Strategies for processing images with 4D-Var data assimilation methods

Isabelle Herlin, Dominique Béréziat, Nicolas Mercier

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Strategies for processing images with 4D-Var data assimilation methods

Isabelle Herlin — Dominique Béréziat — Nicolas Mercier

N° 7495 — version 2

initial version Décembre 2010 — revised version Décembre 2010

Thème NUM

ISSN 0249-6399 ISRN INRIA/RR--7495--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
PARIS - ROCQUENCOURT

Strategies for processing images with 4D-Var data assimilation methods

Isabelle Herlin* , Dominique Béréziat† , Nicolas Mercier*

Thème NUM — Systèmes numériques
Équipes-Projets Clime

Rapport de recherche n° 7495 — version 2 — initial version Décembre 2010 — revised version Décembre 2010 — 20 pages

Abstract: Data Assimilation is a well-known mathematical technic used, in environmental sciences, to improve, thanks to observation data, the forecasts obtained by meteorological, oceanographic or air quality simulation models. It aims to solve the evolution equations, describing the dynamics of the state variables, and an observation equation, linking at each space-time location the state vector and the observations. Data Assimilation allows to get a better knowledge of the actual system's state, named the reference. In this article, we first describe various strategies that can be applied in the framework of variational data assimilation to study various image processing issues. Second, we detail the mathematical setting and the analysis of pros and cons of each strategy for the issue of motion estimation. Last, results are provided on synthetic images and satellite acquisitions.

Key-words: image processing, inverse problems, data assimilation, non linear advection, optical flow

* INRIA, CEREA, joint laboratory ENPC - EDF R&D, Université Paris-Est
† Université Pierre et Marie Curie, LIP6

Centre de recherche INRIA Paris – Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Stratégies pour le traitement d'images avec des méthodes d'assimilation de données 4D-Var

Résumé : L'assimilation de données est un outil largement utilisé dans les sciences de l'environnement pour améliorer, au moyen de données d'observation, les prédictions obtenues par les modèles de simulation. Elle s'applique en météorologie, en océanographie et en qualité de l'air, par exemple. L'assimilation de données permet de résoudre les équations d'évolution, décrivant la dynamique des variables d'état du modèle, et les équations d'observation, qui lient le vecteur d'état et les observations. Dans cet article, nous décrivons plusieurs stratégies d'assimilation d'images, dans le contexte de la formulation faible de l'assimilation variationnelle. Nous détaillons ensuite les équations mathématiques de ces stratégies et nous analysons leurs avantages et défauts respectifs pour une application à l'estimation du mouvement. Des résultats sont fournis sur des données synthétiques et des images satellite.

Mots-clés : traitement d'images, problèmes inverses, assimilation de données, advection non linéaire, flot optique

1 Introduction: images and models

Simulation models and image data are simultaneously available for numerous scientific domains such as meteorology, oceanography, glaciology, atmosphere chemistry, engineering, biology and medicine. In all cases, complex flow models have been settled on the one hand, and huge amounts of different kinds of data, including image acquisitions, provide, on the other hand, a deep insight on the observed phenomenon. Models and images are indeed two complementary descriptions of the same reality. Two questions arise from that remark. From the model viewpoint: how to improve the forecast obtained by a simulation model with a better use of the image information? From the image viewpoint: how to derive accurate characteristic features from images by making an optimal use of the model? This paper is concerned by the description, formalization and analysis of various strategies that can be applied from the image viewpoint. The discussion is illustrated on the issue of motion estimation.

As explained in [3], an ill-posed Image Processing problem, such as estimating a motion field $\mathbf{W}(\mathbf{x},t)$ from a sequence of images $\mathbf{I}(\mathbf{x},t)$, is solved in the data assimilation framework if the following components are available or derived from heuristics:

- a dynamic model, also named evolution equation, for $\mathbf{W}(\mathbf{x},t)$,
- an equation linking $\mathbf{W}(\mathbf{x},y)$ and $\mathbf{I}(\mathbf{x},t)$, even in an implicit way: $\mathbb{I}(\mathbf{W}(\mathbf{x},t),(\mathbf{I}(\mathbf{x},t))=0$.

However, several formulations can be derived from these two equations, which correspond to various definitions of the state vector and observation operator involved in the data assimilation system. Each formulation may be viewed as a strategy for optimally combining images and model, whose performances depend on image properties.

The article is organized as follows. Section 2 briefly describes the variational data assimilation method known as weak formulation of 4D-Var. For a complete understanding of the method, its application to motion estimation is provided in Section 3. Section 4 discusses, in a generic way, several strategies, with the corresponding choices of state vector and observation operator. References from the literature are classified according to these strategies. Illustrations are given on the motion estimation issue in Section 5. Experimental results are displayed and analyzed in the same section. Section 6 ends the paper with some perspectives to this research study.

2 Variational Data Assimilation

Let us first summarize the major principles of the weak formulation of variational data assimilation.

2.1 Mathematical setting

Let $\mathbf{X}$ being a state vector depending on the spatial location $\mathbf{x}$ ($\mathbf{x}=(x,y)$ for 2D image data) and time t. $\mathbf{X}$ is defined on $A=\Omega\times[t_0,t_1]$, Ω being a bounded spatial domain and $[t_0,t_1]$ a temporal domain.

We assume $\mathbf{X}$ is evolving in time according to:

$$\frac{\partial \mathbf{X}}{\partial t}(\mathbf{x},t) + \mathbb{M}(\mathbf{X})(\mathbf{x},t) = \mathcal{E}_m(\mathbf{x},t) \tag{1}$$

$\mathbb{M}$, named *evolution model*, is supposed differentiable. As $\mathbb{M}$ only approximates the effective evolution of the state vector, a *model error* $\mathcal{E}_m$ is introduced to quantify the deviation in space and time. The existence of this *model error* corresponds to the weak formulation of data assimilation.

Observations $\mathbf{Y}(\mathbf{x},t)$, obtained for instance from satellite acquisitions, are available at location $\mathbf{x}$ and date t. They are linked to the state vector through an observation equation:

$$\mathbf{Y}(\mathbf{x},t) = \mathbb{H}(\mathbf{X}(\mathbf{x},t)) + \mathcal{E}_o(\mathbf{x},t) \tag{2}$$

$\mathbb{H}$ is named the *observation operator* and supposed differentiable. Equation (2) is the standard form of the observation equation used in the Data Assimilation literature. However, this formulation is quite restrictive to describe the complex links existing between the observations and the state vector. To be more general, the following, which includes the previous one, is used in the paper:

$$\mathbb{H}(\mathbf{X},\mathbf{Y})(\mathbf{x},t) = \mathcal{E}_o(\mathbf{x},t) \tag{3}$$

The *observation error* $\mathcal{E}_o$ simultaneously represents the imperfection of the observation operator $\mathbb{H}$ and the measurement errors.

We consider having some knowledge on the initial condition of the state vector at $t = t_0$:

$$\mathbf{X}(\mathbf{x},t_0) = \mathbf{X}_b(\mathbf{x}) + \mathcal{E}_b(\mathbf{x}) \tag{4}$$

$\mathbf{X}_b$ is named the *background* and $\mathcal{E}_b$, the *background error*, expresses the uncertainty on the initial state value.

$\mathcal{E}_m$, $\mathcal{E}_o$ and $\mathcal{E}_b$ are assumed to be Gaussian and fully characterized by their covariance matrices Q, R and B [11].

2.2 Variational formulation

In order to solve the System (1), (4), (3) with respect to $\mathbf{X}$ having a maximal a posteriori probability given the observations, the following functional is defined and has to be minimized:

$$\begin{aligned} E(\mathbf{X}) =& \frac{1}{2}\int_A\int_A \left(\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X})\right)^T (\mathbf{x},t) Q^{-1} \left(\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X})\right)(\mathbf{x}',t') d\mathbf{x}dtd\mathbf{x}'dt' \\ &+ \int_A\int_A \mathbb{H}(\mathbf{X},\mathbf{Y})^T(\mathbf{x},t) R^{-1}(\mathbf{x},t,\mathbf{x}',t') \mathbb{H}(\mathbf{X},\mathbf{Y})(\mathbf{x}',t') d\mathbf{x}dtd\mathbf{x}'dt' \\ &+ \int_\Omega\int_\Omega \big(\mathbf{X}(\mathbf{x},t_0) - \mathbf{X}_b(\mathbf{x})\big)^T B^{-1}(\mathbf{x},\mathbf{x}') \big(\mathbf{X}(\mathbf{x}',t_0) - \mathbf{X}_b(\mathbf{x}')\big) d\mathbf{x}d\mathbf{x}' \end{aligned} \tag{5}$$

As $\mathcal{E}_m$, $\mathcal{E}_o$ and $\mathcal{E}_b$ are assumed to be independent, the functional E represents the log-density of the joint probability law [2]. The minimization is carried

out by solving the associated Euler-Lagrange equation. The differential $\frac{\partial E}{\partial \mathbf{X}}$ is obtained by computing the derivative of E with respect to $\mathbf{X}$ in direction η:

$$\frac{\partial E}{\partial \mathbf{X}}(\eta) = \lim_{\gamma \to 0} \frac{d}{d\gamma}(E(\mathbf{X} + \gamma\eta)) \tag{6}$$

and introducing an auxiliary variable λ, named *adjoint variable* in the literature of Data Assimilation, defined by:

$$\lambda(\mathbf{x}, t) = \int_A Q^{-1}(\mathbf{x}, t, \mathbf{x}', t') \left(\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X}) \right) (\mathbf{x}', t') d\mathbf{x}' dt' \tag{7}$$

The determination of the Euler-Lagrange equation associated to Equation (5) leads to the following so-called Optimality System [9]:

$$\lambda(\mathbf{x}, t_1) = 0 \tag{8}$$

$$-\frac{\partial \lambda}{\partial t} + \left(\frac{\partial \mathbb{M}}{\partial \mathbf{X}} \right)^* \lambda = -\int_A \left(\frac{\partial \mathbb{H}}{\partial \mathbf{X}} \right)^* (\mathbf{x}, t) R^{-1} \mathbb{H}(\mathbf{X}, \mathbf{Y})(\mathbf{x}', t') d\mathbf{x}' dt' \tag{9}$$

$$\mathbf{X}(\mathbf{x}, t_0) = \int_\Omega B\lambda(\mathbf{x}', t_0) d\mathbf{x}' + \mathbf{X}_b(\mathbf{x}) \tag{10}$$

$$\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X}) = \int_A Q\lambda(\mathbf{x}', t') d\mathbf{x}' dt' \tag{11}$$

As the initial condition for λ is given at time t_1 (Equation (8)), λ has to be computed backward in time with Equation (9). It makes use of the two *adjoint operators* denoted by $\left(\frac{\partial \mathbb{M}}{\partial \mathbf{X}} \right)^*$ and $\left(\frac{\partial \mathbb{H}}{\partial \mathbf{X}} \right)^*$. Solving the Optimality System is however not straightforward: the state vector is determined from Equations (10) and (11) using the adjoint variable and the adjoint variable is determined from Equations (8) and (9) using the state vector. To break this deadlock, an incremental method is applied and summarized in the next subsection.

2.3 Incremental algorithm

The underlying idea comes from the following lemma:

$$\min_{w \in \mathcal{V}(w_0)} E(w) = \min_{\delta w \in \mathcal{V}(0)} E(w_0 + \delta w)$$

where w_0 denotes a local minimum of E and $\mathcal{V}(w_0)$ a neighborhood of w_0. The state vector $\mathbf{X}$ is therefore replaced by $\mathbf{X}_b + \delta\mathbf{X}$ in Equations (9), (10) and (11). $\mathbf{X}_b$ is the background variable and $\delta\mathbf{X}$ the incremental variable. As $\mathbb{M}$ and $\mathbb{H}$ are non linear (for the applications described in the paper), a first order development is applied:

$$\mathbb{M}(\mathbf{X}) \simeq \mathbb{M}(\mathbf{X}_b + \delta\mathbf{X}) = \mathbb{M}(\mathbf{X}_b) + \left. \frac{\partial \mathbb{M}}{\partial \mathbf{X}} \right|_{\mathbf{X}_b} \delta\mathbf{X} \tag{12}$$

$$\mathbb{H}(\mathbf{X}, \mathbf{Y}) \simeq \mathbb{H}(\mathbf{X}_b + \delta\mathbf{X}, \mathbf{Y}) = \mathbb{H}(\mathbf{X}_b, \mathbf{Y}) + \left. \frac{\partial \mathbb{H}}{\partial \mathbf{X}} \right|_{\mathbf{X}_b} \delta\mathbf{X} \tag{13}$$

Equations (8) to (11) are rewritten as:

$$\lambda(\mathbf{x}, t_1) = 0 \tag{14}$$

$$-\frac{\partial \lambda}{\partial t} + \frac{\partial \mathbb{M}}{\partial \mathbf{X}}\Big|^*_{\mathbf{X}_b} \lambda = -\int_A \frac{\partial \mathbb{H}}{\partial \mathbf{X}}\Big|^*_{\mathbf{X}_b} R^{-1} \left(\mathbb{H}(\mathbf{X}_b, \mathbf{Y}) + \frac{\partial \mathbb{H}}{\partial \mathbf{X}}\Big|_{\mathbf{X}_b} \delta\mathbf{X} \right) d\mathbf{x}' dt' \tag{15}$$

$$\mathbf{X}_b(\mathbf{x}, t_0) = \mathbf{X}_b(\mathbf{x}) \tag{16}$$

$$\frac{\partial \mathbf{X}_b}{\partial t} + \mathbb{M}(\mathbf{X}_b) = 0 \tag{17}$$

$$\delta\mathbf{X}(\mathbf{x}, t_0) = \int_\Omega B\lambda(\mathbf{x}', t_0) d\mathbf{x}' \tag{18}$$

$$\frac{\partial \delta\mathbf{X}}{\partial t} + \frac{\partial \mathbb{M}}{\partial \mathbf{X}}\Big|_{\mathbf{X}_b} \delta\mathbf{X}(\mathbf{x}, t) = \int_A Q(\mathbf{x}, t, \mathbf{x}', t')\lambda(\mathbf{x}', t') d\mathbf{x}' dt' \tag{19}$$

The background variable $\mathbf{X}_b$ is calculated from Equations (16) and (17). The adjoint variable λ is then obtained with Equations (14) and (15). Last, the incremental variable $\delta\mathbf{X}$ is computed using Equations (18) and (19). However this only produces an approximate solution $\mathbf{X}_b + \delta\mathbf{X}$ due to the first order Taylor developments applied in (12) and (13). An iterative method is then applied until convergence.

Section 3 directly illustrates this method on the issue of motion estimation. This allows the Reader to better understand the discussion of Section 4 about the various strategies, which can be applied for image assimilation.

3 Motion estimation

Having described the mathematical setting of variational data assimilation and the incremental algorithm used to derive the solution for the weak formulation, this section addresses their application to motion estimation from a discrete image sequence. The illustration concerns the retrieval of ocean surface currents from Sea Surface Temperature data (SST).

During the last twenty years, many visible and infrared sensors have been launched on different satellites and acquire SST data with a spatial resolution ranging from a few hundred meters to a few kilometers. They display the impact of the mesoscale dynamics on the ocean surface. Structures such as eddies, jets, filaments are visible on such image sequences (see for instance Figure 1[1]).

However, as clouds are not transparent in visible and infrared bands, no data is available if the surface is covered by clouds. As explained in [3], variational data assimilation allows an optimal framework to retrieve a dense motion field from the images, without applying neither a preprocessing nor a spatial regularization, even in case of missing data and cloudy acquisitions, thanks to the dynamic model involved in the mathematical setting and the observation error covariance matrix R.

The next two subsections describe the observation Equation (3), the covariance matrix R and the evolution Equation (1) used in [3] to retrieve motion from image data. The implementation is further described in Subsection 5.1.

[1] Data have been provided by E. Plotnikov and G. Korotaev from the Marine Hydrophysical Institute of Sevastopol, Ukraine.

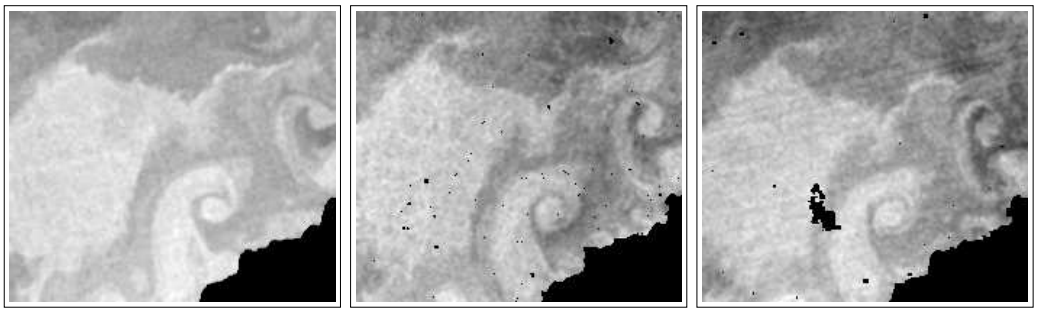

Figure 1: Sea Surface Temperature data acquired over the Black Sea. Black pixels correspond to clouds or ground area.

3.1 Observation equation

Advection of the grey level values by the motion field is the assumption commonly used by most authors of the image processing community studying the issue of motion estimation. This was first proposed by Horn and Schunk [7]. This heuristic is applied in the context of processing satellite images (see, for example [3, 4, 5, 6, 8, 10, 13]). The advection equation is also used, without specific physical justification, by authors retrieving motion from image data with a data assimilation approach. However, in case of the ocean surface circulation, images acquired by satellite sensors, within a short time interval, support the physical assumption of a transport-diffusion of the contrasts induced by the ocean mesoscale processes on the sea surface.

Let $\mathbf{I}$ be the sequence of images defined on Ω. Let $\mathbf{W}(\mathbf{x},t)$ be the velocity vector of a point $\mathbf{x} \in \Omega$ between t and $t+\triangle t$. The advection assumption corresponds to:

$$\mathbf{I}(\mathbf{x}+\mathbf{W}(\mathbf{x},t)\triangle t, t+\triangle t) = \mathbf{I}(\mathbf{x},t) \tag{20}$$

This equation is non linear with respect to $\mathbf{W}$ and its left member is often linearized using a first order Taylor development around $(\mathbf{x},t)$. This provides the so-called optical flow constraint [7], which corresponds to a transport equation:

$$\frac{\partial \mathbf{I}}{\partial t}(\mathbf{x},t) + \nabla \mathbf{I}(\mathbf{x},t) \cdot \mathbf{W}(\mathbf{x},t) = 0 \quad \forall \mathbf{x} \in \Omega \tag{21}$$

with $\nabla \mathbf{I} = \left(\frac{\partial \mathbf{I}}{\partial x} \quad \frac{\partial \mathbf{I}}{\partial y} \right)^T$ and the operator $\cdot$ denoting the dot product of two vectors.
By identifying (21) and (3), we have:

$$\begin{aligned} \mathbf{X}(\mathbf{x},t) &= \mathbf{W}(\mathbf{x},t) \\ \mathbf{Y}(\mathbf{x},t) &= \mathbf{I}(\mathbf{x},t) \\ \mathbb{H}(\mathbf{X},\mathbf{Y})(\mathbf{x},t) &= \frac{\partial \mathbf{I}}{\partial t}(\mathbf{x},t) + \nabla \mathbf{I}(\mathbf{x},t) \cdot \mathbf{W}(\mathbf{x},t) \end{aligned}$$

An appropriate observation error $\mathcal{E}_o$ has to be defined. Its covariance matrix R weights the contribution of the observations in Equation (5). Its inverse should

then have values close to zero when observations should be discarded, which is the case for missing and cloudy data. Using a Dirac covariance formalism, R is written as:

$$R(\mathbf{x}, t, \mathbf{x}', t') = r(\mathbf{x}, t)\delta(\mathbf{x} - \mathbf{x}')\,\delta(t - t')$$

with r a real matrix, whose size being to the number of components of the operator $\mathbb{H}$. Its inverse is:

$$R^{-1}(\mathbf{x}, t, \mathbf{x}', t') = \delta(\mathbf{x} - \mathbf{x}')\,\delta(t - t')\,r^{-1}(\mathbf{x}, t) \tag{22}$$

The matrix r^{-1} characterizes the quality of the observation: a high value indicates that the observation value is relevant while a value close to zero indicates an irrelevant one, which should be given a low weight. Assuming the availability of a function f measuring the confidence in the observation data ($f \in [0, 1]$, $f = 0$ for no confidence), one admissible formulation of r^{-1} is:

$$r^{-1}(\mathbf{x}, t) = r_0^{-1}(1 - f(\mathbf{x}, t)) + r_1^{-1} f(\mathbf{x}, t) \tag{23}$$

$r^{-1}(\mathbf{x}, t)$ is equal to the "minimal value" r_0^{-1} if confidence is 0 and equal to the "maximal value" r_1^{-1} if confidence is 1. Matrices r_0 and r_1 are chosen to be constant in space and time, and invertible. For completeness, f is modeled as:

$$f(\mathbf{x}, t) = f_{\text{sensor}}(\mathbf{x}, t) f_{\mathbb{H}}(\mathbf{x}, t) \tag{24}$$

f_{sensor} indicates the quality of the observation values and is obtained from the metadata provided simultaneously with the image acquisitions in case of satellite data. $f_{\mathbb{H}}$ measures the confidence in the observation model; its value is 0 if the observation equation (Equation (21)) is not verified.

With such definition of the covariance matrix R, observation values with a low confidence have a low weight during the computation of the solution.

3.2 Evolution equation

A simple heuristic on the dynamics is to assume $\mathbf{X} = \mathbf{W}$ being constant over time along a pixel trajectory, which is expressed by:

$$\frac{d\mathbf{X}}{dt} = 0$$

or:

$$\frac{\partial \mathbf{X}}{\partial t} + (\mathbf{X} \cdot \nabla)\mathbf{X} = 0 \tag{25}$$

This is a transport equation of $\mathbf{X}$. Identifying Equations (25) and (1), the evolution model is $\mathbb{M}(\mathbf{X})(\mathbf{x}, t) == (\mathbf{X}(\mathbf{x}, t) \cdot \nabla)\mathbf{X}(\mathbf{x}, t)$.

The covariance matrix Q, associated to the model error $\mathcal{E}_m$, is used in Functional (5) for weighting the term $\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X})$. $\mathbf{X}$ being a two-component vector, the Q matrix is of size 2×2 and chosen as:

$$Q(\mathbf{x}, t, \mathbf{x}', t') = val_Q(\mathbf{x} - \mathbf{x}', t - t') \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{26}$$

with $val_Q(\mathbf{x}, t) = \exp(-\frac{1}{\sigma}(\|\mathbf{x}\| + |t|))$ for ensuring a first-order regularization of $\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X})$ or $val_Q(\mathbf{x}, t) = \delta(\mathbf{x})\delta(t)$ for a zero-order one [3]. In such a way, we force the temporal evolution of $\mathbf{X} = \mathbf{W}$ to be coherent with the heuristics modeled by $\mathbb{M}$.

4 Alternatives for image assimilation

Having shortly described what is variational data assimilation and its application to the issue of motion estimation, we come to the core of the paper, which concerns the discussion on the various strategies which can be applied.

4.1 First strategy

As stated in Section 2, the aim is to retrieve some quantity $\mathbf{X}$ from an image sequence $\mathbf{I}$. This is the usual objective of image processing problems, such as optical flow estimation, image registration, curves or surfaces matching, tracking of multiple objects, segmentation, restoration, deconvolution, denoising and shape from shading.

The first strategy of image assimilation solves the System {(1), (4), (3)}, to retrieve the state vector $\mathbf{X}$ on the space-time domain A by *directly* assimilating image observations $\mathbf{Y} = \mathbf{I}$ in the dynamic model $\mathbb{M}$.

As the evolution equation (Equation (1)) involves a model error $\mathcal{E}_m(\mathbf{x}, t)$, the energy function (5) includes a term $E_1(\mathbf{X})$ (the first one) to minimize this error over the space-time domain A:

$$E_1(\mathbf{X}) = \frac{1}{2} \int_A \int_A \left(\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X}) \right)^T Q^{-1}(\mathbf{x}, t, \mathbf{x}', t') \left(\frac{\partial \mathbf{X}}{\partial t} + \mathbb{M}(\mathbf{X}) \right) d\mathbf{x} dt d\mathbf{x}' dt' \tag{27}$$

This is the weak formulation of 4D-Var, due to the average confidence in the dynamics expressed by $\mathbb{M}$, which is often derived from heuristics. In the best case $\mathbb{M}$ is obtained by simplifying physical models such as atmosphere or ocean models, but does not correctly represent the small scales non linear phenomena visible on the image data. As the model and the image data are two complementary views of the reality, this formulation forces the confidence in data and may be viewed as an image-driven process.

In the initial condition equation (Equation (4)), $\mathbf{X}_b$ is taken null or computed from the image data.

Image data and particularly satellite data may be corrupted by noise: the acquisition may be missing on a pixel or a region, or part of the image may be occluded by a cloud. Applying a preprocessing on the image data is not an optimal solution, due to the difficulty to assess the uncertainty on the result. As described in [3] and summarized in Section 3, a better solution is to include the preprocessing in the definition of the covariance matrix R, in order to discard the noisy data or give them a low weight during the computation.

The links between the vector state $\mathbf{X}$ and the image observations $\mathbf{Y} = \mathbf{I}$ may be complex and result in an implicit observation equation: $\mathbb{H}(\mathbf{X}, \mathbf{Y}) = 0$. This limits the range of admissible comparisons of the state vector $\mathbf{X}$ and the observations $\mathbf{Y}$ to this unique equation. The observation term, $E_2(\mathbf{X})$, measuring the discrepancy of $\mathbf{X}$ and $\mathbf{Y}$, in Equation (5), is:

$$E_2(\mathbf{X}) = \int_A \int_A \mathbb{H}(\mathbf{X}, \mathbf{Y})^T(\mathbf{x}, t) R^{-1}(\mathbf{x}, t, \mathbf{x}', t') \mathbb{H}(\mathbf{X}, \mathbf{Y})(\mathbf{x}', t') d\mathbf{x} dt d\mathbf{x}' dt'$$

Alternatives to this first strategy are described in the following.

4.2 Alternatives

We note $\mathbb{I}(\mathbf{X}, \mathbf{I}) = 0$ the equation that links the image sequence $\mathbf{I}$ to the quantity $\mathbf{X}$ to be retrieved. In Section 2, the state vector has been chosen as $\mathbf{X}$, the observation vector $\mathbf{Y}$ as $\mathbf{I}$, and the observation operator $\mathbb{H}$ has been defined as $\mathbb{H}(\mathbf{X}, \mathbf{Y}) = \mathbb{I}(\mathbf{X}, \mathbf{I})$. Alternatives are obtained by defining in a different way the state vector and the observation operator. An extended state vector $\mathbf{X}_e$ is defined as $\mathbf{X}_e(\mathbf{x}, t) = (\mathbf{X}(\mathbf{x}, t) \quad \mathbf{q}(\mathbf{x}, t))^T$ with $\mathbf{q}(\mathbf{x}, t)$ being a mathematical quantity similar to the image value $\mathbf{I}(\mathbf{x}, t)$. This allows the features of $\mathbf{q}(\mathbf{x}, t)$ to be directly compared with those of $\mathbf{I}(\mathbf{x}, t)$. $\mathbf{Y}(\mathbf{x}, t)$ is again chosen equal to $\mathbf{I}(\mathbf{x}, t)$.

The second strategy of image assimilation directly measures the discrepancy of the extended state vector and the observation by the difference of $\mathbf{q}(\mathbf{x}, t)$ and $\mathbf{Y}(\mathbf{x}, t)$: the observation operator becomes $\mathbf{q}(\mathbf{x}, t) - \mathbf{Y}(\mathbf{x}, t)$. The term $E_2(\mathbf{X})$ is replaced by:

$$E_2(\mathbf{X}_e) = \int_A \int_A (\mathbf{q} - \mathbf{Y})^T(\mathbf{x}, t) R^{-1}(\mathbf{x}, t, \mathbf{x}', t')(\mathbf{q} - \mathbf{Y})(\mathbf{x}', t') d\mathbf{x} dt d\mathbf{x}' dt'$$

computed on the space-time domain. This strategy is equivalent to the assimilation of image values $\mathbf{I}(\mathbf{x}, t)$ in an extended model with state vector $\mathbf{X}_e$.

In the third strategy of image assimilation, a transformation T is first applied to the observation $\mathbf{Y}$, such as Fourier transform, wavelet or curvelet [12] transform. The discrepancy is then measured in the space of the transform coefficients and $E_2(\mathbf{X})$ becomes:

$$E_2(\mathbf{X}_e) = \int_{t_0}^{t_1} \int_{t_0}^{t_1} (T(\mathbf{q}) - T(\mathbf{Y}))^T(t) R^{-1}(t, t')(T(\mathbf{q}) - T(\mathbf{Y}))(t') dt dt'$$

The aim of the transform T is to extract the characteristics, that have to be considered by the model to improve forecasts. Consequently, the data that are assimilated in the model are no more the raw observations acquired by the sensor, but pseudo-observations computed by the numerical process associated with T, whose performances strongly impact the result of the assimilation process. The characteristics of this numerical process have to be taken into account in order to estimate the uncertainties involved in the data assimilation process.

Even if attractive, these second and third strategies include a strong requirement: the design of a dynamic model $\mathbb{F}$ for $\mathbf{q}(t)$:

$$\frac{\partial \mathbf{q}}{\partial t}(\mathbf{x}, t) + \mathbb{F}(\mathbf{q})(\mathbf{x}, t) = \mathcal{E}_f(\mathbf{x}, t) \tag{28}$$

A transport equation is often chosen, as in [12], that corresponds to the optical flow constraint [7] applied in image processing. However, when studying satellite acquisitions, this transport equation and the evolution displayed on the image data may be incompatible: e.g., the appearance and disappearance of clouds on image data represent physical processes which are not modeled by a simple advection.

The design of the observation error covariance matrix R has also to be considered in a different way for the three strategies.

- In the first strategy, R has been defined in [3] and shortly described in Subsection 3.1 from the metadata provided with images and from the observation equation itself.

- In the second strategy, a covariance matrix satisfying Equations (22) and (23) is defined by taking the quality measure f equal to f_{sensor} computed from the metadata provided with images.
- In the third strategy, the observation error covariance matrix has to be defined according to the transformation that is applied to the image sequence and to its numerical implementation. This is the major weakness of this strategy.

5 Application to motion estimation

We designed two methods for motion estimation, corresponding to the first two strategies. The third one is no more discussed in the paper because we focus on the impact of adding the quantity $\mathbf{q}$ to the original state vector $\mathbf{X}$ in order to obtain an extended state vector $\mathbf{X}_e$. The two methods are fully described in Subsections 5.1 and 5.2. They are applied on synthetic data, to quantify the differences in Subsection 5.3, and on satellite data, for a qualitative discussion in Subsection 5.4.

5.1 First strategy: direct assimilation of the observation

The aim of this subsection is to summarize the main components of the first strategy that has been discussed in Section 3 and to describe its implementation. This is mandatory to better understand the differences with the second strategy, explained the next subsection, involving the quantity $\mathbf{q}$ and the extended state vector $\mathbf{X}_e$.

$\mathbf{I}$ denotes the sequence of images defined on the bounded domain Ω. The state vector $\mathbf{X}(\mathbf{x},t)$ is the velocity vector $\mathbf{W}(\mathbf{x},t)$ of a point $\mathbf{x} \in \Omega$ between t and $t+\triangle t$. The transport equation of the image value by the velocity vector, the so-called optical flow constraint equation [7], is considered:

$$\frac{\partial \mathbf{I}}{\partial t}(\mathbf{x},t) + \nabla \mathbf{I}(\mathbf{x},t) \cdot \mathbf{W}(\mathbf{x},t) = 0 \quad \forall \mathbf{x} \in \Omega \tag{29}$$

In this case, the image $\mathbf{I}(\mathbf{x},t)$ constitutes the observation vector $\mathbf{Y}(\mathbf{x},t)$ and the observation model $\mathbb{H}$ is:

$$\mathbb{H}(\mathbf{X},\mathbf{Y})(\mathbf{x},t) = \mathbb{H}(\mathbf{W},\mathbf{I})(\mathbf{x},t) = \frac{\partial \mathbf{I}}{\partial t}(\mathbf{x},t) + \nabla \mathbf{I}(\mathbf{x},t) \cdot \mathbf{W}(\mathbf{x},t) \tag{30}$$

Equations (22) and (23) are used to define the inverse of the observation error matrix R and locate the observation values which must be discarded. The observation model $\mathbb{H}$ being scalar, the matrices r_0^{-1} and r_1^{-1} are scalar and respectively set to ϵ and $1-\epsilon$ with $\epsilon \simeq 10^{-6}$. Equation (24) is used as the observation confidence. f_{sensor} is set to 0 if data are not or wrongly acquired and to 1 otherwise. $f_{\mathbb{H}}$ is chosen from the following remark: the spatio-temporal gradient is null on regions of uniform grey level values and Equation (29) is then reduced to $0=0$ on these pixels. For avoiding further considering them, $f_{\mathbb{H}}$ is defined by:

$$f_{\mathbb{H}}(\mathbf{x},t) = 1 - \exp(-\|\nabla_3 \mathbf{I}(\mathbf{x},t)\|^2) \tag{31}$$

where ∇_3 denotes the spatio-temporal gradient operator.

The transport of the velocity is taken as evolution equation:

$$\frac{\partial \mathbf{W}}{\partial t} + (\mathbf{W} \cdot \nabla)\mathbf{W} = 0 \tag{32}$$

This equation is rewritten as a two-component system:

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} + V\frac{\partial U}{\partial y} = 0 \tag{33}$$

$$\frac{\partial V}{\partial t} + U\frac{\partial V}{\partial x} + V\frac{\partial V}{\partial y} = 0 \tag{34}$$

and the evolution model is:

$$\begin{aligned} \mathbb{M}(\mathbf{X}) = \mathbb{M}(\mathbf{W}) &= \begin{pmatrix} \mathbb{M}_1(\mathbf{W}) & \mathbb{M}_2(\mathbf{W}) \end{pmatrix}^T \\ &= \begin{pmatrix} U\frac{\partial U}{\partial x} + V\frac{\partial U}{\partial y} & U\frac{\partial V}{\partial x} + V\frac{\partial V}{\partial y} \end{pmatrix}^T \end{aligned}$$

$\mathbf{W}$ being a two-component vector, the Q matrix is of size 2×2 and chosen as:

$$Q(\mathbf{x}, t, \mathbf{x}', t') = val_Q(\mathbf{x} - \mathbf{x}', t - t') \begin{pmatrix} Q_U & 0 \\ 0 & Q_V \end{pmatrix} \tag{35}$$

with $val_Q(\mathbf{x}, t) = \exp(-\frac{1}{\sigma}(\|\mathbf{x}\| + |t|))$ or $val_Q(\mathbf{x}, t) = \delta(\mathbf{x})\delta(t)$. We tested both without noticing significant differences on the results.

The background value at $t = t_0$ (Equation (4)) is given a null value. We consider the background error $B(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}') \begin{pmatrix} B_U & 0 \\ 0 & B_V \end{pmatrix}$ for measuring the quadratic discrepancy between $\mathbf{X}(\mathbf{x}, t_0) = \mathbf{W}(\mathbf{x}, t_0)$ and $\mathbf{X}_b(\mathbf{x}) = \mathbf{W}_b(\mathbf{x})$.

5.1.1 Adjoint operators

In order to determine the adjoint operators for $\mathbb{M}$ and $\mathbb{H}$, the directional derivatives are first established. Using the Equation (6), we obtain:
$\frac{\partial \mathbb{M}}{\partial \mathbf{X}}(\eta) = \begin{pmatrix} \frac{\partial \mathbb{M}_1}{\partial \mathbf{X}}(\eta) & \frac{\partial \mathbb{M}_2}{\partial \mathbf{X}}(\eta) \end{pmatrix}^T$ and:

$$\frac{\partial \mathbb{M}_1}{\partial \mathbf{X}}(\eta) = U_x\eta^1 + U\eta^1_x + V\eta^1_y + U_y\eta^2 \tag{36}$$

$$\frac{\partial \mathbb{M}_2}{\partial \mathbf{X}}(\eta) = V_x\eta^1 + V_y\eta^2 + U\eta^2_x + V\eta^2_y \tag{37}$$

with $\eta = \begin{pmatrix} \eta^1 & \eta^2 \end{pmatrix}^T$. The subscripts x and y stand for partial derivatives with respect to x and y. Using the definition of the adjoint operator, integration by parts and considering boundary terms equal to zero, the adjoint operator of $\mathbb{M}$ is defined by:

$$\left(\frac{\partial \mathbb{M}_1}{\partial \mathbf{X}}\right)^*(\lambda) = -U\lambda^1_x - V_y\lambda^1 - V\lambda^1_y + V_x\lambda^2 \tag{38}$$

$$\left(\frac{\partial \mathbb{M}_2}{\partial \mathbf{X}}\right)^*(\lambda) = U_y\lambda^1 - U_x\lambda^2 - U\lambda^2_x - V\lambda^2_y \tag{39}$$

with $\lambda = \begin{pmatrix}\lambda^1 & \lambda^2\end{pmatrix}^T$. In a compact form, the adjoint operator of $\mathbb{M}$ is written:

$$\begin{aligned}\left(\frac{\partial \mathbb{M}}{\partial \mathbf{X}}\right)^* (\lambda) &= -\begin{pmatrix}\lambda_x^1 & \lambda_x^2\\ \lambda_y^1 & \lambda_y^2\end{pmatrix}^T \begin{pmatrix}U\\ V\end{pmatrix} - \begin{pmatrix}V_y & -U_y\\ -V_x & U_x\end{pmatrix}^T \begin{pmatrix}\lambda^1\\ \lambda^2\end{pmatrix}\\ &= -\begin{pmatrix}\nabla\lambda^1 & \nabla\lambda^2\end{pmatrix}^T \mathbf{W} - \begin{pmatrix}\nabla^\perp V & -\nabla^\perp U\end{pmatrix}^T \lambda\end{aligned}$$

with $\nabla^\perp U = \begin{pmatrix}U_y & -U_x\end{pmatrix}^T$.
The directional derivative of the observation operator is:

$$\frac{\partial \mathbb{H}}{\partial \mathbf{X}}(\eta) = \nabla \mathbf{I} \cdot \eta$$

Let μ be a scalar function, we have:

$$\left(\frac{\partial \mathbb{H}}{\partial \mathbf{X}}\right)^* (\mu) = \mu \nabla \mathbf{I}$$

5.1.2 Discretization

In this context, the three PDEs ((17),(15),(19)) become:

$$\frac{\partial \mathbf{W}}{\partial t} + (\mathbf{W}\cdot\nabla)\mathbf{W} = 0 \tag{40}$$

$$-\frac{\partial \lambda}{\partial t} - (\lambda\cdot\nabla)\mathbf{W} - (\mathbf{W}\cdot\nabla^\perp)\lambda = -\nabla \mathbf{I} R^{-1} \star L \tag{41}$$

$$\frac{\partial \delta\mathbf{W}}{\partial t} + (\delta\mathbf{W}\cdot\nabla)\mathbf{W} + (\mathbf{W}\cdot\nabla)\delta\mathbf{W} = Q \star \lambda \tag{42}$$

with $L = \mathbf{I}_t + \nabla \mathbf{I}^T(\mathbf{W} + \delta\mathbf{W})$. The right members of Equations (41) and (42) are expressed as a convolution product. For their evaluation, the knowledge of L and λ over the whole temporal domain is required. However, choosing a Dirac matrix for R and Q reduces the convolution products to a simple multiplication. In this case, the computation of $\lambda(t)$ depends on values at $t+1$ in (41) and $\delta\mathbf{X}(t)$ is obtained from variables at $t-1$ in (42). The algorithm expresses a frame-by-frame process: the whole sequence of state and observation vectors, adjoint and incremental variables may be left on the mass storage, except the frames involved in the computation.

These three equations are discretized using a finite difference technique. $\triangle\mathbf{x}$ and $\triangle t$ denote the space and time periods of discretization. Let us first examine Equation (40): it is a 2D non linear advection equation. It has two Components (33) and (34). The first one combines a term of linear advection in direction y and non linear one in direction x and is expressed as a two-equation system using the splitting method:

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} = 0 \tag{43}$$

$$\frac{\partial U}{\partial t} + V\frac{\partial U}{\partial y} = 0 \tag{44}$$

Equation (43) is rewritten in a conservative form as $\dfrac{\partial U}{\partial t} + \dfrac{\partial F(U)}{\partial x} = 0$ with $F(U) = \dfrac{1}{2}U^2$ and discretized with a Lax-Friedrich scheme:

$$U_{i,j}^{k+1} = \frac{1}{2}(U_{i+1,j}^k + U_{i-1,j}^k) - \frac{\triangle t}{2\triangle\mathbf{x}}(F_{i+1,j}^k - F_{i-1,j}^k)$$

with $U^k_{i,j} = U(x_i, y_i, t_k)$, $F^k_{i,j} = F(U(x_i, y_i, t_k))$. The term $\frac{1}{2}(U^k_{i+1,j} + U^k_{i-1,j})$ stabilizes the scheme by adding a diffusive effect while $\triangle t$ verifies the Courant-Friedrich-Levy condition. The linear advection (44) is discretized using a first-order upwind scheme [1]:

$$U^{k+1}_{i,j} = U^k_{i,j} - \frac{\triangle t}{\triangle \mathbf{x}} \left(\max(V^k_{i,j}, 0) \left(U^k_{i,j} - U^k_{i,j-1} \right) + \min(V^k_{i,j}, 0) \left(U^k_{i,j+1} - U^k_{i,j} \right) \right)$$

In the same way, it can be seen that the second component of (40) (Equation 34) contains a linear advection term in the direction x and a non linear one in the direction y. The same discretization strategy is then applied.

Equation (41) combines a linear advection $((\lambda \cdot \nabla)\mathbf{W})$, a term of reaction $((\mathbf{W} \cdot \nabla)\lambda)$ and a forcing term $(\nabla I R^{-1} \star L)$. Its first component is $-\frac{\partial \lambda^1}{\partial t} - U\lambda^1_x - V_y\lambda^1 - V\lambda^1_y + V_x\lambda^2 = -\mathbf{I}_x R^{-1} \star L$. It is split into two parts. The first part contains the linear advection in the direction x and the reaction term: $-\frac{\partial \lambda^1}{\partial t} - U\lambda^1_x - V_y\lambda^1 = 0$ and is discretized as (44) with an upwind scheme. However, the equation is retrograde and its initial condition is given at time t_1:

$$\begin{aligned} (\lambda^1)^{k-1}_{i,j} &= \left(1 + \frac{\triangle t}{2\triangle \mathbf{x}}(V^k_{i,j+1} - V^k_{i,j-1})\right)(\lambda^1)^k_{i,j} + \\ &\quad \frac{\triangle t}{\triangle \mathbf{x}} \left(\max(U^k_{i,j}, 0)((\lambda^1)^k_{i,j} - (\lambda^1)^k_{i-1,j}) + \min(U^k_{i,j}, 0)((\lambda^1)^k_{i+1,j} - (\lambda^1)^k_{i,j}) \right) \end{aligned}$$

The second part contains the linear advection term in the direction y and the forcing term: $-\frac{\partial \lambda^1}{\partial t} - V\lambda^1_y = -V_x\lambda^2 - \mathbf{I}_x R^{-1} \star L$. Again, an upwind scheme is used:

$$\begin{aligned} (\lambda^1)^{k-1}_{i,j} &= (\lambda^1)^k_{i,j} - \frac{\triangle t}{2\triangle \mathbf{x}} \left(V^k_{i+1,j} - V^k_{i-1,j} \right)(\lambda^2)^k_{i,j} + \triangle t(-\mathbf{I}_x R^{-1} \star L)^k_{i,j} + \\ &\quad \frac{\triangle t}{\triangle \mathbf{x}} \left(\max(V^k_{i,j}, 0)((\lambda^1)^k_{i,j} - (\lambda^1)^k_{i,j-1}) + \min(V^k_{i,j}, 0)((\lambda^1)^k_{i,j+1} - (\lambda^1)^k_{i,j}) \right) \end{aligned}$$

The second component of (41) is discretized with the same method.

The last equation, (42), is similar to Equation (41), we therefore use the same discretization technique.

5.2 Second strategy: assimilation by comparison with a pseudo-quantity

This subsection has the same structure than Subsection 5.1 in order to highlight similarities and differences of both methods and better understand the impact on the results described in the following.

The extended state vector $\mathbf{X}_e(\mathbf{x}, t)$ is defined as $\mathbf{X}_e(\mathbf{x}, t) = (\mathbf{W}(\mathbf{x}, t) \quad \mathbf{q}(\mathbf{x}, t))^T$. It includes a quantity $\mathbf{q}$ which belongs to the same space than the image observations $\mathbf{I}$. The observation vector $\mathbf{Y}$ is again defined as:

$$\mathbf{Y}(\mathbf{x}, t) = \mathbf{I}(\mathbf{x}, t) \tag{45}$$

and the observation equation, used to measure the discrepancy between the state vector $\mathbf{X}_e$ and the observation $\mathbf{Y}$ becomes:

$$\mathbf{q}(\mathbf{x},t) - \mathbf{Y}(\mathbf{x},t) = 0 \tag{46}$$

The observation operator $\mathbb{H}$ is:

$$\mathbb{H}(\mathbf{X}_e, \mathbf{Y})(\mathbf{x},t) = \mathbf{q}(\mathbf{x},t) - \mathbf{Y}(\mathbf{x},t) \tag{47}$$

The discussion of Subsection 5.1 on the observation error matrix R is still valid. However, the confidence function f reduces to f_{sensor}.

The evolution system for $\mathbf{X}_e$ is taken as:

$$\frac{\partial \mathbf{W}}{\partial t} + (\mathbf{W}\cdot\nabla)\mathbf{W} = 0 \tag{48}$$

$$\frac{\partial \mathbf{I}}{\partial t} + \mathbf{W}\cdot\nabla\mathbf{I} = 0 \tag{49}$$

and the evolution model is

$$\begin{aligned}\mathbb{M}(\mathbf{X}_e) &= \begin{pmatrix}\mathbb{M}_1(\mathbf{X}_e) & \mathbb{M}_2(\mathbf{X}_e) & \mathbb{M}_3(\mathbf{X}_e)\end{pmatrix}^T \\ &= \begin{pmatrix}UU_x + VU_y & UV_x + VV_y & U\mathbf{q}_x + V\mathbf{q}_y\end{pmatrix}^T\end{aligned}$$

with $\mathbb{M}_1(\mathbf{X}_e) = \mathbb{M}_1(\mathbf{W})$ and $\mathbb{M}_2(\mathbf{X}_e) = \mathbb{M}_2(\mathbf{W})$ of Section 5.1.

$\mathbf{X}_e$ being a three-component vector, the matrix Q is of size 3×3 and chosen as:

$$Q(\mathbf{x},t,\mathbf{x}',t') = val_Q(\mathbf{x}-\mathbf{x}',t-t')\begin{pmatrix}Q_U & 0 & 0\\ 0 & Q_V & 0\\ 0 & 0 & Q_{\mathbf{q}}\end{pmatrix} \tag{50}$$

with $val_Q(\mathbf{x},t) = \delta(\mathbf{x})\delta(t)$.

The background value at $t = t_0$ (Equation (4)), $\mathbf{X}_{eb}(\mathbf{x})$, is given a null value for $\mathbf{W}$ and the first observation for $\mathbf{q}$. The background error is chosen as $B(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')\begin{pmatrix}B_U & 0 & 0\\ 0 & B_V & 0\\ 0 & 0 & B_{\mathbf{q}}\end{pmatrix}$ for measuring the quadratic discrepancy between $\mathbf{X}_e(\mathbf{x},t_0)$ and $\mathbf{X}_{eb}(\mathbf{x})$.

5.2.1 Adjoint operators and discretization

Given $\eta = (\eta_1 \quad \eta_2 \quad \eta_3)^T$, $\dfrac{\partial \mathbb{M}_1}{\partial \mathbf{X}_e}(\eta)$ and $\dfrac{\partial \mathbb{M}_2}{\partial \mathbf{X}_e}(\eta)$ have the same values than in Equations (36) and (37). The directional derivative of $\mathbb{M}_3$ is defined by:

$$\frac{\partial \mathbb{M}_3}{\partial \mathbf{X}_e}(\eta) = \mathbf{q}_x\eta^1 + \mathbf{q}_y\eta^2 + U\eta^3_x + V\eta^3_y$$

Considering $\lambda = (\lambda^1 \quad \lambda^2 \quad \lambda^3)^T$, the adjoint operators are given by :

$$\begin{aligned}\left(\frac{\partial \mathbb{M}_1}{\partial \mathbf{X}_e}\right)^*(\lambda) &= -U\lambda^1_x - V_y\lambda^1 - V\lambda^1_y + V_x\lambda^2 + \mathbf{q}_x\lambda^3\\ \left(\frac{\partial \mathbb{M}_2}{\partial \mathbf{X}_e}\right)^*(\lambda) &= U_y\lambda^1 - U_x\lambda^2 - U\lambda^2_x - V\lambda^2_y + \mathbf{q}_y\lambda^3\\ \left(\frac{\partial \mathbb{M}_3}{\partial \mathbf{X}_e}\right)^*(\lambda) &= -U_x\lambda^3 - U\lambda^3_x - V_y\lambda^3 - V\lambda^3_y\end{aligned}$$

The same schemes than in Subsubsection 5.1.2 are used to compute U and V. For the component $\mathbf{q}$ of the extended state vector an explicit Euler integration is applied with an upwind scheme:

$$\mathbf{q}_{i,j}^{k+1} = \mathbf{q}_{i,j}^{k} - \Delta t \left[\frac{\max(U_{i,j}^k, 0)\left(\mathbf{q}_{i,j}^k - \mathbf{q}_{i-1,j}^k\right) + \min(U_{i,j}^k, 0)\left(\mathbf{q}_{i+1,j}^k - \mathbf{q}_{i,j}^k\right)}{\Delta \mathbf{x}} + \frac{\max(V_{i,j}^k, 0)\left(\mathbf{q}_{i,j}^k - \mathbf{q}_{i,j-1}^k\right) + \min(V_{i,j}^k, 0)\left(\mathbf{q}_{i,j+1}^k - \mathbf{q}_{i,j}^k\right)}{\Delta \mathbf{x}} \right]$$

5.3 Comparison of the two strategies on synthetic data

A synthetic image sequence is obtained from the initial conditions, displayed in Figure 2, using the transport of brightness and advection of velocity equations. Part of the sequence is used for the image assimilation process and displayed on Figure 3.

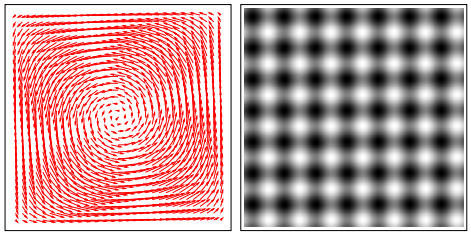

Figure 2: Left: initial motion field; Right: initial value of $\mathbf{q}$.

.

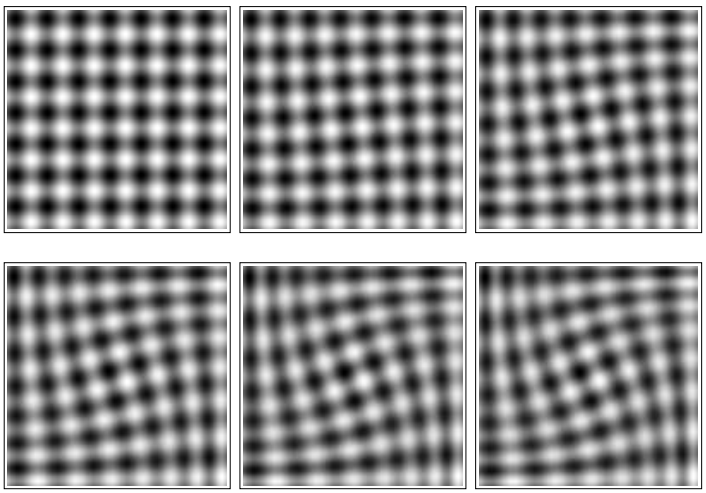

Figure 3: Sequence of images used for assimilation.

The methods described in Section 5.1 and 5.2 are respectively named IME (Image Model-External) and IMI (Image Model-Internal, since a mathematical quantity $\mathbf{q}$, comparable to image acquisition is included in the extended state

	Norm(%)	Orientation(in degree)
IME	0.279	2.128
IMI	0.098	0.792

Table 1: Error statistics.

vector). For both methods, the same images have been used for the assimilation process. The background value of velocity is $\mathbf{W}_b = 0$. For IMI, the background value $\mathbf{q}_b$ is the first image.

Results of motion estimation are displayed on Figure 4 and compared to the real motion field. Statistics are given in Table 1. The first column is the average of the relative error on the euclidian norm. The second column is the average of the absolute error on orientation. Slightly better results are obtained with IMI, due to the richness of the comparison of the state vector with the observations.

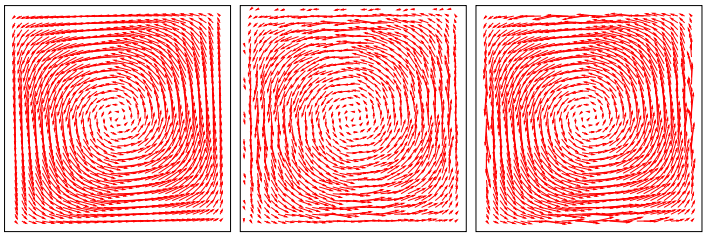

Figure 4: Left: Ground truth; Center: IME; Right: IMI.

.

5.4 Comparison of the two strategies on satellite data

The input data is a sequence of four satellite images acquired over the black Sea[2].

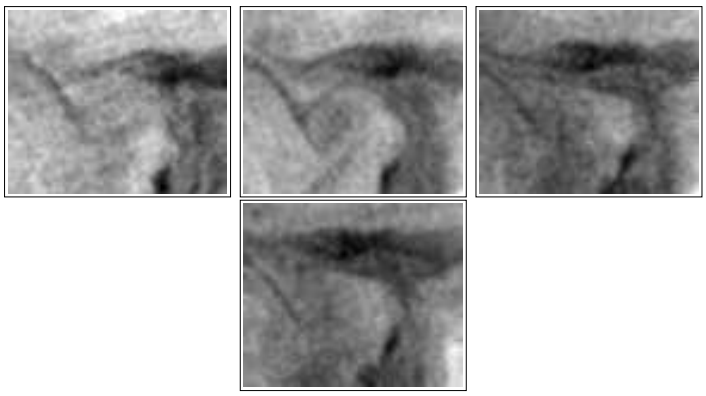

Figure 5: Satellite acquisitions.

[2]Data have been provided by E. Plotnikov and G. Korotaev from the Marine Hydrophysical Institute of Sevastopol, Ukraine.

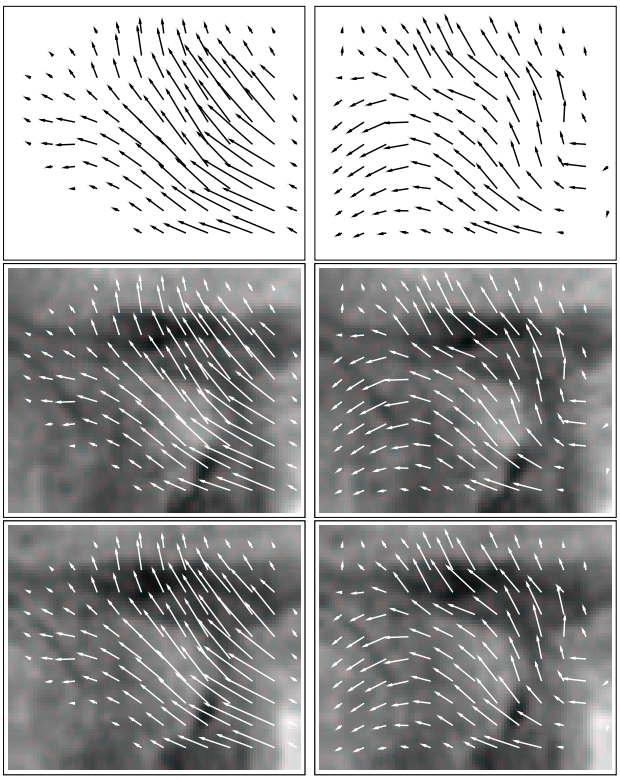

Figure 6: Left: result of IME, Right: result of IMI. Up: motion fields. Center: motion fields superposed to acquisition 3. Bottom: motion fields superposed to acquisition 4.

As previously, the background values are: $\mathbf{W}_b = 0$ for both models; $\mathbf{q}_b =$ *first image* for IMI. The two methods are applied and the retrieved motion fields are displayed on Figure 6. The counter-clockwise rotational motion occurring in the bottom-left part of the image sequence is correctly retrieved by IMI and missed by the IME method. The conclusion on satellite acquisition is the same than on synthetic data.

6 Conclusion

The paper describes and analyses various strategies that can be applied for deriving the dynamics from an image sequence and an image model with a variational data assimilation method.

In a first attempt, the data assimilation system, designed to retrieve a quantity $\mathbf{X}(\mathbf{x},t)$ linked by the equation $\mathbb{I}(\mathbf{X}(\mathbf{x},t),\mathbf{I}(\mathbf{x},t)) = 0$ to image data $\mathbf{I}(\mathbf{x},t)$, uses $\mathbf{X}$ as state vector, $\mathbf{I}(\mathbf{x},t)$ as observation $\mathbf{Y}(\mathbf{x},t)$ and $\mathbb{H}(\mathbf{X},\mathbf{Y}) = \mathbb{I}(\mathbf{X},\mathbf{I})$ as observation operator. This is a direct assimilation of image acquisition.

The alternative is to include a pseudo-image, similar to the image acquisition, in an extended state vector. This allows an easier comparison between the state vector and the observations. Different methods are then be derived, depending on the transformation applied to the real observations and pseudo-images. The main drawbacks are:

- the increase of the memory required for computation, due to the huge size of satellite image sequences.
- the difficulty to correctly handle the dynamics of the pseudo-image and design the observation error covariance matrix.

However results obtained on synthetic and real data show an improvement compared to those obtained with the direct assimilation. Moreover, large displacements can be better handled, because $\mathbf{q}$ is integrated in time with an adaptive time-step value, allowing a more precise quantification of transport.

As a first perspective of this work, a comparison will be done for various dynamic equations of $\mathbf{W}$ within IME and IMI formulations. In a second step, the two strategies will be implemented for a new application such as curve tracking.

References

[1] W.F. Ames. *Numerical Methods for Partial Differential Equations.* Academic Press, 1977.

[2] A. Apte, C.K.R.T. Jones, A.M. Stuart, and J. Voss. Data assimilation: Mathematical and statistical perspectives. *Int. J. Numer. Meth. Fluids*, 56:1033–1046, 2008.

[3] D. Béréziat and I. Herlin. Solving ill-posed image processing problems using data assimilation. *Numerical Algorithms*, 54, 2010.

[4] D. Béréziat, I. Herlin, and L. Younes. A generalized optical flow constraint and its physical interpretation. In *CVPR*, pages 2487–2492, 2000.

[5] I. Cohen and I. Herlin. Optical flow and phase portrait methods for environmental satellite image sequences. In *Proceedings of European Conference on Computer Vision*, Oxford, UK, April 1996.

[6] T. Corpetti, E. Mémin, and P. Pérez. Dense estimation of fluid flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):365–380, March 2002.

[7] B.K.P. Horn and B.G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[8] T. Isambert, J.P. Berroir, and I. Herlin. A multiscale vector spline method for estimating the fluids motion on satellite images. In *Proceedings of European Conference on Computer Vision*, Marseille, France, October 2008. Springer.

[9] F.-X. Le Dimet, I.M. Navon, and D.N. Daescu. Second-order information in data assimilation. *Monthly Weather Rev*, 130:629–648, March 2002.

[10] E. Mémin and P. Pérez. Optical flow estimation and object-based segmentation with robust techniques. *IEEE Trans. on Image Processing*, 7(5):703–719, May 1998.

[11] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005.

[12] O. Titaud, A. Vidard, I. Souopgui, and F.-X. Le Dimet. Assimilation of image sequences in numerical models. *Tellus A*, 62:30–47, 2010.

[13] X. Vigan, C. Provost, R. Bleck, and P. Courtier. Sea surface velocities from Sea Surface Temperature image sequences. *Journal of Geophysical Research*, pages 19499–19514, August 2000.

Centre de recherche INRIA Paris – Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
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 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

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399