

Interpolated swell fields from SAR measurements

Pierre Tandeo, René Garello
and Ronan Fablet

Institut Mines-Telecom - Telecom Bretagne
CNRS UMR 6285 LabSTICC - Pôle CID
Technopôle Brest Iroise - CS 83818
29238 BREST Cedex - FRANCE
Email: pierre.tandeo@telecom-bretagne.eu

Romain Husson and Fabrice Collard
Collecte Localisation Satellites
Technopôle Brest Iroise
29238 Brest Cedex - France

Bertrand Chapron
IFREMER
Spatial Oceanography Laboratory
Technopôle Brest Iroise
29238 Brest Cedex - France

Abstract—Synthetic Aperture Radar (SAR) sensors on-board satellites are very well suited for observing sea surface geophysical parameters such as ocean swell. But on a very large scale, SAR data are too sparse for deriving some global information. From the original work of [4], and following some generic assumption on the physics of the swell propagation in deep water, it was shown that using a back-propagating scheme, it was possible to retrieve the source of the swell system and then generate a propagating field. In this paper, we are proposing a simpler and original approach, by assimilating the SAR data into a given swell field and then using a Kalman Filter/Smoothing technique for updating the main parameters of the swell (wavelength, direction, and significant wave height) within the complete field. This method shows very encouraging results which will be confronted with *in situ* measurements when available.

I. INTRODUCTION

The observation of ocean swell using Synthetic Aperture Radar (SAR) on-board satellites has been demonstrated since the ERS-1 mission in 1992. In particular, ocean swell spectra can be derived from SAR images using a quasi-linear transformation. Unfortunately, SAR data are very sparse, acquired at different times and with heterogeneous quality. This statement demonstrates the need to assimilate SAR data, for instance using a third generation ocean wave model (cf. [1] and [2]). An alternative is also possible. Under deep water and no current assumptions, swell can propagate for thousands of kilometers from the storm source. This idea was exploited in [4] to retro-propagate SAR observations (backward in time) to the most accurate storm identification and then to re-propagate them (forward in time) along the complete storm event within the corresponding basin. This analysis was so-called “fireworks”. Then, [7] uses these pseudo-observations to generate complete spatial fields of sea swell using a spatio-temporal interpolation.

In this work, we propose an original methodology to assimilate the SAR observations and interpolate swell fields. As in the fireworks analysis, we only use the along track SAR observations. Our idea is to use these SAR observations to update the integral parameters of the swell, considered as unknown statistical variables. The evolution of these variables are computed sequentially on the grid analysis. Each pixel is seen as a potential storm source and the swell is forwardly propagated along a direction and a group velocity. Using a Ensemble Kalman Filter (EnKF), we artificially create pseudo

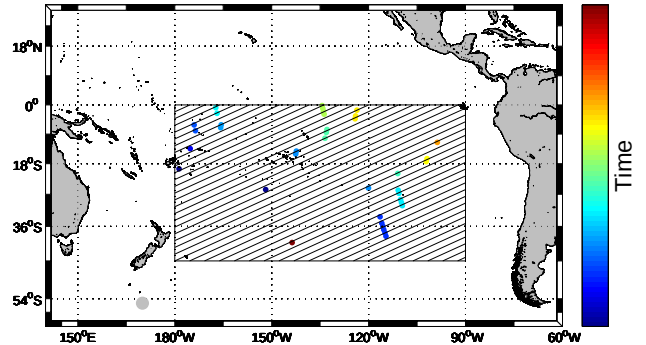


Fig. 1. SAR along track observations from the 13 July (blue) to the 18 July 2004 (red). The swell data are generated by a storm located in 55°S, 170°E (gray dot) into the Southern Pacific Ocean. The data are corresponding to swell partitions with periods close to 17 s.

SAR observations using a random generator and we follow the most probable solution. Then, we proceed in the backward direction using a similar random approach involving an Ensemble Kalman Smoother (EnKS).

The paper is organized as follows. Section II presents the methodology used in this article. It corresponds to a statistical model based on SAR data and resolved by EnKF/EnKS. In Section III, we apply the methodology and evaluate the results. We further discuss and summarize the key results of our investigations in Section IV.

II. METHOD

A. SAR data

We use ocean swell data provided by the Advanced-SAR instrument on-board ENVISAT via the wave mode (cf. [9] and [8]). The swells are characterized by three integral parameters: the significant height H_s , the wavelength λ and the direction θ of the swell. The data used as case-study in this work correspond to a particular storm in July 2004 into the Southern Pacific Ocean (situations studied in [4]). During this event, we keep into account the swell partitions with periods close to 17 s. The SAR observations are very sparse in space and time (cf. Fig. 1).

B. State space model

In this work, we estimate the integral parameters of the swell on a regular spatio-temporal grid. Thus, we define a spatial grid of p pixels into the domain covered by the SAR observations (hatched domain in Fig. 1). Then, for each analysis time t_k , we define the unknown n -dimensional state vector $\mathbf{x}(t_k)$ with $n = 4 \times p$. It corresponds to the fields of significant height H_s , wavelength λ and the zonal and meridional components of the direction, respectively noted θ_u and θ_v . The state evolution in time is given by

$$\mathbf{x}(t_k) = \mathcal{M}(\mathbf{x}(t_{k-1})) + \boldsymbol{\eta}(t_k), \quad (1)$$

where $\mathcal{M}$ corresponds to (i) the swell propagation in deep water with a group velocity $\sqrt{\lambda g}/(2\sqrt{2\pi})$ and (ii) the swell energy decay proportional to $1/(\alpha \sin(\alpha))$ with α the angular great-circle distance from the storm source (supposed to be known) as proposed by [4]. The state evolution given in Eq. (1) is physically realistic but not completely deterministic. For instance, in the current version of the state operator $\mathcal{M}$, we do not take into account either the swell dissipation, the surface current or the island shadow effects. Therefore, at each time t_k , we add a random perturbation noted $\boldsymbol{\eta}(t_k)$. We assume that the perturbations are Gaussianly distributed with zero mean and a constant in time $n \times n$ covariance matrix $\mathbf{Q}$. At the initial time of Eq. (1), we introduce an *a priori* knowledge of the integral parameters. We assume that this background information follows a Gaussian distribution given by the mean $\mathbf{x}^b$ and the covariance matrix $\mathbf{B}$. In this study, the initial condition is given by the fireworks analysis (cf. [7]). It corresponds to a consistent and realistic prior information.

At irregular time, as described in Fig. 1, the integral parameters of the swell are partially observed. They are noted $\mathbf{y}$. The state vector at time t_k is related to the observation by the observation equation defined by

$$\mathbf{y}(t_k) = \mathcal{H}(\mathbf{x}(t_k)) + \boldsymbol{\epsilon}(t_k), \quad (2)$$

where the observation operator $\mathcal{H}$ is the spatial interpolation from the analysis to the observation grid. We know that the SAR observations have a systematic error. Therefore, in Eq. (2), we suppose that the stochastic random vector $\boldsymbol{\epsilon}(t_k)$ is an additive zero mean Gaussian error characterized by the covariance matrix $\mathbf{R}$. In the current version of the algorithm, the covariance matrices $\mathbf{B}$, $\mathbf{Q}$ and $\mathbf{R}$ are supposed to be diagonal and, as explained in [4], correspond to standard deviation of 0.3 m for H_s , 36 m for λ and 17° for θ . Note that for simplicity, Eq. (1) is written with a discrete evolution of the state in time. In reality, this equation is continuous and manages the irregular time sampling of the SAR observations. More precisely, we use an exponential decrease factor for $\mathbf{Q}$, proportional to the time lag between two consecutive observations and/or analysis (see [11] for more details).

C. Ensemble Kalman filter

We use the sequential EnKF algorithm proposed by [3] to propagate the state in a forward way. As in [10], we use a

version of the EnKF where the state evolution $\mathcal{M}$ and the observation $\mathcal{H}$ operators do not need to be linearized. In the initial step of the EnKF algorithm, at time t_1 , an ensemble of $\mathbf{x}$'s composed by N members is randomly generated. The members of the ensemble follow a Gaussian distribution given by the vector mean $\mathbf{x}^b$ and the covariance matrix $\mathbf{B}$. The N initial members are stored in the vectors $\mathbf{x}_i^f(t_1) \forall i \in \{1, \dots, N\}$. Then, we proceed forward from $k = 1$ to $k = K$ using the update step and the analysis step as described below. In the update step, at each time t_k , we randomly generate N samples of $\boldsymbol{\eta}_i$ and $\boldsymbol{\epsilon}_i \forall i \in \{1, \dots, N\}$ with respective covariances $\mathbf{Q}$ and $\mathbf{R}$. Then, following Eq. (1), the i -member of the updated state is given by

$$\mathbf{x}_i^f(t_k) = \mathcal{M}(\mathbf{x}_i^f(t_{k-1})) + \boldsymbol{\eta}_i(t_k), \quad (3)$$

and the mapping from the forecast state space to the observational space of the i -member is computed as

$$\mathbf{y}_i^f(t_k) = \mathcal{H}_k(\mathbf{x}_i^f(t_k)). \quad (4)$$

The N members of the ensemble are used to estimate the sample means of the propagated state in the state space and in the observational space denoted by $\mathbf{x}^f(t_k)$ and $\mathbf{y}^f(t_k)$ respectively. In the analysis step, we follow [10] methodology which avoids the linearization of the observational operator. The Kalman gain is computed from

$$\mathbf{K}(t_k) = \mathbf{P}_{xy}^f(t_k) (\mathbf{P}_{yy}^f(t_k) + \mathbf{R})^{-1}, \quad (5)$$

where $\mathbf{P}_{xy}^f(t_k)$ is the sample cross-covariance matrix and $\mathbf{P}_{yy}^f(t_k)$ is the sample covariance matrix, which are determined by

$$\begin{aligned} \mathbf{P}_{xy}^f(t_k) &= \frac{1}{N-1} \\ &\times \sum_{i=1}^N \left(\mathbf{x}_i^f(t_k) - \mathbf{x}^f(t_k) \right) \left(\mathbf{y}_i^f(t_k) - \mathbf{y}^f(t_k) \right)^\top \end{aligned} \quad (6)$$

and

$$\begin{aligned} \mathbf{P}_{yy}^f(t_k) &= \frac{1}{N-1} \\ &\times \sum_{i=1}^N \left(\mathbf{y}_i^f(t_k) - \mathbf{y}^f(t_k) \right) \left(\mathbf{y}_i^f(t_k) - \mathbf{y}^f(t_k) \right)^\top. \end{aligned} \quad (7)$$

Having $\mathbf{K}(t_k)$ from Eq. (5), the N members of the ensemble are then updated by

$$\mathbf{x}_i^a(t_k) = \mathbf{x}_i^f(t_k) + \mathbf{K}(t_k) \mathbf{d}_i(t_k) \quad (8)$$

where $\mathbf{d}_i(t_k) \forall i \in \{1, \dots, N\}$ are the N innovation vectors in which we use perturbed observations such as $\mathbf{d}_i(t_k) = \mathbf{y}(t_k) + \boldsymbol{\epsilon}_i(t_k) - \mathbf{y}_i^f(t_k)$. Finally, the updated analyzed state is represented by the sample mean $\mathbf{x}^a(t_k)$ and the sample covariance $\mathbf{P}^a(t_k)$.

D. Ensemble Kalman smoother

The backward recursions correspond to the sequential EnKS algorithm proposed by [5]. It uses the results of the EnKF computed above. In the initial step of the EnKS algorithm, at time t_K , we use the members of the filtered state, $\forall i \in \{1, \dots, N\}$, such as $\mathbf{x}_i^s(t_K) = \mathbf{x}_i^a(t_K)$ and $\mathbf{P}^s(t_K) = \mathbf{P}^a(t_K)$. Then, we proceed backward from $k = K - 1$ to $k = 1$ using the analysis step and we compute

$$\mathbf{x}_i^s(t_k) = \mathbf{x}_i^a(t_k) + \mathbf{K}^s(t_k) \left(\mathbf{x}_i^s(t_{k+1}) - \mathbf{x}_i^f(t_{k+1}) \right) \quad (9)$$

where $\mathbf{K}^s(t_k)$ is the Kalman smoother gain matrix given by $\mathbf{P}_{xx}^a(t_k) \left(\mathbf{P}^f(t_{k+1}) \right)^{-1}$ with

$$\begin{aligned} \mathbf{P}_{xx}^a(t_k) &= \frac{1}{N-1} \\ &\times \sum_{i=1}^N \left(\mathbf{x}_i^a(t_k) - \bar{\mathbf{x}}^a(t_k) \right) \left(\mathcal{M}(\mathbf{x}_i^a(t_k)) - \mathcal{M}(\bar{\mathbf{x}}^a(t_k)) \right)^\top. \end{aligned} \quad (10)$$

The Gaussian distribution of the updated state estimate is given by the sample mean and covariance respectively denoted by $\bar{\mathbf{x}}^s(t_k)$ and $\mathbf{P}^s(t_k)$.

III. RESULTS

Here we apply the state space model given in (1)-(2) to assimilate the SAR observations using the EnKF/EnKS with $N = 1000$ members. The time-lag between two analysis is 3 h and the analysis grid is $3^\circ \times 3^\circ$. The results of the estimated integral parameters of the swell for the middle time of the assimilation window are given in Fig. 2. It corresponds to the results of the EnKS, i.e. the $\bar{\mathbf{x}}^s$ vector and the standard deviation of the diagonal elements of the $\mathbf{P}^s$ matrix. In the areas with low standard deviation (East part of the analysis), the results are consistent with the outputs of the WaveWatch 3 model and the fireworks analysis (not shown in this paper). In a future work, we plan to compare precisely these different analysis with buoy data from the National Oceanographic Data Center (NODC).

IV. CONCLUSION AND PERSPECTIVES

In this work, we assumed a stochastic swell field process characterized by integral parameters (significant height, wavelength, direction) sparsely observed by SAR measurements. Compared to the fireworks analysis which generates pseudo-observations using backward and then forward propagations, our methodology only assimilate SAR data once without reprocessing them. Then, our sequential model uses the memoryless Markovian property and each pixel of the grid analysis is seen as a potential storm source where we generate random perturbations of the integral parameters. The results show a good consistency with the fireworks analysis and the wave model outputs. Although the physic used in our model is simple, it avoids the use of a complete wave model which can be critical in a data assimilation scheme with a large number of members.

The perspectives for future works will consist of using another interesting property of the state space model proposed here: the ability to assimilate easily other data sources. For instance, we plan to combine SAR measurements, buoy data from NODC, altimeters and seismographs as proposed in [6]. These data sources can observe totally or partially the integral parameters (for instance, only H_s for altimeters) with different error measurements and space-time sampling.

ACKNOWLEDGMENTS

We would like to thank the GlobWave project for providing SAR data. This study was supported by the Institut Mines-Telecom and the Marine Collaborative Ground Segment (MCGS) project.

REFERENCES

- [1] L. Aouf, J. M. Lefevre, D. Hauser and B. Chapron, "On the combined assimilation of RA-2 altimeter and ASAR wave data for the improvement of wave forecasting," *Proceedings of 15 years of progress in Radar Altimetry Symposium*, vol. 15, pp. 13–18, 2006.
- [2] E. Bauer, K. Hasselmann, I. R. Young and S. Hasselmann, "Assimilation of wave data into the wave model WAM using an impulse response function method," *Journal of Geophysical Research*, vol. 101, pp. 3801–3816, 1996.
- [3] G. Burgers, P. J. Van Leeuwen and G. Evensen, "Analysis scheme in the ensemble Kalman filter," *Monthly Weather Review*, vol. 126, pp. 1719–1724, 1998.
- [4] F. Collard, F. Ardhuin and B. Chapron, "Monitoring and analysis of ocean swell fields from space: New methods for routine observations," *Journal of Geophysical Research*, vol. 114, C7, 2009.
- [5] G. Evensen and P. J. Van Leeuwen, "An ensemble Kalman smoother for nonlinear dynamics," *Monthly Weather Review*, vol. 128, pp. 1852–1867, 2000.
- [6] R. Husson, F. Ardhuin, F. Collard, B. Chapron and A. Balanche, "Revealing forerunners on Envisat's wave mode ASAR using the Global Seismic Network," *Geophysical Research Letters*, vol. 39, pp. L15609, 2012.
- [7] R. Husson, "Development and validation of a global observation-based swell model using wave mode operating Synthetic Aperture Radar," *PhD dissertation*, 2012.
- [8] H. Johnsen, G. Engen, F. Collard, V. Kerbaol and B. Chapron, "ENVISAT ASAR Wave Mode Products-quality assessment and algorithm upgrade," *Proceedings of SEASAR*, 2006.
- [9] V. Kerbaol, H. Johnsen, B. Chapron and B. Rosich, "Quality assessment of ENVISAT ASAR Wave mode products based on regional and seasonal comparisons with WAM model outputs," *Proceedings of ERS/Envisat Symposium*, pp. 6–10, 2004.
- [10] D. T. Pham, "Stochastic methods for sequential data assimilation in strongly nonlinear systems," *Monthly Weather Review*, vol. 129, pp. 1194–1207, 2001.
- [11] P. Tandeo, P. Ailliot and E. Autret, "Linear Gaussian state-space model with irregular sampling: application to sea surface temperature," *Stochastic Environmental Research and Risk Assessment*, vol. 25, pp. 793–804, 2011.

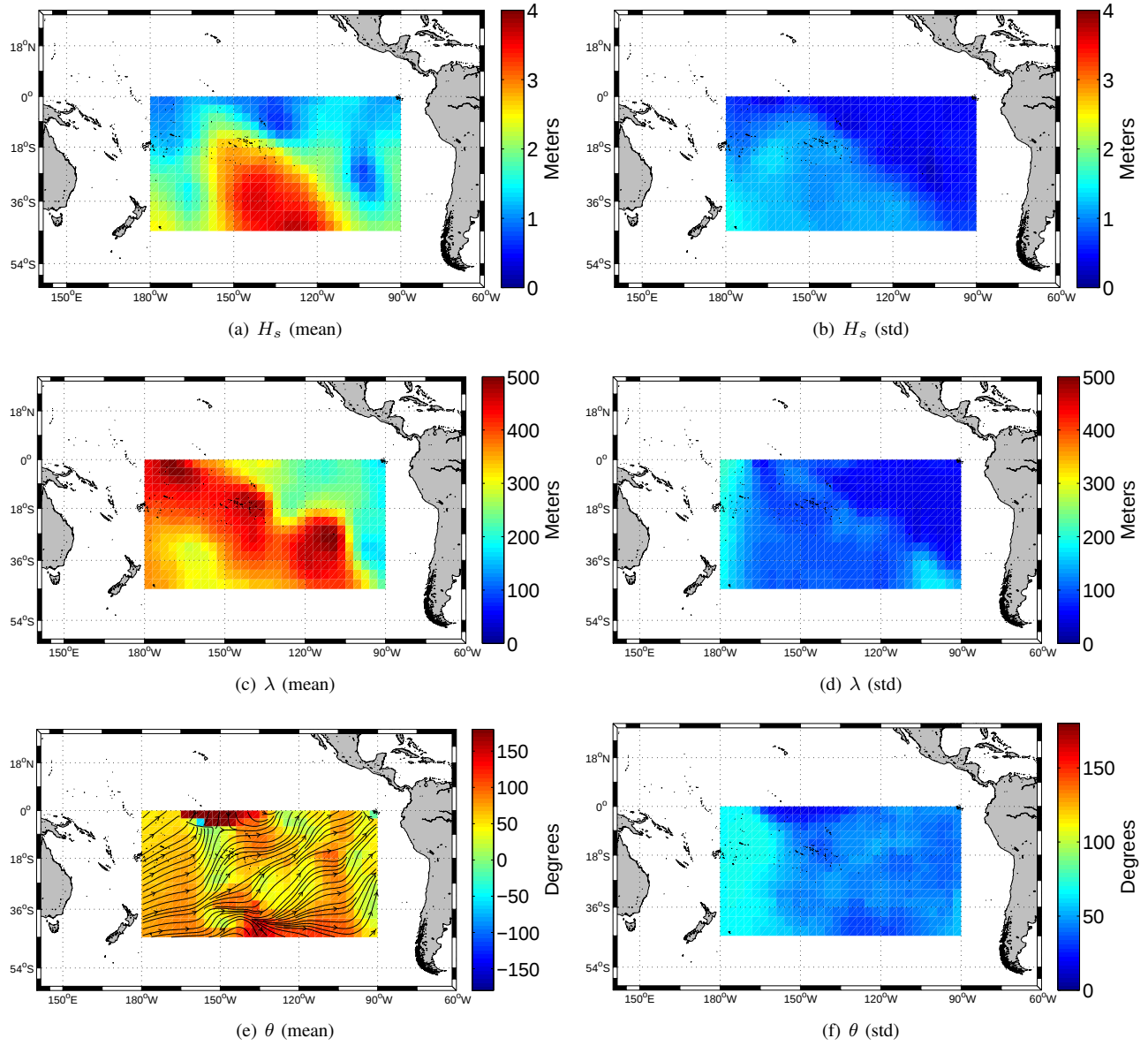


Fig. 2. Estimated mean (first column) and standard deviation (second column) of the significant height (first row), wavelength (second row) and direction with the corresponding streamlines (third row) the 15th of July, 2004 at 1200 UTC.