ANALYSIS OF SST IMAGES BY WEIGHTED ENSEMBLE TRANSFORM KALMAN FILTER

Gorthi Sai, Sébastien Beyou, Etienne Mémin

ANALYSIS OF SST IMAGES BY WEIGHTED ENSEMBLE TRANSFORM KALMAN FILTER

Sai Subrahmanyam Gorthi, Sébastien Beyou and Étienne Mémin

INRIA, Campus Universitaire de Beaulieu, F-35042 Rennes, France.

ABSTRACT

This paper presents a novel, efficient scheme for the analysis of Sea Surface Temperature (SST) ocean images. We consider the estimation of the velocity fields and vorticity values from a sequence of oceanic images. The contribution of this paper lies in proposing a novel, robust and simple approach based on Weighted Ensemble Transform Kalman filter (WETKF) data assimilation technique for the analysis of real SST images, that may contain coast regions or large areas of missing data due to the cloud cover.

Index Terms— Data assimilation, Particle filters, Ensemble Kalman filters, Image motion analysis

1. INTRODUCTION

The analysis of geophysical fluid flows is of the utmost importance in domains such as oceanography, hydrology or meteorology for applications of forecasting, climate changes study, or for monitoring hazards or events. In all these domains orbital or geostationary satellites provide a huge amount of image data with a still increasing spatial and temporal resolution. Since several years there is a growing interest in extracting from those images a sequence of motion fields depicting the evolution of the observed fluid flow. Compared to *in situ* measurement techniques supplied by dedicated probes or Lagrangian drifters, satellite images provide a much more denser observation field. They however offer only an indirect access to the physical quantities of interest, and give consequently rise to difficult inverse problems to estimate characteristic features of the flow such as the velocity fields or its vorticity maps.

Motion estimation are classically based on the temporal conservation of the luminance between two successive images and assume other assumptions to obtain the unicity of the solutions [1] but, although these approaches provide consistent estimates over space, they can fail to be consistent in time and to reconstruct accurately the dynamical evolution law.

Dynamical consistency of the velocity measurements can be enforced by embedding the estimation problem within an image based assimilation process. Variational assimilations of image information have been recently considered for the estimation of consistent atmospheric motion fields [2]. On the other hand the stochastic data assimilation filters are also well known for incorporating temporal consistency. Recently, a data assimilation procedure embedding an Ensemble Kalman filter (EnKF) [3] into the particle filter (PF) framework, referred to as Weighted Ensemble Kalman filter (WEnKF), has been proposed [4] and applied for filtering noisy scenarios of fluid vorticity maps.

2. WETKF FOR SST IMAGES

This paper proposes a particle filter extension of the Ensemble Transform Kalman filter [5, 6]. This filter employs a nonlinear observation model that directly relies on SST images and allows us to directly extract the fluid velocity fields from a SST image sequence. Ensemble Transform Kalman filter (ETKF) is an ensemble implementation of the well known recursive Kalman filter (KF), similar to the Ensemble Kalman Filter [3], which minimizes the mean square error of the state estimates through a set of analytic equations, and contrary to the EnKF it doesn't use perturbed observations.

Data assimilation filters couple a dynamical model describing the state evolution and a measurement model that relates the state variables, vorticity in our case, to the observations. In the proposed approach we employ a stochastic filtered version of vorticity-velocity Navier-Stokes formulation as the dynamical model and rely directly on a nonlinear displaced frame reconstruction error of SST images for the measurement model. Vorticity maps are first sampled by combining the dynamics and the Kalman updates equations at the analysis step and then corrected through an importance sampling weighting provided by the observations likelihood, so the result of the ETKF analysis can viewed as the proposal sampling distribution of the particle filter. We further propose an uncertainty model which enables us to handle regions with missing data, while boundary conditions are imposed on coastal regions. This filter is referred as the Weighted Ensemble Transform Kalman Filter (WETKF).

In the following we describe the WETKF image assimilation process. The dynamics, the image model and the uncer-

tainty model used are successively presented. Finally results obtained on synthetic and real world oceanic SST images are shown and analyzed.

2.1. Assimilation process

The proposed assimilation process can be described as follows. At time $k = 0$, the N particles $\{\boldsymbol{x}_0^{a,(i)}, i = 1, ..., N\}$ are initialized with perturbed optical flow estimates of vorticity, obtained from images I_0 and I_1. From time $k = 1$ to T, we perform the following three steps for the WETKF assimilation of state (vorticity) $\widehat{\boldsymbol{x}}_k$ using the corresponding SST images $\boldsymbol{I}_k, \boldsymbol{I}_{k+1}$.

1. **Prediction:**

   Assuming a quasi-2D flow for which the analytical transport component has been removed from the data, we model the evolution of the SST images through a filtered vorticity-velocity formulation of the Navier-Stokes equation with a stochastic forcing function.

   The ensemble of states $\{\boldsymbol{x}_{k-1}^{a,(i)}, i = 1, ..., N\}$ of previous analysis are propagated through the vorticity dynamics ($d\xi = -\nabla\xi \cdot \boldsymbol{v}dt + \nu\Delta\xi dt + \eta_k dB$), to supply the forecast ensemble members $\boldsymbol{x}_k^{f,(i)}$, expressed at time k as $\boldsymbol{x}_k^{f,(i)} = \sum_{t=k-1}^{k-\delta t}\left(M(\boldsymbol{x}_t^{f,(i)}) + \sigma_{t+\delta t}^{(i)} d\boldsymbol{B}\right); \boldsymbol{x}_{k-1}^{f,(i)} = \boldsymbol{x}_{k-1}^{a,(i)}$.

   Here, the state vector, $\boldsymbol{x}_k$, represents the vorticity, ξ, at each image grid point coordinates, the operator $M(\boldsymbol{x}_t^{f,(i)})$ denotes a discrete scheme of the dynamical operator and the random forcing term is a spatially correlated and temporally uncorrelated Gaussian field [3].

2. **Update:**

   This step aims at modifying the ensemble members sampling through a Kalman correction step incorporating the available observations. As we assume quasi-2D non-diverging flows, the observation model on which we rely is derived from the brightness consistency assumption: $\boldsymbol{I}(\boldsymbol{x}, k) = \boldsymbol{I}(\boldsymbol{x} + \boldsymbol{d}(\boldsymbol{x}), k+1) + \gamma d\boldsymbol{B}$; here $\boldsymbol{d}(\boldsymbol{x}) = \int_{k-1}^{k-\delta t} \boldsymbol{v}dt$, integrated up to the (discrete) time k. The diffusion matrix $\boldsymbol{\gamma}$ is diagonal and corresponds to a Gaussian smoothing of the error. Its components are given by $\gamma(\boldsymbol{x}, \boldsymbol{x}) = g_\sigma \star |I(\boldsymbol{x}, k) - I(\boldsymbol{x} + \boldsymbol{d}(\boldsymbol{x}), k+1)|$, and $d\boldsymbol{B}$ is an uncorrelated n-dimensional white noise.

   It is important to point out that this is a nonlinear measurement model. As already noticed ([7]), for a nonlinear model, all the terms needed in the Kalman update can be performed by replacing $\mathbf{H}\mathbf{X}^f$ with $\mathbf{H}(\mathbf{X}^f)$, where $\mathbf{H}(\mathbf{X}^f) = \frac{1}{\sqrt{N}}\{\mathbf{H}(\mathbf{x}^{f,(i)}) - \overline{\mathbf{H}(\mathbf{x}^{f,(i)})}, i = 1 \ldots N\}$.

The ETKF update [6] is performed with this observation model to supply the proposal ensemble sampling $\boldsymbol{x}_k^{a,(i)}$ (this is referred as the analysis step in the ensemble filters terminology and corresponds in our case to the proposal density sampling stage of our particle filter [4]). For the application of Kalman filter based recursion, the major hurdle in Geophysical applications lies in the inversion of $n \times n$ dimensional covariance term $(\boldsymbol{H}(\boldsymbol{X}^f)\boldsymbol{H}(\boldsymbol{X}^f)^T + \boldsymbol{R})^{-1}$, related to the Kalman gain computation, where n usually is very high.

Following [6] an efficient computation of this inversion is implemented through the Sherman—Morison–Woodbury formula. When the inverse of the measurement error covariance matrix $\boldsymbol{R}$ is available (as in our case since it is diagonal) this strategy requires only inversion of dimension of the squared cardinality of the ensemble members ($N \times N$), which is much smaller than the original covariance dimension (*i.e.*, $N << n$).

3. **Particle Filter Weights and Resampling:**

   After this proposal sampling stage, the ensemble members are weighted with a mean square displaced image difference based likelihood and are resampled according to the importance sampling rules of the particle filter [4]. Finally, an empirical mean of the analysis ensemble members, $\boldsymbol{x}_k^{a,i}$, weighted accordingly to their importance weights, ω_i, defines the vorticity estimate at time k: $\widehat{\boldsymbol{x}}_k = \sum_i \omega_i \boldsymbol{x}_k^{a,i}$. The corresponding velocity field $\widehat{\boldsymbol{v}}_k$ is deduced through the Biot-Savart kernel.

2.2. Observation uncertainties for missing data

High uncertainties at points corresponding to occluded regions (e.g. SST regions hidden by the cloud cover) are imposed explicitly. The strength of these uncertainties is set to a value proportional to the number of occluded points over a centered local window of dimension $(2l + 1) \times (2l + 1)$. The magnitude of the proportionality factor is scaled to the highest uncertainty of the SST data (excluding occluded points). The variance of the Gaussian smoothing in the uncertainty model is proportional to this strength, in order to propagate the boundary uncertainties to occluded regions. At coast regions a mask is created with zero velocity values over land regions (Dirichlet boundary condition). The dynamic update and the measurement correction are frozen on those regions.

3. RESULTS

In this section, we present our WETKF based direct image assimilation results on real oceanic images, with large missing data regions and with coasts (zero motion regions). Previously, in order to assess quantitatively the performances of our approach, we developed synthetic image sequences through backward propagation of real SST images from a

simulated 2D turbulent flow. Furthermore, in order to mimic areas of missing data we randomly introduce "holes" into those sequences. These simulated images strictly follow the Navier-Stokes dynamics and displaced image difference observation model, these sequences comply thus perfectly with the dynamics and the measurement model. Let us note, nevertheless that the forcing sustaining the 2D turbulence is unknown. The initial condition is also unknown.

3.1. Synthetic images with missing data

The size of the synthetic ocean images we have considered is 200x200 pixels, while the square holes were centered in 4 different locations with dimensions of 25, 29, 33, and 43 square pixels. Their locations have been kept constant over two subsequent images, and then randomly altered. Thus we will have a synthetic sequence with these four holes consistently at same location over a pair of subsequent images and randomly changing for the next pair of images and so on. Figure 1(a) and (b) depicts two subsequent images, for which the locations of holes have been changed.

(a)   (b)   (c)

(d)   (e)
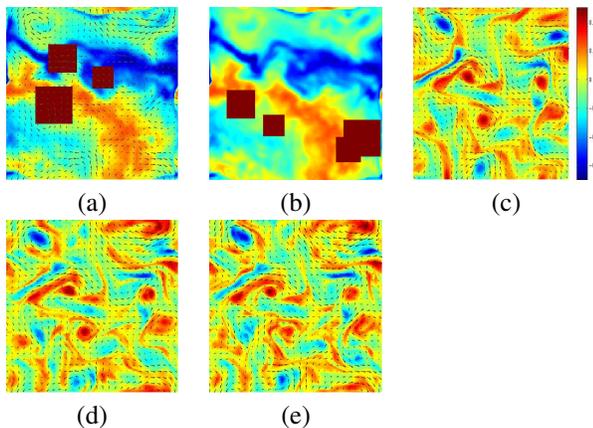
Fig. 1. Simulated ocean images with missing data regions: (a) Image 97 with estimated field (b) Image 98 (c) True vorticity and field; WETKF estimated Vorticity and field (d) without holes (e) with holes

An instance of the true vorticity with its corresponding motion field superimposed is presented in Figure 1(c). The corresponding vorticity and velocity fields extracted by the proposed WETKF based direct image assimilation without and with missing data regions are displayed in Figures 1(d) and (e), respectively. It is very interesting to note that the temporal consistency of the WETKF assimilation has made it possible to extract the vorticity and the motion fields with a very close accuracy, even in presence of larger areas with missing data. The RMSE comparison of the errors in WETKF estimates without and with holes in the synthetic ocean sequence are plotted in Figure 2. Its is noteworthy to highlight that the accuracy in motion field is almost same as without the holes in images.

(a)   (b)
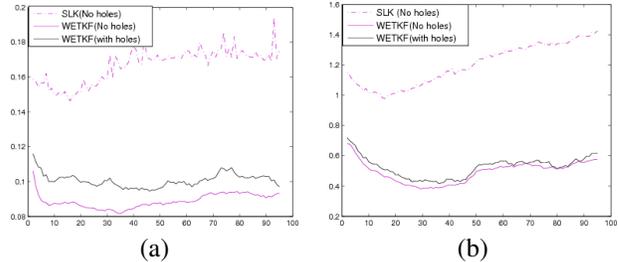
Fig. 2. RMSE comparison of Simulated holes from ocean images: (a) RMSE in vorticity with and without introducing holes (b) RMSE in velocity with and without holes

In our experiments, we noticed that the motion estimates or the vorticity can be extracted with small compromise in accuracy even if the holes remain at same locations even up to 5 images. However, if the holes reside at the same location throughout the sequence, the accuracy of the estimates in that region and its neighborhood is obviously seriously affected. However other regions are not impacted.

3.2. Ocean images

(a)   (b)   (c)

(d)   (e)   (f)
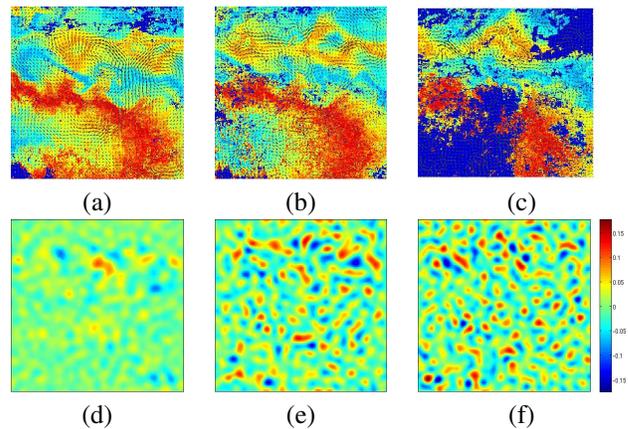
Fig. 3. SST images in the Pacific ocean during a period corresponding to El Niño Southern Oscillation: images superimposed with WETKF direct image assimilation estimated fields of (a) initial image #1 (b) image #10 (c) image #20; and corresponding vorticity of (d) initial image #1 (e) image #10 and (f) image #20

We now present the results on real data, with modified strategies in WETKF assimilation. The oceanic sea surface temperature images we considered corresponds to two areas of the Pacific ocean off the Panama isthmus during an episode of El Niño Southern Oscillation. The intensity of images is being proportional to the temperature of the regions. We have processed a set of 48 images; the spatial resolution is about 10 km/pixel (0.1 degree) and the temporal latency is 24 hours. Some representative results including the initial image

are shown in Figure 3. Figure 3 (a), (b) and (c), shows the estimated fields #1, #10, and #20 superimposed on the corresponding images, while the Figure 3 (d), (e) and (f) show their associated vorticity maps. The initialization is provided by a stochastic Lucas–Kanade approach [8]. It is interesting to note, from the upper half region, that though the motion at the initial image appears to be too smooth and describes only large-scale structures, the intensity pattern is well tracked and the vorticity maps are refined along time. Further, in spite of important missing data regions, the motion on the whole image is recovered.

Finally, we consider a portion of this oceanic data, in Figure 4, with coasts as well as missing data regions. Further, this portion also shows strong vortices. We plot two typical processed images, images #10 and #25 with their motion fields superimposed in Figure 4 (a,b). The corresponding vorticity maps are shown in Figure 4 (c,d), respectively. As is well observed most of the strong vortices corresponds to seemingly proper image regions. It is even more interesting to note that in the first image results (Figure 4 (a,d)) strong vortex at the top middle region has been well recovered.
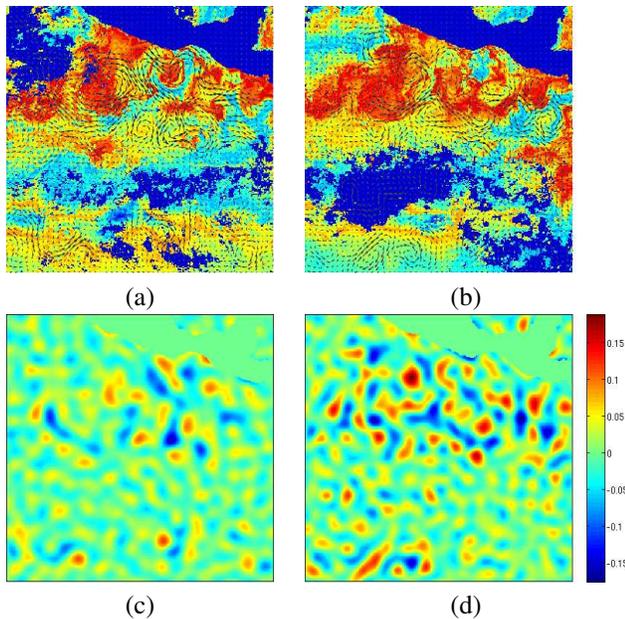
Fig. 4. SST images of the Pacific ocean off the Panama isthmus: velocity fields estimated by WETKF superimposed on the (a) image #10 (b) image #25; corresponding vorticity maps (c,d).

4. CONCLUSION

In this paper we have presented a novel assimilation method defined as a combination of the Ensemble Transform Kalman Filter and a Particle filter. This filter incorporates in this study a nonlinear direct image observation. It provides good results on real SST sequences and particularly shows the ability to handle correctly missing data over large regions. Further developments could include the use of a more realistic oceanic dynamics.

5. REFERENCES

[1] D. Heitz, E. Mémin, and C. Schnörr, "Variational fluid flow measurements from image sequences: synopsis and perspectives," *Experiments in Fluids*, vol. 48, pp. 369–393, 2010, 10.1007/s00348-009-0778-3.

[2] T. Corpetti, P. Héas, E. Mémin, and N. Papadakis, "Pressure image assimilation for atmospheric motion estimation," *Tellus A*, vol. 61, no. 1, pp. 160–178, 2009.

[3] G. Evensen, "The Ensemble Kalman Filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, pp. 343–367, 2003, 10.1007/s10236-003-0036-9.

[4] N. Papadakis, E. Mémin, A. Cuzol, and N. Gengembre, "Data assimilation with the weighted ensemble Kalman filter," *Tellus A*, vol. 62, no. 5, pp. 673–697, 2010.

[5] C. H. Bishop, B. J. Etherton, and S. J. Majumdar, "Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects," *Monthly Weather Review*, vol. 129, no. 3, pp. 420–436, 2001.

[6] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, "Ensemble Square Root Filters," *Monthly Weather Review*, vol. 131, no. 7, pp. 1485–1490, 2003.

[7] P. L. Houtekamer and H. L. Mitchell, "A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation," *Monthly Weather Review*, vol. 129, no. 1, pp. 123–137, 2001.

[8] T. Corpetti and E. Mémin, "Stochastic models for local optical flow estimation," *Conf. on Scale Space and Variational Methods, SSVM'11*, June 2011.