Achimer
http://archimer.ifremer.fr

Space-time models for moving fields with an application to significant wave height fields

Ailliot Pierre [1, *], Baxevani Anastassia [2], Cuzol Anne [3], Monbet Valerie [3], Raillard Nicolas [1, 3, 4]

[1] Univ Europeenne Bretagne, UMR 6205, Math Lab, Brest, France.
[2] Univ Gothenburg, Chalmers Univ Technol, Dept Math Sci, Gothenburg, Sweden.
[3] Univ Europeenne Bretagne, UMR 3192, Lab STICC, Vannes, France.
[4] Laboratoire d'Oc´eanographie Spatiale, IFREMER, France

* Corresponding author : Pierre Ailliot, email address : ailliot@univ-brest.fr

Abstract :

The surface of the ocean, and so such quantities as the significant wave height, equation image, can be thought of as a random surface that develops over time. In this paper, we explore certain types of random fields in space and time, with and without dynamics that may or may not be driven by a physical law, as models for the significant wave height. Reanalysis data is used to estimate the sea-state motion which is modeled as a hidden Markov chain in a state space framework by means of an AR(1) process or in the presence of the dispersion relation. Parametric covariance models with and without dynamics are fitted to reanalysis and satellite data and compared to the empirical covariance functions. The derived models have been validated against satellite and buoy data.

Keywords : space-time model, significant wave height, state-space models

1 Introduction

Spatio-temporal modeling is an important area in statistics that is one of rapid growth at the moment, with various applications in environmental science, geophysical science, biology, epidemiology and others. See for instance [11], [15] and [17].

Especially after all the recent technological advances such as satellite scanning that resulted in increasingly complex environmental data sets, estimating and modeling the covariance structure of a space-time process have been of great interest although not as well developed as methods for the analysis of purely spatial or purely temporal data. Because it is often difficult to think about spatial and temporal variations simultaneously, it is tempting to focus on the analysis of how the covariances at a single point vary across time or how the covariances at a single time vary across space. If these were the only characteristics that mattered, then separable models would suffice. Allowing though the merely spatial and merely temporal covariances to define the space-time dependence is a severe restriction, that is not satisfied by many geophysical processes, such as meteorological systems, rainfall cells, air pollution, etc., that exhibit motion. Hence the need for nonseparable covariance models which include interactions between the spatial and the temporal variability, see [10, 24, 17] among others for recent contributions.

In this paper, we explore methods for constructing models for the significant wave height, a parameter related to the energy of the sea-state, based on fitting random field models to data collected from different sources. A full description of the data, some aspects of its limitations and some assumptions that are reasonable in modeling it are given in the next subsection. The models are then described and interpreted in terms of these assumptions.

An important feature of the significant wave height fields, non-compatible with the assumption of separability, is their motion. The apparent motion of the significant wave height fields is actually the composition of various motions, that of the wind fields (see [1]) that generate the waves and those of the various wave systems that compose each sea-state. In order to simplify the analysis though, we consider that each sea-state moves with a single velocity that is the composition of the ones mentioned above.

In the area of interest, a part of the North Atlantic Ocean, westerly winds are pre-

vailing and low pressure systems are generally moving to the East. As a consequence, the significant wave height fields are also moving to the East, although the important variability in the meteorological conditions in this area implies also an important variability in the speed and direction at which the sea-states are traveling. In this study, we propose the use of the output of numerical weather forecast systems to estimate the resulting sea-state motions. Although these numerical models are sometimes inaccurate, we show that they provide enough information on the state of the atmosphere and ocean in order to get at least a rough estimate of the prevailing motion.

Using a Lagrangian reference frame, instead of a Eulerian (fixed) reference frame, appears to be natural for modeling moving processes. Such an idea has already been used in the literature (see e.g. [17] and references therein), but it is generally assumed that the motion is constant in time (*frozen velocity*) and eventually in space. The Lagrangian reference frame moves with the sea-states, and as a consequence we expect longer range dependence in the temporal domain than in the Eulerian reference frame. Indeed, we show that the main difference between the covariance structures for the two reference frames is a slower decrease to zero with time for the Lagrangian one, and that changing the reference frame to the Langrangian actually leads to more accurate space-time interpolation.

Another originality of this work is that we combine different data sources that provide information at different scales: reanalysis (or *hindcast*) data, satellite data and buoy data a short description of which is given in subsection 1.1. In Section 2, a covariance model with constant velocity is introduced. The method used to estimate the motion of the sea-states and the field covariance model in the associated Lagrangian reference frame are discussed in Section 3. A comparison of the ability of the different models to produce accurate space-time interpolations is presented in Section 4, where is also shown that the model with changing velocity outperforms the one with constant velocity. Finally, some conclusions are presented in Section 5.

1.1 Data

The data used in this work come from three different sources that are briefly described next.

- **Satellite data.** The observations consist of the significant wave height taken at
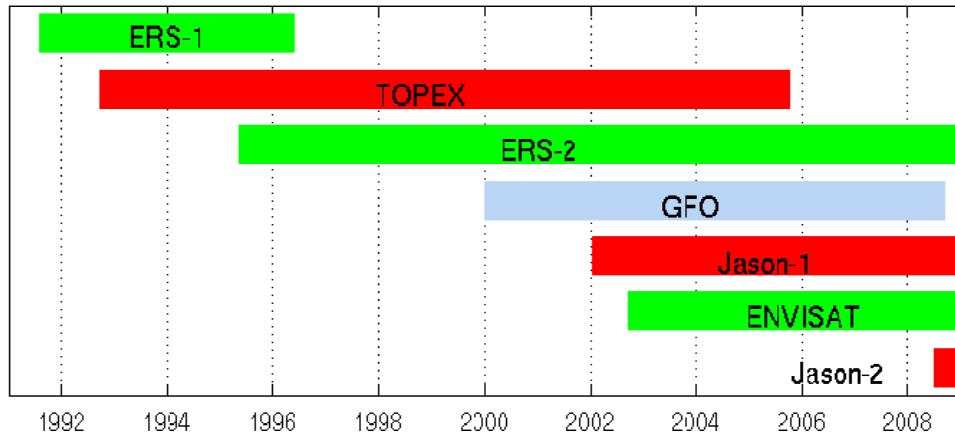
3

Figure 1: Time periods when altimeter data are available for each satellite.

discrete locations along one-dimensional tracks from seven different satellite altimeters that have been deployed progressively since 1991 and whose operation times can be seen in Figure 1.

These tracks have two orientations and those with the same orientation are roughly parallel. For convenience, we shall use the term passage for the observations from one pass along a single track. It should be noted that, for each track, each passage may have observations at slightly different locations (in terms of longitude and latitude), which are consequently neither equidistant nor, from passage to passage, identical. Each observation collected is a summary statistic from a sampling window of about 7 kilometres by 7 kilometres. These windows are about 5.5-7 kilometres apart and so the windows can overlap marginally. This smooths the observations for the passage and introduces some short term dependence. Another characteristic of the data is that the observations have been discretised. This, unfortunately, has removed the very small scale variation and, in many places, especially where the observations for a passage are low, has resulted in flat sections in the sample path.

Moreover, satellite altimeter data have been calibrated using buoy data and the adequacy is generally satisfying (see [29]). More information on the data sets and their particulars can be found in the URL :

`ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves`

- **Hindcast data.** Recently, a wave reanalysis 6-hourly data set on a global 1.5×1.5
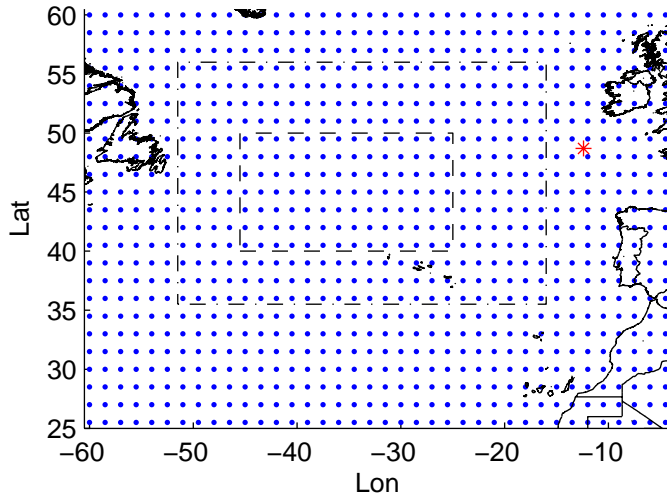
4

Figure 2: Grid of ERA-Interim data and domain of interest D_0 (dashed box). The dashed-dotted box corresponds to the domain D where the velocity fields are computed in Section 3.2 and the * indicates the location of buoy 62029.

latitude/longitude grid (see Figure 2) covering the period of 1957 to 2002 was made available, the ERA-40 data set. This reanalysis was carried out by the European Centre for Medium-Range Weather Forecasts (ECMWF) using its Integrated Forecasting System, a coupled atmosphere-wave model with variational data assimilation. Shortly after, new progress was made by producing the ERA-Interim data set which has progressed beyond the end of ERA-40 data set and now covers the period up to 2007. A distinguished feature of ECMWF's model is its coupling through the wave height dependent Charnock parameter (see [18]), to a third generation wave model, the well-known WAM ([20]), which makes wave data a natural output of both the ERA-40 and the ERA-Interim system, the latter having a variational bias correction using satellite data. A large subset of the complete ERA-Interim data set, including significant wave height estimates, can be freely downloaded and used for scientific purposes at the URL `http://data.ecmwf.int/data/`. In this work we use the ERA-Interim data set from 1992 until 2007.

- **Buoy data.** Buoy data are often considered as a reference to provide ground-truth for hindcast and satellite data. In this study, we use K1 buoy data (station 62029), which is part of the UK Met Office monitoring network. It is located at position

5

(48.701 N, 12.401 W) (see Figure 2) and provides hourly significant wave height estimates. In this work, we only consider data for the period from April 2002 until December 2007.

A typical example of the data coverage over a 24 hour time window can be seen in Figure 3. Hindcast data, like the ERA-Interim data, are sampled over a regular 1.5×1.5 degrees grid at synoptic times every six hours starting at midnight, in contrast to the irregular, both in space and time, sampling provided by the satellite altimeter. However, the ERA-Interim data set tends to underestimate the variability of the significant wave height (see next Section) and it only provides information at a synoptic scale whereas satellite data also give smaller scale information. For the two data sets to become more compatible, the satellite data have been smoothed in order to eliminate small scale components. In practice, a moving average filter with a window of size 10 observations, which covers a distance of about 50 km, has been applied on each track and, to reduce computational coast, the data have been under-sampled so that only one observation every 10 is used.

Different analyses have been presented studying the correlation structure of the sea surface energy as measured by the significant wave height. For example [13] carried out a comparison of model (ERA-WAM, the 15 year first version of the ERA-40) and satellite climatologies (provided by the Southampton Oceanography Centre from altimeter data) to find that the model data showed similar tendencies when compared to altimeter data. Moreover, the ERA-40 data have been extensively validated against observations and other reanalysis data sets, ([8, 7]), and turns out that performs well when compared with measurements from in-situ buoy and global altimeter data. It is reasonable therefore to assume that the dynamics and the shape of spatio-temporal covariance structure of the field is the same as those computed using the significant wave height estimates from the ERA-Interim data set. However, we keep in mind that the satellite altimeter provides with significant wave height estimates that are more accurate, although not as regular, and use these data sets for the final estimation of any parameters entering the covariance structure of the underlying field.

This principle is illustrated in the next section, where an empirical estimate of the spatio-temporal covariance structure is obtained using the ERA-Interim data set and
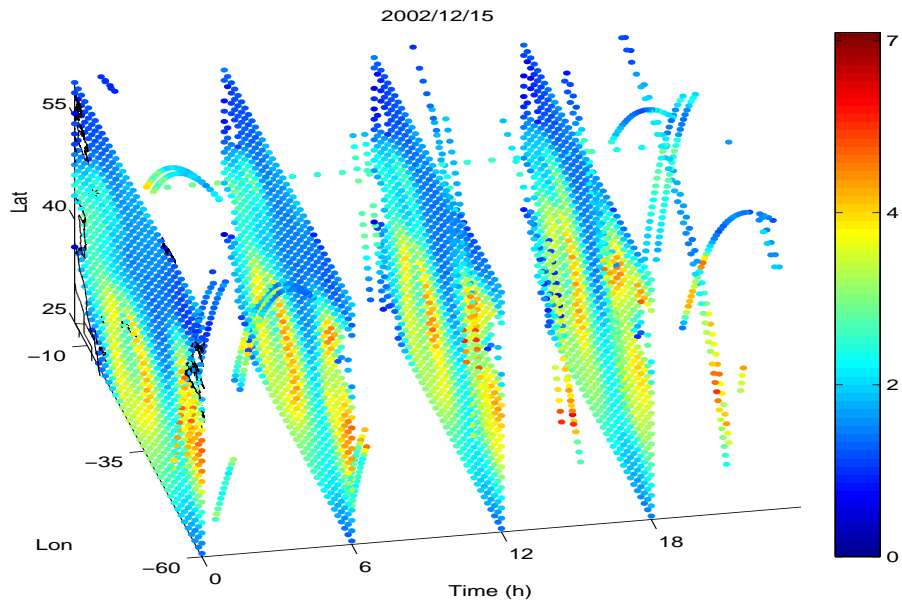
Figure 3: 3D representation of the available data for 2002/12/15. The 2D fields at times 0, 6, 12, 18 correspond to ERA-Interim data, the 24 observations at location (48.701 N, 12.401 W) correspond to buoy data and the other data to the various satellite tracks.

is used to check for the presence of stationarity and isotropy of the underlying field. The affirmative answer simplifies the model considerably making it possible to compute again the empirical covariance of the field using satellite data and then fit an appropriate parametric model.

Buoy data will be used only in Section 4.2 as a reference to compare the accuracy of the significant wave height obtained by interpolating satellite data using the different space-time models. There are some questionable extremely high values (above 20 meters!) in the time series considered, so in order to filter out these values we have applied a moving median filter, with a time-window of width 6 hours. This is also expected to bring the temporal scale closer to the one of ERA-Interim data set.

7

2   Model with constant velocity

The observations consist of the significant wave heights taken at discrete locations either along the one-dimensional satellite tracks or at the ERA-Interim grid points. In order to simplify the initial analysis, we only consider a central region of the North Atlantic, with latitudes ranging between 25W and 45.5W and longitudes ranging between 40N and 50N (see Figure 2), which from now on will be denoted by D_0. It would seem unlikely that the field is stationary in time, its characteristics are expected to change seasonally at least, although probably inter-annual components (see e.g. [2]) are also present. The annual cycle generally dominates the within-year variability of the significant wave height field, especially in the northern hemisphere, see [5]. One way of dealing with seasonal components is to fit an annual cycle to the data. An alternative way, which is employed here, is to focus on one month at a time and consider the data from that month over the different years as independent realizations of the same time-stationary random field. This is a usual assumption for meteorological processes, despite the fact it does not take into account the low frequency variation due to inter-annual variability (trend, NAO, etc). In this paper we present results only for the month of December.

Figure 4 shows the empirical mean and standard deviation at each grid point in the area D_0, computed using the ERA-Interim significant wave height data set from the 16 December months (period 1992-2007). It is apparent that the data are non-stationary in space, the mean values and variability in the north of D_0 appear to be higher than in the south. Notice the correlation between the magnitude of the mean values of significant wave height and variability. High mean values correspond to high variability and areas with calmer conditions like in the south of D_0 have smaller variance. The data have been standardized locally, that is at each grid point, by removing the mean and scaling by the standard deviation.

In order to make the two estimates, from the ERA-Interim and the satellite altimeter, compatible, we have considered 1.5×1.5 degree boxes centered at the ERA-Interim grid points and used all the satellite data that fall inside each one to obtain estimates of the mean and standard deviation, for the 16 December months. These estimates, which can be seen on Figure 5, although they exhibit an important spatial variability, which may
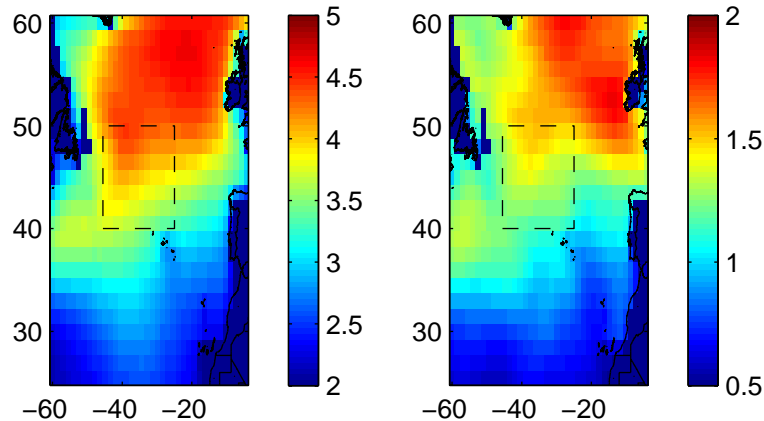
Figure 4: *Mean (left panel) and standard deviation (right panel) of H_s computed using ERA-Interim data (December 1992-2007). The dashed box indicates the domain D_0.*

	ERA-Interim	Satellite
Mean	4.25	4.39
Variance	2.29	3.07

Table 1: Sample mean and sample variance of H_s in D_0 over the months of December for the period 1992-2007.

be due to the low space-time sampling of the satellite data (the sample size for each box can be seen in Figure 5, (right panel)), present the same overall spatial trend with the ERA-Interim data, compare to Figure 4, although the actual values of the mean and variance are quite different as can be seen in Table 1.

This sampling variability could probably be reduced by either increasing the box size or by using some type of spatial smoothing. Here we decided to scale the data by removing the mean-field estimates obtained using the ERA-Interim data to which the difference of the two overall means (i.e. the difference between the two readings in the first row of Table 1) was added, and then scale by dividing by the ERA-Interim standard deviation corrected by the ratio of the two overall standard deviations given in Table 1 (the two readings on the second row).

In the following, we shall consider that the standardized versions of the significant wave heights are partial observations of a random field $Y(\mathbf{p}, t)$ where $\mathbf{p} \in D_0 \subset \mathbb{R}^2$ and
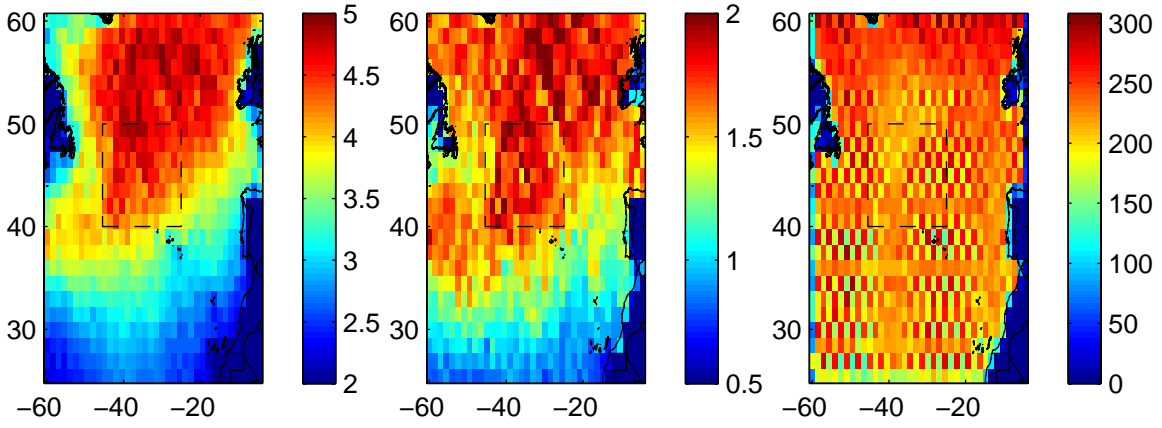
Figure 5: *Mean (left panel) and standard deviation (middle panel) of H_s computed from satellite data (December 1992-2007). The right panel is the sample size for each 1.5×1.5 degree box. The dashed boxes indicate the domain D_0.*

$t \in T$, measured in days. Hence,

$$E(Y(\mathbf{p}, t)) = 0$$
$$\text{Var}(Y(\mathbf{p}, t)) = 1.$$

We further assume that the field is stationary both in space and time, i.e. for all $\mathbf{p}, \mathbf{p}' \in D_0$ and $t, t' \in T$

$$\text{Cov}(Y(\mathbf{p}, t), Y(\mathbf{p}', t')) = C_Y(\mathbf{p}' - \mathbf{p}, t' - t). \tag{1}$$

To check this simplifying assumption, we have computed the covariance $\text{Cov}(Y(\mathbf{p}, t), Y(\mathbf{p}', t'))$ for various values of $\mathbf{p}$ and t as functions of $\mathbf{p}'$ and t'. The resulting estimates seem relatively independent of the choice of $\mathbf{p}$ and t.

An empirical estimate of C_Y for different time lags computed using the ERA-Interim data, can be seen in Figure 6. Notice that the maximum of the spatial covariance $C_Y(\cdot, t)$ drifts to the West as the time lag increases. This is due to the prevailing sea-state motion to the East. Such behavior strengthens the arguments for using a covariance structure that includes some sort of dynamics, like the covariance

$$C_Y(\mathbf{p}' - \mathbf{p}, t' - t) = C(\mathbf{p}' - \mathbf{p} - \mathbf{V}_0(t' - t), t' - t) \tag{2}$$

where $\mathbf{V}_0$ denotes the mean velocity of the sea-state, and which can be thought as the covariance of a static field subordinated by the constant velocity $\mathbf{V}_0$. I.e., let $X(\mathbf{p}, t)$ be a

10

field with spatio-temporal covariance function $C(\mathbf{p}-\mathbf{p}',t-t')$, then the field $X(\mathbf{p}-\mathbf{V}_0 t,t)$ has covariance function

$$\text{Cov}(X(\mathbf{p}-\mathbf{V}_0 t,t),X(\mathbf{p}'-\mathbf{V}_0 t',t')) = C(\mathbf{p}'-\mathbf{p}-\mathbf{V}_0(t'-t),t'-t).$$

Various parametric covariance models have been considered for C, including some usual separable models as well as the non-separable model proposed in [17]. The best fit to the empirical covariance function using standard weighted least-square method (see e.g. [14]), has been obtained for the simple nonseparable rational quadratic model

$$C(\mathbf{p}'-\mathbf{p},t'-t) = (1-C_0)\mathbb{1}_{\{0,0\}}(\mathbf{p}'-\mathbf{p},t'-t) + \frac{C_0}{1 + \frac{d(\mathbf{p},\mathbf{p}')^2}{\theta_S^2} + \frac{|t'-t|^2}{\theta_T^2}}. \tag{3}$$

The parameters θ_S and θ_T describe the spatial and temporal range respectively, $1-C_0$ is the space-time nugget effect and $d(\mathbf{p},\mathbf{p}')$ denotes the geodesic distance on the sphere between locations $\mathbf{p}$ and $\mathbf{p}'$. The nugget effect may model small-scale structures, those usually not observed in the ERA-Interim data because of the space-time sampling resolution, and the measurement error which may be present in the satellite data. Eventually, different nugget effects could be used for the different satellites.

Model (3) has been fitted to the empirical covariance function, see Figure 6. The elliptic level curves of the parametric model seem to fit the overall shape of the empirical covariance function both in space and time and in particular the model seems to be able to describe the mean motion of the sea-state. The estimated mean velocity $\mathbf{V}_0$ corresponds to a mean motion of $27.8\ km.h^{-1}$ to the East and $1.3\ km.h^{-1}$ to the North, which seems to be a physically realistic velocity for that area.

Notice that model (3) becomes isotropic after removing the mean motion, whereas Figure 6 suggests the presence of a small anisotropy. And indeed a slightly better fit was achieved by using an anisotropic model. However, because of the geometry of the satellite tracks that provide information only along two main directions, the assumption of isotropy simplifies considerably the estimation of the covariance function using satellite data (see also [3]).
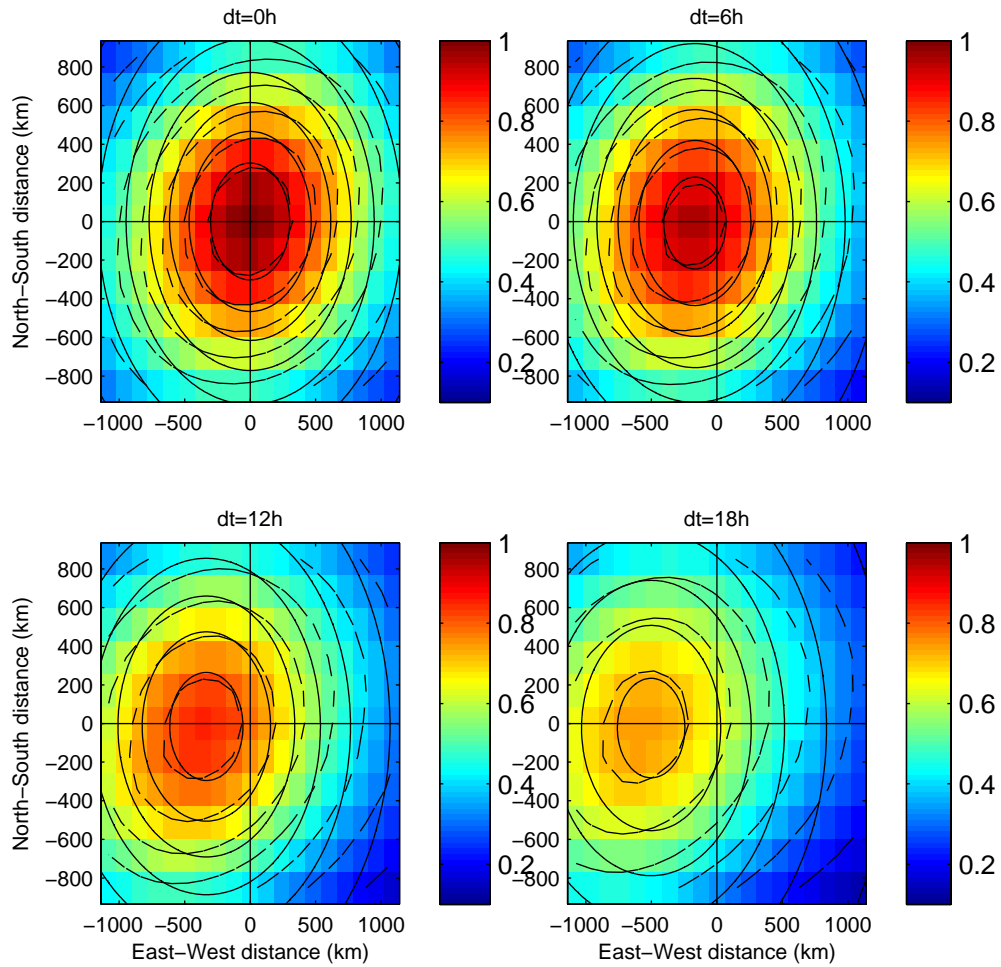
Figure 6: *Empirical covariance function for ERA-Interim data for time lags $|t - t'| = 0$ hour (top left panel), $|t - t'| = 6$ hour (top right panel), $|t - t'| = 12$ hour (bottom left panel) and $|t - t'| = 18$ hour (bottom right panel). The dashed lines represent the levels of the empirical covariance function and the full lines the levels of the fitted parametric model.*

3 Model with dynamic velocity

In the previous section, we incorporated the motion of the sea-states into the model by considering a static field subordinated by a constant velocity. Such a field is still not optimal since the assumption of constant velocity over such large region and for a time span greater than 10 hours, is not realistic, see [4]. In this section, we propose a new approach by subordinating a static field with a dynamically changing velocity. The section is organized as follows: first we define velocity through a flow of diffeomophisms that are the solution to the transport equation and discuss ways to incorporate these velocity fields into the covariance model. Then, we describe a method for estimating the sea-state velocity fields within a state-space model framework, with the aim towards a spatio-temporal consistency.

3.1 Dynamic velocity

In this subsection we introduce sea-state motion using a flow of diffeomorphisms. Denote by $\boldsymbol{\phi}(\mathbf{p}, t, s)$ the motion of the point $\mathbf{p}$ between times t and s in the interval $[t_0 - \Delta T, t_0 + \Delta T]$ for some arbitrary reference point t_0. Assume also that the flow satisfies $\boldsymbol{\phi}(\mathbf{p}, t, t) = \mathbf{p}$ and the flow property $\boldsymbol{\phi}(\cdot, t, s) = \boldsymbol{\phi}(\cdot, u, s) \circ \boldsymbol{\phi}(\cdot, t, u)$ and that the inverse of $\boldsymbol{\phi}(\mathbf{p}, t, s)$ exists and is denoted by $\boldsymbol{\phi}^{-1}(\mathbf{p}, t, s)$. This construction is sensible both from the physical and mathematical point of view. The physical analogy is that the sea-state develops over time. Mathematically for every diffeomorphism $\boldsymbol{\phi}$ there exist a velocity field $\mathbf{V}(\mathbf{p}, t) = (u(\mathbf{p}, t), v(\mathbf{p}, t))$ such that (see [26]):

$$\boldsymbol{\phi}(\mathbf{p}, t, s) = \mathbf{p} + \int_t^s \mathbf{V}(\boldsymbol{\phi}(\mathbf{p}, t, \tau), \tau) \, d\tau. \tag{4}$$

This differential equation has been employed to describe sea-state dynamics in [4] and in [19] to solve the landmark matching problem.

For mathematical convenience we will make assumptions on the flow of diffeomorphisms which may be unrealistic from a physical point of view. For example, the sea-state motions should be differentiable and define one to one transformations: every sea-state at time t should be uniquely matched to a sea-state at time t' and thus different sea-states cannot "cross" as they do in the real world. There may also be some boundary problems

since sea-states appear or disappear from the domain D_0 as they travel. In order to avoid loosing too much data when the sea-states move, in the next section the velocity field is computed on a bigger domain D which includes D_0. In practice, D has been chosen in order to ensure that the flow of diffeomorphisms $\phi(\mathbf{p}, t_0, t)$ is generally defined for all $\mathbf{p} \in D_0$ ad $t \in [t_0 - \Delta T, t_0 + \Delta T]$ for ΔT less than one day.

Using the flow of diffeomorphisms defined in (4), a natural generalization of the frozen covariance model discussed in Section 2 is obtained by assuming that

$$\text{Cov}(Y(\mathbf{p}, t), Y(\mathbf{p}', t')) = C(\phi^{-1}(\mathbf{p}', t_0, t') - \phi^{-1}(\mathbf{p}, t_0, t), t' - t) \tag{5}$$

for the covariance function C given in (3). The physical interpretation of this model is that the covariance function depends on the locations where points $\mathbf{p}'$ and $\mathbf{p}$ were at time t_0 before arriving at times t' and t respectively to their current locations. For velocity field constant in space, i.e. $\mathbf{V}(p, t) = \mathbf{V}_0$ for all $\mathbf{p}, t$, the covariance function in (5) simplifies to the covariance function in (2). Relation (5) is equivalent to assuming C is the covariance function of the field in the Lagrangian reference frame, i.e. the covariance function of the field Z defined by

$$Z(\mathbf{p}, t) = Y(\phi(\mathbf{p}, t_0, t), t). \tag{6}$$

Here we should notice that the covariance given in (5) is generally a non-stationary one, except in the particular case the flow of diffeomorphisms is such that $\phi^{-1}(\mathbf{p}', t_0, t') - \phi^{-1}(\mathbf{p}, t_0, t)$ is only a function of $t' - t$ and $\mathbf{p}' - \mathbf{p}$. This is the case for example when the velocity field is constant in space (i.e. $\mathbf{V}(\mathbf{p}, t) = \mathbf{V}(t)$ for all $\mathbf{p}, t$), case which includes the constant velocity model introduced in Section 2.

This apparent contradiction with the stationarity assumptions which have been made in Section 2 may be better understood if we consider the velocity field as a random field, in which case the covariance function in (5) should be understood as a conditional expectation given the velocity field, and a more correct way of writing the model is

$$\text{Cov}(Y(\mathbf{p}, t), Y(\mathbf{p}', t')) = E[\text{Cov}(Z(\phi(\mathbf{p}, t_0, t), t), Z(\phi(\mathbf{p}', t_0, t'), t'))] \tag{7}$$

where the expectation is taken with respect to the random field ϕ. Writing a proper stochastic model for the velocity field $\mathbf{V}$ for which (7) leads to a second-order stationary space-time covariance function for Y will be the subject of future work.

3.2 Motion estimation

After having introduced the motion of sea-states through a flow of diffeomorphisms that are the solution to the transport equation, in this subsection we present a method for estimating the associated velocity field.

3.2.1 Framework

To estimate the velocity of the sea-state motion we use the ERA-Interim data presented in Section 1.1, since the coverage of the satellite data is generally poor and do not provide enough information to track correctly the motion of the sea-state systems. The regular coverage of the ERA-Interim data does not only provide us with more information but also simplifies the modeling since it allows the use of existing models for processes on regular space-time lattices. The objective is nevertheless to use the estimated velocity fields to model the satellite data.

The most widely used technique for motion estimation is based on maximizing the "local" correlation between two rectangular regions in successive images, and this method has been used in particular for meteorological and oceanographic applications , see e.g. [25, 30]. In parallel, in the computer vision community, the problem of motion estimation has been addressed using differential methods. It consists in solving a partial differential equation (PDE) system which is built assuming conservation of the intensity of the displaced object and a certain spatial regularity of the flow, suggestions for which in the particular case of fluid motion can be found in [12].

One drawback of the above mentioned methods is that the velocity fields are estimated independently of each other, using only the information contained in pairs of successive images. As a consequence the temporal consistency of the estimated velocity fields is not guaranteed although is expected since the wave systems are traveling with almost constant velocity. Indeed the local correlation method when applied to the ERA-Interim data failed to reproduce the temporal consistency that is usual in geophysical flows. Different classes of methods allow to deal with this lack of coherence by introducing dynamical information on the velocity fields using a priori information that may come from physical knowledge. The variational methods for instance, are related to optimal control theory [23]. In that framework, a sequence of motion fields is estimated knowing an initial state, a dynamical

model, and noisy and possibly incomplete observations. Dynamically consistent motion estimates are then obtained on a fixed time interval, given observations over the whole period [21, 28].

In this work, the motion estimation problem is formulated and solved sequentially within a state-space model framework. The hidden state is the velocity field, which is supposed to be a Markovian process with a transition kernel that is parameterized using a simple physical model to warranty that the velocity fields evolve slowly in space and time. The hidden state (velocity field) is related to the observations (ERA-Interim data) through a conservation of the characteristics of the moving sea-states between successive times. Then, the velocity fields are estimated using a particle filter which permits to compute approximations of the distribution of the hidden state given the observations. State-space models have already been used for motion estimation in [1] and [16]. In [1], the method is applied to a similar hindcast data set, although the velocity was supposed to be constant in space and the hidden state was discretized for computational reasons. In [16], the hidden velocity fields are guided by an a priori dynamical law constructed from fluid flow equations.

3.2.2 State-space model

In this subsection, we describe the state-space model which has been used to estimate the sea-state motion.

We commence by introducing some notation. Let $D = (\mathbf{p}_1, ..., \mathbf{p}_N)$ be the set of the $N = 336$ grid points of the ERA-Interim data set which are located between longitude 16.5W and 51W and latitude 36N and 55.5N (see Figure 2). This domain has been chosen empirically so that the motion of the sea states in D_0 can be followed on a ± 24 hour time window. Moreover, let $V(D,t) = (u(\mathbf{p}_1,t), \cdots, u(\mathbf{p}_N,t), v(\mathbf{p}_1,t), \cdots, v(\mathbf{p}_N,t))^T$ be the vector corresponding to the velocity field at time t, with u and v denoting the zonal and meridional components respectively.

Usually for the description of the sea-state we have available only some statistics related to the spectrum governing the sea surface process. Most often these are the significant wave height (H_s), the mean direction of propagation (Θ_m), and the mean period (T_m) which are included among others in the ERA-Interim data set. Let us then

denote by

$$S(D,t) = (H_s(\mathbf{p}_1,t)\cos(\Theta_m(\mathbf{p}_1,t)), \cdots, H_s(\mathbf{p}_N,t)\cos(\Theta_m(\mathbf{p}_N,t)),$$
$$H_s(\mathbf{p}_1,t)\sin(\Theta_m(\mathbf{p}_1,t)), \cdots, H_s(\mathbf{p}_N,t)\sin(\Theta_m(\mathbf{p}_N,t)),$$
$$T_m(\mathbf{p}_1,t), \cdots, T_m(\mathbf{p}_N,t))^T$$

the vector that describes the sea-state conditions at time t over the region D (see [22]). In practice, this information is available at discrete times with a regular time step of six hours.

Dynamics of the hidden velocity field The time evolution of the hidden state V is modeled using a mixture of two models. The first model is a physical approximation of the velocity of a wave group which is valid for sea-states with narrow band spectrum (see e.g. [31]). Hereafter, $\mathbf{V}_{disp}(\mathbf{p}_i,t) = (u_{disp}(\mathbf{p}_i,t), v_{disp}(\mathbf{p}_i,t))^T$ denotes the group velocity at time t and location $\mathbf{p}_i$ and $\mathbf{V}_{disp}(D,t)$ the associated velocity field obtained by concatenating the velocity at the different locations. The group velocity is related to the mean period and direction of the sea-state through the dispersion relation:

$$\mathbf{V}_{disp}(\mathbf{p}_i,t) = \frac{g}{4\pi}T_m(\mathbf{p}_i,t)\begin{pmatrix} \cos(\Theta_m(\mathbf{p}_i,t)) \\ \sin(\Theta_m(\mathbf{p}_i,t)) \end{pmatrix} \tag{8}$$

where g is the gravitational constant. Equation (8) should provide a good approximation of the velocity of the sea-state when a unique swell system and calm wind conditions are dominating. However, such weather conditions are not always prevailing in the North Atlantic in which case a simple AR(1) model may provide a better approximation of the sea-state dynamics.

To be specific, let C_t denote a Bernoulli variable which governs the choice of the model at time t, then

$$V(D,t) = C_t V_{disp}(D,t) + (1 - C_t)\left[\mu + A\left(V(D, t - \Delta t) - \mu\right)\right] + \epsilon_t \tag{9}$$

with $A \in \mathbb{R}^{2N \times 2N}$, $\mu \in \mathbb{R}^{2N}$, and $\{\epsilon_t\}$ a Gaussian white noise sequence with zero mean and covariance matrix Σ^{dyn}. We further assume that $\{C_t\}$ is an i.i.d. sequence of Bernoulli

variables with parameter π_c which is independent from both the velocity and the sea-state fields. Note that particle-based multiple model filters may also introduce a Markov structure on $\{C_t\}$ (see [27] for instance).

Observation equation   In order to relate the hidden velocity fields to the observed sea-state conditions, we assume that the characteristics of a sea-state evolve slowly between two time instants if we follow correctly its motion, i.e. that for all $i \in \{1, ..., N\}$,

$$S(\mathbf{p}_i, t) = S(\mathbf{p}_i - \mathbf{V}(\mathbf{p}_i, t)\Delta t, t - \Delta t) + \eta(\mathbf{p}_i, t). \tag{10}$$

When $\eta(\mathbf{p}_i, t)$ denote the evolution of the sea-state fields between two time steps, then it is further assumed that $\{\eta\}$ is a Gaussian white noise sequence with covariance matrix Σ^{obs}.

Parametrization   Let us now give more details on the parametrization of the various matrices of parameters which appear in the state-space model:

- **Autoregressive matrix** A. We assume a block structure

$$A = \begin{pmatrix} A_{1,1} & 0 \\ 0 & A_{2,2} \end{pmatrix}$$

with $A_{1,1} = A_{2,2}$ and $A_{1,1}(i,j) \propto \exp(-\frac{d(\mathbf{p}_i, \mathbf{p}_j)}{\lambda_A})$ and the normalizing constraint $\sum_{j=1}^{n} A_{1,1}(i,j) = \theta$ for $i \in \{1, ..., n\}$ and some constant $0 < \theta < 1$. The velocity field at time t is thus obtained by smoothing the velocity field at time $t - \Delta t$, using weights which decrease with distance, and satisfy a certain constrain $\theta < 1$ in order to warranty the stability of the AR model. Due to the difficulties presented in setting up an automatic procedure for estimating the unknown parameters, both θ and λ_A, as well as the parameters entering the covariance matrices Σ^{dyn} and Σ^{obs}, have been chosen to agree with the results of previous studies, see [5, 3] and references therein. So the values assigned are $\theta = 0.9$ and $\lambda_A = 200km$.

- **Mean vector** μ. μ denotes the mean of the stationary distribution of the AR(1) model. We assume that this mean is the same at all locations and we used the

parameter values obtained when fitting the model with constant velocity (see Section 2).

- **Covariance matrix Σ^{dyn}.** We assume that the velocity innovations on the zonal and the meridional component are independent, i.e. that Σ^{dyn} has also a block structure

$$
\Sigma^{dyn} = \begin{pmatrix} \Sigma^{dyn}_{1,1} & 0 \\ 0 & \Sigma^{dyn}_{2,2} \end{pmatrix}
$$

where $\Sigma^{dyn}_{1,1} = \Sigma^{dyn}_{2,2}$ are $N \times N$ spatial covariance matrices with standard exponential shape

$$
\Sigma^{dyn}_{1,1}(i,j) = \sigma^2_{dyn} \exp\left(-\frac{d(\mathbf{p}_i, \mathbf{p}_j)}{\lambda_{dyn}}\right) \tag{11}
$$

where σ^2_{dyn} represents the variance of the innovation at each location and $\lambda_{dyn} > 0$ its spatial range. In practice, we have used the parameter values $\sigma_{dyn} = 1$ and $\lambda_{dyn} = 500km$.

- **Covariance matrix Σ^{obs}.** We also use a block structure

$$
\Sigma^{obs} = \begin{pmatrix} \Sigma^{obs}_{1,1} & 0 & 0 \\ 0 & \Sigma^{obs}_{2,2} & 0 \\ 0 & 0 & \Sigma^{obs}_{3,3} \end{pmatrix}
$$

where $\Sigma^{obs}_{1,1} = \Sigma^{obs}_{2,2} = \Sigma^{obs}_{3,3}$ are $N \times N$ spatial covariance matrices with coefficients

$$
\Sigma^{obs}_{1,1}(i,j) = \sigma^2_{obs} \exp\left(-\frac{d(\mathbf{p}_i, \mathbf{p}_j)}{\lambda_{obs}}\right)
$$

In practice, we have used the parameter values $\sigma_{obs} = 0.5$ and $\lambda_{obs} = 500km$.

3.2.3 Particle filter

The model described in the previous section is used to estimate hidden velocity fields that are consistent with the sea-state fields given by the ERA-Interim data set. For that, we compute the conditional expectation

$$
\tilde{V}(D,t) = E\left[V(D,t)|S(D,t_0), \cdots, S(D,t-\Delta t), S(D,t)\right]
$$

of the velocity field at time t given the history of the sea-state fields up until time t. Notice that a possible refinement would consist in computing the smoothing probabilities and then estimate the hidden velocity by the conditional expectation $E[V(D,t)|S(D,t_0), \cdots, S(D,t-\Delta t), S(D,t), S(D,t+1), ..., S(D,T)]$ of the velocity field given both past and future sea-state conditions.

The observation equation (10) is non linear, and in such situation it is usual to compute sequential Monte Carlo approximations of the conditional expectation using a particle filter algorithm. At each time step t, the conditional filtering distribution $p(V(D,t)|S(D,t_0), \cdots, S(D,t-\Delta t), S(D,t))$ is approximated by a set of M weighted particles $\{V^{(i)}(D,t), \omega_t^{(i)}\}_{i=1:M}$ where $V^{(i)}(D,t) \in \mathbb{R}^{2N}$ are velocity fields and $\omega_t^{(i)} \in [0,1]$ the associated weights. The conditional expectation is then approximated by

$$\hat{V}(D,t) = \sum_{i=1}^{M} \omega_t^{(i)} V^{(i)}(D,t).$$

and the weighted set of particles is updated at each time according to a Sequential Importance Sampling and Resampling scheme (see e.g. [9]). The sampling step at time t consists in simulating the velocity field using the dynamical model (9). We first generate independently $\{c_t^{(i)}\}_{i=1:M}$ according to a Bernoulli distribution with probability π_c, and then $V^{(i)}(D,t)$ is simulated using either the group velocity or the AR process depending on the value of $c_t^{(i)}$. The weights $\{\omega_t^{(i)}\}_{i=1:M}$ are then computed as follows :

$$\omega_t^{(i)} \propto \exp\left(-(||S(D,t) - S(D - V^{(i)}(D,t)\Delta t, t - \Delta t)||_{\Sigma_{obs}})^\nu\right) \qquad (12)$$

with $||d||_{\Sigma_{obs}}^2 = d^T \Sigma_{obs}^{-1} d$ and the normalizing constraint $\sum_{i=1}^{M} \omega_t^{(i)} = 1$. The parameter ν controls the decay of the weights with distance between the observed sea-state field at time t and the previous field after moving. The Gaussian model (10) corresponds to the case $\nu = 2$. In practice, due to the high dimension of the observation space, this choice leads to a quick degeneracy of the particles: after a few iterations, one weight is almost equal to one whereas all the others close to zeros (see also [6]). In order to avoid this problem, after several trials ν was chosen equal to $2/3$.

Notice that the computation of the weights using equation (12) requires the sea-state fields $S(D - V^{(i)}(D,t)\Delta t, t - \Delta t)$ at other locations than those given by ERA-Interim data set. In practice, we use a simple linear interpolation of the ERA-interim field at time $t - 1$.
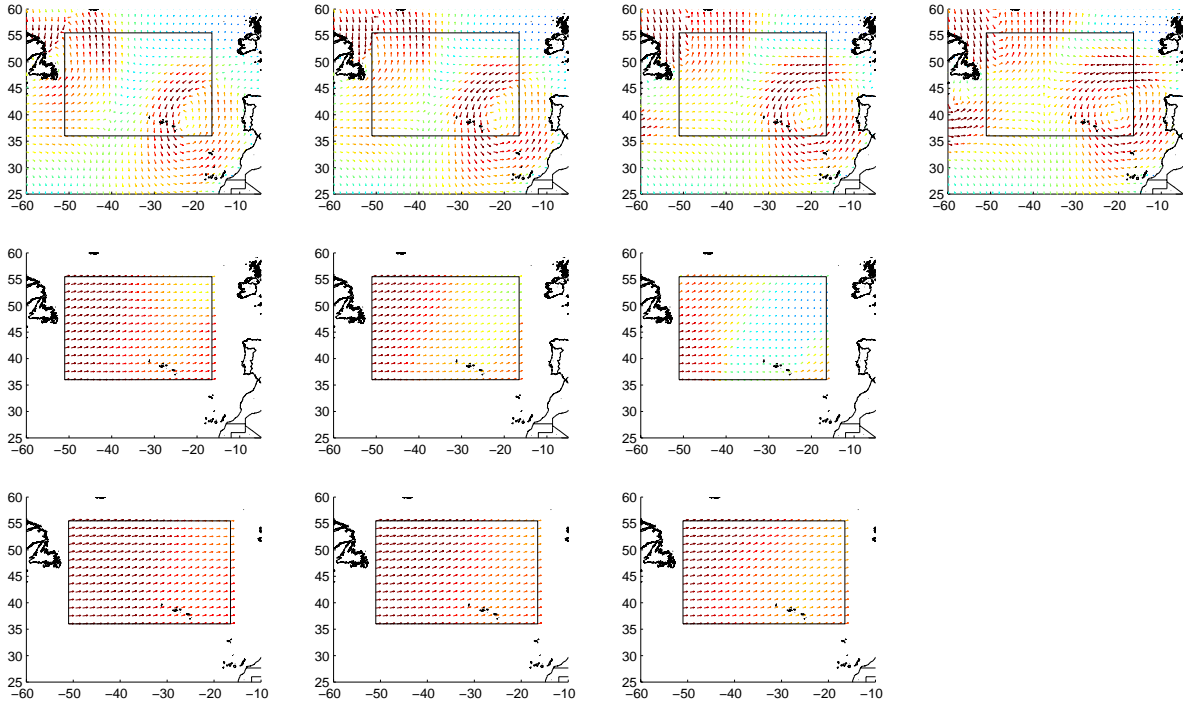
Figure 7: *ERA-Interim fields at consecutive times at 6 h intervals between 2002/12/16 at 18:00 and 2002/12/17 at 12:00 (top panels) and associated motions with the dispersion relation (middle panels), and without the dispersion relation (bottom panels).*

The particle filter was run with $M = 10^4$ particles. It seems to be a good compromise since it leads to reasonable CPU time and robust estimations of the velocity fields.

3.2.4 Visual validation

Before including the velocity fields in the covariance model, we have first performed visual check of their realism. An example of results is shown on Figure 7. It displays a 24 hour sequence of ERA-Interim fields where two low pressure systems can easily be identified: one in the North-West which is moving to the East and one in the South-East which is almost fixed. The velocity fields obtained with a mixing proportion $\pi_c = 1/3$ seem to be able to track these two systems, whereas use of a pure autoregressive model ($\pi_c = 0$) leads to a bad estimation of the velocity of the static system. In general, introducing the dispersion relation in the dynamics leads only to minor improvements except in some specific situations. We do believe however that the benefits would be substantial in areas where swell conditions are more dominating.
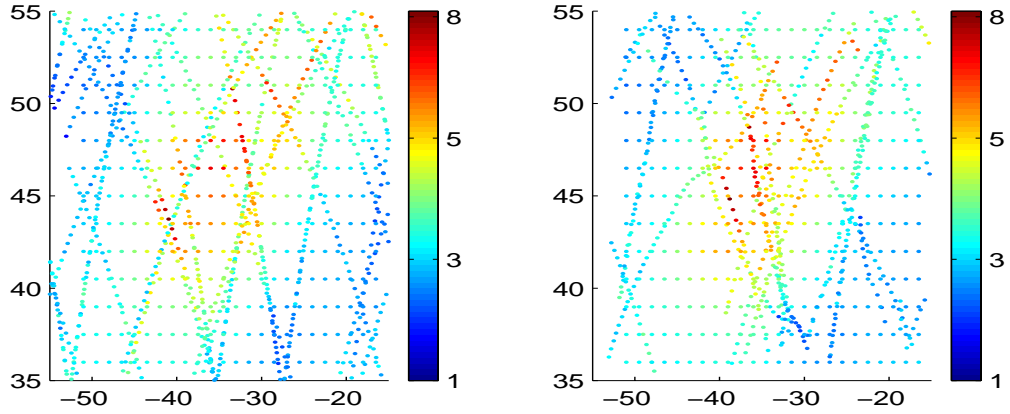
Figure 8: *All available satellite tracks for a period of time of 48h centered at time t_0 with ERA-Interim H_s field at time t_0: 2002/12/17 ; Left : without use of the displacements (Eulerian fields); Right : using the displacements (Lagrangian fields).*

It is also important to check that the estimated velocities are realistic for the significant wave height as measured by the satellites. The left panel on Figure 8 shows all the satellite data $H_s^{sat}(\mathbf{p}, t)$ as a function of $\mathbf{p} \in D$ available at time $t \in I$ where I is a 48 hour time window centered at some arbitrary time t_0. For comparison purposes, we also show the significant wave height field given by the ERA-Interim data set at time t_0. Due to the motion of sea-states, there are important differences between some data $H_s^{sat}(\mathbf{p}, t)$ and $H_s^{sat}(\mathbf{p}', t')$ which are close to each other on the Figure 8 (left panel) $(\mathbf{p} \approx \mathbf{p}')$ but which have been moving differently between time t (resp. t') and t_0. The right panel in Figure 8 shows the same satellite data in the Lagrangian reference frame induced by the estimated velocity field $\hat{V}$. More precisely, if $\boldsymbol{\phi}$ denotes the flow of diffeomorphisms associated to $\hat{V}$, we plot $H_s^{sat}(\boldsymbol{\phi}^{-1}(\mathbf{p}, t_0, t), t)$ as a function of $\mathbf{p}$. Since the flow of diffeomorphisms permits to follow the motion of the sea-states we expected an improved spatial coherence in the new reference frame. And indeed, on the right panel of Figure 8, the high significant wave height values are all located in the same area which corresponds to a storm location at time t_0. The storm can also be seen on the ERA-Interim data.

Figure 9 shows the estimated empirical joint probability density function of the zonal and the meridional components of the velocity at a central location $\mathbf{p}_0 =$ (46.5N,34.5W). The distribution shows again that sea-state systems are generally traveling to the east, but also that there is an important variability which cannot be accommodated by the
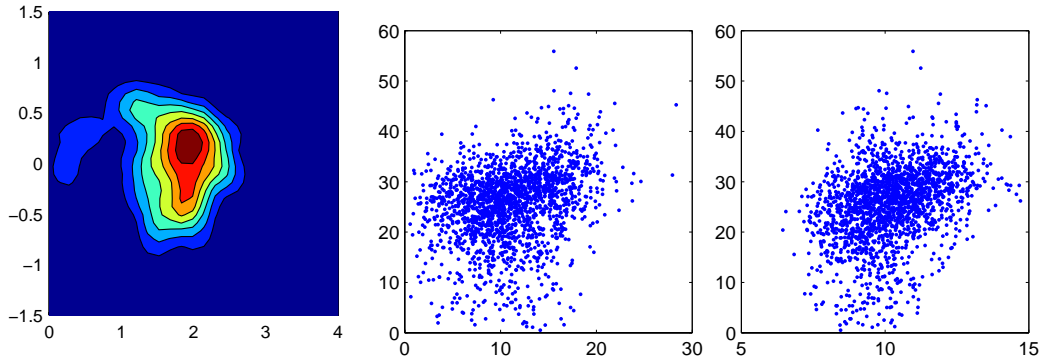
Figure 9: *Left panel : empirical bivariate pdf of $(u(\mathbf{p}_0), v(\mathbf{p}_0))$ at location $\mathbf{p}_0$ with coordinates (34.5W,46.5N). Middle panel: relation between the wind speed in ms^{-1} (x-axis) and the velocity of the sea-state motion in $km.h^{-1}$ (y-axis) at location (46.5N,34.5W). Right panel : relation between the mean period of the sea-state T_m in s (x-axis) and the velocity of the sea-state motion in $km.h^{-1}$ (y-axis) at location (34.5W,46.5N).*

constant velocity model. The mean velocity over the months of December 1992-2007 in D_0 is 19.5 kmh^{-1} to the East and 0.5 kmh^{-1} to the North, and this is slightly smaller than the mean velocity computed in Section 2 by fitting the constant velocity model. Figure 9 also shows that there is a positive dependence between the estimated velocities at location $\mathbf{p}_0$ and both the wind speed and the mean period of the sea state which are given by the ERA-Interim data set. The analogous plots for the direction show also a clear relation between the direction the sea-states are traveling and both the wind direction and the mean direction of the sea-state θ_m. This confirms that the estimated velocities are not only linked to the physical propagation of the waves but also to the motion of the wind fields.

3.3 Covariance estimation

Figure 10 shows empirical estimates of the covariance function computed using ERA-Interim and satellite data in the Lagrangian reference frame, i.e. as a function of $d(\boldsymbol{\phi}^{-1}(\mathbf{p}', t_0, t'), \boldsymbol{\phi}^{-1}(\mathbf{p}, t_0, t))$ and $|t' - t|$ and the fitted parametric model (3). This has been done for three different configurations. First for the static case, that is no dynamics are present, secondly, for constant velocity $\mathbf{V}_0$ and finally for the case where velocity is modeled in the presence of the dispersion relation. Inclusion of velocity into the covari-

23

Velocity	Static model		Constant Velocity		Dynamic ($\pi_c = 0$)		Dynamic ($\pi_c = \frac{1}{3}$)	
Data	ERA	Satellite	ERA	Satellite	ERA	Satellite	ERA	Satellite
θ_S (km)	966	873	997	896	966	913	967	909
θ_T (hour)	26.1	27.1	39.1	36.6	41.2	43.3	43.9	44.7
$1 - C_0$	0.018	0.047	0.029	0.051	0.019	0.047	0.021	0.051

Table 2: Parameter values of the fitted parametric covariance model (3) for different velocity fields.

ance functions results to longer temporal range dependence, even longer when velocity is modeled using dispersion relation, (see also Table 2). This effect was to be expected since the Lagrangian reference system moves along with the sea-states. The rate of the spatial dependence seems to be unaffected by the introduction of dynamics in the model.

Overall, the agreement between the empirical covariance function computed using the ERA-Interim and the satellite altimeter data and the fitted parametric covariance models, seem to be satisfactory as manifested in Figure 10. However, comparing the entries in Table 2, we can draw some further conclusions. There seems to be some inconsistency in the sizes of spatial dependence between the ERA-Interim and the satellite data sets, which is of the order of 100 km (relative difference of about 10%) in the absence of dynamics. This difference however drops to half when dynamics (non-constant velocity) are included in the model.

Summarizing, modeling sea-state motion using the dispersion relation leads to a covariance model with longer temporal dependence which should allow better tracking of the sea-state motion.
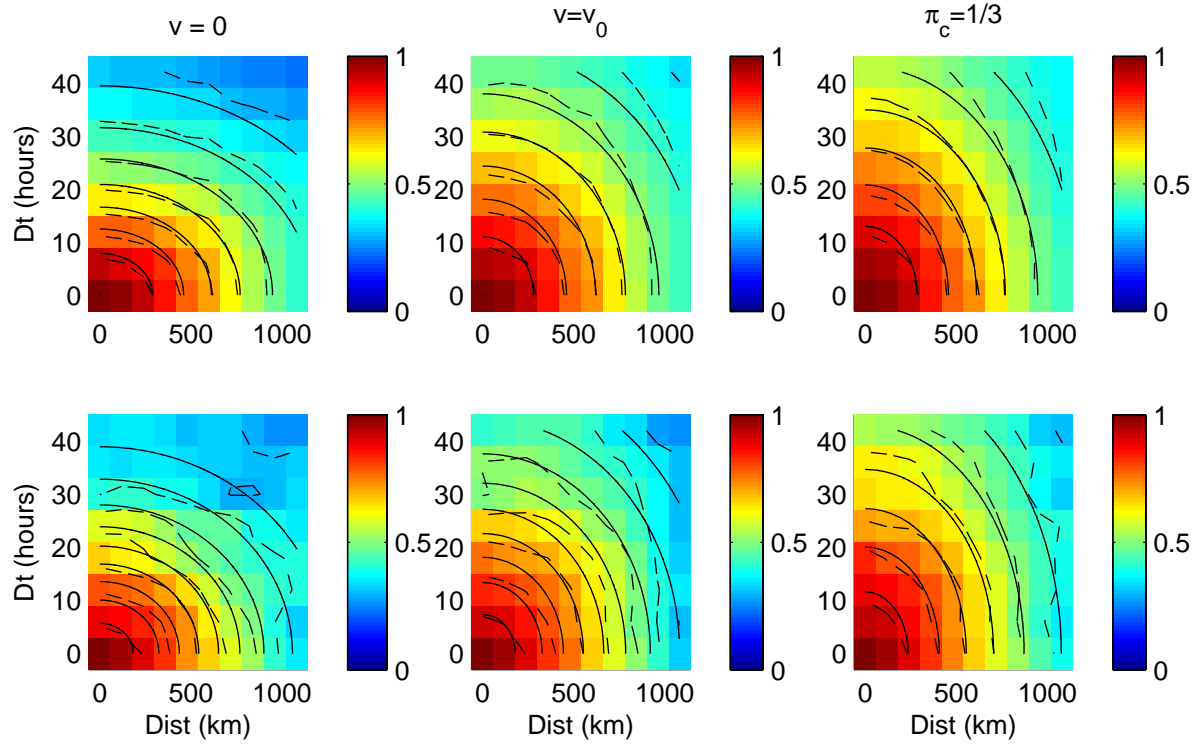
Figure 10: *Empirical correlation function of the Lagrangian field computed using ERA-Interim data (top panels) and satellite data (bottom panels). Left panels : no motion ($\mathbf{V}(\mathbf{p}, t) = 0$), middle panels : constant velocity ($\mathbf{V}(\mathbf{p}, t) = \mathbf{V}_0$), right panels : dynamic velocity with dispersion relation ($\pi_c = 1/3$). The dashed lines represent the levels of the empirical covariance function and the full lines the levels of the fitted parametric model.*

Velocity	Null	Frozen	$\pi_c = 0$	$\pi_c = \frac{1}{3}$
RMSE (m)	0.5250	0.5018	0.4810	0.4801

Table 3: RMSE for different velocity fields computed using cross-validation.

4 Numerical results

In this section, we check the accuracy of the fields obtained by interpolating satellite data using the model presented in this paper and ordinary kriging (see e.g. [14]). This could be useful, in order for example, to produce historical data for metocean studies. We first perform cross-validation on satellite data and then make a comparison with a buoy.

4.1 Cross-validation of satellite data

In this section, the whole methodology is validated using cross-validation on satellite data. For that, each satellite passage which intersects D_0 is predicted using all other satellite passages that are available in a 2 day time window. The global root mean square errors (RMSE) of the difference between the true satellite data and the interpolated values are given in Table 3. The model with dynamic velocity performs best and the introduction of the velocity (even constant) clearly improves the static model. Introducing the dispersion relation in the dynamics does not seem to improve things a lot, but again we expect more benefits in an area dominated by swell conditions.

Figure 11 gives the RMSE for each year: it is obviously correlated to the amount of available satellite data and the RMSE decreases when the number of operational satellites increases (see Figure 1). The inter-annual variability though, could also be due to other factors like meteorological conditions (years with stormy conditions and high significant wave height are more difficult to forecast) and the geometry of the tracks (for example, new satellites are calibrated by working simultaneously with another satellite, and this may reduce the benefit of introducing the motion in the model).

4.2 Virtual buoy

For many offshore applications, it is beneficial to have time series describing significant wave height conditions on a long time period at some specific locations. When there is no
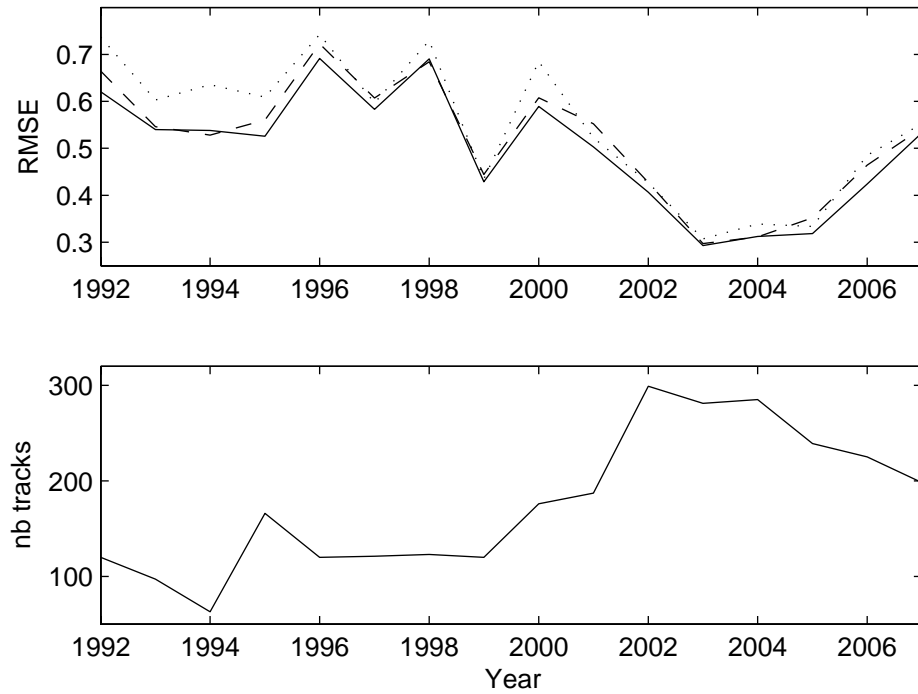
Figure 11: *Top panel : time evolution of the root mean square error computed using cross-validation on the period 1992-2007. Solid line : changing velocity ($\pi_c = 1/3$), dashed line: frozen velocity, dotted line: no velocity. Bottom panel: evolution of the number of satellite tracks which cross the domain D.*
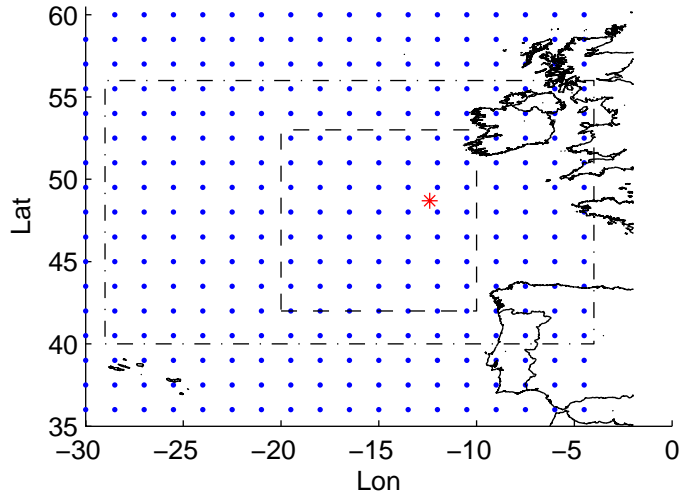
Figure 12: Grid of ERA-Interim data, K1 buoy (*), domains D_0 (dashed box) and D (dash-dotted box).

in-situ information available at the location of interest, hindcast data are generally used in operational applications. However these data have well known limitations such as the under-estimation of the severe sea-states.

Using the model proposed in this paper, we can produce significant wave height time series at any location by interpolating the satellite data. Here, we focus on the location (48.7N,12.4W) for which there is also buoy data for the time period 2002-2007. As the buoy is not inside the domain D_0 considered previously, we have modified the domains D and D_0 as shown on Figure 12. Then, the model has been fitted in the new areas using the method described in the previous sections. Again, we used a time window $\Delta T = \pm 24h$.

Table 4 shows that the method that gives a virtual buoy that best matches the in-situ observations is obtained by interpolating satellite data using the velocity fields computed using particle filter. It slightly improves the results obtained with ERA-Interim data (we use the same interpolation method with changing velocity than for satellite data) and the model with satellite data and the constant velocity. If hindcast data are not available to estimate the motion, this last interpolation method should be favored since it clearly improves the results obtained with a static covariance model.

Figure 13 shows that ERA-Interim data systematically under-estimates high significant wave height values, and this is problematic for extreme value analysis. This is also evident on the quantile-quantile plots on Figure 14. The two examples of time-series shown

Velocity Data	Null Sat.	Constant Sat.	Changing Sat.	Changing ERA
RMSE	0.45	0.36	0.33	0.35

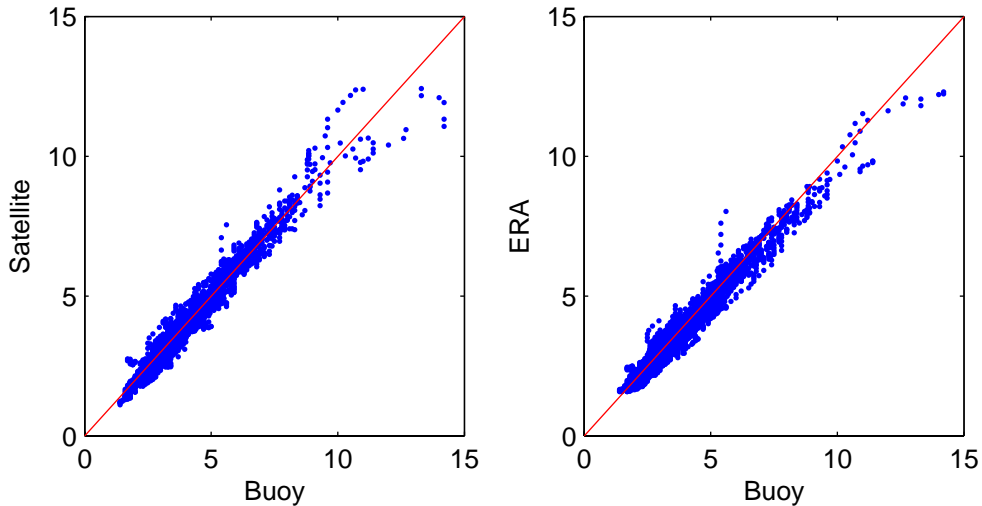Table 4: RMSE of the time series of the buoy for different velocity fields and data sets.

Figure 13: *Scatter plot of H_s measured at the buoy (x-axis) against interpolated value (y-axis) obtained from satellite (left) and ERA-Interim data (right).*

on Figures 15 and 16 show also that ERA-Interim data tend to smooth the temporal variability observed at the buoy and under-estimate the significant wave height in the storms. But, storms are not always well reconstructed using satellite data neither. For example, the biggest storm observed in buoy data occurred around the 12/09/2007 (see Figure 16) and the time-series obtained using satellite data fail to reproduce both the intensity of the event and the date of arrival of the storm. The main reason is probably the poor sampling of this storm by satellites : only one satellite passage crosses the track of the storm and at locations far from the position of the buoy (about 200 km).
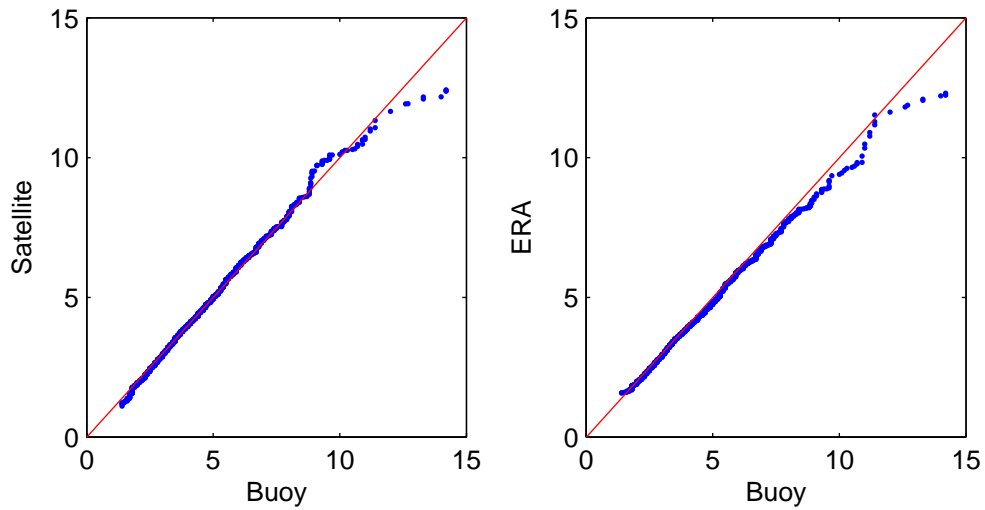
Figure 14: *Quantile-quantile plot of H_s measured at the buoy (x-axis) against interpolated value (y-axis) obtained from satellite (left) and ERA-Interim data (right).*
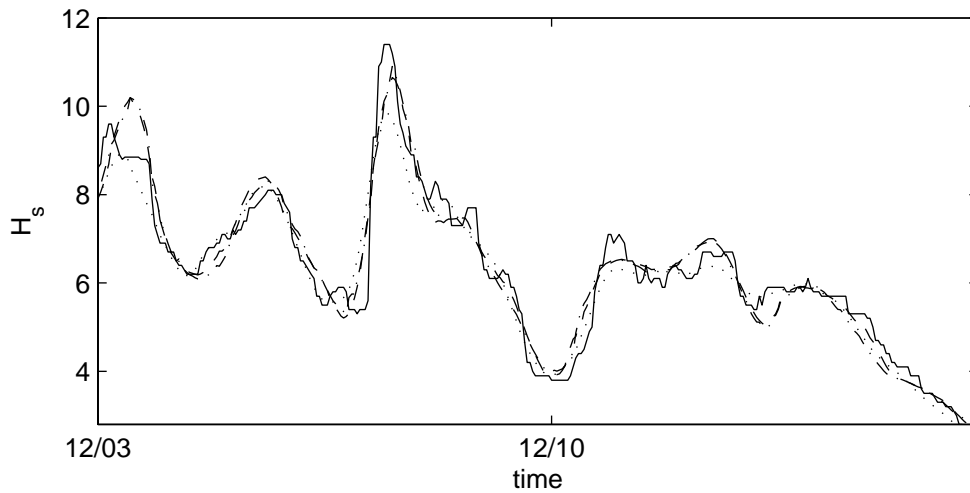

Figure 15: *Example of H_s time series measured at K1 buoy (solid line), compared with interpolated using satellite data and changing velocity [resp. constant velocity] (dashed line [resp. dash-dotted line]), interpolated using ERA-Interim data and changing velocity (dotted line).*

Figure 16: *Example of H_s time series measured at K1 buoy (solid line), compared with interpolated using satellite data and changing velocity [resp. constant velocity] (dashed line [resp. dash-dotted line]), interpolated using ERA-Interim data and changing velocity (dotted line).*
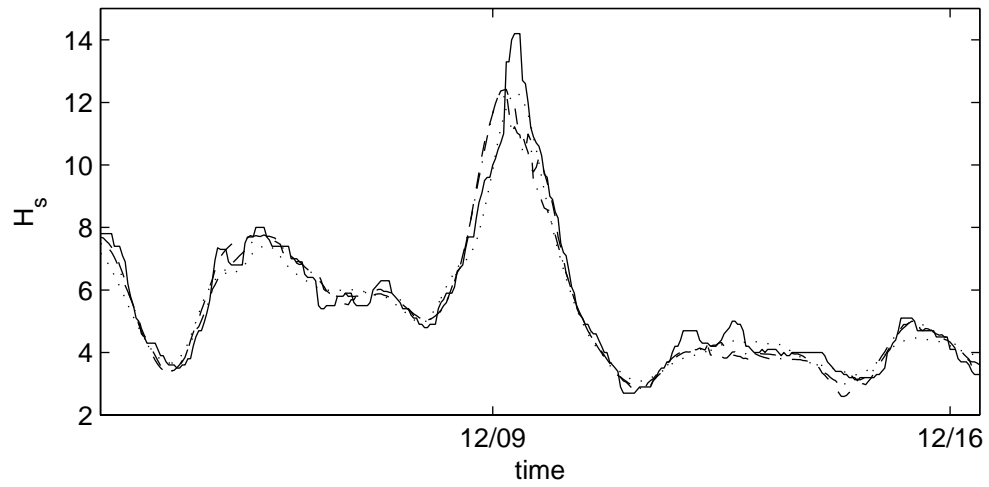
31

5   Conclusions

The contributions of this paper are mainly for applications. First of all, in order to model the evolution of the significant wave height over space and time, three different sources of data, satellite altimeter, reanalysis and buoy data, have been combined. These sources provide information at different scales.

Since the ocean, and so quantities such as the significant wave height, can be thought of as a random surface which develops over time, a static model for the covariance structure of the field does not suffice. Instead, we propose to study the problem using the Lagrangian reference frame; that is, we include dynamics in the model by considering a static field subordinated by a velocity field. Assuming the velocity is constant over space and time, although improving the results obtained using a static model, is still unrealistic and hence a new velocity field introduced through a flow of diffeomorphisms is considered.

The important problem of motion estimation is then formulated and solved sequentially within a state-space model framework. The motion is described by a hidden Markov chain, including a physical a priori law, the dispersion relation, to describe the evolution of the sea-states. The data used are the ERA-Interim data set. On the one hand, the satellite coverage is generally poor for tracking the motion of the sea-states. On the other hand, the sampling regularity of the ERA-Interim data, both in time and space, allows for the use of existing models for processes on regular grids.

The fitted models have been cross-validated against satellite and buoy data. In both cases, the best results are obtained by interpolating satellite data using velocity fields computed by means of the state-space model. If hindcast data are not available for estimating the velocity fields, the model with constant velocity should be favored since it clearly improves the results obtained with the static covariance model. For operational applications, interpolating satellite data using the dynamic covariance model proposed in this paper should lead to a better description of extreme events compared to hindcast data.

Acknowledgments

References

[1] P. Ailliot, V. Monbet, and M. Prevosto. An autoregressive model with time-varying coefficients for wind fields. *Environmetrics*, 17(2):107–117, 2006.

[2] G.A. Athanassoulis and Ch.N. Stephanakos. A nonstationary stochastic model for long-term time series of significant wave height. *Journal of Geophysical Research*, 100(C8):16149–16162, 1995.

[3] A. Baxevani, C. Borgel, and I. Rychlik. Spatial models for the variability of the significant wave height on the world oceans. *International Journal of Offshore and Polar Engineering*, 18(1):1–7, 2008.

[4] A. Baxevani, S. Caires, and I. Rychlik. Spatio-temporal statistical modelling of significant wave height. *Environmetrics*, 20:14–31, 2009.

[5] A. Baxevani, I. Rychlik, and R.J. Wilson. A new method for modelling the space variablilty of significant wave height. *Extremes*, June 2006.

[6] L.M. Berliner and C.K. Wikle. Approximate importance sampling monte carlo for data assimilation. *Physica D*, 230:37–49, 2007.

[7] S. Caires and A. Sterl. A new non-parametric method to correct model data: Application to significant wave height from the era-40 reanalysis. *J. Atmospheric and Oceanic Tech.*, 22(4):443–459, 2005.

[8] S. Caires, A. Sterl, J. R. Bidlot, N. Graham, and V. Swail. Intercomparison of different wind wave reanalyses. *J. Clim.*, 17(10):1893–1913, 2004.

[9] O. Cappé, E. Moulines, and Rydén T. *Inference in hidden Markov models*. Springer-Verlag, New York, 2005.

[10] G. Christakos. *Modern spatiotemporal geostatistics*. Oxford University Press, New York, 2000.

[11] G. Christakos and D. T. Hristopoulos. *Spatiotemporal environmental health modeling: a tractatus stochastic*. Kluwer Academic Publishers, 1998.

[12] T. Corpetti, E. Mémin, and P. Pérez. Dense estimation of fluid flows. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(3):365–380, 2002.

[13] P.D. Cotton, P.G. Challenor, L. Redbourn-Marsh, S. Gulev, A. Sterl, and R.S. Bprtkovskii. An intercomparison of voluntary observing satellite data and modelling wave climatologies. In *In Swail, V. R., editor, Satellite Microwave Remote Sensing*, pages 451–460, WMO, Geneva, Switzerland, 2001.

[14] N. Cressie. *Statistics for spatial data.* Wiley-Interscience; Rev Sub edition, 1993.

[15] N. Cressie and H.C. Huang. Classes of nonseparable, spatiotemporal stationary covariance functions. *J. Am.Stat. Assoc.*, 94:1330–1340, 1999.

[16] A. Cuzol and E. Mémin. A stochastic filtering technique for fluid flow velocity fields tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(7):1278–1293, 2009.

[17] T. Gneiting, M. G. Genton, and P. Guttorp. *Geostatistical space-time models, stationarity, separability and full symmetry. In Statistical Methods for Spatio-Temporal Systems.* Chapman Hall/CRC, 2007.

[18] P.A.E.M. Janssen, J.D. Doyle, J. Bidlot, B. Hansen, L. Isaksen, and P. Viterbo. *Impact and feedback of ocean waves on the atmosphere.* 2002.

[19] S. C. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Trans, Image Process*, 9:1357–1370, 2000.

[20] G.J. Komen, L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P.A.E.M. Janssen. *Dynamics and Modelling of Ocean Waves.* Cambridge Univ. Press, 1994.

[21] G. K. Korotaev, E. Huot, F.X. Le Dimet, I. Herlin, S.V. Stanichny, D.M. Solovyev, and L. Wu. Retrieving ocean surface current by 4-d variational assimilation of sea surface temperature images. *Remote Sensing of Environment*, 112:1464–1475, 2008.

[22] H Krogstad and S. Barstow. Directional distributions in ocean wave spectra. In *Proc. 9 th Int. Offshore and Polar Eng. Conf., ISOPE*, volume III, p 76-89, 1999.

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

[24] C. Ma. Nonstationary covariance functions that model space-time interactions. *Statistics and probability letter*, October 2002.

[25] J. Marcello, F. Eugenio, F. Marqués, A. Hernàndez-Guerra, and A. Gasull. Motion estimation techniques to automatically track oceanographic thermal structures in multisensor image sequences. *IEEE trans. on Geosciences and Remote sensing*, 46(9):2743–2762, 2008.

[26] B. Markussen. Large deformation diffeomorphisms with application to optic flow. *Computer Vision and Image Understanding*, 106(1):97–105, 2007.

[27] S. McGinnity and G.W. Irwin. Maneuvering target tracking using a multiple-model bootstrap filter. *A. Doucet, N. de Fretitas, N. Gordon (Eds) : Sequential Monte Carlo methods in practice*, pages 247–271, 2001.

[28] N. Papadakis, T. Corpetti, and E. Mémin. Dynamically consistent optical flow estimation. In *Proc. Int. Conf. on Computer Vision (ICCV'07)*, Rio de Janeiro, Brazil, October 2007.

[29] P. Queffeulou. Long term validation of wave height measurements from altimeters. *Marine Geodesy*, 27:495–510, 2004.

[30] J. Schmetz, K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, and L. Van De Berg. Operational cloud-motion winds from Meteosat infrared images. *Journal of Applied Meteorology*, 32:1206–1225, 1993.

[31] G.B. Whitham. *Linear and Nonlinear Waves*. New York: John Wiley Sons, 1974.