185–203, August 1981.
- E. Huot, T. Isambert, I. Herlin, J.-P. Berroir, and G. Korotaev. Data assimilation of satellite images within an oceanographic circulation model. *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, 2:II–II, May 2006. ISSN 1520-6149. doi: 10.1109/ICASSP.2006.1660330.
- T. Isambert, J.-P. Berroir, and I. Herlin. Fast and stable vector spline method for fluid apparent motion estimation. *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, 2:II –505–II –508, October 2007. ISSN 1522-4880. doi: 10.1109/ICIP.2007.4379203.
- G. Korotaev, E. Huot, F.X. Le Dimet, I. Herlin, S.V. Stanichny, and D.M. Solovyev. 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*, volume 38, 2007.
- François-Xavier Le Dimet and Olivier Talagrand. Variational algorithms for analysis and assimilation of meteorological observations : Theoretical aspects. *Tellus*, 38A:97–110, 1986.
- B. Luong, J. Blum, and J. Verron. A variational method for the resolution of a data assimilation problem in oceanography. *Inverse Problems*, 14(4):979–997, 1998.
- Jianwei Ma and Gerlind Plonka. Combined curvelet shrinkage and nonlinear anisotropic diffusion. *IEEE Trans. Image Process.*, 16(9):2198–2206, 2007.
- Jianwei Ma, Olivier Titaud, Arthur Vidard, and François-Xavier Le Dimet. Spatio-temporal structure extraction and denoising of geophysical fluid image sequences using 3d curvelet transforms. Technical Report RR-6683, INRIA, 2008.
- Stephane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.
- Maëlle Nodet. *Modélisation mathématique et assimilation de données lagrangiennes pour l’océanographie*. PhD thesis, Université de Nice Sophia-Antipolis, 2005.
- N. Papadakis and E. Mémin. Variational assimilation of fluid motion from image sequences. *SIAM Journal on Imaging Science*, 2008.

- Innocent Souopgui. Vector field regularization by gradient vector flow. Technical report, INRIA, 2008.
- D. Suter. Motion estimation and vector splines. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 939–942, June 1994. doi: 10.1109/CVPR.1994.323929.
- A. N. Tikhonov. Regularization of ill-posed problems. *Dokl. Akad. Nauk SSSR*, 1963.
- Arthur Vidard, Andrea Piacentini, and François-Xavier Le Dimet. Variational data analysis with control of the forecast bias. *Tellus. Series A, Dynamic meteorology and oceanography*, 56:177–188, 2004.
- X. Vigan, C. Provost, R. Bleck, and P. Courtier. Sea surface velocity from sea surface temperature image sequences. *Journal of Geophysical Research*, 2000. Estimation du mouvement dans les images.



Centre de recherche INRIA Grenoble – Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex
